Chebyshev Inequalities with Law Invariant Deviation Measures

Bogdan Grechuk Anton Molyboha Michael Zabarankin*

Abstract

The consistency of law invariant general deviation measures, introduced by Rockafellar et al., with concave ordering has been used to generalize Rao-Blackwell theorem and to develop an approach for reducing minimization of law invariant deviation measures to minimization of the measures on subsets of undominated random variables with respect to concave ordering. This approach has been applied for constructing the Chebyshev and Kolmogorov inequalities with law invariant deviation measures, in particular with mean absolute deviation, lower semideviation and conditional value-at-risk deviation. Also, an advantage of the Kolmogorov inequality with certain deviation measures has been illustrated in estimating the probability of the exchange rate of two currencies to be within specified bounds.

1 Introduction

The notions of risk and deviation, particularly in finance literature, are often used interchangeably since Markowitz [9], who was the first to suggest the use of variance as the measure of risk in portfolio analysis. The recently emerged theory of general deviation measures [10, 15], relying on an axiomatic framework and dual characterization, generalizes the notion of *standard deviation* to measure *"nonconstancy"* in a random variable (r.v.). These measures possess all the main properties of the standard deviation, however, in contrast to the latter, are not necessarily symmetric with respect to the ups and downs of an r.v. Besides standard deviation as the originating example, the most well-known deviation measures include *lower and upper semideviations*, *mean absolute deviation*, *median absolute deviation*, *conditional value-at-risk (CVaR) deviation*, *mixed CVaR deviation*, and *worst-case mixed-CVaR deviation*. In Markowitz's portfolio selection problem [9], Rockafellar et al. [11, 12] replaced standard deviation by a general deviation measure and generalized a number of results in classical portfolio theory including the one-fund theorem and the capital asset pricing model (CAPM); see also [10, 13, 14].

The aim of this work is to generalize the well-known Chebyshev inequality for *law invariant* deviation measures, i.e. those which depend only on distributions of r.v.'s. This class includes all the aforementioned examples of deviation measures and is preferable in engineering applications. In decision making under uncertainty or in the case of limited information, the Chebyshev inequality is often used for estimating the probability of a dread event or disaster. For example, Roy [16] estimated the probability of an uncertain portfolio return X to be in default, i.e., less than a specified threshold ξ, in terms of the mean $\mu = EX$ and standard deviation $\sigma = \sigma(X)$ of X via the classical two-sided (two-tailed) Chebyshev inequality:

$$\mathbb{P}[X \leq \xi] \leq \mathbb{P}[|X-\mu| \geq \mu - \xi] \leq \frac{\sigma^2}{(\mu-\xi)^2}, \qquad \xi \leq \mu. \tag{1}$$

*Department of Mathematical Sciences, Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ 07030; bgrechuk@stevens.edu, amolyboh@stevens.edu, mzabaran@stevens.edu

Estimating this probability is motivated by the principle of Safety First, which is central in the actuarial science and asserts that an individual will seek to reduce the probability of a dread event as much as possible [16]. Remarkably, if ξ is the risk-free rate of return r_f, the right-hand side in (1) is reciprocal to the famous *Sharpe ratio* $(\mu - r_f)/\sigma$, and consequently, the problem of minimizing the probability of portfolio default can be replaced by maximizing the Sharpe ratio (unless X is normally distributed, the problems are not equivalent). Although the estimate of $\mathbb{P}[X \le \xi]$ can be improved by the one-sided (one-tailed) Chebyshev inequality

$$\mathbb{P}[X \le \xi] \le \frac{1}{1 + \left(\frac{\mu - \xi}{\sigma}\right)^2}, \qquad \xi \le \mu, \tag{2}$$

the main motivation question is what if another deviation measure $\mathcal{D}(X)$ of X is either known or preferable in decision making — can we generalize (1) and (2) for an arbitrary law invariant deviation measure and what are the implications of this generalization? For example, Smith [17] generalized (1) in terms of certain moments of X (other than the mean and variance) and illustrated its advantage in various operations research problems including Bayesian statistics and option pricing. Chebyshev inequalities with general deviation measures would complement the recently developed mean-deviation framework [11, 12] for portfolio analysis and enrich decision making techniques. Their application is not, however, limited to the areas of finance and risk analysis where the choice of a particular deviation measure is dictated by agent's risk preferences. Generalized Chebyshev inequalities can be used in statistics to evaluate the probability of how significantly an r.v. deviates from its expected value in terms of customized measures of deviation. They could prove to be invaluable in engineering, in particular in reliability, safety, and quality control.

On an *atomless* (nonatomic) probability space, every law invariant lower-semicontinuous (l.s.c.) deviation measure is consistent with *concave ordering*,[1] i.e., if $X \succcurlyeq_c Y$ for two r.v.'s X and Y, then $\mathcal{D}(X) \le \mathcal{D}(Y)$ for every law invariant deviation measure $\mathcal{D}$. This fact follows from the result of Dana [2] and means that decision making with law invariant deviation measures over random outcomes with equal means conforms with *risk averse* preferences (*second order stochastic dominance (SSD)*). However, significance of this result goes well beyond its implications in decision making and risk analysis. Based on this fact, we generalize *Rao-Blackwell Theorem* for law invariant deviation measures and, what is more important, develop an approach for minimizing law invariant deviation measures with chance constraints. In general, the approach reduces minimization of a law invariant deviation measure on a set $U \subseteq \mathcal{L}^p(\Omega)$ to minimization of the deviation measure on a set U_c of undominated r.v.'s with respect to concave ordering:

$$(a) \quad \inf \mathcal{D}(X) \quad \text{s.t.} \quad X \in U \qquad \Longleftrightarrow \qquad (b) \quad \inf \mathcal{D}(X) \quad \text{s.t.} \quad X \in U_c \tag{3}$$

where the set $U_c \subseteq U$ is called *reduced set* and is defined as the minimal set with the following property: for any $Y \in U$, there exists $X \in U_c$ such that $X \succcurlyeq_c Y$. Obviously, the problems (3a) and (3b) are equivalent. The question of whether (3a) attains its minimum on U is not trivial. In certain cases, it can be readily answered based on the equivalence of (3a) and (3b): when the reduced set U_c is bounded and weakly closed, a law invariant l.s.c. deviation measure $\mathcal{D}$ attains its minimum on U_c; see [6, Theorem 9.2] and [7, Theorem 7.3.4]. However, the suggested approach is especially efficient when U is determined by chance constraints on $\mathcal{L}^p(\Omega)$. In this case, (3b) reduces to a finite parameter optimization problem.

We formulate Chebyshev inequalities with an arbitrary law invariant deviation measure $\mathcal{D}$ as the minimization problem (3a) with chance constraints and using the suggested approach, reduce the problem to (3b). In

[1] If F_X and F_Y are cumulative distribution functions of r.v.'s X and Y with expected values EX and EY, respectively, then, X dominates Y with respect to second-order stochastic dominance (SSD), or $X \succcurlyeq_2 Y$, if $\int_{-\infty}^x F_X(t)dt \le \int_{-\infty}^x F_Y(t)dt$ for all $x \in \mathbb{R}$. X dominates Y with respect to concave ordering, or $X \succcurlyeq_c Y$, if $X \succcurlyeq_2 Y$ and $EX = EY$; see [8]. In particular, $X \succcurlyeq_2 Y$ implies that $E[U(X)] \ge E[U(Y)]$ for all *increasing concave utility functions*, while $X \succcurlyeq_c Y$ implies that $E[U(X)] \ge E[U(Y)]$ for all *concave utility functions* (not necessarily increasing).

particular, for the two-sided Chebyshev inequality, the set U_c turns out to be a one-parameter family of discrete r.v.'s, and thus, (3b) becomes one-parameter optimization over that family of discrete r.v.'s., while for the one-sided Chebyshev inequality, U_c consists only of a single r.v., which just solves (3b). As an illustration, one-sided and two-sided Chebyshev inequalities are constructed for mean-absolute deviation, lower semideviation and CVaR deviation and are then specialized to the case when the distributions of r.v.'s belong to some set of distributions, e.g., when the distributions are symmetric. Based on the Chebyshev inequality with an arbitrary law invariant deviation measure $\mathcal{D}$, we derive a Kolmogorov inequality, which estimates the probability of a discrete-time martingale $S(t)$, $0 \le t \le T$, to stay within specified bounds provided that $\mathcal{D}(S(T))$ is given. The classical Kolmogorov inequality estimates the same probability in terms of standard deviation of $S(T)$. As an example, we show that for a discrete-time martingale with an approximately normal distribution, Kolmogorov inequalities with certain deviation measures provide better estimates for the probability in question than the classical Kolmogorov inequality does.

The paper is organized into six sections. Section 2 reviews main properties of deviation measures. Section 3 presents the technique for solving optimization problems with law invariant deviation measures based on the fact that these deviation measures are consistent with concave ordering. Through the suggested optimization approach, Section 4 generalizes the Chebyshev inequality for an arbitrary law invariant deviation measure and presents examples of the one-sided and two-sided inequalities for lower semideviation, mean absolute deviation, and CVaR deviation. Section 5 derives the generalized Kolmogorov inequality and presents an example illustrating advantage of certain deviation measures in estimating the probability of the exchange rate of two currencies to be within specified bounds. Section 6 concludes the paper.

2 Deviation Measures

This section reviews main properties of deviation measures.

Let $(\Omega, \mathcal{M}, \mathbb{P})$ be a probability space, where Ω denotes the designated space of future states ω, $\mathcal{M}$ is a field of sets in Ω, and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{M})$. We assume that the probability space Ω is *atomless*, i.e., there exists an r.v. with a continuous cumulative distribution function (CDF). This implies existence of r.v.'s on Ω with all possible distribution functions (see, e.g., [3]). We restrict our attention to r.v.'s from $\mathcal{L}^p(\Omega) = \mathcal{L}^p(\Omega, \mathcal{M}, \mathbb{P}), 1 \le p < \infty$, with the norm $\|X\|_p = (E[|X|^p])^{1/p}$. Let $F_X(x) = \mathbb{P}[X \le x]$ and $q_X(\alpha) = \inf\{z \,|\, F_X(z) > \alpha\}$ denote the CDF and an α-quantile of an r.v. X, respectively. The relations between r.v.'s will be understood to hold in the almost sure sense, e.g., we write $X = Y$ if $\mathbb{P}[X = Y] = 1$ and $X \ge Y$ if $\mathbb{P}[X \ge Y] = 1$.

The general deviation measures, introduced by Rockafellar et al. [10, 15], are defined as follows.

Definition 1 (deviation measures) *A* deviation measure[2] *is any functional* $\mathcal{D} : \mathcal{L}^p(\Omega) \to [0, \infty]$ *satisfying*

(D1) $\mathcal{D}(X) = 0$ *for constant* X*, but* $\mathcal{D}(X) > 0$ *otherwise (*nonnegativity*),*

(D2) $\mathcal{D}(\lambda X) = \lambda \mathcal{D}(X)$ *for all* X *and all* $\lambda > 0$ *(*positive homogeneity*),*

(D3) $\mathcal{D}(X + Y) \le \mathcal{D}(X) + \mathcal{D}(Y)$ *for all* X *and* Y *(*subadditivity*),*

(D4) set $\{X \in \mathcal{L}^p(\Omega) \big| \mathcal{D}(X) \le c\}$ *is closed for all* $c < \infty$ *(*lower-semicontinuity*).*

Axiom D1 has the consequence, shown in [10], that

$$\mathcal{D}(X + C) = \mathcal{D}(X) \text{ for all constants } C \text{ (\textit{insensitivity to constant shift})}. \tag{4}$$

[2] In [10, 15], deviation measures are defined on $\mathcal{L}^2(\Omega)$, and axiom D4 was not included in the definition. Deviation measures satisfying D4 were called *l.s.c. deviation* measures.

In general, for two r.v.'s with the same distribution, a deviation measure may assume different values. In this work, we consider only distribution-independent or so-called law invariant deviation measures [10].

Definition 2 (law invariant deviation measures) *A deviation measure* $\mathcal{D} : \mathcal{L}^p(\Omega) \to [0, \infty]$ *is called* law invariant, *if* $\mathcal{D}(X_1) = \mathcal{D}(X_2)$ *for any two r.v.'s* X_1 *and* X_2 *yielding the same distribution function on* $(-\infty, \infty)$.

The most well-known examples of deviation measures are:

(i) standard deviation $\sigma(X) = \|X - EX\|_2$;

(ii) lower and upper semideviations $\sigma_-(X) = \|[X - EX]_-\|_2$ and $\sigma_+(X) = \|[X - EX]_+\|_2$, where

$$[X]_- = \max\{0, -X\}, \qquad [X]_+ = \max\{0, X\}; \tag{5}$$

(iii) *mean absolute deviation* $\mathrm{MAD}(X) = E|X - EX|$;

(iv) *conditional value-at-risk (CVaR) deviation* defined for any $\alpha \in (0,1)$ by

$$\mathrm{CVaR}_\alpha^\Delta(X) \equiv EX - \frac{1}{\alpha}\int_0^\alpha q_X(\beta)d\beta. \tag{6}$$

All these deviation measures are law invariant. See e.g. [10] for more examples.

An important property of the class of law invariant deviation measures is its consistency with *concave ordering*. An r.v. X dominates Y with respect to second-order stochastic dominance (SSD), or $X \succcurlyeq_2 Y$, if $\int_{-\infty}^x F_X(t)dt \le \int_{-\infty}^x F_Y(t)dt$ for all $x \in \mathbb{R}$, and X dominates Y with respect to concave ordering, or $X \succcurlyeq_c Y$, if $X \succcurlyeq_2 Y$ and $EX = EY$; see [8]. A functional $\mathcal{F} : \mathcal{L}^p(\Omega) \to \mathbb{R} \cup \{-\infty\}$ is called *Schur concave* [2] if $X \succcurlyeq_c Y$ implies $\mathcal{F}(X) \ge \mathcal{F}(Y)$. We call a deviation measure $\mathcal{D}$ *consistent* with concave ordering, if $-\mathcal{D}$ is Schur concave, i.e. $X \succcurlyeq_c Y$ implies $\mathcal{D}(X) \le \mathcal{D}(Y)$. The result of Dana [2, Theorem 4.1], restated below, implies that on an atomless probability space every law invariant deviation measure possesses this property.

Proposition 1 *Let* $(\Omega, \mathcal{M}, P)$ *be an atomless probability space and* $\mathcal{D} : \mathcal{L}^p(\Omega) \to [0, \infty]$ *be any law invariant deviation measure. Then* $X \succcurlyeq_c Y$ *implies* $\mathcal{D}(X) \le \mathcal{D}(Y)$.

The consistency of law invariant deviation measures with concave ordering has two implications. First is that decision making with law invariant deviation measures over random outcomes with the same mean conforms with risk-averse preferences (SSD) and second is that it plays a central role in generalizing a variety of classical results, e.g., Rao-Blackwell inequality, Chebyshev inequality, Kolmogorov inequality, for an arbitrary law invariant deviation measure.

The first result that immediately follows from the consistency of law invariant deviation measures with concave ordering is a generalization of Rao-Blackwell theorem[3] for an arbitrary law invariant deviation measure.

Proposition 2 (generalized Rao-Blackwell inequality) *Let* $\mathcal{D} : \mathcal{L}^p(\Omega) \to [0, \infty]$ *be a law invariant deviation measure. Suppose* $X \in \mathcal{L}^p(\Omega)$ *and* $Y : \Omega \to \mathbb{R}$ *are some r.v.'s. Let an r.v.* Z *be defined by* $Z = E[X|Y]$. *Then* $\mathcal{D}(Z) \le \mathcal{D}(X)$.

Proof. Since every law invariant deviation measure is consistent with concave ordering, the proof follows from the fact that $Z \succcurlyeq_c X$ (see, e.g., [3, Corollary 2.62]). □

The consistency of law invariant deviation measures with concave ordering also plays a crucial role for solving optimization problems with deviation measures, in particular nonconvex problems with chance constraints. This is the subject of the next section.

[3] Classical Rao-Blackwell theorem states that $\sigma(E[X|Y]) \le \sigma(X)$ for standard deviation σ.

3 Minimization of Law Invariant Deviation Measures

In this section, we develop an optimization technique for minimizing law invariant deviation measures and illustrate the technique in solving an optimization problem with a chance constraint.

Let $\mathcal{D} : \mathcal{L}^p(\Omega) \to [0,\infty]$ be a law invariant deviation measure and let U be an arbitrary set of r.v.'s in $\mathcal{L}^p(\Omega)$. For an optimization problem

$$\min \mathcal{D}(X) \quad \text{s.t.} \quad X \in U \tag{7}$$

we define a *reduced set* as a set with the following property.

Definition 3 *A set $U_c \subseteq U$ is called a* reduced set *for the problem (7), if for every $X \in U$ there exists $Y \in U_c$, such that $\mathcal{D}(X) \geq \mathcal{D}(Y)$.*

Clearly, if U_c is a reduced set for (7), then (7) is equivalent to the following problem

$$\min \mathcal{D}(X) \quad \text{s.t.} \quad X \in U_c. \tag{8}$$

Since $\mathcal{D}$ is consistent with concave ordering, a reduced set U_c can be chosen as a set of undominated r.v.'s with respect to concave ordering.

Proposition 3 *Let $\mathcal{D} : \mathcal{L}^p(\Omega) \to [0,\infty]$ be a law invariant deviation measure. Let $U_c \subseteq U$ be a set such that for every $X \in U$ there exists $Y \in U_c$, such that $Y \succcurlyeq_c X$. Then U_c is a reduced set for the problem (7).*

Proof. Consistency of $\mathcal{D}$ with concave ordering implies $\mathcal{D}(X) \geq \mathcal{D}(Y)$. □

Proposition 3 reduces the minimization problem over a set $U \subseteq \mathcal{L}^p(\Omega)$ to the minimization problem over a set U_c of undominated r.v.'s, which often has a simpler structure. In particular, the suggested approach is especially efficient when U is determined by chance constraints on $\mathcal{L}^p(\Omega)$. In this case, reduced set U_c is often a one-parameter family, or even a singleton.

The following *sufficient* condition for concave ordering, established by Hanoch and Levy [4, Theorem 3], is central for constructing reduced sets for different problems throughout the paper.

Proposition 4 *Let $X_1 \in \mathcal{L}^p(\Omega)$ and $X_2 \in \mathcal{L}^p(\Omega)$ be r.v.'s with CDFs $F_1(x)$ and $F_2(x)$, respectively, with $EX_1 = EX_2$. If there exists $x_0 \in \mathbb{R}$ such that $F_1(x) \leq F_2(x)$ for $x < x_0$ and $F_1(x) \geq F_2(x)$ for $x \geq x_0$, then $X_1 \succcurlyeq_c X_2$.*

Example 1 *Let $U = \{X \in \mathcal{L}^p(\Omega) \mid EX = 0,\ \mathbb{P}[X \leq -a] \geq \beta\}$ with fixed $a > 0$ and $\beta \in (0,1)$. Then the reduced set for the problem (7) is a singleton $U_c = \{X^\star\}$, where*

$$X^\star = \begin{cases} -a & \textit{with probability} \quad \beta \\ \frac{a\beta}{1-\beta} & \textit{with probability} \quad 1-\beta \end{cases} \tag{9}$$

Consequently, the optimal value of the problem is $\min\limits_{X \in U} \mathcal{D}(X) = \mathcal{D}(X^\star)$.

Detail. Since $\mathbb{P}[X^\star \leq -a] = \beta$ and $EX^\star = 0$, we have $X^\star \in U$. Next we show that for any $X \in U$, $X^\star$ dominates X with respect to concave ordering. Let $X \in \mathcal{L}^p(\Omega)$ be an r.v. such that $EX = 0$ and $\mathbb{P}[X \leq -a] \geq \beta$ for some fixed $\beta \in (0,1)$, and let $F^\star(x)$ and $F(x)$ be CDFs of $X^\star$ and X, respectively. It follows from $F(-a) = \mathbb{P}[X \leq -a] \geq \beta$ that $F^\star(x) \leq F(x)$ for $x < a\beta/(1-\beta)$ and that $1 = F^\star(x) \geq F(x)$ for $x \geq a\beta(1-\beta)$. Thus, by Proposition 4, $X^\star \succcurlyeq_c X$, and Proposition 3 implies that $U_c = \{X^\star\}$ is the reduced set for the problem (7). □

Proposition 3 is a key element in our approach to the optimization problems throughout the paper. In particular, in the following section we use Proposition 3 to construct generalized Chebyshev inequalities with law invariant deviation measures.

4 Chebyshev Inequalities with Law Invariant Deviation Measures

The classical Chebyshev theorem estimates the probability of how significantly an r.v. deviates from its mean in terms of its standard deviation and is formulated as follows (see, e.g., [5]).

Theorem 1 (Chebyshev Theorem for standard deviation) *For any r.v. $X \in \mathcal{L}^p(\Omega)$ and any real number $a > 0$, the Chebyshev inequality holds*

$$\mathbb{P}[|X - EX| \geq a] \leq \frac{\sigma(X)^2}{a^2}. \tag{10}$$

The question of interest is whether we can obtain probability estimates similar to (10) in terms of other deviation measures. We consider the following generalization for (10):

$$\mathbb{P}[X - EX \leq -a \text{ or } X - EX \geq b] \leq g_{\mathcal{D}}(\mathcal{D}(X)), \tag{11}$$

where $\mathcal{D}$ is an arbitrary law invariant deviation measure, the condition "$X - EX \leq -a$ or $X - EX \geq b$" with $a > 0$ and $b > 0$ is an "asymmetric" generalization of the interval $|X - EX| \geq a$, and the function $g_{\mathcal{D}}$ is to be determined. Also, the Chebyshev inequality can be improved, if considered r.v.'s are from a subspace of $\mathcal{L}^p(\Omega)$, e.g., from a subspace of symmetric r.v.'s. We will derive improved generalized Chebyshev inequalities which hold only for r.v.'s from a cone $V \subset \mathcal{L}^p(\Omega)$ and will also derive one-sided versions of all these Chebyshev inequalities.

4.1 Problem Formulation in Optimization Framework

The problem that we address is formulated as follows.

Problem I *Given a law invariant deviation measure $\mathcal{D} : \mathcal{L}^p(\Omega) \to [0, \infty]$, cone $V \subset \mathcal{L}^p(\Omega)$ (i.e. $X \in V$ implies $\lambda X \in V$ for $\lambda > 0$), and constants $a > 0$ and $b > 0$, construct a function $g_{\mathcal{D}}(d)$ such that*

$$\mathbb{P}[X \leq -a \text{ or } X \geq b] \leq g_{\mathcal{D}}(\mathcal{D}(X)) \qquad \text{for all } X \in V \tag{12}$$

under the following two requirements:

(R1) $g_{\mathcal{D}}$ is determined by a, b, V, and $\mathcal{D}$ only (i.e., the form of $g_{\mathcal{D}}$ does not depend on X).

(R2) $g_{\mathcal{D}}$ provides the smallest possible bound in (12), i.e., for every $d > 0$, the Chebyshev inequality (12) turns into equality for some X with $\mathcal{D}(X) = d$. (For example, the trivial solution $g_{\mathcal{D}} \equiv 1$ in most cases fails to satisfy this requirement.)

In particular, the classical Chebyshev inequality (10) is a special case of (12) with $\mathcal{D}(X) = \sigma(X)$, $b = a$, and $V = \{X \mid EX = 0\}$. The Chebyshev inequality with general deviation measure (11) is a version of (12) with $V = \{X \mid EX = 0\}$.

To satisfy the condition (R2) in Problem I, we reformulate the inequality (12) as the maximization problem

Problem II (Optimal $g_{\mathcal{D}}(d)$)

$$\begin{aligned} g_{\mathcal{D}}(d) = \sup_{X \in \mathcal{L}^p(\Omega)} \; & \mathbb{P}[X \leq -a \text{ or } X \geq b] \\ \text{s.t. } & X \in V, \; \mathcal{D}(X) \leq d, \end{aligned} \tag{13}$$

and then proceed with the following complementary problem.

Problem III (complementary problem) *Given a law invariant deviation measure* $\mathcal{D}: \mathcal{L}^p(\Omega) \to [0,\infty]$*, cone* V*, and a real number* $\beta \in [0,1]$*, find*

$$
\begin{aligned}
u_{\mathcal{D}}(\beta) = &\inf_{X \in \mathcal{L}^p(\Omega)} \mathcal{D}(X) \\
&\text{s.t. } X \in \mathcal{C}_\beta = \{X \mid X \in V,\ \mathbb{P}[X \le -a \ \ or\ \ X \ge b] \ge \beta\}\,.
\end{aligned}
\tag{14}
$$

Problem III is an optimization problem with a chance constraint, and, in general, is nonconvex. For some β the feasible set $\mathcal{C}_\beta$ can be empty, and in this case we define $u_{\mathcal{D}}(\beta) = +\infty$. The function $u_{\mathcal{D}}(\beta)$ in (14) is nondecreasing with respect to β, since $\mathcal{C}_{\beta_1} \supseteq \mathcal{C}_{\beta_2}$ for any $\beta_1 < \beta_2$. It can be shown that a solution to Problem I is an inverse to $u_{\mathcal{D}}(\beta)$.

Proposition 5 *Let* $u_{\mathcal{D}}(\beta)$ *be the optimal value of Problem III. Then a solution to Problem I is the inverse function to* $u_{\mathcal{D}}(\beta)$ *given by*

$$
g_{\mathcal{D}}(d) = \sup\{\beta \mid u_{\mathcal{D}}(\beta) \le d\}. \tag{15}
$$

Proof. We should prove that for the function (15), the inequality (12) holds. Let X be an arbitrary r.v. such that $X \in V$. Let $\beta = \mathbb{P}[X \le -a \ \text{ or } \ X \ge b]$ and $d = \mathcal{D}(X)$. According to (14), $X \in \mathcal{C}_\beta$ and $u_{\mathcal{D}}(\beta) \le \mathcal{D}(X) = d$. Consequently, it follows from (15) that $g_{\mathcal{D}}(d) \ge \beta$. In other words, $g_{\mathcal{D}}(\mathcal{D}(X)) \ge \mathbb{P}[X \le -a \ \text{ or } \ X \ge b]$. □

Thus, given an optimal value, $u_{\mathcal{D}}(\beta)$, of Problem III, the function $g_{\mathcal{D}}(d)$ defined by (15) solves (12). Now, we establish a sufficient condition for $g_{\mathcal{D}}(d)$ to satisfy the requirement (R2) in Problem I.

Proposition 6 *Let the infimum in* (14) *be attained on* $\mathcal{C}_\beta$ *for all* $\beta \in [0,1]$*, and let* $u_{\mathcal{D}}(\beta)$ *be l.s.c. for all* $\beta \in [0,1]$*. Then the function* $g_{\mathcal{D}}$*, given by* (15)*, satisfies the requirement (R2) in Problem I.*

Proof. For some fixed $d > 0$, we should construct an r.v. $X \in V$ with $\mathcal{D}(X) = d$, yielding equality in (12). The case $g_{\mathcal{D}}(d) = 0$ is trivial, and thus, we assume $g_{\mathcal{D}}(d) > 0$. Let X^* solve the problem (14) for $\beta^* = g_{\mathcal{D}}(d)$, i.e. $u_{\mathcal{D}}(\beta^*) = \mathcal{D}(X^*)$. Lower-semicontinuity of $u_{\mathcal{D}}(\beta)$ along with (15) implies $\mathcal{D}(X^*) = u_{\mathcal{D}}(\beta^*) \le d$. For the r.v. $\widehat{X} = X^* \cdot d/\mathcal{D}(X^*)$, we have $\mathcal{D}(\widehat{X}) = d$ and $\widehat{X} \in V$, and since $d/\mathcal{D}(X^*) \ge 1$, we also obtain

$$
\mathbb{P}[\widehat{X} \le -a \ \text{ or } \ \widehat{X} \ge b] \ge \mathbb{P}[X^* \le -a \ \text{ or } \ X^* \ge b] \ge \beta^* = g_{\mathcal{D}}(d),
$$

where the last inequality follows from the fact that $X^* \in \mathcal{C}_{\beta^*}$. This implies that $\widehat{X}$ yields equality in (12), and the proof is finished. □

The problem (14) has the form (7) and can be solved by the technique developed in Section 3.

Algorithm 1 (Constructing two-sided generalized Chebyshev inequality)

1. *Given* $a, b > 0$ *and a cone* $V \subset \mathcal{L}^p(\Omega)$*, for the problem (14), construct a reduced set* U_c*, based on Proposition 3.*

2. *Given a deviation measure* $\mathcal{D}$*, solve*

$$
u_{\mathcal{D}}(\beta) = \min_{X \in U_c} \mathcal{D}(X). \tag{16}
$$

3. *Verify that* $u_{\mathcal{D}}(\beta)$ *is l.s.c. and that the infimum in (16) is attained for all* $\beta \in [0,1]$*, i.e. all assumptions of Proposition 6 hold.*

4. *For the function* $u_{\mathcal{D}}(\beta)$*, construct its inverse* $g_{\mathcal{D}}(d)$ *according to (15).*

Generalized two-sided Chebyshev inequality takes the form

$$\mathbb{P}[X \le -a \ \text{ or } \ X \ge b] \le g_{\mathcal{D}}(\mathcal{D}(X)) \quad \text{for all } X \in V. \tag{17}$$

The correctness of (17) follows from Proposition 5, the requirement (R1) is obviously holds, and the requirement (R2) follows from Proposition 6.

Similarly, one-sided version of the Chebyshev inequality can be formulated as follows.

Problem IV *Given a law invariant deviation measure $\mathcal{D} : \mathcal{L}^p(\Omega) \to [0,\infty]$, cone $V \subset \mathcal{L}^p(\Omega)$, and constant $a > 0$, construct a function $g^-_{\mathcal{D}}(d)$ such that*

$$\mathbb{P}[X \le -a] \le g^-_{\mathcal{D}}(\mathcal{D}(X)) \quad \textit{for all } X \in V, \tag{18}$$

where $g^-_{\mathcal{D}}(\mathcal{D}(X))$ satisfies the requirements (R1) and (R2) for (18).

Similarly to Problem I, this one reduces to an optimization problem.

Problem V *Given a law invariant deviation measure $\mathcal{D}$, cone V, and a real number $\beta \in (0,1)$, find*

$$\begin{aligned} u^-_{\mathcal{D}}(\beta) = &\inf_{X \in \mathcal{L}^p(\Omega)} \mathcal{D}(X) \\ &\text{s.t. } X \in \mathcal{C}^-_\beta = \{X \mid X \in V,\ \mathbb{P}[X \le -a] \ge \beta\}. \end{aligned} \tag{19}$$

Proposition 7 *Let $u^-_{\mathcal{D}}(\beta)$ be the optimal value of Problem V for a given deviation measure $\mathcal{D}$ and cone V. Then the function*

$$g^-_{\mathcal{D}}(d) = \sup\{\beta \mid u^-_{\mathcal{D}}(\beta) \le d\} \tag{20}$$

satisfies (18) *and the requirement (R1). If, in addition, the infimum in* (19) *is attained on $\mathcal{C}^-_\beta$ for all $\beta \in (0,1)$, and $u^-_{\mathcal{D}}(\beta)$ is l.s.c. for all $\beta \in (0,1)$, then $g^-_{\mathcal{D}}$ satisfies the requirement (R2) for Problem IV.*

Proof. The proof of this fact follows from the proofs of Propositions 5 and 6. □

To construct one-sided generalized Chebyshev inequalities, we use the following algorithm.

Algorithm 2 (Constructing one-sided generalized Chebyshev inequality)

1. *Given $a > 0$ and a cone $V \subset \mathcal{L}^p(\Omega)$, for the problem (19), construct a reduced set U^-_c, based on Proposition 3.*

2. *Given a deviation measure $\mathcal{D}$, for every $\beta \in (0,1)$, solve*

$$u^-_{\mathcal{D}}(\beta) = \min_{X \in U^-_c} \mathcal{D}(X). \tag{21}$$

3. *Verify that $u^-_{\mathcal{D}}(\beta)$ is l.s.c. and that the infimum in (21) is attained for all $\beta \in (0,1)$, i.e. all assumptions of Proposition 7 hold.*

4. *Construct the inverse function $g^-_{\mathcal{D}}(d)$ according to (20).*

Generalized one-sided Chebyshev inequality takes the form

$$\mathbb{P}[X \le -a] \le g^-_{\mathcal{D}}(\mathcal{D}(X)) \quad \text{for all } X \in V. \tag{22}$$

Proposition 7 guarantees that (22) is correct and satisfies the requirements (R1) and (R2) for Problem IV.

4.2 Chebyshev Inequalities with Law Invariant Deviation Measures

We begin with constructing two-sided generalized Chebyshev inequality (11) for any law invariant deviation measure $\mathcal{D}$. By Proposition 5, the inequality reduces to the optimization problem (14) with $V = \{X \mid EX = 0\}$. The following proposition constructs a reduced set for this problem.

Proposition 8 *For the problem (14) with $V = \{X \mid EX = 0\}$, a reduced set is a one-parameter family $U_c = \left\{X_x | x \in [0,\beta] \cap [\beta - a/(a+b), b/(a+b)]\right\}$, where*

$$X_x = \begin{cases} -a & \text{with probability} \quad x, \\ \dfrac{ax - b(\beta - x)}{1-\beta} & \text{with probability} \quad 1-\beta, \\ b & \text{with probability} \quad \beta - x. \end{cases} \tag{23}$$

Proof. It follows from (23) that $EX_x = 0$ and $\mathbb{P}[-a < X < b] \le 1-\beta$. Consequently, in (14), $X_x \in \mathcal{C}_\beta$ for every $x \in [0,1]$. To show that for any $X \in \mathcal{C}_\beta$ there exists $x \in [0,\beta] \cap [\beta - a/(a+b), b/(a+b)]$ such that $X_x \succcurlyeq_c X$, we consider two cases.

In the first case, let $\mathbb{P}[X \le -a] \le b/(a+b)$ and $\mathbb{P}[X \ge b] \le a/(a+b)$. Then for $x = \min\left\{\beta, \mathbb{P}[X \le -a]\right\}$, the inequality $0 \le \beta - x \le \mathbb{P}[X \ge b]$ implies that $x \le b/(a+b)$ and $\beta - x \le a/(a+b)$, and consequently, $x \in [0,\beta] \cap [\beta - a/(a+b), b/(a+b)]$.

Also, it is straightforward to verify, that under these conditions we have $-a \le (ax - b(\beta - x))/(1-\beta) \le b$. Let $F(t)$ and $F_*(t)$ be the CDFs of X and X_x, respectively. Then $F_*(t) \le F(t)$ for $t < (ax - b(\beta - x))/(1-\beta)$, and $F_*(t) \ge F(t)$ for $t \ge (ax - b(\beta - x))/(1-\beta)$. Thus, based on Proposition 4, we conclude that $X_x \succcurlyeq_c X$.

In the second case, let either $\mathbb{P}[X \le -a] > b/(a+b)$ or $\mathbb{P}[X \ge b] > a/(a+b)$ hold. For the r.v. X_0 defined by

$$X_0 = \begin{cases} -a & \text{with probability} \quad b/(a+b), \\ b & \text{with probability} \quad a/(a+b), \end{cases}$$

with the CDF $F_0(t)$, we have $EX_0 = 0$, $\mathbb{P}[-a < X_0 < b] = 0$, $\mathbb{P}[X_0 \le -a] \le b/(a+b)$ and $\mathbb{P}[X_0 \ge b] \le a/(a+b)$. Consequently, as shown in the first case, $X_x \succcurlyeq_c X_0$ for some x.

If $\mathbb{P}[X \le -a] > b/(a+b)$, then $F_0(t) \le F(t)$ for $t < b$, and $F_0(t) \ge F(t)$ for $t \ge b$. Proposition 4 implies that $X \succcurlyeq_c X_0$ and thus, $X_x \succcurlyeq_c X_0 \succcurlyeq_c X$. Similarly, if $\mathbb{P}[X \ge b] > a/(a+b)$, then $F_0(t) \le F(t)$ for $t < -a$, and $F_0(t) \ge F(t)$ for $t \ge -a$, whence $X_x \succcurlyeq_c X_0 \succcurlyeq_c X$. It is left to apply Proposition 3, and (23) follows. $\square$

By Proposition 8, a solution to the problem (14) with $V = \{X \mid EX = 0\}$ can be represented by

$$u_{\mathcal{D}}(\beta) = \min_{x \in [0,\beta] \cap [\beta - \frac{a}{a+b}, \frac{b}{a+b}]} h_\beta(x), \tag{24}$$

where $h_\beta(x) = \mathcal{D}(X_x)$, with X_x given by (23).

Similarly, constructing one-sided version of the Chebyshev inequality with law invariant deviation measures, namely,

$$\mathbb{P}[X - EX \le -a] \le g_{\mathcal{D}}^{-}(\mathcal{D}(X)), \tag{25}$$

reduces to the optimization problem (19) with $V = \{X \mid EX = 0\}$. By virtue of Example 1, a reduced set for this problem is a single r.v. $X^\star$, determined by (9). Thus, $u_{\mathcal{D}}^{-}(\beta) = \mathcal{D}(X^\star)$.

As an illustration for the developed optimization approach, we construct two-sided symmetric[4] and one-sided Chebyshev inequalities for mean absolute deviation, lower semideviation, CVaR deviation, and a symmetrization of CVaR deviation given by $\mathcal{D}(X) = \max\left\{\mathrm{CVaR}_\alpha^\Delta(X), \mathrm{CVaR}_\alpha^\Delta(-X)\right\}$.

[4]Two-sided symmetric inequalities correspond to the case $b = a > 0$.

Example 2 (mean absolute deviation) *For $MAD = E|X-EX|$ and $a>0$, two-sided symmetric and one-sided Chebyshev inequalities take the form*

$$\mathbb{P}[|X-EX|\geq a] \leq \frac{\mathrm{MAD}(X)}{a} \quad \textit{and} \quad \mathbb{P}[X\leq EX-a] \leq \frac{\mathrm{MAD}(X)}{2a}, \tag{26}$$

respectively.

Detail. For X_x, given by (23), the function $h_\beta(x)$ in (24) reduces to

$$h_\beta(x) = \mathrm{MAD}(X_x) = |-a|\cdot x + \left|a\cdot\frac{2x-\beta}{1-\beta}\right|\cdot(1-\beta)+|a|\cdot(\beta-x) = a\cdot(\beta+|2x-\beta\,|\,).$$

It attains its minimum at $x=\beta/2$, and (24) implies that $u_{\mathcal{D}}(\beta)=a\beta$. From (15), we obtain $g_{\mathcal{D}}(d)=d/a$. Consequently, the two-sided Chebyshev inequality reduces to the first formula in (26).

To construct a one-sided version of the inequality, we use Example 1 and obtain

$$u_{\mathcal{D}}^{-}(\beta) = \mathrm{MAD}(X^\star) = |-a|\cdot\beta + \left|a\cdot\frac{\beta}{1-\beta}\right|\cdot(1-\beta) = 2a\beta,$$

where $X^\star$ is given by (9). Then we deduce from (20) that $g_{\mathcal{D}}^{-}(d)=d/2a$, and consequently, the one-sided Chebyshev inequality is given by the second formula in (26). □

Example 3 (lower semideviation) *For lower semideviation $\sigma_-^2(X)=\|[X-EX]_-\|_2$ and $a>0$, the two-sided Chebyshev inequality takes the form*

$$\mathbb{P}[|X-EX|\geq a] \leq \begin{cases} \dfrac{16}{9}k+\dfrac{1}{9}, & k\geq\dfrac{1}{20} \\ \dfrac{1}{2}\left(\sqrt{k^2+4k}-k\right), & k<\dfrac{1}{20} \end{cases} \tag{27}$$

where $k=\sigma_-^2(X)/a^2$.

The one-sided version of the inequality is given by

$$\mathbb{P}[X\leq EX-a] \leq \frac{\sigma_-(X)^2}{a^2}. \tag{28}$$

Detail. For X_x, defined by (23), the function $h_\beta(x)$ in (24) reduces to

$$h_\beta(x) = \sigma_-(X_x) = \begin{cases} a\sqrt{(-1)^2\cdot x + \left(\dfrac{2x-\beta}{1-\beta}\right)^2\cdot(1-\beta)}, & x\leq\beta/2, \\ a\sqrt{x}, & x>\beta/2. \end{cases}$$

The minimum of $h_\beta(x)$ over $x\in[0,\beta]\cap[\beta-1/2,1/2]$ is attained at $x=(5\beta-1)/8$ if $\beta\geq 1/5$ and at $x=0$ if $\beta<1/5$. Since $(5\beta-1)/8\leq\beta/2$ for $\beta\leq 1$, the value of the minimum is

$$u_{\mathcal{D}}(\beta) = \begin{cases} \dfrac{1}{4}a\sqrt{9\beta-1}, & \beta\geq 1/5, \\ \dfrac{a\beta}{\sqrt{1-\beta}}, & \beta<1/5. \end{cases}$$

Consequently, it follows from (15) that

$$g_{\mathcal{D}}(d) = \sup\{\beta \mid u_{\mathcal{D}}(\beta) \le d\} = \begin{cases} \dfrac{16d^2 + a^2}{9a^2}, & d \ge \dfrac{a}{\sqrt{20}}, \\ \dfrac{\sqrt{d^4 + 4d^2a^2} - d^2}{2a^2}, & d < \dfrac{a}{\sqrt{20}}. \end{cases}$$

With $k = d^2/a^2 = \sigma_-^2(X)/a^2$, we obtain the Chebyshev inequality (27).

To construct a one-sided version of the inequality, we use Example 1 and obtain $u_{\mathcal{D}}^-(\beta) = \sigma_-(X^\star) = a\sqrt{\beta}$. This together with (20) yields

$$g_{\mathcal{D}}^-(d) = \sup\{\beta \mid a\sqrt{\beta} \le d\} = \frac{d^2}{a^2},$$

and (28) follows. □

Example 4 (conditional value-at-risk deviation) *Let* $d = \mathrm{CVaR}_\alpha^\Delta(X)/a$. *Then for* $\alpha \le 1/2$, *the two-sided Chebyshev inequality is represented by*

$$\mathbb{P}[|X - EX| \ge a] \le \begin{cases} \dfrac{d}{1+d}, & d \le \dfrac{1/2-\alpha}{1/2+\alpha}, \\ \dfrac{1}{2} + \alpha d - \sqrt{\alpha(1-d)(1-\alpha(1+d))}, & d \in \left[\dfrac{1/2-\alpha}{1/2+\alpha}, 1\right), \\ 1, & d \ge 1. \end{cases} \tag{29}$$

For $\alpha \ge 1/2$, *the two-sided Chebyshev inequality has similar form and can be obtained by using the relation* $\alpha \cdot \mathrm{CVaR}_\alpha^\Delta(X) \equiv (1-\alpha) \cdot \mathrm{CVaR}_{1-\alpha}^\Delta(-X)$.

The one-sided version of the inequality for all $\alpha \in (0,1)$ *takes the form*

$$\mathbb{P}[X \le EX - a] \le \begin{cases} \dfrac{\alpha \cdot \mathrm{CVaR}_\alpha^\Delta(X)}{a + \alpha(\mathrm{CVaR}_\alpha^\Delta(X) - a)}, & \mathrm{CVaR}_\alpha^\Delta(X) < a, \\ 1, & \mathrm{CVaR}_\alpha^\Delta(X) \ge a. \end{cases} \tag{30}$$

Detail. We only sketch the derivation of the Chebyshev inequality (29) for the case $\alpha \le 1/2$. For X_x, given by (23), the function $h_\beta(x)$ in (24) reduces to

$$\frac{h_\beta(x)}{a} = \frac{\mathrm{CVaR}_\alpha^\Delta(X_x)}{a} = \begin{cases} \dfrac{1-\alpha}{\alpha}, & x \le \alpha + \beta - 1, \\ \dfrac{2x^2 + (1 - 2\alpha - 2\beta)x + \alpha\beta}{\alpha(1-\beta)}, & \alpha + \beta - 1 \le x \le \alpha, \\ 1, & \alpha \le x. \end{cases}$$

It can be shown that the optimal value of the optimization problem (24) is determined by

$$\frac{u_{\mathcal{D}}(\beta)}{a} = \begin{cases} \dfrac{\beta}{1-\beta}, & \beta \le 1/2 - \alpha, \\ \dfrac{1 - (1-2\alpha)^2 - (1-2\beta)^2}{8\alpha(1-\beta)}, & 1/2 - \alpha \le \beta \le 1/2 + \alpha, \\ 1, & 1/2 + \alpha \le \beta. \end{cases}$$

Finally, we compute $g_{\mathcal{D}}(d)$ according to (15) and obtain the two-sided Chebyshev inequality (29).

For a one-sided version of the inequality, we use Example 1 and obtain

$$u_{\mathcal{D}}^{-}(\beta) = \mathrm{CVaR}_{\alpha}^{\Delta}(X^{\star}) = \begin{cases} a\dfrac{\beta(1-\alpha)}{(1-\beta)\alpha}, & \beta \le \alpha, \\ a, & \beta > \alpha. \end{cases}$$

Then for $d < a$, equation (20) yields

$$g_{\mathcal{D}}^{-}(d) = \sup\{\beta \mid u_{\mathcal{D}}^{-}(\beta) \le d\} = \frac{\alpha d}{a + \alpha(d-a)}.$$

For $d = \mathrm{CVaR}_{\alpha}^{\Delta}(X)$, the above formula reduces to the right-hand side of (30). For $d \ge a$, the best estimate is 1. □

Example 5 *For the deviation measure* $\mathcal{D}(X) = \max\left\{\mathrm{CVaR}_{\alpha}^{\Delta}(X), \mathrm{CVaR}_{\alpha}^{\Delta}(-X)\right\}$*, the two-sided Chebyshev inequality takes the form*

$$\mathbb{P}[|X - EX| \ge a] \le \begin{cases} \dfrac{2\alpha\mathcal{D}(X)}{a}, & \mathcal{D}(X) < a, \\ 1, & \mathcal{D}(X) \ge a. \end{cases} \tag{31}$$

Detail. For X_x, defined by (23), the minimum of $h_\beta(x) = \mathcal{D}(X_x)$ over $x \in [0,\beta] \cap [\beta - 1/2, 1/2]$ (see (24)) is attained at $x = \beta/2$. We have

$$u_{\mathcal{D}}(\beta) = \begin{cases} \dfrac{a\beta}{2\alpha}, & \beta \le 2\alpha, \\ a, & \beta > 2\alpha. \end{cases}$$

Consequently, from (15) we obtain

$$g_{\mathcal{D}}(d) = \sup\{\beta \mid u_{\mathcal{D}}(\beta) \le d\} = \begin{cases} 2\alpha d/a, & d < a, \\ 1, & d \ge a. \end{cases}$$

□

Obviously, in all the examples above the assumptions in Propositions 6 and 7 hold. The following proposition guarantees that this will be the case for an arbitrary law invariant deviation measure $\mathcal{D}$.

Proposition 9 *Let* $\mathcal{D} : \mathcal{L}^p(\Omega) \to \mathbb{R}$ *be a law invariant deviation measure, and let* $V = \{X \in \mathcal{L}^p(\Omega) : EX = 0\}$.

(i) $u_{\mathcal{D}}^{-}(\beta)$*, given by (19), is l.s.c.*

(ii) for every $\beta \in [0,1]$, $\mathcal{D}$ *in (14) attains its minimum on* $\mathcal{C}_\beta$ *and* $u_{\mathcal{D}}(\beta)$ *is l.s.c.*

Proof. (i) For every $\beta \in [0,1]$, let X_β be a solution to (19) determined by (9). Let a sequence β_n converge from the left to some fixed $\beta \in (0,1)$. Without loss of generality, we may assume that each X_{β_n} is comonotone[5] with X_β (such r.v.'s exist by virtue of [2, Lemma 4.2]). In this case, (9) implies that the sequence X_{β_n} is uniformly bounded and converges to X_β a.s. This implies convergence in $\mathcal{L}^p(\Omega)$, and D4 implies that $u_{\mathcal{D}}^{-}(\beta) = \mathcal{D}(X_\beta) \le \liminf_{n\to\infty} \mathcal{D}(X_{\beta_n}) = \liminf_{n\to\infty} u_{\mathcal{D}}^{-}(\beta_n)$.
(ii) For every $\beta \in [0,1]$ and $x \in [0,\beta] \cap [\beta - a/(a+b), b/(a+b)]$, let $X_{\beta,x}$ be given by (23). As in (i), we can show that the function $h(\beta,x) = \mathcal{D}(X_{\beta,x})$ is l.s.c. Let a sequence $\left\{(\beta_n, x_n)\right\}_{n\in\mathbb{N}}$ converge to (β,x) for some arbitrary $\beta \in [0,1]$ and $x \in [0,\beta] \cap [\beta - a/(a+b), b/(a+b)]$. By virtue of [2, Lemma 4.2], we may assume that each

[5]Two r.v.'s $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ are *comonotone*, if there exists a set $A \subseteq \Omega$ such that $\mathbb{P}[A] = 1$ and $(X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \ge 0$ for all $\omega_1, \omega_2 \in A$.

X_{β_n,x_n} is comonotone with $X_{\beta,x}$. In this case, (23) implies that X_{β_n,x_n} converges to $X_{\beta,x}$ a.s. Since the sequence X_{β_n,x_n} is uniformly bounded, this implies convergence in $\mathcal{L}^p(\Omega)$. Thus, lower semicontinuity D4 implies that $h(\beta,x) = \mathcal{D}(X_{\beta,x}) \leq \liminf_{n\to\infty} \mathcal{D}(X_{\beta_n,x_n}) = \liminf_{n\to\infty} h(\beta_n,x_n)$. In turn, lower semicontinuity of $h(\beta,x)$ implies that $u_{\mathcal{D}}(\beta) = \min_x h(\beta,x)$ is l.s.c., and for every β, there exists x such that $u_{\mathcal{D}}(\beta) = h(\beta,x)$. □

Proposition 9 proves that for an arbitrary law invariant deviation measure all the assumptions in Propositions 6 and 7 hold. Thus, the developed approach can be used for constructing a generalized Chebyshev inequality for any law invariant deviation measure.

4.3 Chebyshev Inequalities for Special Distributions

It is known that the ordinary one-sided Chebyshev inequality can be improved for symmetric distributions. In this section, we apply the developed optimization technique to derive improved generalized one-sided and two-sided Chebyshev inequalities, if we know that distribution belongs to some class, such as symmetric, log-concave, etc. To construct two-sided and one-sided improved Chebyshev inequalities, we solve problems (14) and (19) with the corresponding set V.

Let us first consider the case when V is the set of all r.v.'s with symmetric distribution, i.e. X is such that $F_X(-x) = 1 - F_X(x)$ for almost all $x \in \mathbb{R}$.

Proposition 10 *Let V be the set of symmetric r.v.'s with zero mean.*

(a) For any $\beta \in [0,1]$, a reduced set for the problem (14) *is a one-parametric family $U_c = \{X_x | \max\{0, \beta - 1/2\} \leq x \leq \beta/2\}$, where*

$$X_x = \begin{cases} -\max\{a,b\} & \textit{with probability} \quad x, \\ -\min\{a,b\} & \textit{with probability} \quad \beta - 2x, \\ 0 & \textit{with probability} \quad 1 - 2(\beta - x), \\ \min\{a,b\} & \textit{with probability} \quad \beta - 2x, \\ \max\{a,b\} & \textit{with probability} \quad x, \end{cases} \tag{32}$$

(b) In the problem (19), feasible set is empty for $\beta > 1/2$; otherwise a reduced set for the problem is the set U_c consisting of a single r.v.

$$X^\star = \begin{cases} -a & \textit{with probability} \quad \beta, \\ 0 & \textit{with probability} \quad 1 - 2\beta, \\ a & \textit{with probability} \quad \beta. \end{cases} \tag{33}$$

Proof. (a) The condition $X \in C_\beta$ in (14) implies that X is symmetric, $EX = 0$, and $\mathbb{P}[X \leq -a \text{ or } X \geq b] \geq \beta$. It is straightforward to verify that $U_c \subseteq C_\beta$. We prove first that $X_x \succcurlyeq_c X$ for any r.v. $X \in C_\beta$ and some x, and then apply Proposition 3. Without loss of generality, we can assume that $a \geq b$. Then $\mathbb{P}[X \geq b] \leq 1/2$ for $X \in C_\beta$ and $c = -\max\{a,b\} = -a$, and thus, $\mathbb{P}[X \leq c] = \mathbb{P}[X \leq -a] = \mathbb{P}[X \leq -a \text{ or } X \geq b] - \mathbb{P}[X \geq b] \geq \beta - 1/2$. Consequently, $\max\{0, \beta - 1/2\} \leq x \leq \beta/2$ for $x = \min\{\beta/2, \mathbb{P}[X \leq c]\}$, i.e., $X_x \in U_c$. Let $F(t)$ and $F_*(t)$ be the CDFs of X and X_x, respectively. Then $F(-a) = \mathbb{P}[X \leq -a] \geq x = F_*(-a)$. Also, $F(-b) = \mathbb{P}[X \geq b] \geq \beta - \mathbb{P}[X \leq -a]$ and $F(-b) = \frac{1}{2}\mathbb{P}[X \leq -b \text{ or } X \geq b] \geq \frac{1}{2}\mathbb{P}[X \leq -a \text{ or } X \geq b] \geq \beta/2$, which imply that $F(-b) \geq \beta - x = F_*(-b)$. Consequently, $F_*(t) \leq F(t)$ for $t < 0$, and, using the symmetry condition, we obtain $F_*(t) \geq F(t)$ for $t \geq 0$. Thus, based on Proposition 4, we conclude that $X_x \succcurlyeq_c X$. It is left to apply Proposition 3.

(b) It is straightforward to verify that $X^\star$ given by (33) satisfies $X^\star \in C_\beta$ in (19), and that $X^\star \succcurlyeq_c X$ for any $X \in C_\beta$. It is left to apply Proposition 3. □

Proposition 10 constructs the reduced sets for the problems (14) and (19) for the r.v.'s with symmetric distributions. As in Proposition 9, we can show that the solutions $u_{\mathcal{D}}(\beta)$ and $u^{-}_{\mathcal{D}}(\beta)$ to these problems are l.s.c. and can be represented in an explicit form for any law invariant deviation measure $\mathcal{D}$. This result paves the way for constructing two-sided and one-sided Chebyshev inequalities for the r.v.'s with symmetric distributions. We illustrate this approach for lower semideviation and CVaR deviation.

Example 6 (lower semideviation) *For lower semideviation $\sigma^2_-(X) = \|[X-EX]_-\|_2$ and $a > 0$, the two-sided Chebyshev inequality for the r.v.'s with symmetric distributions takes the form*

$$\mathbb{P}[|X-EX| \geq a] \leq \frac{2\sigma_-(X)^2}{a^2}. \tag{34}$$

Example 7 (conditional value-at-risk deviation) *For CVaR deviation $\mathcal{D}(X) = \mathrm{CVaR}^{\Delta}_{\alpha}(X)$ and $a > 0$, the two-sided and one-sided Chebyshev inequality for the r.v.'s with symmetric distributions are given by*

$$\mathbb{P}[|X-EX| \geq a] \leq \begin{cases} \dfrac{2\alpha \mathrm{CVaR}^{\Delta}_{\alpha}(X)}{a}, & \dfrac{\mathrm{CVaR}^{\Delta}_{\alpha}(X)}{a} < \dfrac{\min\{\alpha, 1-\alpha\}}{\alpha}, \\ 1, & \dfrac{\mathrm{CVaR}^{\Delta}_{\alpha}(X)}{a} \geq \dfrac{\min\{\alpha, 1-\alpha\}}{\alpha}, \end{cases} \tag{35}$$

$$\mathbb{P}[X \leq EX - a] \leq \begin{cases} \dfrac{\alpha \mathrm{CVaR}^{\Delta}_{\alpha}(X)}{a}, & \dfrac{\mathrm{CVaR}^{\Delta}_{\alpha}(X)}{a} < \dfrac{\min\{\alpha, 1-\alpha\}}{\alpha}, \\ \dfrac{1}{2}, & \dfrac{\mathrm{CVaR}^{\Delta}_{\alpha}(X)}{a} \geq \dfrac{\min\{\alpha, 1-\alpha\}}{\alpha}. \end{cases} \tag{36}$$

The estimates for the probabilities in (34)–(36) are tighter than those (general ones) derived in Examples 3 and 4.

Next we derive the one-sided Chebyshev inequality for the r.v.'s with log-concave density. This means that $X \in V$ if and only if X has the density $f_X(x)$, for which $\log f_X(x)$ is a concave function.

Proposition 11 *Let V be the set of r.v.'s X with $EX = 0$ and a log-concave density, and let $a > 0$. Then for the problem (19), a reduced set U_c can be chosen as the set of r.v.'s X which satisfy the conditions: (i) $EX = 0$, (ii) $\mathbb{P}[X \leq -a] \geq \beta$, and (iii) $1/q'_X(\alpha)$ is a linear function of α, where $q_X(\alpha)$ is the α-quantile of X.*

Proof. Let us choose some X with a log-concave density such that $EX = 0$ and $\mathbb{P}[X \leq -a] \geq \beta$, or $q_X(\beta) \leq -a$. Since the density of X is log concave, the function $g_X(\alpha) = 1/q'_X(\alpha)$ is concave. For $k \in [g_X(\beta)/(\beta - 1), g'_X(\beta-)]$, let

$$g_k(\alpha) = \begin{cases} g_X(\alpha), & 0 < \alpha \leq \beta, \\ g_X(\beta) + k(\alpha - \beta), & \beta \leq \alpha < 1, \end{cases}$$

and let Y_k be an r.v. such that $q_{Y_k}(\alpha) = q_X(\alpha)$ for $\alpha < \beta$ and such that $1/q'_{Y_k}(\alpha) = g_k(\alpha)$. For $k = g'_X(\beta-)$ and all $\alpha \in (0,1)$, $g_k(\alpha) \geq g_X(\alpha)$, whence $q_{Y_k}(\alpha) \leq q_X(\alpha)$, and therefore, $EY_k \leq EX$. On the other hand, for $k = g_X(\beta)(\beta - 1)$, concavity of $g_X(\alpha)$ implies $g_k(\alpha) \leq g_X(\alpha)$ for all $\alpha \in (0,1)$, which implies $q_{Y_k}(\alpha) \geq q_X(\alpha)$ and therefore, $EY_k \geq EX$. Thus, there exists some k_0 such that for the corresponding Y_0, we have $EY_0 = EX$. Now concavity of $g_X(\alpha)$ implies that there exists some $\alpha_0 \geq \beta$ such that $g_{k_0}(\alpha) \leq g_X(\alpha)$ for all $\alpha \leq \alpha_0$ and $g_{k_0}(\alpha) \geq g_X(\alpha)$ for all $\alpha \geq \alpha_0$. This implies that $q'_{Y_0}(\alpha) \geq q'_X(\alpha)$ for all $\alpha \leq \alpha_0$ and $q'_{Y_0}(\alpha) \leq q'_X(\alpha)$ for all $\alpha \leq \alpha_0$. This, together with $q_{Y_0}(\alpha) = q_X(\alpha)$ for $\alpha < \beta$, and $EY_0 = EX$ implies that there exists some $\alpha_1 \geq \alpha_0$, such that $q_{Y_0}(\alpha) \geq q_X(\alpha)$ for all $\alpha \leq \alpha_1$ and $q_{Y_0}(\alpha) \leq q_X(\alpha)$ for all $\alpha \geq \alpha_1$. Thus, based on Proposition 4, we conclude that $Y_0 \succcurlyeq_c X$.

Now let $g^\star(\alpha) = g(\beta) + k_0(\alpha - \beta)$, and let Z_0 be an r.v. such that $1/q'_{Z_0}(\alpha) = g^\star(\alpha)$ and $EZ_0 = EY_0$. Then $g^\star(\alpha) \geq g_{k_0}(\alpha)$ for all $\alpha \leq \beta$, and $g_{k_0}(\alpha) = g_X(\alpha)$ for all $\alpha \geq \beta$. This implies $q'_{Z_0}(\alpha) \leq q'_{Y_0}(\alpha)$ for all $\alpha \leq \beta$

and $q'_{Z_0}(\alpha) = q'_{Y_0}(\alpha)$ for all $\alpha \geq \beta$. Thus, the condition $EZ_0 = EY_0$ implies $q_{Z_0}(\beta) \leq q_{Y_0}(\beta) = q_X(\beta) \leq -a$, so that the probability constraint is satisfied for Z_0. This implies that $Z_0 \in U_c$. On the other hand, the condition $q'_{Z_0}(\alpha) \leq q'_{Y_0}(\alpha)$ for all $\alpha \in (0,1)$ implies that there exists some α^*, such that $q_{Z_0}(\alpha) \geq q_{Y_0}(\alpha)$ for all $\alpha \leq \alpha^*$ and $q_{Z_0}(\alpha) \leq q_{Y_0}(\alpha)$ for all $\alpha \geq \alpha^*$. By Proposition 4, the last conditions together with $EZ_0 = EY_0$ guarantee that $Z_0 \succcurlyeq_c Y_0 \succcurlyeq_c X$, and the proof is finished. □

Now Algorithm 2 can be used to construct the one-sided Chebyshev inequality for r.v.'s with log-concave distributions for any law invariant deviation measure. Indeed, Proposition 11 reduces (21) to a one-dimensional optimization problem, whose solution $u^-_{\mathcal{D}}(\beta)$ can be found numerically, and then the function inverse to $u^-_{\mathcal{D}}(\beta)$ provides the estimate for the probability in question.

Example 8 (standard deviation) *The one-sided Chebyshev inequality for standard deviation* σ *and the r.v.'s with log-concave distributions takes the form*

$$\mathbb{P}[X - EX \leq -a] \leq \phi(\sigma(X)/a), \quad a > 0, \tag{37}$$

where the function $\phi(t)$ *is calculated numerically and is shown on Figure 1.*

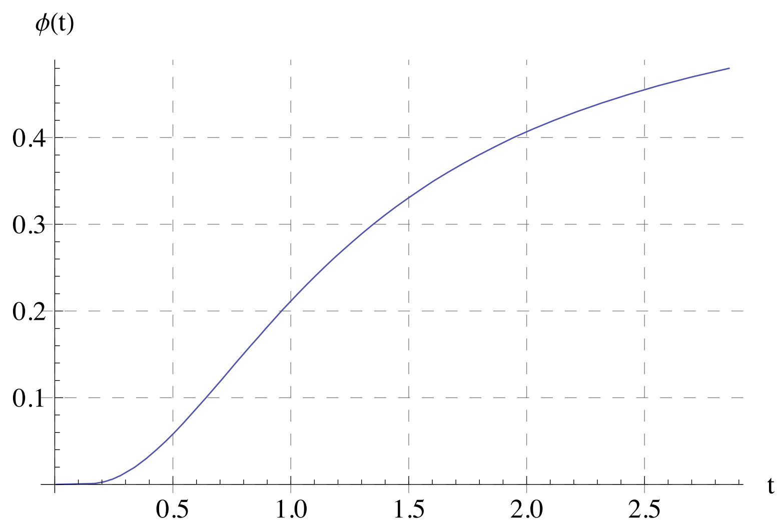

Figure 1: The function $\phi(t)$ entering the one-sided Chebyshev inequality (37) with standard deviation for the r.v.'s with log-concave distributions.

Obviously, for the class of r.v.'s with log-concave distributions, including uniform distribution, normal distribution, exponential distribution, Gamma distribution $f_X(x) = x^{m-1}\theta^m e^{-x\theta}/\Gamma(m)$ with $m > 1$, beta-distribution $f_X(x) = x^{a-1}(1-x)^{b-1}/B(a,b)$, $x \in (0,1)$ with $a \geq 1$, $b \geq 1$, etc. (see [1]), the Chebyshev inequality (37) significantly improves the estimate for the probability in question compared to the ordinary one-sided Chebyshev inequality $\mathbb{P}[X - EX \leq -a] \leq \sigma(X)^2/(\sigma(X)^2 + a^2)$ (the reader may compare the function $\phi(t)$ in Figure 1 with $t^2/(t^2+1)$, corresponding to $\sigma(X)$).

Similarly, using Proposition 11, we can construct Chebyshev inequalities for the r.v.'s with log-concave distributions for an arbitrary law invariant deviation measure. Also, the suggested approach (Algorithms 1 and 2) encompasses constructing two-sided and one-sided versions of Chebyshev inequalities for r.v.'s with distributions from other families.

5 Kolmogorov Inequality with Law Invariant Deviation Measures

Let a sequence of r.v.'s $S_1, S_2, \ldots, S_n$ be a discrete-time martingale, i.e., $E[S_i \mid S_1, \ldots, S_{i-1}] = S_{i-1}$ for $i = 2, \ldots, n$. In particular, if $X_1, X_2, \ldots, X_n$ are independent r.v.'s with zero mean, the sum $S_k = \sum_{i=1}^{k} X_i$, $k = 1, \ldots, n$ is a martingale. Suppose that $ES_1 = 0$. The classical Kolmogorov inequality for martingales states that

$$\mathbb{P}\big[\max_{1\le k\le n} |S_k| \ge \lambda\big] \le \frac{\sigma^2(S_n)}{\lambda^2}. \tag{38}$$

It estimates the probability of large deviations of $S_1, S_2, \ldots, S_n$ in terms of standard deviation of S_n. The same inequality holds for continuous-time martingales $S(t), t \in [t_0, t_1]$

$$\mathbb{P}\big[\max_{t} |S(t)| \ge \lambda\big] \le \frac{\sigma^2(S(t_1))}{\lambda^2}. \tag{39}$$

5.1 Generalized Kolmogorov Inequality

The following proposition generalizes (38) and (39) for an arbitrary law invariant deviation measure $\mathcal{D}$.

Proposition 12 *Let $S_1, S_2, \ldots, S_n$ be a discrete-time martingale such that $ES_1 = 0$, let $S(t)$, $t \in [t_0, t_1]$ be a continuous-time martingale with continuous sample paths and $E[S(t_0)] = 0$, and let $\lambda > 0$. Let $\mathcal{D}$ be a law invariant deviation measure, and let function $g_{\mathcal{D}}(d)$ be nondecreasing and such that the generalized Chebyshev inequality (11) with $a = b = \lambda$ holds for every r.v. X. Then the generalized Kolmogorov inequalities are determined by*

$$\mathbb{P}\big[\max_{1\le k\le n} |S_k| \ge \lambda\big] \le g_{\mathcal{D}}\big(\mathcal{D}(S_n)\big), \qquad \mathbb{P}\big[\max_{t} |S(t)| \ge \lambda\big] \le g_{\mathcal{D}}\big(\mathcal{D}(S(t_1))\big). \tag{40}$$

Similarly, if for nondecreasing function $g_{\mathcal{D}}^{-}(d)$, the one-sided Chebyshev inequality (18) *holds, then*

$$\mathbb{P}\big[\min_{1\le k\le n} S_k \le -a\big] \le g_{\mathcal{D}}^{-}\big(\mathcal{D}(S_n)\big), \qquad \mathbb{P}\big[\min_{t} S(t) \le -a\big] \le g_{\mathcal{D}}^{-}\big(\mathcal{D}(S(t_1))\big). \tag{41}$$

Proof. Let $N = \min\{k \mid |S_k| \ge \lambda\} \wedge n$ be the smallest index such that $|S_N| \ge \lambda$, or, if $\max\limits_{1\le k\le n-1} |S_k| < \lambda$, then $N = n$. Then N is an r.v., and the event $\{N = k\}$ depends only on values $S_1, \ldots, S_k$. This r.v. is called a *stopping time* with respect to the sequence $S_1, S_2, \ldots, S_n$. Since $N \le n$, Doob's optional sampling theorem states that $E[S_n | S_N] = S_N$, which implies $S_N \succcurlyeq_c S_n$. By Proposition 1, $\mathcal{D}(S_n) \ge \mathcal{D}(S_N)$. Then from (11), it follows that

$$\mathbb{P}\big[\max_{1\le k\le n} |S_k| \ge \lambda\big] = \mathbb{P}[|S_N| \ge \lambda] \le g_{\mathcal{D}}\big(\mathcal{D}(S_N)\big) \le g_{\mathcal{D}}\big(\mathcal{D}(S_n)\big).$$

Similarly, $S_{N^-} \succcurlyeq_c S_n$ for the stopping time $N^- = \min\{k \mid S_k \le -a\} \wedge n$, and thus, $\mathcal{D}(S_n) \ge \mathcal{D}(S_{N^-})$. Consequently, (18) implies

$$\mathbb{P}\big[\min_{1\le k\le n} S_k \le -a\big] = \mathbb{P}[S_{N^-} \le -a] \le g_{\mathcal{D}}^{-}\big(\mathcal{D}(S_N)\big) \le g_{\mathcal{D}}^{-}\big(\mathcal{D}(S_n)\big).$$

For a continuous-time martingale with stopping times $T = \min\{t \mid |S(t)| \ge \lambda\} \wedge t_1$ and $T^- = \min\{t \mid S(t) \le -a\} \wedge t_1$, the Kolmogorov inequalities are proved similarly. □

In Section 4, we have developed the algorithm for constructing the one-sided and two-sided generalized Chebyshev inequalities for an arbitrary law invariant deviation measure $\mathcal{D}$ (formulas (18) and (11), respectively). Since the obtained functions $g_{\mathcal{D}}^{-}(d)$ and $g_{\mathcal{D}}(d)$ are nondecreasing, they can be used in the Kolmogorov inequalities (41) and (40) with this $\mathcal{D}$.

Observe that if (11) reduces to an equality for some r.v. X_0, so does (40) for $S_1 = S_2 = S_3 = \ldots = S_n = X_0 - EX_0$. This means that (40) cannot be tightened provided that $g_{\mathcal{D}}(d)$ in (11) satisfies the requirement (R2).

The following examples are similar to those in Section 4.

Example 9 (two-sided Kolmogorov inequalities) *Let $S_1, S_2, \ldots, S_n$ be a martingale with $ES_1 = 0$. Then, for MAD, lower semideviation, CVaR deviation and the deviation measure in Example 5, the Kolmogorov inequalities are determined by*

(i) MAD*:*

$$\mathbb{P}\big[\max_{1\le k\le n} |S_k| \ge \lambda\big] \le \frac{\mathrm{MAD}(S_n)}{\lambda}. \tag{42}$$

(ii) lower semideviation*:*

$$\mathbb{P}\big[\max_{1\le k\le n} |S_k| \ge \lambda\big] \le \begin{cases} \frac{16}{9}d + \frac{1}{9}, & d \ge \frac{1}{20}, \\ \frac{1}{2}\left(\sqrt{d^2+4d} - d\right), & d < \frac{1}{20}, \end{cases} \tag{43}$$

where $d = \sigma_-^2(S_n)/\lambda^2$.

(iii) CVaR deviation *for $\alpha \le 1/2$:*

$$\mathbb{P}\big[\max_{1\le k\le n} |S_k| \ge \lambda\big] \quad \le \begin{cases} \frac{d}{1+d}, & d \le \frac{1/2-\alpha}{1/2+\alpha}, \\ \frac{1}{2} + \alpha d - \sqrt{\alpha(1-d)(1-\alpha(1+d))}, & d \in \left[\frac{1/2-\alpha}{1/2+\alpha}, 1\right), \\ 1, & d \ge 1. \end{cases} \tag{44}$$

where $d = \mathrm{CVaR}_\alpha^\Delta(S_n)/\lambda$. For $\alpha \ge 1/2$, the Kolmogorov inequality follows from (44) *and the relation* $\alpha \cdot \mathrm{CVaR}_\alpha^\Delta(S_n) \equiv (1-\alpha)\cdot \mathrm{CVaR}_{1-\alpha}^\Delta(-S_n)$.

(iv) $\mathcal{D}_\alpha(X) = \max\left\{\mathrm{CVaR}_\alpha^\Delta(X), \mathrm{CVaR}_\alpha^\Delta(-X)\right\}$*:*

$$\mathbb{P}\big[\max_{1\le k\le n} |S_k| \ge \lambda\big] \le \begin{cases} \frac{2\alpha\mathcal{D}_\alpha(S_n)}{\lambda}, & \mathcal{D}_\alpha(S_n) < \lambda, \\ 1, & \mathcal{D}_\alpha(S_n) \ge \lambda. \end{cases} \tag{45}$$

Detail. These inequalities follow from (40) and the Chebyshev inequalities in Examples 2–5. □

Similarly, the one-sided Chebyshev inequalities in Examples 2–5 result in the corresponding one-sided Kolmogorov inequalities. The same inequalities hold for continuous-time martingales.

5.2 Application of Generalized Kolmogorov Inequality

Application of the generalized two-sided Kolmogorov inequality can be illustrated by the following example. Let $S(t)$ be a discrete-time martingale with $t = 0, 1, 2, \ldots$ such that $S(0) = 0$ and the increments $X_t = S(t) - S(t-1)$ are independent and identically distributed r.v.'s with mean 0 and finite variance σ^2. The process $S(t)$ can be used to model various real-life processes, such as the logarithm of an exchange rate of two currencies or the logarithm of the rate of return of a stock.

Suppose $S(t)$ is the logarithm of the exchange rate of two currencies, and we are interested in estimating $\mathbb{P}[|S(t)| < \lambda$ for all $t \le n]$. For illustrative purposes, let $\lambda = \sigma\sqrt{n}$. When n is sufficiently large, we can assume the distribution of $S(n)$ to be approximately normal, i.e. $S(n) \sim N\left(0, n\sigma^2\right)$.

From the Kolmogorov inequality (38) with standard deviation, we obtain

$$\mathbb{P}\big[\max_{1\le t\le n} |S(t)| \ge \lambda\big] \le \frac{\sigma^2(S(n))}{\lambda^2} = 1,$$

which, in fact, provides no information.

Similarly, from the Kolmogorov inequality (42) with MAD, we have

$$\mathbb{P}\big[\max_{1\le t\le n}|S(t)|\ge\lambda\big]\le\frac{\mathrm{MAD}(S(n))}{\lambda}=\sqrt{\frac{2}{\pi}}<0.8.$$

Indeed, since $S(n)/\lambda\sim N(0,1)$, we can write

$$\mathrm{MAD}(S(n)/\lambda)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}|x|\mathrm{e}^{-x^2/2}dx=\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}x\,\mathrm{e}^{-x^2/2}dx=\sqrt{\frac{2}{\pi}}.$$

Thus, in this case, MAD is more informative than standard deviation.

The question arises whether we can obtain a better estimate, using another deviation measure. The next example addresses this issue.

Example 10 *The Kolmogorov inequality* (45) *with the deviation measure*

$$\mathcal{D}_\alpha(X)=\max\left\{\mathrm{CVaR}_\alpha^\Delta(X),\mathrm{CVaR}_\alpha^\Delta(-X)\right\}$$

results in the estimate

$$\mathbb{P}\big[\max_{1\le t\le n}|S(t)|\ge\lambda\big]\le 2\alpha_0\approx 0.764, \tag{46}$$

where α_0 *is such that* $\mathrm{CVaR}_{\alpha_0}^\Delta(Z)=1$ *for* $Z\sim N(0,1)$. *The estimate (46) cannot be improved by using another deviation measure in the two-sided Kolmogorov inequality.*

Detail. To obtain a non-trivial estimate in (45), we require $\mathcal{D}_\alpha(S(n))<\lambda$. Equivalently, $\mathrm{CVaR}_\alpha^\Delta(S(n)/\lambda)<1$, or $\alpha>\alpha_0$, where α_0 is the root of the equation $\mathrm{CVaR}_\alpha^\Delta(S(n)/\lambda)=1$, which reduces to

$$\mathrm{CVaR}_{\alpha_0}^\Delta(S(n)/\lambda)\equiv-\frac{1}{\alpha_0\sqrt{2\pi}}\int_{-\infty}^{\Phi^{-1}(\alpha_0)}x\,\mathrm{e}^{-x^2/2}dx=-\frac{\mathrm{e}^{-\left(\Phi^{-1}(\alpha_0)\right)^2/2}}{\alpha_0\sqrt{2\pi}}=1,$$

where Φ is the CDF of the standard normal distribution. From this equation, we find $\alpha_0\approx 0.382$ numerically.

Now, from the Kolmogorov inequality (45) for every $\alpha>\alpha_0$, it follows that

$$\mathbb{P}\big[\max_{1\le t\le n}|S(t)|\ge\lambda\big]\le\frac{2\alpha}{\lambda}\max\left\{\mathrm{CVaR}_\alpha^\Delta(S(n)),\mathrm{CVaR}_\alpha^\Delta(-S(n))\right\},$$

which for $\alpha\to\alpha_0$ and $\lambda=\sigma\sqrt{n}$ reduces to

$$\mathbb{P}\big[\max_{1\le t\le n}|S(t)|\ge\lambda\big]\le\frac{2\alpha_0}{\lambda}\sigma\sqrt{n}=2\alpha_0\approx 0.764.$$

Finally, we show that no other law invariant deviation measure $\mathcal{D}$ provides better estimate. Let Y be an r.v. given by

$$Y=\begin{cases}-\sigma\sqrt{n} & \text{with probability} \quad \alpha_0,\\ 0 & \text{with probability} \quad 1-2\alpha_0,\\ \sigma\sqrt{n} & \text{with probability} \quad \alpha_0.\end{cases} \tag{47}$$

Then substituting Y into the Chebyshev inequality (17) for a deviation measure $\mathcal{D}$, we obtain $g_{\mathcal{D}}(\mathcal{D}(Y))\ge P(|Y|\ge\sigma\sqrt{n})=2\alpha_0$. On the other hand, Proposition 2 implies that $\mathcal{D}(Y)\le\mathcal{D}(S(n))$. Then $g_{\mathcal{D}}(\mathcal{D}(S(n)))\ge g_{\mathcal{D}}(\mathcal{D}(Y))\ge 2\alpha_0$, and thus, the estimate (46) cannot be improved by using another deviation measure in the Kolmogorov inequality. □

This approach can be readily extended to the case when X_i are not identically distributed, or when $S(n)$ is not normally distributed. We only need to assume that either the distribution of $S(n)$ is known or α_1 and α_2 such that $\mathrm{CVaR}^{\Delta}_{\alpha_1}(S(n)) = \lambda$ and $\mathrm{CVaR}^{\Delta}_{\alpha_2}(-S(n)) = \lambda$ are available.

As another illustration, suppose that $X(t)$ is the rate of return of a portfolio at time t with $S(t) = \ln(X(t)+1)$ to be a martingale process. If t_0 is the present time moment then the generalized Kolmogorov inequality (41) estimates the probability that $X(t) \leq -c$ at some $t \in [t_0,t_1]$ for some $c \in (0,1)$:

$$\mathbb{P}\big[\min_t X(t) \leq -c\big] = \mathbb{P}\big[\min_t S(t) \leq \ln(1-c)\big] \leq g^{-}_{\mathcal{D}}(\mathcal{D}(S(t_1))).$$

In particular, if $\ln(X(t_1)+1) \sim N(0,\sigma_0^2)$ and $c = 1-\mathrm{e}^{-\sigma_0}$, then $S(t_1) \sim N(0,\sigma_0^2)$ and with the standard deviation, the above inequality reduces to

$$\mathbb{P}\big[\min_t X(t) \leq -c\big] \leq \frac{\sigma^2(S(t_1))}{\sigma^2(S(t_1))+\sigma_0^2} = \frac{1}{2}.$$

The next example shows that this estimate can be improved.

Example 11 *The one-sided Kolmogorov inequality with CVaR deviation provides*

$$\mathbb{P}\big[\min_t X(t) \leq -c\big] \leq \alpha_0 \approx 0.382, \tag{48}$$

where α_0 *solves* $\mathrm{CVaR}^{\Delta}_{\alpha_0}(Z) = 1$ *for* $Z \sim N(0,1)$. *The estimate (48) cannot be improved by using another deviation measure in the one-sided Kolmogorov inequality.*

Detail. By Proposition 12, the right-hand side in the above Kolmogorov inequality coincides with the one in the Chebyshev inequality (30). To obtain a non-trivial estimate in (30), we require $\mathrm{CVaR}^{\Delta}_{\alpha}(S(t_1)) < -\ln(1-c) = \sigma_0$. Equivalently, $\alpha > \alpha_0$, where α_0 is the root of the equation $\mathrm{CVaR}^{\Delta}_{\alpha}(S(t_1)) = \sigma_0$, which reduces to $\mathrm{CVaR}^{\Delta}_{\alpha}(Z) = 1$. From this equation, we find $\alpha_0 \approx 0.382$ numerically.

From the Kolmogorov inequality for CVaR deviation for $\alpha > \alpha_0$ and $a = -\ln(1-c)$, we have

$$\mathbb{P}\big[\min_t X(t) \leq -c\big] \leq \frac{\alpha \cdot \mathrm{CVaR}^{\Delta}_{\alpha}(S(t_1))}{a+\alpha(\mathrm{CVaR}^{\Delta}_{\alpha}(S(t_1))-a)},$$

which for $\alpha \to \alpha_0$ and $c = 1-\mathrm{e}^{-\sigma_0}$ reduces to

$$\mathbb{P}\big[\min_t X(t) \leq -c\big] \leq \frac{\alpha_0 \cdot \sigma_0}{\sigma_0+\alpha_0(\sigma_0-\sigma_0)} = \alpha_0 \approx 0.382.$$

The fact that no other law invariant deviation measure $\mathcal{D}$ provides a better estimate can be shown similarly to that in Example 10. □

6 Conclusions

We have observed that on an atomless probability space every law invariant deviation measure is consistent with concave ordering, i.e. $X \succcurlyeq_c Y$ implies $\mathcal{D}(X) \leq \mathcal{D}(Y)$. An immediate consequence of this fact is that decision making with law invariant deviation measures over random outcomes with equal means conforms with risk-averse preferences. However, it also implies that Rao-Blackwell theorem holds for an arbitrary law invariant deviation measure and that minimization of law invariant deviation measures can be reformulated as minimization over sets of undominated r.v.'s with respect to concave ordering. Using the latter fact, we have developed an approach for reducing minimization of law invariant deviation measures with certain chance constraints to finite parameter optimization problems. We have applied this approach to obtain Chebyshev and Kolmogorov inequalities with an arbitrary law invariant deviation measure in two cases: (i) for r.v.'s with all possible distributions on $\mathcal{L}^p(\Omega)$, and (ii) for the r.v.'s with distributions from a given set, in particular for the r.v.'s with symmetric distributions. As an illustration, we have derived Chebyshev and Kolmogorov inequalities for mean-absolute deviation, lower semideviation, and conditional value-at-risk deviation. Also, we have demonstrated that in the example of a discrete-time martingale with an asymptotically normal distribution, the Kolmogorov inequality with deviation measures other than standard deviations, e.g., mean absolute deviation, provides better estimates for the probability in question.

References

[1] Bergstrom, T., Bagnoli, M., Log-concave probability and its applications. Economic Theory **26**, 445–469 (2005)

[2] Dana, R.-A.: A representation result for concave Schur-concave functions. Mathematical Finance **15**(4), 613–634 (2005)

[3] Follmer, H., Schied, A.: Stochastic finance, (2nd ed.). Berlin New York: de Gruyter 2004.

[4] Hanoch, G., Levy, H.: The efficiency analysis of choices involving risk. Rev. economic studies **36**, 335–346 (1969)

[5] Hogg, R.V., Craig, A., McKean, J.W.: Introduction to mathematical statistics, (6th ed.). New York: Prentice Hall 2004

[6] Kantorovich, L.V., Akilov, G.P.: Variational Methods for the Study of Nonlinear Operators. Holden-Day, Inc. 1964.

[7] Kurdila, A., Zabarankin, M.: Convex Functional Analysis. Series: Systems and Control: Foundations and Applications, Birkhauser, Switzerland, 2005.

[8] Levy, H.: Stochastic dominance. Boston-Dordrecht-London: Kluwer Academic Publisher 1998

[9] Markowitz, H.M.: Portfolio selection. The Journal of Finance **7**(1), 77–91 (1952)

[10] Rockafellar, R.T., Uryasev, S., Zabarankin, M.: Generalized deviations in risk analysis. Finance and Stochastics **10**(1), 51–74 (2006)

[11] Rockafellar, R.T., Uryasev, S., Zabarankin, M.: Optimality conditions in portfolio analysis with general deviation Measures. Mathematical Programming **108**(2-3), 515–540 (2006)

[12] Rockafellar, R.T., Uryasev, S., Zabarankin, M.: Master funds in portfolio analysis with general deviation measures. The Journal of Banking and Finance **30**(2), 743–77 (2006)

[13] Rockafellar, R.T., Uryasev, S., Zabarankin, M.: Equilibrium with investors using a diversity of deviation measures. The Journal of Banking and Finance **31**(11), 3251–3268 (2007)

[14] Rockafellar, R.T., Uryasev, S., and Zabarankin, M.: Risk tuning with generalized linear regression. Mathematics of Operations Research **33**(3), 712–729 (2008)

[15] Rockafellar, R.T., Uryasev, S., Zabarankin, M.: Deviation measures in risk analysis and optimization. Report 2002-7, ISE Dept., Univ. of Florida (2002)

[16] Roy, A.: Safety first and the holding of assets. Econometrica **20**, 431–449 (1952)

[17] Smith, J.: Generalized Chebyshev inequalities: Theory and applications in decision analysis. Operations Research **43**(5), 807–825 (1995)