Department of Economics and Management
University of Brescia
Italy

WORKING PAPER

Discrete Conditional Value-at-Risk

Carlo Filippi
Wlodzimierz Ogryczak
M. Grazia Speranza

WPDEM 2/2016

Via S. Faustino 74/b, 25122 Brescia – Italy
Tel 00390302988742 Fax 00390302988703
email: paqricerca.dem@unibs.it

Discrete Conditional Value-at-Risk

Carlo Filippi [1] *Wlodzimierz Ogryczak* [2] *M. Grazia Speranza* [3]

(1) *Department of Economics and Management, University of Brescia, Italy, carlo.filippi@unibs.it*

(2) *Institute of Control and Computation Engineering, Warsaw University of Technology, Poland, ogryczak@ia.pw.edu.pl*

(3) *Department of Economics and Management, University of Brescia, Italy, grazia.speranza@unibs.it*

Abstract

The Conditional Value-at-Risk (CVaR) has become a very popular concept to measure the risk of an investment. In fact, though originally proposed for adoption in a financial context, the concept has potential for a broad range of applicability. In this paper, we consider problems that are formulated as mixed integer linear programming (MILP) models and show that a discrete version of the CVaR, that we call Discrete CVaR (DCVaR), can be adopted. The DCVaR mediates between a conservative Minimax/Maximin criterion and an aggressive minimum cost/maximum profit criterion, in all cases where uncertainty or variability matters. We show that the Discrete CVaR satisfies properties that make it an attractive measure. In particular, the models resulting from the adoption of the Discrete CVaR remain MILP models. To illustrate the relevance of the proposed model, we apply the DCVaR measure to several instances of the multidimensional knapsack problem and the p-median problem.

Keywords: Conditional Value-at-Risk; discrete optimization; mixed integer linear programming.

1 Introduction

The trade-off between risk and return in financial investments has been deeply investigated, starting from Markowitz model Markowitz [1952], where risk was measured with the variance of the portfolio rate of return. In general, the risk of an investment is associated with the uncertainty on the future return. It is well known that, to avoid risk, one should invest in assets with stable (certain) rates of return and that large expected returns can be achieved only through investment in risky assets. Several measures have been proposed in finance to measure the uncertainty of the portfolio rate of return. The *Value-at-Risk* (VaR) and the *Conditional Value-at-Risk* (CVaR) have recently become very popular concepts. Given a specified probability level α, the α-VaR of a portfolio is the lowest amount ν such that, with probability α, the loss will not exceed v, whereas the α-CVaR is the conditional expectation of losses above that amount v (see Rockafellar and Uryasev [2000]). Whereas the VaR has some undesirable characteristics, the CVaR has all the desirable characteristics needed to be defined a coherent risk measure (see Artzner et al. [1999]). The CVaR is widely used in finance to measure and compare portfolios. The concept of CVaR is built upon the assumption that the rates of return of the assets follow a known probability distribution.

The CVaR can be used as an objective function that drives portfolio optimization. The value of the parameter α defines the trade-off between the risk related to the uncertainty of the returns and the quality of the solution, measured through the expected loss (that should be small) or its opposite, the expected return (that should be large). Large values of α indicate a strong risk aversion. When α is close to 1, the VaR is large and the CVaR, the average value of losses larger than the VaR, is small. On the opposite, for small values of α the VaR is small and the CVaR is large. Thus, in the former case, the investor cares about large losses only, while in the latter cares about much smaller losses (in the extreme case about all possible losses).

The control of uncertainty is vital in finance but is extremely important in many other areas. Often parameters of an optimization model are not known with certainty and deterministic models are built upon estimations. The coefficients of the objective function are usually recognized as the most relevant parameters. Several efforts are being made to model uncertainty. One of the most commonly adopted techniques is scenario generation Kaut and Wallace [2007]. Let us assume that S scenarios for the coefficients of the objective function of a Mixed Integer Linear Programming (MILP) model are generated. All scenarios are equally probable and we aim at being protected against a certain number, say k, of the worst scenarios, that is against a fraction k/S of the worst possible values of the objective function. In most discrete optimization problems, the objective is cost minimization. The situation is then similar to the case of a portfolio optimization where the focus is on loss minimization. In this context, a scenario is bad when the objective function is large. Thus, a sensible objective becomes the minimization of the average value of the objective function in the k worst scenarios. If we have a statistical percentage α, instead of a fraction k/S, we can consider the $\lceil (1-\alpha)S \rceil$ worst scenarios. One can see that this value is conceptually close to the CVaR. We call this measure *Discrete CVaR* (DCVaR). If $k = 1$, then $\alpha = 1 - 1/S$ and the problem becomes the minimization of the objective function in the worst scenario, that is, in the scenario where the objective function is largest. This corresponds to adopting the classical Wald's Minimax criterion Wald [1950]. If $k = S$, then $\alpha = 0$ and the problem becomes the minimization of the average value of the objective function over all the scenarios. In the former case we aim at protecting the decision against the worst possible scenario, in the latter we aim at optimizing the average value (we will call it payoff, in general) ignoring uncertainty. The value of $1-k/S$, or in general of α, sets the trade-off between protection against uncertainty and payoff of the objective function.

The Minimax, or Maximin for a maximization problem, models are traditionally used to protect the decision maker against risk intended as the worst possible outcome. Such models are very conservative and may cause a relevant worsening of the total or average payoff of the solution. In many applications, the outcome is determined by the realization of an uncertain event. However, there are Minimax/Maximin models that are completely deterministic. In the latter case, usually the decision maker aims at being protected against variability and the DCVaR concept can be used to model the trade-off between variability and payoff of a solution.

The goal of this paper is to show that the DCVaR can be adopted in problems that are formulated through MILP models to control the trade-off between the amount of uncertainty or variability of a solution and its payoff. In the rest of this introductory section, we first present some examples of potential application of the DCVaR. Then, we review the state of the art and, finally, we discuss in more detail the contribution of this paper and its content.

1.1 Examples

We consider some classical optimization problems that are formulated through MILP models and show how the issue of controlling uncertainty or variability may arise.

Uncertainty. Let us consider the multidimensional knapsack problem (MKP) Kellerer et al. [2004]. A set of n activities is given. A profit (for example, the return of the investment) p_j is associated with each activity $j,\ j=1,\dots,n$. A set of m resources are used by the activities, each available in a maximum amount b_i, $i=1,\dots,m$. Each activity j requires a quantity a_{ij} of resource i. The problem is to select a subset of the activities, with maximum total profit, while guaranteeing that the constraints on the maximum availability of the resources are satisfied. The problem is formulated as a binary linear program, with a binary variable x_j associated with each activity j. The classical knapsack problem is the MKP with one constraint only.

It is often the case that the returns of the investment in the activities are not known with precision. Combining market forecasts, historical data, opinions of experts, one may build up a set of scenarios describing the uncertainty. Activities with average high profit may have a very variable (very low in some scenarios and very high in others) profit. The goal becomes to control the trade-off between the average profit of the activities selected and the uncertainty associated with them.

Variability. Let us consider the p-median problem Hakimi [1964], where p locations have to be chosen among a set of m potential locations to serve n customers. The distance of customer j from location i is c_{ij}. The goal is to minimize the total (average) distance between customers and their closest facilities.

The optimal solution of the p-median problem may be such that some of the customers are very far from the closest facility with respect to others. Thus, the solution may be unfair and unsatisfactory. In fact, to take this concern into account, in the p-center problem Hakimi [1964] the maximum distance between a customer and its closest facility is minimized. This objective corresponds to a Minimax approach to the problem. Whereas the p-median pursues the efficiency objective, the p-center pursues the objective of service fairness and equitability among customers (see Kostreva and Ogryczak [1999], Ogryczak [2009]). A trade-off solution would be of interest.

1.2 Literature

The Wald's criterion has been extended to several different risk measures protecting against uncertainty in decision making. The variance, introduced by Markowitz in his classical earth-breaking model for portfolio optimization (see Markowitz [1952]), gives rise to a quadratic model. The classical measures, such as the variance, aim at capturing the dispersion around the mean and are minimized. Performance measures that still aim at protecting the investor against risk, while being maximized, are called safety measures. For an overview and classification of risk and safety measures we refer to Mansini et al. [2003] or Mansini et al. [2015].

Generalization of the Wald's criterion leads to the quantile shortfall risk measure. Such second order quantile risk measure was introduced in different ways by several authors (see Artzner et al. [1999], Embrechts et al. [1997], Ogryczak [1999], Ogryczak and Ruszczyński [2002], Rockafellar and Uryasev [2000]). The measure, now commonly called CVaR, after Rockafellar and Uryasev [2000], represents the mean shortfall at a specified confidence level. The CVaR at α level, or α-CVaR, is the expected return of the portfolio in the $1-\alpha$ worst cases and is also called expected shortfall, average value-at-risk, tail value-at-risk. In the finance and banking literature the α-quantile is usually called VaR. The CVaR evaluates the value (or risk) of an investment focusing on the less profitable outcomes. It is a coherent measure of financial portfolio risk, that is, it satisfies properties of a risk measure that capture the preferences of a rational investor Artzner et al. [1999]. The CVaR has become the most commonly adopted measure. It is theoretically appealing and allows an investor to express the risk aversion level through the value of α. Application of the CVaR concept to a discrete random variable has been considered in Benati [2004], where computational aspects are also analyzed.

Outside the financial literature, the CVaR has received very limited attention. The CVaR measure is used in Toumazis and Kwon [2013] to mitigate risk in routing hazardous materials, while in Sarin et al. [2014] as a measure for stochastic scheduling problems. Apart from the area of decisions under risk, tail means have been already applied to location problems Ogryczak and Zawadzki [2002], to fair resource allocations in networks Ogryczak and Śliwiński [2002], Ogryczak et al. [2014] as well as to planning aperture modulation for radiation therapy treatment Romeijn et al. [2006] and to statistical learning problems Takeda and Kanamori [2009]. In radiation therapy treatment planning problems the CVaR measures are used to set the dose volume constraints Romeijn et al. [2006].

In a simplified form, as the criterion of the k largest outcomes minimization, the DCVaR concept was considered in discrete optimization much earlier than the CVaR measure was introduced. In particular, the criterion of the k largest distances minimization was introduced in location problems in Slater [1978] as the so called k-centrum model. If $k = 1$ the model reduces to the standard p-center model while with $k = m$ it is equivalent to the classical p-median model. Early works on the concept Andreatta and Mason [1985a,b], Slater [1978] were focused on the case of the discrete single facility location on tree graphs. Later, Tamir Tamir [2001] has presented polynomial time algorithms for solving the multiple facility k-centrum on path and tree graphs, while Ogryczak and Tamir Ogryczak and Tamir [2003] have shown that the criterion can be modeled with an auxiliary set of simple linear inequalities, thus simplifying several k-centrum models. Computational aspects of related problems have been studied in Grygiel Grygiel [1981], Punnen Punnen [1992] and Punnen and Aneja [1996]. In machine scheduling the problem of minimizing the sum of the k, among n, maximum tardiness jobs on a single machine without preemption was shown to be polynomially solvable for any fixed value k Woeginger [1991]. The partial sum criterion was also considered for Steiner trees in graphs and shortest paths Duin and Volgenant [1997], Garfinkel et al. [2006] as well as for fair resource allocations in networks Ogryczak and Śliwiński [2002].

1.3 Contribution and structure

In this paper, we focus on optimization problems that can be formulated as Mixed Integer Linear Programming (MILP) models. We consider two different settings where one may find the optimization of the average or total payoff and the minimax/maximin approach unsatisfactory, the first being biased towards the payoff and in no way sensitive to uncertainty or variability and the second being biased in the opposite direction.

In the first setting, the coefficients of the objective function are uncertain and uncertainty is described through a finite set of possible scenarios. In the second setting, all input data is known but the variability of the payoffs (coefficients) in the objective function is a concern for the decision maker. We will refer to these two settings as the *uncertainty setting* and the *variability setting*, respectively. In fact, we will see that the two settings are identical from the abstract and formal point of view and are modeled in the same way. However, they describe quite different practical problems. The DCVaR is, in the uncertainty setting, the average payoff in the k worst scenarios, whereas it is the average of the k worst payoffs in the variability setting. We assume that all scenarios are equally probable and that the individual payoffs have identical importance.

The main goal of this paper is to unify the two above settings and to show the broad applicability of the DCVaR concept. We derive the DCVaR measure in a simple and self-contained way, relying only on LP theory and basic statistical concepts. We show that, if we interpret the individual payoffs as the only, equally probable, values of a discrete random variable, the DCVaR is equivalent to the classical CVaR, and that it has several desirable properties. We also show that, if we start with a mixed integer linear program, adopting the DVCaR leads to another mixed integer linear program. We complement the theoretical part with two examples of application, namely the MKP, as an example of uncertainty, and the p-center/p-median problem, as an example of variability. We formulate the models resulting from the adoption of the DCVaR and solve them for different values of k on benchmark instances.

The paper is structured as follows. In Section 2, we introduce the general model we study, that is a MILP problem, and the notation that allows us to capture in the same model the two settings. The DCVaR measure is presented in Section 3, together with the relations with the CVaR and desirable properties that are valid for the CVaR and make the DCVaR as attractive in a discrete setting as the CVaR in a continuous setting. In Section 4, we show how the adoption of the DCVaR leads to another mixed integer program. In Section 5, to illustrate the relevance of the proposed model, we apply the DCVaR measure to a couple of classic combinatorial problems. In Section 6, we briefly discuss the case of not equal weights. Finally, some conclusions are drawn in Section 7.

2 Unifying uncertainty and variability

We consider a general MILP model in minimization form:

$$\begin{aligned} \min \quad & c^\top x \\ \text{subject to} \quad & Ax = b \\ & x_B \in \{0,1\}^{|B|} \\ & x_N \geq 0, \end{aligned} \tag{1}$$

where c is a real n-vector, A is a real $(m \times n)$-matrix, b is a real m-vector, x is the n-vector of variables. The index set $\{1,2,\dots,n\}$ is partitioned in two subsets B and N where B contains the indices of the binary variables and N contains the indices of the continuous variables. If B is empty, the model is a plain LP model. Notice that general integer variables can always be modeled using binary variables. We will show later how the concepts can be adapted to the maximization form

For the sake of simplicity, we assume that the feasible set

$$X = \{x \in \mathbb{R}^d : Ax = b, x_B \in \{0,1\}^{|B|}, x_N \geq 0\} \tag{2}$$

is nonempty and bounded.

We assume that c is the average over a finite number S of *states*. In the uncertainty setting, a state ℓ represents a *scenario*, whereas in the variability setting it represents the ℓ-th *agent*, for example a customer or a period of a discretized time horizon. We call c^ℓ the ℓ-th *realization* of c, $\ell = 1,\dots,S$. Then,

$$c = \frac{1}{S}\sum_{\ell=1}^{S} c^\ell.$$

We assume that the realizations have equal weights. This assumption is often satisfied in practical applications and allows for better properties. Then, we define

$$y_\ell(x) = (c^\ell)^\top x,$$

with $\ell = 1,\dots,S$. Note that $y_\ell(x)$ represents the value of the objective function of problem (1) when the ℓ-th state is considered. For any $x \in X$, $y_\ell(x)$ is called the ℓ-th *outcome* of x.

In (1) the average outcome is optimized. In the uncertainty setting, the average is taken over all the scenarios. In the variability setting, it is taken over all the agents. We denote the average outcome of a solution x as

$$M(x) = c^\top x = \frac{1}{S}\sum_{\ell=1}^{S}(c^\ell)^\top x = \frac{1}{S}\sum_{\ell=1}^{S} y_\ell(x).$$

We call the model where $M(x)$ is minimized the *Minavg model*. Here, we search for a solution that has the best possible average outcome. This corresponds to adopting a Bayesian approach (see Berger [1985]). An optimal solution x^{avg} of the Minavg model is such that

$$M(x^{\text{avg}}) = \min\{M(x) : x \in X\}.$$

As X is nonempty and bounded, the above expression is well defined.

If the Minimax criterion is adopted, then the objective function becomes

$$\mu(x) = \max\{(c^\ell)^\top x : \ell = 1, \ldots, S\} = \max\{y_\ell(x) : \ell = 1, \ldots, S\}.$$

This model, that we call *Minimax model*, selects a solution $x^{\max}$ such that

$$\mu(x^{\max}) = \min\{\mu(x) : x \in X\}.$$

The Minimax model is only focused on the worst outcome, and thus, takes into account only a small part of the available information. On the other hand, the Minavg model uses all the available information but it does not consider variability in the outcome and the existence of very poor possible results. Let us interpret the concepts in the uncertainty and variability settings.

Uncertainty. In (1), we have S scenarios of the coefficients of the objective function and c^ℓ is the vector of the coefficients in scenario ℓ. The outcome $y_\ell(x)$ measures the performance of solution x in the ℓ-th scenario. In the Minavg model, an optimal solution x^{avg} is such that the average outcome, taken over all the scenarios, is optimized. In the Minimax model, an optimal solution $x^{\max}$ is such that the outcome in the worst scenario is optimized.

Variability. In (1), we have S agents, for example customers or time periods. The vector c^ℓ is the vector of the coefficients of the objective function for agent ℓ. The outcome $y_\ell(x)$ measures the performance of the

ℓ-th agent in solution x. The Minavg model optimizes the average performance of all the agents, whereas the Minimax model optimizes the performance of the worst agent.

In summary, in the uncertainty setting the Minavg model implies risk indifference, whereas the Minimax model implies as high as possible risk aversion. In the variability setting, the Minavg model implies that the only goal is system efficiency and there is no concern about fairness among agents, whereas opposite concerns are taken into account in the Minimax model.

3 The Discrete Conditional Value-at-Risk

In this section, we formally define the DCVaR, that is the average outcome over a given fraction $1-\alpha$ of the worst outcomes. Henceforth, we shall refer to $\beta = 1-\alpha$, the complement to 1 of α. In other words, we will refer directly to the fraction of the worst outcomes. The main reason is that our focus is on the control of the worst outcomes rather than on the good ones. Note that the same convention is also adopted in Mansini et al. [2015]. The expression Discrete Conditional Value-at-Risk may not be ideal to capture the underlying concept related to uncertainty or variability in a general MILP model. However, we adopt it due to the popularity of the CVaR and to the equivalence of the measure with the CVaR in a discrete setting. We recall that we are considering MILP problem (1) which is in minimization form.

Definition 1. For any fraction β, with $\beta \in (0,1]$, and for any $x \in X$, the *Discrete Conditional Value-at-Risk* (DCVaR) $M_\beta(x)$ is the largest average outcome attainable by a subset of $\lceil \beta S \rceil$ states in solution x.

Note that, when $\beta = k/S$, then the DCVaR is the largest average outcome attainable by a subset of k states. The β-DCVaR optimization model aims at finding a solution x^* such that

$$M_\beta(x^*) = \min\{M_\beta(x) : x \in X\}. \tag{3}$$

When $\beta = 1$ the 1-DCVaR model becomes the Minavg model, whereas for $\beta = 1/S$ the model is the Minimax model.

For any $x \in X$, let $y(x) = (y_1(x), y_2(x), \ldots, y_S(x))^\top$ be the vector of outcomes. Then, let

$$\tau(1), \tau(2), \ldots, \tau(S)$$

be a *valid permutation* of the outcome indices, that is a permutation such that

$$y_{\tau(1)}(x) \geq y_{\tau(2)}(x) \geq \ldots \geq y_{\tau(S)}(x). \tag{4}$$

The permutation τ actually depends on $x \in X$, but, to ease notation, we take such a dependence for granted. Let $y^{\tau}(x)$ be the ordered vector of outcomes. Whereas $y_j(x)$ indicates the j-th component of the original vector $y(x)$, $y_{\tau(j)}(x)$ indicates the j-th component of the ordered vector $y^{\tau}(x)$.

We can now express the β-DCVaR, with $\beta \in (0,1]$, in formulae as

$$M_{\beta}(x) = \frac{1}{\lceil \beta S \rceil} \sum_{j=1}^{\lceil \beta S \rceil} y_{\tau(j)}(x). \tag{5}$$

Example 1. Suppose that $S = 10$ and that, for a given solution $x \in X$, we have:

$$y(x) = (12,3,1,7,18,9,4,12,15,13)^{\top}.$$

Then, we consider the valid permutation $\tau : 5,9,10,1,8,6,4,7,2$, which orders the vector of outcomes. The corresponding ordered vector is

$$y^{\tau}(x) = (18,15,13,12,12,9,7,4,3,1)^{\top}.$$

We note that an alternative valid permutation $\tau' : 5,9,10,8,1,6,4,7,2$ where indices 1 and 8 are switched would result in the same ordered vector.

Using formula (5) we can compute $M_{\beta}(x)$ for different values of β. If $\beta = 2/10$, then $M_{0.2}(x) = (y_5 + y_9)/2 = 16.5$. If $\beta = 4/10$, then $M_{0.4}(x) = (y_5 + y_9 + y_{10} + y_1)/4 = 14.5$. If β is not a multiple of $\frac{1}{S}$, we round the value up to the closest multiple. For example, if $\beta = 0.35$, which is included between $3/10$ and $4/10$, then $\lceil \beta S \rceil = 4$ and we obtain $M_{0.35}(x) = M_{0.4}(x) = 14.5$. □

3.1 Properties

We derive several properties of the DCVaR. First, we point out that the DCVaR is a well-defined measure, in the sense that it does not depend on the particular valid permutation of outcomes used. Then, we describe the relation between the DCVaR and the classical CVaR. We also show that DCVaR is consistent with the preferences of a risk-averse decision maker, and that the DCVaR is a measure of equitability and fairness. Finally, we show that the DCVaR preserves linearity and convexity of optimization problems.

3.1.1 The DCVaR is a well-defined measure

The DCVaR is not affected by the specific valid permutation of outcomes. Indeed, directly from Definition 1 and expression (5), the following statement follows.

Proposition 1. *For any given β, $\beta \in (0,1]$, the β-DCVaR is a symmetric function of $y_1(x), y_2(x), \ldots, y_S(x)$, i.e., given two valid permutations τ' and τ'', the value $M_\beta(x)$ obtained with permutation τ' is identical to the value obtained with permutation τ''.*

3.1.2 Relation between DCVaR and CVaR

In order to clarify the relation between the DCVaR and the CVaR, we need to define some functions. We start from the right-continuous function:

$$F_{y(x)}(\xi) = \sum_{\ell=1}^{S} \frac{1}{S} \delta_\ell(\xi), \quad \text{where} \quad \delta_\ell(\xi) = \begin{cases} 1 & \text{if } y_\ell(x) \leq \xi \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

which, for any real value ξ, provides the fraction of outcomes smaller than or equal to ξ. Note that function $F_{y(x)}(\xi)$ can be seen as the cumulative distribution function of a discrete random variable that takes the values $y_1(x), y_2(x), \ldots, y_S(x)$, each with the same probability $1/S$. In Figure 1, we show function $F_{y(x)}(\xi)$ for Example 1.

Similarly, the left-continuous right tail function

$$\overline{F}_{y(x)}(\xi) = \sum_{\ell=1}^{S} \frac{1}{S} \overline{\delta}_\ell(\xi) \quad \text{where} \quad \overline{\delta}_\ell(\xi) = \begin{cases} 1 & \text{if } y_\ell(x) \geq \xi \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

can be defined which, for any real value ξ, provides the fraction of outcomes greater than or equal to ξ. Note that $\overline{F}_{y(x)}(\xi) = 1 - F_{y(x)}(\xi)$ for $\xi \notin \{y_1(x), y_2(x), \ldots, y_S(x)\}$.

Next, we introduce the quantile function $F^{(-1)}_{y(x)}$ as the left-continuous inverse of $F_{y(x)}(x)$, i.e., $F^{(-1)}_{y(x)}(\alpha) = \inf\ \{\xi : F_{y(x)}(\xi) \geq \alpha\}$ for $\alpha \in (0,1]$. The quantile function $F^{(-1)}_{y(x)}$ is the nondecreasing stepwise function

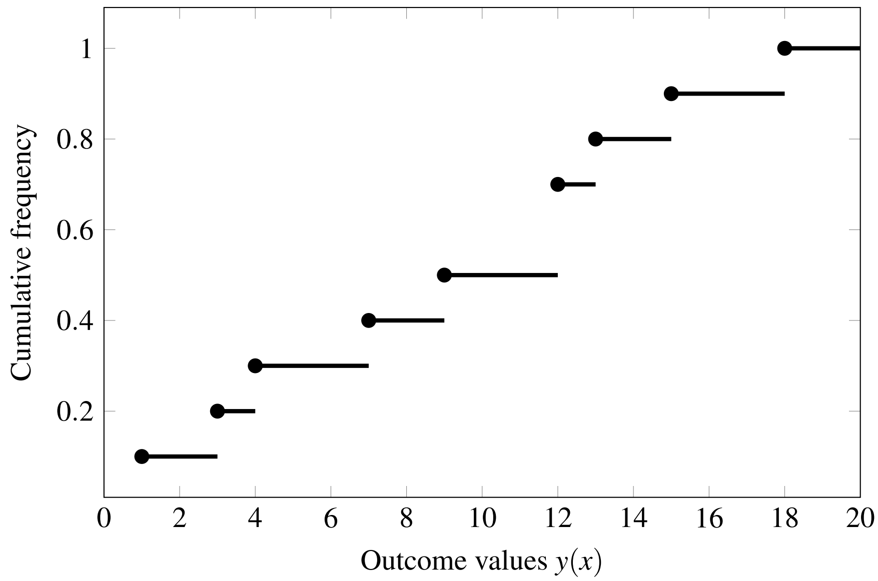

Figure 1: *Plot of function $F_{y(x)}(\xi)$ for $y(x)$ as in Example 1.*

$F^{(-1)}_{y(x)}(\alpha) = y_{\tau(S-j+1)}(x)$ for $(j-1)/S < \alpha \leq j/S$.

Similarly, we introduce the right tail quantile function $\overline{F}^{(-1)}_{y(x)}$ as the left-continuous inverse of $\overline{F}_{y(x)}$, i.e., $\overline{F}^{(-1)}_{y(x)}(\alpha) = \sup\{\xi : \overline{F}_{y(x)}(\xi) \geq \alpha\}$ for $\alpha \in (0,1]$. Actually, $\overline{F}^{(-1)}_{y(x)}(\alpha) = F^{(-1)}_{y(x)}(1-\alpha)$. Note that the function $\overline{F}^{(-1)}_{y(x)}$ is the nonincreasing stepwise function $\overline{F}^{(-1)}_{y(x)}(\alpha) = y_{\tau(j)}(x)$ for $(j-1)/S < \alpha \leq j/S$. In Figure 2, we show function $F^{(-1)}_{y(x)}(\xi)$ for Example 1.

Variability of the outcome distribution can be described with the Lorenz curve (see, e.g., Ogryczak and Ruszczyński [2002]). The (convex) absolute Lorenz curve for any distribution may be viewed as an integrated quantile function

$$F^{(-2)}_{y(x)}(\alpha) = \int_0^\alpha F^{(-1)}_{y(x)}(\nu)d\nu, \tag{8}$$

with $\alpha \in (0,1]$. Similarly, the upper (concave) absolute Lorenz curve may be used which integrates the right tail quantile function

$$\overline{F}^{(-2)}_{y(x)}(\alpha) = \int_0^\alpha \overline{F}^{(-1)}_{y(x)}(\nu)d\nu. \tag{9}$$

Actually,

$$\overline{F}^{(-2)}_{y(x)}\left(\frac{k}{S}\right) = \frac{1}{S}\sum_{j=1}^{k} y_{\tau(j)}(x).$$

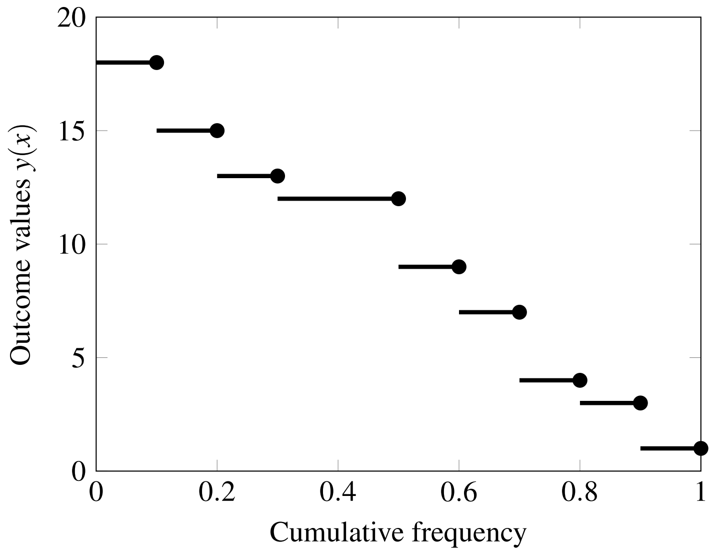

Figure 2: Plot of function $\overline{F}^{(-1)}_{y(x)}(\xi)$ for $y(x)$ as in Example 1.

Recalling expression (5), we obtain

$$M_\beta(x) = \frac{1}{\lceil \beta S \rceil} \sum_{j=1}^{\lceil \beta S \rceil} y_{\tau(j)}(x) = \frac{1}{\lceil \beta S \rceil} \sum_{j=1}^{\lceil \beta S \rceil} \overline{F}^{(-1)}_{y(x)}(\frac{j}{S}) = \frac{S}{\lceil \beta S \rceil} \overline{F}^{(-2)}_{y(x)}(\frac{\lceil \beta S \rceil}{S}). \tag{10}$$

Given a specified probability level β, we recall that the $(1-\beta)$-CVaR measure $\overline{\mu}_\beta(x)$ is represented by the right tail mean (of loss distribution) (see Rockafellar and Uryasev [2000]), that is

$$\overline{\mu}_\beta(x) = \frac{1}{\beta} \overline{F}^{(-2)}_{y(x)}(\beta). \tag{11}$$

Since $S/\lceil \beta S \rceil \leq 1/\beta$ and $\overline{F}^{(-2)}_{y(x)}$ is a nondecreasing function, (10) and (11) imply the following.

Proposition 2. *For all $\beta \in (0,1]$, we have $M_\beta(x) \leq \overline{\mu}_\beta(x)$. In particular,*

$$M_\beta(x) = \overline{\mu}_{\frac{\lceil \beta S \rceil}{S}}(x),$$

that is, the β-DCVaR coincides with the $\left(1 - \frac{\lceil \beta S \rceil}{S}\right)$-CVaR.

The above proposition justifies the name Discrete CVaR we have adopted. In Figure 3, we plot functions $M_\beta(x)$ and $\overline{\mu}_\beta(x)$ with respect to parameter β for the outcome vector of Example 1.

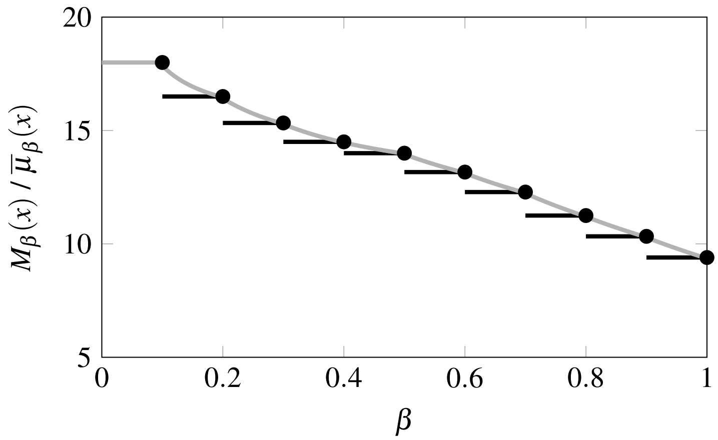

Figure 3: Plot of $M_\beta(x)$ (black line) vs. $\overline{\mu}_\beta(x)$ (gray line) for $y(x)$ as in Example 1.

3.1.3 The DCVaR is consistent with risk-aversion

We recall that the Maximin solution is considered risk-averse in the decision under risk setting (what we call uncertainty setting) as well as equitable (or fair) in the pure deterministic setting (that we call variability setting) Ogryczak [2009]. These characteristics essentially follow from the consistency with the increasing convex order, which is more commonly known as the Second order Stochastic Dominance (SSD) Müller and Stoyan [2002]. The stochastic dominance can be expressed on the cumulative distribution functions. Having introduced the function (7), one may further integrate it to get, for $\eta \in \mathbb{R}$, the second order cumulative distribution function

$$\overline{F}^{(2)}_{y(x)}(\eta) = \int_{\eta}^{\infty} \overline{F}_{y(x)}(\xi)d\xi,$$

which represents the average excess over any real target η. Graphs of functions $\overline{F}^{(2)}_{y(x)}(\eta)$ (with respect to η) take the form of convex decreasing curves Ogryczak and Ruszczyński [1999]. By definition, the pointwise comparison of the second order cumulative distribution functions provides a characterization of the SSD. Formally, $y(x')$ SSD dominates $y(x'')$ (written as $y(x') \preceq_{SSD} y(x'')$) if and only if $\overline{F}^{(2)}_{y(x')}(\eta) \leq \overline{F}^{(2)}_{y(x'')}(\eta)$ for all η where at least one strict inequality holds.

By the theory of convex conjugate functions Boyd and Vandenberghe [2004], the pointwise comparison of the absolute Lorenz curves $\overline{F}^{(-2)}_{y(x)}$ provides an alternative characterization of the SSD relation Ogryczak and Ruszczyński [2002]. This means that $y(x')$ SSD dominates $y(x'')$ if and only if $\overline{F}^{(-2)}_{y'(x)}(\alpha) \leq \overline{F}^{(-2)}_{y(x'')}(\alpha)$ for all $\alpha \in (0, 1]$ where at least one strict inequality holds.

Proposition 3. *For any $\beta \in (0, 1]$, the DCVaR is SSD consistent, in the sense that*

$$y(x') \preceq_{SSD} y(x'') \quad \Rightarrow \quad M_\beta(x') \leq M_\beta(x'').$$

The above proposition is straightforward. Indeed, if $y(x') \preceq_{SSD} y(x'')$, then $\overline{F}^{(-2)}_{y'(x)}(\alpha) \leq \overline{F}^{(-2)}_{y(x'')}(\alpha)$ for all $\alpha \in (0, 1]$, and thereby, for any $\beta \in (0, 1]$,

$$M_\beta(x') = \frac{S}{\lceil \beta S \rceil} \overline{F}^{(-2)}_{y(x')} \left(\frac{\lceil \beta S \rceil}{S} \right) \leq \frac{S}{\lceil \beta S \rceil} \overline{F}^{(-2)}_{y(x'')} \left(\frac{\lceil \beta S \rceil}{S} \right) = M_\beta(x'').$$

3.1.4 The DCVaR is a measure of fairness

As, from (5), the DCVaR can be expressed as Ogryczak and Śliwiński [2003]

$$M_\beta(x) = \frac{1}{\lceil \beta S \rceil} \max_{\pi \in \Pi(S)} \sum_{j=1}^{\lceil \beta S \rceil} y_{\pi(j)}(x), \tag{12}$$

where $\Pi(S)$ is the set of all permutations of S, the following statement is valid.

Proposition 4. *For any given S and for any $\beta \in (0, 1]$, the DCVaR is a convex piecewise linear function of $y_1(x), y_2(x), \ldots, y_S(x)$.*

Due to Proposition 1, the DCVaR is actually a Schur-convex function of $y(x)$ (see Marshall and Olkin [1980]).

Proposition 5. *For any given S and for any $\beta \in (0, 1]$, the DCVaR is a Schur-convex function of $y_1(x), y_2(x), \ldots, y_S(x)$.*

The latter confirms that the DCVaR is an order-preserving function. As a consequence, minimization of the DCVaR represents equitable preferences Kostreva et al. [2004] or fair optimization Ogryczak et al. [2014] with respect to the outcomes $y(x)$.

3.1.5 Piecewise linearity

Finally, from Proposition 4 we deduce a convexity property of the DCVaR with respect to the decision vector x.

Proposition 6. *For any given S and for any $\beta \in (0,1]$, if outcomes $y_\ell(x) = f_\ell(x)$ are defined by linear functions of x, then the DCVaR of $y(x)$ is a convex piecewise linear function of x.*

This guarantees the possibility to model the DCVaR optimization problem with an LP expansion (additional linear constraints and variables) of the original MILP or LP problem.

4 Optimizing the Discrete Conditional Value-at-Risk

The optimization of the DCVaR defined in (3) with $M_\beta(x)$ formulated as in (5) is not practical, since it relies on a permutation of the outcomes, that is, a function of $x \in X$. We can, however, find a different formulation that leads to a MILP model.

Proposition 7. *For any β, with $\beta \in (0,1]$, the DCVaR is given by*

$$M_\beta(x) = \max\left\{ \frac{1}{\lceil \beta S \rceil} \sum_{\ell=1}^{S} y_\ell(x) z_\ell : \sum_{\ell=1}^{S} z_\ell = \lceil \beta S \rceil,\ z_\ell \in \{0,1\}, \ell = 1,\ldots,S \right\}. \tag{13}$$

The above proposition is straightforward. Indeed, the binary variables describe the incidence vectors of the outcome subsets. The unique constraint imposes that only subsets of cardinality $\lceil \beta S \rceil$ are chosen. The objective function maximizes the average outcome of the selected subset.

Proposition 7 allows a practical formulation for the DCVaR model. As the matrix of the coefficients of the constraint in problem (13) is totally unimodular, we can relax the integrality constraints, obtaining the equivalent linear program

$$M_\beta(x) = \max\left\{ \frac{1}{\lceil \beta S \rceil} \sum_{\ell=1}^{S} y_\ell(x) z_\ell : \sum_{\ell=1}^{S} z_\ell = \lceil \beta S \rceil,\ 0 \le z_\ell \le 1, \ell = 1,\ldots,S \right\}. \tag{14}$$

Problem (14) is feasible and bounded for any $\beta \in (0,1]$ and $x \in X$. Thus, we can switch to the dual, obtaining

$$M_\beta(x) = \min\left\{ \lceil \beta S \rceil u + \sum_{\ell=1}^{S} v_\ell : \lceil \beta S \rceil (u + v_\ell) \ge y_\ell(x),\ v_\ell \ge 0, \ell = 1,\ldots,S \right\}.$$

Replacing the outcomes with their explicit expression in terms of $x \in X$, we are able to express the

DCVaR model for the minimization problem (1) as

$$
\begin{aligned}
\min \quad & \lceil \beta S \rceil u + \sum_{\ell=1}^{S} v_\ell \\
\text{subject to} \quad & \lceil \beta S \rceil (u + v_\ell) \geq (c^\ell)^\top x \quad \ell = 1, \dots, S \\
& Ax = b \\
& v_\ell \geq 0 \qquad\qquad\qquad\quad\; \ell = 1, \dots, S \\
& x_B \in \{0,1\}^{|B|}, x_N \geq 0.
\end{aligned} \tag{15}
$$

We remark that (15) is still a MILP model. With respect to the original (1), it has $S+1$ additional continuous variables and S additional inequality constraints. Moreover, the objective function and the new constraints have integer coefficients, provided the realizations are integer valued.

4.1 Case of maximization

We consider here the case where the original MILP problem is formulated in maximization form as

$$
\begin{aligned}
\max \quad & c^\top x \\
\text{subject to} \quad & Ax = b \\
& x_B \in \{0,1\}^{|B|} \\
& x_N \geq 0.
\end{aligned} \tag{16}
$$

In this case, the DCVaR is the smallest average outcome attainable by a subset of states whose cardinality is at least k. The simplest way to derive the DCVaR optimization model is to transform problem (16) in minimization form and then use the DCVaR optimization model (15). With some algebraic calculations, considering that u is a free variable, we obtain the following DCVaR optimization problem for problem (16)

$$
\begin{aligned}
\max \quad & \lceil \beta S \rceil w - \sum_{\ell=1}^{S} v_\ell \\
\text{subject to} \quad & \lceil \beta S \rceil (w - v_\ell) \leq (c^\ell)^\top x \quad \ell = 1, \dots, S \\
& Ax = b \\
& v_\ell \geq 0 \qquad\qquad\qquad\quad\; \ell = 1, \dots, S \\
& x_B \in \{0,1\}^{|B|}, x_N \geq 0.
\end{aligned} \tag{17}
$$

In the above model, the role of w is the same as $-u$ in (15).

5 Computational experiments

The aim of this section is to illustrate the concepts presented above by using two classical combinatorial optimization problems. The first one is the MKP, which gives rise to MILP models with maximization objective, and will be used as an example of control of the uncertainty in the objective function coefficients. The second one is the p-median/p-center problem, which gives rise to MILP models with a minimization objective, and will be used as example of control of variability. The models are tested on benchmark instances taken from the OR-library Beasley [1990], as detailed below. All used instances are available at `http://people.brunel.ac.uk/~mastjjb/jeb/orlib/files`.

All tests have been performed on a 64-bit Windows machine, with Intel Xeon processor E5-1650, 3.50 GHz, and 16 GB of RAM. All models have been built and solved using IBM ILOG CPLEX Optimization Studio 12.6 (64 bit version) with default solver settings, except for the maximum running time, set to 7200 seconds for each model-instance pair.

5.1 Example of uncertainty: The multidimensional knapsack problem

The MKP is a well-known combinatorial optimization problem whose feasible set may be expressed as follows

$$X = \left\{ x \in \{0,1\}^m : \sum_{j=1}^{n} a_{ij} x_j \le b_i \ (i = 1, \dots, m) \right\}.$$

The MKP is formulated as Kellerer et al. [2004]

$$\text{MKP: } \max \left\{ \sum_{j=1}^{n} p_j x_j : x \in X \right\},$$

where all input parameters are integers.

We assume that the profits p_j are not known exactly and that S scenarios are given to capture the uncertainty. In other words, we assume that S vectors $p^\ell = (p_{\ell 1}, p_{\ell 2}, \dots, p_{\ell n})$ are given, for $\ell = 1, \dots, S$, and that a probability $\pi_i = 1/S$ is assigned to each vector. For consistency, we assume also that in MKP we define

$p_j = \sum_{\ell=1}^{S} \pi_\ell p_{\ell j}$ for all $j = 1,\ldots,n$. In this setting, the Maximin model related to the MKP is

$$\text{MKP-MAXMIN:} \ \max\left\{u : u \leq \sum_{j=1}^{n} p_{\ell j} x_j, \ell = 1,\ldots,S; x \in X\right\}.$$

The DCVaR model aims at maximizing the total profit earned under the $k = \lceil \beta S \rceil$ worst scenarios. This leads to the following model

$$\text{MKP-MAXMIN}(\beta):$$
$$\max\left\{ku - \sum_{\ell=1}^{S} \nu_\ell : ku - k\nu_\ell \leq \sum_{j=1}^{n} p_{\ell j} x_j, \ell = 1,\ldots,S; \nu_\ell \geq 0, \ell = 1,\ldots,S; x \in X\right\}.$$

We compare the solutions obtained by the three models above on a set of 10 medium size benchmark instances taken from the OR-library, namely instances #11–20 from file `mknapcb1.txt`, where $n = 100$ and $m = 5$. These are randomly generated instances where $b_i = \frac{1}{2}\sum_{j=1}^{n} a_{ij}$, $i = 1,\ldots,m$, and the profit p_j is correlated to $\sum_{i=1}^{m} a_{ij}$ for all $j = 1,\ldots,n$ (see Chu and Beasley [1998] for details).

For each instance, $S = 100$ scenarios are generated. We assume that for each object j, the original profit p_j is actually the maximum possible one, and we generate independently 100 values by drawing uniformly distributed integers in the interval $[-0.75p_j, p_j]$. A similar scheme has been used in Ogryczak and Śliwiński [2003] to generate rates of return in a portfolio optimization problem. The four models MKP-MAXMIN, MKP-MAXMIN(0.10), MKP-MXMIN(0.30), and MKP are run on each instance. The obtained solutions are compared in Table I.

In Table I, the solution returned by each model is evaluated with respect to the objective of the other models. More precisely, we have four rows per instance, each row corresponding to a model. Column *Min* contains the percentage differences between the minimum total profit obtained under the worst scenario and the optimal value of the MKP-MAXMIN model. Similarly, column $M_{0.10}$ (resp. $M_{0.30}$) contains the percentage differences between the average profit under the $\lceil 0.10 \times S \rceil$ (resp. $\lceil 0.30 \times S \rceil$) worst scenarios and the optimal value of the MKP-MAXMIN(0.10) (resp. MKP-MAXMIN(0.30)) model. Column *Avg* contains the percentage differences between the average profit under all scenarios and the optimal value of the MKP model. In other words, these four columns contain the percentage *regrets* that the solution of each model implies with respect to every considered objective. Such regrets are computed in negative form, to emphasize the fact that they correspond to a reduction with respect to the optimal value. For example, it turns out that in instance #11 the

Table I: Multidimensional knapsack instances with uncertain profits and 100 scenarios: percentage deviations on the optimal objective values.

Instance	*Model type*	*Min*	$M_{0.10}$	$M_{0.30}$	*Avg*	*Time*	*Gap*
11	MKP-MAXMIN	–	0.00	-6.15	-13.26	1553	–
...	MKP-MAXMIN(0.10)	0.00	–	-6.15	-13.26	35	–
...	MKP-MAXMIN(0.30)	-92.74	-11.40	–	-9.69	22	–
...	MKP	-129.56	-60.40	-20.13	–	0	–
12	MKP-MAXMIN	–	-1.45	-9.98	-13.82	3184	–
...	MKP-MAXMIN(0.10)	-16.20	–	-5.20	-7.02	141	–
...	MKP-MAXMIN(0.30)	-103.73	-13.55	–	-9.11	27	–
...	MKP	-249.80	-77.95	-26.57	–	0	–
13	MKP-MAXMIN	–	-2.16	-7.90	-15.01	7201	4.11
...	MKP-MAXMIN(0.10)	-23.03	–	-1.19	-10.51	249	–
...	MKP-MAXMIN(0.30)	-64.67	-14.43	–	-10.42	69	–
...	MKP	-139.48	-65.92	-25.90	–	0	–
14	MKP-MAXMIN	–	-1.31	-6.87	-9.07	5796	–
...	MKP-MAXMIN(0.10)	-18.23	–	-4.63	-10.78	237	–
...	MKP-MAXMIN(0.30)	-94.52	-21.91	–	-6.71	22	–
...	MKP	-130.27	-52.82	-19.26	–	0	–
15	MKP-MAXMIN	–	-2.10	-8.22	-18.74	7201	6.07
...	MKP-MAXMIN(0.10)	-6.34	–	-3.96	-14.50	2994	–
...	MKP-MAXMIN(0.30)	-32.29	-14.32	–	-9.98	40	–
...	MKP	-130.93	-67.78	-32.40	–	0	–
16	MKP-MAXMIN	–	-1.36	-8.21	-12.64	7200	4.42
...	MKP-MAXMIN(0.10)	-15.93	–	-3.97	-12.77	547	–
...	MKP-MAXMIN(0.30)	-48.82	-12.45	–	-8.16	60	–
...	MKP	-132.76	-55.08	-20.84	–	0	–
17	MKP-MAXMIN	–	-1.98	-7.13	-14.14	7201	5.03
...	MKP-MAXMIN(0.10)	-1.94	–	-4.49	-13.72	476	–
...	MKP-MAXMIN(0.30)	-74.33	-14.90	–	-11.69	22	–
...	MKP	-167.20	-71.10	-24.71	–	0	–
18	MKP-MAXMIN	–	-2.00	-6.05	-12.38	7201	5.82
...	MKP-MAXMIN(0.10)	-10.14	–	-3.85	-12.55	2140	–
...	MKP-MAXMIN(0.30)	-81.91	-17.64	–	-8.40	225	–
...	MKP	-183.47	-68.97	-28.57	–	0	–
19	MKP-MAXMIN	–	-1.56	-6.88	-14.12	3540	–
...	MKP-MAXMIN(0.10)	-7.00	–	-4.89	-13.47	150	–
...	MKP-MAXMIN(0.30)	-32.34	-7.82	–	-11.34	26	–
...	MKP	-205.90	-83.99	-33.68	–	1	–
20	MKP-MAXMIN	–	-2.37	-3.86	-15.02	7201	5.90
...	MKP-MAXMIN(0.10)	-6.95	–	-3.75	-15.73	809	–
...	MKP-MAXMIN(0.30)	-81.51	-18.95	–	-10.95	53	–
...	MKP	-111.51	-74.29	-31.74	–	0	–

worst possible profit for the solution returned by the MKP-MAXMIN(0.30) is 92.74% lower than the worst possible profit for the solution returned by the MKP-MAXMIN.

Finally, column *Time* reports the CPU time in rounded seconds required to obtain the solutions, and column *Gap* that contains percentage optimality gap for instances not solved to optimality within the time limit of 7200 seconds. These are all instances of the MKP-MAXMIN model, which turns out to be very hard to solve for CPLEX. On the other extreme, CPLEX is very fast in solving the classical MKP model. MKP-MAXMIN(0.10) and MKP-MAXMIN(0.30) instances are all solved to optimality with a CPU time ranging from 22 to 2994 seconds. Figures in column *Time* suggest a strong inverse correlation between the β value and the hardness of the instance for CPLEX. It is also interesting to note that for instance #11, models MKP-MAXMIN and MKP-MAXMIN(0.10) return the same solution, but while CPLEX required 1553 seconds to prove optimality with respect to the former model, it requires just 35 seconds to prove optimality with respect to the latter model.

Column *Min* suggests a strong negative correlation between the β value and the value of the worst possible profit, though the magnitude of the regret value is clearly related to the adopted perturbation. The positive correlation between the β value and the value of the average profit is less pronounced, and there are some cases of non-monotonicity.

To give a better view of the differences among the solutions returned by the different models, in Figure 4 we show the cumulative frequency function (6) of the four solutions for instance #15, taken as an example. More specifically, the values on the horizontal axis are the possible profits, while the values on the vertical axis are fractions of the scenarios. For example, in the MKP-MAXMIN solution, 65% of the scenarios imply a profit not greater than 5000. In the MKP-MAXMIN(0.10) solution, the percentage reduces to 59% and in the MKP-MAXMIN(0.30) solution a 49% is reached. Finally, in the MKP solution, 41% of the scenarios imply a profit not exceeding 5000. To improve readability, we fill the discontinuities with vertical segments.

The plot shows a common pattern: the distributions for models MKP-MAXMIN and MKP-MAXMIN(0.10) are quite similar, with a large number of outcomes concentrated on the right of the worst outcome, though the solutions of the MKP-MAXMIN(β) models tend to slightly dominate the solution of the MKP-MAXMIN, in the sense that the percentage of scenarios with profit not exceeding δ is smaller in the MKP-MAXMIN(0.10) solution. The distribution corresponding to the MKP is very different, with a bold left tail that includes a fraction of bad outcomes. The distribution corresponding to the MKP-MAXMIN(0.30) tends to mediate between these two patterns.

5.2 Example of variability: A facility location problem

Let us consider a set of S customers, each with unit demand, a set of m potential locations for p facilities, and an $m \times S$ matrix $C = [c_{i\ell}]$ of distances from potential locations to customers. The p-median problem is to select p potential facilities in order to minimize the total distance from customers to selected facilities, assuming that each customer is supplied from the closest selected facility Drezner and Hamacher [2004]. The feasible set of the (uncapacitated) p-median problem is described as

$$XY = \left\{ x \in \{0,1\}^{m\times S}, y \in \{0,1\}^m : \sum_{i=1}^{m} x_{i\ell} = 1, \ell = 1,\dots,S; \right.$$
$$\left. \sum_{i=1}^{m} y_i = p,\ x_{i\ell} \le y_i, i = 1,\dots,m, \ell = 1,\dots,S \right\}.$$

The classical p-median problem is described by the following Minavg model:

$$p\text{-MEDIAN: } \min \left\{ \sum_{i=1}^{m} \sum_{\ell=1}^{S} c_{i\ell} x_{i\ell} : (x,y) \in XY \right\}.$$

The p-center problem requires to minimize the maximum distance between a customer and its closest facility. This corresponds to using a Minimax model, formulated as

$$p\text{-CENTER: } \min \left\{ u : u \ge \sum_{i=1}^{m} c_{i\ell} x_{i\ell}, \ell = 1,\dots,S;\ (x,y) \in XY \right\}.$$

In this context, a DCVaR model aims at minimizing the total distance travelled by the $k = \lceil \beta S \rceil$ customers with largest distances from their closest facilities. This leads to the following model

$$p\text{-CENTER}(\beta):$$
$$\min \left\{ ku + \sum_{\ell=1}^{n} \nu_\ell : ku + k\nu_\ell \ge \sum_{i=1}^{m} c_{i\ell} x_{i\ell}, \ell = 1,\dots,S; \nu_\ell \ge 0, \ell = 1,\dots,S;\ (x,y) \in XY \right\}.$$

We compare the solutions obtained by the three models above on a set of 10 benchmark instances taken from the OR-library, namely those described in files pmed$z$.txt with $z = 1,2,\dots,10$. The first five instances have 100 customers, the remaining instances have 200 customers. In all instances, the customer set and the potential location set are the same. On each instance, we test four models: p-CENTER, p-CENTER(0.05), p-CENTER(0.10), and p-MEDIAN. The value of p is fixed to 5 for all instances. The obtained solutions are compared in Table II.

The structure of Table II is similar to Table I, except for column *s*, reporting the number of customers in the instance, and column *Gap*, that has been omitted as all instances are solved within the time limit of 7200 seconds.

Columns *Max–Avg* are filled similarly to columns *Min–Avg* in Table II, but taking into account that we start from a minimization problem. The objective of the *p*-CENTER model is to minimize the maximum possible distance of a customer from the closest active facility, the objective of the *p*-CENTER(β) is to minimize M_β, i.e., the average of the $\lceil \beta S \rceil$ worst distances from the closest facility, and the objective of the *p*-MEDIAN model is to minimize the distance from the closest facility averaged over all customers. To emphasize the fact that in this case the regrets correspond to increments with respect to the optimal value, they are computed in positive form.

According to the results, models *p*-CENTER(0.05) and *p*-CENTER(0.10) always offer a compromise solution between *p*-CENTER and *p*-MEDIAN in terms of regrets. More precisely, the *p*-CENTER(0.05) and *p*-CENTER(0.10) models ensure a much better performance than the *p*-CENTER model in terms of average objective at the expense of a reasonable increment of the largest distance from an active facility.

From Table II it is clear that the *p*-MEDIAN and the *p*-CENTER models are much easier to solve for CPLEX. This phenomenon may be explained by the fact that CPLEX default parameters and settings are optimized over sets of benchmark instances for classical models like *p*-MEDIAN and *p*-CENTER.

In Figure 5, we plot the cumulative frequency distribution (6) for instance `pmed6`, taken as an example. The pattern is found in all instances: for most distances δ from the closest facility, the solution of the *p*-MEDIAN model dominates the solution of the *p*-CENTER model, in the sense that the percentage of customers with distance not exceeding δ is larger in the *p*-MEDIAN solution. However, for the largest distances, the role is swapped, and the *p*-MEDIAN solution allows a certain fraction of customers to have a very large distance from their facilities. From the figures it is also clear that the *p*-CENTER(0.05) and the *p*-CENTER(0.10) solutions mediate between the Minimax criterion and the Average criterion, not only in terms of values (average and maximum) but also in terms of distribution.

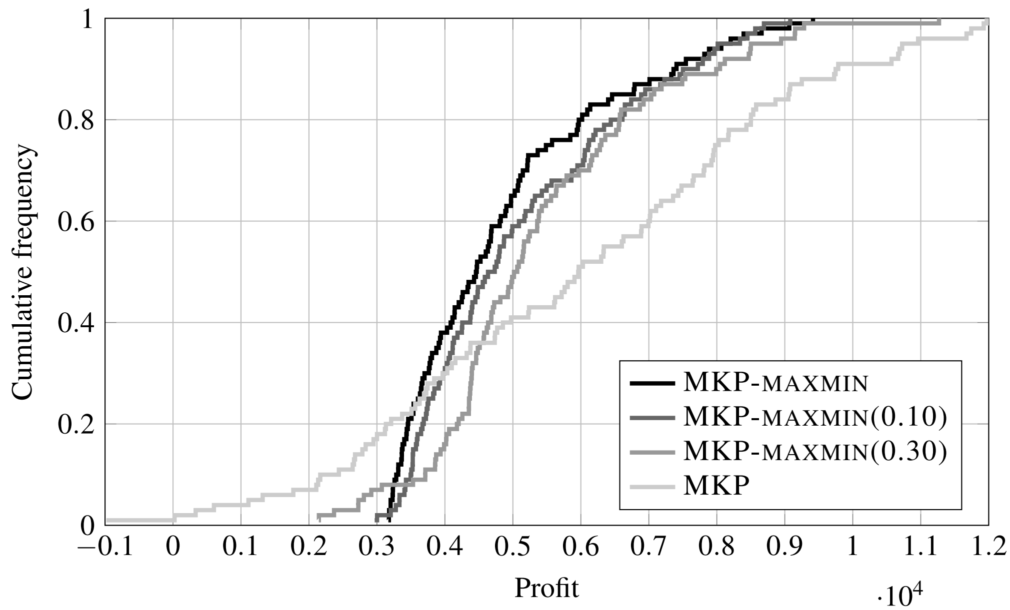

Figure 4: Cumulative frequency plots for instance #15 in `mknapbc1`.

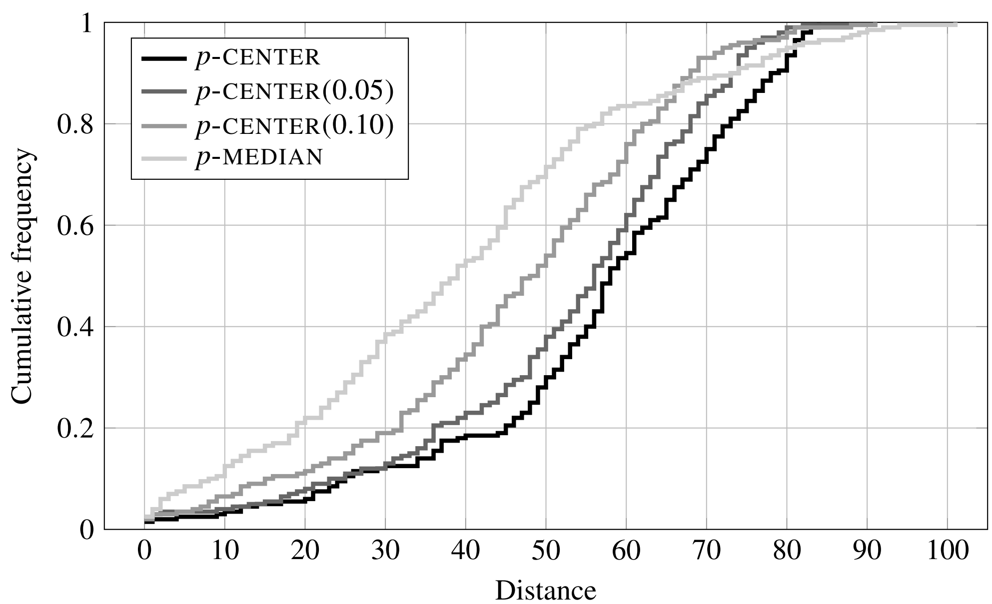

Figure 5: Cumulative frequency plots for instance `pmed6`.

Table II: Facility location instances with 100 and 200 customers from the OR-library.

Instance	*s*	*Model*	*Max*	$M_{0.05}$	$M_{0.10}$	*Avg*	*Time*
pmed1	100	*p*-CENTER	–	7.08	11.33	47.45	7
...	100	*p*-CENTER(0.05)	0.00	–	1.50	21.12	251
...	100	*p*-CENTER(0.10)	0.79	0.34	–	21.60	199
...	100	*p*-MEDIAN	4.72	2.02	2.04	–	1
pmed2	100	*p*-CENTER	–	4.68	7.92	44.40	70
...	100	*p*-CENTER(0.05)	1.56	–	2.02	21.81	935
...	100	*p*-CENTER(0.10)	7.81	2.68	–	22.09	936
...	100	*p*-MEDIAN	19.53	17.56	14.51	–	5
pmed3	100	*p*-CENTER	–	2.13	2.96	35.22	4
...	100	*p*-CENTER(0.05)	1.57	–	0.42	24.40	143
...	100	*p*-CENTER(0.10)	0.00	0.00	–	27.18	159
...	100	*p*-MEDIAN	59.06	27.00	15.13	–	1
pmed4	100	*p*-CENTER	–	1.68	5.10	40.15	5
...	100	*p*-CENTER(0.05)	3.70	–	0.72	21.96	238
...	100	*p*-CENTER(0.10)	11.85	1.38	–	18.44	267
...	100	*p*-MEDIAN	31.85	6.88	2.95	–	1
pmed5	100	*p*-CENTER	–	6.51	10.98	43.07	5
...	100	*p*-CENTER(0.05)	2.59	–	0.79	25.59	382
...	100	*p*-CENTER(0.10)	8.62	2.79	–	21.40	376
...	100	*p*-MEDIAN	29.31	13.20	7.62	–	1
pmed6	200	*p*-CENTER	–	3.26	8.19	43.32	3
...	200	*p*-CENTER(0.05)	8.33	–	2.53	33.53	76
...	200	*p*-CENTER(0.10)	8.33	0.75	–	17.34	80
...	200	*p*-MEDIAN	20.24	11.42	10.39	–	1
pmed7	200	*p*-CENTER	–	6.04	10.40	41.79	70
...	200	*p*-CENTER(0.05)	5.00	–	1.00	36.27	3081
...	200	*p*-CENTER(0.10)	7.50	0.67	–	26.39	3255
...	200	*p*-MEDIAN	23.75	9.53	6.27	–	6
pmed8	200	*p*-CENTER	–	7.62	14.46	65.91	36
...	200	*p*-CENTER(0.05)	13.33	–	1.95	35.67	1516
...	200	*p*-CENTER(0.10)	13.33	0.00	–	24.41	774
...	200	*p*-MEDIAN	28.89	12.33	10.51	–	4
pmed9	200	*p*-CENTER	–	4.61	10.87	72.95	60
...	200	*p*-CENTER(0.05)	6.10	–	2.73	41.64	814
...	200	*p*-CENTER(0.10)	2.44	1.41	–	20.18	540
...	200	*p*-MEDIAN	28.05	10.24	6.49	–	6
pmed10	200	*p*-CENTER	–	6.46	13.63	71.89	33
...	200	*p*-CENTER(0.05)	20.00	–	2.08	47.96	1093
...	200	*p*-CENTER(0.10)	24.29	3.38	–	33.01	1093
...	200	*p*-MEDIAN	37.14	9.69	5.32	–	5

6 The case of not equal weights

The models and theory of the previous sections were presented under the assumption that the outcomes $y_1(x), y_2(x), \ldots, y_S(x)$ have equal weights, that is, that each weight is $1/S$. This implies that, in the uncertainty setting, all scenarios have the same probability and, in the variability setting, all agents have the same importance. Whereas this may be the case in many applications, the case of different weights would be a natural and interesting extension. Let us assume that $y_\ell(x)$ has weight $w_\ell \geq 0$, $\ell = 1, \ldots, S$, with $\sum_{\ell=1}^{S} w_\ell = 1$.

In case of general weight, we should generalize Definition 1 as follows.

Definition 2. For any fraction β, with $\beta \in (0, 1]$, and for any $x \in X$, the *Discrete Conditional Value-at-Risk* (DCVaR) $M_\beta(x)$ is the largest average outcome attainable by a subset of states whose total weight is at least β.

According to the latter definition, the DCVaR might be computed as

$$M_\beta(x) = \max \left\{ \frac{\sum_{\ell=1}^{S} y_\ell(x) w_\ell z_\ell}{\sum_{\ell=1}^{S} w_\ell z_\ell} : \sum_{\ell=1}^{S} w_\ell z_\ell \geq \beta, z_\ell \in \{0, 1\}, \ell = 1, \ldots, S \right\}. \tag{18}$$

Note that problem (18) is harder to solve than problem (14). More specifically, problem (18) is NP-hard (see Benati [2004]). Besides the computational issue, Definition 2 and problem (18) pose an inconsistency issue, illustrated in the following example.

Example 2. In the context of a minimization problem, consider the situation depicted in Table III, where $S = 7$.

Table III: An inconsistency example.

w	0.2	0.1	0.1	0.1	0.1	0.2	0.2
$y(x^1)$	0	1	1	1	1	2	5
$y(x^2)$	0	1	1	1	2	1	5

Taking into account weights w_ℓ, the distribution of $y(x^2)$ is better than the distribution of $y(x^1)$ for any rational decision maker, since the former stochastically dominates the latter. However, if we choose $\beta = 0.30$, according to Definition 2 and problem (18), we get $M_{0.3}(y(x^1)) = 3.5 < 4 = M_{0.3}(y(x^2))$. As a consequence, we should prefer $y(x^1)$ to $y(x^2)$. □

Due to the above inconsistency, the DCVaR concept cannot be extended to the case of general weights, while maintaining its properties.

7 Conclusions

We have shown that concepts developed in financial risk management and fair optimization lead to an innovative measure for general MILP problems. Such a measure, that we called Discrete Conditional Value-at-Risk (DCVaR), aims at mediating between a conservative Minimax/Maximin criterion and an aggressive minimum total cost/maximum total profit criterion. We have developed the concepts in a simple and self-contained way, relying only on LP theory and basic statistical concepts. Modeling approaches based on scenario generation to capture uncertainty can naturally evolve towards adoption of the DCVaR as performance measure. We believe that the proposed concepts and models have a great potential for application in many different areas where uncertainty or fairness is a concern.

Several research directions remain to be explored. The DCVaR optimization model can be adopted for many specific optimization problems, as location problems, scheduling problems with uncertain processing times, and others. As our illustrative examples suggest, the DCVaR optimization problem poses new computational issues, concerning the most effective exact solution method, or the most appropriate cuts. Moreover, in the context of heuristic methods, the analysis of local search techniques for DCVaR models is also worth of investigation, for its consequences on the development of efficient metaheuristics.

References

G. Andreatta and F. Mason. k-eccentricity and absolute k-centrum of a probabilistic tree. *European Journal of Operational Research*, 19(1):114–117, 1985a. doi: http://dx.doi.org/10.1016/0377-2217(85)90315-7.

G. Andreatta and F. M. Mason. Properties of the k-centra in a tree network. *Networks*, 15(1):21–25, 1985b. doi: 10.1002/net.3230150103.

Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999. ISSN 1467-9965. doi: 10.1111/1467-9965.00068.

J. E. Beasley. OR-library: distributing test problems by electronic mail. *Journal of the Operational Research Society*, 41(11):1069–1072, 1990.

Stefano Benati. The computation of the worst conditional expectation. *European Journal of Operational Research*, 155(2):414–425, 2004. doi: http://dx.doi.org/10.1016/S0377-2217(02)00905-0.

James O. Berger. *Statistical decision theory and Bayesian Analysis*. Springer Series in Statistics. Springer-Verlag, New York, 1985. ISBN 0387960988.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004. ISBN 0521833787.

P.C. Chu and J.E. Beasley. A genetic algorithm for the multidimensional knapsack problem. *Journal of Heuristics*, 4(1):63–86, 1998. doi: 10.1023/A:1009642405419.

Z. Drezner and H. W. Hamacher, editors. *Facility Location: Applications and Theory*. Springer, Berlin, 2004.

C.W. Duin and A. Volgenant. The partial sum criterion for steiner trees in graphs and shortest paths. *European Journal of Operational Research*, 97(1):172–182, 1997. doi: http://dx.doi.org/10.1016/S0377-2217(96)00113-0.

P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events: For Insurance and Finance*. Applications of mathematics. Springer, 1997. ISBN 9783540609315.

Robert Garfinkel, Elena Fernández, and Timothy J. Lowe. The k-centrum shortest path problem. *TOP*, 14 (2):279–292, 2006. doi: 10.1007/BF02837564.

G. Grygiel. Algebraic $\sum_k$ assignment problem. *Control and Cybernetics*, 10(3-4):155–165, 1981.

S. L. Hakimi. Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12(3):450–459, 1964. doi: 10.1287/opre.12.3.450.

Michal Kaut and Stein W. Wallace. Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2):257–271, 2007.

Hans Kellerer, Ulrich Pferschy, and David Pisinger. *Knapsack Problems*. Springer-Verlag, Berlin, 2004.

Michael M. Kostreva, WÅĆodzimierz Ogryczak, and Adam Wierzbicki. Equitable aggregations and multiple criteria analysis. *European Journal of Operational Research*, 158(2):362–377, 2004. doi: http://dx.doi.org/10.1016/j.ejor.2003.06.010.

M.M. Kostreva and W. Ogryczak. Equitable approaches to location problems. In J.-C. Thill, editor, *Spatial Multicriteria Decision Making and Analysis: A Geographic Information Sciences Approach*, pages 103–126. Ashgate, Brookfield, 1999.

R. Mansini, W. Ogryczak, and M. G. Speranza. LP solvable models for portfolio optimization: A classification and computational comparison. *IMA Journal of Management Mathematics*, 14:187–220, 2003.

R. Mansini, W. Ogryczak, and M. G. Speranza. *Linear and Mixed Integer Programming for Portfolio Optimization*. EURO Advanced Tutorials on Operational Research. Springer-Verlag, New York, 2015. ISBN 978-3-319-18482-1.

H. M. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

Albert W. Marshall and Ingram Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York, 1980.

Alfred Müller and Dietrich Stoyan. *Comparison Methods for Stochastic Models and Risks*. John Wiley & Sons, New York, 2002.

W. Ogryczak. Inequality measures and equitable locations. *Annals of Operations Research*, 167(1):61–86, 2009. ISSN 0254-5330. doi: 10.1007/s10479-007-0234-9.

W. Ogryczak and A. Ruszczyński. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, 13(1):60–78, 2002. doi: 10.1137/S1052623400375075.

W. Ogryczak and M. Zawadzki. Conditional median: A parametric solution concept for location problems. *Annals of Operations Research*, 110(1-4):167–181, 2002. ISSN 0254-5330. doi: 10.1023/A:1020723818980.

W. Ogryczak, H. Luss, M. Pióro, D. Nace, and A. Tomaszewski. Fair optimization and networks: A survey. *Journal of Applied Mathematics*, 2014(Article ID 612018):1–25, 2014.

Wlodzimierz Ogryczak. Stochastic dominance relation and linear risk measures. In Andrzej M.J. Skulimowski, editor, *Financial Modelling – Proc. 23rd Meeting EURO WG Financial Modelling, Cracow, 1998*, pages 191–212. Progress & Business Publisher, 1999.

Wlodzimierz Ogryczak and Andrzej Ruszczyński. From stochastic dominance to mean-risk models: Semideviations as risk measures. *European Journal of Operational Research*, 116(1):33–50, 1999. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/S0377-2217(98)00167-2.

Wlodzimierz Ogryczak and Tomasz Śliwiński. On equitable approaches to resource allocation problems: The conditional minimax solutions. *Journal of Telecommunications and Information Technology*, 2002 (3):40–48, 2002.

Wlodzimierz Ogryczak and Tomasz Śliwiński. On solving linear programs with the ordered weighted averaging objective. *European Journal of Operational Research*, 148(1):80–91, 2003. doi: http://dx.doi.org/10.1016/S0377-2217(02)00399-5.

Wlodzimierz Ogryczak and Arie Tamir. Minimizing the sum of the k largest functions in linear time. *Information Processing Letters*, 85(3):117–122, 2003. doi: http://dx.doi.org/10.1016/S0020-0190(02)00370-8.

Abraham P. Punnen. K-sum linear programming. *The Journal of the Operational Research Society*, 43(4): 359–363, 1992.

Abraham P. Punnen and Y.P. Aneja. On k-sum optimization. *Operations Research Letters*, 18(5):233–236, 1996. ISSN 0167-6377. doi: http://dx.doi.org/10.1016/0167-6377(95)00059-3.

R. Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2: 21–41, 2000.

H. Edwin Romeijn, Ravindra K. Ahuja, James F. Dempsey, and Arvind Kumar. A new linear programming approach to radiation therapy treatment planning problems. *Operations Research*, 54(2):201–216, 2006. doi: 10.1287/opre.1050.0261.

S. C. Sarin, H. D. Sherali, and L. Liao. Minimizing conditional-value-at-risk for stochastic scheduling problems. *Journal of Scheduling*, 17:5–17, 2014.

Peter J. Slater. Centers to centroids in graphs. *Journal of Graph Theory*, 2(3):209–222, 1978. doi: 10.1002/jgt.3190020304.

Akiko Takeda and Takafumi Kanamori. A robust approach based on conditional value-at-risk measure to statistical learning problems. *European Journal of Operational Research*, 198(1):287–296, 2009. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/j.ejor.2008.07.027.

Arie Tamir. The k-centrum multi-facility location problem. *Discrete Applied Mathematics*, 109(3):293–307, 2001. doi: http://dx.doi.org/10.1016/S0166-218X(00)00253-5.

I. Toumazis and C. Kwon. Routing hazardous materials on time-dependent networks using conditional value-at-risk. *Transportation Research Part C*, 37:73–92, 2013.

A. Wald. *Statistical Decision Functions*. John Wiley, New York, 1950.

Gerhard Woeginger. On minimizing the sum of k tardiness. *Information Processing Letters*, 38(5):253–256, 1991. ISSN 0020-0190. doi: http://dx.doi.org/10.1016/0020-0190(91)90067-R.