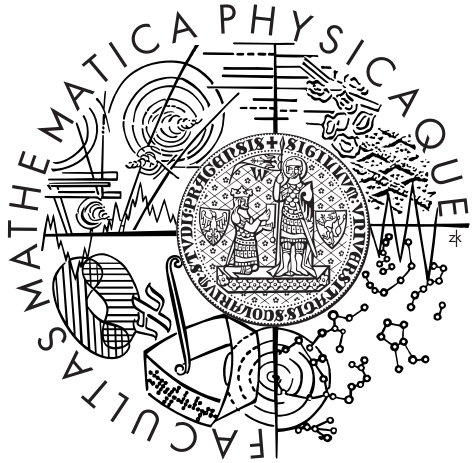
Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

AUTOREFERÁT DISERTAČNÍ PRÁCE



Mgr. Michal Houda

Stabilita a aproximace pro úlohy stochastického programování

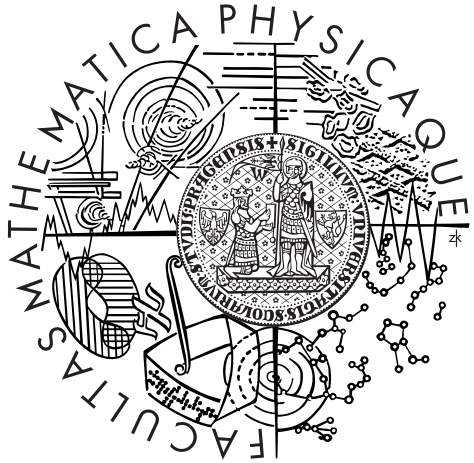
Katedra pravděpodobnosti a matematické statistiky

Školitel: RNDr. Vlasta Kaňková, CSc.
Studijní obor: M5 – Ekonometrie a operační výzkum

Praha, 2009

Charles University in Prague
Faculty of Mathematics and Physics

ABSTRACT OF DOCTORAL THESIS



Mgr. Michal Houda

Stability and approximations for stochastic programs

Department of Probability and Mathematical Statistics

Supervisor: RNDr. Vlasta Kaňková, CSc.
Branch: M5 – Econometrics and Operations Research

Prague, 2009

Dizertační práce byla vypracována na základě výsledků získaných během doktorského studia, které uchazeč absolvoval na Katedře pravděpodobnosti a matematické statistiky Matematicko-fyzikální fakulty Univerzity Karlovy v Praze v letech 2001–2009.

Uchazeč: Mgr. Michal Houda
Katedra aplikované matematiky a informatiky
Ekonomická fakulta Jihočeské univerzity
Studentská 13, 370 05 České Budějovice

Školitel: RNDr. Vlasta Kaňková, CSc.
Ústav teorie informace a automatizace
Akademie věd České republiky
Pod Vodárenskou věží 4, 182 08 Praha 8

Oponenti: Prof. RNDr. Jitka Dupačová, DrSc.
Katedra pravděpodobnosti a matematické statistiky
Matematicko-fyzikální fakulta Univerzity Karlovy
Sokolovská 83, 186 75 Praha 8

Dr. René Henrion
Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstr. 39, D-10117 Berlin, Germany

Autoreferát byl odeslán dne 21. 5. 2009

Obhajoba disertační práce se koná dne 22. 6. 2009 v 9.30 hodin před komisí pro obhajoby disertačních prací oboru M5 – Ekonometrie a operační výzkum na Matematicko-fyzikální fakultě Univerzity Karlovy v posluchárně KPMS-praktikum, Sokolovská 83, Praha 8.

S disertační prací je možno se seznámit na studijním oddělení doktorského studia MFF UK, Ke Karlovu 3, Praha 2.

Předseda RDSO M5: Prof. RNDr. Jitka Dupačová, DrSc.
KPMS MFF UK
Sokolovská 83, 186 75 Praha 8

Contents

1 Introduction

Stochastic programming is large part of the theory of optimization in which the problems with random parameters occur. We search for solutions of such problems, investigate their properties and properties of the problems themselves. All mentioned is important especially in cases in which the randomness cannot be neglected. To begin our investigation at a mathematically correct base let us start with a formal description of the uncertainty in mathematical programming.

Consider the following general optimization problem

$$\text{minimize } c(x;\xi) \text{ subject to } x \in X,\ f(x;\xi) \leq 0 \tag{1}$$

where $\xi \in \mathbb{R}^s$ is a data element of the problem, $x \in X \subset \mathbb{R}^n$ is a decision vector, the dimensions n, s, m, and the mappings $c\colon \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$ and $f\colon \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^m$ are structural elements of the problem. For our purpose, we extend the description of the framework by the following characterization:

1. The knowledge of the data is insufficient (uncertain). All that is known about the data vector ξ (at least at the very beginning) is that it belongs to a given *uncertainty set* $\Xi \subset \mathbb{R}^s$.

2. The objective function of (1) is required to be the best possible given the actual realization (instance) of $\xi \in \Xi$.

3. The constraints of problem (1) are required to be satisfied as much as possible given the actual realization of $\xi \in \Xi$.

If the realization of ξ is known and fixed in advance (before the decision has to be taken), and the elements of the problem have suitable properties, standard algorithms of

deterministic optimization can be used to solve problem (1) (for example convex programming). But in practice, *uncertainty* of the data is typical property and inevitably has to be considered during the course of building the modelling framework. There are several approaches to deal with uncertainty.

1. **Sensitivity analysis.** This is a traditional way to deal with uncertainty. At the stage of building and solving the optimization problem, the uncertainty is simply ignored and the data is replaced by some nominal value. The accuracy of the optimal solution is inspected ex-post for a single, generated solution and it is limited only to small perturbations of the nominal data.

2. **Parametric programming.** The uncertainty is introduced to the model via a parameter – member of a specified parametric space.

3. **Stochastic programming.** Stochastic programming handles the uncertainty of stochastic nature: we consider ξ in (1) to be a random vector and assume that we are able to identify its underlying probability distribution. The idea of stochastic programming is to incorporate available information about data through its probability distribution and solve the modified, deterministic problem.

4. **Robust programming.** Techniques of *robust programming* are looking for a solution to optimization program (1) that satisfies its constraints for all possible realizations of ξ, i. e., that is feasible for any member of the problems belonging to the family (1).

The dissertation thesis deals with some questions of stochastic and partly of robust programming. It is divided into three main parts. The first part is devoted to the stability in stochastic programming problems, involving the necessary theoretical background on probabilistic metrics and the survey of literature. This part ends with Chapter 5 concerning approximations in stochastic programming. Second part numbered as Chapter 6 deals with convexity in chance-constrained problems, especially for the case of weakly dependent constraint rows. Last part of the thesis, Chapter 7, compares stochastic and robust programming as the two disciplines solving one original optimization problem by different means. The results of the thesis are published by the author's proceedings and papers, the list of publications is included.

2 Stochastic programming problems

A stochastic programming problem can be introduced in the following general formulation:

$$\inf_{x\in X} \int_{\Xi} F_0(x;\xi)\ \mu(\mathrm{d}\xi) \text{ subject to } \int_{\Xi} F_j(x;\xi)\ \mu(\mathrm{d}\xi) \leq 0,\ j = 1,\ldots,d \tag{2}$$

In addition to general assumptions from (1), we assume that $\Xi \subset \mathbb{R}^s$ is a closed set, $\mu \in \mathcal{P}(\Xi)$ is a probability measure with support Ξ, X is a closed set not depending on

μ, $F_j : \mathbb{R}^n \times \Xi \to \bar{\mathbb{R}}$ are for all $j = 0, \ldots, d$ extended real random lower semicontinuous functions.

In what follows we use the following notation:

$$X(\mu) := \left\{ x \in X : \int_\Xi F_j(x;\xi)\ \mu(\mathrm{d}\xi) \leq 0,\ j = 1, \ldots, d \right\},$$
$$\varphi(\mu) := \inf_{x \in X(\mu)} \int_\Xi F_0(x;\xi)\ \mu(\mathrm{d}\xi),$$
$$\psi(\mu) := \left\{ x \in X(\mu) : \int_\Xi F_0(x;\xi)\ \mu(\mathrm{d}\xi) = \varphi(\mu) \right\}.$$

Here, $X(\mu)$ is the set of constraints, $\varphi(\mu)$ is the optimal value and $\psi(\mu)$ is the optimal solution set of the problem (2). Different classes of stochastic programming problems, fall into this general framework.

Recourse (penalization) problems falls into the frame (2) with $d = 0$; the problem constraints do not exhibit any probabilistic nature and the problem reduces to

$$\inf_{x \in X} \int_\Xi F_0(x;\xi)\ \mu(\mathrm{d}\xi). \tag{3}$$

In the usual application framework, the function F_0 is split up into two parts. The first-stage objective function $c(x)$ does not depend on ξ and gives us the (first stage) decision x. After some decision x is chosen, random variable ξ is realized and some compensation/penalization action is taken evaluated by the cost function $\mathcal{Q}(x;\xi)$ taken into consideration in the first stage decision through its expected value:

$$\inf_{x \in X} \left(c(x) + \int_\Xi \mathcal{Q}(x;\xi)\ \mu(\mathrm{d}\xi) \right) \tag{4}$$

where $c\colon \mathbb{R}^n \to \mathbb{R}$, and $\mathcal{Q}\colon \mathbb{R}^n \times \mathbb{R}^s \to \bar{\mathbb{R}}$ The cost function $\mathcal{Q}$ is usually given by a second-stage optimization problem depending on the realization of ξ as well as on the first-stage decision x. Many possible variants of $\mathcal{Q}$are examined by stochastic programming.

Chance-constrained programming problems (in their default form) consider $d = 1$ and set $F_0(x;\xi) = c(x)$ where $c\colon \mathbb{R}^n \to \mathbb{R}$. Further, let $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^s$ be a multifunction representing stochastic constraints of the problem, and

$$F_1(x;\xi) := p - \chi_{H(x)}(\xi),$$

with $p \in [0;1]$ being a prescribed probability level, and $\chi_A(\cdot)$ the characteristic function of a set A. The easy transformation of (2) leads to the formulation of *(joint) chance-constrained programming problem:*

$$\min c(x) \text{ subject to } x \in X,\ \mu(H(x)) \geq p \tag{5}$$

There are also more complicated models falling into the frame of (2) and involving probabilistic terms can be formulated; there is some possibilities mentioned in the thesis.

Almost all methods of stochastic programming rely on the fact that the distribution μ is known and fixed in advance. In practice, this is not the case, the true distribution is rarely completely known and some kind of approximation or statistical *estimates* of μ should be used. Mathematically, we replace the "original" distribution μ by the "new" one denoted ν. The key question of the theory of stability is how the optimal value and optimal solution of (2) *change* in this case. We can split up such large question into several areas of interest: in *qualitative stability* we are looking for suitable qualitative properties of the optimal value and the optimal solution set, such as (semi-) continuity, differentiability, or persistence. *Quantitative stability* tries to quantify previous properties, for example to find a convenient upper bounds for $|\varphi(\mu) - \varphi(\nu)|$ and $\mathbb{D}(\psi(\mu), \psi(\nu))$ where $\mathbb{D}$ is suitably selected set-distance. More formally we look for functions m_φ, m_ψ having convenient properties (e.g. Lipschitz, Hölder continuity, etc.) such that

$$|\varphi(\mu) - \varphi(\nu)| \leq m_\varphi(d(\mu, \nu))$$
$$\mathbb{D}(\psi(\mu), \psi(\nu)) \leq m_\psi(d(\mu, \nu))$$

where d is some function measuring "difference" between distributions μ and ν. Natural question arises: how to measure (quantify) the "difference" between two distributions, the question is examined in Chapter 3.

3 Probability metrics

Denote $\mathcal{P}(\Xi)$ the space of all probability measures defined on $\Xi \subset \mathbb{R}^s$ and $\mathcal{P}_\mathcal{F} \subset \mathcal{P}(\Xi)$ some of its subset depending on a chosen class $\mathcal{F}$ of nonlinear real functions defined on Ξ. Natural choice of a distance for the general stochastic programming problem (2) is a distance that uniformly compares expectations on a variety of nonlinear function $\mathcal{F}$. Such distance is defined for $\mu, \nu \in \mathcal{P}_\mathcal{F}$ by

$$d_\mathcal{F}(\mu, \nu) := \sup_{F \in \mathcal{F}} \left| \int_\Xi F(\xi)\mu(\mathrm{d}\xi) - \int_\Xi F(\xi)\nu(\mathrm{d}\xi) \right| \tag{6}$$

and known as the *distance having ζ-structure* or *Zolotarev pseudometric* on the space $\mathcal{P}_\mathcal{F}$ of probability measures. $d_\mathcal{F}$ is always a pseudometric; if it has finite value and the class $\mathcal{F}$ is rich enough to ensure that $d_\mathcal{F}(\mu, \nu) = 0$ implies $\mu = \nu$, then it is a metric on $\mathcal{P}_\mathcal{F}$.

Consider an open set $U \subset \mathbb{R}^n$ and define the class of functions

$$\mathcal{F}_U := \left\{ F_j(x; \cdot) : x \in X \cap \operatorname{cl} U,\ j = 0, \ldots, d \right\}$$

and the space of probability metrics corresponding to $\mathcal{F}_U$ as

$$\mathcal{P}_{\mathcal{F}_U}(\Xi) := \Bigg\{ \nu \in \mathcal{P}(\Xi) : \\ -\infty < \int_\Xi \inf_{x \in X \cap r\mathbb{B}} F_j(x;\xi) \mathrm{d}\nu(\xi) \ \text{ for each } r > 0, \\ \sup_{x \in X \cap \mathrm{cl}\, U} \int_\Xi F_j(x;\xi) \mathrm{d}\nu(\xi) < +\infty \ \text{ for each } j = 0, \dots, d \Bigg\}.$$

The conditions on ν in $\mathcal{P}_{\mathcal{F}_U}(\Xi)$ are such that expectations in (6) exist and are finite so the distance is well defined.

Definition 1. The distance

$$d_{\mathcal{F}_U}(\mu,\nu) := \sup_{\substack{j=0,\dots,d \\ x \in X \cap \mathrm{cl}\, U}} \Big| \int_\Xi F_j(x;\xi)(\mu)(\mathrm{d}\xi) - \int_\Xi F_j(x;\xi)(\nu)(\mathrm{d}\xi) \Big|$$

is called *minimal information (m. i.) metric*.

The distance $d_{\mathcal{F}_U}$ is a distance having ζ-structure. It is the "minimal distance" for the stability of the model (2) in sense that it takes information only from specific properties of the functions F_j. General stability theorems deal with this probability distance.

Minimal information metric is difficult to handle. In order to find more friendly results we look for another distance having ζ-structure by enlarging the class $\mathcal{F}$ and thus bounding $d_{\mathcal{F}_U}$ from above. Sometimes, it is also necessary to reduce the class $\mathcal{P}_\mathcal{F}$ of acceptable probability measures to ensure existence and finiteness of expectations. The common strategy of constructing an enlarged class $\mathcal{F}_{id}$ is such that $\mathcal{F}_{id}$ contains all the functions $CF_j(x;\cdot)$ for all $x \in X \cap \mathrm{cl}\, U$ and some $C > 0$ and share some typical analytical properties of F_j. We call such probability metric as an *ideal probability metric* for the predefined class of stochastic programming problems.

Recourse functionals exhibit often a local Lipschitz property. For example, we define the class of 1-*Lipschitz continuous functions* as
$\mathcal{F}_1 := \{F : \Xi \to \mathbb{R} : L_1(F) \le 1\}$ where $L_1(F)$ is a usual Lipschitz constant of the function F. Denote further $\mathcal{P}_1(\Xi)$ the class of probability measures having finite the first absolute moment, and let $\mu, \nu \in \mathcal{P}_1(\Xi)$ We define so-called *Wasserstein metric* by

$$W_1(\mu,\nu) := d_{\mathcal{F}_1} = \Big| \int_\Xi F(\xi)\mu(\mathrm{d}\xi) - \int_\Xi F(\xi)\nu(\mathrm{d}\xi) \Big|.$$

Wasserstein metric is a distance that has ζ-structure and is ideal for recourse problems where recourse functional is Lipschitz continuous. In one-dimensional space, the calculation is simple, as it simplifies to

$$W_1(\mu,\nu) = \int_{-\infty}^{+\infty} |F(t) - G(t)| \mathrm{d}t.$$

But the distance and the results do not work well with distributions having *heavy tails*; the example is given in the thesis.

General recourse functionals can grow faster than linearly; then the generalized notion of p-th order Lipschitz continuity applies resulting in *p-th order Fortet-Mourier metric* ζ_p used by more complicated recourse models (for example multistage programs); $W_1(\mu, \nu)$ coincides with $\zeta_1(\mu, \nu)$.

For **chance-constrained problems** (c.f. (5)), $\mathcal{B}$-discrepancies (or α-metrics) apply. Consider a class $\mathcal{B} = \mathcal{B}(\Xi)$ of Borel subsets of Ξ such that it contains all multifunctions H_j (i.e., constraint multifunctions; in (5), we consider only $j = 1$). We define$\mathcal{F} = \mathcal{F}_{\mathcal{B}} := \{\chi_B : B \in \mathcal{B}\}$ Due to the nature of indicator function χ the corresponding class of convenient probability measures is the whole set $\mathcal{P}(\Xi)$. The distance

$$\alpha_B = d_{\mathcal{F}_{\mathcal{B}}} = \sup_{B \in \mathcal{B}} |\mu(B) - \nu(B)|$$

is called *α-(pseudo-)metric* or, *$\mathcal{B}$-discrepancy*. A specific distance can be defined according to the nature and properties of $H_j(x)$; for example, polyhedral discrepancy, half-space discrepancy.

A special case of $\mathcal{B}$-discrepancy distance can be formed starting from a problem with random right-hand side only:

$$\min c'x \text{ subject to } x \in X, \mu\{\xi \in \Xi; g(x) \geq \xi\} \geq p$$

where $\Xi \subset \mathbb{R}^s$, $g \colon \mathbb{R}^n \to \mathbb{R}^s$, $p \in [0; 1]$. The corresponding class $\mathcal{B}$ is formed by s-dimensional "semi-closed intervals" $\mathcal{B}_{\mathrm{K}}(\Xi) := \{(-\infty; \xi], \xi \in \Xi\}$. The corresponding $\mathcal{B}$-discrepancy is called *Kolmogorov* or *Kolmogorov-Smirnov metric*; it is simply the difference of probability distributions functions F, G corresponding to μ, ν:

$$\mathcal{K}(\mu, \nu) = \sup_{B \in \mathcal{B}_{\mathrm{K}}(\Xi)} |\mu(B) - \nu(B)| = \sup_{z \in \Xi} |F(z) - G(z)|.$$

A considerable advantage of this metrics is the computational simplicity when it is applied to workaday optimization problems. Hence, it is often used in practice instead more sophisticated distances even it is not always an ideal metric. On the other hand, the Kolmogorov metric is not good choice when we deal, for example, with approximation of unknown mass points of a discrete distribution. But it deals well with distributions having heavy tails.

4 Stability in stochastic programming problems

Before stating a general stability theorem we need some additional notions to be defined. Consider a nonempty open set $U \subset \mathbb{R}^n$, $\nu \in \mathcal{P}_{\mathcal{F}_U}(\Xi)$, and denote

$$\begin{aligned} X_U(\nu) &:= \left\{ x \in X \cap \operatorname{cl} U : \int_\Xi F_j(x;\xi)\ \nu(\mathrm{d}\xi) \le 0,\ j = 1,\dots,d \right\} \\ \varphi_U(\nu) &:= \inf_{x \in X_U(\nu)} \int_\Xi F_0(x;\xi)\nu(\mathrm{d}\xi) \\ \psi_U(\nu) &:= \left\{ x \in X_U(\nu) : \int_\Xi F_0(x;\xi)\nu(\mathrm{d}\xi) = \varphi_U(\nu) \right\}. \end{aligned} \tag{7}$$

Definition 2. A nonempty set $S \subset \mathbb{R}^n$ is called *complete local minimizing (CLM) set* of the optimization problem (7) relative to U if $S = \psi_U(\nu) \subset U$.

Definition 3. Let $(\mathcal{P}, d)$ be a (semi-) metric space. We say that a set-valued mapping $S\colon (\mathcal{P}, d) \rightrightarrows \mathbb{R}^n$ is (*Berge*) *upper semicontinuous* at the point $\mu \in \mathcal{P}$ if for every open set $\mathcal{O} \subset \mathbb{R}^n$ with $S(\mu) \subset \mathcal{O}$ it holds that $S(\nu) \subset \mathcal{O}$ for each $\nu \in \mathcal{P}$ with sufficiently small value of $d(\mu, \nu)$.

Definition 4. For $\tau \ge 0$ we define the *growth function* by

$$\psi_\mu(\tau) := \min \left\{ \int_\Xi F_0(x;\xi)\mu(\mathrm{d}\xi) - \varphi(\mu) : \operatorname{dist}(x;\psi(\mu)) \ge \tau,\ x \in X_U(\mu) \right\}$$

and the *associated growth function* by $\Psi_\mu(\eta) := \eta + \psi_\mu^{-1}(2\eta),\ \eta \in \mathbb{R}_+$, where $\psi_\mu^{-1}(2\eta) := \sup\{\tau \ge 0 : \psi_\mu(\tau) \le 2\eta\}$.

Special constructions of growth functions are known under their names of linear or quadratic growth. Stronger growth conditions are also used, as strong convexity condition.

Definition 5. Given some subset $U \subset \mathbb{R}^n$, the function $c\colon \mathbb{R}^n \to \mathbb{R}$ is said to be *strongly convex* if there exists some $\kappa > 0$ such that for all $x_1, x_2 \in U$ and all $\lambda \in [0;1]$,

$$c(\lambda x_1 + (1-\lambda)x_2) \le \lambda c(x_1) + (1-\lambda)c(x_2) - \frac{1}{2}\kappa\lambda(1-\lambda)\|x_1 - x_2\|^2$$

In problems with probabilistic constraints the stability property for constraint sets $X(\mu)$ are needed. To state it we will use a notion of the special metric regularity condition as introduced in Rockafellar and Wets [12], Section 9.G.

Denote

$$\begin{aligned} X_y(\mu) &:= \left\{ x \in X : \int_\Xi F_j(x;\xi)\mu(\mathrm{d}\xi) \le y_j,\ j = 1,\dots,d \right\} \\ X_x^{-1}(\mu) &:= \left\{ y \in \mathbb{R}^d : x \in X_y(\mu) \right\} \end{aligned}$$

Definition 6. We say that the mapping $x \mapsto X_x^{-1}(\mu)$ is *metrically regular* at some pair $(\bar{x}, 0) \in \mathbb{R}^n \times \mathbb{R}^d$, $\bar{x} \in X_0(\mu)$ if there are $a \geq 0$ and $\varepsilon > 0$ such that for all $x \in X$ and $y \in \mathbb{R}^d$ with $\|x - \bar{x}\| \leq \varepsilon$ and $\max\limits_{j=1,\dots,d} |y_j| \leq \varepsilon$ it holds that

$$\operatorname{dist}(x, X_y(\mu)) \leq a \max_{j=1,\dots,d} \max \left\{ 0, \int_{\Xi} F_j(x;\xi)\mu(\mathrm{d}\xi) - y_j \right\}. \tag{8}$$

We are now ready to formulate a *general stability theorem* for the problem (2); the theorem was introduced in Henrion and Römisch [5].

Theorem 1 (Henrion and Römisch [5])**.** *Consider the stochastic programming problem* (2) *with its general assumptions. Let* $P \in \mathcal{P}_{\mathcal{F}_U}$ *and assume that* $\psi(\mu) \neq \emptyset$, U *is open* bounded *neighbourhood of* $\psi(\mu)$, $x \mapsto \int_{\Xi} F_0(x;\xi)\mu(\mathrm{d}\xi)$ *is Lipschitz on* $X \cap \operatorname{cl} U$ *if* $d \geq 1$, *and* $x \mapsto X_x^{-1}(\mu)$ *is metrically regular at each pair* $(\bar{x}; 0)$, $\bar{x} \in \psi(\mu)$. *Then* $\psi_U \colon (\mathcal{P}_{\mathcal{F}_U}, d_{\mathcal{F}_U}) \rightrightarrows \mathbb{R}^m$ *is (Berge) upper semicontinuous at* μ, *and there exist* $\delta > 0$, $L > 0$, $\hat{L} \geq 1$ *such that for* $\nu \in \mathcal{P}_{\mathcal{F}_U}(\Xi)$, $d_{\mathcal{F}_U}(\mu, \nu) < \delta$, *we have that*

$$\begin{aligned}
&|\varphi_U(\mu) - \varphi_U(\nu)| \leq L d_{\mathcal{F}_U}(\mu, \nu) \\
&\psi_U(\nu) \neq \emptyset \text{ and it is a CLM set wrt. } U \\
&\psi_U(\nu) \subset \psi(\mu) + \Psi_\mu \left(\hat{L} d_{\mathcal{F}_U}(\mu, \nu) \right) \mathbb{B}
\end{aligned}$$

This general theorem states several stability results in one with respect to the minimal information metric. Different estimates for upper bounds with more suitable metrics than the last for specific structures of stochastic programming problem are derived in the literature.

Early stability results for **recourse problems** was obtained with the β-metrics applied. Later, the assumptions have been weakened and better – Lipschitz estimates with the Wasserstein metric have been obtained. The case of the complete linear recourse is studied in Römisch and Schultz [13]. The following theorem is a possible non-linear modification based on these results. Further, the continuity properties and compactness ensure existence and boundedness of optimal solution of the original (unperturbed) problem and, at the same time, we can leave out the localization property.

Theorem 2 (Houda [7],Houda [8])**.** *Consider the program* (3), *where the following assumption are fulfilled:* X *is a compact set, the function* F_0 *is uniformly continuous on* $\mathbb{R}^n \times \mathbb{R}^s$, *and* $F_0(x;\cdot)$ *is Lipschitz continuous for all* $x \in X$ *with constant* L *not depending on* x. *Then* ψ *is (Berge) upper semicontinuous at* μ *with respect to* $(\mathcal{P}_1(\Xi), W_1)$, *and for any* $\nu \in \mathcal{P}_1(\Xi)$ *we have that* $\psi(\nu) \neq \emptyset$ *and*

$$|\varphi(\mu) - \varphi(\nu)| \leq L\, W_1(\mu, \nu).$$

Quantitative stability of optimal solution sets can be achieved by direct application of Theorem 1. Stringent assumptions as the strong convexity condition were also used earlier in the literature; generally, such bounds are not the best but could have better computation aspects.

Theorem 3 (Houda [7], Kaňková and Houda [10]). *Let assumptions of Theorem 2 be fulfilled. Let X be convex and for all $\xi \in \Xi$ let $F_0(\cdot;\xi)$ be strongly convex on $\mathbb{R}^n$ with parameter $\kappa > 0$. Then*

$$\|\psi(\mu) - \psi(\nu)\|^2 \leq \frac{8}{\kappa} L\, W(\mu,\nu) \tag{9}$$

Stability results for **chance constrained programs** are direct corollaries of Theorem 1 where appropriate $\mathcal{B}$-discrepancies are applied. The crucial assumption is the metric regularity; specific conditions implying the metric regularity are examined in context of chance-constrained programming by several authors. For example, consider a (possibly) nonlinear chance-constrained program with random right-hand side only:

$$\min c(x) \text{ subject to } x \in X,\ \mu\{\xi \in \Xi : g(x) \geq \xi\} \geq p \tag{10}$$

where $g\colon \mathbb{R}^n \to \mathbb{R}^s$, $c\colon \mathbb{R}^n \to \mathbb{R}$, and $p \in [0;1]$. X and Ξ is assumed to be as in the general stochastic programming model.

Theorem 4 (Henrion [4]). *In (10), assume c convex, X closed and convex, g having concave components, and $\mu \in \mathcal{P}(\Xi)$ is r-concave for some $r < 0$. In addition, assume that $\psi(\mu)$ is nonempty and bounded, and there exists some $\hat{x} \in X$ such that $F_\mu(g(\hat{x}) > p$ where F_μ is the distribution function of μ (Slater condition). Then $\psi(\cdot)$ is (Berge) upper semicontinuous at μ, and there exist constants $L, \delta > 0$ such that for any $\nu \in \mathcal{P}(\Xi)$ with $\mathcal{K}(\mu,\nu) < \delta$ we have that $\psi(\nu) \neq \emptyset$ and*

$$|\varphi(\mu) - \varphi(\nu)| \leq L\mathcal{K}(\mu,\nu).$$

The Slater condition and convexity assumptions ensure the metric regularity condition from the general theorem. Applying a strong convexity condition one can arrive at the stability of optimal solution set that will be of Hölder type.

5 Approximations in stochastic programming problems with recourse

Recall the bound of Theorem 2 concerning optimal values of the problems

$$|\varphi(\mu) - \varphi(\nu)| \leq L\, W_1(\mu,\nu). \tag{11}$$

Our motivation here is the following: we see the stability bound "double structured": the Lipschitz constant, representing the model structure itself; and the randomness expressed by the value of the Wasserstein metric.

If the probability measure μ, needed for succesful solution of the stochastic optimization problem, is not available, but we have empirical data at our disposition we can use them instead and replace the original distribution with the empirical version. Let $\xi_1, \xi_2, \ldots, \xi_N, \ldots$ be independent random variables with the same probability distribution μ. For notation simplicity, we denote its distribution function by F instead of F_μ.

Definition 7. The random function

$$F_N(t) = \frac{1}{N}\sum_{i=1}^{N} \chi_{(-\infty;t]}(\xi_i), \quad t \in \mathbb{R} \tag{12}$$

is called **empirical distribution function** based on the sample $\xi_1, \ldots, \xi_N$; χ_A is again the characteristic (indicator) function of the set A.

For each realization of the sample, $F_N(t)$ is actually a distribution function; we denote associated probability measure as μ_N and call it *empirical measure*. By Glivenko-Cantelli theorem and the law of large numbers, a sequence of empirical distribution functions F_N converges almost surely to the distribution function F under general conditions as N goes to infinity. Considering the definition of the Wasserstein and Kolmogorov metric, values of these metrics for F and F_N converge too.

One-dimensional Wasserstein metric involves integral computation of difference between distribution functions. To find its convergence rate at infinity we are interested in behaviour of the associated empirical process.

Definition 8. The **integrated empirical process** is defined by

$$\sqrt{N}\, W(\mu_N, \mu) = \int_{-\infty}^{+\infty} \sqrt{N}\, |F_N(t) - F(t)|\, \mathrm{d}t. \tag{13}$$

Definition 9. A stochastic process $\mathbb{U}$ is called **Brownian bridge** if it is continuous Gaussian process having mean function $\mathbb{E}\mathbb{U}(t) = 0$ and covariance function $\mathrm{cov}(\mathbb{U}(s), \mathbb{U}(t)) = \min(s,t) - st$ where $s, t \in [0;1]$.

We find the Brownian bridge as the (weak) limit of the integral (13) in case of the uniform distribution on $[0;1]$ (see Section 3.8 of Shorack and Wellner [14]):

$$\int_{-\infty}^{+\infty} \sqrt{N} \left| \frac{1}{N}\sum_{i=1}^{N} \chi_{(-\infty;t]}(\xi_i) - t \right| \mathrm{d}t \to_d \int_{-\infty}^{+\infty} |\mathbb{U}(t)| \mathrm{d}t \tag{14}$$

In this case the probability distribution of the limit of (14) is also known explicitly. To find a weak limit for distributions that are not uniform is not so easy. The Inverse Theorem with a simple substitution does not suffice because one has to involve the derivative of $F(x)$ into consideration. The condition under which the convergence is valid was found by del Barrio et al. [3].

Proposition 5 (del Barrio et al. [3], Theorem 2.1)**.** *The limit theorem* $\sqrt{N}(F_N(t) - F(t)) \to_w \mathbb{U}(F(t))$ *in* $L_1(\mathbb{R})$ *is valid if and only if* $\int_{-\infty}^{+\infty} \sqrt{F(t)(1-F(t))}\mathrm{d}t < +\infty$

From this it easily follows that, under the last condition, the following convergence result is valid:

Theorem 6. *If* $\int_{-\infty}^{+\infty} \sqrt{F(t)(1-F(t))}\mathrm{d}t < +\infty$ *then*

$$\int_{-\infty}^{+\infty} \sqrt{N} \left| \frac{1}{N} \sum_{i=1}^{N} \chi_{(-\infty;t]}(\xi_i) - F(t) \right| \mathrm{d}t \to_d \int_{-\infty}^{+\infty} |\mathbb{U}(F(t))| \mathrm{d}t \tag{15}$$

The convergence results for some "representative" distributions are illustrated. Furthermore, to extend the results, we focus not only on the independent case but also on the case of weakly dependent data. For simplicity we are not going in depth and choose most simple variant of dependency, M-dependent sequences.

Definition 10. Let $\{\xi_t\}_{-\infty}^{+\infty}$ be a random sequence defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let, moreover, $\mathcal{B}(-\infty, a)$ be the σ-algebra generated by $\ldots, \xi_{a-1}, \xi_a$, and $\mathcal{B}(b, +\infty)$ be the σ-algebra generated by $\xi_b, \xi_{b+1}, \ldots$. The sequence is said to be M**-dependent** if $\mathcal{B}(-\infty, a)$ and $\mathcal{B}(b, +\infty)$ are independent for $b - a > M$.

We can prove by the techniques employed in Kaňková [9] that the case of M-dependent samples can be transformed to the independent one and the asymptotic results remain valid.

Theorem 7. *For every natural* N *there exists* $k \in \{0, 1, \ldots\}$ *and* $r \in \{1, \ldots, M\}$ *such that* $N = Mk + r$*, and*

$$|F_N(t) - F(t)| \leq \sum_{j=1}^{M} \frac{N_j}{N} |F_{N_j}(t) - F(t)|, \quad t \in \mathbb{R} \tag{16}$$

where F_{N_j} *are empirical distribution functions determined by sequences of* N_j *independent random variables, and*

$$N_j = \begin{cases} k+1 & \text{for } j = 1, \ldots, r \\ k & \text{for } j = r+1, \ldots, M \end{cases}$$

The numerical study confirms the expected results: both Wasserstein and Kolmogorov metrics vanish to zero as number of samples grows. It is true even for cut Cauchy distribution representing a distribution with heavy tails. Convergence of Kolmogorov's empirical process's is confirmed to be the same regardless the original distribution (even for Cauchy distribution). For Wasserstein's integrated empirical process, the rate of convergence is stabilizing quickly in the case of the uniform, normal, and exponential distributions. The problem of Cauchy distribution is easily justifiable: although cutting allowed us to use the Wasserstein distance, numerical properties of the distance would to be hardly favourable. The results with dependent data are seen to differ slightly but one can appreciated the difference as not very important. As members of used MA(1) process are actually weakly dependent, this behaviour is expected again.

6 Convexity and structural dependence in chance-constrained problems

In Chapter 6 of the thesi we focus our attention to a slightly modified formulation of chance-constrained problem:

$$\min c(x) \text{ subject to } \mathbb{P}\{h(x;\xi) \geq 0\} \geq p \tag{17}$$

where $x \in \mathbb{R}^n$, $\xi : \Omega \to \mathbb{R}^s$ is a random vector with the distribution μ and distribution function F, $c\colon \mathbb{R}^n \to \mathbb{R}$, $h\colon \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^d$, and $p \in [0;1]$ is a prescribed probability level. To emphasize importance of the value p in this chapter, we denote the constraint set as

$$M(p) := \big\{x \in \mathbb{R}^n : \mathbb{P}\{h(x;\xi)\} \geq p\big\}.$$

A matter of importance both in theoretical and practical applications of the chance-constrained programming is to determine when the set $M(p)$ of feasible solutions is convex. It is trivially known that the sets $M(0)$, $M(1)$ are convex if $h(\cdot,\xi)$ are concave functions for all $\xi \in \mathbb{R}^s$. Classical result of Prékopa (see Prékopa [11] for overview) states that if μ is absolutely continuous (with respect to Lebesgue measure), log-concave measure (or r-concave with $r \geq -1/s$), and the one-dimensional components of h are quasi-concave functions of (x,ξ) then $M(p)$ is a convex set.

The quasi-concavity property is not preserved under addition. Considering a problem with random right-hand side in the form

$$\min c(x) \text{ subject to } \mathbb{P}\{g(x) \geq \xi\} \geq p, \tag{18}$$

h defined as $h(x;\xi) = g(x) - \xi$ is quasi-concave if $g(x)$ is concave. Recently, Henrion and Strugarek [6] proposed an alternative approach to deal with this problem: their idea is to relax concavity condition of g and make more stringent concavity condition on the probability distribution μ.

Definition 11. A function $g\colon \mathbb{R}^n \to (0;+\infty)$ is called r-concave for some $r \in [-\infty;+\infty]$ if

$$g(\lambda x + (1-\lambda)y) \geq \big[\lambda g^r(x) + (1-\lambda)g^r(y)\big]^{1/r}$$

for all $x, y \in \mathbb{R}^n$ and all $\lambda \in [0;1]$. Cases $r = -\infty, 0$, and $+\infty$ are treated by continuity.

Definition 12 (Henrion and Strugarek [6], Definition 2.2)**.** A function $f\colon \mathbb{R} \to \mathbb{R}$ is called *r-decreasing* for some $r \in \mathbb{R}$ if it is continuous on $(0;+\infty)$, and there exists a threshold $t* > 0$ such that $t^r f(t)$ is strictly decreasing for all $t > t*$.

Theorem 8 (Henrion and Strugarek [6], Theorem 3.1)**.** *If there exist $r_i > 0$ such that the components g_i of g are $(-r_i)$-concave, the components ξ_i of ξ have $r_i + 1$-decreasing densities, and the components ξ_i of ξ are independently distributed, then $M(p)$ is convex for all $p > p^* := \max_i F_i(t_i^*)$ where F_i denotes the distribution function of ξ_i and t_i^* refer to the definition of $r_i + 1$-decreasing probability density.*

In the sequel we ask for the relaxation of the independence condition in Theorem 8. To do this, we define an α' coefficient of dependence by the following definition.

Definition 13. For a random vector ξ we define a coefficient α' of (weak) dependence as

$$\alpha' := \sup_z \left| F(z) - \prod_i F_i(z_i) \right| \tag{19}$$

where F is the distribution function of the vector ξ, F_i are the corresponding one-dimensional marginal distribution functions and $z = (z_1, \ldots, z_s) \in \mathbb{R}^s$.

Denote $M'(p) = \left\{ x \in X : \prod_{i=1}^s F_i(g_i(x)) \geq p \right\}$. If the components ξ_i of ξ are independently distributed, then $M(p) = M'(p)$. This is not true in case of weak dependence, but the following proposition is valid:

Proposition 9. *If the components ξ_i of ξ in* (18) *are α-dependent (in the sense of Definition 13) then*

$$M'(p+\alpha) \subset M(p) \subset M'(p-\alpha) \subset M(p-2\alpha). \tag{20}$$

If p is sufficiently high, then possibly non-convex $M(p)$ is bounded from both side by convex sets and the following theorem is valid.

Theorem 10. *If there exist $r_i > 0$ such that the components g_i of g are $(-r_i)$-concave, the components ξ_i of ξ have $r_i + 1$-decreasing densities, the components ξ_i of ξ are α-dependently distributed, and $p > \max_i F_i(t_i^*) + \alpha$ (with t_i^* and F_i as in Theorem 8), then $M(p)$ is bounded (from both sides) by* ***convex*** *sets $M'(p+\alpha)$ and $M'(p-\alpha)$.*

Under additional assumptions on objective and constraint functions, we can prove the following theorem.

Theorem 11. *Consider the problem* (18) *and let assumption 1–4 of Theorem 10 be fulfilled. Further assume c to be Lipschitz continuous function on $\mathbb{R}^n$, the mapping $x \mapsto \{y \in \mathbb{R} | x \in M'(p+\alpha-y)\}$ be metrically regular at $(\bar{x}'(p+\alpha), 0)$, and α-dependence coefficient satisfyigng $\alpha < \frac{1}{2}\varepsilon$, where ε is provided by metric regularity condition. Then there exists a constant $L > 0$ such that*

$$|\varphi'(p+\alpha) - \varphi(p)| \leq L \max\left\{0, p + \alpha - \prod F_i(g_i(\bar{x}'(p-\alpha)))\right\} \tag{21}$$

is fulfilled. Here, $\varphi(p)$, $\varphi'(p+\alpha)$ are the optimal values, and $\bar{x}'(p-\alpha)$, $\bar{x}'(p+\alpha)$ are the optimal solutions of the dependent and the independent problem respectively.

A simple example originating in Henrion and Strugarek [6] shows that for the chosen normal distribution, convexity of the feasible set is assured theoretically at the probability level of 0.921 for the independent case (the actual probability level seems to be around the value of 0.7). For the weak dependent case, these tresholds (theoretical and actual) are shifted towards the center of feasibility sets (center of image), and the optimal values and optimal solutions remain stable as the value of α-coefficient is small.

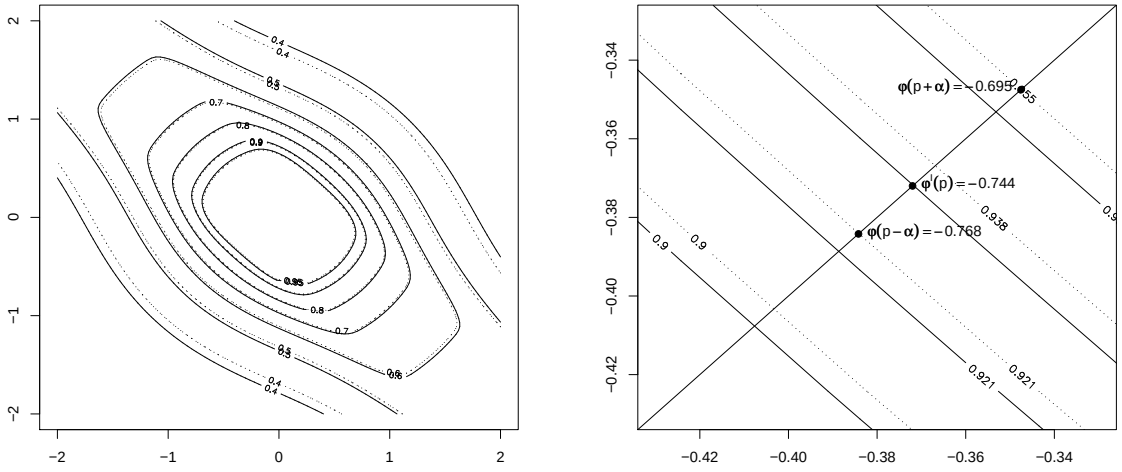


Figure 1: Contour lines for $M'(p)$ (solid) and $M(p)$ (dotted) sets

7 Chance-constrained problems and robust programming approach

Techniques of *robust programming* have became a popular alternative to stochastic programming during the last two decades. The robust programming handles the uncertainty in an optimization model through the so-called "worst case" analysis: the constraints of the general optimization model (1) are required to be satisfied for all possible realizations of uncertainty parameter ξ, and we optimize the worst-case objective function among all robust solutions. We consider the following "standardized" form of the *uncertain convex program*:

$$\text{minimize } c'x \text{ subject to } x \in X, f(x;\xi) \leq 0, \tag{22}$$

where $x \in X \subset \mathbb{R}^n$, $c \in \mathbb{R}^n$, and $\xi \in \Xi \subset \mathbb{R}^s$ is a parameter (data) vector. We assume further that X is convex and closed set, $f\colon X \times \Xi \to \mathbb{R}$ is convex in x for all $\xi \in \Xi$, and Ξ is a prescribed set of instances. Without lost of generality, the objective function is linear and the constraint function f is scalar-valued. If the realization of ξ is known and fixed, the problem (22) is deterministic convex program and we can use techniques of convex programming to solve the problem.

Assuming that ξ is a random vector defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with known probability distribution $\mu \in \mathcal{P}(\Xi)$, in *chance-constrained program* (PCP) we require the constraints of (22) to be fulfilled with a prescribed level of probability $1-\varepsilon$. The problem reads

$$\text{minimize } c'x \text{ subject to } x \in X_\varepsilon := \big\{x \in X : \mu\{\xi \in \Xi : f(x;\xi) > 0\} \leq \varepsilon\big\}. \tag{23}$$

The problem (23) or its approximation need not to be convex even if f is convex in x for all ξ, and to evaluate the probability in the definition of X_ε one often has to calculate values of multidimensional integrals.

In *robust programming approach*, also known as 'min-max' or 'worst-case' approach, we look for a solution which is feasible for *all* possible instances of ξ; this approach leads to the *robust convex problem* (RCP):

$$\min_{x\in X} c'x \text{ subject to } f(x;\xi) \leq 0 \text{ for all } \xi \in \Xi. \tag{24}$$

The problem (24) is convex but it has an infinite number of constraints and so it is numerically hard to solve. Another disadvantage of robust programming in its original form is the fact that it allocate the same weight to all values of the parameter regardless possible different importance of individual instances of ξ.

Consider a set of independent samples $\xi_1, \ldots, \xi_N$ distributed according to μ, the original distribution of the parameter ξ. The problem (23) is approximated, for the given sample, by replacing the original probability distribution μ by the empirical probability measure μ_N associated with the empirical distribution function F_N, and then the problem reads

$$\min_{x\in X} c'x \text{ subject to } x \in X[\varepsilon, N] := \left\{x \in X; \frac{1}{N}\,\mathrm{card}\{i; f(x;\xi_i) > 0\} \leq \varepsilon\right\} \tag{25}$$

where card denotes the cardinality of the argument. We refer to the problem (25) as to the *chance-constrained sampled problem* (PCP_N). The essential idea of (25) is that the relative frequency of constraint violations corresponds to the desired upper level of infeasibility in (23). (25) is the program with a single constraint and in some simple cases it is computationally tractable.

Recently, Calafiore and Campi [2] proposed the following approximation to the robust convex program (24) by the following *robust sampled convex problem* (SCP_N):

$$\min_{x\in X} c'x \text{ subject to } X[N] := \left\{x \in X; f(x;\xi_i) \leq 0 \text{ for } i = 1, \ldots, N\right\}. \tag{26}$$

This is a relaxation of the original robust convex problem: we do not require the original constraints to be satisfied for all realizations of $\xi \in \Xi$ but only for a certain finite but sufficiently large number of samples which are moreover the most probable to happen. Calafiore and Campi [1] found a rule to set up N in order to have the optimal solution of (26) feasible in (23). The problem is convex, it has a finite number of constraints and it is effectively computable. In addition, it incorporates weights to the individual parameter instances of ξ and these are such that are most probable to happen.

The two mentioned approaches to solve uncertain convex program are based on a different "philosophy" how to understand the uncertainty. Consider the following uncertain convex program

$$\text{minimize } x \text{ subject to } x \geq \xi, x \in \mathbb{R} \tag{27}$$

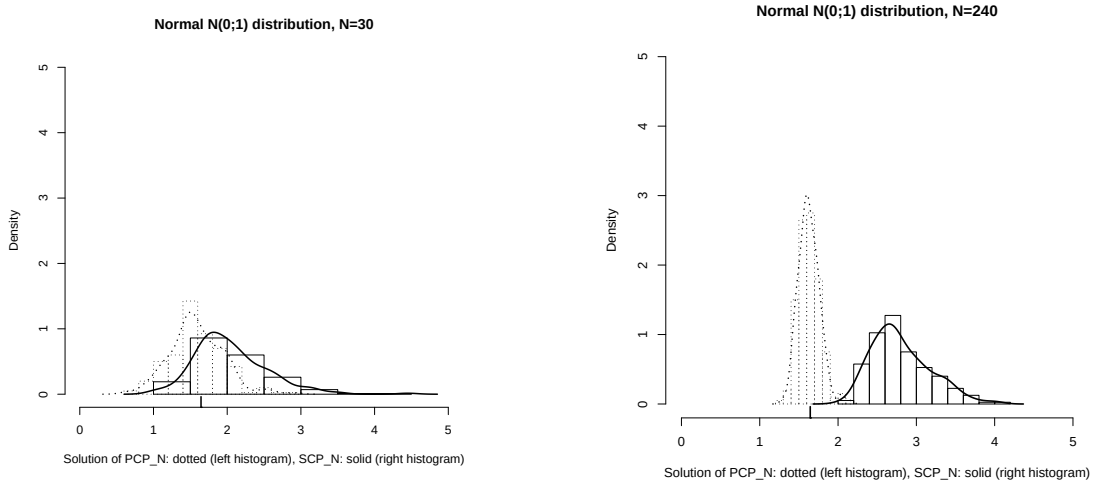


Figure 2: Convergence of optimal values for (SCP_N) and (PCP_N)

where $\xi \in \mathbb{R}$ is distributed according to a standard normal distribution $\mathrm{N}(0;1)$; We denote F its distribution function. Further, we define the following deterministic programs

$$\min_{x\in\mathbb{R}} x \text{ subject to } x \geq F^{-1}(1-\varepsilon), \tag{28}$$

$$\min_{x\in\mathbb{R}} x \text{ subject to } \frac{1}{N}\,\mathrm{card}\{i; x < \xi_i\} \leq \varepsilon, \tag{29}$$

$$\min_{x\in\mathbb{R}} x \text{ subject to } x \geq \max_{i=1,\ldots,N} \xi_i. \tag{30}$$

(28) is chance-constrained problem (PCP), (29) is sampled version of it (PCP_N), and (30) is robust sampled problem (SCP_N). Normal distribution is defined on an unbounded set, hence the missing robust program (RCP) is not well defined – there is no real solution feasible to all the instances of ξ.

For each of the three above problems the optimal solution coincides with the lower boundary of their feasibility sets. That is, the optimal solution of the chance-constrained problem (PCP) related to (27) is $1-\varepsilon$ quantile of F, the optimal solution of the chance-constrained sampled problem (PCP_N) is computed as the lowest from N samples which are greater or equal to the 0.95 sample quantile, and the optimal solution to (SCP_N) is the highest of N samples considered.

To compare the approaches we set up $\varepsilon = \beta = 0.05$. The optimal solution of (28), marked by a small tickmark on x-axis, is approximately 1.64. The number of samples assuring that the optimal solution of (SCP_N) is 0.05-feasible with probability 0.95 is about 240. Thus, we consider three values of the sample size $N = 30, 240, 3000$, representing low, "accurate", and large sample size.

The optimal solution of the chance-constrained sampled problem (29) converges, as N goes to infinity, to the solution of (28) (marked by the tickmark). The sampling method related to (25) is useful especially if the number N of samples is high, as a possible error in estimating optimal solution of the chance-constrained problem (23) decreases. On the other hand, the robustly sampled optimal solutions of (30) (solid histograms) are getting away from the point of the optimal solution of (PCP) as far as the number of samples increases, and are going to the upper boundary of the support of F (i. e. to the infinity in both of our cases), but with rapidly decreasing rate. If N is greater than 241, the optimal value of the (SCP_N) program is feasible in (PCP) with probability of 0.95.

The presented approaches exhibit very different numerical results. The models of stochastic programming have in common that we estimate the probability distribution by means of observations from the past. The resulting solution in (PCP_N) is an approximation to the (unknown) solution of (PCP) and the approximation is better as the number of samples (observations) is higher. Furthermore, our solution of the chance-constrained "sampled" problem is only approximatively ε-feasible for a given level ε. On the other hand, this level is usually not crucial for real applications if our preferences are pointed towards costs saving solutions.

The optimal solutions to robust problems hedge against the worst-case realization of uncertain parameters regardless their "importance". The randomized (sampled) approach incorporate the information about importance to the model via probability distribution of samples so that the optimal solution of the sampled problem does not have to satisfy the constraints for all possible realizations of the parameter. At the same time, the probability of such violation is small and for a given ε one could easily compute the number of samples to generate in order to obtain an ε-feasible solution. Indeed, if the number of samples is significantly greater than required, the optimal solution of the sampled problem also hedges against the parameters with the smaller probability of occurrence. This could be the task if the risk of constraint violation has to be minimized as much as possible and costs of doing that are of smaller importance.

Choosing the approximation method to an uncertain convex program is ambiguous. The selected method has to fill up the needs of practical dimension of the problem:

- how much the probability of violation of the constraints is crucial,
- how many samples one has at disposition or can generate.

Getting an answer to the first question stronger, one's preferences have to be directed towards the robust sampled problems. On the other hand the solution of the stochastic program could be useful in cases where the $1 - \varepsilon$ level is not crucial and our preferences are pointed more likely towards costs savings solutions.

References

[1] Calafiore, G. and Campi, M. C. (2004). Decision making in an uncertain environment: the scenario based optimization approach. In Andrýsek, J., Kárný, M., and Kracík, J., editors, *Multiple Participant Decision Making*, pages 99–111. Advanced Knowledge International.

[2] Calafiore, G. and Campi, M. C. (2005). Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming, Ser. A*, 102(1):25–46.

[3] del Barrio, E., Giné, E., and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *The Annals of Probability*, 27(2):1009–1071.

[4] Henrion, R. (2000). Qualitative stability of convex programs with probabilistic constraints. In Nguyen, V. H., Strodiot, J.-J., and Tossings, P., editors, *Optimization, Lecture Notes in Economics and Mathematical Systems*, volume 481, pages 164–180. Springer, Berlin.

[5] Henrion, R. and Römisch, W. (1999). Metric regularity and quantitative stability in stochastic programs with probabilistic constraints. *Mathematical Programming*, 84(1):55–88.

[6] Henrion, R. and Strugarek, C. (2008). Convexity of chance constraints with independent random variables. *Computational Optimization and Applications*, 41(2):263–276.

[7] Houda, M. (2001). Stability and estimates in stochastic programming problems (special cases). Diploma thesis, Charles University in Prague, Faculty of Mathematics and Physics. In Czech.

[8] Houda, M. (2002). Probability metrics and the stability of stochastic programs with recourse. *Bulletin of the Czech Econometric Society*, 9(17):65–77.

[9] Kaňková, V. (1994). A note on estimates in stochastic programming. *Journal of Computational and Applied Mathematics*, 56(1-2):97–112.

[10] Kaňková, V. and Houda, M. (2003). A note on quantitative stability and empirical estimates in stochastic programming. In Leopold-Wildburger, U., Rendl, F., and Wascher, G., editors, *Operations Research Proceedings 2002*, pages 413–418, Berlin. Springer.

[11] Prékopa, A. (1995). *Stochastic Programming*. Akadémiai Kiadó, Budapest.

[12] Rockafellar, R. T. and Wets, R. J. B. (1997). *Variational Analysis*. Springer-Verlag, Berlin.

[13] Römisch, W. and Schultz, R. (1993). Stability of solutions for stochastic programs with complete recourse. *Mathematics of Operations Research*, 18(3):590–609.

[14] Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

List of author's publications

[15] Houda, M. (2002). Stability and metrics in stochastic programming problems. In Šafránková, editor, *Proceedings of Week of Doctoral Studies 2002, Part I – Mathematics and Computer Sciences*, pages 204–209, Prague. Charles University, Faculty of Mathematics and Physics.

[16] Houda, M. (2003). On quantitative stability in stochastic programming problems with recourse. In Houška, M., editor, *Proceedings of the 21st International Conference Mathematical Methods in Economics 2003*, pages 117–121, Prague. Czech University of Agriculture.

[17] Houda, M. (2004). Wasserstein metrics and empirical distributions in stability of stochastic programs. In Lukáčik, M., editor, *Proceedings of the International Conference Quantitative Methods in Economics (Multiple Criteria Decision Making XII)*, pages 71–77, Bratislava. University of Economics.

[18] Houda, M. (2005a). Estimation in chance-constrained problem. In Skalská, H., editor, *Proceedings of the 23rd International Conference Mathematical Methods in Economics 2005*, pages 134–139, Hradec Králové. Gaudeamus.

[19] Houda, M. (2005b). Using metrics in stability of stochastic programming problems. *Acta Oeconomica Pragensia*, 13(1):128–134.

[20] Houda, M. (2006a). Approximations in stochastic and robust programming problems. In Lukáš, L., editor, *Proceedings of the 24th International Conference Mathematical Methods in Economics 2006*, pages 249–254, Pilsen. University of West Bohemia.

[21] Houda, M. (2006b). Comparison of approximation in stochastic and robust programming. In Hušková, M. and Janžura, M., editors, *Proceedings of the 7th Prague Symposium on Asymptotic Statistics and 15th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes Prague Stochastics 2006*, pages 418–425, Prague. MatfyzPress.

[22] Houda, M. (2007a). Convexity and dependence in chance-constrained programming. Research Report 2190, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic.

[23] Houda, M. (2007b). Role of dependence in chance-constrained and robust programming. *Acta Oeconomica Pragensia*, 15(4):111–120.

[24] Houda, M. (2008). Weak structural dependence in chance-constrained programming. In Řehořová, P. and Maršíková, K., editors, *Proceedings of the 26th International Conference on Mathematical Methods in Economics 2008*, pages 198–203, Liberec. Technical University of Liberec, Faculty of Economics.

[25] Houda, M. and Kaňková, V. (2006). Empirical processes in stochastic programming. In Hušková, M. and Janžura, M., editors, *Proceedings of the 7th Prague Symposium on Asymptotic Statistics and 15th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes Prague Stochastics 2006*, pages 426–436, Prague. MatfyzPress.

[26] Kaňková, V. and Houda, M. (2003). A note on quantitative stability and empirical estimates in stochastic programming. In Leopold-Wildburger, U., Rendl, F., and Wascher, G., editors, *Operations Research Proceedings 2002*, pages 413–418, Berlin. Springer.

[27] Kaňková, V. and Houda, M. (2006). Dependent samples in empirical estimation of stochastic programming problems. *Austrian Journal of Statistics*, 35(2&3):271–279.