Annals of Operations Research 27 (1990) 115–142

STABILITY AND SENSITIVITY-ANALYSIS FOR STOCHASTIC PROGRAMMING

Jitka DUPAČOVÁ

Charles University, Prague, Czechoslovakia

Abstract

Stability and sensitivity studies for stochastic programs have been motivated by the problem of incomplete information about the true probability measure through which the stochastic program is formulated and in connection with the development and evaluation of algorithms. The first part of this survey paper briefly introduces and compares different approaches and points out the contemporary efforts to remove and weaken assumptions that are not realistic (e.g., strict complementarity conditions). The second part surveys recent results on qualitative and quantitative stability with respect to the underlying probability measure and describes the ways and means of statistical sensitivity analysis based on Gâteaux derivatives. The last section comments on parallel statistical sensitivity results obtained in the parametric case, i.e., for probability measures belonging to a parametric family indexed by a finite dimensional vector parameter.

Keywords: Qualitative stability for SP, quantitative stability for SP, statistical sensitivity analysis, Gâteaux derivatives, asymptotic behavior.

1. Introduction

A detailed insight into the origin of surprisingly many linear programming problems reveals that the assumption of fixed, completely known coefficient values is not justified in practice. This fact was already recognized in the fifties and stimulated the development of stochastic linear programming (e.g. [5,7,9,65]). First of all, it was necessary to clarify carefully the meaning of a linear program in which random coefficients appear and to introduce completely new solution concepts. In 1955–1965, the basic ideas for the development of different approaches were elaborated and extended to stochastic nonlinear programming problems. At the same time, the first applications were successfully solved. In the next decade remarkable theoretical results were achieved. These have been collected in various works, e.g. in the monograph of Kall [30]. For new developments see e.g. Wets [72]. The main interest appears to be concentrated on designing efficient algorithms (cf. Ermoliev and Wets [19]), on proper treatment of incomplete information about the probability distribution of random parameters and on developing dynamic stochastic programming models.

Present knowledge gives a good basis for nontrivial applications of stochastic programming in which the complexity of stochastic programming problems is of concern. The applications have to reflect the interplay between the available statistical data, the chosen model and the numerical approaches based on approximation techniques. Results on stability and sensitivity can help to obtain (probabilistic) estimates of errors and to support development of new algorithms; cf. Robinson and Wets [50] in connection with scenario analysis and designing discretization schemes, Dantzig and Glynn [10] in connection with implementation of sampling techniques, etc.

Consider the nonlinear programming problem

$$\begin{aligned} &\text{minimize} && c_0(x, \omega) \\ &\text{subject to} && c_i(x, \omega) \leqq 0, \quad i = 1, \ldots, r \\ &&& c_i(x, \omega) = 0, \quad i = r+1, \ldots, r+s \\ &&& x \in M_0 \subset \mathbb{R}^n, \end{aligned} \tag{1}$$

where ω is a random vector and an optimal decision $\hat{x} \in M_0$ has to be chosen *before* a realization of ω is observed. For this purpose, (1) has to be reformulated in a meaningful way. The commonly used *basic assumptions* are:

(i) $(\Omega, \mathscr{B}, P)$ is a given probability space with $\Omega \subset \mathbb{R}^l$, $\mathscr{B}$ the corresponding Borel σ-field on Ω and P a known probability measure. Moreover, P does not depend on x.

(ii) $M_0 \subset \mathbb{R}^n$ is a given Borel set and $c_i : M_0 \times \Omega \to \mathbb{R}^1$, $i = 0, \ldots, r+s$ are given functions such that $c_i(x, \cdot)$, $i = 0, \ldots, r+s$ are random variables for all $x \in M_0$.

To build a decision model corresponding to (1) means to specify the set of feasible decisions, say $M(P) \subset M_0$ and to define a real objective function $f(\cdot, P)$: $M(P) \to \mathbb{R}^1$ that generates (independently of individual realizations of ω) a preference relation on $M(P)$. Finally, the mathematical program

$$\text{minimize } f(x, P) \text{ on the set } M(P) \tag{2}$$

is solved to get the optimal decision.

There are numerous a priori equally proper ways to get program (2): We mention only the *models with probability (or chance) constraints* with $M(P) \subset M_0$ defined through conditions

$$g_i(x, P) := \alpha_i - P\{c_k(x, \omega) \leqq 0,\ k \in I_i\} \leqq 0, \quad i = 1, \ldots, m, \tag{3}$$

with $I_i \subset \{1, \ldots, r\}$, $\alpha_i \in (0, 1)$, $i = 1, \ldots, m$, and the *penalty models* that include also the two stage or recourse stochastic programs and for which $M(P) = M \subset M_0$ is a fixed set independent of P. The objective function in penalty models

$$f(x, P) = E_P\{c_0(x, \omega) + q(x, \omega)\} \tag{4}$$

contains a penalty term $q(x, \omega)$ that evaluates the loss due to violation of constraints by a chosen decision $x \in M$ for an observed realization of ω.

The function $q: M \times \Omega \to \mathbb{R}^1$ can be given explicitly or implicitly as the optimal value function of a second stage program. We shall mention two examples of this type that will appear later in the description of stability results; for other examples see e.g. Kall [30].

The penalty function

$$q(x, \omega) := \min\left\{ q^{\mathrm{T}} y \colon \sum_{j=1}^{k} w_{ij} y_j = c_i(x, \omega),\ i = 1, \ldots, r+s,\ y \geqq 0 \right\} \tag{5}$$

with a given matrix $W = (w_{ij})$ and a given k-dimensional vector q such that

$$\{ y : Wy = z,\ y \geqq 0 \} \neq \emptyset \quad \text{for every } z \in \mathbb{R}^{r+s}$$

and that

$$\{ u : W^{\mathrm{T}} u \leqq q \} \neq \emptyset,$$

corresponds to the *complete recourse problem*.

For (1) with *linear* constraints, say

$$Ax = b, \quad x \in M_0 \subset \mathbb{R}^n,$$

where M_0 is a nonempty convex polyhedral set and ω contains all random elements of matrix A and vector b, a *quadratic recourse function*

$$q(x, \omega) = \max_{y \in Y} \left\{ (b - Ax)^{\mathrm{T}} y - \tfrac{1}{2} y^{\mathrm{T}} B y \right\} \tag{6}$$

was introduced by Rockafellar and Wets [53]. In (6), Y is a nonempty convex polyhedral set and B is a given, symmetric positive definite matrix.

The choice of the model should stem from the nature of the solved problem and, of course, it is often influenced by the structure of data, by the software available and by the decision maker's attitude as well. To solve program (2), one can in principle rely on the well-known nonlinear programming solution techniques. However, the dependence of the objective function and/or of the set of feasible decisions on the probability measure P means that even the evaluation of the function values can be a rather demanding procedure. That is why in many algorithms the probability measure P is approximated by a simpler one and an approximate program is solved instead of the original one (see e.g. [19,31,71]).

The common belief is that a small change of probability measure P (due to approximation or estimation) does not cause a large change of the optimal value

$$\phi(P) = \inf\{ f(x, P) \colon x \in M(P) \}$$

and of the set of optimal solutions

$$X(P) = \{ x \in M(P) \colon f(x, P) = \phi(P) \}$$

of program (2). Unfortunately, this does not come true in general, which does not surprise anyone familiar with parametric programming or with robust estimation in statistics.

The need for analysis of the behaviour of optimal solutions and of the optimal value of approximated or perturbed nonlinear programming problems was one of the impulses for the development of parametric programming (see e.g. [23,46]). In spite of a similar motivation, the first stability and sensitivity studies for stochastic programming with respect to the probability measures developed independently. They were raised by the fact that in real life situations, the probability measure P is hardly known completely so that the program (2) is mostly solved for an estimate P' of the true probability measure P.

The first attempts [28,73] treated simple penalty models with probability measure P belonging to a specified set $\mathscr{P}^*$ of probability measures defined by prescribed values of moments. The original results were later essentially extended and utilized in designing algorithmic procedures [6,31,34,35].

The *statistical approach* followed in the seventies [11,36,70]:

Based on statistical sample data, the probability measure P can be estimated by an empirical probability measure P_N and the problem (2) solved with P_N in place of P. Consequently, the optimal value $\phi(P_N)$ and the set of optimal solutions $X(P_N)$ can be considered as estimates of the true $\phi(P)$ and $X(P)$. The quality of such an estimate is related to its large sample properties such as consistency (considered to be the minimal requirement), asymptotic normality, etc. Small sample properties are of interest as well, but this area has not yet been explored in connection with stochastic programming.

The consistency of $\phi(P_N)$ was one of the first results [11,36]:

For $f(x, P) := \int_\Omega h(x, \omega)P(d\omega)$ with h bounded and continuous, for $M(P) = M \neq \emptyset$ convex, compact and for P_N the empirical probability measure based on independent identically distributed observations (or, more generally, on an ergodic sequence of observations) of ω, one gets

$$P\{\phi(P_N) \to \phi(P) \text{ as } N \to \infty\} = 1. \tag{7}$$

For h Lipschitzian in x with the Lipschitz constant independent of ω it is even possible to prove an *exponential rate of convergence* of the estimate [36].

The results were recently extended by Kaňková to the case of probability constraints [39] and to more complicated empirical probability measures [38]. Nevertheless, her approach is based on uniform convergence of the functions involved. (The pointwise convergence of function values does not imply the convergence of the optimal values; for a discussion see Kall [32]). Another line of methodological attack uses the concept of epi-convergence that proved to be quite powerful and convenient; for a very general consistency result for the optimal values $\phi(P_N)$ and for the sets of optimal solutions $X(P_N)$, see Dupačová and Wets [18].

As to the rate of convergence of optimal solutions, Tsybakov [66] and Vogel [67] gave conditions under which the probability of large deviations between the sets of optimal solutions $X(P)$ and $X(P_N)$ is of a given order. See also Tamm [64] for a result in this direction.

The results concerning the asymptotic distribution of $\phi(P_N)$ and $X(P_N)$ are still in progress, see e.g. Dupačová and Wets [18], King [40–42], Shapiro [61]; one of the goals is constructing confidence intervals and regions for the true optimal value and for the true optimal solution.

Results on stability and sensitivity for program (2) with respect to the *parameters* of the underlying probability measure are connected with certain results in parametric programming, some of which will be extensively reviewed in this volume, see also Armacost and Fiacco [1], Dupačová [15], Attouch and Wets [2]. As the parameter values are typically statistical estimates of the true ones, the results of parametric programming have been complemented by statistical inference [12,15,60,67]:

Similarly as in parametric programming, the first results were obtained for $M(P)$ given explicitly by inequality and equation constraints, under linear independence condition and under a suitable second order sufficient condition (to guarantee uniqueness of the optimal solutions and of the corresponding Lagrange multipliers) and under the strict complementarity conditions that reduce the local stability studies for the given problem to stability studies of a program constrained to an affine subspace [48]. For a survey of these results see Dupačová ([14], section 2).

However, the assumptions listed above are too strong. They apply to stochastic program (2) with the true probability measure (indexed by the true but unknown parameter vector), so that in contrast to many parametric programming problems these assumptions cannot be verified. Besides, in contrast to estimation in statistics, even the assumed uniqueness of the "true" optimal solution cannot be fully accepted. For recent results that aim at removing some of the mentioned assumptions see Shapiro [60], Dupačová [16,17] and section 3 of this paper.

From the point of view of our problem setting, the most natural idea is to consider *probability measure P in program (2) as the parameter*. Equipped with the weak topology, the space $\mathscr{P}$ of probability measures on $(\Omega, \mathscr{B})$ is a complete, separable space that can be metrized by Prohorov or bounded Lipschitz metric (see Huber [27]). However, it is not a linear space so that results of parametric programming with parameters belonging to a linear metric space cannot be applied. We can rely only on qualitative stability results for parametric programming with parameters belonging to a general metric or topological space [4,49] such as continuity of the optimal value $\phi(P)$ and of the set of optimal solutions $X(P)$. Moreover, for a suitable choice of a subset of probability measures, *quantitative stability results* based on a Lipschitz or Hölder property can be obtained [54–56]; see also section 2.2. Also, stability results with respect to a pair of parameters (P, p), $P \in \mathscr{P}$ and p a real vector, seem to be of increasing importance; cf. Robinson and Wets [50] for penalty models, Römisch and Schultz [56] for models with probability constraints.

For sensitivity results or for postoptimality analysis, one can compute *Gâteaux derivatives* of $\phi(P)$ and of the (unique) optimal solutions under additional assumptions about program (2) and to get a link with the statistical approach, cf. Dupačová [13,14], Shapiro [62,63] and section 2.3.

There are intrinsic connections between stability and sensitivity results obtained by means of the three seemingly different approaches mentioned above: Results on asymptotic properties of statistical estimators $\phi(P_N)$ and $x(P_N)$ can be applied to the case of estimated parameters for the given parametric family of probability measures, Gâteaux derivatives can be obtained as a special sensitivity result concerning a scalar parameter and, on the other hand, they can be used as a heuristic tool for statistical approach.

The developments up to 1985, except for the statistical approach, are mostly covered by the survey paper by Dupačová [14]. The present paper continues in a similar direction and gives a survey of recent developments in stability and sensitivity analysis for stochastic programming that are connected with results obtained in parametric programming; the statistical approach falls beyond the scope of the paper. In spite of the evident progress during the last years, there are still many open problems, for instance in connection with stability for models with probability constraints and with sensitivity analysis for optimal solutions.

2. Stability and sensitivity analysis with respect to the probability measure

2.1. SELECTED QUALITATIVE STABILITY RESULTS

To study the stability of program (2) with respect to probability measure P means to get at first qualitative results about the behavior of the optimal value $\phi(P)$ and of the set of optimal solutions $X(P)$ such as their "continuity" at P; for different concepts of semicontinuity of multifunctions such as Hausdorff upper and lower semicontinuity (H-u.sc., H-l.sc.) or Berge upper and lower semicontinuity (B-u.sc., B-l.sc.) consult, e.g., Bank et al. [4].

In our case the parameter space is a subset of the space $\mathscr{P}$ of probability measures on $(\Omega, \mathscr{B})$ endowed with the topology of weak convergence of probability measures (see definition 1 below). We can rely on general results such as theorems 4.2.2, 4.2.3 and 4.3.3 of Bank et al. [4] formulated below for program (2) as theorems 1–3.

THEOREM 1

Let M be continuous at $P_0 \in \mathscr{P}$ with $M(P_0) \neq \emptyset$ and compact. Let f be lower semicontinuous on $M(P_0) \times \{P_0\}$ and such that f is upper semicontinuous at a point (x_0, P_0) with $x_0 \in X(P_0)$.

Then ϕ is continuous at P_0 and X is B-u.sc. at P_0.

The assumption of $M(P_0)$ compact can be weakened, e.g., for $f(x, P) = f(x)$ independent of P or in the convex case:

THEOREM 2

Let $M(P_0)$ be closed and let M be B-u.sc. at P_0. Let $f(x, P) = f(x)$ be independent of P and lower semicontinuous on $M(P_0)$. Then ϕ is lower semicontinuous at P_0.

THEOREM 3

Let multifunction M be B-l.sc. and closed at P_0 with $M(P)$ convex for all $P \in \mathscr{P}$. Let $X(P_0) \neq \emptyset$ and bounded. Let f be lower semicontinuous on $\mathbb{R}^n \times \{P_0\}$ and such that f is upper semicontinuous at a point (x_0, P_0) with $x_0 \in X(P_0)$. Furthermore, let $f(\cdot, P)$ be quasiconvex on $\mathbb{R}^n$ for any fixed P. Then ϕ is continuous at P_0 and X is B-u.sc. at P_0.

Theorem 1 can be generalized to stability of local minimizers. Another generalization is based on the concept of epi-semicontinuity that can be applied to extended real functions. Problem (2) can then be formulated as a seemingly unconstrained minimization problem

$$\text{minimize } \tilde{f}(x, P) \text{ for } x \in \mathbb{R}^n \tag{8}$$

where

$$\begin{aligned} \tilde{f}(x, P) &= f(x, P) && \text{if } x \in M(P), \\ &= +\infty && \text{if } x \notin M(P). \end{aligned}$$

Such an extended real function $\tilde{f}: \mathbb{R}^n \times \mathscr{P} \to (-\infty, +\infty]$ is said to be *proper* if it is not identically $+\infty$.

The next theorem gives a very general result on stability and persistence of local minimizers obtained by Robinson [49].

THEOREM 4

Let $\tilde{f}: \mathbb{R}^n \times \mathscr{P} \to (-\infty, +\infty]$; let $G \subset \mathbb{R}^n$ be an open set and let $\tilde{f}(\cdot, P_0)$ be proper on G. Define for $P \in \mathscr{P}$:

$$\phi_G(P) = \inf_{x \in \mathrm{cl} G} \tilde{f}(x, P)$$

$$X_G(P) = \{x \in \mathrm{cl}\, G : \tilde{f}(x, P) = \phi_G(P)\}$$

and assume that $X_G(P_0) \subset G$, cl G is compact and $\tilde{f}$ is lower semicontinuous on cl $G \times \mathscr{P}$. Moreover, let at some $x_0 \in X_G(P_0)$, with $\mathscr{V}(x_0)$ the system of neighborhoods of x_0,

$$\tilde{f}(x_0, P_0) \geq \sup_{v \in \mathscr{V}(x_0)} \limsup_{P \to P_0} \inf_{x \in V} \tilde{f}(x, P) \tag{9}$$

(epi-upper semicontinuity) hold true.

Then $\phi_G(P_0)$ is finite and ϕ_G is continuous at P_0. X_G is closed at P_0 and B-u.sc. there. Further, there is a neighborhood U of P_0 such that for each $P \in U$, $\tilde{f}(\cdot, P)$ is proper on cl$G$, $X_G(P)$ is nonempty, compact and $X_G(P) \subset G$.

Application of theorem 4 to stochastic programming problems as done by Kall [33], Robinson and Wets [50] is discussed in what follows. The assumption of $M(P)$ fixed means an essential simplification. The reason is that, except for very special cases discussed, e.g., by Kall [33], there are no easily verifiable assumptions that guarantee epi-upper semicontinuity of f in (9) (or lower semicontinuity of $M(P)$ at P_0); see also Wang [69], Römisch and Schultz [56] for conditions that guarantee lower semicontinuity of special sets $M(P)$ defined by probability constraints (3).

Lower semicontinuity of objective function (8) for $M(P) = M$, a nonempty fixed convex closed set, can be ascertained by a number of conditions (see e.g. Wets [72]) that cannot be clarified here in detail. The following definition indicates that just the weak convergence cannot fully guarantee continuity:

DEFINITION 1

Let P_N, $N = 1, 2, \dots, P_0$ be probability measures on the same probability space $(\Omega, \mathscr{B})$. Then P_N is said to *converge weakly* to P_0 as $N \to \infty$ if for any *bounded* continuous function h: $\Omega \to \mathbb{R}^1$,

$$\int_\Omega h(\omega) P_N(\mathrm{d}\omega) \to \int_\Omega h(\omega) P_0(\mathrm{d}\omega).$$

To obtain the desired continuity or epi-continuity of the expectation functionals

$$f(x, P) = \int_\Omega h(x, \omega) P(\mathrm{d}\omega)$$

means, e.g., to restrict the class of considered functions $h(x, \cdot)$ or to restrict the set $\mathscr{P}$ of probability measures to a subset $\mathscr{P}_h$ with respect to which the functions $h(x, \cdot)$ are uniformly integrable. This was done in detail, e.g., in Robinson and Wets [50] (see also Kall [33]) for stochastic programs with complete recourse (5) for which objective function (8) can be written in the form

$$\tilde{f}(x, P) = \tilde{c}(x) + \int_\Omega q(x, \omega) P(\mathrm{d}\omega).$$

They assume that $\tilde{c}$ is a (nonrandom) lower semicontinuous extended real function epi-upper semicontinuous at $x_0 \in X_G(P_0)$ with $\tilde{c}(x_0)$ finite and that functions c_i, $i = 1, \dots, r+s$ in (5) are continuous and uniformly integrable for all $x \in \mathrm{cl}\ G$. The later assumption implies continuity and uniform integrability of penalty function q.

A different continuity result was obtained in Dupačová and Wets ([18], theorem 3.9) in connection with the statistical approach. It is based on epigraphical approach, too, and it has led to the concept of epi-consistency of lower semicontinuous real functions [43] defined as follows:

DEFINITION 2

A sequence $\{\tilde{f}_N\}$ of random lower semicontinuous functions is epi-consistent if there is a random (necessarily) lower semicontinuous function $\tilde{f}$ such that $\tilde{f}_N$ epi-converges to $\tilde{f}$ with probability 1.

Applied to problem (8), epi-consistency of the sequence $\{\tilde{f}(x, P_N)\}$ with limit $\tilde{f}(x, P_0)$ implies consistency of $\phi(P_N)$. Moreover, all cluster points of sequences of (local) minimizers of $\tilde{f}(x, P_N)$ are almost surely minimizers of $\tilde{f}(x, P_0)$.

2.2. QUANTITATIVE STABILITY RESULTS

We shall assume now that a stability and persistence result is available. It surely gives a feeling of certainty in approximating (or estimating) P. In spite of this, the question of magnitude of the error connected with an approximation or estimation scheme is very important. From now on, only *real* functions or functionals will be used.

If the true probability measure P is known one can get bounds of the error due to approximating P by another (simpler) measure utilizing *quantitative stability results*. The idea is simple: choose in $\mathscr{P}$ a suitable metric d that metrizes, at least locally, the weak convergence. If function ϕ and multifunction X enjoy a Lipschitz property, then

$$d(P, Q) < \epsilon \Rightarrow |\phi(P) - \phi(Q)| < K\epsilon,$$

respectively

$$d(P, Q) < \epsilon \Rightarrow \mathrm{dist}_{\mathrm{H}}(X(P), X(Q)) < K'\epsilon,$$

where the Lipschitz constants K, K' depend on the chosen metric d, and the Hausdorff distance $\mathrm{dist}_{\mathrm{H}}(X, Y) = \sup(\delta(X, Y), \delta(Y, X))$ with $\delta(X, Y) = \sup_{x \in X} \inf_{y \in Y} \|x - y\|$ is used.

One of the first results on Lipschitz continuity of the unique optimal solutions can be found in Robinson [47]. As we want to avoid uniqueness assumptions (such as X singleton on a neighborhood of P) as much as possible, we shall rely on quantitative stability results based on the notion of pseudo-Lipschitz property for multifunctions introduced by Aubin [3].

In connection with our problem (2), the following two theorems on quantitative stability were proved by Klatte [44]:

THEOREM 5

Assume that

(i) $X(P) \neq \emptyset$ bounded, multifunction M is closed-valued and closed at P;

(ii) M is pseudo-Lipschitzian at each pair $(x_0, P) \in X(P) \times \{P\}$;

(iii) f is Lipschitzian jointly with respect to $x \in X(P)$ and Q belonging to a neighborhood of P in the following sense: There are real numbers $\beta \in (0, 1]$ $L_f > 0$ and $\delta_f > 0$ such that

$$|f(x, P) - f(z, Q)| \leqq L_f\left(\|x - z\| + d(P, Q)^{\beta}\right).$$

Then X is u.sc. at P and ϕ is Lipschitzian, i.e., there are positive numbers δ_{ϕ}, L_{ϕ} such that $X(Q) \neq \emptyset$ and

$$|\phi(P) - \phi(Q)| \leqq L_{\phi} d(P, Q)^{\beta}$$

whenever $d(P, Q) < \delta_{\phi}$.

THEOREM 6

Let the assumptions of theorem 5 be supplemented by:

(iv) For program (2), there exists a strict local minimizer $x(P)$ of order $q \geqq 1$, i.e., there exist $r > 0$, $\Delta > 0$ such that

$$f(x, P) > f(x(P), P) + \Delta \|x - x(P)\|^{q}$$

for all $x \in [M(P) \cap B(x(P), r)] - \{x(P)\}$, where $B(x(P), r)$ denotes the closed ball around $x(P)$ with radius r.

Then there are positive numbers L, δ such that

$$\|x(P) - x\|^{q} \leqq L d(P, Q)^{\beta}$$

for all Q such that $d(P, Q) < \delta$ and for all $x \in X(Q)$.

Again, theorems 5 and 6 can be modified to cover the local minimization case. For their application to quantitative stability in stochastic programming, the choice of metric d on the space $\mathscr{P}$ is essential. It turned out (cf. [57]) that the bounded Lipschitz metric d_{BL} and the total variation distance d_{TV} are especially suitable for penalty problems and for problems with probability constraints, respectively. They are defined as follows:

$$d_{\mathrm{BL}}(P, Q) = \sup\left|\int_{\Omega} h(\omega) P(\mathrm{d}\omega) - \int_{\Omega} h(\omega) Q(\mathrm{d}\omega)\right|,$$

where the supremum is taken over all functions h satisfying the Lipschitz condition

$$|h(u) - h(v)| \leqq d(u, v)$$

for a metric d in $\mathbb{R}^{l}$ that is bounded by 1. The metric d_{BL} metrizes the weak topology, whereas the total variation distance

$$d_{\mathrm{TV}}(P, Q) = \sup_{B \in \mathscr{B}^*} |P(B) - Q(B)|$$

where $\mathscr{B}^*$ is a subset of the σ-field of Borel sets in $\mathbb{R}^{l}$, does not generate the weak topology. It reduces for

$$\mathscr{B}^* = \{\emptyset, (-\infty, z], z \in \mathbb{R}^{l}\}$$

to the Kolmogorov distance

$$d_K(P, Q) = \sup_{z \in \mathbf{R}^l} |F_P(z) - F_Q(z)|,$$

where F_P and F_Q denote distribution functions corresponding to the probability measures P and Q, respectively. (For details consult, e.g., Huber [27].)

For *stochastic programs with recourse* (4), $M(P) = M$ is a fixed closed set and the objective function $f(x, P)$ is linear in P. In Römisch and Schultz [54], the bounded Lipschitz metric $d = d_{BL}$ was applied to the class $\mathscr{P}(\Omega, \nu, K) \subset \mathscr{P}$ defined through moments conditions

$$\mathscr{P}(\Omega, \nu, K) := \left\{ P \in \mathscr{P} : \int_\Omega \| \omega \|^{2\nu} P(\mathrm{d}\omega) \leqq K \right\}.$$

The Lipschitz property (iii) of theorem 5 was verified for $P \in \mathscr{P}(\Omega, \nu, K)$, for stochastic linear programs with $c_0(x, \omega) = c^T x$ in (4), with linear complete recourse (see (5) for linear constraints) and for quadratic recourse (6). Accordingly, it was shown that the objective value function is (locally) Hölder continuous at $P \in \mathscr{P}(\Omega, \nu, K)$ with exponent $\beta = 1 - (1/\nu)$ if $X(P)$ is nonempty and bounded. Moreover, for problems with linear complete recourse, Hölder continuity of optimal solutions is fulfilled at P if P has continuous and positive density.

The case of *probability constraints* (3) was studied in Römisch and Schultz [54,56]. They considered the case of f independent of P and $M(P)$ given by

$$M(P) := \{ x \in M_0 : P(H(x)) \geqq \alpha \},$$

with H a closed set-valued mapping, e.g.,

$$H(x) = \{ \omega : c_i(x, \omega) \leqq 0,\ i = 1, \dots, r \}$$

for c_i continuous, $i = 1, \dots, r$. Their general stability results assert that if f is Lipschitzian on compact sets and the multifunction

$$p \mapsto \{ x \in M_0 : P(H(x)) \geqq p \} \tag{10}$$

is pseudo-Lipschitzian at each $(x_0, \alpha) \in X(P) \times \{\alpha\}$, the mapping X is upper semicontinuous at P and ϕ is (locally) Lipschitzian at P (with respect to the metric d_{TV} for $\mathscr{B}^*$ chosen so that it contains all sets $H(x)$, $x \in M_0$).

The pseudo-Lipschitz property of multifunction (10) was verified for the convex case, i.e., for M_0 convex, H of convex graph, the Slater condition fulfilled for the probability constraint and for P belonging to a convexity class, such as the class of log-concave probability measures (cf. Prékopa [45]). For the (generally nonconvex) case of $H(\omega) := \{ x : Ax \geqq \omega \}$, sufficient conditions for the pseudo-Lipschitz property were obtained, too. One condition is, for instance, that P is absolutely continuous and its density is bounded below by a positive number on a neighborhood related to the set $X(P)$.

The results on qualitative stability provide a deeper insight into the structural properties of the solved stochastic program (2) in their dependence on small perturbations of the probability measure P. To apply them for computing error bounds one should be able to compute at least an upper bound to the considered distance $d(P, Q)$ and to the Lipschitz or Hölder constants. There exist some bounds on $d(P, Q)$, e.g., for $Q = P_N$ – the empirical measure; in this case quantitative stability results provide rates of convergence of the optimal value $\phi(P_N)$, cf. Römisch and Schultz [54]. However, computing bounds on the Lipschitz or Hölder constants seems to be intractable.

2.3. DIFFERENTIABILITY AND STATISTICAL SENSITIVITY ANALYSIS

If the true probability measure P is not known completely, *statistical sensitivity analysis* enables one to draw statistical conclusions about the error due to estimation of P.

Assume first that M is a nonempty fixed closed subset of $\mathbb{R}^n$ and that the objective function has the form

$$f(x, P) = \int_\Omega h(x, \omega) P(\mathrm{d}\omega).$$

Its optimal value $\phi(P) = \inf_{x \in M}(x, P)$ is assumed to be continuous on a convex neighborhood U of P_0, with $X(P_0) \neq \emptyset$ and $\phi(P_0)$ finite.

Due to the fact that $f(x, P)$ is linear in P (on the restriction of the linear space of finite sign measures to $\mathscr{P}$) we have for all $P, Q \in U$ and $0 \leqq t \leqq 1$:

$$\phi((1-t)P + tQ) = \inf_{x \in M} [(1-t) f(x, P) + t f(x, Q)] \geqq (1-t)\phi(P) + t\phi(Q),$$

so that ϕ *is concave on* U. The Gâteaux derivative of ϕ at $P \in U$ in the direction of $Q - P$, $Q \in U$ is defined as

$$\phi'(P; Q-P) = \lim_{t \to 0^+} \frac{\phi(P_t) - \phi(P)}{t}, \tag{11}$$

where $P_t = (1-t)P + tQ$. For a *fixed* $Q \in U$, $\phi'(P; Q-P)$ is nothing but the right-hand derivative of the concave function

$$\phi_Q(t) := \phi(P_t), \quad \phi_Q\colon [0, 1] \to \mathbb{R}^1$$

at $t = 0$. This means, inter alia, that the limit (11) exists and

$$\phi'(P; Q-P) \geqq \phi(Q) - \phi(P).$$

To obtain an explicit formula for $\phi'(P; Q-P)$ we can apply directly, e.g., the theorem of Danskin ([8], chap. 2, theorem 1).

To this purpose notice that the derivative $\partial f(x, P_t)/\partial t$ exists for all $x \in M$ and equals the difference quotient:

$$\frac{\partial}{\partial t} f(x, P_t) = \frac{f(x, P_t) - f(x, P)}{t} = f(x, Q) - f(x, P). \tag{12}$$

THEOREM 7

Let M be compact, let $f(x, P)$, $f(x, Q)$ be finite continuous functions of x. Then

$$\phi'(P; Q-P) = \min_{x\in X(P)} [f(x, Q) - f(x, P)] = \min_{x\in X(P)} f(x, Q) - \phi(P). \tag{13}$$

In the convex case (see assumptions of theorem 3 with $M(P) = M$ and with $f(\cdot, P)$ continuous for all $P \in U$), the same result and formula (13) follow, e.g., from theorem 16 of Gol'shtein [24]. Using a similar approach, it is even possible to prove it for local minimizers.

THEOREM 8

Let the assumptions of theorem 4 be fulfilled for $P_0 = P$ and let $f(x, P)$ be continuous on G. Let Q be an arbitrary element of a neighborhood U of P for which $X_G(Q) \neq \emptyset$, $X_G(Q) \subset G$ and $f(\cdot, Q)$ is finite on G. Then the Gâteaux derivative of ϕ_G at P in the direction of $Q - P$ exists and is given by

$$\phi_G'(P; Q-P) = \min_{x\in X_G(P)} f(x, Q) - \phi_G(P). \tag{14}$$

Proof

Put $P_t = (1-t)P + tQ$. Due to the upper semicontinuity of X_G at P, $X_G(P_t) \subset G$ for t small enough, say for $t \in (0, t_0)$, $t_0 > 0$. According to (12) we can write for any $x_0 \in X_G(P)$, $x_t \in X_G(P_t)$

$$\begin{aligned}\phi_G(P_t) &= f(x_t, P_t) = f(x_t, P) + t\frac{\partial}{\partial t} f(x_t, P) \leqq f(x_0, P_t)\\ &= f(x_0, P) + t\frac{\partial}{\partial t} f(x_0, P_t) = \phi_G(P) + t[f(x_0, Q) - f(x_0, P)],\end{aligned} \tag{15}$$

so that

$$\frac{\phi_G(P_t) - \phi_G(P)}{t} \leqq f(x, Q) - f(x, P) \quad \text{for all } x \in X_G(P). \tag{16}$$

Using the trivial inequality $f(x_t, P) \geqq \phi_G(P)$ together with the first part of (15), we get

$$\phi_G(P_t) \geqq \phi_G(P) + t[f(x_t, Q) - f(x_t, P)]$$

and, consequently,

$$\frac{\phi_G(P_t) - \phi_G(P)}{t} \geqq f(x, Q) - f(x, P) \quad \text{for all } x \in X_G(P_t). \tag{17}$$

Inequalities (16) and (17) imply that

$$\begin{aligned}\min_{x\in X_G(P_t)} [f(x, Q) - f(x, P)] &\leqq \frac{\phi_G(P_t) - \phi_G(P)}{t}\\ &\leqq \min_{x\in X_G(P)} [f(x, Q) - f(x, P)].\end{aligned} \tag{18}$$

Denote $m(t) = \min_{x \in X_G(P_t)}[f(x, Q) - f(x, P)]$; in this parametric optimization problem, the assumptions of theorem 2 are fulfilled, so that $m(t)$ is lower semicontinuous on $0 \leqq t \leqq t_0$. Accordingly, (18) implies

$$\limsup_{t \to 0} m(t) \leq m(0) = \liminf_{t \to 0} m(t),$$

so that the limit

$$\lim_{t \to 0} \frac{\phi_G(P_t) - \phi_G(P)}{t} = \phi_G'(P; Q - P)$$

exists and equals (14). □

Computing the Gâteaux derivative at P in the direction of $Q - P$ means solving two stochastic programs: to obtain the whole set $X(P)$ of optimal solutions of stochastic program (2) for the given probability measure P and to evaluate the minimum of the objective function

$$f(x, Q) = \int_\Omega h(x, \omega) Q(\mathrm{d}\omega)$$

on the set of optimal solutions $X(P)$.

The function ϕ is said to be Gâteaux differentiable at P if there is a *linear* functional l_P such that

$$\phi'(P; Q - P) = l_P(Q - P) \quad \forall Q \in U. \tag{19}$$

If this functional is continuous it can be represented as

$$\phi'(P; Q - P) = \int_\Omega \psi_P(\omega) Q(\mathrm{d}\omega),$$

with ψ_P standardized so that

$$\int_\Omega \psi_P(\omega) P(\mathrm{d}\omega) = 0.$$

If $X(P) = \{x(P)\}$ *is a singleton*, we can evidently put

$$\psi_P(u) = h(x(P), u) - \phi(P). \tag{20}$$

Moreover, for a special choice of $Q = \delta_u$ – the degenerated probability measure concentrated at the point u – we get

$$\phi'(P; \delta_u - P) = \psi_P(u),$$

which is the *influence function* suggested by Hampel [25]. It measures the influence of an observation u toward the approximate estimation error $\phi'(P; P_N - P)$ when the empirical probability measure P_N is used instead of P. Notice that due to the linearity of $\phi'(P; \cdot)$ we have

$$\phi(P_N) - \phi(P) \cong \phi'(P; P_N - P) = \frac{1}{N} \sum_{k=1}^{N} \psi_P(u_k).$$

If ψ_P is unbounded, an outlier may cause great discrepancies. The maximum absolute value of the influence function

$$\gamma^* = \sup_u |\psi_P(u)| \tag{21}$$

is called the *gross error sensitivity*. According to Hampel [25] it measures the worst approximate influence which a fixed amount of contamination can have on the value of the estimate. Another characteristics that may be relevant in the context of stochastic programming concerns the worst approximate effect of wiggling or rounding of observations: the *local-shift sensitivity*

$$\lambda^* = \sup_{u \neq v} \frac{|\psi_P(u) - \psi_P(v)|}{|u - v|}. \tag{22}$$

The influence function provides a simple heuristic insight into the asymptotic properties of the estimate, i.e., of $\phi(P_N)$ in our context.

If the function ϕ was Fréchet differentiable at P, i.e.,

$$\frac{|\phi(Q) - \phi(P) - \phi'(P;\, Q - P)|}{d(P,\, Q)} \to 0 \quad \text{as} \quad d(P,\, Q) \to 0, \tag{23}$$

with $\phi'(P;\, Q - P)$ linear and continuous in the increments, and if the empirical probability measure P_N converged to the true one at the rate $N^{-1/2}$, i.e.,

$$\sqrt{N}\, d(P_N,\, P) \text{ bounded in probability as } N \to \infty, \tag{24}$$

one could get a simple proof of asymptotic normality for $\phi(P_N)$:

$$\sqrt{N}\,(\phi(P_N) - \phi(P)) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \psi_P(u_k) + \sqrt{N}\, o(d(P_N,\, P)),$$

so that the limit distribution of $\sqrt{N}\,(\phi(P_N) - \phi(P))$ would be given by that of the sum

$$\frac{1}{\sqrt{N}} \sum_{k=1}^{N} \psi_P(u_k)$$

of independent identically distributed variables $\psi_P(u_k)$. Provided that the variance

$$\sigma_P^2 = \int_\Omega \psi_P^2(\omega) P(\mathrm{d}\omega) < +\infty, \tag{25}$$

$\phi(P_N)$ would be asymptotically normal

$$\sqrt{N}\,(\phi(P_N) - \phi(P)) \sim \mathcal{N}(0,\, \sigma_P^2). \tag{26}$$

Unfortunately, (24) does not hold true in general (see Huber [27]). As to Fréchet differentiability of ϕ, if the Fréchet differential $\phi'(\mathrm{P};\, \mathrm{Q} - \mathrm{P})$ in (23) exists, it equals the Gâteaux differential (19) that can be obtained by routine calculations that do not involve a metric. Computing Gâteaux derivatives is thus a proper starting point of a statistical analysis. As to the linearity of Gâteaux differentials $\phi'(P; Q-P)$ with respect to Q we have to turn our attention once more to the properties of the stochastic program (13) or (14) whose objective function $f(x, Q)$ is linear with respect to parameter values $Q \in U$ and whose set of feasible solutions $X(P)$ does not depend on the parameter Q. Hence, $\phi'(P; Q-P)$ given by (13) or (14) is concave on U in general. Its linearity is guaranteed in case $X(P)$ is a singleton.

Summary

If the set of feasible solutions is fixed, $M(P)=M \neq \emptyset$ and fulfils together with objective functions

$$f(x, P)=\int_\Omega h(x, \omega)P(\mathrm{d}\omega)$$

the assumptions of theorems 7 or 8 (or of a similar theorem that also guarantees persistence and stability) and if the set of optimal solutions $X(P)=\{x(P)\}$ is a singleton, then the optimal value function ϕ is Gâteaux differentiable at P and its Gâteaux differential can be represented as

$$\phi'(P; Q-P)=\int_\Omega \psi_P(\omega)Q(\mathrm{d}\omega),$$

with

$$\psi_P(u)=h(x(P), u)-\phi(P).$$

Provided that the asymptotic variance

$$\sigma_P^2=\int_\Omega \psi_P^2(\omega)(\mathrm{d}\omega)<+\infty,$$

one can try to prove the asymptotic normality of $\phi(P_N)$, i.e., the statement that

$$\sqrt{N}\left(\phi(P_N)-\phi(P)\right)>\mathcal{N}\left(0, \sigma_P^2\right).$$

To this purpose, the statistical approach mentioned in the introduction is needed to get rigorous statistical results in terms of convergence in distribution. In the general case, i.e., for $X(P)$ containing more than one point, one cannot expect asymptotically normal behavior of the estimate $\phi(P_N)$.

Example

A stochastic programming formulation of the newsboy problem leads to the program

$$\text{maximize}\left[(s-p)x-sE_P(x-\omega)^+\right] \quad \text{for } 0 \leqq x \leqq b.$$

Here, s is the sale price, p is the purchasing price, ω is the random demand and x is the amount of newspapers to be ordered. Accordingly, we have $0<p<s$ and we put $\alpha=p/s$.

Assume that the probability measure P is carried by a compact support, say $[\underline{d}, \bar{d}]$ with $0 \leqq \underline{d}<\bar{d} \leqq b$ and that it is absolutely continuous. This implies that the objective function to be *maximized* can be written as

$$
\begin{aligned}
f(x, P) &= (1-\alpha)x && \text{for } x < \underline{d},\\
&= (1-\alpha)x - \int_{\underline{d}}^{x} (x-\omega)P(\mathrm{d}\omega) && \text{for } \underline{d} \leqq x \leqq \bar{d},\\
&= -\alpha x + E_P\omega && \text{for } x > \bar{d}.
\end{aligned}
$$

It attains its maximum on the interval $[\underline{d}, \bar{d}]$ at the point

$$x(P) = u_{1-\alpha}(P),$$

the 100(1 − α)% quantile of P. The optimal value

$$\phi(P) = \max_{x\in[0,\, b]} f(x, P) = x(P)(1-\alpha) - \int_{\underline{d}}^{x(P)} F(y)\,\mathrm{d}y,$$

where $F(y) = P\{\omega \leqq y\}$ is the distribution function corresponding to P.

Let Q be another probability measure carried by a subset of $[0, +\infty)$. Then

$$
\begin{aligned}
\phi'(P; Q-P) &= f(x(P), Q) - \phi(P)\\
&= (1-\alpha)x(P) - \phi(P) - \int_{\Omega} (x(P)-\omega)^{+} Q(\mathrm{d}\omega).
\end{aligned}
$$

Denote

$$A_P = (1-\alpha)x(P) - \phi(P) \left(= \int_{\underline{d}}^{x(P)} F(y)\,\mathrm{d}y \geqq 0 \right).$$

The influence function (20) (for $u \geqq 0$!)

$$
\begin{aligned}
\psi_P(u) &= A_P - (x(P)-u)^{+}\\
&= A_P && \text{for } u \geqq x(P),\\
&= u + A_P - x(P) && \text{for } u < x(P).
\end{aligned}
$$

Accordingly, the gross error sensitivity (21):

$$
\begin{aligned}
\gamma^* &= \max\left[A_P, \max_{0\leqq u\leqq x(P)} |u + A_P - x(P)| \right]\\
&= \max[A_P, |A_P - x(P)|] = \max[A_P, \phi(P) + \alpha x(P)] < +\infty.
\end{aligned}
$$

The local shift sensitivity (22) approximately equals the slope of the influence function

$$\lambda^* = \sup_{u\neq v} \frac{|\psi_P(v) - \psi_P(u)|}{|u-v|} = 1,$$

and the asymptotic variance

$$
\begin{aligned}
\sigma_P^2 &= E_P\left[(x(P)-\omega)^{+} \right]^2 - A_P^2\\
&= \int_{\underline{d}}^{x(P)} (x(P)-\omega)^2 P(\mathrm{d}\omega) - A_P^2 < +\infty.
\end{aligned}
$$

One can thus expect that

$$\phi(P_N) = \max_{0 \leqq x \leqq b} \left[(1-\alpha)x - \frac{1}{N} \sum_{k=1}^{N} (x - u_k)^+ \right]$$

is asymptotically normal

$$\sqrt{N}(\phi(P_N) - \phi(P)) \sim \mathcal{N}(0, \sigma_P^2). \tag{27}$$

Moreover, if (27) holds true, then σ_P^2 can be replaced by its empirical estimate $\sigma_{P_N}^2$.

Similar results can be obtained for P an absolutely continuous distribution over $[0, \infty)$ that possesses the mean value $E_P\omega$.

It is also possible to prove the existence of Gâteaux derivatives of the optimal value function ϕ in the case that the set of feasible solutions depends on P and is given by explicit constraints. However, to interpret the results, there is no longer at our disposal a direct parallelism with the statistical inference based on the notion of the influence function.

Assume that

$$M(P) = \{x \in M_0 : g_i(x, P) \leqq 0, \quad i = 1, \ldots, m\}, \tag{28}$$

with g_i linear with respect to the parameter P and with $M_0 \subset \mathbb{R}^n$ nonempty closed convex. Evidently, the models with probabilistic constraints can be written in this form, too. Notice that the perturbed program

$$\text{minimize } f(x, P_t) \tag{29}$$

on the set

$$M(P_t) = \{x \in M_0 : g_i(x, P_t) \leqq 0, \ i = 1, \ldots, m\} \tag{30}$$

that corresponds to the contaminated probability measure

$$P_t = (1-t)P + tQ$$

depends (for P, Q fixed) *linearly* on the scalar real parameter $t \in [0, 1]$. This means that results of parametric programming for finite dimensional parameters can be used, cf. Gol'shtein [24], Rockafellar [51], Fiacco and Kyparisis [22].

Denote $L(x, v, P)$ the Lagrange function corresponding to the problem

$$\text{minimize } f(x, P) \text{ on the set } M(P) \text{ given by (28)}, \tag{31}$$

i.e.,

$$L(x, v, P) = f(x, P) + \sum_{i=1}^{m} v_i g_i(x, P). \tag{32}$$

For P fixed, L is defined on $M_0 \times \mathbb{R}_+^m$. Consider further the dual problem

$$\text{maximize } \inf_{x \in M_0} L(x, v, P) \text{ on the set } \mathbb{R}_+^m \tag{33}$$

and denote by $X(P)$ and $V(P)$ the sets of optimal solutions of (31) and (33), respectively. The following theorem is a simple adaptation of Gol'shtein's result to our problem.

THEOREM 9 ([24], theorem 17)

Let $M_0 \neq \emptyset$ be convex, closed and $P, Q \in \mathscr{P}$. Let $f(\cdot, P)$, $g_i(\cdot, P)$, $1 \leqq i \leqq m$ be convex continuous on M_0. Let for programs (31), (33) the sets $X(P)$ and $V(P)$ be nonempty and bounded. Assume further that the functions $f(\cdot, Q)$, $g_i(\cdot, Q)$, $1 \leqq i \leqq m$ are convex and finite for all x belonging to a neighborhood of $X(P)$. Then the Gâteaux derivative of $\phi(P)$ in the direction of $Q - P$ exists and

$$\phi'(P; Q-P) = \min_{x \in X(P)} \max_{v \in V(P)} \left(L(x, v, Q) - L(x, v, P)\right)$$
$$= \max_{v \in V(P)} \min_{x \in X(P)} \left(L(x, v, Q) - L(x, v, P)\right).$$

Again, for $X(P)$, $V(P)$ consisting of one point only, one gets linearity of $\phi'(P; Q-P)$ on a whole convex neighborhood U of P that contains probability measures Q for which the assumptions of theorem 9 are fulfilled. Furthermore, in accordance with our previous results, uniqueness of Lagrange multipliers $v \in V(P)$ is not relevant if the set $M(P)$ is fixed, independent of P.

In the nonconvex case with $M_0 = \mathbb{R}^n$, one can impose different assumptions that imply uniqueness of the Kuhn–Tucker point $[x(P), v(P)]$ that corresponds to the local minimizer $x(P)$ of (31) such as the linear independence condition (see, e.g., Fiacco [21]). If they are fulfilled, the optimal value function

$$\phi(P_t) = \phi((1-t)P + tQ) = \phi_Q(t)$$

is differentiable at $t = 0$ with

$$\phi_Q'(0) = \phi'(P; Q-P) = L(x(P), v(P), Q) - L(x(P), v(P), P).$$

Of course, the formulation of the second order sufficient conditions assumes that the objective function $f(\cdot, P)$ and the constraints $g_i(\cdot, P)$, $i = 1, \ldots, m$, are twice continuously differentiable at $x(P)$. We are not going to discuss these smoothness assumptions; for relevant results in this direction see Wang [68].

If, in addition, strict complementarity conditions are fulfilled and the functions $g_i(x, Q)$, $i = 1, \ldots, m$, $f(x, Q)$ are differentiable at $x(P)$, then ϕ_Q is twice continuously differentiable on a right neighborhood of 0 with

$$\phi_Q''(0) = -\left(\nabla_z L(z(P), Q) - \nabla_z L(z(P), P)\right)^{\mathrm{T}} C^{-1} \left(\nabla_z L(z(P), Q) - \nabla_z L(z(P), P)\right), \tag{34}$$

where $z(P) = [x(P), v(P)]$ and $C = \nabla_{zz} L(z(P), P)$. Without strict complementarity conditions, this result is no longer valid. Nevertheless, it is still possible to provide a second-order approximation of $\phi(P_t)$, cf. Fiacco and Kyparisis [22], Shapiro [60,62].

Let us turn our attention now to *Gâteaux differentiability of optimal solutions*. Quite naturally, this concept can be directly applied if the optimal solution of program (31) is unique and if multifunction X is in fact a function on a neighborhood of P. The first results in this direction – see Dupačová [13] – were obtained under corresponding differentiability assumptions, for $M(P)$ fixed, independent of P, under the linear independence condition and the strong second order sufficient condition fulfilled for (31). These conditions guarantee the existence of unique optimal solution $x(P_t)$ and of the unique vector of Lagrange multipliers $v(P_t)$ for program (29), (30) for all t belonging to a right neighborhood of 0. The result can be easily extended to the set of feasible solutions given by (28). In this case, the Gâteaux derivative

$$x'(P;\, Q-P) = \lim_{t\to 0^+} \frac{x(P_t) - x(P)}{t}$$

exists and equals the (necessarily unique) optimal solution of the following convex quadratic program that stems from Jittorntrum [29]:

$$\text{minimize} \qquad \tfrac{1}{2}x^{\mathrm{T}}\nabla_{xx}L(x(P),\, v(P),\, P)x + x^{\mathrm{T}}\nabla_x L(x(P),\, v(P),\, Q) \tag{35}$$

$$\text{subject to} \qquad x^{\mathrm{T}}\nabla_x g_i(x(P),\, P) + g_i(x(P),\, Q) = 0, \quad i \in I^+(P), \tag{36}$$

$$x^{\mathrm{T}}\nabla_x g_i(x(P),\, P) + g_i(x(P),\, Q) \leqq 0, \quad i \in I^0(P), \tag{37}$$

where $i \in I^+(P)$ are indices of those constraints of (28) that are active at $x(P)$ with a positive Lagrange multiplier $v_i(P)$, whereas $I \in I^0(P)$ correspond to the remaining constraints of (28) that are active at $x(P)$. Moreover, Gâteaux derivatives $v_i'(P;\, Q-P)$ of nonzero Lagrange multipliers for (31) equal Lagrange multipliers corresponding to the optimal solution of the quadratic program (35)–(37).

The Gâteaux derivatives are linear with respect to Q only if strict complementarity conditions hold true, i.e., $I^0(P) = \emptyset$. For explicit formulas see Dupačová [13,14].

Relaxation of strict complementarity still gives $x'(P;\, Q-P)$ continuous but no longer linear in Q. To interpret this result, consider *M(P) fixed*, independent of P, so that

$$g_i(X(P),\, Q) = 0 \quad \text{for } i \in I^+(P) \cup I^0(P)$$

in constraints (36), (37). For $Q = P_N$, the empirical probability measure concentrated in N points, say, ω_k, $k = 1, \ldots, N$, we get that the Gâteaux derivative $x'(P;\, P_N - P)$ equals the optimal solution of the convex quadratic program whose set of feasible solutions is a convex polyhedral cone (see (36), (37)):

$$K = \{x : x^{\mathrm{T}}\nabla_x g_i(x(P)) = 0,\ i \in I^+(P),\ x^{\mathrm{T}}\nabla_x g_i(x(P)) \leqq 0,\ i \in I^0(P)\}$$

and whose vector of coefficients in the linear term of the objective function (see (35)) equals

$$\frac{1}{N}\sum_{k=1}^{N}\left[\nabla_x h(x(P),\, \omega_k) + \sum_{i\in I^+(P)\cup I^0(P)} v_i(P)\nabla_x g_i(x(P))\right]$$

and, being an average of independent identically distributed random vectors, it is asymptotically normal provided that the variance matrix of $\nabla_x h(x(P), \omega)$ is of a bounded norm. It gives again some heuristic for the asymptotic distribution of the optimal solutions $x(P_N)$ of (31) with P replaced by its estimate P_N: one can expect that $\sqrt{N}(x(P_N) - x(P))$ is asymptotically equivalent to the solution of the corresponding quadratic program with asymptotically normal coefficients in the linear term of the objective function.

A similar result has been obtained by Shapiro [63] who also studies the existence and properties of Gâteaux derivatives $x'(P;\, Q - P)$ under Mangasarian–Fromowitz constraint qualification. He shows that Gâteaux derivatives can be discontinuous if Lagrange multipliers are not unique.

Example (continued)

In our example, we have got that the optimal solution $x(P) \neq 0$, $x(P) = u_{1-\alpha}(P)$ and moreover, $x(P)$ is an unconstrained maximizer of $f(x, P)$. This means that the strict complementarity condition is fulfilled. Asymptotic normality of the sample quantiles

$$x(P_N) = u_{1-\alpha}(P_N)$$

for $0 < \alpha < 1$ is a well known result; see e.g. Serfling [59] for the statistical background and Wets [70] for an application to this example.

3. Remarks on parametric case

Assume now that the true probability measure P in stochastic program (2) is known to belong to a parametric family $\mathscr{P} = \{P_y,\ y \in Y\}$ of probability measures on $(\Omega, \mathscr{B})$ that is indexed by a parameter vector y belonging to an open set $Y \subset \mathbb{R}^q$. This means that stability and sensitivity for stochastic program (2) can be treated via existing techniques of parametric programming with parameters belonging to a subset of Euclidean space. For one of the first papers in this direction see Armacost and Fiacco [1].

In the context of stochastic programming, the true parameter values are often not known but estimated by means of sample data. The estimates usually enjoy quite convenient statistical properties. A natural question is how far are these properties inherited by the optimal value function and by the set of optimal solutions obtained by solving the substitute stochastic program that corresponds to the estimated parameter values.

In the parametric case we can assume that the stochastic program (2) has the form

$$\text{minimize } f(x, y) \text{ on a set } M(y) \tag{38}$$

with the optimal value function

$$\phi(y) = \inf\{ f(x, y) \colon x \in M(y)\}$$

and with the set of optimal solutions

$$X(y) = \{x \in M(y) \colon f(x, y) = \phi(y)\}.$$

We shall denote by η the true parameter vector.

We shall assume that the conditions under which the continuity (local stability) results hold true are already known from parametric programming.

The weakest continuity property is needed in the following result due to Vogel [67].

THEOREM 10

Let the multifuction X be H-u.sc. at η. Assume further that there exists an estimate y_N of η such that for all $\delta > 0$ and for a $k > 0$,

$$P\{ \| y_N - \eta \| \geqq \delta \} = o(N^{-k}) \quad \text{for } N \to \infty.$$

Then for all $\epsilon > 0$,

$$P\{\exists \hat{x} \in X(y_N) \text{ with } d(\hat{x}, X(\eta)) \geqq \epsilon\} = o(N^{-k}).$$

If the optimal value function is continuous at η we can use properties of transformed random sequences (see, e.g., Serfling [59]) to obtain statistical properties of the estimate $\phi(y_N)$ of the true optimal value $\phi(\eta)$. The same idea also applies to the estimated optimal solutions provided that the optimal solution $x(y)$ of (38) is almost surely unique and continuous on a neighborhood of η.

THEOREM 11 ([59], theorem 1.7)

Let $\phi: Y \to \mathbb{R}^1$, $x: Y \to \mathbb{R}^n$ be Borel functions that are continuous on a neighborhood U of η with P probability 1. Then we have

(a) $P\left(\lim_{N\to\infty} y_N = \eta\right) = 1 \Rightarrow P\left(\lim_{N\to\infty} \phi(y_N) = \phi(\eta)\right) = 1$ and

$$P\left(\lim_{N\to\infty} x(y_N) = x(\eta)\right) = 1$$

(convergence with probability 1).

(b) $\lim_{N\to\infty} P\{ \| y_N - \eta \| < \epsilon\} = 1, \forall \epsilon > 0 \Rightarrow$

$$\lim_{N\to\infty} P\{ | \phi(y_N) - \phi(\eta) | < \epsilon\} = 1, \forall \epsilon > 0$$

and $\lim_{N\to\infty} P\{ \| x(y_N) - x(\eta) \| < \epsilon\} = 1, \forall \epsilon > 0$

(convergence in probability).

Consistency of the optimal value $\phi(y_N)$ and of the (unique) optimal solution $x(y_N)$ of the substitute program

$$\text{minimize } f(x,\ y_N) \text{ on the set } M(y_N) \tag{39}$$

thus follows from the consistency of the estimate y_N and from the continuity of the functions ϕ and x. For

$$M(y) = \{x \in \mathbb{R}^n:\ g_i(x,\ y) \leq 0,\ i = 1, \ldots, m\} \tag{40}$$

it is sufficient to this purpose if the linear independence condition and the strong second order sufficient condition of Robinson [47] are fulfilled for the true program (38) with $y = \eta$.

The assertions of theorem 11 can be completed by a rate of convergence if the functions ϕ and x are continuously differentiable at $y = \eta$. We shall discuss here only the case of asymptotically normal estimates y in which the delta-theorem (cf. Serfling [59], theorem 3.3.A and its corollary) can be applied as was done in Dupačová ([12], [14], section 2).

THEOREM 12

Let y_N be an asymptotically normal estimate of η. i.e., $\sqrt{N}(y_N - \eta)$ converges in distribution to a random vector with distribution $\mathcal{N}(0,\ \Sigma)$, briefly $\sqrt{N}(y_N - \eta) \sim \mathcal{N}(0,\ \Sigma)$.

(a) Let ϕ be continuously differentiable at η with $\nabla\phi(\eta) \neq 0$. Then $\phi(y_N)$ is asymptotically normal:

$$\sqrt{N}\left(\phi(y_N) - \phi(\eta)\right) \sim \mathcal{N}\left(0,\ (\nabla\phi(\eta))^{\mathrm{T}}\Sigma\nabla\phi(\eta)\right). \tag{41}$$

(b) Let $X(y) = \{x(y)\}$ be singleton for y belonging to a neighborhood of η and let x be continuously differentiable at η with $(\nabla x(\eta))^{\mathrm{T}}\Sigma(\nabla x(\eta)) \neq 0$. Then $x(y_N)$ is asymptotically normal:

$$\sqrt{N}\left(x(y_N) - x(\eta)\right) \sim \mathcal{N}\left(0,\ (\nabla x(\eta))^{\mathrm{T}}\Sigma\nabla x(\eta)\right). \tag{42}$$

As we know from parametric programming, substantially weaker conditions are needed for differentiability of the optimal value function than for differentiability of optimal solutions. Even under the linear independence condition and the strong second order sufficient condition (but without strict complementarity conditions fulfilled for (38) at the true parameter vector η), only directional differentiability of x can be proved [29,48]. Directional derivatives $x'(\eta;\ z)$ and $v'(\eta;\ z)$ can again be obtained as the optimal solution and the corresponding multipliers of the convex quadratic program (compare (35)–(37), [29,60])

$$\text{minimize } \tfrac{1}{2}z^{\mathrm{T}}\nabla_{yy}L(x(\eta),\ v(\eta),\ \eta)z + x^{\mathrm{T}}\nabla_{xy}L(x(\eta),\ v(\eta),\ \eta)z + \tfrac{1}{2}x^{\mathrm{T}}\nabla_{xx}L(x(\eta),\ v(\eta),\ \eta)x \tag{43}$$

subject to

$$z^{\mathrm{T}}\nabla_y g_i(x(\eta), \eta) + x^{\mathrm{T}}\nabla_x g_i(x(\eta), \eta) = 0, \quad i \in I^+(\eta), \tag{44}$$

$$z^{\mathrm{T}}\nabla_y g_i(x(\eta), \eta) + x^{\mathrm{T}}\nabla_x g_i(x(\eta), \eta) \leqq 0, \quad i \in I^0(\eta), \tag{45}$$

where

$$L(x, v, y) = f(x, y) + \sum_{i=1}^{m} v_i g_i(x, y)$$

is the Lagrange function of program (38), (40); $I^+(\eta)$ is the set of indices of the constraints $g_i(x, \eta) \leqq 0$ that are active at $x(\eta)$ with positive Lagrange multipliers $v_i(\eta)$; and $I^0(\eta)$ is the set of indices of the remaining constraints that are active at $x(\eta)$. Detailed stability analysis of this program with respect to its parameter $z = y_N - \eta$ helps to describe the generally nonnormal asymptotic distribution of the optimal solutions $x(y_N)$; see Dupačová [16] for a fixed polyhedral set M of feasible solutions. The quadratic program (43)–(45) depends on the unknown true parameter vector η and on the true optimal solution $x(\eta)$. Nevertheless, it is possible to prove [17] that for N large enough, the coefficients can be replaced by their sample counterparts (i.e., η replaced by y_N, etc.) without influencing the asymptotic results.

Higher order differentiability results help to obtain further statistical information such as Berry–Esséen rate of convergence in (41) or (42) – for an example see Dupačová [12], or a higher order asymptotic expansion for the density of $\sqrt{N}(x(y_N) - x(\eta))$ by the method suggested by Vogel [67]. However, besides the corresponding smoothness properties of the functions f, g_i, $i = 1, \ldots, m$, results of this type are conditioned by the strict complementarity conditions.

The linear independence condition can again be replaced by the Mangasarian–Fromowitz constraint qualification [62], the assumptions of a unique optimal solution of (38) on a whole neighborhood of η and those of second order differentiability for the true program (38) with $y = \eta$ can be relaxed to a certain extent by means of the concept of generalized derivative suggested by Rockafellar [52] and complemented by results on convergence in distribution for measurable multifunctions and selections by Salinetti and Wets [58]. This gives a possibility to generalize the asymptotic result (42) as done in detail for the linear–quadratic stochastic program with quadratic recourse function (6) and with $c_0(x, \omega) = c(x)$, a nonrandom quadratic function [40].

Extension of this methodology to more general topological spaces (see King [40–42]) also opens up quite new perspectives to the statistical approach and to the statistical sensitivity analysis with respect to the probability measure. Also in this case, quadratic programs with randomly perturbed linear term in the objective function play an essential role.

Acknowledgement

I gratefully acknowledge the helpful comments of Alan King, Werner Römisch, Alexander Shapiro and Roger Wets.

References

[1] R.L. Armacost and A.V. Fiacco, Computational experience in sensitivity analysis for nonlinear programming, Math. Programming 6 (1974) 301–326.

[2] H. Attouch and R.J.-B. Wets, Epigraphical approach, Annales Institut H. Poincaré, Analyse Non Linéaire (1989).

[3] J.-P. Aubin, Lipschitz behavior of solutions to convex minimization problems, Math. Oper. Res. 9 (1984) 87–111.

[4] B. Bank et al., *Non-Linear Parametric Optimization* (Akademie-Verlag, Berlin, 1982).

[5] B.M. Beale, On minimizing a convex function subject to linear inequalities, JRSS 17B (1955) 173–184.

[6] J.R. Birge and R.J.-B. Wets, Computing bounds for stochastic programming problems by means of a generalized moment problem, Math. Oper. Res. 12 (1987) 149–162.

[7] A. Charnes and W.W. Cooper, Chance-constrained programming, Manag. Sci. 5 (1959) 73–79.

[8] J.M. Danskin, *Theory of Max-Min, Econometrics and Operations Research*, vol. 5 (Springer, 1967).

[9] G.B. Dantzig, Linear programming under uncertainty, Manag. Sci. 1 (1955) 197–206.

[10] G.B. Dantzig and P.W. Glynn, Parallel processors for planning under uncertainty, *5th Int. Conf. on Stochastic Programming*, Ann Arbor (August 13–18, 1989).

[11] J. Dupačová, Experience in stochastical programming models, in: A. Prékopa (ed.), *Survey of Mathematical Programming*, Proc. 9th Int. Mathematical Programming Symp., Budapest 1976 (Akadémiai Kiado, Budapest, 1979) pp. 99–105.

[12] J. Dupačová, Stability in stochastic programming with recourse-estimated parameters, Math. Programming 28 (1984) 72–83.

[13] J. Dupačová, Stability in stochastic programming with recourse-contaminated distributions, Math. Programming Study 27 (1986) 133–144.

[14] J. Dupačová, Stochastic programming with incomplete information: A survey of results on postoptimization and sensitivity analysis, Optimization 18 (1987) 507–532.

[15] J. Dupačová, On some connections between parametric and stochastic programming, in: J. Guddat et al. (eds.), *Parametric Optimization and Related Topics*, Mathematical Research Band 35 (Akademie-Verlag, Berlin, 1987) pp. 74–81.

[16] J. Dupačová, On nonnormal asymptotic behavior of optimal solutions of stochastic programming problems: The parametric case, in: P. Mandl and M. Hušková (eds.), *Proc. 4th Prague Symp. on Asymptotic Statistics*, Charles University, Prague (1989) pp. 205–214. See also WP-88-19 IIASA, Laxenburg.

[17] J. Dupačová, On statistical sensitivity analysis in stochastic programming, *5th Int. Conf. on Stochastic Programming*, Ann Arbor (August 13–18, 1989).

[18] J. Dupačová and R.J.-B. Wets, Asymptotic behaviour of statistical estimators and of optimal solutions of stochastic optimization problems, Ann. Statist. 16 (1988) 1517–1549.

[19] Yu. Ermoliev and R.J.-B. Wets (eds.), *Numerical Techniques for Stochastic Optimization Problems* (Springer, Berlin, 1988).

[20] A.V. Fiacco, Sensitivity analysis for nonlinear programming using penalty methods, Math. Programming 10 (1976) 287–311.

[21] A.V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming* (Academic Press, New York, 1983).

[22] A.V. Fiacco and J. Kyparisis, Sensitivity analysis in nonlinear programming under second order assumptions, in: A.V. Balakrishnan and M. Tona (eds.), *Systems and Optimization*, Lecture Notes in Control and Information Sciences 66 (Springer, Berlin, 1985) pp. 74–97.

[23] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming Sequential Unconstrained Minimization Techniques* (Wiley, New York, 1968).

[24] E.G. Gol'shtein, *Vypukloje Programmirovanije, Elementy Teoriji* (Nauka, Moscow, 1970) [*Theory of Convex Programming*, Translations of Mathematical Monographs 36; American Mathematical Society, Providence, RI (1972)].

[25] F.R. Hampel, The influence curve and its role in robust estimation, J. Am. Statist. Assoc. 69 (1974) 383–397.

[26] F.R. Hampel et al., *Robust Statistics. The Approach Based on Influence Functions* (Wiley, New York, 1986).

[27] P.J. Huber, *Robust Statistics* (Wiley, New York, 1981).

[28] M. Iosifescu and R. Theodorescu, Sur la programmation linéaire, C.R. Acad. Sci. Paris 256 (1963) 4831-4833.

[29] K. Jittorntrum, Solution point differentiability without strict complementarity in nonlinear programming, Math. Programming Study 21 (1984) 127–138.

[30] P. Kall, *Stochastic Linear Programming* (Springer, Berlin, 1976).

[31] P. Kall, Computational methods for solving two-stage stochastic linear programming problems, J. Appl. Math. Phys. 30 (1979) 261–271.

[32] P. Kall, Approximation to optimization problems: An elementary review, Math. Oper. Res. 11 (1986) 9–18.

[33] P. Kall, On approximations and stability in stochastic programming, in: J. Guddat et al. (eds.), *Parametric Optimization and Related Topics*, Math. Research Band 35 (Akademie-Verlag Berlin, 1987) pp. 387–407.

[34] P. Kall, Stochastic programming with recourse: Upper bounds and moment problems, in: J. Guddat et al. (eds.), *Advances in Mathematical Optimization*, Mathematical Research vol. 45 (Akademie-Verlag, Berlin, 1988) pp. 86–103.

[35] P. Kall and D. Stoyan, Solving stochastic programming problems with recourse including error bounds, Math. Oper. Forsch. Statist. Ser. Optimization 13 (1982) 431–447.

[36] V. Kaňková, Optimum solution of a stochastic optimization problem with unknown parameters, *Trans. 7th Prague Conf. 1974* (Academia, Prague, 1977) pp. 239–244.

[37] V. Kaňková, An approximate solution of a stochastic optimization problem, *Trans. 8th Prague Conf. 1978* (Academia, Prague, 1978) pp. 349–353.

[38] V. Kaňková, Empirical estimates in stochastic programming, *Trans. 10th Prague Conf. 1986* (Academia, Prague, 1988) pp. 21–28.

[39] V. Kaňková, Estimates in stochastic programming – Chance constrained case, to appear in Problems of Control and Information Theory.

[40] A.J. King, Asymptotic behaviour of solutions in stochastic optimization: Nonsmooth analysis and the derivation of non-normal limit distribution, Dissertation, Univ. of Washington (1986).

[41] A.J. King, Generalized delta theorems for multivalued mappings and measurable selections, WP-88-56, IIASA, Laxenburg (1988), in Math. Oper. Res. (1989).

[42] A.J. King, Asymptotic distributions for solutions in stochastic optimizaton and generalized M-estimation, WP-88-58, IIASA, Laxenburg (1988).

[43] A.J. King and R.J.-B. Wets, Epi-consistency for convex stochastic programs, RC 14640 (# 65642) 5/26/89, IBM Research Division (1989).

[44] D. Klatte, A note on quantitative stability results in nonlinear optimization, in: K. Lommatzsch (ed.), *Proc. 19. Jahrestagung "Mathematische Optimierung"*, Seminarbericht No. 90, Humboldt-Universität, Berlin (1987) pp. 77–86.

[45] A. Prékopa, Logarithmic concave measures with application to stochastic programming, Acta Sci. Math. 32 (1971) 301–316.

[46] S.M. Robinson, Perturbed Kuhn–Tucker points and rates of convergence for a class of nonlinear programming algorithms, Math. Programming 7 (1974) 1–16.

[47] S.M. Robinson, Strongly regular generalized equations, Math. Oper. Res. 5 (1980) 43–62.

[48] S.M. Robinson, Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity, Math. Programming Study 30 (1986) 45–66.

[49] S.M. Robinson, Local epi-continuity and local optimization, Math. Programming 37 (1987) 208–222.

[50] S.M. Robinson and R.J.-B. Wets, Stability in two-stage stochastic programming, SIAM J. Control Optim. 25 (1988) 1409–1416.

[51] R.T. Rockafellar, Directional differentiability of the optimal value function in a nonlinear programming problem, Math. Programming Study 21 (1984) 213–226.

[52] R.T. Rockafellar, Proto-differentiability of set-valued mappings and its applications in optimization, Ann. Inst. H. Poincaré: Analyse Non Linéaire (1989).

[53] R.T. Rockafellar and R.J.-B. Wets, A Lagrangian finite generation technique for solving linear–quadratic problems in stochastic programming, Math. Programming Study 28 (1986) 63–93.

[54] W. Römisch and R. Schultz, Distribution sensitivity in stochastic programming, Preprint No. 160, Humboldt-Universität zu Berlin (1987), submitted to Math. Programming.

[55] W. Römisch and R. Schultz, Stability of solutions for stochastic programs with complete recourse having $C^{1,1}$ data, Inst. für Operations Research der Universität Zürich (1989).

[56] W. Römisch and R. Schultz, Distribution sensitivity for the chance constrained model of optimal load dispatch (1989) submitted to J. Optim. Theory Appl.

[57] W. Römisch and A. Walkobiner, Obtaining convergence rates for approximations in stochastic programming, in: J. Guddat et al. (eds.), *Parametric Optimization and Related Topics*, Mathematical Research Band 35 (Akademie-Verlag Berlin, 1987) pp. 327–343.

[58] G. Salinetti and R.J.-B. Wets, On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima, Math. Oper. Res. 11 (1986) 385–419.

[59] R.J. Serfling, *Approximation Theorems in Mathematical Statistics* (Wiley, New York, 1980).

[60] A. Shapiro, Second order sensitivity analysis and asymptotic theory of parametrized nonlinear programs, Math. Programming 33 (1985) 280–299.

[61] A. Shapiro, Asymptotic properties of statistical estimators in stochastic programming, Research Report 55/87 (17), University of South Africa, Ann. Statist. 17 (1989) 841–858.

[62] A. Shapiro, Sensitivity analysis of nonlinear programs and differentiability of metric projections, SIAM J. Control Optim. 26 (1988).

[63] A. Shapiro, On differential stability in stochastic programming, Math. Programming (1990).

[64] E. Tamm, On solvability of the nonlinear programming problem with a random parameter, ENSV TA Toimetised Füüsika Matemaatika 30 (1981) 220–224 (in Russian).

[65] G. Tintner, Stochastic linear programming with applications to agricultural economics, in: *Proc. 2nd Symp. on Linear Programming*, Washington (1955) pp. 197–207.

[66] A.B. Tsybakov, Error bounds for the method of minimization of empirical risk; transl. from Problemy Peredachi Informatsii 17 (1981) 50–61.

[67] S. Vogel, Stability results for stochastic programming problems, Optimization 19 (1988) 269–288.

[68] J. Wang, Distribution sensitivity analysis for stochastic programs with complete recourse, Math. Programming 31 (1985) 286–297.

[69] J. Wang, Continuity of feasible solution sets of probabilistic constrained programs, submitted to J. Optim. Theory Appl.

[70] R.J.-B. Wets, A statistical approach to the solution of stochastic programs with (convex) simple recourse, Working Paper, University of Kentucky (1979).

[71] R.J.-B. Wets, Stochastic programming: Solution techniques and approximation schemes, in: A. Bachem et al. (eds.), *Mathematical Programming – The State of the Art*, Bonn, 1982 (Springer, Berlin, 1983) pp. 566–603.

[72] R.J.-B. Wets, Stochastic programming, to appear in: *Handbook on Operations Research*.

[73] J. Žáčková, On minimax solutions of stochastic linear programming problems, Čas. pěst. mat. 91 (1966) 423–430.