STABILITY RESULTS FOR STOCHASTIC PROGRAMS AND SENSORS, ALLOWING FOR DISCONTINUOUS OBJECTIVE FUNCTIONS

by

Zvi Artstein†*

Department of Theoretical Mathematics
The Weizmann Institute of Science
Rehovot 76100, Israel

and

Roger J-B Wets†

Department of Mathematics
University of California Davis
Davis, CA 95616, U.S.A.

SIAM J. on Optimization, 4 (1995), 537-550

Abstract. The paper examines the stability of the optimal value and the solutions of stochastic programming problems. We check stability with respect to variations in both the problem formulation and the probability distribution that describes the uncertainty. Of particular interest is the case where the payoff functions may be discontinuous. We apply these results to analyze the stability of sensors, that model the possibility of making inquiries to improve the probabilistic information available about the uncertain quantities.

†Research supported by a grant from the U.S.-Israel Binational Science Foundation.
*Incumbent of the Hettie H. Heineman Professorial Chair in Mathematics.

1. INTRODUCTION

In this paper we study the stability of the value and solutions of stochastic programs. The stability checked is with respect to perturbations in both the payoff functional and the probability distribution that describes the uncertainty. For the probability distribution we consider perturbations in the sense of the weak convergence of measures. It is natural to check the stability with respect to this convergence, as it reflects errors in measurements or in sampling. For the payoff functional we come up with a new convergence notion for integrands. This convergence could be interpreted as epi-convergence on a function space equipped with a convergence notion related to graph convergence. It allows discontinuities, thus enabling us to deal with a richer class of problems than has been possible so far. We provide examples demonstrating the need for discontinuous payoff functionals, and show how the convergence introduced in this paper applies. The results are then used to verify the stability of the value of sensors.

We work with an abstract formulation. The stochastic optimization problem considered here is of the form

$$(*) \qquad \operatorname*{maximize}_{x\in X} \int_\Xi f(x,\xi)P(d\xi)$$

where X and Ξ are complete separable metric spaces, and P is a probability measure on Ξ; the latter space is equipped with its Borel structure. The payoff functional f is a mapping

$$f(x,\xi) : X \times \Xi \longrightarrow [-\infty,\infty] \ ,$$

and it is assumed to be, at least, measurable in the ξ variable. For a fixed x the integral is taken in the extended sense, and if not defined, it is set equal to $-\infty$.

As mentioned, the stability of the problem $(*)$ is examined with respect to variations both in P and in f. Related questions were examined in the literature, and from various aspects, see e.g. Kall [9], Wang [20], Robinson and Wets [14], Dupačová [7], [8], Römisch and Schultz [16], [17]. These papers consider variations in P to be in the sense of weak convergence. The payoff functionals in these papers are assumed to be continuous. The analysis of [14] assumes, for instance, continuity of the integrands, and use uniform convergence on compacta as the convergence mode. Stability with discontinuous functionals was considered by Langen [11] and by Schultz [18]; we comment on these papers in the body of the text. What the present paper offers is a general framework for checking the continuity of the value with respect to functional convergence of discontinuous payoff functions, and the weak convergence of the underlying probabilities. Aspects other than continuity are examined in Kall [9], Dupačová [7], [8] and Römisch and Schultz [16], [17], notably the differentiability of the value. The setting of our paper is more general, but we examine only continuity properties.

The paper is organized as follows. In the next section we display two examples where discontinuities in the payoff functional occur. These examples are used in later sections to illustrate and check the abstract results. In Section 3 we examine semicontinuity of the integration in $(*)$ for a fixed decision x, and in Section 4 the continuity of this integral is studied. These results form a major step in the analysis of the stability of the value, which we come to in Section 5. We also provide there a result on the stability of the solution set, and a result on robustness of optimal solutions. The applications of the stability analysis to the examples of Section 2 are explored in Section 6. Sensors were introduced in [2] as a tool to evaluate inquiries into the structure of the uncertainty; the application of the stability analysis to the stability of sensors is given in Section 7, along with a telegraphic introduction to the sensors model.

2. EXAMPLES

Dicontinuous payoff functions arise in a variety of contexts, such as linear stochastic programs with random recourse matrix, stochastic programs with discontinuous recourse costs, multistage stochastic programs with chance constraints, etc.. A particularly rich collection of examples is supplied by stochastic integer programming problems, see e.g., Laporte, Louveaux and Mercure [12], Schultz [18] [19] and Klein Haneveld, Stougie and van der Vlerk [10]. We present here two simple examples with discontinuous payoff functions; they are used in the sequel as illustrations for, and tests of, the abstract results.

Example 2.1. The newsboy problem with mandated backorder. This is a simple inventory problem in the form of the classical newsboy problem, as treated, e.g., in Richmond [13, Chapter 7].

A newsboy has to place an order for a number, say x, of copies. Each copy costs him c dollars, while he sells the newspapers at r dollars apiece. These newspapers are sold until the demand is met. If, however, the demand is greater than x, then the newsboy is obliged to backorder the missing copies. This extra effort costs him K dollars, a cost which does not depend on the number of papers to backorder (say, when the newsboy has to send out a truck to collect the backorder).

Thus, for a given demand ξ and decision x, the newsboy's payoff is

$$f(x,\xi) = \begin{matrix} r\xi - cx & \text{if} \quad \xi \le x \\ (r-c)\xi - K & \text{if} \quad \xi > x \ . \end{matrix} \tag{2.1}$$

If the demand ξ is known, then $x = \xi$ would maximize the payoff. But in the situation we analyze, the demand ξ is a random variable whose realization is not known to the newsboy when the order x is placed. The probability distribution P that governs ξ is, however, available to the newsboy. Thus, if the newsboy wishes to maximize expected payoff, his decision problem is

$$\text{(2.2)} \qquad \underset{0\le x<\infty}{\text{maximize}} \int_0^\infty f(x,\xi)P(d\xi)$$

with f given in (2.1).

Stability of the problem (2.2) means that the value and solutions do not vary too much when only small perturbations occur in the data. One perturbation that should be considered is that the true probability distribution is not exactly equal to P. Another deviation may be that the fixed backorder cost is not K, but another number, hopefully close to K. The place of the discontinuity, namely the constraint $\xi \le x$ in (2.1), could also be only an approximation of the true value, as a small number of unsatisfactory customers may be allowed. There may be limits not known in advance on the size of the backorder; then the second line in (2.1) is only an approximation of a term

$$(r-c)\, min(\xi, L) - K \quad \text{if} \quad \xi > x$$

with L large, and not known in advance. In all the preceding situations, it is clear what should be meant by "small deviations" in the payoff. The convergence introduced in this paper agrees with this intuitive perception.

Example 2.2. A mixed integer-linear stochastic recourse problem. We exhibit a simplified version of the general two-stage stochastic program with the integer-linear recourse process. (The example was added after a referee had kindly pointed out to us the paper by Schultz [18].) The problem is as follows (we index the parameters by nought for later use).

$$\text{(2.3)} \qquad \underset{-\infty<x<\infty}{\text{maximize}} \; c_0(x) - \int_{-\infty}^\infty \Phi_0(|x-\xi|)P_0(d\xi)$$

where the function Φ_0 is determined by the optimization procedure

$$\text{(2.4)} \qquad \Phi_0(b) = \min\{q_0 y + r_0 z : \alpha_0 y + \beta_0 z = b,\; y \in R_+,\; z \in Z_+\}\ .$$

Here $R_+ = [0, \infty)$ and $Z_+ = \{0, 1, 2, \ldots\}$; the constants q_0, r_0, α_0 and β_0 are assumed positive.

The interpretation is that the decision x is made first, with payoff $c_0(x)$. Then a random quantity ξ is realized, and a recourse action is performed by choosing y and z such that $\alpha_0 y + \beta_0 z = |x-\xi|$. The cost in turn is $q_0 y + r_0 z$, and at the recourse stage it should be minimized. The quantity z is thought of as the cheaper one (this would be reflected by the condition $\frac{q_0}{\alpha_0} > \frac{r_0}{\beta_0}$, but we do not need this); in turn, z is restricted to be an integer. As we see in Section 6, it is this nature of recourse that generates discontinuities in the integrand $\Phi_0(\cdot)$.

We are interested in the stability of the value of (2.3), with respect to perturbations in the data $c_0(\cdot), q_0, r_0, \alpha_0, \beta_0$ and the underlying distribution P_0.

3. SEMICONTINUITY OF THE INTEGRAL

In this section we consider a preliminary, yet a crucial, step toward the stability analysis. Let Ξ again be a complete separable metric space, endowed with its Borel structure. Consider the integral

$$I(g,P) = \int_{\Xi} g(\xi)P(d\xi) \tag{3.1}$$

where P is a probability measure on Ξ, and $g : \Xi \to [-\infty, \infty]$ is a measurable function. We allow $I(g,P) = +\infty$ or $I(g,P) = -\infty$, and when $I(g,P)$ is still not defined, we set it to be $-\infty$. We are interested in the semicontinuity properties of $I(\cdot,\cdot)$ in its two variables. Specifically, for convergence modes yet to be chosen, we look for conditions on a pair (g_0, P_0) such that $I(g,P)$ is semicontinuous at (g_0, P_0) ,namely, either lower or upper semicontinuous. Note that the integration in (3.1) is a degenerate case of the optimization problem $(*)$, namely when the decision space is a singleton.

Langen [11, Theorem 3.3] established semicontinuity results for $I(g,P)$, in the spirit of our result Theorem 3.5 (we thank a referee for pointing this out.) The convergence on sequences of functions g_k used in [11] could be called lower semicontinuous convergence, namely $\xi_k \to \xi_0$ implies $\liminf\ g_k(\xi_k) \geq g_0(\xi_0)$. When continuity is sought, and both upper and lower semicontinuous convergences are assumed, then one gets the continuous convergence, i.e., $\xi_k \to \xi_0$ implies $g_k(\xi_k) \to g_0(\xi_0)$, which is the convergence used by Robinson and Wets [14, Theorem 2.1]. In particular, in the case of continuity, both [11] and [14] need a continuous limit function $g_0(\cdot)$. Here and in the next section we relax the notion of convergence, allowing in particular discontinuous limit integrands $g_0(\cdot)$. We propose a type of graph lower convergence for g_k. For the measures P we use, as do [11] and [14], weak convergence of measures. Both notions are presented now.

We denote by $\mathcal{P}$ the family of probability measures on Ξ, and by $\mathcal{G}$ the family of measurable functions $g : \Xi \to [-\infty, \infty]$.

Definition 3.1. (See e.g. Billingsley [5].) The sequence P_k of probability measures on Ξ converges weakly to P_0 if for every bounded and continuous function $h : \Xi \to (-\infty, \infty)$, the integrals $I(h, P_k)$ converge to $I(h, P_0)$.

Since Ξ is a complete separable space, the weak convergence on $\mathcal{P}$ is metrizable. The Prohorov metric induces weak convergence, see [5]. The relevance of the weak convergence to variational problems can be observed through the structure of the Prohorov metric. The Wasserstein metrics, used by Römisch and Schultz [16], [17], have a similar structure, and could be used as well.

We need the following notation. For $g \in \mathcal{G}$ we denote by $lc\ g$ the lower closure of g, namely $(lc\ g)(\xi) = \liminf\ g(\eta)$ as $\eta \to \xi$. Similarly, $uc\ g$ denotes the upper closure of g, namely $(uc\ g)(\xi) = \limsup\ g(\eta)$ as $\eta \to \xi$.

Definition 3.2. The function g_0 in $\mathcal{G}$ is an epi-sublimit of the sequence g_k if

$$(lc\ g_0)(\xi) \ \leq \ \liminf\ g_k(\xi_k) \quad \text{as} \quad \xi_k \to \xi \ ,$$

for all $\xi \in \Xi$. The function g_0 in $\mathcal{G}$ is a hypo-suplimit of g_k if $-g_0$ is an epi-sublimit of the sequence $-g_k$, or equivalently, if

$$(uc\ g_0)(\xi) \ \geq \ \limsup\ g_k(\xi_k) \quad \text{as} \quad \xi_k \to \xi$$

for all $\xi \in \Xi$.

In the case that g_0 is finite-valued and continuous, and g_k continuously converge to g_0, then g_0 is both an epi-sublimit and a hypo-suplimit of a sequence g_k. It is easy to see that more generally, if the graphs of the functions g_k converge to the closure of the graph of g_0 in the sense of set-convergence in $\Xi \times [-\infty, \infty]$, then g_0 is both an epi-sublimit and a hypo-suplimit of the sequence g_k. Notice that g_0 is an epi-sublimit of g_k if the epigraphs of g_k, i.e. $\{(\xi, r) : r \geq g_k(\xi)\}$, lower converge in the Painlev'e-Kuratowski sense to the closure of the epigraph of g_0. The closure operation, however, implies that even if sublimit is replaced by limit (i.e., inequality is replaced by equality in the definition), then the limit is not necessarily unique. For instance, let g be the Dirichlet function, namely $g(r) = 1$ for r rational in $[0,1]$, and $g(t) = 0$ otherwise. Then the constant sequence $g, g, \ldots$ has $1 - g$ as both epi-sublimit and hypo-suplimit.

We are aiming at continuity results for $I(g, P)$ when $\mathcal{P}$ is endowed with weak convergence, and in $\mathcal{G}$ we consider a convergence based on the notions of epi-sublimit and hypo-suplimit. Additional conditions have to be posed in order to guarantee the semi-continuity. One issue that may cause problems is the unboundedness of the integrands. Indeed, even for g fixed, $I(g, P)$ is not semicontinuous with respect to the variable P if g is unbounded. Imposing uniform boundedness on g is too restrictive in the framework of variational problems, as objective functions often assume the value $-\infty$. We follow here Robinson and Wets [14] and Kall [9], and restrict ourselves to equi-tight (uniformly integrable in the terminology of [14]) families, as follows.

Definition 3.3. Let $\mathcal{W}$ be a subset of pairs (g, P) in $\mathcal{G} \times \mathcal{P}$. The family $\mathcal{W}$ is equi-tight if for every $\epsilon > 0$ a compact set K_ϵ, and a bound b_ϵ exist, such that for all $(g, P) \in \mathcal{W}$ the following three conditions hold.

(1) $P(\Xi \backslash K_\epsilon) < \epsilon$

(2) $|g(\xi)| \leq b_\epsilon$ if $\xi \in K_\epsilon$

(3) $\int_{\Xi \backslash K_\epsilon} |g(\xi)| P(d\xi) < \epsilon$

The preceding definition is a slight generalization of Robinson and Wets [14]; the latter does not require explicitly the boundedness (2), but the boundedness holds in [14], and follows from the continuity condition on the g_k and the uniform convergence; also, [14] considers families $\mathcal{W}$ of the form $\mathcal{G}_1 \times \mathcal{P}_1$ with $\mathcal{G}_1 \subset \mathcal{G}$ and $\mathcal{P}_1 \subset \mathcal{P}$.

Equi-tightness does not suffice to guarantee semicontinuity; one needs in addition, some information on the measure of the discontinuity points. To this end we need the following terminology.

Notation 3.4. We say that ξ is a lower continuity point of $h : \Xi \to [-\infty, \infty]$, if $\liminf\ h(\eta) \geq h(\xi)$ as $\eta \to \xi$. The set of points that are not lower continuity points of h is denoted by $LDisc\ h$. Likewise, ξ is an upper continuity point of h if it is a lower continuity point of $-h$. The set of points which are not upper continuity points is denoted $UDisc\ h$.

The sets $LDisc\ h$ and $UDisc\ h$ may not be measurable. When h itself is measurable, these sets are analytic. Indeed, $LDisc\ h$ is given by

$$\bigcup_{\epsilon>0} \bigcap_{\delta>0} proj_1\{(\xi,\eta) : d(\xi,\eta) < \delta,\ h(\eta) - h(\xi) > -\epsilon\} \tag{3.2}$$

where $proj_1$ is the projection on the first coordinate of (ξ, η), and $d(\cdot,\cdot)$ is the metric on Ξ. The subset of $\Xi \times \Xi$ appearing in (3.2) is then measurable, hence its projection is analytic (see the projection theorem in Castaing and Valadier [6, III.4.23]). Since the intersection and the union in (3.2) can be performed denumerably, the analyticity of the resulting set follows. Being analytic, the set $LDisc\ g$ for $g \in \mathcal{G}$ belongs to the completion of the Borel field of Ξ with respect to any measure P in $\mathcal{P}$. In particular, the values $P(LDisc\ g)$, for $g \in \mathcal{G}$ and $P \in \mathcal{P}$, are well defined.

The following is the main result of the present section.

Theorem 3.5. Suppose that g_0 is an epi-sublimit of g_k and suppose that P_k converge weakly to P_0 in $\mathcal{P}$. Suppose that $\mathcal{W} = \{(g_k, P_k) : k = 0, 1, 2, \ldots\}$ is equi-tight, and that $P_0(LDisc\ g_0) = 0$. Then $\liminf\ I(g_k, P_k) \geq I(g_0, P_0)$.

Proof. Let $\epsilon > 0$ be fixed. Let K_ϵ and b_ϵ be provided by the equi-tightness of $\mathcal{W}$ (Definition 3.3). Since $P_0(LDisc\ g_0) = 0$, it follows that for every $\delta > 0$ an open set D_δ exists such that D_δ contains $LDisc\ g_0$, and $P_0(D_\delta) < \delta$. We choose δ and D_δ such that $\delta b_\epsilon \leq \epsilon$.

Consider the set $K = K_\epsilon \backslash D_\delta$. This set is compact and contains only lower continuity points of g_0. In particular, g_0 is lower semicontinuous on K. As such, the restriction of g_0 to K is the pointwise limit of a monotone increasing sequence of continuous functions on K. (See Beer [4] for a proof and a useful algorithm for constructing such a sequence.) Therefore a continuous function, say h_0, on K exists such that

$$\int_K |g_0(\xi) - h_0(\xi)| P_0(d\xi) < \epsilon\ . \tag{3.3}$$

We now extend h_0 to the entire space Ξ by letting $h_0(\xi) = g_0(\xi)$ if $\xi \notin K$. Then, since $h_0(\xi) \le g_0(\xi)$ on K, it follows that each $\xi \in K$ is a lower continuity point of h_0, and h_0 is an epi-sublimit of g_k. We prove now

$$\liminf \; I(g_k, P_k) - I(h_0, P_0) \ge -8\epsilon \; . \tag{3.4}$$

Together with (3.3), and the observation that ϵ is arbitrarily small, (3.4) would complete the proof.

To verify (3.4), we choose a continuous function h_1 defined on an open neighborhood, say Q_1, of K, and such that $h_1(\xi) = h_0(\xi)$ for $\xi \in K$. Such a function h_1 clearly exists. We claim that an open neighborhood Q of K can be found, such that $Q \subset Q_1$ and such that the following three conditions are satisfied.

(i) $\int_{Q\setminus K}(|h_0(\xi)| + |h_1(\xi)|)P_0(d\xi) < \epsilon$

(ii) For some k_0, $g_k(\xi) - h_1(\xi) \ge -\epsilon$ for $\xi \in Q$ and $k \ge k_0$

(iii) $P_0(\partial Q) = 0$ (where ∂Q denotes the boundary of Q).

To verify the existence of such a Q we use a compactness argument as follows. The estimate for the integral in (i) is satisfied for some neighborhood of K since $P_0(K) = 0$, therefore $P_0(Q'\setminus K)$ converge to zero as $Q'\setminus K$ gets smaller. To establish the estimates in (ii) and (iii) consider first a point ξ_0 in K, and establish (ii) and (iii) for a neighborhood Q_{ξ_0} of ξ_0 . To find a neighborhood of ξ_0 for which (ii) holds, notice that the condition $h_1(\xi_0) \le (lc\; g_0)(\xi_0)$ and the epi-sublimit property imply that there exists a neighborhood V_{ξ_0} of ξ_0 and a k_0 such that $g_k(\xi') \ge h_1(\xi_0) - \frac{\epsilon}{2}$ if $k \ge k_0$ and $\xi' \in V_{\xi_0}$. On the other hand, the continuity of $h_1(\cdot)$ in a neighborhood of K implies that a neighborhood V'_{ξ_0} of ξ_0 exists such that $h_1(\xi_0) \ge h_1(\xi') - \epsilon/2$ for all $\xi' \in V'_{\xi_0}$. Then for $\xi' \in V_{\xi_0} \cap V'_{\xi_0}$, the inequality $g_k(\xi') \ge h_1(\xi') - \epsilon$ holds for $k \ge k_0$, hence $V_{\xi_0} \cap V'_{\xi_0}$ is a neighborhood of ξ_0 for which (ii) is satisfied. A standard argument could produce a neighborhood Q_{ξ_0} of ξ_0 which is a subset of $V_{\xi_0} \cap V'_{\xi_0}$, and such that $P(\partial Q_{\xi_0}) = 0$. For this Q_{ξ_0} both (ii) and (iii) are satisfied. The compactness of K implies that the union of a finite number of such Q_{ξ_0}, with $\xi_0 \in K$, covers K. This union is a neighborhood of K for which (ii) and (iii) are satisfied, and its intersection with the neighborhood satisfying (i) is the desired set Q.

The conditions $P_k \to P_0$, and $P_0(\partial Q) = 0$, and the continuity of h_1 on Q, imply that

$$\lim \left| \int_Q h_1(\xi)P_k(d\xi) - \int_Q h_1(\xi)P_0(d\xi) \right| = 0 \; . \tag{3.5}$$

Therefore, by (i) and (ii),

$$\liminf \left(\int_Q g_k(\xi)P_k(d\xi) - \int_Q h_0(\xi)P_0(d\xi) \right) \ge -2\epsilon \; . \tag{3.6}$$

This is not quite what is required in (3.4) as the integration in the latter is on the entire space Ξ. We note, however, that the equi-tightness implies that

$$\int_{\Xi\backslash K_\epsilon} |g_k(\xi)|P_k(d\xi) < \epsilon \tag{3.7}$$

for all $k = 0, 1, 2, \ldots$ (and recall that $h_0 = g_0$ on $\Xi\backslash K_\epsilon$). Also, since $K_\epsilon\backslash Q$ is a compact set included in D_δ, it follows that $P_0(K_\epsilon\backslash Q) \leq \delta$, and since $P_k \to P_0$ it follows that

$$\limsup \int_{K_\epsilon\backslash Q} |g_k(\xi)|P_k(d\xi) \leq \delta b_\epsilon \leq \epsilon \ . \tag{3.8}$$

Combining (3.6), (3.7), (3.8), and noting that $Q \cup (\Xi\backslash K_\epsilon) \cup (K_\epsilon\backslash Q) = \Xi$, the estimate in (3.4) is verified. This completes the proof.

The following is the upper semicontinuity result, analogous to the previous one; it follows immediately from the definition of hypo-suplimit, and the lower semicontinuity result.

Theorem 3.6. Suppose g_0 is a hypo-suplimit of g_k, and that P_k converge weakly to P_0. Suppose that the family $\mathcal{W} = \{(g_k, P_k) : k = 0, 1, \ldots\}$ is equi-tight, and $P_0(UDisc\, g_0) = 0$. Then $\limsup\ I(g_k, p_k) \leq I(g_0, P_0)$.

It is clear that the assumptions $P_0(LDisc\ g_0) = 0$ and $P_0(UDisc\ g_0) = 0$ cannot be removed from the previous results. Such an assumption is not used in Langen [11, Theorem 3.3], as the semicontinuous convergence used in [11] (see the early part of this section) makes it irrelevant where the discontinuities occur.

4. CONTINUITY OF THE INTEGRAL

We use the two theorems of the previous section to derive conditions guaranteeing the continuity of $I(g, P)$. to this end we need the following terminology.

Notation 4.1. For $h : \Xi \to [-\infty, \infty]$ we denote by $Disc\ h$ the set of points ξ in Ξ at which h is discontinuous.

Note that $Disc\ h$ is the union of $LDisc\ h$ and $UDisc\ h$. The set $Disc\ h$ is Borel even if h itself is not measurable. Indeed, for a bounded h the set $Disc\ h$ is a denumerable union of closed sets $Disc_\epsilon\ h$, consisting of the points ξ in Ξ at which the discontinuity is at least ϵ, i.e., $\limsup |h(\eta) - h(\zeta)| \geq \epsilon$ as η and ζ converge to ξ. For unbounded h we notice that $Disc\ h$ is preserved under the transformation $h \to h(1 + |h|)^{-1}$ (with $\infty \cdot \infty^{-1} = 1$), which transforms h into a bounded mapping.

Theorem 4.2. Suppose that g_0 is both an epi-sublimit and a hypo-suplimit of the sequence g_k, and that P_k converges weakly to P_0. Suppose that the family $\mathcal{W} = \{(g_k, P_k) : k = 0, 1, \ldots\}$ is equi-tight, and $P_0(Disc\ g_0) = 0$. Then $I(g_k, P_k)$ converge to $I(g_0, P_0)$.

Proof. The conditions of both Theorem 3.5 and Theorem 3.6 hold, and the two semicontinuity conclusions imply the continuity.

Remark 4.3. A referee pointed out to us that Theorem 4.2 has an alternative direct proof, using Rubin's Theorem (see Billingsley [5, Theorem 5.5]). A proper change of variables in our problem reduces the convergence to the case of Rubin's result. We rely on the semicontinuity of the preceding section. This semicontinuity is of interest for its own sake, and our proof yields also the estimate in Proposition 4.5. In turn, it might be of interest to interpret our semicontinuity result in the framework of Rubin's Theorem.

Remark 4.4. As we mentioned already, the continuity result improves available results; e.g., Robinson and Wets [14] where all integrands are assumed continuous (hence $P_0(Disc\ g_0) = 0$ since $Disc\ g_0$ is empty), and the convergence $g_k \to g_0$ is uniform on compact sets. Our setting is more relaxed, but notice that under the conditions of Theorem 4.2, the functions g_k converge to g_0 uniformly on compact subsets of the continuity points of g_0. The assumption $P_0(Disc\ g_0) = 0$ that we add is natural in the presence of discontinuities (it is used in Billingsley [5, Chapter 5] in the analysis of weak convergence, which is used in turn in Schultz [18], in his study of stability in stochastic programs; it was also used in [1] in the context of variational limits). If this assumption is dropped, the convergence may fail even for g_0 fixed. To get an example, let $g_0(\xi) = 0$ for $0 \leq \xi < 1$ and $g_0(\xi) = 1$ for $1 \leq \xi \leq 2$, let P_0 have an atom at $\xi = 1$ and let P_k have an atom at $\xi = 1 - k^{-1}$. Similarly, the continuity may then fail for P_0 fixed. For instance, let $g_0(\xi) = 0$ for $0 \leq \xi \leq 1$, $g_0(\xi) = 1$ for $1 < \xi \leq 2$ and P_0 as before, and $g_k(\xi) = min(\xi^k, 1)$.

Without the assumption $P_0(Disc\ g_0) = 0$, the continuity of the integral may fail. Our method of proof, however, provides an estimate for the discontinuity that may result. We present here a result along this direction, in the case where the integrands are bounded; analogous results can be derived for the equi-tight case.

Proposition 4.5. Suppose that g_0 is both an epi-sublimit and a hypo-suplimit of g_k, and $P_k \to P_0$. Suppose that g_k are uniformly bounded by the same bound β. Let $\Delta = P_0(Disc\ g_0)$. Then

$$\limsup |I(g_k, P_k) - I(g_0, P_0)| \leq 2\beta\Delta\ .$$

Proof. We follow the proof of the lower semicontinuity in Theorem 3.5, but choose D_δ to be such such that $P_0(D_\delta) \leq \Delta + \delta$. This affects the analysis when (3.8) is derived, and $\delta\beta_\epsilon$ there should be replaced by $(\delta + \Delta)\beta$. This change is carried back to (3.4) and the effect is that -8ϵ is replaced by $-2\Delta\beta - 8\epsilon$, and since ϵ is arbitrarily small we are left with $-2\Delta\beta$. The analogous argument for the upper semicontinuity completes the proof.

5. STABILITY ANALYSIS

In this section we use the semicontinuity and the continuity results of the previous sections in order to derive stability results for the stochastic optimization problem $(*)$. We consider a sequence (f_k, P_k) of pairs of integrands and probability measures. We find conditions under which the values of the corresponding stochastic problems converge to the value associated with the pair (f_0, P_0). To this end let

$$val(f_k, P_k) = \operatorname*{supremum}_{x \in X} \int_\Xi f_k(x, \xi) P_k(d\xi), \tag{5.1}$$

where the integral is set to be equal to $-\infty$ if not defined.

As a preliminary step toward the stability we consider the convergence of the functionals γ_k defined by

$$\gamma_k(x) = \int_\Xi f_k(x, \xi) P_k(d\xi) \ . \tag{5.2}$$

Notice that $val(f_k, P_k)$ is the supremum over $x \in X$ of $\gamma_k(x)$.

A major tool in the analysis of stability of minimization problems is the concept of epi-convergence of the cost functional, which becomes hypo-convergence in the case of maximization problems; for a detailed account, cf. Attouch [3] or Rockafellar and Wets [15]. Recall that the sequence $\gamma_k : X \to [-\infty, \infty]$ hypo-converges to γ_0 if at every point x, lim sup $\gamma_k(x_k) \le \gamma_0(x_0)$ when $x_k \to x_0$, and there exists at least one sequence $y_k \to x_0$ such that lim $\gamma_k(y_k) = \gamma_0(x_0)$. (Hypo-convergence can be characterized in terms of the set-convergence of the hypographs, see [15].)

We state now the main connections between convergence properties of the pairs (f_k, P_k) and the associated upper semicontinuity properties, and hypo-convergence for the functions γ_k.

Proposition 5.1. Suppose that whenever $x_k \to x_0$ in X, the function $f(x_0, \cdot)$ is a hypo-suplimit of $f(x_k, \cdot)$, and $\mathcal{W} = \{(f(x_k, \cdot), P) : k = 0, 1, \ldots\}$ is equi-tight. Suppose also that $P(UDisc\ f(x_0, \cdot)) = 0$ for all x_0. Then γ is upper semicontinuous.

Proof. The assumptions of Theorem 3.6 hold for $g_k(\cdot) = f(x_k, \cdot)$ and $P_k = P$ for all k. The conclusion then of Theorem 3.6 amounts to the upper semicontinuity of γ.

Theorem 5.2. Suppose that P_k converge to P_0 and that whenever $x_k \to x_0$, the function $f_0(x_0, \cdot)$ is a hypo-suplimit of $f_k(x_k, \cdot)$, and that for each x_0, there is at least one sequence $y_k \to x_0$ such that $f_0(x_0, \cdot)$ is an epi-sublimit of $f_k(y_k, \cdot)$. Suppose that whenever $x_k \to x_0$ the set $\mathcal{W} = \{(f_k(x_k, \cdot), P_k) : k = 0, 1, \ldots\}$ is equi-tight, and that $P_0(Disc\ f_0(x_0, \cdot)) = 0$ for all x_0. Then γ_k hypo-converge to γ_0.

Proof. By Theorem 3.6, $limsup\ \gamma_k(x_k) \le \gamma_0(x_0)$ whenever $x_k \to x_0$, and by Theorem 3.5, $liminf\ \gamma_k(y_k) \ge \gamma_0(x_0)$ for the specific sequence y_k guaranteed by the conditions. This completes the proof.

Notice the structure in the previous result. The conditions resemble a kind of hypo-convergence for the functional $f(x, \cdot)$, however with respect to the hypo-suplimit and the epi-sublimit notions for functions defined on Ξ, rather than with respect to convergence of real numbers. The additional conditions (equi-tightness and the measure zero condition) guarantee that this hypo-limit notion for function-valued functionals $f_k(x, \cdot)$, is translated into the standard hypo-limit of the functions γ_k.

Once the upper semicontinuity and the hypo-convergence of γ_k are established, existence results and the stability of the value follow in a direct manner. We display here several such results. A more elaborate discussion, which could be applied here, can be found in Rockafellar and Wets [15].

Theorem 5.3. Under the conditions of Proposition 5.1, if X is compact, then a solution to $(*)$ exists.

Proof. An upper semicontinuous function on a compact set attains its maximum, hence γ attains its maximum, and the point x at which the maximum is attained, is an optimal solution.

Theorem 5.4. Under the conditions of Theorem 5.2, if X is compact, then $val(f_k, P_k)$ converges to $val(f_0, P_0)$.

Proof. Let x_k be such that $val(f_k, P_k) - \gamma_k(x_k) \le k^{-1}$. Compactness implies that x_k has a converging subsequence, say the whole sequence converges to x_0. The conditions imply that $f_0(x_0, \cdot)$ is a hypo-suplimit of $f_k(x_k, \cdot)$. By Theorem 3.6 $\limsup\ \gamma_k(x_k) \le \gamma_0(x_0)$. Since $\gamma_0(x_0) \le val(f_0, P_0)$, the choice of x_k concludes the first part of the result. For the second part note that the conditions of Theorem 3.5 are fulfilled for $g_k(\cdot) = f_k(y_k, \cdot)$ and P_k, with $y_k \to x_0$ and x_0 such that $val(f_0, P_0) - \gamma_0(x_0) \le \epsilon$, and ϵ arbitrarily small. Then $\gamma_k(y_k) \to \gamma_0(x_0)$, which proves that $\liminf\ val(f_k, P_k) \ge val(f_0, P_0)$, and together with the first part, the proof is complete.

We present a result concerning the behavior of the solution set of $(*)$. Denote by $\sigma(f_k, P_k)$ the set of x such that $\gamma_k(x) = val(f_k, P_k)$, namely, $\sigma(f_k, P_k) = argmax\ \gamma_k$.

Theorem 5.5. Suppose that the conditions of Proposition 5.1 hold for each f_k, $k = 0, 1, \ldots$. Suppose also that the conditions of Theorem 5.2 hold. If X is compact, then $\limsup\ \sigma(f_k, P_k)$ is included in $\sigma(f_0, P_0)$.

Proof. From Proposition 5.1, γ_k is upper semicontinuous, and from Theorem 5.2 γ_k hypo-converge to γ_0. The semicontinuity of the set-valued mapping $argmax\ \gamma_k$ then follows in a standard way.

We conclude our stability analysis with the question of robustness, namely, suppose that an optimal solution for the data (f_0, P_0) is applied to the problem with the data (f_k, P_k). Is the resulting error small if k is large?

Theorem 5.6. Suppose that the conditions of Theorem 5.2 hold, and let x^* be an optimal solution for the data (f_0, P_0). Suppose also that $f_0(x^*, \cdot)$ is an epi-sublimit of $f_k(x^*, \cdot)$. Then x^* is an approximate solution for the data (f_k, P_k) for k large, namely $val(f_k, P_k) - \gamma_k(x^*)$ converge to 0 as $k \to \infty$.

Proof. The continuity result of Theorem 5.2 together with x^* being an optimal solution imply that

$$val(f_k, P_k) - \gamma_0(x^*) \longrightarrow 0 \ ,$$

and since $\gamma_k(x^*) \leq val(f_k, P_k)$ it follows that

$$\limsup \ \gamma_k(x^*) - \gamma_0(x^*) \leq 0 \ .$$

Applying Theorem 3.5 with the present conditions implies

$$\liminf \ \gamma_k(x^*) - \gamma_0(x^*) \geq 0 \ .$$

Hence $\gamma_k(x^*)$ converges to $\gamma_0(x^*)$ and together with the convergence of $val(f_k, P_k)$ to $\gamma_0(x^*)$, the proof is complete.

We note that a bound for the discontinuity gap, as presented in Proposition 4.5 for integrals, is valid also in the case of the optimization problems.

6. ON THE EXAMPLES

In this section we examine the examples introduced in Section 2, in light of the stability results of the previous section.

Consider first Example 2.1. For the sake of the illustration we consider the cost functional

$$f_0(x,\xi) \ = \begin{array}{ll} r\xi - cx & \text{if} \ \ \xi \leq x \\ (r-c)\xi - K_0 & \text{if} \ \ \xi > x \end{array} \tag{6.1}$$

where ξ is governed by the exponential distribution with parameter λ_0, namely $P_0(d\xi) = \lambda_0 e^{-\lambda_0 \xi} d\xi$. A straightforward computation shows then that the maximization problem

$$\underset{0 \leq x < \infty}{\text{maximize}} \int_0^\infty f_0(x,\xi) \lambda_0 e^{-\lambda_0 \xi} d\xi \tag{6.2}$$

translates into the problem

$$\underset{0 \leq x < \infty}{\text{maximize}} \left(\frac{r}{\lambda_0} - cx - (K_0 + \frac{c}{\lambda_0}) e^{-\lambda_0 x} \right) \ . \tag{6.3}$$

Solving (6.3) is a simple exercise (note that the payoff function is concave on $[0, \infty)$), and the optimal solution is

$$x^* = \frac{1}{\lambda_0} log(1 + \frac{\lambda_0 K_0}{c}) \ . \tag{6.4}$$

We consider now the possibility that the payoff functional (6.1) is only an approximation of the true payoff functional. In view of the discussion in Section 2, consider the payoff functionals

$$(6.5) \qquad f_k(x,\xi) = \begin{cases} r\xi - cx & \text{if } \xi \le x + \epsilon_k(x) \\ (r-c)\, min(\xi, L_k) - K_k & \text{if } \xi > x + \epsilon_k(x) \ . \end{cases}$$

Recall that $\epsilon_k(x)$ reflects the possibility that a number of customers may remain unsatisfied; the deviation K_k from K_0 reflects an error in the cost of the backorder, and L_k reflects the possibility that the backorder is limited. We also wish to consider the possibility that the parameter λ_0 of the underlying probability measure in (6.2) is only an approximation of the true exponent λ_k of the probability distribution. Thus, we consider (6.2) an approximation of

$$(6.6) \qquad \underset{0 \le x < \infty}{\text{maximize}} \int_0^\infty f_k(x,\xi)\lambda_k e^{-\lambda_k \xi} d\xi \ .$$

Theorem 6.1. Suppose that $\lambda_k \to \lambda_0$, $K_k \to K_0$, $L_k \to \infty$ and $\epsilon_k(x)$ converge uniformly to 0. Then the value of (6.6) converges to the value of (6.2). Furthermore, if the optimal solution x^* of (6.2), given in (6.4), is applied to the problem (6.6) with k sufficiently large, then the resulting payoff is a good approximation of the optimal payoff.

Proof. The result follows from Theorem 5.4 and Theorem 5.6, once the conditions of these results are verified. We comment on these conditions. The condition $P_0(Disc\ f_0(x,\cdot)) = 0$ holds since the discontinuity of the payoff function (6.1) occurs on a set of Lebesgue measure zero (at one point actually), and P_0 is absolutely continuous. The equi-tightness follows since $|f_0(x,\cdot)|$ grows linearly in ξ, while $\lambda_k e^{-\lambda_k \xi}$ decays exponentially, uniformly as $\lambda_k \to \lambda_0$. The graph-type convergence, i.e., epi-sublimit and hypo-suplimit, are easy to verify under the convergence conditions. Finally, it is easy to verify that the search for optimal solutions can be restricted to a compact set of possible orders. Checking all that completes the proof.

We examine now Example 2.2. We are interested in the stability of the value for a sequence of problems, determined by the data $D_k = (c_k(x), q_k, r_k, \alpha_k, \beta_k, P_k)$ for $k = 0, 1, 2, \ldots$. Denote the value of the problem (2.3) with D_k instead of D_0 by v_k. We find conditions guaranteeing that $v_k \to v_0$.

Before addressing the stability problem we wish to identify the points of discontinuity of $\Phi_k(|x-\xi|)$; the discontinuities are caused by the structure of the recourse stage. The problem we choose is simple enough to allow the computation of the discontinuities (see Remark 6.3). Indeed, for b fixed, only a finite number of $z \in Z_+$ participate in the minimization (2.4); namely those with $\beta_0 z \le b$. Therefore, for x fixed, the only possible discontinuities are at $|x-\xi| = \beta_k j$ for $j = 0, 1, \ldots$.

We identify the following two ensembles. Let Δ denote the family of probability measures P on $(-\infty, \infty)$ such that $\int_{-\infty}^{\infty} |\xi| P(d\xi) \leq \delta_0$, with δ_0 a prescribed constant. Let Γ denote the family of continuous functions $c(x) : (-\infty, \infty) \to R$, such that $c(x) \leq \kappa_0(x)$, with $\kappa_0(\cdot)$ a prescribed function satisfying $\kappa_0(x) \to -\infty$ as $|x| \to \infty$.

Theorem 6.2. Suppose that $q_k \to q_0$, $r_k \to r_0$, $\alpha_k \to \alpha_0$ and $\beta_k \to \beta_0$. Suppose that all $c_k(\cdot)$ belong to Γ, and that $c_k(x) \to c_0(x)$ uniformly on compact sets. Suppose that all P_k belong to Δ, and that $P_k \to P_0$ weakly. Finally suppose that P_0 is atomless. Then $v_k \to v_0$. Furthermore, if x^* is an optimal solution of (2.3), and x^* is used as the decision for the data D_k, then k large implies that the resulting cost is close to the optimal one.

Proof. The claims follow from Theorems 5.4 and 5.6, once the conditions are verified. We comment on these conditions. The definitions of Δ and Γ imply that the search for an optimal x can be limited to a compact interval. Then the required tightness of P_k follows from the linear bound on the growth of the value in the recourse stage. The functional convergence of $\Phi_k(|x - \xi|)$ is transparent at all continuity points. The structure of discontinuities that we displayed earlier shows that the required functional convergence holds at all points. Finaly, for P_0 atomless, the requirement $P_0(Disc\ \Phi_k(|x - \cdot|))$ holds since the set of discontinuities is discrete. This completes the proof.

Remark 6.3. The preceding example is a much simplified version of the general mixed integer-linear two-stage stochastic program analyzed by Schultz [18]. We thank a referee for pointing out this reference. We added then the example to show how the considerations of Schultz [18] can be incorporated into our framework. Indeed, Schultz verifies the continuity of the value as $P_k \to P_0$ (and P_k belong to a family of a type similar to Δ), when P_0 is absolutely continuous with respect to the Lebesgue measure. This is indeed the right general condition, as in the case with several dimensions analyzed in [18], one can only guarantee that the discontinuities of $\Phi(|x - \cdot|)$ occur in a denumerable number of hyperplanes; in our case these are points, so a condition of being atomless suffices.

7. STABILITY OF SENSORS

Sensors were introduced in [2] as a tool to evaluate the worth of an inquiry into the structure of uncertainty. We recall here the basic idea and definition of the concept (the interested reader can check [2] for the theory and examples). We then verify the stability of sensors with possibly discontinuous objective functions. In fact, given the preceding results in the paper, the stability of sensors follow quite easily.

Consider the maximization problem $(*)$. In many real situations, the decision-maker has an option of acquiring information about the random event ξ before choosing x. This information may not determine ξ completely; rather, a more accurate probability may be obtained. For instance, the newsboy in Example 2.1 may conduct a market research, say by sampling the prospective buyers. The given probability P may be improved by conditioning on the result, say r, of the market research. Let us denote the emerging probability by P_r. Before deciding on the market research (which may involve a cost), the decision-maker knows what possible outcomes P_r may emerge, and he may also know the probability that governs the distribution of the results r, hence he knows the distribution that governs the possible probabilities P_r. This probability on probabilities we call a sensor. Thus a sensor reflects the possible outcomes of a market research, before such research is conducted. But note that different market research procedures may lead to different sensors, and the decision-maker may need to decide which one to follow. To this end we introduce the following.

Definition 7.1. A sensor, say S, is a probability distribution on $\mathcal{P}$. The value of the sensor S, when applied to the problem $(*)$, is

$$val(f,S) = \int_{\mathcal{P}} val(f,P)S(dP) \ . \tag{7.1}$$

The rationale of defining the value of the sensor S as the expectation in (7.1) is as follows. The value of a fixed probability P, given f, is $val(f,P)$. The occurrence of P is governed probabilistically by the sensor S, hence expectation of $val(f,P)$ with respect to the measure S is the value of S to a decision-maker who wishes to maximize expected payoff.

Applying a sensor S may be subject to errors and deviations generated, e.g., by errors in the sampling that determines S. We are interested in the stability of $v(f,S)$ with respect to variations both in f and in S. For convergence in the space of sensors we take the weak convergence of measures. In [2], continuity of the value with respect to S was established for a payoff function f fixed and continuous. Here we use the analysis of the previous section and establish continuity with respect to the pair (f,S), with payoff functionals possibly discontinuous.

We denote by $supp\ S$ the support of S, namely the smallest closed set M of probability measures in $\mathcal{P}$, such that $S(M)=1$.

Theorem 7.2. Suppose that the functions $f_k(x,\xi)$ are uniformly bounded, $k = 0,1,2,\ldots$. Suppose that whenever $x_k \to x_0$, the function $f_0(x_0,\cdot)$ is a hypo-suplimit of $f_k(x_k,\cdot)$, and that for every x_0 a sequence $y_k \to x_0$ exists such that $f_0(x_0,\cdot)$ is an epi-sublimit of $f_k(y_k,\cdot)$. Suppose that S_0 is a sensor such that $P(Disc\ f_0(x_0,\cdot)) = 0$ for every x_0 and every $P \in supp\ S_0$. Then, if S_k converge to S_0, the values $val(f_k,S_k)$ converge to $val(f_0,S_0)$.

Proof. Denote $v_k(P) = val(f_k,P)$, namely, the value of $(*)$ with f_k the payoff function and P the probability distribution. A particular case of Theorem 5.2 implies that $v_0(P)$ is continuous on $supp\ S_0$. The general case of Theorem 5.2 implies that $v_k(P_k)$ converges to $v_0(P_0)$ if P_k converges weakly to P_0. With these properties, the Lebesgue dominated convergence theorem implies

$$\int v_k(P)S_k(dP) \quad \text{converge to} \quad \int v_0(P)S_0(dP) \ .$$

In view of (7.1), this convergence is the desired conclusion and the proof is complete.

The preceding result is stated for f_k uniformly bounded which makes the equi-tightness condition trivial. Unbounded functionals can be incorporated, with the proper conditions of equi-tightness for $(f_k(x, \cdot), P_k)$ whenever $P_k \to P$ and $P \in \mathit{supp}\ S_0$. We leave out the details.

REFERENCES

[1] Z. Artstein, Chattering variational limits of control systems. *Forum Mathematicum*, to appear.

[2] Z. Artstein and R.J-B Wets, Sensors and information in optimization under stochastic uncertainty. *Mathematics of Operations Research*, to appear.

[3] H. Attouch, *Variational Convergence for Functions and Operators*, Pitman, Boston, 1984.

[4] G. Beer, A geometric algorithm for approximating semicontinuous functions. *J. Approximation Theory* 49 (1987), pp. 31-40.

[5] P. Billingsley, *Convergence of Probability Measures.* Wiley, New York, 1968.

[6] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions.* Lecture Notes in Mathematics 580, Springer-Verlag, Berlin, 1977.

[7] J. Dupačová, Stability and sensitivity analysis for stochastic programming. *Annals of Operations Research* 27 (1990), pp. 115-147.

[8] J. Dupačová, On statistical sensitivity analysis in stochastic programming. *Annals of Operations Research* 30 (1991), pp. 199-214.

[9] P. Kall, On approximation and stability in stochastic programming. In *Parametric Optimization and Related Problems*, J. Guddat et al. eds., Mathematische Research Band 35, Akademie-Verlag, Berlin 1987, pp. 387-407.

[10] W.K. Klein Haneveld, L. Stougie and M.H. van der Vlerk, Stochastic integer programming with simple recourse. Research Memorandum no. 455, Institute of Economic Research, Rijksuniversiteit Groningen, 1991.

[11] H.J. Langen, Convergence of dynamic programming models. *Mathematics of Operations Research* 6 (1981), pp. 493-512.

[12] G. Laporte, F. Louveaux and H. Mercure, The vehicle routing problem with stochastic travel times. Technical Report no.96, Faculté Notre-Dame de la Paix, 1989.

[13] S.B. Richmond. *Operations Research for Management Decisions.* Ronald Press Co., New York, 1968.

[14] S.M. Robinson and R.J-B Wets, Stability in two-stage stochastic programming. *SIAM J. on Control and Optimization* 25 (1987), pp. 1407-1416.

[15] R.T. Rockafellar and R.J-B Wets, Variational Analysis, an introduction. In *Multifunctions and Integrands, Stochastic Analysis and Optimization*, G. Salinetti ed., Lecture Notes in Mathematics 1091, Springer-Verlag, Berlin 1984, pp.1-54.

[16] W. Römisch and R. Schultz, Distribution sensitivity in stochastic programming. *Mathematical Programming* 50 (1991), pp. 197-226.

[17] W. Römisch and R. Schultz, Stability analysis for stochastic programs. *Annals of Operations Research* 30 (1991), pp. 241-266.

[18] R. Schultz, Continuity and stabilization in two-stage stochastic integer programming. In *Stochastic Optimization, Numerical Methods and Technical Applications*, K. Marti ed., Lecture Notes in Economics and Mathematical Systems 379, Springer-Verlag, Berlin 1992, pp. 81-92.

[19] R. Schultz, Continuity properties of expectation functionals in stochastic integer programming. *Mathematics of Operations Research*, to appear.

[20] J. Wang. Distribution sensitivity analysis for stochastic programs with complete recourse. *Mathematical Programming* 31 (1985), pp. 286-297.