WORKING PAPER

ON NONNORMAL ASYMPTOTIC BEHAVIOR OF OPTIMAL SOLUTIONS OF STOCHASTIC PROGRAMMING PROBLEMS: THE PARAMETRIC CASE

Jitka Dupačová

March 1988
WP-88-19

IIASA
International Institute
for Applied Systems Analysis

ON NONNORMAL ASYMPTOTIC BEHAVIOR OF OPTIMAL SOLUTIONS OF STOCHASTIC PROGRAMMING PROBLEMS: THE PARAMETRIC CASE

Jitka Dupačová

March 1988
WP-88-19

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

Under incomplete information about the parameters of the true distribution of the random coefficients, the optimal solutions to stochastic programs can be only approximated. This paper extends the previous results of the author to the case when strict complementarity conditions need not to be assumed.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

CONTENTS

ON NONNORMAL ASYMPTOTIC BEHAVIOR OF OPTIMAL SOLUTIONS OF STOCHASTIC PROGRAMMING PROBLEMS: THE PARAMETRIC CASE

Jitka Dupačová

Department of Mathematical Statistics
Charles University, Prague

1. INTRODUCTION

In stochastic programming problems, we are supposed to get optimal solutions of the program

$$\text{minimize } E\{f(x, \xi)\} = \int_{\Xi} f(x, \xi) P(d\xi) \tag{1}$$

on a given set $S \subset R^n$

or, at least, to get their approximation. The later situation appears in cases of incomplete knowledge of the probability measure P, when our decisions are based mostly on sample information only, or it can arise when P has to be approximated in course of numerical procedure, see e.g. Birge and Wets (1986), Kall (1987). Under the both mentioned circumstances, the properties of the approximate solutions are of great interest. Similar problems are treated in statistical estimation theory mostly for small n and under assumption of S open.

In this paper we shall study the asymptotic distribution of the approximate optimal solutions for the case when the incomplete information concerns the parameters of the probability measure P; we refer to Dupačová and Wets (1986, 1987, 1988), King and Rockafellar (1986), King (1986, 1987) for the nonparametric approach.

Let P in (1) be a probability measure that is known to belong to a given parametric family $\{P_y, y \in Y\}$ of probability measures on (Ξ, A), $\Xi \subset R^s$ and $Y \subset R^q$ is a given open set. Denote by η the true, unknown vector parameter and put

$$g(x, y) := E_{P_y}\{f(x, \xi)\} = \int_{\Xi} f(x, \xi) P_y(d\xi) \ . \tag{2}$$

Using this notation, program (1) becomes

$$\text{minimize } g(x, \eta) \text{ on a given set } S \ . \tag{3}$$

Assume that the true parameter vector η has been estimated by y^ν, $\nu = 1, 2, \cdots$ using sample information; the index ν reflects the dependence on the (increasing) sample size. Accordingly, we use the optimal solution

$$x(y^\nu) \in \arg\min_{x \in S} g(x, y^\nu) \tag{4}$$

to estimate (or to approximate) the true optimal solution

$$x(\eta) \in \arg\min_{x \in S} g(x, \eta) \tag{5}$$

of the stochastic program (3) and the value of $\min_{x \in S} g(x, y^\nu) = g(x(y^\nu), y^\nu)$ to estimate (or to approximate) the optimal value of the objective function $g(x, \eta)$ in (3)).

In a common situation, y^ν enjoys "good" properties such as consistency, asymptotic normality, asymptotic efficiency. The question is how far are these properties inherited by $x(y^\nu)$ and by $\min_{x \in S} g(x, y^\nu)$. The relatively easy case is connected with differentiability property of the optimal solutions $x(y) \in \arg\min_{x \in S} g(x, y)$ and of the optimal value $g(x(y), y) = \min_{x \in S} g(x, y)$ for y belonging to a neighborhood of η: If $x(y)$ (resp. $g(x(y), y)$) is differentiable on $O(\eta)$ and if y^ν is asymptotically normal,

$$\sqrt{\nu}(y^\nu - \eta) \sim N(0, \Sigma) \ ,$$

then the well known results on smooth transformations of multinormal variables (see e.g. Serfling 1980) can be used to get asymptotic normality of $x(y^\nu)$ (resp. of $g(x(y^\nu), y^\nu)$) – see e.g. Dupačová (1984), (1987), Shapiro (1985).

We shall use results of parametric programming to get asymptotic behavior of $x(y^\nu)$ and of $g(x(y^\nu), y^\nu)$. We shall concentrate to the results that are devoted to the differentiability property and we shall summarize them for S defined by explicitely given constraints

$$S(y) = \{x \in R^n : g_i(x, y) \geq 0, i = 1, \ldots, m, g_i(x, y) = 0, i = m+1, \ldots, m+p\} \ . \tag{6}$$

For $y \in Y$ consider the parametric program

$$\text{minimize } g_0(x, y) \text{ on the set } S(y) \ . \qquad P(y)$$

Let

$$L(x, u, y) = g_0(x, y) + \sum_{i=1}^{m+p} u_i g_i(x, y)$$

be the corresponding Lagrange function defined on $R^n \times R_+^m \times R^p$ and denote $u(y)$ the vector of Lagrange multipliers that correspond to the optimal solution $x(y)$ of $P(y)$, i.e.,

$$\begin{aligned} &\nabla_x L(x(y), u(y), y) = 0 \\ &x(y) \in S(y),\ u_i(y) \geq 0,\ 1 = 1, \ldots, m \ , \\ &u_i(y) g_i(x(y), y) = 0,\ i = 1, \ldots, m \end{aligned} \tag{7}$$

holds true for the pair $[x(y), u(y)]$. We assume that $P(y)$ has an optimal solution for the true parameter η. The basic method for analyzing $P(y)$ can be found already in Fiacco and Mc Cormick (1968). It uses the following assumptions:

A1 - Differentiability

The functions $g(x, y)$, $i = 0, 1, \ldots, m + p$ are twice continuously differentiable with respect to (x, y) on a neighborhood of $[x(\eta), \eta]$.

A2 - Linear independence condition

Let $I(\eta) \subset \{1, \ldots, m\}$ contains indices of active inequality constraints, i.e.,

$$g_i(x(\eta), \eta) = 0,\ i \in I(\eta) \ ,$$

then

$$\nabla_x g_i(x(\eta), \eta),\ i \in I(\eta),\ \nabla_x g_i(x(\eta), \eta),\ i = m + 1, \ldots, m + p$$

are linearly independent.

A3 - Strict complementarity conditions

For $i = 1, \ldots, m$,

$$u_i(\eta) = 0 \Longleftrightarrow i \notin I(\eta) \ .$$

A4 - The second order sufficient condition

The inequality

$$z^T \nabla_{xx}^2 L(x(\eta), u(\eta), \eta) z > 0 \tag{8}$$

holds true for each $z \neq 0$ such that

$$z^T \nabla_x g_i(x(\eta), \eta) = 0, \; i = m+1, \ldots, m+p \ ,$$

$$z^T \nabla_x g_i(x(\eta), \eta) = 0 \ \ \forall i \in I(\eta) \text{ for which } u_i(\eta) > 0 \ ,$$

$$z^T \nabla_x g_i(x(\eta), \eta) \geq 0 \ \ \forall i \in I(\eta) \text{ for which } u_i(\eta) = 0 \ .$$

Assertion 1 (Fiacco 1976, 1983) Under assumptions A1–A4

a) For all y in a neighborhood of η, there is a unique optimal solution $x(y)$ of $P(y)$ and a unique vector of Lagrange multipliers $u(y)$ such that (6) holds true.

b) The functions $x(y)$ and $u(y)$ are continuously differentiable at η.

c) The optimal value function $g_0(x(y), y)$ is twice continuously differentiable at η.

The differentiability property b) depends heavily on the assumed strict complementarity conditions A3. At the same time, A3 is related to the "true" program $P(\eta)$ and in our context, it can be hardly fully verified. In optimization problems of mathematical statistics, assumptions A2 and A3 are mostly respected by assuming that S is open or that $x(\eta)$ is an interior point of S. (It means, that no constraints are taken into account.)

If the assumptions of strict complementarity are dropped, one uses the strengthened form of the second order sufficient condition introduced by Robinson (1980):

A5 - The strong second order sufficient condition

The inequality (8) holds true for each $z \neq 0$ such that

$$z^T \nabla_x g_i(x(\eta), \eta) = 0, \; i = m+1, \ldots, m+p \ ,$$

$$z^T \nabla_x g_i(x(\eta), \eta) = 0 \ \ i \in I(\eta) \text{ for which } u_i(\eta) > 0 \ .$$

Assertion 2 (Robinson 1980) Under assumptions A1, A2, A5, the optimal solution $x(y)$ and the vector of Lagrange multipliers exist, are unique and Lipschitz continuous on a neighborhood of η.

Assertion 3 (Jittorntrum 1984) Under assumptions A1, A2, A5, the optimal solution $x(y)$ and the vector of Lagrange multipliers are directionally differentiable at the point η.

Assertion 4 (Rockafellar 1984) Under assumptions A1, A2, A5, the optimal value function $g_0(x(y), y)$ is continuously differentiable at η.

Very often, the differentiability assumption A1 can be weakened slightly by dropping the assumption on the existence of the second order derivatives with respect to y. An essential relaxation makes use of Lipschitz continuity of the first order derivatives only. In this case, it is possible to prove Lipschitz continuity of the optimal solution; see e.g. Robinson (1974) and in the context of stochastic programming Wang (1985). As to the linear independence assumption A2, it can be replaced by Mangasarian–Fromowitz constraint qualification, see e.g. Fiacco and Kyparisis (1985) and Shapiro (1985).

We shall follow the detailed analysis by Robinson (1986) according to which the desired differentiability property of the optimal solutions of $P(y)$ cannot hold true without the strict complementarity conditions A3, whereas in the general case we have only

Assertion 5 (Robinson 1986) Under assumptions A1, A2, A5, the optimal solutions $x(y)$ of $P(y)$ are Bouligand differentiable at η.

2. BOULIGAND DIFFERENTIABILITY OF OPTIMAL SOLUTIONS

We shall apply the results by Robinson (1986) for to get asymptotic behavior of $x(y^\nu)$ under assumptions A1, A2, A5. To simplify the explanations we shall concentrate to the case when no explicite constraints are spelled out in $P(y)$; it is fully in line with our original problem (1). Moreover, the reasonings of Robinson (1984) show that in local stability studies of the parametric program $P(y)$ one can get rid of the explicite constraints provided that the assumption A2 (or the more general nondegeneracy assumption) holds true. We shall briefly delineate Robinson's approach for the parametric program

$$\text{minimize } g(x, y) \text{ on a set } S \tag{9}$$

where $g: R^n \times R^q \to R^1$ and $S \subset R^n$ under the following assumptions:

B1 – There exist continuous derivatives

$$\frac{\partial^2 g(x, y)}{\partial x_i \partial y_k}, \frac{\partial^2 g(x, y)}{\partial x_i \partial x_j} \text{ for } i, j = 1, \ldots, n, k = 1, \ldots, q$$

on a neighborhood of $[x(\eta), \eta]$ and the matrix

$$\left[\frac{\partial^2 g(x(\eta), \eta)}{\partial x_i \partial x_j}\right]_{1 \leq i, j \leq n}$$

is positive definite.

B2 – The set S is convex, polyhedral with int $S \neq \emptyset$.

To avoid trivial situations, we shall assume that program (9) has an optimal solution $x(\eta)$ for $y = \eta$ and that $x(\eta) \notin \text{int } S$.

Under assumption B2, the set $S - x(\eta)$ can be replaced near the origin by its tangent cone T at the origin. It means that the optimality condition

$$0 \in \nabla_x g(x(\eta), \eta) + N_S(x(\eta)) \tag{10}$$

for the "true" program

$$\text{minimize } g(x, \eta) \text{ on the set } S \tag{11}$$

can be replaced by

$$0 \in \nabla_x G(0, \eta) + N_T(0) \tag{12}$$

where

$$G(x, y) = g(x - x(\eta), y)$$

and $N_T(z_0)$ denotes the normal cone to T at the point z_0, i.e., $N_T(z_0) = \{\mu : (z - z_0)^T \mu \leq 0 \ \forall z \in T\}$.

Consider now the generalized equation

$$0 \in \nabla_x G(x, y) + N_T(x) \tag{13}$$

that expresses the necessary optimality condition for the program

$$\text{minimize } G(x, y) \text{ on } T \ . \tag{14}$$

To get a solution of (13), we use the linearization technique of Robinson (1980) that leads to the *linear* generalized equation

$$0 \in \nabla_x G(0, y) + \nabla^2_{xx} G(0, \eta) x + N_T(x) \tag{15}$$

whose solution for y near η is near to the solution of the nonlinear generalized equation (13). (See Robinson 1980, Theor. 2.3).

For small perturbations $y - \eta$ we have approximately

$$\nabla_x G(0, y) \doteq \nabla_x G(0, \eta) + \nabla^2_{xy} G(0, \eta)(y - \eta) \ .$$

The solution of (13) can be thus further approximated by the solution of the generalized equation

$$w \in \nabla_x G(0, \eta) + \nabla^2_{xx} G(0, \eta)x + N_T(x) \tag{16}$$

with

$$w = - \nabla^2_{xy} G(0, \eta)(y - \eta) \ , \tag{17}$$

or, equivalently, by the solution of the quadratic program

$$\text{minimize } x^T(\nabla_x G(0, \eta) - w) + \frac{1}{2} x^T \nabla^2_{xx} G(0, \eta)x \tag{18}$$

on T.

Thanks to B1, the quadratic program (18) has a unique, locally Lipschitzian solution $X(w)$ that equals for w near to 0 (i.e., for y near to η) to the Bouligand derivative $\Delta x(\eta; y - \eta)$ of the optimal solution of (9) at the true parameter value η applied to $y - \eta$ (see Robinson 1986, Theor. 3.5). It means that

$$\begin{aligned} x(y) - x(\eta) &= \Delta x(\eta; y - \eta) + o(y - \eta) \\ &= X(w) + o(y - \eta) \end{aligned} \tag{19}$$

where w is defined by (17), what explains the word "approximated" used in connection with solutions of (16) or (18) and their relationship to the solutions of generalized equations (13) or (15) for y near to η.

The necessary and sufficient condition for to get affine Bouligand derivative (i.e., to get differentiability of $x(y)$ at the point $y = \eta$) reads

$$- \nabla_x g(x(\eta), \eta) \in \text{rint } N_S(x(\eta)) \ . \tag{20}$$

For a polyhedral set S defined through explicitely given (linear) constraints, condition (20) is fulfilled if and only if the corresponding strict complementarity conditions A3 hold true. (Robinson 1986). In the context of our paper it means that in the general case, the asymptotic distribution of $\sqrt{\nu}[x(y^\nu) - x(\eta)]$ need not be normal in spite of asymptotic normality of $\sqrt{\nu}(y^\nu - \eta)$. We have only

THEOREM 1 *Let assumptions B1 and B2 hold true for the program $P(y)$ and let y^ν be an asymptotically normal estimate of η, i.e.,*

$$\sqrt{\nu}(y^\nu - \eta) \sim N(0, \Sigma) \ . \tag{21}$$

Then the asymptotic distribution of $\sqrt{\nu}(x(y^\nu) - x(\eta))$ is that of $\sqrt{\nu}X(w^\nu)$, where $X(w^\nu)$ is the optimal solution of the quadratic program (18) corresponding to an asymptotically normal perturbation $w^\nu = - \nabla^2_{xy} G(0, \eta)(y^\nu - \eta)$ in the linear part of the objective function.

3. THE PERTURBED QUADRATIC PROGRAM

To develop the asymptotic result in detail we have to study the perturbed quadratic program (18). We shall modify the results of Guddat (1976) to our case. To simplify the notation, put

$$C = \nabla^2_{xx} G(0, \eta),\ B = \nabla^2_{xy} G(0, \eta) \text{ and } p = \nabla_x G(0, \eta) \quad . \tag{22}$$

Assume that $x(\eta) \notin \operatorname{int} S$ so that the tangent cone $T \neq R^n$. It can be written as

$$T = \{x \in R^n : Ax = 0,\ x \geq 0\} \text{ where } A \text{ is an } (m, n) \text{ matrix} \quad .$$

According to B1, the quadratic program

$$\begin{aligned} &\text{minimize } (p - w)^T x + \frac{1}{2} x^T C x \\ &\text{subject to } Ax = 0,\ x \geq 0 \end{aligned} \tag{23}$$

that corresponds to (18) has a unique optimal solution $X(w)$ for an arbitrary $w \in R^n$ and $X(w)$ is a Lipschitz continuous vector function on R^n.

The set T can be decomposed into its vertex (if any), its interior and into finitely many relatively open faces of T. Each face, say Σ, is determined by a subset $J \subset \{1, \ldots, n\}$ in the following way:

$$\Sigma(J) = \{x \in R^n : Ax = 0,\ x_j = 0 \text{ for } j \in J \text{ and } x_j > 0 \text{ for } j \notin J\} \quad . \tag{24}$$

To each of the faces, the *stability set* $\sigma(J)$ can be constructed. It is by definition the set of all parameter vectors $p - w$ for which the optimal solution $X(w)$ of (23) belongs to $\Sigma(J)$. According to Guddat (1976), the stability sets form a decomposition of the parameter space R^n for which, inter alia, the following properties hold true:

(i) For any of subsets $J \subset \{1, \ldots, n\}$, the function $X(w)$ is linear on clo $\sigma(J)$.

(ii) If $\Sigma(J)$ is the vertex of T, then $X(w)$ is constant on clo $\sigma(J)$.

Let us compute the optimal solution $X(w)$ of (23) directly assuming that $p - w \in \sigma(J)$. The necessary and sufficient conditions for $x \in \Sigma(J)$ to be the optimal solution of (23) can be written as

$$\begin{aligned} &Ax = 0,\ x_j > 0 \text{ for } j \notin J,\ x_j = 0 \text{ for } j \in J \\ &Cx + A^T u - v = w - p,\ v_j \geq 0 \text{ for } j \in J,\ v_j = 0 \text{ for } j \notin J \quad . \end{aligned} \tag{25}$$

Through conditions (25), the set $\sigma(J)$ is defined.

Choose now an arbitrary index set $J \subset \{1,\ldots,n\}$ and consider the system of equations

$$Ax = 0,\ x_j = 0 \text{ for } j \in J \ . \tag{26}$$

Let k be the rank of the matrix of the system (26), i.e., the rank of

$$\begin{bmatrix} A \\ I_J \end{bmatrix}$$

where I_J denotes the reduced n dimensional identity matrix I_n that contains only the rows corresponding to indices $j \in J$. It means that the system (25) can be equivalently written as

$$\tilde{x} = A_J y,\ y \in R^{n-k},\ \tilde{x} \in R^k \text{ with a } (k,\ n-k) \text{ matrix } A_J \ .$$

Assume that $\tilde{x}$ consists of first k components of x, $\tilde{x} = (x_1, \ldots, x_k)^T$ and substitute

$$x = \begin{bmatrix} A_J y \\ y \end{bmatrix} = \begin{bmatrix} A_J \\ I_{n-k} \end{bmatrix} y \tag{27}$$

into the objective function of (23):

$$(p - w)^T \begin{bmatrix} A_J y \\ y \end{bmatrix} + \frac{1}{2} y^T (A_J^T \vdots I_{n-k})\, C \begin{bmatrix} A_J \\ I_{n-k} \end{bmatrix} y \ .$$

Denote

$$p_J = (A_J^T \vdots I_{n-k})(p - w),\ C_J = (A_J^T \vdots I_{n-k}) C \begin{bmatrix} A_J \\ I_{n-k} \end{bmatrix}$$

and instead of minimizing the original objective function of (23) subject to (26), solve the ***unconstrained*** quadratic program

$$\text{minimize } p_J^T y + \frac{1}{2} y^T C_J y \ . \tag{28}$$

Evidently, C_J is positive definite again, so that there is a unique optimal solution of (28) for an arbitrary p_J, namely,

$$y(p_J) = -\ C_J^{-1} p_J \ .$$

Accordingly, the optimal solution $\hat{x}(w)$ of the program

$$\text{minimize } (p - w)^T x + \frac{1}{2} x^T C x$$

$$\text{subject to } Ax = 0,\ x_j = 0 \text{ for } j \in J \tag{29}$$

has the form (see (27))

$$\begin{aligned} \hat{x}(w) &= \begin{bmatrix} A_J \\ I_{n-k} \end{bmatrix} y(p_J) = - \begin{bmatrix} A_J \\ I_{n-k} \end{bmatrix} C_J^{-1}(A_J^T : I_{n-k})(p - w) \\ &= R_J w + r_J \ . \end{aligned} \tag{30}$$

Formula (30) gives the general form of optimal solutions of (29) for an arbitrary $w \in R^n$. According to the notion of stability sets σ, (30) is the general form of optimal solutions of the program

$$\text{minimize } (p - w)^T x + \frac{1}{2} x^T C x \tag{31}$$

on the set

$$\Sigma(J) = \{x \in R^n : Ax = 0,\ x_j = 0 \text{ for } j \in J,\ x_j > 0 \text{ for } j \notin J\}$$

for all $p - w \in \text{clo}\, \sigma(J)$. As a result, we have

THEOREM 2 *The optimal solution $X(w)$ of (28) is a piecewise linear continuous vector function on R^n. For $p - w \in \text{clo}\, \sigma(J)$, it has the form (30). Moreover, for all $J \subset \{1, \ldots, n\}$ $X(w)$ is differentiable if $p - w \in \text{int}\, \sigma(J)$.*

4. THE ASYMPTOTIC DISTRIBUTION

Let y^ν, $\nu = 1, \cdots$, be asymptotically normal estimates of the true parameter vector η, i.e.,

$$\sqrt{\nu}(y^\nu - \eta) \sim N(0, \Sigma) \ . \tag{32}$$

We shall use the results of Sections 2 and 3 to get the asymptotic distribution of the approximate optimal solutions $x(y^\nu)$ of the "true" program

$$\text{minimize } g(x, \eta)$$

on a given convex polyhedral set $S \subset R^n$

under assumption B1 only.

As we already know (see 19), $\sqrt{\nu}[x(y^\nu) - x(\eta)]$ is asymptotically equivalent to $\sqrt{\nu}X(w^\nu)$, where $X(w^\nu)$ denotes the (unique) optimal solution of the quadratic program

$$\begin{aligned} &\text{minimize } (p - w^\nu)^T x + \frac{1}{2} x^T C x \\ &\text{subject to } Ax = 0,\ x \geq 0 \end{aligned} \tag{33}$$

with $w^\nu = - B(y^\nu - \eta)$.

According to (30), $X(w^\nu)$ can be written for $p - w^\nu \in \text{clo}\ \sigma(J)$ as

$$X(w^\nu) = R_J w^\nu + r_J = - R_J B(y^\nu - \eta) + r_J \ . \tag{34}$$

We are going to use $X(w^\nu)$ for to approximate $x(\eta)$ in case that y^ν is near to η, i.e., w^ν is near to 0. Our reference point is thus the optimal solution of the quadratic program

$$\begin{aligned} &\text{minimize } p^T x + \frac{1}{2} x^T C x \\ &\text{subject to } Ax = 0,\ x \geq 0 \ . \end{aligned} \tag{35}$$

Let us distinguish two cases:

(i) $p \in \text{int}\ \sigma(J)$ for an index set $J \subset \{1,\ldots,n\}$. Then there is a neighborhood O of zero such that for $w \in O$, $p - w \in \text{int}\ \sigma(J)$ holds true. In this case, X is differentiable at 0 and

$$\sqrt{\nu}(x(y^\nu) - x(\eta)) \sim N(0, R_J B \Sigma B^T R_J^T) \tag{36}$$

Notice that in this case, strict complementarity conditions are fulfilled for (35).

(ii) $p \in \text{bound}\ \sigma(J)$ for an index set $J \subset \{1,\ldots,n\}$. It means that $p \in \text{bound}\ \sigma(J')$ for $J' \neq J$, $J' \subset \{1,\ldots,n\}$, too. Let

$$p \in \bigcap_{h=1}^{H} \text{bound}\ \sigma(J_h) \ .$$

Then for each of index sets J_h, $h = 1,\ldots,H$, there is a different representation (34) of the optimal solutions. For w near to 0,

$$p - w \in \bigcup_{h=1}^{H} \text{clo}\ \sigma(J_h)$$

and we can compute the distribution function according to

$$P\{X(w) \leq a\} = \sum_{h=1}^{H} P\{X(w) \leq a | p - w \in \sigma(J_h)\} P\{p - w \in \sigma(J_h)\}$$

$$= \sum_{h=1}^{H} P\{R_{J_h} w + r_{J_h} \leq a | p - w \in \sigma(J_h)\} P\{p - w \in \sigma(J_h)\} \quad . \tag{37}$$

For w^ν asymptotically normal we thus get an asymptotic distribution of $X(w^\nu)$ that is a *mixture of normal distributions conditioned by convex polyhedral sets*. Notice that one of these distributions can be degenerated if the optimal solution of (35) is in the vertex of the cone $T = \{x \in R^n : Ax = 0, x \geq 0\}$.

5. DISCUSSION

Our result is in full agreement with that given by King and Rockafellar (1986) and King (1986) for linear-quadratic stochastic programs and nonparametric approach. Moreover, it indicates that their special assumption about the objective function might be relaxed.

As to the asymptotic behavior of the optimal value function $g(x(y^\nu), y^\nu)$, the situation is much simpler. Thanks to Assertion 4 of Section 1, it is asymptotically normal for asymptotically normal estimates y^ν of η and, on the top of it one can get additional results on its bias, see Shapiro (1985).

One can consider the probability measure P to be a parameter in program (1); for continuity results on the optimal value function and on the optimal solution set see Kall (1987), Robinson and Wets (1987), Römisch and Schulz (1987).

REFERENCES

Birge, J. and R. Wets (1986): Designing approximation schemes for stochastic problems, in particular for stochastic progarms with recourse. Math. Programming Study 27, 54–102.

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

Dupačová, J. (1987): On some connection between parametric and stochastic programming. In: Guddat J. et al (eds): Parametric optimization and Related Topics. Math. Research Band 35, Akademie-Verlag, Berlin, 74–81.

Dupačová, J. and R. J-B Wets (1986): Asymptotic behavior of statistical estimators and optimal solutions for stochastic optimization problems. IIASA WP-86-41, Laxenburg, Austria.

Dupačová, J. and R. J-B Wets (1987): Asymptotic behavior of statistical estimators and of optimal solutins of stochastic optimization problems, II. IIASA WP-87-9, Laxenburg, Austria.

Dupačovaá, J. and R. J-B Wets (1988): Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. Accepted for Annals of Statistics.

Fiacco, A.V. (1976): Sensitivity analysis in nonlinear programming using penalty methods, Math. Programming 10, 287–311.

Fiacco, A.V. (1983): Introduction to sensitivity and stability analysis in nonlinear programming. Academic Press, New York.

Fiacco, A.V. and G.P. Mc Cormick (1968): Nonlinear programming: Sequential unconstrained minimization techniques. Wiley, New York.

Fiacco, A.V. and J. Kyparisis (1985): Sensitivity analysis in nonlinear programming under second order assumptions. In: Bachem, A. and H. Th. Jongen (eds): Systems and optimization. Springer Verlag, Berlin, 74–97.

Guddat, J. (1976): Stability in convex quadratic parametric programming. Math. Operationsforsch. u. Statist. 7, 223–245.

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

Kall, P. (1987): On approximations and stability in stochastic programming. In: Guddat, J. et al. (eds): Parametric optimization and related topics. Math. Research Band 35, Akademic-Verlag Berlin, 387–407.

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

King, A.J. (1987): Central limit theory for Lipschitz mappings. IIASA, WP-87-127, Laxenburg, Austria.

King, A.J. and R.T. Rockafellar: Non-normal asymptotic behaviour of solution estimates in linear-quadratic stochastic optimization. Lecture presented at the Int. Conference on Stochastic Programming, Prague, Sept. 14–19.

Robinson, S.M. (1974): Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear-programming algorithms. Mathematical Programming 7, 1–16.

Robinson, S.M. (1980): Strongly regular generalized equations. Math. of Operations Research 5, 43–63.

Robinson, S.M. (1984): Local structure of feasible sets in nonlinear programming. Part II: Nondegeneracy. Mathematical Programming Study 22, 217–230.

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

Robinson, S.M. and R. J-B Wets (1987): Stability in two-stage stochastic programming . SIAMJ. Control and Optimization 25, 1409–1416.

Rockafellar, R.T. (1984): Directional differentiability of the optimal value function in a non-linear programming problem. Mathematical Programming Study 21, 213–226.

Römisch, W. and R. Schultz (1987): Distribution sensitivity in stochastic programming. WP160, Sektion Mathematik, Humboldt University, Berlin; presented in parts at the Int. Conference on Stochastic Programming, Prague, Sept. 14–19.

Serfling, R.J. (1980): Approximation theorems in mathematical statistics, Wiley.

Shapiro, A. (1985): Second order sensitivity analysis and asymptotic theory of parametrized nonlinear programs. Mathematical Programming 33, 280–299.

Wang, J. (1985): Distribution sensitivity analysis for stochastic programs with complete recourse. Mathematical Programming 31, 286–297.