The statistical method of empirical means is applied to solve the general stochastic programming problem with compound risk functions. The convergence of the method is established in probabilistic terms.

Optimization problems under uncertainty often occur in practice. They are formalized and analyzed in stochastic programming theory. Stochastic optimization problems differ from nonlinear programming in that they contain functions of a special form (risk functions) that are Lebesgue integrals (expectations) of the random parameters of the problem. Various methods are available for solving such problems (see Ermoliev and Wets [1]). One of the solution methods is traceable to a common statistical approach. Assume that several independent identically distributed observations of the random parameters of the stochastic programming problem are given, i.e., some statistical data are available about the parameters. Then the expectations of these parameters can be approximately replaced with the empirical means. Thus, the original problem is replaced with an approximate problem, whose solution is a statistical estimator of the solution of the original problem. This approach gives rise to natural questions of consistency, asymptotic normality, and rate of convergence of the estimators (depending on the number of observations).

This approach to the solution of stochastic programming problems will be called the empirical mean method or the statistical method. It is attractively simple, permits using standard mathematical software for problem solving, and also allows the empirical means and their gradients to be computed concurrently by parallelized processing.

Note that many statistical problems (e.g., the estimation of the parameters of statistical models by least squares, maximum likelihood, or robust methods, see Huber [2]) are essentially stochastic programming problems of a special kind and the empirical mean method is traditionally applied to these problems as a matter of course. Consistency, efficiency, and the asymptotic properties of the empirical mean method have been studied in statistics. For instance, Korkhin [3] has established the limiting (nonnormal) distribution for the estimators of the regression parameters in the presence of prior inequality constraints on the parameters.

The empirical mean method began to be used for solving stochastic programming problems by Yubi [4], Tamm [5], Kankova [6], Dupačová [7], and Wets [8]. Stochastic programming problems are characterized by inequality constraints, nonsmooth and even discontinuous risk functions, and nonuniqueness of optimal solutions. The consistency of the empirical mean method under general conditions has been investigated by Salinetti and Wets [9], Dupačová and Wets [10], King and Wets [11], Knopov [12], and Norkin [13]. The asymptotic properties of the method (the limiting properties of the estimators and the rate of convergence) have been studied by Vapnik [14], King [15], Shapiro [16, 17], Ermoliev and Norkin [18], and Norkin [13].

King [15] and Shapiro [17] applied the $\delta$-method of generating limit theorems for transformed random variables to study the empirical mean method. For nonsmooth (Hadamard-differentiable) transformations, the $\delta$-method has been substantiated by Reeds [19], Fernholz [20], Grubel [21], and Gill [22]. The rate of convergence of the empirical mean method has been studied in [18] and [13] on the basis of a new concept of normalized convergence of random variables, investigated in detail by Ermol’ev and Norkin [23].

In this paper, we generalize the results of [18, 13, 17] by focusing on a more general stochastic programming problem with compound risk functions. The convergence of the empirical mean method is interpreted as a consequence of stability of
some functional parametric programming problem. The rate of convergence of the method is examined using the notion of normalized convergence of random variables.

**STATEMENT OF THE PROBLEM**

Consider the general stochastic programming problem [24]

\[ F_0(x) = E\Phi_0(x, \theta_0, \ldots, \theta_m(x, \theta), \theta) \rightarrow \min, \]

\[ F_i(x) = E\Phi_i(x, \theta_i, \ldots, \theta_m(x, \theta), \theta) \leq 0, \quad x \in X, \quad i \in I = \{1, \ldots, l\}, \tag{1} \]

where

\[ \Phi_j(x) = E\Psi_j(x, \theta) = \int_\Theta \Psi_j(x, \theta) P(d\theta), \quad j \in J = \{1, \ldots, m\}, \]

\[ g_i(x, y) = \Phi_i(x, y, \theta) = \int_\Theta \Phi_i(x, y, \theta) P(d\theta), \quad y \in Y \subseteq \mathbb{R}^m, \quad i \in I, \tag{2} \]

\[ X \text{ is a compactum in some topological space, } x \text{ is a deterministic variable, } Y \text{ is a set in } \mathbb{R}^m, \theta \in \Theta, (\Theta, \Sigma, P) \text{ is some probability space, } \varphi_j: X \times \Theta \to \mathbb{R}^l \text{ and } \Phi_j: X \times Y \times \Theta \to \mathbb{R}^l \text{ are functions integrable for each } x \in X \text{ and } y \in Y, i \in I \cup 0, j \in J \cup 0. \]

In particular cases, we may have

\[ F_0(x) = E\Psi_0(x, \theta), \quad F_i(x) = E\Psi_i(x, \theta) \rightarrow c_i(x) = E\Psi_i(x, \theta) \rightarrow \]

\[ - (E\Psi_i(x, \theta))^2 \rightarrow c_i(x), \quad F_i(x) = \max_{j \in J_i} E\Psi_j(x, \theta) \text{ etc.}. \]

The empirical mean method (or the statistical method) for solving the problem (1) can be described as follows. The problem (1) is replaced with a sequence of problems of the form

\[ F_0^*(x, \omega) = \frac{1}{S} \sum_{k=1}^S \Phi_0\left(x, \frac{1}{S} \sum_{h=1}^S \varphi_0(x, \theta_h), \ldots, \frac{1}{S} \sum_{h=1}^S \varphi_m(x, \theta_h), \theta_0\right) \rightarrow \min, \]

\[ F_i^*(x, \omega) = \frac{1}{S} \sum_{k=1}^S \Phi_i\left(x, \frac{1}{S} \sum_{h=1}^S \varphi_0(x, \theta_h), \ldots, \frac{1}{S} \sum_{h=1}^S \varphi_m(x, \theta_h), \theta_i\right) \]

\[ \ldots, \frac{1}{S} \sum_{h=1}^S \varphi_m(x, \theta_h), \theta_k \right) \leq \theta, \quad i \in I, \quad x \in X, \tag{3} \]

where \( \theta_k \) are independent identically distributed random variables (observations), whose distribution is determined by the measure \( P \).

The solution of the problem (3) depends on the vector random variable \( \omega \equiv (\theta_1, \ldots, \theta_s) \). We assume that all \( \theta^s \) and \( \theta_k \) are defined on the same probability space \( (\Omega, \Sigma_\omega, P_\omega) \) — a direct countable product of instances of the original space \( (\Theta, \Sigma, P) \). Elementary outcomes in \( (\Omega, \Sigma_\omega, P_\omega) \) are sequences \( \omega = (\theta_1, \theta_2, \ldots) \) of elementary outcomes in \( \Theta \). We assume that \( \theta_k(\omega) \) are independent identically distributed random variables in \( (\Omega, \Sigma_\omega, P_\omega) \). The solution of the problem (3) thus depends on \( \omega \).

Denote by \( F_0^*(\omega), F_i^*(\omega) \) and \( X_0^*(\omega), X_i^*(\omega) \) the optimal values and the optimal solutions of the problems (1), (3), and by

\[ X_0^* = \{ x \in X | F_0(x) \leq F_0^* + \epsilon, \quad i \in I \}, \]

\[ X_i^* = \{ x \in X | F_i(x, \omega) \leq F_i^*(\omega) + \epsilon, \quad i \in I \}, \]

the approximate solutions of the problems (1), (3).

Our goal is to investigate the convergence, in some probabilistic sense, of the random variables \( F_i^*(\omega), X_i^*(\omega), X_\omega^* \) to the corresponding values \( F_i^*, X_i^*, X_\omega^* \).