Mathematical Programming Study 9 (1978) pp. 43–58

RELIABILITY TYPE INVENTORY MODELS BASED ON STOCHASTIC PROGRAMMING

A. Prékopa
Technological University of Budapest and Computer and Automation Institute of the Hungarian Academy of Sciences, Budapest, Hungary

P. Kelle
Computer and Automation Institute of the Hungarian Academy of Sciences, Budapest, Hungary

Received: 4 April 1977
Revised manuscript received 12 October 1977

Abstract

The models discussed in the present paper are generalizations of the models introduced previously by A. Prékopa [6] and M. Ziermann [13]. In the mentioned papers the initial stock level of one basic commodity is determined provided that the delivery and demand process allow certain homogeneity (in time) assumptions if they are random. Here we are dealing with more than one basic commodity and drop the time homogeneity assumption. Only the delivery processes will be assumed to be random. They will be supposed to be stochastically independent. The first model discussed in this paper was already introduced in [9]. All these models are stochastic programming models and algorithms are used to determine the initial stock levels rather than simple formulas. We have to solve nonlinear programming problems where one of the constraints is probabilistic. The function and gradient values of the corresponding constraining function are determined by simulation. A numerical example is detailed.

Keywords: Stochastic Programming, Reliability Models, Inventory Models, Chance Constrained Programming, Application in Economics.

1 Introduction

The models discussed in the present paper are generalizations of the models introduced previously by A. Prékopa [6] and M. Ziermann [13]. In those papers the initial stock level of one basic commodity is determined, provided that the delivery and demand process allow certain homogeneity (in time) assumptions if they are random. Here we are dealing with more than one basic commodity and drop the time homogeneity assumption. Only the delivery processes will assume to be random. They will be supposed to be stochastically independent. Out of the models discussed in this paper the first one was already

introduced in [9]. All these models are stochastic programming models and algorithms serve in determining the initial stock levels instead of simple formulas. We have to solve nonlinear programming problems where one of the constraints is probabilistic. The function and gradient values of the corresponding constraining function are determined by simulation.

The numerical computation for these models (as introduced in the present paper) is more sophisticated than that required for the earlier models of Prékopa and Ziermann; however, if the delivery process is inhomogeneous then with the present methodology we can get closer to reality and can handle many delivery processes simultaneously.

The most general model introduced in [6] is the following. Let M denote the initial stock level, $(0,T)$ the investigated time interval, α_t the amount of the basic commodity delivered up to time t and β_t the cumulative demand at time t, where $0 \le t \le T$. The initial stock level is the smallest M satisfying

$$\mathbf{P} = \left(\inf_{0 \le t \le T} (M + \alpha_t - \beta_t) > 0 \right) \ge 1 - \epsilon, \tag{1.1}$$

where ϵ is a prescribed, small positive number; e.g., $\epsilon = 0.05$. Under the assumptions introduced in [6] in connection with the random processes α_t, β_t, (1.1) holds with equality for the optimal initial stock. Thus an equation can be used to determine M. This is called the *reliability equation.*

For the easier understanding of the generalizations we introduce in this paper, we need to review the modelling of the random processes α_t, β_t introduced in [6]. Since the model for β_t is the same as that of α_t, only its parameters are different, it will be sufficient to deal with α_t.

Let λ be a real number satisfying $0 \le \lambda \le 1$ and $t_1, \dots, t_n$ further $\tau_1, \dots, \tau_{n-1}$ be independent samples taken from the population uniformly distributed in $(0,1)$. Let $\tau_1^* \le \tau_2^* \le \cdots \le \tau_{n-1}^*$ be the ordered sample corresponding to τ_i, $i = 1, \dots, n-1$ and put $\tau_0^* = 0$, $\tau_n^* = 1$. Now we define α_t in the following manner:

$$\alpha_t = c\lambda\nu/n + c(1-\lambda)\tau_\nu^*, \qquad 0 \le t \le T, \tag{1.2}$$

where ν is the number of those t_i which are smaller than t, c is a positive constant, cT equals the total demand occurring in the time interval $(0,T)$ and this is supposed to be equal the total amount delivered in the same time interval. If $\lambda = 1$, then α_t is the empirical probability distribution function belonging to the sample $t_1, \dots, t_n$. In connection with β_t we use m instead of n and μ instead of λ.

In [10] it is proved that the following limit relations hold:

$$\begin{aligned}
&\lim_{\substack{m\to\infty \\ n\to\infty}} \mathbf{P}\left(\left(\frac{mn}{m+n+m(1-\lambda)^2+n(1-\mu)^2} \right)^{1/2} \sup_{0\le t\le 1} (\alpha_t - \beta_t) < y \right) \\
&= \lim_{\substack{m\to\infty \\ n\to\infty}} \mathbf{P}\left(\left(\frac{mn}{m+n+m(1-\lambda)^2+n(1-\mu)^2} \right)^{1/2} \sup_{0\le t\le 1} (\beta_t - \alpha_t) < y \right) \\
&= \begin{cases} 1 - \exp(-2y^2) & \text{if} \quad y > 0, \\ 0, & \text{if} \quad y \le 0. \end{cases}
\end{aligned} \tag{1.3}$$

Here we have fixed $T = 1$ for the sake of simplicity. This choice does not restrict the generality.

If we assume the left-hand sides of (1.3) approximately equal the right-hand side value, then for a given ϵ it turns out that $M = M_{\lambda,\mu}$ is an approximate solution of the Reliability Equation:

$$M_{\lambda,\mu} = c\left[\frac{1}{2}\left(\frac{1+(1-\lambda)^2}{n} + \frac{1+(1-\mu)^2}{m}\right)\log\frac{1}{\epsilon}\right]^{1/2}. \tag{1.4}$$

If α_t is a deterministic process and $\alpha_t = ct$ $(0 \le t \le 1)$, then the corresponding M value can be obtained from (1.4) if we take the limit $n \to \infty$. We proceed similarly if β_t is deterministic. We remark that the minimal amount δ delivered at one delivery time and λ satisfy the following relation: $\lambda = n\delta/c$. Similar relation holds for the parameters of the process β_t.

2 Generalization of the delivery and demand processes

In this section we repeat the generalization of the delivery process as it is given in [9].

In Section 1 we mentioned the following assumptions in connection with the delivery process:

(a) the number of delivery times is fixed, this was denoted by n;

(b) the n delivery time points are so distributed in the interval $(0,1)$ as the elements of a sample of size n taken from a population uniformly distributed in the same interval;

(c) the total amount delivered is constant and is equal to c which is also the total demand;

(d) The random vector the components of which are the random delivered amounts is stochastically independent of the random vector of the delivery time points;

(e) denoting by δ the smallest amount to be delivered if a delivery occurs, the model for the distribution of the remaining amount among the n delivery time is the following:

Divide the interval $(0, c-n\delta)$ into n parts by choosing $n-1$ independent and uniformly distributed random points and assign the quantities equal to the lengths of the subintervals to the n delivery times. In what follows we maintain the assumptions (a), (c), (d), and modify the assumptions (b), (e).

For the modeling of the delivery process we choose L uniformly distributed independent random points in the interval $(0, c-n\delta)$, where $L > n-1$. Let $y_1^*, \ldots, y_L^*$ denote the ordered sample formed from the L random points. Out of this ordered sample we select those which have subscripts $k_1 < k_2 < \cdots < k_{n-1}$ and add to the fixed delivery amounts the following

$$\eta_1 = y_{k_1}^*, \quad \eta_2 = y_{k_2}^* - y_{k_1}^*, \quad \ldots, \quad \eta_n = c - n\delta - y_{k_{n-1}}^*. \tag{2.1}$$

Thus the amounts delivered at the delivery times will be

$$\delta + \eta_1, \quad \delta + \eta_2, \quad \ldots, \quad \delta + \eta_n. \tag{2.2}$$

A similar model is used for the delivery time points. To the fixed amount to be delivered at one occasion there correspond a fixed time γ as the minimal distance between two consecutive delivery time points ($0 \le \gamma \le 1/n$). The delivery time points are selected from an ordered sample $x_1^* \le x_2^* \le \cdots \le x_N^*$ of a sample of size N taken from a population uniformly distributed in $(0, 1-n\gamma)$, so that we select those elements which have subscripts $j_1 < j_2 < \cdots < j_n$ form the random variables

$$\xi_1 = x_{j_1}^*, \quad \xi_2 = x_{j_2}^* - x_{j_1}^*, \quad \ldots, \quad \xi_n = x_{j_n}^* - x_{j_{n-1}}^* \tag{2.3}$$

and finally take the partial sums of the random variable

$$\gamma + \xi_1, \quad \gamma + \xi_2, \quad \ldots, \quad \gamma + \xi_n. \tag{2.4}$$

This partial sum represents the n delivery time points.

Let $s(z_1, \ldots, z_{n-1})$ denote the joint probability density function of the random variables $\eta_1, \ldots, \eta_{n-1}$. It is easy to see that this function has the following form

$$\begin{aligned} & s(z_1, \ldots, z_{n-1}) \\ & \quad = \left(\frac{1}{c - n\delta}\right)^n \frac{\Gamma(L+1)}{\Gamma(k_1)\Gamma(k_2 - k_1)\cdots\Gamma(k_{n-1} - k_{n-2})\Gamma(L + 1 - k_{n-1})} \\ & \qquad \times \left(\frac{z_1}{c - n\delta}\right)^{k_1 - 1} \left(\frac{z_2}{c - n\delta}\right)^{k_2 - k_1 - 1} \cdots \left(\frac{z_{n-1}}{c - n\delta}\right)^{k_{n-1} - k_{n-2} - 1} \\ & \qquad \left(1 - \frac{z_1 + \cdots + z_{n-1}}{c - n\delta}\right)^{L - k_{n-1}} \end{aligned} \tag{2.5}$$

if $z_i > 0$, $i = 1, \ldots, n-1$; $z_1 + \ldots + z_{n-1} < c - n\delta$ and $s(z_1, \ldots, z_{n-1}) = 0$ otherwise. $\Gamma(y) = \int_0^\infty x^{y-1} e^{-x}\, \mathrm{d}x$. $\xi_1, \ldots, \xi_n$ have a similar joint density function.

Thus the random vectors $(\eta_1, \ldots, \eta_{n-1})$ and $(\xi_1, \ldots, \xi_n)$ have Dirichlet distributions. For properties of this multivariate distribution the reader is referred to [12].

3 The inventory models

Model I [9]

The model for the delivery process is that one discussed in Section 2. The demand is assumed to have constant intensity i.e., the demand occurring in the interval $(0, t)$ is equal to ct where c is a constant. M denotes the initial stock level. The demand will be met continuously in the whole interval $(0, 1)$ if and only if the following relations hold

$$\begin{aligned} & M && \ge \gamma + \xi_1, \\ & M + \delta + \eta_1 && \ge 2\gamma + \xi_1 + \xi_2, \\ & && \vdots \\ & M + (n-1)\delta + \eta_1 + \cdots + \eta_{n-1} && \ge n\gamma + \xi_1 + \cdots + \xi_n. \end{aligned} \tag{3.1}$$

Let us introduce the notations:

$$\begin{aligned}
\zeta_1 &= \xi_1, \\
\zeta_2 &= \xi_1 + \xi_2 - \eta_1, \\
&\vdots \\
\zeta_4 &= \xi_1 + \cdots + \xi_n - \eta_1 - \cdots - \eta_{n-1}.
\end{aligned} \tag{3.2}$$

The random vectors $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_{n-1})$ and $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ are independent and their probability density functions are logconcave functions in $\mathbb{R}^{n-1}$ resp. $\mathbb{R}^n$. It follows that the altogether $2n-1$ components have a logconcave joint density in $\mathbb{R}^{2n-1}$.

The notation of a logconcave probability measure was introduced in [7]. A probability measure P defined on the measurable subsets of $\mathbb{R}^m$ is said to be logconcave if for every pair A, B of convex subsets, of $\mathbb{R}^m$ and every $0 < \lambda < 1$ the following inequality holds:

$$\mathbf{P}(\lambda A + (1-\lambda)B) \geq [\mathbf{P}(A)]^{\lambda}[\mathbf{P}(B)]^{1-\lambda}. \tag{3.3}$$

The main theorem of [7] says that if a probability measure is generated by a logconcave probability density, then it is a logconcave measure. On the other hand any linear transform of a random vector having logconcave distribution has again logconcave distribution [9, Theorem 3]. Thus the random vector $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_n)$ has logconcave probability distribution.

We can write the Reliability Equation in our case, by taking into account only one basic commodity, as follows:

$$h(M) = \mathbf{P}(\zeta_i \leq M + (i-1)\delta - i\gamma, \quad i = 1, \ldots, n) = p, \tag{3.4}$$

where $0 < p < 1$ and $p \approx 1$ in practice (see Fig. 1).

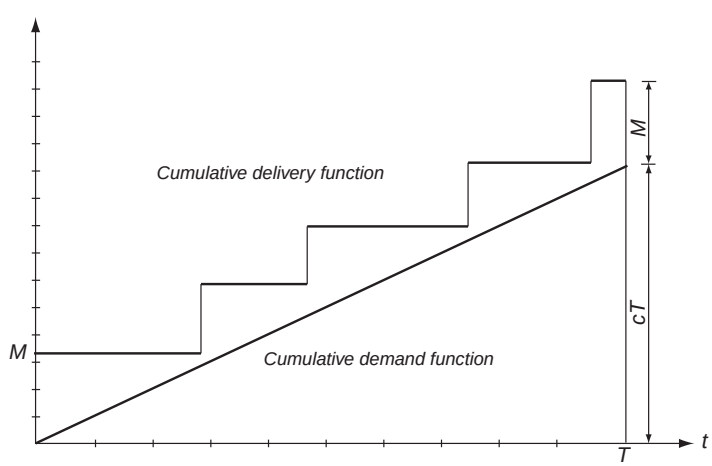

Fig. 1: The cumulative delivery function should be above the cumulative demand function with a prescribed high probability (for each material).

The function $h(M)$ is logconcave on the halfline $[0,\infty)$ because the joint probability distribution function of a random vector having logconcave probability distribution is a logconcave point function [7]. Model I for more than one basic material consists of the following stochastic programming problem:

$$\begin{aligned} &\text{minimize} \quad d^{(1)}M^{(1)} + \cdots + d^{(l)}M^{(l)}, \\ &\text{subject to } h(\boldsymbol{M}) = h_1(M^{(1)}) \cdots h_l(M^{(l)}) \geq p, \\ &\qquad\qquad \boldsymbol{M} \geq \mathbf{0}, \qquad \boldsymbol{M} \in D, \end{aligned} \tag{3.5}$$

where $\boldsymbol{M} = (M^{(1)}, \ldots, M^{(l)})$ and D is a subset of $\mathbb{R}^l$ determined by some constraints such as: the components of $\boldsymbol{M}$ be smaller than or equal to certain upper bounds or that the initial stock amounts do not take more room than a certain upper limit and do not require more financial investment than a further upper limit, etc. The numbers $d^{(1)}, \ldots, d^{(l)}$ are nonnegative and they are some valuations of units of goods to be determined on the basis of local knowledge. Sometimes the objective function turns out to be nonlinear.

In the above discussion we assumed that the demand function is linear. Of course we can drop this assumption and use that model for demand processes introduced in Section 2.

Stochastic programming models with independent joint constraints were considered first by Miller and Wagner [4].

Model II.

This model differs from the previous one in that further constraints containing conditional expectation appear. With this we prescribe not only the rarity of the occurrence of unsatisfied demand but also prescribe upper bound for the average magnitude of the unsatisfied demand. Thus upper bound may depend on the basic material. We assume that unsatisfied demand will not be lost. Thus the model works with backorders.

If one of the inequalities (3.1) is violated, it means there was a shortage just prior to the considered delivery time. The shortage is proportional to the unsatisfied demand. To assume a deterministic demand process with constant intensity means that the length of the time interval in which unsatisfied demand exist is proportional to the magnitude of the violation. Here we did take into account that no demand is lost. Our model consists of the problem formulated below. The superscripts refer to the various basic commodities:

$$\begin{aligned} &\text{minimize} \quad d^{(1)}M^{(1)} + \cdots + d^{(l)}M^{(l)}, \\ &\text{subject to } h(\boldsymbol{M}) \geq p, \\ &\qquad \mathbf{E}(\zeta_i^{(j)} - M^{(j)} - (i-1)\delta^{(j)} + i\gamma^{(j)} \mid \zeta_i^{(j)} - M^{(j)} - (i-1)\delta^{(j)} + i\gamma^{(j)} > 0) \leq g_i^{(j)}, \\ &\qquad i = 1, \ldots, n; \qquad j = 1, \ldots, l, \\ &\qquad \boldsymbol{M} \geq 0, \qquad \boldsymbol{M} \in D, \end{aligned} \tag{3.6}$$

where the $g_i^{(j)}$ are constant and $\mathbf{E}$ stands for expectation. The conditional expectation type constraints may even replace the probabilistic constraint. For every i, j the random variable $\zeta_i^{(j)}$ has a logconcave probability density. It follows from this (see [8]) that

the constraining functions in the conditional expectation constraints are monotonically decreasing functions of the variables $M^{(j)}$ and every such constraint is simply equivalent to a lower bound for the variable $M^{(j)}$ appearing in the constraint. We return to this question at the end of the section.

Model III.

Again we assume that no demand will be lost. The difference between this model and (3.6) consists in a penalty term what we introduce now. Let us introduce the random variables

$$\kappa_i^{(j)} = \begin{cases} q_i^{(j)}(\zeta_i^{(j)} - M^{(j)} - (i-1)\delta^{(j)} + i\gamma^{(j)}) & \text{if } \zeta_i^{(j)} - M^{(j)} - (i-1)\delta^{(j)} + i\gamma^{(j)} > 0, \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

$i = 1, \ldots, n$; $j = 1, \ldots, l$ where $q_i^{(j)} \geq 0$ for every i and j. It is easy to show that $\mathbf{E}(\kappa_i^{(j)})$ is a convex function of the variable $M^{(j)}$. To this it is enough to know that $\zeta_i^{(j)}$ has a continuous probability distribution. Since Model III has only a new objective function as compared to the model given by (3.6), it will be enough to formulate the new objective function, given by:

$$\sum_{j=1}^{l} d^{(j)} M^{(j)} + \sum_{j=1}^{l} \sum_{i=1}^{n} \mathbf{E}(\kappa_i^{(j)}). \tag{3.8}$$

The construction of the above three models are in correspondence with the three general model constructions given in [8].

Model III contains model I and Model II as special cases. We obtain Model II by setting $q_i^{(j)} = 0$, and Model I by setting $g_i^{(j)} = \infty$, $i = 1, \ldots, n$; $j = 1, \ldots, l$.

Now we return to the conditional expectation contained in Problem (3.6) and the expectations $\mathbf{E}(\kappa_i)$. Let $f_i^{(j)}$ resp. $F_i^{(j)}$ denote the probability density and the probability distribution functions of the random variable $\zeta_i^{(j)}$. For the sake of simplicity the superscripts will be omitted in the sequel. It is easy to see that if ζ is a continuously distributed random variable and a is a constant, then the following equality holds:

$$\begin{aligned} \mathbf{E}(\zeta - a \mid \zeta - a > 0) &= \int_a^\infty [1 - F(x)]\,\mathrm{d}x / [1 - F(a)] \\ &= \int_a^\infty x f(x)\,\mathrm{d}x / [1 - F(a)] - a, \end{aligned} \tag{3.9}$$

where F is the probability distribution function of ζ. In view of this we have that

$$\begin{aligned} &\mathbf{E}(\zeta_i - M - (i-1)\delta + i\gamma \mid \zeta_i - M - (i-1)\delta + i\gamma > 0) \\ &\quad = \frac{1}{1 - F_i(M + (i-1)\delta - i\gamma)} \int_{M+(i-1)\delta - i\gamma}^{1-n\gamma} x f_i(x)\,\mathrm{d}x - M - (i-1)\delta + i\gamma. \end{aligned} \tag{3.10}$$

Similarly we obtain that

$$\mathbf{E}(\kappa_i) = q_i \int_{M+(i-1)\delta-i\gamma}^{1-n\gamma} [1 - F_i(x)]\,\mathrm{d}x. \tag{3.11}$$

A simple argument shows that

$$\begin{aligned} f_i(x) = {} & \frac{\Gamma(N+1)\Gamma(L+1)}{\Gamma(j_i)\Gamma(N-j_i+1)\Gamma(k_{i-1})\Gamma(L-k_{i-1}+1)(c-n\delta)(1-n\gamma)} \\ & \times \int_0^b \left(\frac{x+u}{1-n\gamma}\right)^{j_i-1} \left(1-\frac{x+u}{1-n\gamma}\right)^{N-j_i} \left(\frac{u}{c-n\delta}\right)^{k_{i-1}-1} \\ & \left(1-\frac{u}{c-n\delta}\right)^{L-k_{i-1}} \mathrm{d}u, \end{aligned} \tag{3.12}$$

(where $b = \min\{1-n\gamma-x,\ c-n\delta\}$) if $0 < x < 1-n\gamma$ and $f_i(x) = 0$ otherwise, $i = 2, 3, \ldots, n$; further

$$f_1(x) = \frac{\Gamma(N+1)}{\Gamma(j_1)\Gamma(N-j_1+1)(1-n\gamma)} \left(\frac{x}{1-n\gamma}\right)^{j_1-1} \left(1-\frac{x}{1-n\gamma}\right)^{N-j_1} \tag{3.13}$$

if $0 < x < 1-n\gamma$ and $f_1(x) = 0$ otherwise.

As we already remarked, the ith conditional expectation-type constraint in Problem (3.6) can be converted into the simple inequality $M^{(j)} \geq M_i^{(j)}$ where $M_i^{(j)}$ is that value of $M^{(j)}$ for which the constraint holds with equality. This value can be determined by numerical integration of the function f_j.

4 Solution of the problems

In this section we present a solution method to the problems discussed in the previous section. We restrict ourselves to the problem of Model I since the solution of the two further problems requires only slight modification.

For the sake of simplicity let us agree that the constraint $M \in D$ be specialized so that it consist in the system of inequalities $M^{(j)} \leq 1$, $j = 1, \ldots, l$. These are, on the other hand, no real restrictions, because the equalities

$$h_j(1) = 1, \qquad j = 1, \ldots, l \tag{4.1}$$

hold trivially and these imply that the optimal $M^{(1)}, \ldots, M^{(l)}$ values are automatically smaller than or equal to 1. The upper bounding of the $M^{(1)}, \ldots, M^{(l)}$ values has the only consequence to allow us to refer to well-known convergence theorem concerning the SUMT method which we plan to use here. [2].

We apply the interior point version of the SUMT. Consider the following penalty function

$$G(r, \boldsymbol{M}) = \sum_{j=1}^{l} d^{(j)} M^{(j)} - r \left\{ \log \left[\prod_{j=1}^{l} h_j(M^{(j)}) - p \right] \right.$$
$$\left. + \sum_{j=1}^{l} \log M^{(j)}(1 - M^{(j)}) \right\}, \tag{4.2}$$

where r is a fixed positive number. It is easy to see that for every fixed $p > 0$ the function $h_1(M^{(1)}) \ldots h_l(M^{(l)}) - p$ is also logconcave which implies that for every fixed $r > 0$ the function $G(r, \boldsymbol{M})$ is convex on the set $\{\boldsymbol{M} \mid \boldsymbol{M} \geq \boldsymbol{0}\}$. From this we only need the fact that $G(r, M)$ is convex on the l-dimensional unit cube $\{\boldsymbol{M} \mid 0 \leq M^{(j)} \leq 1, j = 1, \ldots, l\}$. The SUMT interior point method works so that we take a sequence $r_1 > r_2 > \cdots$ consisting of positive numbers, tending to 0 and minimize $G(r_k, M)$ with respect to $\boldsymbol{M}$ (in principle) for every r_k. If $\boldsymbol{M}_k$ is the minimizing vector then $G(r_k, M_k)$ tends to the minimum value of the objective function in Model I. Thus $\boldsymbol{M}_k$ is an approximate optimal solution to the problem if k is large enough.

The referred to convergence is ensured if the set of feasible solutions is bounded, the constraining functions as well as the objective function are continuous on the set of feasible solutions, further there exists interior point of this set and all inequalities hold as strict inequalities at every interior point. As regards Model I here we have the constraints $0 \leq M^{(j)} \leq 1$, $j = 1, \ldots, t$ and the probabilistic constraint which restricts further the unit cube. The former ones hold strictly at every interior point of the unit cube thus it will be enough to consider the probabilistic constraint. Let $\boldsymbol{M}_1$ be an interior point in the set of feasible solutions. We shall show that $h(\boldsymbol{M}_1) > p$ where p is fixed and $0 < p < 1$ the line section connecting the points $\boldsymbol{1}$ and $\boldsymbol{M}_1$ is entirely feasible because the set of feasible solutions is convex. Let $\boldsymbol{M}_0$ be a feasible point on the line connecting $\boldsymbol{1}$ and $\boldsymbol{M}_1$ lying outside the section so that $\boldsymbol{M}_1$ be between $\boldsymbol{1}$ and $\boldsymbol{M}_0$. Then there exists a $0 < \lambda < 1$ so that

$$\boldsymbol{M}_1 = \lambda \boldsymbol{1} + (1 - \lambda) \boldsymbol{M}_0$$

from which, using the logconcavity of the function $h(M)$, it follows that

$$h(\boldsymbol{M}_1) \geq [h(\boldsymbol{1})]^{\lambda} [h(\boldsymbol{M}_0)]^{1-\lambda} \geq p^{1-\lambda} > p. \tag{4.3}$$

Thus we have shown that the SUMT interior point method is convergent in case of Model I.

Many general unconstrained minimization technique can be applied for the function (4.2). Some of them use only function values, some use gradient values too. In order to facilitate the application of methods belonging to the latter category, we present a methodology to compute the gradient values. Since $h(\boldsymbol{M})$ is the product of the functions $h_j(M^{(j)})$, $j = 1, \ldots, l$ it will be enough to consider the derivatives of the functions $h_j(M^{(j)})$. Let us omit the j, for the sake of simplicity, The function (3.4) is the joint probability distribution function of the random variables $\zeta_1, \ldots, \zeta_n$ at the point with coordinates $M + (i-1)\delta - i\gamma$, $i = 1, \ldots, n$.

We remark that if $F(z)$ is the probability distribution function corresponding to a continuous probability distribution, then the following relation holds:

$$\frac{\partial F(z)}{\partial z_i} = F(z_j,\ j \neq i \mid z_i) f_i(z_i), \qquad i = 1, \dots, n, \tag{4.4}$$

where $f_1, \dots, f_n$ are the probability density functions of the one-dimensional marginal distributions and $F(\cdot \mid z_i)$ is the $(n-1)$-dimensional conditional probability distribution function given that the ith random variable equals z_i.

We assume that $n\delta < c$, $n\gamma < 1$ (if one of the equalities $n\delta = c$, $n\gamma = 1$ holds, our procedure can essentially be simplified). To compute the derivative of the function (3.4), first we take the partial derivatives with respect to all $z_1, \dots z_n$ of the function

$$\mathbf{P}(\zeta_i \leq z_i + (i-1)\delta - i\gamma,\ i = 1, \dots, n) \tag{4.5}$$

and put $z_1 = \dots = z_n = M$. The sum of these equals the derivative of $h(M)$. The partial derivative of the function (4.5) with respect to z_i can be obtained by using the formula (4.4). Putting $z_1 = \dots = z_n = M$ we obtain

$$\begin{aligned}
&\mathbf{P}(\zeta_i \leq M + (j-1)\delta - j\gamma,\ j \neq i \mid \zeta_i = M + (i-1)\delta - i\gamma) f_i(M + (i-1)\delta - i\gamma) \\
&= f_i(M + (i-1)\delta - i\gamma) \int_0^v \mathbf{P}(\zeta_j \leq M + (j-1)\delta - j\gamma;\ j \neq i \mid \xi_1 + \dots + \xi_i \\
&= M + (i-1)\delta - i\gamma + x_1\eta_1 + \dots + \eta_{i-1} = x) \\
&\quad \times \frac{\Gamma(N+1)}{\Gamma(j_i)\Gamma(N+1-j_i)(1-n\gamma)} \left(\frac{M + (i-1)\delta - i\gamma + x}{1 - n\gamma} \right)^{j_i - 1} \\
&\quad \times \left(1 - \frac{M + (i-1)\delta - i\gamma + x}{1 - n\gamma} \right)^{N - j_i} \frac{\Gamma(L+1)}{\Gamma(k_{i-1})\Gamma(L+1-k_{i-1})} \frac{1}{c - n\delta} \\
&\quad \times \left(\frac{x}{c - n\delta} \right)^{k_{i-1}-1} \left(1 - \frac{x}{c - n\delta} \right)^{L - k_{i-1}} \mathrm{d}x, \\
&v = \min\{1 - M - (i-1)\delta - (n-i)\gamma,\ c - n\delta\},
\end{aligned} \tag{4.6}$$

where $f_i(z)$ is the probability density function of the random variable ζ_i. The probability in the second row of (4.6) can be expressed as an absolute probability and thus we obtain an expression similar to (3.4).

We recall that the random variables $\xi_1, \dots, \xi_n$ arise from a sample of size N, taken from a population uniformly distributed in the interval $(0, 1)$, in a way described in Section 2. The joint distribution of $\xi_1, \dots, \xi_n$ given that $\xi_1 + \dots + \xi_i = u = M + (i-1)\delta - i\gamma + x$ coincide with the joint distribution of two independent random vectors. These vectors consists of $j_i - 1$ resp. $N - j_i$ components and in both cases the joint densities are given by expressions of the type (**??**). In case of the first vector N, n, $1 - n\gamma$ should be replaced by $j_i - 1$, $i - 1$, u and in case of the second vector, by $N - j_i$, $n - i$, $1 - n\gamma - u$, respectively. Similar is the situation concerning the random variables $\eta_1, \dots, \eta_{n-1}$.

We apply simulation for the computation of the probability $h(\boldsymbol{M})$. The computation of the gradient values is more sophisticated because beyond simulation numerical integration is also needed. Hence it seems to be more economic to apply gradient free minimization technique when carrying out the unconstrained minimization of the penalty function.

5 Simulation technique for the computation of the values of the function $h(M)$

Two methods are proposed. The first one follows the modelling of the delivery processes. We take many samples of size N resp. L, order them and select the required elements. This method has the great disadvantage that the ordering of the sample elements requires much computer time. It is known that the ordering time of N elements increases in the order of magnitude of $N \log_2 N$.

The second method is more effective than the just mentioned former one. It is based on the fact that any Dirichlet distribution can be represented as the joint distribution of random variables $y_1, \ldots, y_n$ by

$$y_i = \frac{x_i}{x_1 + \cdots + x_{n+1}}, \qquad i = 1, \ldots, n, \tag{5.1}$$

where $x_1, \ldots, x_{n+1}$ are independent, standard gamma distributed random variables with parameters $\vartheta_1 > 0, \ldots, \vartheta_{n+1} > 0$, i.e. x_i has the following probability density:

$$\frac{z^{\vartheta_i - 1} e^{-z}}{\Gamma(\vartheta_i)}. \tag{5.2}$$

In fact, the joint density of the random variables (5.1) is given by

$$\frac{\Gamma(\vartheta_1 + \cdots + \vartheta_{n+1})}{\Gamma(\vartheta_1) \cdots \Gamma(\vartheta_{n+1})} z_1^{\vartheta_1 - 1} \ldots z_n^{\vartheta_n - 1} (1 - z_1 - \cdots - z_n)^{\vartheta_{n+1} - 1} \tag{5.3}$$

if $z_i > 0$, $i = 1, \ldots, n$; $z_1 + \cdots + z_n < 1$ and is 0 otherwise. Thus by a suitable choice of $\vartheta_1, \ldots, \vartheta_{n+1}$, we get the required Dirichlet density.

Ahrens and Dieter [1] gave effective simulation technique for the simulation of the gamma distribution. Their method is particularly effective when the ϑ parameter is large or is not an integer.

The probability density functions, (2.5) and (**??**) slightly differ from the density function (5.3). The simulation technique described above requires only very simple modification in both cases. Let us consider the gamma probability density function

$$\frac{\lambda^{\vartheta} z^{\vartheta - 1} e^{-\lambda z}}{\Gamma(\vartheta)}, \qquad z > 0. \tag{5.4}$$

If $x_1, \ldots, x_{n+1}$ are independent and gamma distributed random variables with parameter pairs $\lambda_1 \vartheta_1; \cdots; \lambda, \vartheta_{n+1}$; where $\lambda = 1 - n\gamma$, $\vartheta_1 = j_1$, $\vartheta_2 = j_2 - j_1, \ldots, \vartheta_n = j_n - j_{n-1}$, $\vartheta_{n+1} = N - j_{n+1} + 1$ then the random variables defined by (5.1) have the same joint probability distribution as $\xi_1, \ldots, \xi_n$ do. On the other hand x_i can be represented as the sum of ϑ_i independent and exponentially distributed random variables with the same parameter λ, for every $i = 1, \ldots, n+1$. Finally, the exponentially distributed random variables can be represented as negative logarithms of random variables uniformly distributed in the interval $(0, 1)$. The simulation of the joint distribution of the random variables can be carried out in a similar way.

Table 1:

p	0.09	0.045	0.025	0.01
0.5	258	879	2645	15606
0.8	195	616	1783	10209
0.9	150	417	1120	6016
0.95	–	306	727	3481

In case of this second simulation technique we only take logarithms of the $N+1$ sample elements but we do not order them. The required computer time is much smaller than in the first case.

The probabilities are approximated by relative frequencies. The sample size ensuring prescribed precision can be determined by the inequality of Bernstein, that is: if ν_m denotes the frequency of an event of probability p, in the course of m independent experiments and ϵ is a given positive number, then

$$\mathbf{P}(|(\nu_m/m) - p| \geq \epsilon) \leq 2\exp\left\{-\frac{1}{2p(1-p)[1+\epsilon/(2p(1-p))]^2}\right\} \tag{5.5}$$

for $0 < \epsilon \leq p(1-p)$.

If the probability on the left-hand side equals δ, then for m we obtain the inequality

$$m \geq (1/\epsilon^2)2p(1-p)[1+\epsilon/(2p(1-p))]^2 \log(2/\delta) \tag{5.6}$$

if $0 < \epsilon \leq p(1-p)$. For fixed δ and ϵ the largest value of the right-hand side of (5.6) corresponds to $p = \frac{1}{2}$, and it is a monotonically decreasing function of p for $\frac{1}{2} < p < 1$ (and monotonically increasing for $0 < p < \frac{1}{2}$, provided $\epsilon \leq p(1-p)$. In such a way we can get a lower bound for m which is good for every p. This is important because our aim is to approximate the probability p. Sometimes we have certain bounds for p. This is the case in connection with such stochastic programming problems where we have probabilistic constraint, i.e., lower bound for the probability.

In our models we use at least 0.8 as lower bound for the function $h(M)$. In practice this means that the factors are greater than or equal to 0.9. Using this information, the required sample size is much smaller than would be the case without any previous information. Table 1 will illustrate the variation of the lower bound for m as a function of ϵ and p when δ is fixed at 0.1.

6 Numerical example

As an example we consider a product whose production involves two basic commodities needed and we want to determine the initial stock levels of the two basic commodities ensuring continuous production. The shortage in each of them stops the production and

the cost of such an event is relatively high so that one of our main objectives is to avoid shortage with prescribed probability near one.

We assume that the demands for both basis commodities are uniform in time and the unsatisfied demand us carried over i.e. the production plan has to be fulfilled. We assume that the basic commodities are delivered by two different sources so that the two delivery processes can be supposed stochastically independent. The $(0,1)$ time interval is now a quarter of a year, 90 days in other terms. According to long term statistics deliveries occur 4 resp. 5 times concerning the first resp. second basic material during one period (90 day). Table 2 shows actual delivery days for six past periods concerning the first basic material.

Table 2:

Deliveries Periods	1	2	3	4
1	23	41	61	82
2	27	48	73	88
3	30	39	60	90
4	19	48	68	89
5	24	50	65	78
6	28	42	71	82
Column averages	25.17	44.66	66.33	84.83

The minimum distance between two consecutive deliveries is 9 days. Since 90 days form a time interval of length 1, this means that the mentioned minimum distance is $\gamma^{(1)} = 0.1$. For the average delivery times we get in the same way

$$z_1 = 0.28, \qquad z_2 = 0.49, \qquad z_3 = 0.73, \qquad z_4 = 0.94.$$

Using our modeling of the delivery time process, we can write

$$\overline{z}_i = i\gamma^{(1)} + \mathbf{E}(x^*_{j_i}), \qquad i = 1, \ldots, n,$$

where $x^*_{j_i}$ denotes the j_ith element of the ordered sample of size N taken from the population uniformly distributed in the interval $(0, 1 - n\gamma^{(1)})$. We have to find integers N, $j_1, \ldots, j_n$ for which the following equalities hold at least approximately:

$$\mathbf{E}(x^*_{j_i}) = \frac{j_i(1 - n\gamma^{(1)})}{N} = \overline{z}_i - i\gamma^{(1)}, \qquad j_i = \frac{\overline{z}_i - i\gamma^{(1)}}{1 - n\gamma^{(1)}} N, \qquad i = 1, \ldots, n.$$

Since the $\overline{z}_i$, $i = 1, \ldots, n$ and $\gamma^{(1)}$ are rationals in practice, such integers N, $j_1, \ldots, j_n$ always exist. It is not worth always to require that the above equalities hold exactly. In fact, if we work with large numbers, then the computer time will considerably be increased. In the above example the values of

$$(\overline{z}_i - i\gamma^{(1)})/(1 - n\gamma^{(1)}), \qquad i = 1, 2, 3, 4$$

are 0.298; 0.493; 0.728; 0.903 and choosing $N = 10$, $j_1 = 3$, $j_2 = 5$, $j_3 = 7$, $j_4 = 9$ we get a good approximation for the above equalities.

Table 3 shows the delivered amounts of the first basic material in the same past six periods

Table 3:

Deliveries Periods	1.0	2.0	3.0	4.0	Totals
1	630	400	670	800	2500
2	700	500	600	900	2700
3	730	580	550	740	2600
4	720	620	650	1010	3000
5	760	580	760	1100	3200
6	750	650	780	920	3100

Dividing the rows by the sums of the rows we get Table 4. From this table we see that the minimal delivered amount is $\delta^{(1)} = 0.16$.

Table 4:

	0.252	0.16	0.268	0.32
	0.259	0.185	0.222	0.333
	0.28	0.223	0.211	0.248
	0.24	0.206	0.216	0.336
	0.237	0.181	0.238	0.344
	0.242	0.21	0.252	0.297
Column averages	0.252	0.194	0.234	0.319

If the column averages are denoted by $\overline{u}_1, \dots, \overline{u}_n$ and we introduce the further notation $v_i = \overline{u}_1 + \cdots + \overline{u}_i$, $i = 1, \dots, n$, then similarly to the case of the delivery times we write the equalities

$$v_i = i\delta^{(1)} + \mathbf{E}(y^*_{k_i}), \qquad i = 1, \dots, n-1,$$

where $y^*_{k_i}$ denotes the k_ith element of a sample of size L taken from a population uniformly distributed in the interval $(0, 1 - n\delta^{(1)})$. We want to determine integers L, $k_1, \dots, k_{n-1}$ so that the following equalities hold at least approximately. In our case the values of

$$(v_i - i\delta^{(1)})/(1 - n\delta^{(1)}), \qquad i = 1, 2, 3$$

are 0.255; 0.35; 0.555. Thus the choices $L = 20$, $k_1 = 5$, $k_2 = 7$, $k_3 = 11$ provide good approximations.

We can proceed in a similar way concerning the second basic commodity. Assume that we obtained the following values: $n = 5$, $N = 10$, $j_1 = 2$, $j_2 = 3$, $j_3 = 5$, $j_4 = 7$, $j_5 = 9$,

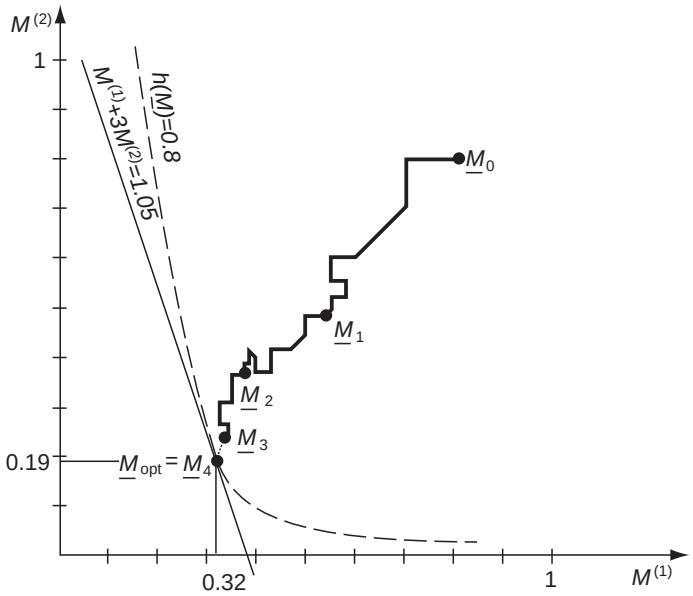

Fig. 2: The steps of the optimization procedure, where M_k is the optimal value of the kth unconstrained optimization.

$\gamma^{(2)} = 0.15$. $L = 10$, $k_1 = 2$, $k_2 = 5$, $k_3 = 7$, $k_4 = 8$, $\delta^{(2)} = 0.12$. As regards the objective function, we have chosen $d^{(2)} = 3d^{(1)}$.

The SUMT interior point method started with values between 0.6 and 0.8. The $r_1, r_2, \ldots$ sequence was chosen to be $1, \frac{1}{5}, \frac{1}{25}, \ldots$ and the initial values of the $k+1$st unconstrained minimization were the optimal values of the kth unconstrained minimization. The method was stopped when the change in the optimal values of the penalty function was less than 0.01. The method of Hooke and Jeeves [3] was applied for the minimization of the penalty function. In our numerical example the minimizing $M^{(1)}$ and $M^{(2)}$ belonging to $r = \frac{1}{25}$ were accepted as optimal solutions of the problem. These are $M^{(1)} = 0.32$ and $M^{(2)} = 0.19$. This means that 32% of the total demand of the first material and 19% of the total demand of the second material will serve the production without shortage with probability $p = 0.8$ and the cost will be minimum among all feasible alternatives.

The test programs written in FORTRAN run between 1.5 and 2.5 minutes on a CDC 330 computer. Further unconstrained optimization methods were also tested such as the method of Rosenbrock [11] and Powell [5] (see Fig. 2). The best computer time was produced by the method of Hooke and Jeeves, however. This method was successfully applied also in other stochastic programming problems where function values were determined by simulation.

References

[1] J.H. Ahrens and K. Dieter, "Computer methods for sampling from gamma, beta, Poisson and binomial distributions", *Computing* 12 (1974) 223–246.

[2] A.V. Fiacco and G.P. McCormick, *Nonlinear programming unconstrained minimization technique* (Wiley, New York, 1968).

[3] R. Hooke and T.A. Jeeves, "Direct search solution of numerical and statistical problems", *Journal of the Association for Computing Machinery* 8 (1959) 215–229.

[4] B.L. Miller and H.M. Wagner, "Chance constrained programming with joint constraints", *Operations Research* 13 (1965) 930–945.

[5] M.J.D. Powell, "An iterative method for finding the minimum of a function of several variables without calculating derivatives", *The Computer Journal* 7 (1964) 155–162.

[6] A. Prékopa, "Reliability equation for an inventory problem and its asymptotic solution", in: A. Prékopa, ed., *Colloquium on application of mathematics to economics* (Akadémiai Kiadó, Budapest, 1965) pp. 317–327.

[7] A. Prékopa, "On logarithmic concave measures with application to stochastic programming", *Acta Universitatis Szegedienis* 32 (1971) 301–316.

[8] A. Prékopa, "Contributions to the theory of stochastic programming", *Mathematical Programming* 4 (1973) 202–221.

[9] A. Prékopa, "Stochastic programming models for inventory control and water storage problems", in: A. Prékopa, ed., *Inventory control and water storage*, Colloquia Mathematica Societatis János Bolyai, 7 (Bolyai János Mathematical Society, Budapest, and North-Holland, Amsterdam, 1973) pp. 229–246.

[10] A. Prékopa, "Generalizations of the theorem of Smirnov with applications to a reliability type inventory problem", *Mathematische Operationsforschung und Statistik* 4 (1973) 283–297.

[11] H.H. Rosenbrock, "An automatic method for finding the greatest or least value of a function", *The Computer Journal* 3 (1960) 175–184.

[12] S.S. Wilks, *Mathematical statistics* (Wiley, New York, 1962).

[13] M. Ziermann, "Anwendung des Smirnov'schen Sätzen auf Lagerhaltungsproblem", Publications of the Mathematical Institute of the Hungarian Academy of Sciences, Series B, 8 (1965) 509–518.