
A Stochastic Allocation Problem

Author(s): Haim Mendelson, Joseph S. Pliskin, Uri Yechiali

Source: *Operations Research*, Vol. 28, No. 3, Part 2 (May - Jun., 1980), pp. 687-693

Published by: INFORMS

Stable URL: <http://www.jstor.org/stable/170035>

Accessed: 28/06/2009 05:08

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=informs>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Operations Research*.

A Stochastic Allocation Problem

HAIM MENDELSON, JOSEPH S. PLISKIN and URI YECHIALI

Tel Aviv University, Tel Aviv, Israel

(Received August 1977; accepted March 1979)

A given quantity of a resource is to be allocated to several activities. The amount of the resource allocated to each activity is used to supply stochastic demands occurring randomly. The system operates as long as all the demands can be met. Whenever the demand of any one activity exceeds its allotment the system fails. The problem is to find the allocation which maximizes the expected time until failure. It is shown that when the available quantity of the resource is large, the optimal allocation is proportionate to the expected demand rate of each activity. An application to a multi-installation inventory problem is presented.

WE CONSIDER the allocation of a given quantity of a resource to various activities. The work was motivated by a computer storage-allocation problem [10] for a system that uses several files where storage space is limited. Each transaction (or job) generates a demand vector, designating the space required in each file for record addition. Whenever one file runs out of space, the system must be reorganized. The problem is to allocate a given amount of storage space to the files so as to maximize the expected time until reorganization.

A natural suggestion is to allocate the resource in proportion to the expected demand rate of each activity. It turns out that such an allocation is not necessarily optimal. However, we show that, when the available quantity of the resource is large, the proportionate allocation rule is asymptotically optimal. That is, the quantity allocated to an activity divided by its demand rate is the same for all activities.

The allocation problem is formally presented in Section 1. An essential role is played by the function $m(\bar{x})$ = expected number of demands fulfilled until system failure. This multidimensional renewal function is analyzed in Section 2. The asymptotic form of $m(\bar{x})$ is used in Section 3 to show that a proportionate allocation is optimal. In Section 4 the ideas and results of the preceding sections are applied to solve a multi-installation inventory problem.

1. THE ALLOCATION PROBLEM

The system consists of n activities which require a common resource. Demands for the resource arise at instants $0 < t^1 < t^2 < \dots < t^k < \dots$,

where the times between demands, $\tau^k = t^k - t^{k-1}$, are i.i.d. random variables. We assume that $\tau^k \sim \tau$, and $E\tau < \infty$. A demand at time t^k is an n -dimensional vector $\underline{V}^k = (V_1^k, V_2^k, \dots, V_n^k)$, where V_i^k is the amount of resource required for activity i . \underline{V}^k are independent random variables identically distributed as a random variable \underline{V} whose distribution function is $F(\underline{v}) = P(\underline{V} \leq \underline{v})$. We do *not* assume independence among the components of \underline{V} . \underline{V} is nonnegative, and possesses a finite mean $E(\underline{V}) = \underline{\mu} = (\mu_1, \mu_2, \dots, \mu_n)$.

Let S be the amount of resource available for allocation. At time $t^0 = 0$, S is allocated to the various activities. Let $x_i \geq 0$ be the allotment for activity i ($i = 1, 2, \dots, n$). Let $D = \{x \mid x \in R^n, x \geq \underline{0}, \sum_{i=1}^n x_i \leq S\}$ denote the set of all feasible allocation vectors. Let $\underline{S}^k = (S_1^k, S_2^k, \dots, S_n^k)$ denote the vector of total demand up to time t^k , where the cumulative demand of activity i is $S_i^k = \sum_{j=1}^k V_i^j$. The system operates as long as $\underline{S}^k \leq x$. The system fails at the first moment t^k when $S_i^k > x_i$ for some i . That is, the system's lifetime $T(x)$ is given by $T(x) = \min\{t^k \mid S_i^k > x_i \text{ for some } i\}$. The problem is to find an allocation vector $x^* \in D$ which maximizes $ET(x)$.

Another measure of system performance is $N(x)$, the number of demands the system is capable of handling before it fails at time $T(x)$. That is, $N(x) = \max\{k \mid \underline{S}^k \leq x\}$. We have

LEMMA 1. $ET(x) = [EN(x) + 1] \cdot E\tau$.

Proof. $T(x) = t^{N(x)+1} = \sum_{k=1}^{N(x)+1} \tau^k$. The result readily follows using Wald's lemma.

It follows from Lemma 1 that $ET(x)$ is maximized together with $EN(x)$. Thus, the problem is to find $x^* \in D$ which maximizes $m(x) = EN(x)$.

A natural suggestion for an optimal allocation vector is to divide S in proportion to the expected demand of each activity—i.e., to let $x_i = \mu_i S / (\sum_{j=1}^n \mu_j)$. Unfortunately, this allocation is not necessarily optimal, as the following example shows. Suppose $\underline{V} = (V_1, V_2)$, where V_1, V_2 are i.i.d., and let $P(V_1 = 2) = P(V_1 = 4) = 1/2$. If $S = 6$, then a proportionate allocation yields $x_1 = x_2 = 3$. A simple computation shows that for such allocation $m(3, 3) = 1/4$, whereas for a nonproportionate allocation vector $x = (4, 2)$ one obtains $m(4, 2) = 1/2$. Thus, since $m(4, 2) > m(3, 3)$, a proportionate allocation is not optimal. However, we will show in the sequel that, for sufficiently large S , a proportionate allocation is asymptotically optimal.

2. ANALYSIS OF $m(x)$

In order to solve the allocation problem, we first analyze the objective function $m(x) = EN(x)$. The analysis is carried out in the following steps:

We derive a renewal-type equation for $m(\underline{x})$. From this equation we obtain the multidimensional Laplace-Stieltjes (L-S) transform of $m(\underline{x})$, which is then used to study the asymptotic behavior of $m(\underline{x})$ for large values of \underline{x} .

Problems of renewal theory in two dimensions have been studied by Hunter [7, 8] and by Bickel and Yahav [3]. The following lemma is a simple generalization of Hunter's "integral equation of two-dimensional renewal theory" [7, p. 387].

LEMMA 2. $m(\underline{x})$ satisfies the integral equation

$$m(\underline{x}) = F(\underline{x}) + \int_{\underline{v} \leq \underline{x}} m(\underline{x} - \underline{v}) dF(\underline{v}). \tag{1}$$

The proof follows from standard renewal-theoretic arguments by conditioning on \underline{V}^1 .

As in the one-dimensional case, equation (1) may be solved in terms of Laplace-Stieltjes transforms. For $\underline{s} = (s_1, s_2, \dots, s_n) > \underline{0}$, we define the multidimensional L-S transforms: $\tilde{m}(\underline{s}) = \int_{\underline{x} \geq \underline{0}} e^{-\underline{s} \cdot \underline{x}} dm(\underline{x})$, and $\tilde{V}(\underline{s}) = E[e^{-\underline{s} \cdot \underline{V}}] = \int_{\underline{v} \geq \underline{0}} e^{-\underline{s} \cdot \underline{v}} dF(\underline{v})$.

By transforming equation (1) we derive

$$\tilde{m}(\underline{s}) = \tilde{V}(\underline{s})/[1 - \tilde{V}(\underline{s})]. \tag{2}$$

In principle, one may invert $\tilde{m}(\underline{s})$ to obtain the renewal function $m(\underline{x})$. However, the actual computation is difficult and the results depend on the specific form of $\tilde{V}(\underline{s})$.

In various applications, the asymptotic behavior of $m(\underline{x})$ for large values of \underline{x} is of interest. Allocation problems usually deal with a situation where each activity is capable of handling numerous demands—i.e., $x_i \gg \mu_i$ for all i .

It is well known [5] that, in the one-dimensional case, the asymptotic behavior of $m(x)$ as $x \rightarrow \infty$ may be derived from the limiting behavior of $\tilde{m}(s)$ as $s \rightarrow 0$ ($s > 0$). We do not know of a multivariate analogue. In order to develop a Tauberian lemma that will serve our purposes, we first prove

LEMMA 3. Let $k(\underline{x})$ be a positive measure with L-S transform $\tilde{k}(\underline{s})$. The function $k(\underline{x})$ is homogeneous of degree r if and only if $\tilde{k}(\underline{s})$ is homogeneous of degree $(-r)$.

Proof. The L-S transform of $k(t\underline{x}) = t^r k(\underline{x})$ is $\tilde{k}(\underline{s}/t) = t^r \tilde{k}(\underline{s})$.

LEMMA 4. Let $m(\underline{x})$ and $k(\underline{x})$ be two positive measures with L-S transforms $\tilde{m}(\underline{s})$ and $\tilde{k}(\underline{s})$, respectively. Assume that the function $\tilde{k}(\underline{s})$ is homogenous of degree $(-r)$. If

$$\lim (\tilde{m}(\underline{s})/\tilde{k}(\underline{s})) = 1 \tag{3}$$

when $\underline{s} \rightarrow \underline{0}$ along any fixed positive direction, then

$$\lim(m(\underline{x})/k(\underline{x})) = 1 \tag{4}$$

where \underline{x} tends to infinity along any fixed positive direction.

Proof. Let $\underline{\lambda} > \underline{0}$ be an arbitrary positive vector, and let $\underline{s} = \underline{\lambda}/t$. Use of (3) and the homogeneity of $\bar{k}(\underline{s})$ yields $\lim_{t \rightarrow \infty} t^{-r} \bar{m}(\underline{\lambda}/t) = \bar{k}(\underline{\lambda})$. It follows that for each sequence $t_n \rightarrow \infty$, $\lim_{n \rightarrow \infty} t_n^{-r} \bar{m}(\underline{\lambda}/t_n) = \bar{k}(\underline{\lambda})$. But $t_n^{-r} \bar{m}(\underline{\lambda}/t_n)$ is the L-S transform of $t_n^{-r} m(t_n \underline{x})$ (evaluated at $\underline{s} = \underline{\lambda}$). Therefore, the continuity of the L-S transform yields $\lim_{n \rightarrow \infty} t_n^{-r} m(t_n \underline{x}) = k(\underline{x})$. This holds for every sequence $t_n \rightarrow \infty$, hence

$$\lim_{t \rightarrow \infty} t^{-r} m(t \underline{x}) = k(\underline{x}). \tag{5}$$

By virtue of Lemma 3, equation (5) may also be written as $\lim_{t \rightarrow \infty} m(t \underline{x})/k(t \underline{x}) = 1$.

Now let $\underline{\lambda} > \underline{0}$ be an arbitrary positive unit vector in \underline{s} -space. The directional derivative of $\bar{V}(\underline{s})$ in the direction of $\underline{\lambda}$ at $\underline{s} = \underline{0}$ exists (since $\underline{\mu} = E(\underline{V})$ is finite), and it is equal to $\underline{\lambda} \cdot \nabla \bar{V}(\underline{0}) = -\underline{\lambda} \cdot \underline{\mu}$. Hence, there exists a linear expansion $\bar{V}(\delta \cdot \underline{\lambda}) = \bar{V}(\underline{0}) + \delta \underline{\lambda} \cdot \nabla \bar{V}(\underline{0}) + o(\delta)$, where $o(\delta)/\delta \rightarrow 0$ as $\delta \rightarrow 0$ ($\delta > 0$). Since $\bar{V}(\underline{0}) = 1$ and $-\nabla \bar{V}(\underline{0}) = \underline{\mu}$, we have $\bar{V}(\delta \underline{\lambda}) = 1 - \delta(\underline{\lambda} \cdot \underline{\mu}) + o(\delta)$. Substitution in (2) and multiplication by $\delta(\underline{\lambda} \cdot \underline{\mu})$ yields,

$$\begin{aligned} (\underline{\mu} \cdot \delta \underline{\lambda}) \bar{m}(\delta \underline{\lambda}) &= \delta(\underline{\lambda} \cdot \underline{\mu}) [1 - \delta(\underline{\lambda} \cdot \underline{\mu}) + o(\delta)] / (\delta(\underline{\lambda} \cdot \underline{\mu}) + o(\delta)) \\ &= (1 - \delta(\underline{\lambda} \cdot \underline{\mu}) + o(\delta)) / (1 + (1/(\underline{\lambda} \cdot \underline{\mu})) \cdot o(\delta)/\delta), \end{aligned}$$

which tends to unity as $\delta \rightarrow 0$ ($\delta > 0$). It follows that if we define $\bar{k}(\underline{s}) = 1/(\underline{s} \cdot \underline{\mu})$, which is homogenous of degree $-r = -1$, the conditions of Lemma 4 are satisfied. Thus,

$$\lim_{t \rightarrow \infty} (1/t) m(t \underline{x}) = k(\underline{x}) \tag{6}$$

for all \underline{x} .

Our aim now is to find a measure $k(\cdot)$ on the nonnegative quadrant of R^n , whose L-S transform is $\bar{k}(\underline{s}) = \int_{v \geq 0} e^{-\underline{s} \cdot \underline{v}} dk(\underline{v}) = 1/(\underline{s} \cdot \underline{\mu})$. In the one-dimensional case, $\bar{k}(s) = 1/(s \cdot \mu)$, which implies that $dk(v) = dv/\mu$. Thus, we have the well-known elementary renewal theorem [5]: $m(x) \approx x/\mu$. However, the solution for higher dimensionality is somewhat less obvious.

THEOREM 1. $m(\underline{x}) \approx \min_{i=1,2,\dots,n}(x_i/\mu_i)$ (as \underline{x} tends to infinity along an arbitrary positive direction).

Proof. Let $B = \{v \mid v \in R^n, v = \alpha \underline{\mu}, \alpha \geq 0\}$. We define the measure $k(\cdot)$ on Borel-sets in R^n as follows: (i) For each section $[\underline{0}, \alpha \underline{\mu}] \subset B$, set $k([\underline{0}, \alpha \underline{\mu}]) = \alpha$. This clearly defines $k(\cdot)$ on all Borel-sets in B . (ii) For

each Borel-set $E \subset R^n$, we define $k(E) = k(E \cap B)$, where the right-hand side is defined by (i). We claim that $k(\cdot)$ satisfies (4). Indeed, $\bar{k}(g) = \int_{v \in B} e^{-g \cdot v} dk(v) = \int_0^\infty e^{-g \cdot \alpha \underline{\mu}} d\alpha = 1/(g \cdot \underline{\mu})$. It follows that $k(x)$, which is the k -measure of the cube $C_x = \{v \mid v \in R^n, 0 \leq v \leq x\}$, is equal to $k(C_x) = k(C_x \cap B) = k(\{\alpha \underline{\mu} \mid 0 \leq \alpha \underline{\mu} \leq x\}) = k([0, \alpha^* \underline{\mu}]) = \alpha^*$ where $\alpha^* = \alpha(x) = \min_{1 \leq i \leq n} (x_i/\mu_i)$. The result follows from Lemma 4.

3. OPTIMAL ALLOCATION

In most applications, S is sufficiently large to justify the assumption that $x_i \gg \mu_i$ for all i . The results of the previous section show that in such situations, the expected number of demands before system failure, $m(x)$, may be approximated by $\min_{i=1,2,\dots,n}(x_i/\mu_i)$. The optimization problem then becomes

$$\begin{aligned} \max\{m(x) = \min_{i=1,2,\dots,n}(x_i/\mu_i)\} \\ \text{s.t. } \sum_{i=1}^n x_i \leq S \quad \text{and} \quad x_i \geq 0, \quad i = 1, 2, \dots, n. \end{aligned} \tag{7}$$

THEOREM 2. *The optimal allocation vector is $x^* = S/(\sum_{j=1}^n \mu_j) \cdot \underline{\mu}$, i.e., x_i^* is proportional to μ_i .*

Proof. The problem (7) is equivalent to the following linear programming problem: maximize $\{\theta\}$ such that

$$\begin{aligned} \theta \mu_i - x_i \leq 0, \quad \text{for } i = 1, 2, \dots, n \quad \text{and} \quad \sum_{i=1}^n x_i \leq S \\ \theta \geq 0; x_i \geq 0, \quad \text{for } i = 1, 2, \dots, n. \end{aligned} \tag{8}$$

An immediate application of duality theory proves that $x_i^* = \mu_i S / \sum_{j=1}^n \mu_j$, $\theta^* = S / \sum_{j=1}^n \mu_j$ is an optimal solution of (8).

Denoting by y_i ($i = 1, 2, \dots, n + 1$) the corresponding dual variables of (8), it is readily seen that the optimal dual solution is $y_i^* = 1 / \sum_{j=1}^n \mu_j$ for $i = 1, 2, \dots, n + 1$. In particular we have

COROLLARY 1. $\partial m(x^*) / \partial S = y_{n+1}^* = 1 / \sum_{j=1}^n \mu_j$.

That is, a unit increase in the amount of the resource available for allocation increases the expected lifetime of the system by $E\tau / (\sum_{j=1}^n \mu_j)$.

4. APPLICATIONS

The type of stochastic allocation problems studied in this paper may arise in several areas. We have already mentioned the computer storage allocation problem [10] which motivated this work. The solution of this problem is obtained by a straightforward application of our previous results—i.e., allotment proportional to expected demand.

Another area of application is inventory systems. Consider a single-

commodity inventory system consisting of n centrally controlled outlets. Procurements are made on a system-wide basis by the central control point, which is kept constantly informed of the inventory levels at all locations. Demands for the commodity are independent random vectors $\underline{V}^k \sim \underline{V}$, where V_i^k is the quantity needed for the k th demand at installation i . \underline{V}^k may include up to $n - 1$ zeroes, so that a demand vector may represent demand at a single installation. The times between consecutive demands $\tau^k \sim \tau$ are i.i.d. with finite mean $E\tau$. When the total demand in one of the installations exceeds the inventory at hand, an order is placed and immediately delivered. The order is used to fulfill unmet demand and to replenish *all* the installations. No redistribution of stock among outlets is allowed between orders. Costs are of two types: an order cost, A , incurred at each replenishment; and an inventory-holding cost which is charged continuously over time at a constant rate c per unit of commodity per unit time. Procurement and transportation costs are not considered, since they do not affect the optimal solution.

The system just described belongs to the class of multi-installation inventory control systems. Some related inventory allocation models are described in references 1, 2, 4, 6, 9, and 13. The application studied here is a dynamic multi-installation continuous-review model. We consider the joint determination of the total order-quantity and its allocation among the outlets.

Let y_i be the level of inventory at installation i just before replenishment ($y_i < 0$ for at least one i). Let \underline{x} denote the vector of inventory levels immediately after replenishment—that is, $x_i - y_i$ is the quantity delivered to outlet i . Since the reorder point is not a decision variable, the problem is to find the optimal value of \underline{x} which minimizes the long run average cost per unit time. Note that here we impose no upper bound on $\sum_{i=1}^n x_i$.

The underlying process is regenerative [12]. Hence, the long run average cost tends with probability 1 to the expected total cost per cycle (= time between successive orders) divided by the expected length of a cycle.

Let $I(\underline{x})$ be the expected inventory holding cost per cycle, and let $T(\underline{x})$ be the length of a cycle. Then the problem is to minimize $\varphi(\underline{x}) = [I(\underline{x}) + A]/ET(\underline{x})$. By conditioning on \underline{V}^1 , we obtain

LEMMA 5. $I(\underline{x})$ satisfies the integral equation

$$I(\underline{x}) = c \left(\sum_{i=1}^n x_i \right) E\tau + \int_{\underline{v} \leq \underline{x}} I(\underline{x} - \underline{v}) dF(\underline{v}). \quad (9)$$

The techniques of Section 2 may now be used to approximate $I(\underline{x})$ and $ET(\underline{x})$. It follows (see details in [11]) that the minimization of $\varphi(\underline{x})$ may

be separated into two subproblems; (i) Determination of the total order quantity, $S = \sum_{i=1}^n x_i$, which is given by a variant of Wilson's lot-size formula:

$$S = \sqrt{2(\sum_{i=1}^n \mu_i)A/(cE\tau)}.$$

(ii) The allocation of S among the installations, which is proportionate, by virtue of Theorem 2.

REFERENCES

1. S. G. ALLEN, "Redistribution of Total Stock over Several User Locations," *Naval Res. Logist. Quart.* **5**, 337-345 (1958).
2. S. G. ALLEN, "A Redistribution Model with Set-up Charge," *Mgmt. Sci.* **8**, 99-108 (1961).
3. P. J. BICKEL AND J. A. YAHAV, "Renewal Theory in the Plane," *Ann. Math. Stat.* **36**, 946-955 (1965).
4. J. L. BRENNER, "Stock Control in a Many Depot System," *Naval Res. Logist. Quart.* **16**, 359-379 (1969).
5. W. FELLER, *An Introduction to Probability Theory and its Applications*, Vol. 2, John Wiley & Sons, New York, 1971.
6. G. HADLEY AND T. M. WHITIN, "An Inventory Transportation Model with N Locations," Chapt. 5, in *Multistage Inventory Models and Techniques*, pp. 116-142, H. Scarf, D. Gilford and M. Shelly (eds.), Stanford University Press, Stanford, Calif., 1963.
7. J. J. HUNTER, "Renewal Theory in Two Dimensions: Basic Results," *Adv. Appl. Prob.* **6**, 376-391 (1974).
8. J. J. HUNTER, "Renewal Theory in Two Dimensions: Asymptotic Results," *Adv. Appl. Prob.* **6**, 546-562 (1974).
9. D. IGLEHART AND A. LALCHANDANI, "An Allocation Model," *SIAM J. Appl. Math.* **15**, 303-323 (1967).
10. H. MENDELSON, J. S. PLISKIN AND U. YECHIALI, "Optimal Storage Allocation for Serial Files," *Comm. ACM* **22**, 124-130 (1979).
11. H. MENDELSON, J. S. PLISKIN AND U. YECHIALI, "A Stochastic Allocation Problem," Working Paper, Department of Statistics, Tel Aviv University, 1977.
12. S. M. ROSS, *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, 1970.
13. K. F. SIMPSON, JR., "A Theory of Allocations of Stocks to Warehouses," *Opns. Res.* **7**, 797-805 (1959).