Average costs in a continuous review (s,S) inventory system with exponentially distributed lead time

2. G. HADLEY AND T. M. WHITIN, *Analysis of Inventory Systems*, Prentice-Hall, Englewood Cliffs, N. J., 1963.
3. M. J. LIBERATORE, "A Stochastic Lead Time Inventory Model," unpublished doctoral dissertation, University of Pennsylvania, Philadelphia, 1976.
4. H. M. WAGNER, *Principles of Operations Research*, Ed. 2, Prentice-Hall, Englewood Cliffs, N. J., 1975.
5. A. R. WASHBURN, "A Bi-Modal Inventory Study with Random Lead Times," NTIS Report No. AD769404, Naval Postgraduate School, Monterey, Calif., September 1973.
6. W. I. ZANGWILL, *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Englewood Cliffs, N. J., 1969.

Average Costs in a Continuous Review (*s,S*) Inventory System with Exponentially Distributed Lead Time

J. WIJNGAARD and E. G. F. VAN WINKEL
University of Technology, Eindhoven, The Netherlands
(Received January 1977; accepted March 1978)

We describe a very elementary direct numerical method to find the average number of backorders, costs and related quantities in a continuous review (*s,S*) inventory system with exponentially distributed lead time. This method can also be used in the study of *E/M/C* queues with state-dependent service and arrival rates.

IN 1959 Galliher, Morse and Simond [2] investigated the steady-state behavior of an inventory system with Poisson arrival and negative exponential leadtime under the assumption that an (*s,q*)-ordering policy is used. They derived explicit expressions for the steady-state probabilities. These probabilities can be used to calculate the average number of backorders, the average inventory costs, and the average ordering costs. We consider the same inventory model but allow the arrival and service to be state dependent. A very elementary direct numerical method is described to find the above-mentioned quantities. For the non-state-dependent case we compared this direct method with the calculation of the explicit expressions of [2] and found about the same computing times.

Various types of continuous time (*s,S*) inventory models have been considered by Feeney and Sherbrooke [1], Gross and Harris [3][4], Higa, Feyerherm and Machado [5], Rose [6], Sherbrooke [7], Tijms [8], and Van der Genugten [9]. The leadtime in most of these models is constant

0030-364X/79/2702-0396 $01.25

or negative exponential and the demand is Poisson or compound Poisson.

The inventory system considered here is equivalent to an $M/M/.$ queue with bulk service and state-dependent arrival and service rates. Gross and Harris [4] give a solution procedure for a very special case. Their procedure consists of the determination of a number of roots of some polynomial and the solution of a set of linear equations. The method described here is simpler but can be used for a general class of state-dependent bulk service system.

1. PRELIMINARIES

The model consists of the following assumptions:

1. customers arrive one by one, and each customer demands one unit;
2. unfulfilled demands are backlogged;
3. orders for replenishing the inventory are placed as soon as the inventory position (= the inventory on hand plus on order minus backorders) reaches a level s;
4. the reorder quantity is $q = S - s$.

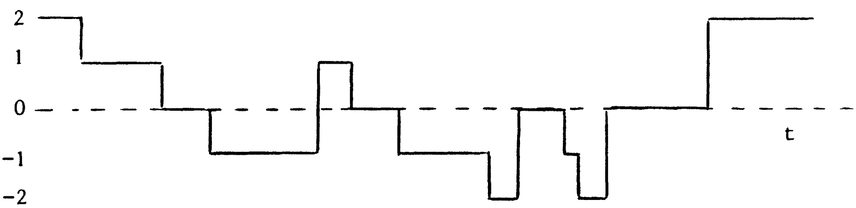

Figure 1. Behavior of the net inventory.

The behavior of the net inventory (= inventory on hand minus back-orders) is illustrated in Figure 1 for $s = 0$, $q = 2$. Since the net inventory does not exceed $s + q$, it can be denoted by $s + q - n$, where n is said to be the state of the system. Backorders occur if $n > s + q$. The number of outstanding orders is $[n/q]$, where $[x]$ is the largest integer smaller than or equal to x.

5. The interarrival time in state n is negative exponential with mean $1/\lambda_n$.
6. In state n the probability of a special order being filled in an infinitesimal interval of length Δt is $\mu_n \Delta t + o(\Delta t)$. That means that the probability of some order being filled is $[n/q]\mu_n \Delta t + o(\Delta t)$. (If $\mu_n = \mu$, $n = 0, 1, \ldots$ this is equivalent to the case where the procurement lead time is negative exponential with mean $1/\mu$.)
7. The cost per period in state n is c_n.

The expected value of the time it takes to move from state n into state 0 is denoted by t_n, $n = 1, 2, \ldots$, and the expected cost during the same period is denoted by w_n, $n = 1, 2, \ldots$, (see [10] for a proof of the existence of these expectations). A cycle is defined as the time between two successive entries into state 0. Hence the expected cycle time is $1/\lambda_0 + t_1$ and the expected cycle costs are $c_0/\lambda_0 + w_1$. This implies that the average cost per unit of time, C, is equal to $(c_0/\lambda_0 + w_1)/(1/\lambda_0 + t_1)$. Notice that by a suitable choice of c_n the quantity C can be made equal to the average number of backorders, the average inventory on hand, the average ordering costs, the stationary probabilities, and so on.

If $c_n = 0$ for $n \leqq s + q$ and $c_n = n - s - q$ for $n > s + q$, then w_1 is equal to the expected total waiting time of all customers until the first visit to 0 and C is equal to the average number of backorders. If $c_n = s + q - n$ for $n \leqq s + q$ and $c_n = 0$ for $n > s + q$, then C is equal to the average inventory on hand. If $c_n = 1$ for $n = k$ and $c_n = 0$ elsewhere, then C is equal to the stationary probability of being in state k.

2. APPROXIMATION OF t_1 AND w_1

One can divide t_n in two parts—the expected time until the first transition and the expected time from the first transition until the end of the cycle. The expected time until the next transition is λ_n^{-1} if $n = 1, \ldots, q - 1$ and $(\lambda_n + [n/q]\mu_n)^{-1}$ if $n \geqq q$. The probability that the next transition is a demand is $\lambda_n/(\lambda_n + [n/q]\mu_n)$; the probability that it is an arrival of an order is $[n/q]\mu_n/(\lambda_n + [n/q]\mu_n)$. Hence

$$t_n = \lambda_n^{-1} + t_{n+1} \quad \text{for} \quad n = 1, \ldots, q - 1 \tag{1}$$

$$t_n = (\lambda_n + [n/q]\mu_n)^{-1} + [n/q]\mu_n(\lambda_n + [n/q]\mu_n)^{-1}t_{n-q} + \lambda_n(\lambda_n + [n/q]\mu_n)^{-1}t_{n+1} \quad \text{for} \quad n \geqq q. \tag{2}$$

The same can be done for w_n,

$$w_n = c_n\lambda_n^{-1} + w_{n+1} \quad \text{for} \quad n = 1, \ldots, q - 1 \tag{3}$$

$$w_n = c_n(\lambda_n + [n/q]\mu_n)^{-1} + n/q\mu_n(\lambda_n + [n/q]\mu_n)^{-1}w_{n-q} + \lambda_n(\lambda_n + [n/q]\mu_n)^{-1}w_{n+1} \quad \text{for} \quad n \geqq q. \tag{4}$$

Let $y_0 \equiv t_1, y_n \equiv t_{n+1} - t_n$, for $n = 1, 2, \ldots, z_0 \equiv w_1, z_n \equiv w_{n+1} - w_n$, for $n = 1, 2, \ldots$. Then

$$y_n = -\lambda_n^{-1}, \quad \text{for} \quad n = 1, \ldots, q - 1. \tag{5}$$

From (2) it follows that $(\lambda_n + [n/q]\mu_n)t_n = 1 + [n/q]\mu_n t_{n-q} + \lambda_n t_{n+1}$; hence $\lambda_n(t_{n+1} - t_n) = [n/q]\mu_n(t_n - t_{n-q}) - 1$ and

$$y_n = [n/q]\mu_n\lambda_n^{-1} \sum_{k=1}^{q} y_{n-k} - \lambda_n^{-1} \quad \text{for} \quad n \geqq q. \tag{6}$$

In the same way we have for z_n,

$$z_n = -c_n\lambda_n^{-1}, \quad \text{for} \quad n = 1, \ldots, q - 1, \tag{7}$$

TABLE I

$y_0 = 1.5$	$y_0 = 2$
$y_1 = 0.5$	$y_1 = 1$
$y_2 = 0$	$y_2 = 1$
$y_3 = -1$	$y_3 = 2$
$y_4 = -5$	$y_4 = 7$

$$z_n = [n/q]\mu_n\lambda_n^{-1} \sum_{k=1}^{q} z_{n-k} - c_n\lambda_n^{-1} \quad \text{for} \quad n \geqq q. \tag{8}$$

From (5)–(8) it is clear that the sequences $\{y_n\}$ and $\{z_n\}$ are determined by y_0 and z_0. Let $\{y_n(y)\}$ and $\{z_n(z)\}$ be the sequences determined by (5)–(8) and $y_0 = y$, $z_0 = z$. Table I shows the first part of the sequences $\{y_n(y)\}$ for $\lambda_n = \mu_n = 1$, $n = 1, 2, \ldots$, $q = 1$ and $y = 1.5$ and 2. The sequences are given by $y_n = ny_{n-1} - 1$, $n = 1, 2, \ldots$. We see that for $y = 1.5$, $\lim_{n\to\infty} y_n(y) = -\infty$ and for $y = 2$, $\lim_{n\to\infty} y_n(y) = +\infty$. However, it is possible to prove that the sequence $\{y_n\}$, which is the sequence $\{(y_n(t_1))\}$, is bounded [10]. Since $y_n(y)$ is monotonically increasing in y, this implies $1.5 < t_1 < 2$. In the same way one can find that $1.7 < t_1 < 1.8$ (see Table II). This approach requires a lot of work to find good approximations for t_1 and w_1, but it is possible to reduce the computational work by using the linearity of $y_n(y)$ and $z_n(z)$ in y and z. We show this for the more general case of state-dependent arrival and service rates. From (5)–(8) it follows that we can write $y_n(y) = \alpha_n y + \beta_n$, $z_n(z) = \alpha_n z + \gamma_n$ where $\{\alpha_n\}$, $\{\beta_n\}$, $\{\gamma_n\}$ are given by $\alpha_0 = 1$, $\beta_0 = 0$, $\gamma_0 = 0$, $\alpha_n = 0$, $\beta_n = -\lambda_n^{-1}$, $\gamma_n = -c_n\lambda_n^{-1}$ for $n = 1, 2, \ldots, q-1$, and $\alpha_n = [n/q]\mu_n\lambda_n^{-1} \sum_{k=1}^{q} \alpha_{n-k}$, $\beta_n = [n/q]\mu_n\lambda_n^{-1} \sum_{k=1}^{q} \beta_{n-k} - \lambda_n^{-1}$, $\gamma_n = [n/q]\mu_n\lambda_n^{-1} \sum_{k=1}^{q} \gamma_{n-k} - c_n\lambda_n^{-1}$, for $n \geqq q$. The expression for α_n shows that $\alpha_n \to \infty$ for $n \to \infty$ if λ_n/μ_n is bounded in n. That means that $y_n(y)$ and $z_n(z)$ can be bounded for only one value of y and z. In [10] it is shown under rather general conditions on λ_n, μ_n and c_n (for instance, λ_n/μ_n and c_n/n bounded) that $y_n(t_1)$ and $z_n(w_1)$ are bounded in n. This implies that t_1 and w_1 are the unique values of y and z for which $y_n(y)$ and $z_n(z)$ are bounded. Dividing $y_n(t_1) = \alpha_n t_1 + \beta_n$, $z_n(w_1) = \alpha_n w_1 + \gamma_n$ by α_n and using the boundedness of $y_n(t_1)$ and $z_n(w_1)$ in n yield $t_1 = -\lim \beta_n/\alpha_n$, $w_1 = -\lim \gamma_n/\alpha_n$. Hence the calculation of α_n, β_n, γ_n for n large enough gives good approximations for t_1 and w_1. An important point, of course, is the error that is made by approximating t_1 and w_1 in this way. In the next

TABLE II

$y_0 = 1.7$	$y_0 = 1.8$
$y_1 = 0.7$	$y_1 = 0.8$
$y_2 = 0.4$	$y_2 = 0.6$
$y_3 = 0.2$	$y_3 = 0.8$
$y_4 = -0.2$	$y_4 = 2.2$
$y_5 = -2$	

lemma an upper bound on this error expressed in α_n, β_n, γ_n is given. A proof is given in [10].

LEMMA. *Let $N \geqq q$ be such that $c_n(\lambda_n + [n/q]\mu_n)^{-1}$ and $(\lambda_n + [n/q]\mu_n)^{-1}$ are bounded on $[N,\infty)$ and $[n/q]\mu_n \geqq r\lambda_n$ for some $r > 1$. Let β_n/α_n attain its minimum on $[N - q, N - 1]$ in k_l and its maximum in k_h and let m_l and m_h be the corresponding numbers for γ_n/α_n. Then $-\beta_{k_h}/\alpha_{k_h} \leqq y_0 \leqq -\beta_{k_l}/\alpha_{k_l} + (1 + r)A/r\alpha_{N-1}$, where A is an upper bound of $(\lambda_n + [n/q]\mu_n)^{-1}$ on $[N,\infty)$, and $-\gamma_{m_h}/\alpha_{m_h} \leqq z_0 \leqq -\gamma_{m_l}/\alpha_{m_l} + (1 + r)B/r\alpha_{N-1}$, where B is an upper bound of $c_n(\lambda_n + [n/q]\mu_n)^{-1}$ on $[N,\infty)$.*

The procedure to approximate t_1 and w_1, with an error of at most, say, ϵ, consists of the recursive calculation of α_n, β_n, γ_n. According to the expressions for α_n, β_n, γ_n it is sufficient to store the last q values of each of these sequences. From the lemma these last q values of α_n, β_n, γ_n also give an upper bound of the error (the difference between t_1 and $-\beta_n/\alpha_n$ and the difference between w_1 and $-\gamma_n/\alpha_n$). As soon as this upper bound is smaller than ϵ, the calculation can be stopped and the actual values of $-\beta_n/\alpha_n$ and $-\gamma_n/\alpha_n$ can be taken as approximations of t_1 and w_1.

3. EXTENSIONS

The higher moments of the costs per cycle can be calculated in the same way. Define ω_n as the costs until the end of the cycle if the system is now in state n and define ω_n' as the costs until the first transition. Let $w_n \equiv E(\omega_n)$ and $x_n \equiv E(\omega_n^2)$. Then

$$\begin{aligned} x_n &= E(\omega_n^2) = E((\omega_n' + \omega_n - \omega_n')^2) = E((\omega_n')^2) + 2E(\omega_n'(\omega_n - \omega_n')) \\ &+ E((\omega_n - \omega_n')^2) = E((\omega_n')^2) + 2E(\omega_n')E(\omega_n - \omega_n') + E((\omega_n - \omega_n')^2) \\ &= E((\omega_n')^2) + 2E(\omega_n')\{[n/q]\mu_n(\lambda_n + [n/q]\mu_n)^{-1}w_{n-q} \\ &+ \lambda_n(\lambda_n + [n/q]\mu_n)^{-1}w_{n+1}\} \\ &+ [n/q]\mu_n(\lambda_n + [n/q]\mu_n)^{-1}x_{n-q} + \lambda_n(\lambda_n + [n/q]\mu_n)^{-1}x_{n+1}. \end{aligned}$$

Hence, if w_n is known, one can apply the same methods to get $x_1 = E(\omega_1^2)$.

If $\lambda_n = \lambda$, $n = 0, 1, \ldots$ and $\mu_n = \mu$, $n = 0, 1, \ldots$, then the state in the inventory process corresponds to the number of phases in an $E_q/M/\infty$ system. If the interarrival times in the inventory process are Erlang-k distributed, the system is equivalent to an $E_{kq}/M/\infty$ queue. This extension gives no new difficulties.

It is not essential that from state n only transitions to state $n + 1$ and to state $n - q$ are possible. We can use the same methods when transitions to all states $0, 1, \ldots, n, n + 1$ are possible [10].

REFERENCES

1. G. FEENEY AND C. C. SHERBROOKE, "The $(S - 1, S)$ Inventory Policy under Compound Poisson Demand," *Management Sci.* **12,** 391–411 (1966).

2. H. P. Galliher, P. M. Morse, and H. Simond, "The Dynamics of Two Classes of Continuous-Review Inventory Systems," *Opns. Res.* **7,** 362–384 (1959).
3. D. Gross, and C. M. Harris, "On One-for-One Ordering Inventory Policies with State-Dependent Lead Times," *Opns. Res.* **19,** 735-760 (1971).
4. D. Gross, and C. M. Harris, "Continuous-Review (*s,S*) Inventory Models with State Dependent Lead Times," *Management Sci.* **19,** 567-574 (1973).
5. I. Higa, A. M. Feyerherm, and A. L. Machado, "Waiting Time in an ($S - 1, S$) Inventory System," *Opns. Res.* **23,** 674-680 (1975).
6. M. Rose, "The ($S - 1, S$) Inventory Model with Arbitrary Back-Ordered Demand and Constant Delivery Times," *Opns. Res.* **20,** 1020–1032 (1972).
7. C. C. Sherbrooke, "Waiting Time in an ($S - 1, S$) Inventory System-Constant Service Time Case," *Opns. Res.* **23,** 819–820 (1975).
8. H. C. Tijms, "Analysis of (*s,S*) Inventory Models," Math. Centre Tracts, no. 40, Amsterdam, 1971.
9. B. B. van der Genugten, "An (*s,S*)-Inventory System with Exponentially Distributed Lead Times," Research Memorandum, EIT/52, Tilburg Institute of Economics, The Netherlands.
10. J. Wijngaard, "A Direct Numerical Method for a Class of Queueing Problems," *Management Sci.* **24** (1978).

On the Estimation of Convex Functions

CHARLES A. HOLLOWAY

Stanford University, Stanford, California

(Received October 1975; accepted May 1978)

We consider the estimation of a convex or concave relationship from a set of limited observations without prior specification of a functional form. A concave programming problem is shown to provide a "best" estimate for an arbitrary norm and n independent variables. The problem is shown to be well suited to a solution using the computational strategy of relaxation (a variant of generalized programming). An example illustrates the procedure and demonstrates the relationship to a procedure for $n = 1$ suggested by Dent.

ESTIMATION of a relationship among dependent and independent variables from a limited set of observations usually involves hypothesizing a parameterized function, followed by determining parameters that optimize a selected criterion. If the hypothesized function is linear in the parameters to be estimated and the least-squares criterion is used, the resulting estimation problem is a quadratic minimization problem. Economists and others may be interested in estimating a general convex function (constant or increasing returns to scale) without desig-

0030-364X/79/2702-0401 $01.25