International Series in
Operations Research & Management Science

Steven Nahmias

Perishable Inventory Systems

Springer

International Series in Operations Research & Management Science

Volume 160

For further volumes:
http://www.springer.com/series/6161

Steven Nahmias

Perishable Inventory Systems

Springer

Steven Nahmias
Santa Clara University
OMIS Department
Santa Clara, California
USA
snahmias@scu.edu

ISSN 0884-8289
ISBN 978-1-4419-7998-8 e-ISBN 978-1-4419-7999-5
DOI 10.1007/978-1-4419-7999-5
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011928073

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

For Bill Pierskalla, who opened the door

Preface

Inventory control has emerged as a leading application of operations research. The Survey of Current Business reported that the dollar value of inventories in the USA alone exceeded $1.3 trillion at the end of 2010. Cost-effective control of inventories can cut costs significantly, and contribute to the efficient flow of goods and services in the economy. Many techniques can be brought to bear on the inventory management problem. Linear and nonlinear programing, queueing, and network flow models, are some examples. However, most inventory control packages are based on the methodology of inventory theory. Inventory theory is an important subfield of operations research that addresses the specific questions: when should an order be placed, and for how much?

Inventory theory had its roots in the well-known EOQ formula, first discovered by Ford Harris nearly 100 years ago (Harris 1915). Harris, working as a young engineer at the Westinghouse Corporation in Pittsburgh, was able to see that a simple formula for an optimal production batch size could be obtained by properly balancing holding and set-up costs. The EOQ formula, first derived by Harris, is amazingly robust – it still serves as an effective approximation for much more complex models. After Harris's work, the development of inventory theory was largely stalled until after World War II. The success of operations research in supporting the war effort was the spur needed to get the field off the ground. It seems that the newsvendor model of inventory choice under uncertainty was developed around this time, although it appears that the fundamental approach of balancing overage and underage costs under uncertainty was really first derived by Edgeworth (1888) in the context of banking.

Serious research into stochastic inventory models began around 1950. An early landmark paper was Arrow, Harris, and Marschak (1951). They were the first researchers to provide a rigorous analysis of a multiperiod stochastic inventory problem. Three significant books on the theory stimulated substantial interest in inventory theory research: Whitin (1957), Arrow, Karlin, and Scarf (1958), and Hadley and Whitin (1963). The 1960s saw an explosion of papers in inventory theory.

None of the books or hundreds of papers on inventory control written up to this time addressed an important class of problems. In every case, a tacit assumption was made that items stored in inventory had an infinite lifetime and unchanging utility. That is, once placed into stock, items would continue to have the same value in the marketplace in perpetuity. In truth, there is a very large class of inventories for which this assumption is wrong. These include inventories subject to decay, obsolescence, or perishability.

Let us define our terms. Decay (or exponential decay) means that a fixed fraction of the inventory is lost every planning period (this has also been referred to as age independent perishability). In continuous time, this translates to the size of the inventory decreasing at an exponential rate. Very few real systems are accurately described by exponential decay. For example, suppose the local grocery store discards an average of 10% of its production each day due to spoilage. In actuality though, some days it will not have to discard any product and some days it will have to discard much more than 10%. Assuming a 10% loss each day is a convenient approximation of a more complex process. Exponential decay has been proposed as a model for evaporation of volatile liquids, such as alcohol and gasoline. But how often are these substances stored in open containers, so that they would be subject to evaporation? Radioactive substances (such as radioactive drugs) are one example of true exponential decay. However, inventory management of radioactive substances is a rather specialized narrow problem. While exponential decay has been proposed as an approximation for fixed life perishables, there are better approximations.

A related problem is that of managing inventory subject to obsolescence. What distinguishes obsolescence from perishability is the following. Obsolescence typically occurs when an item has been superseded by a better version. Electronic components, maps, and cameras are examples of items that become obsolete. Notice that in each case, the items themselves do not change. What changes is the environment around them. As a result of the changing environment, the utility of the item has declined. In some cases, the utility goes to zero, and unsold items are salvaged or discarded. However, it is often the case that utility does not decrease to zero. Declining utility can result in declining demand and/or decreasing prices. For example, older electronic items, such as a prior generation of PDAs or hard drives, continue to be available for some time, but are typically sold at reduced prices. From a modeling perspective, the point at which an item becomes obsolete cannot be predicted in advance. Hence, obsolescence is characterized by uncertainty in the useful lifetime of the product.

Finally, we come to perishability. We assume the following definition of perishability throughout this monograph. A perishable item is one that has constant utility up until an expiration date (which may be known or uncertain), at which point the utility drops to zero. This includes many types of packaged foods, such as milk, cheese, processed meats, and canned goods. It also includes virtually all pharmaceuticals and photographic film. This writer's interest in this area was originally sparked by blood bank management. Whole blood has a legal lifetime of 21 days, after which time it must be discarded due to the buildup of contaminants. When uncertainty of the product lifetime is assumed, the class of items one can model is substantially larger. For example, perishable inventory with an uncertain lifetime can accurately describe many types of obsolescence.

Considering the large number of perishable items in the economy, why was this important class of problems ignored for so long? The short answer is that the problems are difficult to analyze. Interestingly, Pete Veinott, a major figure in inventory theory, wrote his doctoral thesis (in the early 1960s) on various deterministic models for ordering and issuing perishable inventories, but never published this work. When this writer inquired why, he said that the notation was so complex and awkward, and he preferred putting the work aside and move on to other problems (Veinott 1978). Van Zyl's (1964) important work on the two period lifetime case with uncertain demand remained largely unknown, as it was never published in the open literature. (This author became aware of Van Zyl's work after completing his doctoral thesis on the subject).

Contents

Chapter 1
Preliminaries

Contents

1.1 Deterministic Demand

When demand is known with certainty, the problem of managing perishables is straightforward for the most part. Consider first the basic EOQ model. Suppose the demand rate is λ, the fixed cost of placing new orders is K, and the holding cost per unit time is h. Then, it is well known that the optimal order size is

$$Q^* = \sqrt{\frac{2K\lambda}{h}}$$

and the optimal time between placement of orders is $T^* = Q^*/\lambda$. Suppose now that the item has a usable lifetime of m. All deliveries are assumed to be of fresh units only. Then, there are two cases: (a) $T^* \leq m$ and (b) $T^* > m$. In case (a) the optimal policy remains the same, since in each order cycle, all units are consumed by demand before they expire. However, in case (b) if Q^* is ordered at the beginning of the cycle, there will be positive inventory on hand at time m, which will have outdated and must be discarded, and a new order placed at that time. Notice, however, that if we reduce the order quantity from $Q^* = \lambda T^*$ to $Q = \lambda m < Q^*$, then the cycle length will remain at m, no units will expire and holding costs will be reduced, since average inventory will be reduced from $Q^*/2$ to $Q/2$. Hence, the modification of the standard EOQ model to include perishability is straightforward.

However, not all deterministic perishable inventory problems are solved so easily. In particular, consider the deterministic nonstationary production planning

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_1,

problem. Demands over an n period planning horizon are known constants, say $(r_1, r_2, \ldots, r_n)$. Costs include holding, h_i, set-up, K_i, and marginal production cost, C_i in each period. Then, Wagner and Whitin (1958) showed (in the infinite lifetime case) that an optimal policy has the following structure. If starting inventory is zero, then the order quantity in each period is either zero or exact requirements – namely, the sum of requirements in the current to some future period. Furthermore, an optimal policy only orders in periods when the starting inventory is zero. As a consequence of this result, one only needs to determine the periods in which ordering takes place, thus reducing the calculations significantly.

It turns out that an exact requirements policy may not be optimal when perishability is introduced. A counterexample is presented in Chap. 7, and methods for resolving this problem is discussed there. When demand is nonstationary, finding optimal order policies for a fixed life inventory is not trivial. However, the vast majority of the research on ordering policies for perishables has focused on stochastic demand – a significantly more difficult problem.

1.2 Periodic Review Versus Continuous Review

Demand uncertainty and (fixed life) perishability combine to result in challenging and complex problems. Stochastic perishable inventory problems fall into one of the two basic categories: periodic review or continuous review.

Most of the research in inventory theory assumes inventory levels are reviewed periodically. This means that the state of the system (on hand inventory) is known only at discrete points in time. This assumption is appropriate, for example, if inventory levels are checked once a day, once a week, etc. The landmark collection of Arrow et al. (1958) assumed periodic review in every case considered, and set the stage for much of the subsequent research on inventories. From a practice point of view, it is probably true that most inventory systems were periodic review 50 years ago. Today, however, point-of-sale scanners and automated inventory control systems have made true continuous review more common.

There are two reasons why continuous review has grown in importance. First, with automated inventory control systems computers can automatically trigger orders when inventory levels hit predetermined levels. Second, continuous review models often are able to provide simple approximations to complex problems that are difficult to solve with periodic review formulations.

Perishable inventory research has also evolved along the two separate tracks of periodic review and continuous review. The periodic review track generalizes the kind of models considered by Arrow et al. (1958), among many others, to incorporate perishability. The continuous review track is largely an outgrowth of the theory of queues with impatient servers. An impatient customer is one who leaves the queue if they have not been served by a fixed time. Queueing models with impatient customers are discussed in detail in Chap. 9.

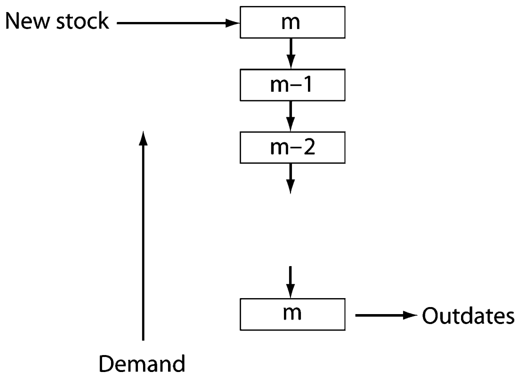
1.3 Periodic Review Preliminaries

As noted earlier, we assume that the on hand inventory level is known only at discrete points in time, which are labeled periods. Assume that periods are numbered 1, 2,.... Demands in successive periods are not known, but are assumed to be random variables, $D_1, D_2, \ldots$ with a known probability distribution. For convenience, assume that the demand distribution is continuous with cumulative distribution function (CDF) $F(x)$ and probability density function (PDF) $f(x)$. (Note that basic results have been shown to carry over to the discrete demand case as well. Also, virtually all of the results carry over to nonstationary demand. Stationarity of the demand distribution is assumed for notational convenience.)

Assume that new orders are always of fresh units that have a usable lifetime of m periods. A little reflection should convince the reader that it is necessary to track the entire age distribution of the on hand inventory in order to determine outdates each period. Hence, the system state is described by a vector $\mathbf{x} = (x_{m-1}, x_{m-2}, \ldots, x_1)$ where x_i is the number of units on hand with i useful periods of life remaining. Note that there are many notation options for the state vector. The state could be defined in terms of age rather than remaining lifetime and the vector could be numbered in order of oldest to youngest rather than vice versa, as we have done. This convention was chosen to reflect that aging occurs in the direction left to right, like the flow of the English language, and that the decision variable, y, can be equated to x_m and placed in the proper position in the vector. Note that x_0 would represent the number of units on hand that have just expired or outdated. We do not need to carry x_0 in the state vector since outdated units are assumed to leave the system. We use the convention throughout that boldface $\mathbf{x}$ is the vector of on hand inventories of each age level, and $x = \sum_{i=1}^{m-1} x_i$ is a scalar quantity representing the total on hand inventory.

The necessity to define a vector valued state variable is only one of the things that separate the perishable inventory problem from the conventional nonperishable problem. As we see, several other concerns arise as a result of perishability. One is the sequence that items are issued to meet demand. Note that there is a substantial literature on optimal issuing policies independent of the ordering problem. The appropriate assumption concerning issuing policies depends on whether the producer or consumer chooses which items satisfy demand. If the producer determines the issuing policy, it is clearly in his interest to issue items on an oldest first basis (known in accounting parlance as FIFO for first in first out). If the consumer determines the issuing policy, it is likely that the consumer will choose the freshest items, resulting in units issued in last in first out (LIFO) sequence. The vast majority of the perishable inventory literature assumes FIFO issuing, and we do so as well unless stated otherwise. Clearly, FIFO is most cost-efficient and results in minimum outdating. (A third alternative, which might be appropriate in some contexts, is to issue the items in a random order. To our knowledge, random issuing policies have not been considered in the context of optimal ordering policies for perishables.)

Fig. 1.1 The flow of demand and product in an FIFO fixed life perishable inventory system



If items are issued according to FIFO, then the aging and demand processes travel in opposite directions. To see what this means, consider the representation of the system state in Fig. 1.1. Bins are labeled $m-1$, $m-2$, . . ., 1, where the contents of bin i are the number of units on hand with i useful periods of life remaining. At the end of each period, all contents of a bin are moved to the next lower bin, and the contents of bin 1 outdate and must be discarded (or salvaged). Because of the FIFO assumption, demand depletes first from bin 1, then from bin 2, etc. Excess demands may be lost or backordered. If excess demands are backordered, then this is reflected in a negative value of x_{m-1}.

Consider now the ordering policy. The optimal number of units to order each period is function of the state vector, $\mathbf{x}$. We represent this function as $y(\mathbf{x})$. As we see, $y(\mathbf{x})$ is a complex nonlinear function of the state variable x. We assume that costs are assessed in the usual way for finite horizon periodic review inventory systems. At the end of each period, the total inventory is determined. If it is positive, assess a holding cost of h per unit held per period. If it is negative (which occurs when excess demands are backordered), assess a cost of p per unit of unsatisfied demand. Furthermore, we assume a marginal order cost only. That is, there is a cost of c per unit ordered. For now, assume that there is no fixed order cost. Finally, we come to the issue of how to assess the cost for items that must be discarded due to outdating. Let θ be the cost of disposing of outdated units. If D is the demand in a period, then the number of units outdating at the end of the period when starting inventories are $\mathbf{x}$, is $\max(x_1 - D, 0)$, which we represent as $(x_1 - D)^+$.

We now face the first issue. The astute reader will notice that the outdating cost, $\theta E(x_1 - D)^+$, is independent of the decision variable y. That means that any single period model ignores the effects of outdating. In fact, one would need to churn through at least m periods of a dynamic programing formulation before the outdating penalties of over ordering would be reflected in the optimal order policy. (This was, in fact, the approach taken in Fries 1975.)

Suppose, however, that one were interested in constructing a one period model that reflected the outdating penalties of over ordering. How could this be done? Let $D_1, D_2, \ldots$ represent demands in successive periods, starting with the current period. Then, the current order, y, would not outdate until m periods into the future, if it had not been consumed by demand by that time. Consider how one would determine the expected outdating of the current order y, m periods into the future.

Define the following sequence of random variables:

$$R_1 = y + \sum_{i=1}^{m-1} x_i$$

$$R_2 = [y + \sum_{i=2}^{m-1} x_i - (D_1 - x_1)^+]^+$$

$$R_3 = [y + \sum_{i=3}^{m-1} x_i - (D_2 + (D_1 - x_1)^+ - x_2)^+]^+$$

etc.

Then, R_i represents the amount of the current inventory on hand i periods into the future, assuming we start with $\mathbf{x}$ and order y. Each value of R_i is a random variable, since it is a function of the future demands. Note the use of the imbedded + functions are necessary to keep track of units lost due to outdating.

To simplify the notation, define the following sequence of random variables recursively:

$$\begin{aligned} B_0 &= 0 \\ B_1 &= (D_1 - x_1)^+ \\ &\vdots \\ B_j &= (D_j + B_{j-1} - x_j)^+ \text{ for } 1 \le j \le m-1. \end{aligned}$$

Interpret the random variable B_j as the total unsatisfied demand in period j after depleting the on hand inventory that would have outdated in period j. It follows that

$$R_m = [y - (D_m + B_{m-1})]^+$$

represents the amount of the current order y, that outdates in m periods.

The goal of the analysis is to compute the expected value of R_m and incorporate this into the one period model.

Define $G_n(t;\, \mathbf{w}_{n-1}) = P\{D_n + B_{n-1} \le t)$ where $\mathbf{w}_i = (x_i, x_{i-1}, \ldots, x_1)$. Note that $\mathbf{w}_{m-1} = \mathbf{x}$.

We present the first result without proof, which is based on a standard induction argument. Details can be found in Nahmias (1972).

Theorem 1.1. $G_n(t;\, \mathbf{w}_{n-1}) = \int_0^t G_{n-1}(v + x_{n-1};\, \mathbf{w}_{n-2}) f(t-v)\mathrm{d}v.$

Theorem 1.2. $E(R_m) = \int_0^y G_m(t; \mathbf{x})\mathrm{d}t$

Proof. It is well known that for any nonnegative random variable, X, the expectation may be computed two ways:

$$E(X) = \int_0^\infty x f(x)\mathrm{d}x = \int_0^\infty [1 - F(x)]\mathrm{d}x.$$

We have $P\{R_m \leq t\} = P\{y - (D_m + B_{m-1}) \leq t\} = 1 - G_m(y - t;\, \mathbf{x})$ for $t \geq 0$ Since R_m is a nonnegative random variable, the result follows from the second representation of the expected value above (after a change of variable). □

1.4 A One Period Newsvendor Perishable Inventory Model

Most readers should be familiar with the classic newsvendor model. A newsvendor must decide at the beginning of each day how many newspapers to purchase. Daily demand is not known, but is assumed to follow a known probability distribution. Let y be the number of newspapers purchased and D the demand. There are two penalties: overage (ordering too much) and underage (ordering too little).

Now, let us consider the perishable inventory model. The penalty for ordering too much is the future penalty of outdating, at θ per unit, and the penalty for ordering too little is penalty cost for excess demand, at p per unit. Hence, a sensible expected one period cost function for the perishable inventory problem is:

$$L(\mathbf{x}, y) = p \int_{x+y}^{\infty} [t - (x + y)]\, f(t)\mathrm{d}t + \theta \int_{0}^{y} G_m(t; \mathbf{x})\mathrm{d}t.$$

It is easy to show that $L(\mathbf{x}, y)$ is convex in y (and is strictly convex as long $f(t) > 0$ for all $t > 0$. Hence, the optimal order quantity, y, for this simple model satisfies:

$$\frac{\partial L(\mathbf{x}, y)}{\partial y} = -p(1 - F(x + y)) + \theta G_m(y;\, \mathbf{x}) = 0.$$

The optimal one period solution, say $y^*(\mathbf{x})$, is a nonlinear function of the entire state vector, $\mathbf{x}$. In this case, $y^*(\mathbf{x}) > 0$ for all positive real vectors $\mathbf{x}$. In addition, as we see in the analysis of the dynamic problem, $y^*(\mathbf{x})$ is decreasing in each component of the state vector, $\mathbf{x}$, but at less than unit rate.

Somewhat sharper results can be obtained when we add holding and marginal order costs.

Theorem 1.3. Suppose that in addition to penalty and outdate costs, we also include marginal order cost at c per unit ordered, and a unit holding cost, h, charged against each unit on hand at the end of the period. Then, the optimal solution has the following form: If $x < \bar{x}$ order $y^*(\mathbf{x})$ solving

$$c + hF(x + y) + p(1 - F(x + y)) + \theta G_m(y;\, \mathbf{x}) = 0$$

where $\bar{x}$ solves

$$c + hF(\bar{x}) - p(1 - F(\bar{x})) = 0.$$

If $x \geq \bar{x}$, no order is placed.

Proof. The expected one period cost function is now:

$$L(\mathbf{x},y)=cy+h\times\int_{0}^{x+y}(x+y-t)f(t)\mathrm{d}t+p\int_{x+y}^{\infty}(t-(x+y))f(t)\mathrm{d}t+\theta\int_{0}^{y}G_m(t;\mathbf{x})\mathrm{d}t.$$

Convexity in y is easy to show so that the minimizing value of y occurs, where the partial derivative of $L(\mathbf{x},y)$ vanishes. The partial derivative is given by:

$$\frac{\partial L(\mathbf{x},y)}{\partial y}=c+hF(x+y)-p(1-F(x+y))+\theta G_m(y;\,\mathbf{x})$$

thus giving the definition of the optimal ordering quantity. Notice that If $x<\bar{x}, \left.\frac{\partial L(\mathbf{x},y)}{\partial y}\right|_{y=0}<0$ and if $x\geq\bar{x}, \left.\frac{\partial L(\mathbf{x},y)}{\partial y}\right|_{y=0}\geq 0$, thus establishing that $y^*(\mathbf{x})$ is positive in the region $x<\bar{x}$ and zero in the region $x\geq\bar{x}$. $\square$

In the next chapter, we extend this approach to a multiperiod dynamic model.

Chapter 2
The Basic Multiperiod Dynamic Model

Content

Chapter 1 outlined an approach for constructing a single-period newsvendor-type model that explicitly accounts for future outdating of the current order. In this chapter, we present the extension of this one-period model to a finite horizon dynamic model. We present only the general m-period lifetime case, but the reader may be interested in reviewing the simpler case for $m = 2$, as this requires only a single-dimensional state variable.

The problem dynamics are described by the one-period transfer function. Given the current state of the system, $\mathbf{x}$, the quantity of fresh product ordered, y, and the realization of demand t, the transfer function, $\mathbf{s}(y, \mathbf{x}, t)$, gives the vector of starting inventories of the next period. The logic behind the transfer function dynamics is very similar to the logic required to derive the expected outdating function $\int_0^y G_m(t; \mathbf{x})\mathrm{d}t$. The one-period transfer function is:

$$s_i(y, \mathbf{x}, t) = [x_{i+1} - (t - \sum_{j=1}^{i} x_j)^+]^+$$

and

$$s_{m-1}(y, \mathbf{x}, t) = \begin{cases} y - (t - x)^+ & \text{if excess demand is backordered,} \\ [y - (t - x)^+]^+ & \text{if excess demand is lost.} \end{cases}$$

Note the similarity of the form of the transfer function to the definition of the sequence of random variables $B_0, B_1, \ldots,$ defined in the previous chapter. This is, of course, not coincidental. In fact, both are just two different ways of showing the system dynamics. This is shown precisely in the following result which is central to the analysis of the dynamic problem.

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_2,

Theorem 2.1. $G_n(y;\mathbf{w}_{n-1}) = \int_0^\infty G_{n-1}[s(y;\mathbf{w}_{n-1},t)f(t)\mathrm{d}t.$

Proof. The proof is somewhat tedious, but conceptually straightforward. Note that by defining $G_0(t) = 1$ for all t, the theorem easily holds for $n = 1$. One proceeds by induction, assuming that the theorem is true for $n - 1$ and showing that this leads to the theorem for n. Again, the details appear in Nahmias (1972).

2.1 The Functional Equations for the General Dynamic Problem

Following the usual approach for dynamic programming analysis, define $C_n(\mathbf{x})$ as the minimum expected discounted cost when n periods remain. Then $C_n(\mathbf{x})$ satisfies the following system of functional equations:

$$C_n(\mathbf{x}) = \min_{y\geq 0}\left\{L(\mathbf{x},y) + \alpha\int_0^\infty C_{n-1}(\mathbf{s}(y,\mathbf{x},t)f(t)\mathrm{d}t\right\},$$

which we write as

$$C_n(\mathbf{x}) = \min_{y\geq 0}\{B_n(\mathbf{x},y)\}.$$

In order to establish the existence and to define the properties of an optimal policy, the key result we need to establish is the convexity of $B_n(\mathbf{x},y)$ in y for every set of starting inventories $\mathbf{x}$. In addition, the main theorem describes several important properties of the optimal order function, $y_n(\mathbf{x})$, when n periods remain in planning horizon. The main theorem requires 17 steps and is proven via a complex induction argument.

We will not present the proof of the main theorem here, but the interested reader can refer to Nahmias (1974) for an outline of the proof or Nahmias (1972) for a detailed exposition.

It is interesting to note that the traditional approach to proving convexity of $B_n(\mathbf{x},y)$ for the standard nonperishable problem breaks down in this case. Typically, one shows that $C_n(\mathbf{x})$ is convex in $\mathbf{x}$, that convexity is preserved via the transfer function, and sums of convex functions are convex, thus easily giving the required convexity of $B_n(\mathbf{x},y)$ in the decision variable, y.

Unfortunately, this straightforward approach does not work for the perishable inventory problem. Consider the case of $m = 2$. Here, the decision variable has only a single dimension. A necessary and sufficient conditions for $C_n(x)$ to be convex are that $C_n{}''(x) \geq 0$ for all $x \geq 0$. We can demonstrate that, in fact,

$C_1''(x) < 0$ for some value of x. As is shown in Nahmias and Pierskalla (1973) (and Nahmias 1972),

$$C_1'(x) = -\theta F(x)F(y_1(x))$$

giving

$$C_1''(x) = -\theta F(y_1(x))f(x) - f(y_1(x))F(x)y_1'(x),$$

where $y_1(x)$ is the optimal order quantity when x is the (one period old) on-hand inventory, and one planning period remains in the horizon. Note that the sign of $C_1''(x)$ is not obvious, since $y_1'(x) < 0$. Consider, however, the following special case. Let us assume that the periodic demand follows the negative exponential distribution with parameter λ. That is $f(x) = \lambda e^{-\lambda x}$ and $F(x) = 1 - e^{-\lambda t}$. Since $y_1(0) > 0$ and the function $y_1(x)$ is continuous, there must exist at least one value of x, say $\tilde{x} > 0$, such that $y_1(\tilde{x}) > \tilde{x}$. Because the exponential density is monotonically decreasing in the region $x \geq 0$ and the cumulative distribution function for the exponential is monotonically increasing in this same region, we have that $F(y_1(\tilde{x})) > F(\tilde{x})$ and $f(y_1(\tilde{x})) < f(\tilde{x})$. In addition, it has been shown that $y_1' \times (\tilde{x}) \geq -1$. Combining these results gives $C_1''(\tilde{x}) < 0$. Hence, we conclude that $C_1(x)$ is not convex in x. However, it turns out that the *degree* of nonconvexity (as measured by a lower bound on $C_1''(x)$) is not very great, and we can show that $B_2(x, y) = L(x, y) + \alpha \int_0^\infty C_1(s(y, x, t)f(t)\mathrm{d}t$ is convex in y. That is, the nonconvexity of $C_1(x)$ is more than compensated for by the convexity of $L(x, y)$. For the general m-period problem, it is the convexity of $B_n(\mathbf{x}, \mathrm{y})$ in y that allows us to establish the existence and basic properties of the optimal order function, $y_n(\mathbf{x})$.

We will assume the following notational convention. For any vector valued function, $g(\mathbf{x})$, $g^{(i)}(\mathbf{x})$ is the first partial derivative of g with respect to the ith variable, and $g^{(i,j)}(\mathbf{x})$ is the second partial derivative with respect to the ith and jth variables, respectively.

In the general m-period lifetime problem, the key result that allows us to prove convexity of $B_n(\mathbf{x}, y)$ is $C_n{}^{(1,1)}(\mathbf{x}) \geq -\theta G_{m-1}{}^{(1)}(\mathbf{x})$ (where the differentiation is done with respect to the first variable in the vector $\mathbf{x}$, which is x_{m-1}). Establishing the validity of this lower bound via induction is extremely complex, requiring a network of inequalities on the first and second partial derivatives of the optimal return functions, $C_n(\mathbf{x})$. To provide the reader with an appreciation of the complexity of this problem, we provide a complete statement of the theorem required to prove convexity. As noted, the proof will not be presented here.

Theorem 2.2. Assume that demands in each period form a sequence of independent identically distributed random variables (although the theorem also holds for nonstationary demands) and that:

(a) The demand distribution, F, possesses a bounded continuous density f with the property that $f(t) > 0$ if $t > 0$ and $f(t) = 0$ if $t < 0$.

(b) Future costs are discounted by a discount factor α, where $0 < \alpha < 1$. Then:

1. $B_n(\mathbf{x}, y)$ is convex in y for all $\mathbf{x} \in R_{m-1}$ and is strictly convex in a neighborhood of the global minimum.
2. $\lim_{y\to 0} \frac{\partial B_n(\mathbf{x},y)}{\partial y} < 0$ and $\lim_{y\to\infty} \frac{\partial B_n(\mathbf{x},y)}{\partial y} > 0$ for all $\mathbf{x}$.
3. There is a unique function $y_n(\mathbf{x})$ given as the solution to $\left.\frac{\partial B_n(\mathbf{x},y)}{\partial y}\right|_{y=y_n(\mathbf{x})} = 0$ and $0 < y_n(\mathbf{x}) < \infty$. In addition $y_n{}^{(i)}(\mathbf{x})$ exists and is continuous for all $\mathbf{x}$, $1 \le i \le m-1$.
4. $$C_n^{(i)}(\mathbf{x}) = -\theta \sum_{j=1}^{i} G_{m-j}(\mathbf{x}(m-j)) H_j(y_n(\mathbf{x}); \bar{\mathbf{x}}(m-j)) + \alpha \sum_{j=1}^{m-i} \int_{w_{j-1}}^{w_j} \{C_{n-1}^{(i+1)}[\mathbf{z}_j(t)]$$
$$- C_{n-1}^{(1)}[\mathbf{z}_j(t)]\} f(t)\mathrm{d}t + \alpha \sum_{j=m-i+1}^{m-1} \int_{w_{j-1}}^{w_j} \{C_{n-1}^{m-j+1}[\mathbf{z}_j(t)] - C_{n-1}^{(1)}[\mathbf{z}_j(t)]\} f(t)\mathrm{d}t,$$
where $\mathbf{z}_j(t) = (y, x_{m-1}, ..., x_{j+1}, \sum_{i=1}^{j} x_i - t, 0, ...0)$ and $w_j = \sum_{i=1}^{j} x_i$, and $C_n^{(m)}(\mathbf{x}) \equiv 0$. The result holds for $1 \le i \le m-1$.
5. $-1 \le y_n^{(1)}(\mathbf{x}) \le y_n^{(2)}(\mathbf{x}) \le \cdots \le y_n^{m-1}(\mathbf{x}) < 0$.
6. (a) $C_n^{(i,k)}(\mathbf{x})$ exists and is continuous for all $\mathbf{x} \in R^{m-1}$ and $1 \le k, i \le m-1$. However, $C_n^{(1,1)}(t; \mathbf{x}(m-2))$ will be discontinuous at $t = 0$ whenever $f(t)$ is discontinuous at $t = 0$.
   (b) $C_n^i[\bar{\mathbf{x}}(m-i), \mathbf{0}] - C_n^{i-1}[\bar{\mathbf{x}}(m-i), \mathbf{0}] = 0$ for $2 \le i \le m-1$. The notation is meant to be interpreted as the last $m-i$ components being zeros.
7. (a) $C_n^{(1,j)}(\mathbf{x}) \ge -\theta G_{m-1}^{(j)}(\mathbf{x}) \quad 1 \le j \le m-1$.
   (b) $C_n^{(i,j)}(\mathbf{x}) - C_n^{(i-1,j)}(\mathbf{x}) \ge -\theta G_{m-i}^{(j+i+1)}(\mathbf{x}(m-i))[1 - \sum_{k=1}^{i-1} H_k\ (x_{m-i+k'}, \ldots, x_{m-i+1})]$ for $m-1 \ge j \ge i \ge 1$.
   (c) $C_n^{(1,i)}(\mathbf{x}) - C_n^{(1,i-1)}(\mathbf{x}) \le \theta[G_{m-1}^{(i-1)}(\mathbf{x}) - G_{m-1}^{(i)}(\mathbf{x})]$ for $m-1 \ge i \ge 2$.
   (d) $$[C_n^{(i,j)}(\mathbf{x}) - C_n^{(i-1,j)}(\mathbf{x})] - [C_n^{(i,j-1)}(\mathbf{x}) - C_n^{(i-1,j-1)}(\mathbf{x})] \le \theta[G_{m-i}^{(j-i)}(\mathbf{x}(m-i))$$
   $$-G_{m-i}^{(j-i+1)}(\mathbf{x}(m-i))][1 - \sum_{k=1}^{i-1} H_k(x_{m-i+k}, \ldots, x_{m-i+1})]$$
   for $m-1 \ge j > i \ge 2$.
8. (a) $$-\theta \sum_{j=1}^{i} G_{m-j}(\mathbf{x}(m-j))[1 - \sum_{k=1}^{j-1} H_k(x_{m-j+k}, \ldots, x_{m-j+1})] \le C_n^{(i)}(\mathbf{x}) \le 0$$
   for $1 \le i \le m-1$ and for all $\mathbf{x}$.
   (b) $C_n^{(i)}(\mathbf{x}) - C_n^{(j)}(\mathbf{x}) \le \theta \sum_{k=j+1}^{i} G_{m-k}(\mathbf{x}(m-k))$
   $[\sum_{q=k-j}^{k-1} H_q(x_{m-k+q}, \ldots, x_{m-k+1})]$ for $1 \le j < i \le m-1$ and for all $\mathbf{x}$.

(c) $C_n^{(i)}(\mathbf{x}) - C_n^{(j)}(\mathbf{x}) \geq -\theta \sum_{k=j+1}^{i} G_{m-k}(\mathbf{x}(m-k))[1 - \sum_{q=1}^{k-1} H_q(x_{m-k+q}, \ldots, x_{m-k+1})]$
for $1 \leq j < i \leq m-1$ and for all $\mathbf{x}$.

(d) $C_n^{(i)}(\mathbf{x}) = 0$ for $\mathbf{x} = (x_{m-1}, \mathbf{0})$ and $x_{m-1} \leq 0$.

9. (a) $\lim_{x_i \to \infty} y_n(\mathbf{x}) = 0 \quad 1 \leq i \leq m-1.$

(b) $\lim_{x_j \to \infty} C_n^{(i)}(\mathbf{x}) = 0 \quad 1 \leq i, j \leq m-1.$

Note that the functions H_k referred to in steps 7 and 8 are used as a convenience for representing derivatives of the outdating function, G_m. As they add nothing to the exposition, we will not discuss them further here. Aside from all of the machinery involving the derivatives of the optimal value functions, what does the theorem tell us about the behavior of the optimal policy function? The key piece of information we obtain from the theorem is step 5: $-1 \leq {y_n}^{(1)}(\mathbf{x}) \leq {y_n}^{(2)}(\mathbf{x}) \leq \cdots \leq {y_n}^{(m-1)}(\mathbf{x}) < 0$. This says two things. First, since all partial derivatives are negative, the optimal order quantity decreases as starting inventories increase. More importantly, it characterizes the sensitivity to starting stocks of different ages. The larger the derivative of $y_n(\mathbf{x})$ in absolute value, the greater the sensitivity of the optimal order function to changes in starting stock. This means that increasing the on-hand quantity of newer stock has a larger effect on optimal order quantities than increasing the on-hand quantity of older stock.

This is a fundamental property of perishable inventory systems that separates such systems from traditional nonperishable systems. One is concerned not only with the amounts of on-hand inventory, but also their ages as well. Because the dimension of the state variable is proportional to the lifetime of the stock in periods, computing an optimal policy is feasible only for relatively short lifetimes. One quickly faces the "curse of dimensionality" that plagues many dynamic programming formulations. For product lifetimes much more than two or three periods, it is unlikely one would use optimal policies. Also, implementation of optimal policies would be complicated by the fact that one needs to keep track of the age distribution of stock. Approximations that depend only on the total stock on-hand are of interest as they are easy to compute and easy to implement. Methods of finding simple approximations for the periodic review problem will be the subject of the next chapter.

An interesting question is whether or not these results hold when demand is discrete rather than continuous, as is assumed in Theorem 2.2. To try to prove convexity of the functions $B_n(\mathbf{x}, y)$ directly under discrete demand would be extremely tedious, if even possible. To circumvent these difficulties, Nahmias and Schmidt (1986) considered a very novel approach to the discrete demand problem. They considered an infinite sequence of continuous demand distributions, F_1, F_2,. . ., that converged weakly to the discrete distribution of demand, F. We know all of the results of Theorem 2.2 hold for each of the continuous distributions F_1, F_2,. . .. Without going into the mathematical details, the authors show that the essential results of Theorem 2.2 carry over in the limit for the discrete case. We do not believe that this approach has been used before or since in the context of a dynamic inventory problem.

It should be noted before closing this section that Fries (1975) also provided a rigorous analysis of the perishable inventory problem, but did not work with the outdate function $G_m(y, \mathbf{x})$. Instead, he developed a straightforward dynamic programming formulation that required m periods before the effects of outdating appeared in the optimal order function. This approach is equally valid as ours outlined here, and preferred for computing optimal policies, as the functional equations are somewhat simpler. As Nahmias (1977a) shows, the two methods give the same policy if one is sufficiently far from the end of the planning horizon, and the discount factor is adjusted in a suitable fashion.

Chapter 3
Extensions of the Basic Multiperiod Dynamic Model

Contents

The basic model treated in Chap. 2 is the building block for several extensions. In many cases, these extensions are suggested by specific applications. Other extensions provide a wide range of applications of the basic theory.

3.1 Random Lifetime

A natural extension of the basic model is to allow for uncertainty in the product lifetime. The basic model is applicable for products with a known expiration date, since the time of outdating is known in advance. However, there are many classes of items whose useful lifetime cannot be predicted in advance. Fresh produce, meat, fowl, and fish would fall into this category. Even in blood banking, local blood banks may receive transfers from other banks that are not completely fresh, making the remaining lifetime uncertain. Perishable inventory models with uncertainty in the product lifetime could be used to model some classes of items subject to obsolescence as well.

Providing an accurate model of lead time uncertainty has long been difficult problem in inventory management (see Hadley and Whitin (1963), pp. 200–204, for a discussion of the issues involved.) The main issue is order crossing. That is, are orders placed on Tuesday allowed to arrive before orders placed on Monday? If successive lead times are independent random variables, then order crossing is inevitable. However, if one is placing orders with the same supplier, then order crossing does not make sense. If we do not allow order crossing, then successive lead times are dependent random variables, making the analysis very complex.

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_3,

The simpler results for (Q,R) models with stochastic lead times (such as the formula presented in Hadley and Whitin on page 153), are based on simultaneously assuming independence of lead times and no order crossing.

As noted, if lead times are random and orders do not cross, then successive lead times are dependent random variables. Allowing for dependency is difficult, and is rare in operations research. A very innovative approach to this problem was developed by Kaplan (1970). Let $A_1, A_2, \ldots$ be a sequence of IID discrete random variables defined on the set $\{1, 2, \ldots, \tau\}$ where τ is the maximum value of the lead time in any period. In any period, when the event $\{A_i = k\}$ is realized, all orders k or more old periods arrive in the next period.

Nahmias (1977c) adapted this approach to the perishable inventory problem. Here, one assumes that if $\{A_i = k\}$ is realized, then all on hand inventory remaining in stock (after satisfying demand) that is at least k periods old at the end of period i will outdate at that time. The random variables $A_1, A_2, \ldots$ are *not* realizations of the lifetimes of successive orders, but the lifetime distribution can be derived from them.

Suppose that $P_k = P\{A_n = k\}$ independent of n. If the event $\{A_n = k\}$ is realized, then the one-period transfer function $\mathbf{s}(y, \mathbf{x}, t, k) = (s_{m-1}(y, \mathbf{x}, t, k), \ldots, s_1(y, \mathbf{x}, t, k))$ is:

for $k = 1$:

$$\begin{aligned} s_j(y, \mathbf{x}, t, 1) &= 0 \qquad \text{if } 1 \le j \le m-2 \\ s_{m-1}(y, \mathbf{x}, t, 1) &= -(t - x - y)^+ \end{aligned}$$

and for $2 \le k \le m$:

$$s_j(y, \mathbf{x}, t, k) = \begin{cases} 0 & \text{if } 1 \le j \le m-k \\ (x_{j+1} - (t - w_j)^+)^+ & \text{if } m-k+1 \le j \le m-1 \\ y - (t - x)^+ & \text{if } j = m-1 \end{cases}$$

where $w_j = \sum_{i=1}^{j} x_i,\ x = \sum_{i=1}^{m-1} x_i.$

The sequence of events in a period is given as follows. First, the current state is observed, and based on that an order is placed for fresh stock. After the order arrives, the demand for the current period is realized, which then is satisfied to the extent possible on an FIFO basis. After demand is either satisfied or backordered (or lost), the aging process random variable, A_n is realized, and outdating determined. To understand the mechanics of the transfer function, consider the following example. Fix the current period at period 10. Suppose that $m = 4$, and the starting state in period 10 is $\mathbf{x} = (18, 12, 20)$ (i.e., 18 units remain from the order delivered one period ago, 12 units remain from the order delivered two periods ago, and 20 units remain from the order delivered three periods ago). Assume that the order quantity of fresh stock in period 10 is 23 units and the realization of demand is 28 units.

Consider how the system state evolves given all of the possible realizations of the aging process random variable, A_{10}. By FIFO, all of the 20 units are consumed first, followed by 8 units of the 12 units of two-period old stock. The vector of on hand inventories in period 10 after the arrival of fresh stock and the realization of the demand is (23, 18, 4, 0). If $A_{10} = 4$, there is no outdating, since all inventories of age 4 was depleted by demand, and the starting state in period 11 is (23, 18, 4). If $A_{10} = 3$, then the on hand inventory ordered three periods ago will outdate (4 units), and the starting state in period 11 is (23, 18, 0). If $A_{10} = 2$, then there will be 18 + 4 = 22 units outdating at the end of period 10, and the starting state in period 11 is (23, 0, 0), and finally if $A_{10} = 1$ is realized all 45 units on hand outdate, and period 11 starts with zero inventory.

The functional equations defining an optimal policy are similar to those in Chap. 2, and are now given by:

$$C_n(\mathbf{x}) = \min_{y\geq 0}\{cy + L(x+y) + \theta\sum_{k=1}^{m} P_k \int_0^{w_{m-k+1}} (w_{m-k+1} - t)f(t)\mathrm{d}t$$
$$+ \alpha\sum_{k=1}^{m} P_k \int_0^{\infty} C_{n-1}\,\mathbf{s}(y, \mathbf{x}, t, k)f(t)\mathrm{d}t\}.$$

The optimal policy when n periods remain in the horizon, $y_n(\mathbf{x})$, possesses virtually all of the same properties of the optimal policy discussed in Chap. 2 for the deterministic problem. The full statement of the appropriate theorem for analyzing the optimal policy can be found in Nahmias (1977c). Note that these functional equations charge the outdating cost in the current period rather than m periods into the future when the current order outdates (similar to Fries 1975). This was done to make the dynamic problem tractable, for as we see, computing the expected outdating of the current order is a very complex problem in itself.

If the lifetime of successive orders were independent, it would not be possible to compute the expected outdating of the current order, since future orders could outdate before the current order. However, due to our assumption that orders outdate in the same sequence that they enter stock, we are guaranteed that this will not occur. Hence, given a current order y and state vector $\mathbf{x}$, it should be possible to determine the expected outdating of the current order (which could be 1, 2,. . ., m periods into the future).

To understand the mechanics of the outdating process when the product lifetime is uncertain, consider another simple example. Let $j_1, j_2, \ldots$ correspond to successive realizations of the lifetime process, A_n. Suppose that $m = 5$, $j_1 = 4$, $j_2 = 4$, $j_3 = 2$, and the starting state in period 1 is (x_4, x_3, x_2, x_1). The goal is to compute the expected outdating of the order placed in period 1, y. Based on the realizations of the aging process, the remaining amounts of x_1 and x_2 (after satisfying demand in period 1) outdate at the end of period 1. The remaining amount of x_3 on hand at the end of period 2 after satisfying demand outdates at that time, and the remaining amounts of y and x_4 outdate at the end of period 3. The expected amount of $y + x_4$ outdating at the end of period 3 is given by $\int_0^{y+x_4} G_3(u, x_3, x_1 + x_2)\mathrm{d}u$ so that the expected outdating of y only at the end of period 3 is $\int_0^{y+x_4} G_3(u,\, x_3,\, x_1 + x_2)\mathrm{d}u - \int_0^{x_4} G_3(u,\, x_3,\, x_1 + x_2)\mathrm{d}u = \int_{x_4}^{y+x_4} G_3(u,\, x_3,\, x_1 + x_2)\mathrm{d}u$. This simple example should give the reader a glimpse of the complexities in the general calculation.

Define $r_k = \sum_{i=1}^{k} P_i$. Then, it can be shown that the expected outdating of the current order for y units is

$$\sum_{k=1}^{m} r_k \sum_{j_1=2}^{m} \cdots \sum_{j_{k-1}=k}^{m} \prod_{i=1}^{k-1} P_{j_i} \int_{v_k}^{y+v_k} G_k(u, \mathbf{v}(k-1))\mathrm{d}u,$$

where the vector terms $\mathbf{v}(k-1)$ are appropriate partial sums of the state vector $\mathbf{x}$ based on the realizations of the aging process, A_n. This is a very tedious calculation and it is unlikely that one would resort to trying to find an optimal policy except for small values of m. However, an explicit representation of the expected outdating of the current order can be used to construct simple approximations, similar to the manner that approximations were constructed for the deterministic lifetime problem.

The essence of the calculation is to find an expression for the unconditional probability that any new order has a lifetime of exactly k periods for all $1 \leq k \leq m$. We call this probability q_k. Based on the definition of the aging process, it follows that

$$q_k = \sum_{k=1}^{m} r_k \sum_{j_1=2}^{m} \cdots \sum_{j_{k-1}=k}^{m} \prod_{i=1}^{k-1} P_{j_i}.$$

It is not difficult to show that this expression can be simplified to:

$$q_k = r_k \prod_{i=1}^{k-1} (1 - r_i).$$

The idea is to use an approximate expression for the expected outdating of the current order, y, and utilize this expression to approximate the expected one-period cost function as well as the one-period transfer function, much as was done in Chap. 3 for the fixed life case. Again, one can use either the Nahmias's bounds or the Chazan and Gal bounds. Calculations reported in Nahmias (1977c) indicate that the Chazan and Gal bounds give slightly better results, and appear to improve as m increases while Nahmias's bounds degrade as m increases.

Using the Chazan and Gal bounds, the myopic expected one-period cost function one obtains is:

$$W(z) = c(1-\alpha)z + L(z) + \sum_{k=1}^{m} q_k(\alpha^{k-1}\theta + \alpha c)\mu_k(z).$$

Since $W(z)$ is quasi-convex in z, the minimizing value occurs where the first derivative is zero so that z^* satisfies:

$$c(1-\alpha) + L'(z) + \sum_{k=1}^{m} q_k(\alpha^{k-1}\theta + \alpha c)\left\{\frac{F(z/k) + F^{k^*}(z)}{2k}\right\} = 0.$$

Comparisons of this approximation with the optimal policy for lifetimes of 2 and 3 periods, and a variety of costs and demand distributions shows close agreement with the optimal.

3.2 Inclusion of a Set-Up Cost

All of the work reviewed thus far on the periodic review perishable inventory problem assumes there is no fixed cost (set-up cost) for placing an order. Fixed order costs can rarely be ignored in practice. In rare occasions, when deliveries are pre-scheduled on a daily or weekly basis, fixed order costs are sunk costs. However, it is more common that fixed costs determine the optimal frequency of orders, and cannot be ignored.

The results of this section are based on Nahmias (1978). Both optimal and approximate ordering policies are considered. The optimal policy is derived only for the one-period problem, but it is likely the multiperiod version possesses the same policy structure. The optimal policy requires calculation of two functions of the starting state, $s(\mathbf{x})$ and $y(\mathbf{x})$ with $s(\mathbf{x}) \le y(\mathbf{x})$. These functions determine the optimal policy as follows: if $s(\mathbf{x})>0$, the optimal policy is to place an order for $y(\mathbf{x})$, and if $s(\mathbf{x}) \le 0$, then no order should be placed. As in the case of no set-up cost, $y(\mathbf{x})$ minimizes the expected one-period cost function, $B(\mathbf{x}, y)$. The function $s(\mathbf{x})$ solves $B(\mathbf{x}, s(\mathbf{x})) = B(\mathbf{x}, y(\mathbf{x})) + K$, and $s(\mathbf{x}) \le y(\mathbf{x})$. Since $B(\mathbf{x}, y)$ is convex in y, these equations uniquely define $s(\mathbf{x})$. Note that Nahmias is able to derive only first period results, but given the fact that the structure of the optimal policy is the same for the first period problem as the multiperiod problem when there is no set-up cost present, it is reasonable to speculate that this is the case here as well.

Except for product lifetimes of two or three periods, the optimal policy is too complex to be practical. Hence, approximations are of particular interest. The natural form of an approximate policy is (s, S), since this policy is known to be optimal for the nonperishable problem, and is easy to implement in practice. (An (s, S) policy is one in which an order is placed up to S if and only if the starting inventory position is less than s.) The approach to constructing an approximation follows closely the method outlined in Chap. 3 for the case of $K = 0$. That is, the expected outdate function is approximated by a function of the total on hand inventory using either the Nahmias or Chazan and Gal bounds, and this approximation is then used to construct a surrogate problem whose optimal solution is used to approximate the original problem. The surrogate problem has only a single dimensional state variable, and is solved by dynamic programing.

Extensive computations reported in Nahmias (1978) indicate that the approximation generally gives costs within 1% of the optimal, although because of the difficulty of computing an optimal policy, the comparisons were done for $m = 2$ only. However, the results would seem to indicate that this is a valid approach for approximating this problem.

3.3 Multiproduct Models of Perishables

Blood bank applications, in particular, have given rise to several multiproduct perishable inventory models. The earliest was explored by Nahmias and Pierskalla (1976). Blood banks store both whole blood and frozen packed red blood cells. Since frozen red blood cells have a lifetime of a year or longer, one can assume that the usable lifetime of frozen blood is infinite. Although frozen blood can be used in place of fresh blood when necessary, one would only want to use the frozen blood as a last resort. This is because the cost of extracting the red blood cells, freezing them, and then subsequently thawing them and combining them with plasma is considerably more expensive than the cost of extracting and storing fresh blood.

The model developed is a direct extension of the approach applied by Nahmias (1975a). Assume that at the beginning of each planning period prior to the realization of demand, orders for both product 1 (the perishable product) and product 2 (the nonperishable product) must be placed. There is a single demand source that depletes first from product 1 according to FIFO, and then product 2 until the demand is either satisfied or backordered. (Note: in the context of blood banking, it is probably more accurate to assume that excess demand is lost rather than backordered. This would mean that excess demand is made up by emergency shipments from other blood banks. The backorder assumption is not crucial to the analysis.)

Costs are charged separately for holding at h_1 for product 1 and h_2 for product 2. Outdating of product 1 is charged at θ and there is a penalty cost of p for unsatisfied demand. Let $\mathbf{x}$, x, and y be defined as earlier for the single product case. Define x^2 as the on hand inventory of product 2 at the start of any period, and let z be the on hand inventory of product 2 after ordering in any period so that the order quantity of product 2 is $z - x^2$. Then, the one-period expected cost function has the form:

$$L(\mathbf{x}, y, z) = c_1 y + h_1 \int_0^{x+y} (x + y - t) f(t)\mathrm{d}t + \theta \int_0^{y} G_m(u;\, \mathbf{x})\mathrm{d}u + c_2(z - x^2)$$
$$+ h_2 z F(x + y) + h_2 \int_{x+y}^{x+y+z} (x + y + z - t) f(t)\mathrm{d}t + p \int_{x+y+z}^{\infty} (t - x - y - z) f(t)\mathrm{d}t.$$

Assume the following relationships among the cost parameters:

(i) $0 \le h_2 \le h_1$

(ii) $0 < c_1 < c_2$

(iii) $p > (1 - \alpha)c_2$

(iv) $0 \le (1 - \alpha)(c_2 - c_1) + (h_2 - h_1) \; < \; \theta.$

The rationale for these assumptions is given as follows. Assumption (1) says that the cost of storing the nonperishable product is at least as large as the cost of storing the perishable product. Assumption (2) says that the cost of acquiring the nonperishable product exceeds the cost of acquiring the perishable product. Assumption (3) is common for single product inventory systems. It says that the unit cost of excess demand exceeds the one-period discounted cost of acquiring new nonperishable product. This guarantees that it is not more economical to incur stockouts than purchase new stock – otherwise, there would be no motivation to operate the system. The expression in the final case corresponds to the cost of acquiring one unit of product 2 and salvaging it one period later minus the cost of acquiring one unit of product 1 and salvaging it one period later plus the difference in the holding costs. If this term were negative, one would never order product 1. If this term exceeded the outdate cost, it would never be optimal to order product 2.

The functional equations defining an optimal policy for the multiperiod version of the problem are

$$C_n(\mathbf{x}, x^2) = \min_{y \ge 0,\, z \ge x^2} \left\{ L(\mathbf{x}, y, z) - c_2 x^2 + \alpha \int_0^{\infty} C_{n-1}(\mathbf{s}_1(\mathbf{x}, y, t), s_2(x + y, z, t)) f(t)\mathrm{d}t \right\}.$$

In this case, the transfer function is m dimensional. The first $m - 1$ components, $\mathbf{s}_1(\mathbf{x}, y, t)$, are essentially the same as the single product form. The mth component is $s_2(x + y, z, t) = z - (t - (x + y))^+$ in the backorder case and $s_2(x + y, z, t) = [z - (t - (x + y))^+]^+$ in the lost sales case.

The structure of the optimal policy is shown in Fig. 3.1. When $m = 2$, the state vector is two dimensional, so the optimal policy can be represented graphically.

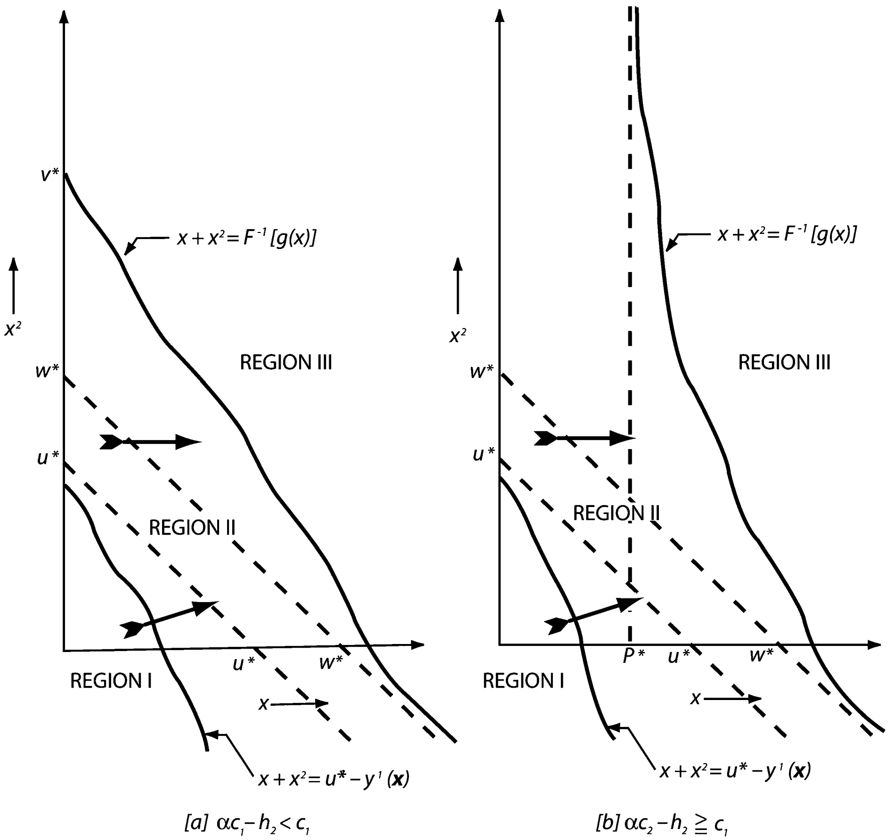


Fig. 3.1 The optimal ordering regions for a two-product perishable/nonperishable problem

Note that figure includes two cases: (a) where $\alpha c_1 - h_2 < c_1$ and (b) where $\alpha c_1 - h_2 \geq c_1$. The structure of the optimal policy is essentially the same in both cases, but the boundary between Regions II and III differ slightly. Region I corresponds to ordering both the perishable and nonperishable products. In Region II, one orders only the perishable product, and in Region III, one does not order. Note that in Region I, one orders so that the total stock on hand after ordering is u^* where

$$u^* = F^{-1}\left[\frac{p - c_2(1-\alpha)}{p + h_2}\right].$$

What is unusual is that the line $x + x^2 = u^*$ lies completely outside of Region I. Typically, when there are ordering regions in a multiproduct inventory problem, one orders to the boundary of the region. In Region I, however, one orders to the interior of Region II. Just the opposite is true in Region II. When ordering only the perishable product, one orders so that one remains short of the boundary between Regions II and III. This unusual behavior of the optimal policy is a direct result of the perishability of product 2. If both products were nonperishable, then the boundary between Regions I and II would be the line $x + x^2 = u^*$. However, the perishability adds a penalty to orders of product 2, which has the effect of tightening the boundary between Regions I and II. In the same way, when one is in Region II, the additional penalty of perishability results in a decreased order size so that one does not order to the boundary of Regions II and III. Note that there an apparent instability in Region II. Since one orders short of the boundary, it implies that one would want to order again, since after ordering the total stock on hand still lies within Region II. However, note that this is not exactly the case, since the state consists of one-period old inventory, and the order quantity consists of fresh inventory.

Note that Nahmias (1982) shows how the approximation techniques discussed in Chap. 3 can be applied to this problem. Although these approximations have never been tested, the success of the approximations in the single product case suggests that this approximation should be effective as well.

Deuermeyer (1979) considered a multiproduct perishable inventory model also suggested by a blood banking application. Consider two production processes, A and B. Process A produces two products (1 and 2) while Process B produces only product 2. Both products are perishable, and both face independent stochastic demands. In the blood banking application, product 1 is platelets (a blood component) and product 2 is whole blood. Process A separates some of the whole blood into platelets and packed red cells while Process B just produces whole blood.

Suppose that the lifetime of product 1 is m periods and the lifetime of product 2 is l periods. The state vector consists of the starting stocks of each product at each age level, represented as $(\mathbf{x}, \mathbf{w})$ having dimension $(m - 1) + (l - 1)$. The decision variable is the pair (a, b), which corresponds to the production quantities for Process A and Process B, respectively. Assume that a units of Type A production yields $\eta_1 a$ of product 1 and $\eta_2 a$ of product 2 while b units of Type B production yield b units of Product 2.

Deuermeyer develops a one-period model using the expected future outdating of both products in the one-period cost function as we did in Chap. 2. This makes it likely that the structure of the optimal one-period policy is the same as the structure of the optimal multiperiod policy. (Unfortunately, due to the complexity of the problem, he was able to obtain first period results only.)

The optimal policy structure is pictured in Fig. 3.2 when $m = l = 2$. The policy is defined by four regions in the $(\mathbf{x}, \mathbf{w})$ space. In Region I, both processes are used. The graph pictures the increase in the inventories of both products first due to Process A and then the increase in product 2 inventory due to Process B. In Region II, one only applies Process A (which increases the inventories of both products) while in Region III one applies only Process B so that only product 2 inventory is increased. Finally, in Region IV, neither process is applied.

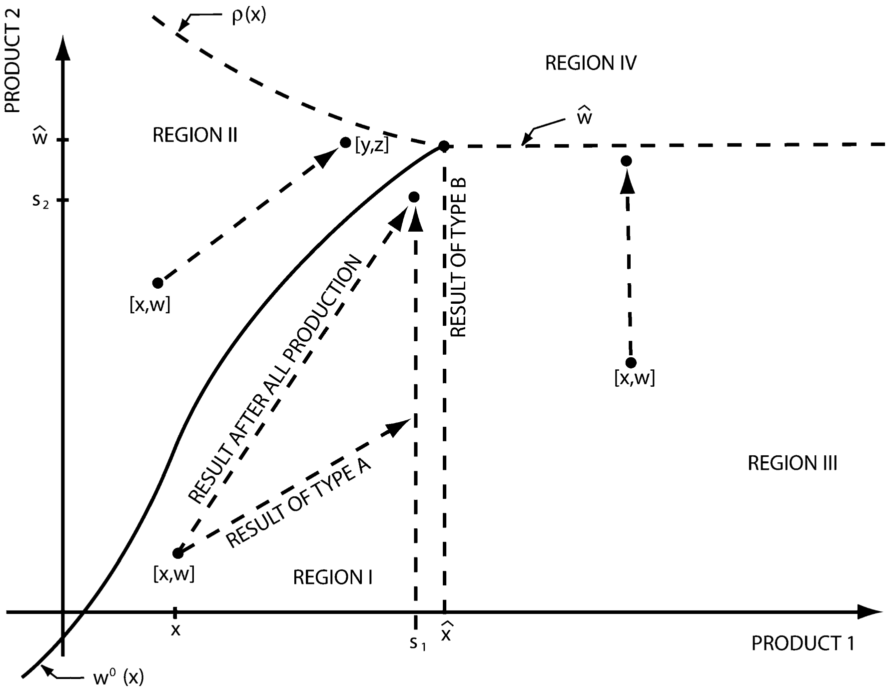


Fig. 3.2 Characterization of the optimal policy for the Devermeger two-product model

Deuermeyer (1980) considers a system consisting of n products all of which are perishable with respective lifetimes of l_i for $1 \leq i \leq n$, and assumes that the system has the substitution property. That is, the order quantity of each product is a nonincreasing function of the on hand inventories of the other products. Again, he is able to characterize the optimal policy for a single-period problem only.

Chapter 4
Continuous Review Perishable Inventory Models

Contents

Continuous review means that the state of the inventory system is known at all times, as opposed to periodic review where the state of the system is known only at discrete points in time labeled periods. Surprisingly, very little is known about optimal ordering policies for fixed life perishables under continuous review. The difficulty is that when an order lead time is present, aging is applied only to the units on hand, and not to the units on order. Furthermore, since units may be ordered at any point in time, there is no limit to the number of different orders that comprise the on-hand inventory at any point in time. Hence, the vector describing the on-hand inventory of each age level has an unlimited number of dimensions.

One way to circumvent this problem is to assume zero order lead time. We briefly review the two papers that take this approach. Before doing so, note that we believe that the simultaneous assumptions of continuous review and zero order lead time are unrealistic, and it is likely that these models have little practical use. There is a big difference between assuming zero order lead time in periodic review systems and in continuous review systems. In the former case, it only means that the lead time is less than a review period, as orders are assumed to be placed at the beginning of the planning period, and are assumed to arrive at the end of the planning period. In the case of continuous review, it means that orders arrive instantaneously. For that reason, the policies obtained are not likely to be useful in a practical setting.

It is worth noting that the primary driver of the continuous review heuristic (Q, R) inventory models that form the basis for most commercial inventory control systems is uncertainty of demand over the replenishment lead time. Where lead

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_4,

times reduced to zero, optimal order quantities would collapse to the simple EOQ formula and reorder levels would be zero. (See Nahmias (2009), Chap. 5, for example).

When there is a positive lead time for ordering, the form of an optimal order policy for perishables under continuous review is not known. Part of the problem is defining the state of the system. As noted above, the state could be an infinite dimensional vector of all previous orders placed and their ages. Given the complexity of the continuous review problem, a reasonable starting point is to find the best policy from a prespecified class of policies.

4.1 One for One (*S*–1, *S*) Policies

The only rigorous analysis of a continuous review perishable inventory system with positive order lead time of which we are aware is Schmidt and Nahmias (1985) (and an extension by Perry and Posner (1998)). They assumed the so-called $(S-1, S)$ policy in which the inventory position (stock on hand plus stock on order) is maintained at a fixed level S. Assume that demands occur one-at-a-time, which would be the case, for example, if demands were generated by a stationary Poisson process. For nonperishables, a $(S-1, S)$ policy places an order for one unit at each occurrence of a demand. With perishability, orders are placed at both the occurrence of demands and outdating. $(S-1, S)$ policies are optimal for very high value items, and are common in military resupply and repair systems. See Nahmias (1981) for a review of models for repairable item systems. A more comprehensive and up-to-date discussion of these systems can be found in Muckstadt (2005)). It is common in both military and civilian applications that equipment goes in for maintenance either when the equipment fails, or at fixed intervals, whichever comes first. Since unplanned failures occur at random, these can be modeled as a stochastic process, and may be labeled the demand process. If an item reaches age m and has not failed (i.e., been demanded), the system for routine maintenance is taken out, and thus "outdates." Hence, the extension of traditional $(S-1, S)$ policies to the case of fixed life perishable inventories, is potentially a very useful extension for modeling maintenance systems.

As noted above, what makes the problem difficult is the interaction of perishability and the order lead time. Assume that demands, (i.e., failures), occur one-at-a-time completely at random according to a stationary Poisson process with rate λ. Costs are charged in the usual way against proportional ordering at c per unit, holding at h per unit held per unit time, outdating at θ per unit, and lost sales at p per unit of unsatisfied demand. Assume a positive order lead time of τ.

Define the multidimensional stochastic process $\xi(t) = (\xi_1(t), \xi_2(t), ..., \xi_S(t))$ as the amount of time elapsed since the last S orders were placed. That is, $\xi_S(t)$ is the elapsed time since the last order was placed, $\xi_{S-1}(t)$ is the elapsed time since the second to last order was placed, etc. This process is the superposition of the demand and outdating processes. The approach is to derive the stationary distribution of $\xi(t)$

and use that to develop an explicit expression for the expected system cost rate as a function of S.

Let $p(t, x_1, x_2, ..., x_S)$ be the probability density of $\xi(t)$. Utilizing the partial differential equations of the process, the authors show that

$$\lim p(t, x_1, x_2, ..., x_S) = \begin{cases} Ke^{-\lambda\tau} & \text{for } x_1 < \tau \\ Ke^{-\lambda x_1} & \text{for } x_1 \geq \tau \end{cases}$$

Where

$$K = \left(e^{-\lambda\tau}\tau^S/S! + \int_\tau^{\tau+m} \left[{x_1}^{S-1} e^{-\lambda x_1}/(S-1)! \right] \mathrm{d}x_1 \right)^{-1}.$$

From these results one can derive the steady state distribution of the on-hand inventory at a random point in time, say $(P_0, P_1, ..., P_S)$.. The expected cost rate as a function of S is:

$$C(S) = c\lambda + (p-c)\lambda P_0 + h\sum_{j=1}^{S} jP_j + (c+\theta)\pi,$$

Where π is the expected outdating rate given by:

$$\pi = Ke^{-\lambda(\tau+m)}(\tau+m)^{S-1}/(S-1)!.$$

Although these equations appear straightforward, computation of the stationary state probabilities, $(P_0, P_1, ..., P_S)$ is complex, requiring numerical integration. Although the authors were unable to prove convexity of $C(S)$, numerical tests suggested that it is at least quasi-convex.

One of the interesting aspects of the problem revealed by numerical tests is the relationship between S and m. One would expect that S would be an increasing function of m. That is, as the lifetime of the product increases, the optimal stocking level would also increase, since the penalty for outdating decreases as m increases. Consider the following case: Fix $\tau = 1, \theta = 100, p = 600$ and $\lambda = 50$. The optimal values of S and the optimal expected cost, $C(S)$ for various values of m observed in this case are:

m	S^*	$C(S^*)$
0.005	0	30.00
0.007	7	29.98
0.008	25	29.62
0.009	38	29.01
0.01	47	28.30
0.02	72	21.79

(continued)

(continued)

m	S^*	$C(S^*)$
0.03	73	17.92
0.04	71	15.49
0.05	70	13.82
0.08	66	10.95
0.10	64	9.86
0.30	62	6.45
0.50	64	5.71
0.80	68	5.45
1.00	68	5.43

As one would expect, the expected cost rate decreases as the product lifetime increases. However, note the unusual behavior of S^* as m increases. The reason that S^* is 0 when the product lifetime is very small is that at this value of m, it is more economical not to run the system at all, since virtually all new items outdate before being able to satisfy demand. As m increases, the likelihood that a unit is able to satisfy demand before outdating increases, and so it becomes economical to hold positive stock. Because we chose a very high stockout penalty cost, it is desirable to maintain a larger inventory to decrease the likelihood of stockouts when m has a moderately low value. The value of S^* reaches a maximum in this case when $m = 0.03$. As m continues to increase, the likelihood of outdating decreases, so less stock is needed to maintain the same risk of stocking out. Note that for $m = 1$, the lifetime is essentially infinite, and the solution coincides with the optimal value of S for the nonperishable problem.

Perry and Posner (1998) developed the following extension of this model. Suppose that customers arriving to the system when it is out of stock are willing to wait a random amount of time Y for the next arrival of a unit. The rationale for this extension is that one can keep track of the times of orders, so one can determine the instance of the next arrival. If this is imminent, it makes sense that a customer would wait rather than leave the system unsatisfied.

As an example, suppose that H is the CDF of Y. They suggest the following form for H:

$$H(y) = q + p1_{\{y \geq \tau\}}$$

Where $1_{\{y \geq \tau\}}$ is the indicator function of the event $\{y \geq \tau\}$. This means that a customer is willing to wait for τ units of time for the next arrival (which must come within that time) with probability p. Since $H(y) = q$ for $y \leq \tau$, Y has mass q at 0 so that $P\{0 < Y \leq \tau\} = p$). The authors also consider several other forms of H.

Their analysis is based on computing the joint distribution of the process W given by

$$W = \{W_1(t), W_2(t), ..., W_S(t) : t \geq 0\}$$

where $W_i(t)$ is time to outdate the ith youngest item if the demand process were stopped at time t. While this extension is interesting, and could be potentially useful in some circumstances, the authors do not provide any calculations to compare the results of their model to those of Schmidt and Nahmias (1985). Until that is done, there is no way to tell if their extension leads to policies that differ significantly from those computed assuming customers do not wait when the system is out of stock.

4.2 Continuous Review Models with Zero Lead Time

As noted above, we feel the simultaneous assumptions of continuous review and zero lead time are unrealistic. Nevertheless, we briefly review the major results in this area. The first to look at this problem was Weiss (1980). Weiss assumed that demands are generated by a stationary Poisson process with fixed rate λ. His main result is that, in the case of lost sales, the form of the optimal policy is to either never order, or to order to a fixed level S when the inventory level drops to zero. Notice that the issue of perishability never really comes up, as order cycles are completely independent. One simply waits until the inventory level drops to zero (whether through demand or outdating) and replenishes to a fixed level (much as one does in the simple EOQ model).

Weiss's results were generalized by Liu and Lian (1999). They assumed full backordering of demand and generalized from a Poisson demand process to a stationary renewal process. They showed that the form of the optimal policy is (s, S) with $s \leq -1$, and provide explicit formulas for computing the optimal policy parameters. (A correction of their results appears in Gurler and Ozkaya (2003).) Gurler and Ozkaya (2008) extended their model to the case where the lifetime of a batch is a random variable. Note that for all of these models, the form of the optimal policy (namely that one only orders when the system is backordered), depends heavily on the assumption that order lead times are zero. It is unlikely that one would use such a policy in the real world, where order lead times are always positive. Hence, it would appear that these results do not provide much insight into the general continuous review perishable inventory problem.

4.3 Optimal (Q, r) Policies with Positive Lead Time

Short of finding optimal policies, a common method in inventory control is to determine the best policy from a class of policies. This is the approach taken by Berk and Gurler (2008). It is well known that for many nonperishable continuous review systems, the optimal policy is a (Q, r) policy. That is, when the inventory position hits r, an order for size Q is placed. As the form of the optimal policy is not known, it would seem natural to consider the best (Q, r) policy in the perishable inventory setting. In this version, the authors assume that $r<Q$, which means that there is at most a single order outstanding. (Future work considers the generalization to the case where multiple orders are outstanding.) They also assume throughout that

demands are generated from a stationary Poisson process, thus allowing them to make use of the memoryless property of the exponential interarrival time distribution.

Because items can perish, a slight modification of the standard policy is required. In nonperishable continuous review systems, one is guaranteed that an order can be placed at the instant the inventory level (or inventory position when more than one outstanding order is allowed) hits the reorder point, r. However, if only a single order is in inventory, the entire order will perish at the same instant, thus dropping the inventory level immediately to zero. Hence, the policy must be modified to: One orders Q units whenever the inventory level hits r or zero, whichever comes first.

This modification points out the inadequacy of this policy, however. Consider a blood bank storing a rare form of blood (AB negative, for example). The lead time for replenishment is three weeks owing to the difficulty of finding donors with this blood type. Suppose that there is a substantial amount of AB negative blood on hand due to outdate in one day. According to this policy, the blood bank would sit on their current supply until it outdates the next day, dropping their inventory to zero. During the three week lead time, there would be no AB negative blood available, precipitating a crisis situation. (Note that this problem does not arise in the model considered by Schmidt and Nahmias (1985), since all orders are for size one only, so the on-hand inventory level can only drop by one unit at a time independent of whether the drop is due to outdating or filling demand.)

Clearly, perishability fundamentally changes the form of the optimal policy. The blood bank seeing the supply of this rare blood type would soon expire and would take steps to replenish the inventory far in advance of the outdating. Hence, this type of (Q, r) policy does not make sense in this context. Before one can consider effective heuristics for this problem, a better understanding of the nature of the optimal policy is required.

A modification of the (Q, R) policy which ameliorates this problem to some extent is considered by Tekin et al. (2001). They suggest a (Q, R, T) policy. The policy is implemented in the following way: A replenishment order of size Q is placed whenever the on-hand inventory level drops to r, or when T units of time have elapsed since the last instance at which the inventory level hit Q, whichever occurs first. For this policy to make sense, they require the unusual assumption that the aging of items in a batch begins only after all of the units of the previous batch are exhausted either by demand or outdating. It seems likely that this would be true only in rare circumstances. With this assumption, epochs at which the inventory hits Q are regeneration points of the system. The analysis can then proceed using renewal reward processes. Their main finding is that when service levels are high, the value of r is not increased as much as it would be in an ordinary (Q, r) system. This policy only makes sense when their assumption about aging is accurate, however.

4.4 An Alternative Approach

As noted above, the difficulty with trying to characterize the optimal policy for a continuous review perishable inventory system is properly defining the state variable. If the state is defined in terms of on-hand inventory (as is almost always done in inventory modeling), then the optimal solution would be a function of an infinite dimensional state variable. This would correspond to all of the on-hand orders and their ages. An interesting question is whether or not there is a single dimensional surrogate variable that might provide reasonable operating policies. We speculate that there is such a surrogate variable. It is the expected remaining lifetime or virtual lifetime of the on-hand inventory. That is, given all of the on-hand orders and their ages, and knowledge of the demand process, what is the expected time that this inventory will be depleted either by demand or by outdating?

Consider the following two scenarios. Product lifetime is 10 days, and expected inter-demand time is 1 day. In case one, assume that there are 100 units of 9-day-old inventory on hand, and in case two, there are 100 units of 1-day-old inventory on hand. The expected remaining lifetime of the on-hand stock in the first case is essentially one day, and in the second case it is slightly less than 9 days (since it is possible for 100 demands to occur before the nine days have elapsed). The expected remaining lifetime measure is far more useful than just knowing the number of units on hand, and a policy based on its value should perform reasonably well. The policy we suggest is the following: Let $E(L)$ be the expected remaining lifetime of the current on-hand inventory. Place an order for Q units at the first instance that $E(L)$ drops below a trigger level, r. (This policy assumes only one outstanding order. If multiple orders are outstanding, the form of the policy would have to be modified to take into account unit on order as well as on hand.)

A problem with this idea is that computation of the expected remaining lifetime for an arbitrary stockpile of items appears to be very difficult. An alternative approach would be to use some type of virtual lifetime measure. This would operate much like the virtual waiting time in queueing theory. At each arrival of new stock, the virtual lifetime process would jump upward by an amount related to the size of the order, the lifetime of the product, and the demand process. One would reorder when the virtual lifetime crossed a critical threshold. Conceptually, this is similar to the suggestion above to base the reorder decision on the expected remaining lifetime in the stockpile, but could be easier to compute and implement.

We are not aware of any work that has explored these ideas, but they could provide a way of obtaining reasonable control policies for perishables in a continuous review environment.

Chapter 5
Approximate Order Policies

Contents

Because optimal policies require the solution of an m – 1 dimensional dynamic program, finding optimal policies is feasible only for moderately small values of m. For that reason, approximate policies are of particular interest. The first issue to be addressed when trying to find approximations is the form of the approximate policy. When perishability is not present, we know that an optimal policy is a function of the total inventory on hand. A reasonable starting point is to restrict attention to policies that are only a function of the total starting stock in any period, independent of the age distribution of stock. (As we see, this approach has pitfalls as well.)

5.1 Forms of Approximate Policies

The obvious choice for an approximation based only on total starting inventory is the simple order-up-to or S policy. That is, if x is the starting inventory in any period, the optimal order quantity is $\max(S - x, 0)$. However, there are other choices as well. Nahmias (1975b) compared the performance of three simple approximations all based on total inventory, x. These are the S-type policy defined above (called the critical number policy in the reference), the linear policy and the modified critical number policy. Notice that in addition to the fact that all three approximations are based on x, they all require determination of a single parameter value: x^* for the critical number policy, β for the linear policy, and w^* for the

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_5,

modified critical number policy. The latter two policies are suggested by the form of the optimal policy.

Extensive comparisons of these three policies were done via simulation for a product lifetime of two periods. The main result was that policies 1 and 2 were competitive, with no significant cost difference observed at the 5% level of significance in most cases tested, and policy 3 performed significantly more poorly in most cases tested. These results suggest that the S policy should be a good approximation in most circumstances. We also discuss a fixed order quantity approximation, suggested by the operation of a blood bank. Fixed order quantity policies are not cost-efficient, as they do not respond to demand fluctuations (as does an S policy).

5.2 *S* Policy Approximations

Having identified the S policy as a suitable form for an approximation, one would like to be able to find the optimal S policy. Finding optimal S policies was considered by Van Zyl (1964) for the case of $m = 2$ and for arbitrary $m > 2$ by Cohen (1976).

The basis of the analysis is to show that the state (that is, the vector of inventories of each age level) under a stationary S policy is a Markov chain. Hence, there exists a stationary or invariant distribution for the system state. When $m = 2$, Van Zyl showed that the stationary distribution of the system state, say $G(y)$, is

$$G(y) = F(y)[1 - F(S - y)]/[1 - F(y)F(S - y)].$$

Given an explicit form of G, finding the optimal S is straightforward. While this is a useful result, the reader should keep in mind that when $m = 2$, the dynamic program defining an optimal policy only has a single dimensional state variable so that calculating a true optimal policy in this case is also straightforward. It is more desirable to use an optimal policy than to use the best S policy, which is known to be suboptimal.

The real question is whether this method provides easily computable S policies for larger values of m. Unfortunately, Cohen was not able to obtain explicit results for the stationary distribution of the starting state when $m > 2$. He suggests using successive approximations to find the stationary distribution, but this is comparable to the difficulty of finding a true optimal solution, so is of questionable practical value.

Hence, there is value to considering simpler methods for computing approximate S policies. One of the key advantages of our approach of first developing a meaningful one period model is that this approach is more convenient for constructing approximations. Nahmias (1976) developed a unique approach for finding approximate order policies for the perishable inventory problem that requires essentially the same computational effort for any value of m. In Chap. 2, we observed that there are two reasons why a simple critical number order policy is not optimal: the outdating function,$G_m(t;\mathbf{x})$, and the one period transfer function, $s(y,\mathbf{x},t)$, are functions of the entire state vector, **x,** and not simply the total inventory, x. Nahmias's approach requires finding approximations of both the expected outdating and the one-period transfer function that depend only on **x** only through x. It turns out that either of two sets of bounds for the expected outdating can be used to construct the approximation. The first, below in Lemma 1, are due to Nahmias.

Lemma 1. $\int_0^y F^{m^*}(t)\mathrm{d}t \leq \int_0^y G_m(t;\mathbf{x})\mathrm{d}t \leq \int_0^y F^{m^*}(t+x)$, *where* F^{m^*} *is the m-fold convolution of the one period demand distribution.*

Proof. In Chap. 2, we have that $G_m(t;\mathbf{x}) = P\{D_m + B_{m-1} \leq t\}$ where the random variables $B_0, B_1, \ldots$ follow the recursion $B_j = (D_j + B_{j-1} - x_j)^+$ for $1 \leq j \leq m-1$. Using the fact that for any function $w^+ = \max(w, 0) \geq w$, it follows that $D_m + B_{m-1} = D_m + (D_{m-1} + B_{m-2} - x_{m-1})^+ \geq D_m + D_{m-1} + B_{m-2} - x_{m-1} \geq \ldots \geq \sum_{i=1}^{m} D_i - \sum_{i=1}^{m-1} x_i$, which gives the right hand side of the Lemma. The left hand side of the Lemma is obtained by noting that $F^{m^*}(t) = G_m(t;\mathbf{0}) \leq G_m(t;\mathbf{x})$ for any vector $\mathbf{x} \geq 0$.

Note that the lower bound is independent of x. Since the expected outdating is an increasing function of the state vector, the upper bound is likely to be more accurate. An entirely different approach to bounding the expected outdating was developed by Chazan and Gal (1977). Like Cohen, they assume that a stationary S policy is followed in each period so that the system state evolves according to a Markov chain. Their bounds are based on the following observation. Suppose that the demand per period is bounded above by S/m (where m is the product lifetime). Let D be a random variable representing one period demand and μ the mean demand. In this case, the outdating of each period has expectation $S/m - \mu$. Why? Consider the following simple example. Suppose $m = 2$, $S = 10$, and suppose the initial state vector is (5, 5). (Note that here the state vector is of dimension m since we track both fresh and one period old inventory.) Then by assumption, demand is bounded above by $S/m = 5$. It follows that demand depletes from the oldest inventory first, and the outdating in period 1 is simply $5 - D_1$. The order quantity is 5, and the state returns to (5, 5). The expected outdating of each period is thus $5 - E(D)$. As another example, suppose the starting state is (10, 0), the demand is fixed at one unit each period. Then, the reader should check that the outdating in successive periods follows the pattern 0, 8, 0, 8,... and the average outdating per period is $4 = 5 - E(D)$.

When demand is not bounded above by S/m, this argument provides an upper bound on the expected outdating per period. The lower bound is obtained as follows: Let $\bar{D}$ be the average demand over m periods and $\bar{D}_T$ the average demand over m periods truncated at S/m. That is, $\bar{D}_T = \min(\bar{D}, S/m)$. Then, the lower bound on the expected outdating per period is given by $S/m - E(\bar{D}_T)$. A proof of the validity of the lower bound appears in Chazan and Gal (1977). We do not present it here, but note that their main result is:

Lemma 2. If one orders to S each period, then the expected outdating per period, say $E(O)$, is bounded by:

$$S/m - E(\bar{D}_T) \leq E(O) \leq S/m - E(D_T).$$

Chazan and Gal (1977) also develop bounds in the special case where daily demands follow a Poisson distribution. The Poisson case provides a tighter lower bound than the one indicated in Lemma 2.

Chazan and Gal do not consider the problem of optimizing the value of S, the order up to point. However, the methods developed by Nahmias (1976) can be utilized to find S using either the bounds in Lemma 1 or Lemma 2. Suppose we use the upper bound in Lemma 1. Then, one replaces the one period outdate function as follows:

$$\int_0^y G_m(u; \mathbf{x})\mathrm{d}x \approx \int_0^y F^{m^*}(u+x)\mathrm{d}u = H(x+y) - H(x)$$

where $H(t) = \int_0^t F^{m*}(u)\mathrm{d}u$ so that the one period expected cost function is now expressed in the form:

$$L(x,z) = c(z-x) + L(z) + \theta H(z) - \theta H(x)$$

and

$$z = x + y.$$

The approximation technique is based on the approach developed by Veinott (1965). Veinott observed that under reasonably general circumstances, when there is no fixed order cost, the expected cost function for the standard multiproduct, multiperiod, dynamic inventory problem can be decomposed into the sum of N independent one period cost functions, and the optimal solution has the property of being "myopic."

Suppose we approximate the one period transfer function, $\mathbf{s}(y, \mathbf{x}, t)$, with the simpler form $s(z,t)$. Then, the total discounted n period cost can be represented in the form $\sum_{n=1}^{N} \alpha^{n-1} E(W(z_n)) - cx_1 - \theta H(x_1)$, where α is the one period discount factor. The optimal policy for this modified model is to order up to S each period, where S minimizes $W(z)$. It was Veinott's observation that the traditional backward dynamic programing approach could be replaced by this much simpler forward approach under the right circumstances. As the N period model is decomposed into N one period models, one needs only look forward one period to achieve an optimal policy. Hence, the term "myopic" describes these policies.

To utilize the myopic approach to constructing approximations, it was indicated that both the one period expected cost function and the transfer function need to be approximated by functions of $z = x + y$, and x. The approximate transfer function takes the form

$$s(z,t) = z - t - H(z) + H(z - t)$$

based on the following argument: if the total on-hand inventory in the current period is z and the demand realization is t, then the total inventory on-hand one period hence is $z - t$ – the amount of inventory outdating at the end of the current period. The expected outdating is approximated by the final two terms using arguments similar to the above.

Utilizing this approximate transfer function and approximate one period cost function, we obtain the following for the myopic one period expected cost function:

$$W(z) = (1 - \alpha)cz + L(z) + (\theta + \alpha c)H(z) - \alpha(\theta + c)\int_0^z H(z - t)f(t)\mathrm{d}t.$$

The ease of optimizing $W(z)$ depends of course, on it shape. Nahmias (1976) showed that when the demand density is a PF_2 density (Polya Frequency function of order 2), that this form of $W(z)$ is quasi convex in z, and is thus easily minimized by binary gradient search. (More specifically, he showed that $W'(z)$ changes sign once from minus to plus as z increases from 0.) The class of PF_2 densities is large, including the exponential, gamma, normal, truncated normal, and lognormal to name a few. Computations reported in Nahmias (1976) for a variety of parameter settings resulted in cost errors of under 1% in most cases tested. Tested cases included lifetimes of two and three periods and exponential and Erlang-2 demands. Because of the difficulty of determining an optimal policy, the accuracy of the approximation for longer lifetimes is difficult to determine.

As indicated earlier, the same approach can be applied to the Chazan and Gal bounds in Lemma 2. We can express the bounds in Lemma 2. as $S/m - a(S)$ and $S/m - b(S)$ where

$$a(S) = \int_0^{S/m} mxf^{m^*}(mx)\mathrm{d}x + (z/k)(1 - F^{m^*}(S))$$

$$b(S) = \int_0^{S/m} xf(x)\mathrm{d}x + (S/m)(1 - F(S/m))$$

Since we have no reason to believe that either the lower or upper bound is more accurate, a reasonable approximation for the expected outdating based on these bounds is simply the average of the upper and lower bounds. That means we would approximate the outdating function as:

$$O(z) = z/m - 0.5(a(z) + b(z))$$

Using this bound, and the same logic described above using Lemma 1 bounds, we obtain the following form for $W(z)$:

$$W(z) = c(1 - \alpha) + L(z) + (\alpha^{m-1}\theta + \alpha c)O(z).$$

As with the other form of $W(z)$ using the Nahmias approximation, the minimizing value of $W(z)$ is also easy to find. Calculations reported in Nahmias (1977c) for a more general version of the problem, indicate that the Chazan and Gal bounds provide more accurate results, especially for larger values of m. This approach should provide very accurate S policy approximations for ordering perishables.

S type approximations were also considered by Nandakumar and Morton (1993) and Cooper (2001). Nandakumar and Morton (1993) applied methods developed by the second author for approximating the standard lead time lost sales problem to the perishable inventory problem. Their approach is relatively simple, and considers bounds on the optimal policy obtained if one considers an infinite lifetime and if one considers a very short lifetime with high likelihood of perishing. They compared the performance of their heuristic to those of Nahmias (1976) described in detail above, and the approximation obtained using Chazan and Gal bounds, also described above. Their computational results indicated virtually no difference in performance between the three methods. We speculate that their parameter set was in too narrow a range to distinguish the methods. Based on this writer's experience, we speculate that the method described above using Chazan and Gal's bounds would perform best in a more extensive numerical test.

Cooper (2001) considered an extensive analysis of the stochastic process that generates outdates and was able to obtain a slightly tighter lower bound on outdates than prior bounds. However, the improvement in expected costs his method resulted in was very small, and the expressions he obtained were quite complex. In the opinion of this writer, the benefit was not justified by the additional complexity of his heuristic.

5.3 Higher Order Approximations

While an S policy based only on total starting stock each period might be adequate for many applications, it is important to remember that it is suboptimal, and ignores the age distribution of starting stock. This can cause serious problems in some situations. For example, consider the blood bank application. Whole blood has a legal lifetime of 21 days. Suppose a large supply of 20-day-old blood is available for a particular blood type, and no other blood of this type is available at the blood bank. An approximation based only on total stock on hand would indicate that there are sufficient stocks in the system, and the recommended order quantity would likely be small. However, on-hand inventories would go to zero one day later due to outdating. Hence, such approximations must always be viewed in the context of the application. If there is a substantial lead time for new procurements (as there would be in the blood bank setting), this could lead to a disastrous situation. One might want to use an optimal policy in this situation. However, an optimal policy requires a state variable with 20 dimensions, so is not really feasible. However, there is a middle ground between an optimal policy and a simple approximation, such as an S type policy, based only on total stock.

This issue was addressed in Nahmias (1977b). The basic idea is to approximate an m period lifetime problem with a k period lifetime problem, where $k < m$. He suggests collapsing the state vector from $(x_{m-1}, x_{m-2}, \ldots, x_1)$ to $(x_{m-1}, x_{m-2}, \ldots, x_{k+1}, \sum_{i=1}^{k} x_i)$ based on the fact that the optimal policy is more sensitive to changes in newer rather than older inventory. He then proceeds to show how one can construct a dynamic program with dimensionality k as an approximation to the original problem. For large values of m, other aggregation schemes might make more sense. For example, in the case of the blood bank, we might define a two dimensional state variable, (x_1, x_2) where x_1 is the on-hand inventory of age 5 days or less, and x_2 is the on-hand inventory of age 6–20 days. Another alternative, described in Nahmias (1975b) is to use a multiple critical number policy, where each critical number would correspond to the on-hand inventory of a particular age grouping. For example, in the blood bank case, one critical number would correspond to the inventory of age 5 days or less, and another of age 6 to 20 days. Still another approach would be to simply change the defined length of a period. In the blood bank example, if a period were defined as a week instead of a day, the lifetime would be three periods, and it would be feasible to compute an optimal policy, as the state variable would have dimension 2.

5.4 Fixed Order Quantity Approximation

Consider a system where a fixed quantity of fresh product is added to the inventory each period. This approach was suggested by Brodheim et al. (1975) to model the operation of a blood bank. In the blood banking context, a fixed replenishment policy assumption might be reasonable if the bank has a relatively constant pool of donors. However, fixed order quantity policies are likely to be inferior to S policies in terms of cost performance, since they do not respond to demand. Since an S policy orders exactly the previous period's demand, it is responsive to highs and lows in the pattern of demand realizations.

However, the analysis is significantly simplified under a fixed order quantity policy. Suppose q units are brought into the system each period. Then, the age distribution of on-hand inventory is of the form (q, q, q,. . ., r, 0, 0,. . ., 0) where $0 \leq r \leq q$. Notice that this means that knowing only the total stock on hand uniquely defines the state vector. If the product lifetime is m periods, and the total inventory on hand starting any period is i, then every age category will contain exactly q units up until age category $[i/m]$, which will contain exactly $r = i - [i/m]q$ units. Interpret $[x]$ as the largest integer less than or equal to x. (Note that the original article contained several typos in this part of the description.)

This means that knowing only the total inventory on hand provides a complete description of the system state, and the total starting stock forms a Markov chain of one dimension. This obviates the need to keep track of the age distribution of stock, and significantly simplifies the analysis. The range of possible values of state is 0 to qm. Since under reasonable assumptions, the resulting Markov chain is recurrent, there exists a stationary distribution $\boldsymbol{\pi} = (\pi_1, \pi_2, ..., \pi_{qm})$ satisfying $\boldsymbol{\pi} P = \boldsymbol{\pi}$, where $\mathbf{P}$ is the one step transition matrix for the chain.

Even though the state is only one dimensional, the number of states is likely to be large for most applications. In the blood banking application, $m = 21$, and for a value of q of 100 the resulting chain has 2,100 states. As this is a bit unwieldy, approximations are considered which provide reasonably tight bounds on the probability of shortage and the expected daily outdating. The authors do not consider the problem of optimizing the value of q. While potentially useful in the right environment, this approach is of limited interest (as noted above) due to the fact that a fixed order quantity policy does not respond to demand fluctuations. In the blood banking context, a sudden surge in demand brought on by a catastrophe would certainly be caused to seek larger supplies of fresh blood.

Chapter 6
Inventory Depletion Management

Contents

6.1 Preliminaries

The inventory depletion management problem is discussed in this chapter. Items are stored in a stockpile for the purpose of serving a demand in the field. Items typically age at different rates in the stockpile and in the field. The stockpile may or may not be replenished, and there may be different costs to replenish the stockpile and make emergency replacements in the field. The vast majority of the analysis on this problem is to determine conditions under which either FIFO (i.e., issuing the oldest item next) or LIFO (issuing the newest item next) is optimal.

How does this problem relate to the problem of ordering perishable inventory, which is the primary topic of this monograph? The problems are similar in that they both deal with optimal management of perishable inventory. But that is where the similarity ends. In the basic perishable inventory problem, items are assumed to be of uniform utility to the field irregardless of their ages at issue, as long they have not expired in inventory. This means that from the inventory management point of view, it is optimal to issue the items in FIFO order. The classic inventory depletion problem is static in most cases: most of the research assumes a fixed stockpile of items with no opportunity for replenishment. There is also no external demand, other than the "field." Items are simply issued on a one-at-a-time basis as they die in the field. For the ordering problem, once items are "issued" (i.e., leave inventory to satisfy demand) they leave the system. One is not concerned with how long these items last in the field. In the inventory depletion management problem, the key driver is the relative rates of aging in the stockpile and in the field.

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_6,

Hence there is little, if any, crossover between the two problems. It is clear that there is a large body of real problems accurately described by the perishable inventory problem. Many products are stamped with an expiration date. An item is considered acceptable as long as the expiration date has not been reached, and useless on or after the expiration date. The inventory depletion problem is much more specialized, and would appear to have limited applicability in the real world.

6.2 Deterministic Field Life Functions

Irregardless, there is a fairly large body of research on the depletion problem. The first published work on the depletion problem was Greenwood (1955). Greenwood's interest in the problem clearly arose from military applications. As he notes in his introduction, the prevailing thinking of the time was that FIFO was the preferred issuing order, but that LIFO might be a better choice in many cases. Greenwood notes that the basic trade-off is that under FIFO, issuing occurs frequently, but fewer items expire in the stockpile, while under LIFO, one observes longer field lives, but many items might expire in storage. Greenwood's basic observation is that if the field life function is linearly decreasing with slope one, both LIFO and FIFO are optimal, but if the field life function is convex, LIFO is optimal and if it is concave FIFO is optimal. He notes that when the field life function is convex and decreasing, FIFO requires substantially more replenishment of the stockpile. Formal proofs of these results were not provided.

Derman and Klein (1958) considered a rigorous analysis of a more narrowly defined problem. Virtually, all of the research that followed is based on Derman and Klein's formulation, which is as follows: Assume a fixed stockpile of n items of varying ages, say $(S_1, S_2, \ldots, S_n)$. Items are issued one-at-a-time to the field upon expiration of the previously issued item. An item issued at age S has a field life of $L(S)$, where $L(S)$ is a known function (later extensions would consider the case where field life is a random variable). The objective is to determine the sequence in which to issue the items so as to maximize the total field life (or total expected field life) of the stockpile. While Greenwood (1955) considered only simple linear forms for $L(S)$, he raised some interesting questions that do not appear to have been considered again in the operations research literature. For example, one can consider a model in which there are relative costs of failure in the stockpile versus failure in the field. It is likely that failure in the field is far more costly in most applications. Another aspect of the problem considered by Greenwood is to incorporate ongoing replenishment of the stockpile. Greenwood assumed that new items replenish the stockpile at a steady rate of θ units per unit time. An interesting question would be to consider optimization of θ, or to allow for more complex replenishment mechanisms.

To understand how the problem might arise in the real world, consider the following. Most portable electronic items, such as digital cameras, mp3 players, and flashlights use disposable batteries. As batteries fail, they are replaced.

Batteries age in storage, but much more slowly than while in use. Given a stockpile of batteries of varying ages, in what order should the batteries be issued to maximize the total useful lifetime of the stockpile?

As noted above, in the case of the standard perishable inventory ordering problem, the issue of field life never arises. Items are assumed to age at unit rate while in storage, but once issued to meet demand, they leave the system. In the basic perishable inventory model, a new item with remaining lifetime m has the same utility to the buyer as an item with one unit of time remaining in its lifetime. To this writer's knowledge, the only study which considers both ordering and issuing of perishable items is the unpublished work of Veinott (1960), but as noted in an earlier chapter, only deterministic demand was considered in this study. It seems likely that a stochastic model which treats both ordering and issuing would be very difficult to analyze.

Consider the field life function, which defines the relative rates of aging in the stockpile and in the field. The simplest assumption is that items age at the same rate in the field as in the stockpile. This would correspond to a field life function of the form $L(S) = m - S$, where m is the useful lifetime of a new item. Note that in this case, we would require that $L(S) = 0$ for $S > m$. As we see, this is a trivial case in which all issuing policies are optimal. However, if items age at different rates in the stockpile and in the field, the field life function will not have unit slope. Suppose, for example, that items in the stockpile are refrigerated and age at exactly half the rate of items in the field. In this case, the appropriate field life function is $L(S) = 0.5(m - S)$, where m should be interpreted as the lifetime of a new item that remains in the stockpile until it expires. We would expect that many real cases can be expressed in the form $L(S) = b(m - S)$, where the slope term b indicates the relative rate of aging of items in the stockpile and the field. In virtually all of the real world applications, we can envision $0 \leq b \leq 1$ implying that items age more quickly in the field than in the stockpile.

One can also envision cases where the field life function is nonlinear. For example, automobile batteries age relatively slowly when not in use. Suppose a brand new battery can be expected to last 5 years. A battery which has been stored for 1 year lasts 4.5 years, and one which has been stored 2 years lasts 4.25 years. This would give rise to a convex decreasing field life function with the property that $L'(S) > -1$. In fact, one can envision field life functions that are nonlinear and nonmonotonic. Consider a family that consumes wines one bottle at a time from their cellar. It is well known that many wines improve with age. Suppose that there is an optimal time, say t^* at which a bottle should be decanted. If $L(S)$ measures the utility of a bottle of wine decanted at age S, the field life function would be concave with the maximum occurring at t^*.

Consider once more the case of a linear field life function with slope -1. This means that

$$L(S) = m - S \text{ for } 0 \leq S \leq m$$
$$L(S) = 0 \text{ for } S \geq m.$$

In this case, all issuing policies yield exactly the same field life for the stockpile. As a simple example, consider a stockpile of five items, labeled 1, 2, 3, 4, 5 with initial ages 1, 2, 3, 4, 5. Suppose that each item has a useful lifetime of $m = 10$. It is easy to show that in this example, no matter what order one issues the items, the total field life of the stockpile is 9. If the items are issued in LIFO order, then item 1 is issued first and lasts 9 units of time in the field. By this point, the other four items in the stockpile have expired, so the total field life of the stockpile is 9. If the items are issued in FIFO order, then the first item issued (item 5) would have a field life of 5 units, and each of the remaining four items would have a field life of exactly one unit, thus totaling 9 units again. One can easily check that issuing the items in a random order also gives a total field life of the stockpile of 9 units.

Hence, when items age at the same rate in the stockpile as in the field, all issuing policies are optimal. However, we know that FIFO minimizes outdating in the standard perishable inventory problem. Hence, these two problems have a fundamentally different structure, and the results from one do not necessarily carry over to the other.

Consider another example. Suppose that items age twice as fast in the field as they do in the stockpile. As noted earlier, the field life function for this case is

$$L(S) = 0.5(m - S) \text{ for } 0 \leq S \leq m$$
$$L(S) = 0 \text{ for } S \geq m.$$

In this case, the order of issue makes a difference. Again, suppose that the stockpile consists of five items of ages 1, 2, 3, 4, 5 and assume $m = 10$. Issuing the newest first (LIFO) results in the following: Item 1 at age one issued at time 0 has field life of 4.5. Item 2 at that point has age 6.5 (and remaining lifetime 3.5) and hence has field life of 1.75 and expire at time 6.25. Item 3 is then 9.25 units old and has field life of 0.375. Items 4 and 5 expire by this point, thus giving a total field life of 6.625 for LIFO.

Consider now issuing the items according to FIFO. Item 5 issued at time zero has field life of 2.5. At this point, item 4 has age 6.5 and field life of 1.75. Item 4 expires at time 4.25. Thus, item 3 has age 7.25 at issue and has field life of 1.375 taking us to time 5.625. Item 2 has age 7.625 and field life of 1.1875 taking us to time 6.8125, and finally the remaining field life of item 1 is $L(7.8125) = 1.09375$, giving a total field life of the stockpile from FIFO of 7.90625. In this case, FIFO is optimal.

Although increasing field life functions have also been treated in the literature, it would seem that there are few real world applications of this case. In most contexts, we think of items deteriorating as they age rather than improving with age.

Greenwood only considered field life functions of the form $L(S) = k(m - S)$ for $k > 0$. If $0 < k < 1$, then items age more slowly in the stockpile than in the field, and if $k > 1$, then items age more quickly in the stockpile than in the field. However, Greenwood also assumed that new items are added to the stockpile at a constant rate of θ. He derives several measures of system performance under this scenario, including average age of items at the time they are sent to the field, number of items that die in the field per unit time, average age of items in the field, average number of items maintained in storage, rate of items dying in storage per unit time, and average age of stored items. Expressions are derived for the various measures of system performance as functions of several input parameter values.

As noted earlier, Derman and Klein defined the problem more narrowly than Greenwood. This allowed them to develop rigorous proofs of their results for more general forms of $L(S)$. Their paper was significant in several respects. One was their problem definition. More importantly, they discovered what would become the standard approach for analyzing the issuing problem – namely, an induction argument on n, the number of items in the stockpile. Derman and Klein's main result is:

Theorem 6.1. If

1. $L(S)$is a positive convex function and
2. LIFO is optimal when $n = 2$,

then LIFO is optimal when $n \geq 3$.

Proof. The proof is by induction on n, the number of items in the stockpile. The total field life of the stockpile may be written in the form $Q(x) = x + L(x + S^*)$, where x is the total field life from the first $n - 1$ items and S^* is the initial age of the last (n-th) item issued. Since Q is the sum of a convex function and a linear function, it is also convex. The objective is to maximize Q. It is well known that the maximum of a convex function over a convex set occurs at the extreme points of that set. Hence, $Q(x)$ is maximized at either $x = 0$ or $x =$ its maximum value. Clearly the first case, $x = 0$, is suboptimal – that would mean never issuing the first $n - 1$ items. But the induction assumption says that x is maximized by an LIFO policy on the first $n - 1$ items issued. That said, either the last item issued is the oldest item in the stockpile or not. If it is not, then it must be younger than the next to the last item issued. But that contradicts assumption (1) which says that for any two items, LIFO is optimal. Hence, by switching the order of the last two items, the total field life can be increased. Hence, LIFO must be optimal for the entire stockpile.

Zehna (1962) points out that one must qualify the assumptions of this theorem. It is untrue, for example, if $L(S)$ is an increasing function, or is not monotone. As noted earlier, this is unlikely to occur in the real world, as it implies that items improve in the field. He also points out that the assumption made by Derman and Klein that issuing a single item is never optimal may, in fact, not be true in all circumstances. Zehna provides a proof of the slightly modified theorem that does not require this assumption. It is:

Theorem. If $L(S)$ is a convex, nonincreasing function and LIFO is optimal for $n = 2$, then LIFO is optimal for all $n \geq 2$.

In order for this result to be useful, one still needs to determine those convex functions for which LIFO is optimal when $n = 2$. Two examples presented by Derman and Klein are (a) $L(S) = a/(b + S)$ for $a > 0, b \geq 0$ and (b) $L(S) = ce^{-ks}$ $(c, k>0)$. (Note that the first case was misprinted in the original as $L(S) = (a/b + S)$). In both cases, the field life function is monotonically decreasing, so they do correspond to perishable items. Both functions are also nonlinear.

In a follow-up note to their original paper, Derman and Klein (1959), show that no matter what the form of the field life function, the issuing problem can be formulated and solved as an assignment problem. Since algorithms for the assignment problem are very efficient, this implies that any real problem can easily be solved if the exact form of the field life function is known. From a practical standpoint, this obviates much of the need for further theoretical results for the deterministic problem, although the advantage of their earlier results is that one can determine optimality of LIFO or FIFO only knowing the shape of the field life function, and not necessarily having to know its exact form.

The first extension of the Derman and Klein work was due to Lieberman (1958) who showed that if $L'(S) \geq -1$ and LIFO is optimal for $n = 2$, then LIFO is optimal for $n > 2$, and if FIFO is optimal for $n = 2$, then FIFO is optimal for $n > 2$.

Zehna also provided several extensions of Derman and Klein's and Lieberman's results. He considers cases where $L'(S) < -1$ and shows that LIFO is optimal in both cases where L is convex or concave. However, this is an unlikely case, as it implies that items age faster in stockpile than in the field. Zehna was also the first to consider stochastic field life functions, but obtained few results for this case.

As noted earlier, Derman and Klein's original work sparked a series of papers. In the opinion of this writer, most of these are of limited practical interest because of Derman and Klein's earlier result that if the form of the field life function is known, all inventory problems can be solved as assignment problems. One of these minor extensions was due to Bomberger (1961), who extended Derman and Klein's results based on the properties of the inverse function $L^{-1}(S)$.

Eilon (1961) focused on the problem of whether LIFO or FIFO is optimal for the case $n = 2$. This is of interest since most of the earlier work used the Derman and Klein induction argument, which requires one to assume that LIFO or FIFO is optimal for $n = 2$. One interesting idea was that if one expressed the field life function as a power series expansion, one can find conditions for the optimality of LIFO or FIFO for two items based on these expansions. In a later work, Eilon (1963) focused on the problem of providing closed form expressions for the total field life of a stockpile under several circumstances. He considered cases of identical items, and nonidentical items. He also considered the case where items flow into the stockpile at a constant rate. The replenishments could be of identical or nonidentical items.

Pierskalla (1967a) considered several extensions of earlier work. In particular, he considered the case where there are multiple demand sources that draw from the stockpile. He focused attention on the case where the field life function, $L(S)$, is a continuous nonincreasing function, and considers primarily FIFO issuing for the demand sources. Unfortunately, because of the complexity of the problem, he is able to prove the optimality of FIFO only in special cases. He also considers the addition of penalty costs – namely, there is a penalty cost incurred each time an item is issued from the stockpile. Again, the analysis focuses on cases where FIFO issuing is optimal. Finally, he treats S-shaped field life functions. Such functions are neither convex nor concave, so do not fall into categories previously studied.

6.3 Stochastic Field Life Functions

An extension to the basic issuing problem is the case where the field life is uncertain. The problem is easy to state. As in the deterministic case, assume a stockpile of n items with varying ages. An item of age S issued to the field has a field life $L(S)$ that is a random variable whose distribution depends on S. When an item expires, another is issued from the stockpile until it is depleted. The goal is to maximize the total expected field life of the stockpile. Although easy to state, this problem is essentially unsolved. The difficulty lies in the fact that when items are issued one-at-a-time, the age of an item issued to the field depends on the realizations of the lifetimes of the previous items issued. Hence, successive lifetimes of items issued to the field are dependent random variables, whose distributions are very difficult to find. For this reason, researchers have been able to obtain results in special cases only.

The first to consider stochastic field lives was Zehna (1962). For the general problem, he was only able to obtain very limited results. He claimed that under the general assumption that $L(S)$ has an expectation that is a decreasing function of S, there is very little one can say about the optimality of any issuing policy. He did note, however, that in the case where $E(L(S)) = a + bS$ with $a, b > 0$, FIFO is optimal. Of course, this case is not very interesting, as one would almost always expect that $E(L(S))$ is decreasing in S. He also obtained explicit results for the case where items are issued on a fixed schedule, rather than upon expiration of the previously issued item, and also where there are only two items in the stockpile. Neither case is very interesting from a practical point of view.

The next to consider stochastic field lives was Pierskalla (1967b). He assumes that there is a family of field life functions of the form $L_i(S) = a_i S + b_i$, where $a_i < 0$, and $b_i > 0$ such that $L_i(S_0) = 0$ for all values of i. That is, each field life function is a decreasing linear function with the same truncation point. He then assumes that upon issue an item of age S has field life $L_i(S)$ with probability p_i. This is, of course, a very restrictive assumption, but it overcomes the difficulties that Zehna had dealing with truncation points. Pierskalla goes on to show that FIFO is optimal under such a setting. He extends his results to the case of multiple item demands on the stockpile.

While Pierskalla's model allows him to obtain explicit results, it contradicts the purpose of assuming uncertainty in the field lives. Assuming all items die at the same age is a very restrictive assumption. One would think that the purpose of considering uncertainty in this context would be to allow different items to have different lifetimes when issued at the same age.

An entirely different approach was taken by Nahmias (1974), which was based on the concept of stochastic ordering. Consider two random variables, X and Y. Then, X is said to be stochastically smaller than Y (written $X \subset Y$) if and only if $F_X(t) \geq F_Y(t)$ for all $t \geq 0$, where $F_X(t)$ and $F_Y(t)$ are the cumulative distribution functions of the random variables X and Y, respectively. Nahmias defines a new criterion based this concept.

Let $p(n)$ be a permutation of the integers $\{1, 2, \ldots, n\}$. Without the loss of generality, assume that the n items in the stockpile are arranged such that $0 < S_1 < S_2 < \ldots < S_n$, and let $p_L(n) = \{1, 2, 3, \ldots, n\}$ and $p_F(n) = \{n, n-1, n-2, \ldots, 1\}$ so that these represent LIFO and FIFO issuing policies, respectively. Furthermore, let $Z(p(n))$ be the field life of the stockpile when issued in the order $p(n)$. He assumes that $Z(p(n))$ is a random variable with an arbitrary distribution. We say that an issuing policy $p^*(n)$ is strongly optimal if $Z(p(n)) \subset Z(p^*(n))$ for all issuing policies $p(n)$. Nahmias shows that if either FIFO or LIFO is strongly optimal for $n = 2$, then under fairly general assumptions, it will be strongly optimal for $n > 2$. The proof follows the logic of Derman and Klein's original proof.

He then considers two specific cases. Let X be a nonnegative random variable whose range is restricted to the interval $(0, S_0)$. Suppose that $X(S)$ is the field life of an item issued at age S. Then, define $X(S) = \max[X - h(S), 0]$, where $h(S)$ is a known function defining the rate of deterioration. In this scenario, the realization of a single random variable determines the lifetimes of all of the items in the stockpile. He considers two versions of this scenario, and is able to show that when $h(S) = S$ (i.e., items age at the same rate in the field as in the stockpile) then both LIFO and FIFO are strongly optimal. In a modification of this model to guarantee items do not die in the stockpile, he shows that only FIFO is strongly optimal. Other forms of $h(S)$ are not considered.

While Nahmias's approach skirts some of the problems with models of Zehna and Pierskalla, it does bring up a host of other issues. Are there common cases where either LIFO or FIFO are strongly optimal for $n = 2$? And what results can one obtain for more general forms of $h(S)$ in the simpler case, where the lifetimes of all of the items in the stockpile are determined by the realization of a single random variable?

Albright (1976) approaches the problem differently. As noted above, the difficulty with this problem is that if items are issued one-at-a-time upon the death of the previously issued item, successive field lives are dependent random variables. Albright avoids this problem by assuming that items are issued at random times independent of the realizations of successive field lives. This is essentially a different problem. Perhaps this would be appropriate if the stockpile were to serve a large number of distinct customers. In this case, the failure process is the superposition of a large number of failure processes, and might look like an independent renewal process. For this case (which he calls the independent case), he is able to obtain interesting results assuming increasing failure rate (IFR) distributions, and a deterministic mechanism that measures aging in storage relative to aging in the field. He does consider the dependent case, but was able to obtain results here only for $n = 2$ items.

Chapter 7
Deterministic Models

Contents

7.1 The Basic EOQ Model with Perishability

As we have seen, managing perishables when demand is uncertain is a challenging problem. When demand is known, one can always find an ordering rule that guarantees no outdating. Hence, at first glance, it would appear that the deterministic demand case is trivial. However, this is not the true when demand is nonstationary.

Consider first the stationary demand case (namely, the basic EOQ model). In this case, incorporating perishability is straightforward. We follow the notation from Nahmias (2009). It is well known that in the EOQ setting, the optimal number of units to order each cycle is given by the formula

$$Q^* = \sqrt{\frac{2K\lambda}{h}},$$

where K is the fixed order cost, h is the holding cost measured on a unit per unit time basis, and λ is the fixed rate of consumption also in units per unit time. The cycle time, T, is the time between placement of orders, given by

$$T = Q^{*/\lambda} = \sqrt{\frac{2K}{\lambda h}}.$$

The model is based on the assumption that goods are durable (i.e., have an infinite lifetime). Now, let us suppose that new orders have a lifetime of m units of time. If $T \leq m$, there is no modification of the policy required. All units are consumed by demand prior to outdating. Consider the case $T > m$. Here, all units on hand at time

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_7,

m (which will be $Q^* - \lambda m$) expire, and one must immediately reorder to avoid shortages. It is obvious that in this situation, reducing the order size from Q^* to λm eliminates outdating and reduces the holding cost and marginal order cost and has no effect on fixed costs. Hence, the optimal policy is to order $\min(Q^*, \lambda m)$.

It follows that in the EOQ environment one orders so that outdating never occurs. It appears safe to assume that this property of an optimal solution carries over to *all* deterministic perishable inventory problems. This includes problems with nonstationary costs, capacity restrictions, etc.

7.2 Dynamic Deterministic Model with Perishability

In what has become a seminal paper, Wagner and Whitin (1958) provided the first analysis of the economic lot scheduling (ELS) problem. Consider a set of known requirements over N planning periods, say $\mathbf{r} = (r_1, r_2, \ldots r_N)$. They assumed stationary only fixed order costs and holding costs, and furthermore that these costs are not changing with time. The key result which allowed them to construct an efficient solution algorithm was that an optimal policy only ordered in periods where starting inventory was zero. (We refer to this as the zero inventory property.) This means that if starting inventory at the beginning of the planning horizon is assumed to be zero, an optimal policy is completely specified by knowing the periods in which ordering occurs. It also means that every order quantity is the sum of requirements for some set of future periods (this is known as an exact requirements policy). Because of this result, the problem can be formulated as a shortest path through an N node network and solved efficiently either by forward or backward dynamic programing. Their paper sparked a great deal of interest among researchers, and led to several generalizations. It has become known as the ELS problem.

Consider the extension of the Wagner Whitin ELS problem, but assume that the usable lifetime of the product is m periods. Clearly, it is now optimal to restrict attention to policies that allow no outdating. Suppose that $\mathbf{y} = (y_1, y_2, \ldots, y_N)$are the production (or purchase) quantities over the N period planning horizon, and further suppose that the order quantity y_i results in outdating of k units in period $i + m$. Then, replacing y_i with $y_i' = y_i - k$ must result in costs less than or equal to those incurred by ordering y_i, since (as with the EOQ model) fixed order costs are unaffected, but both holding costs and marginal order costs as well as outdate costs, are reduced. Hence, one might think that this implies that the ELS problem with perishability is uninteresting. This is certainly not the case, however.

The first to consider the extension of the ELS problem to the case of perishable inventory appears to have been Smith (1975). Smith assumed that the zero inventory property carried over to the perishable case (which is not true). He also seems to have ignored the property that (in the fixed life case) an optimal policy always has zero outdating.

Friedman and Hoch (1978) noted that Smith's algorithm was flawed, since it was based on incorrect assumptions. They provided the following example to show that when fixed life perishability is included, the zero inventory property does not always hold. Consider a problem with a three period planning horizon, and nonstationary costs. Suppose that requirements in each period are for one unit, set-up costs are 0.5 each period, holding costs are \$1 per unit held each period, and marginal production (or order) costs are respectively 8, 10, and 12. Furthermore, suppose that the product lifetime is two periods. The optimal solution to this problem is to order two units in period 1 and one unit in period 2. The cost of this policy is the following: \$1 for setting up in two periods, \$2 for holding since there is one unit on hand at the end of periods 1 and 2, and marginal production costs of \$16 + \$10, for a total cost of \$29 over the three period planning horizon. Since this policy never orders more than two periods of demand, no units outdate. The key point is that this policy results in placing an order in period 2 when the entering inventory is one unit. Hence, the zero inventory property is violated. (Notice, that if we removed the two period lifetime restriction, the optimal policy would call for ordering 3 units in period 1.)

The stochastic problem is difficult for several reasons. One is that it is necessary to define a multidimensional state variable to keep track of the on hand inventories of each age level. This is not necessary in the deterministic case, however. Friedman and Hoch discovered a very clever way to track the age of the inventory without having to employ a multidimensional state variable. If one keeps track of both the period in which production (or ordering) occurs for a unit, and the period in which that unit is used to satisfy demand, then one can easily find the age of the unit in every period it is stored in the system. In this way, one can compute the total cost associated with that unit during its lifetime in the system. This cost matrix (which excludes fixed costs) is the main driver of the computational algorithm.

They assume on hand inventory of each age level is subject to decay at the end of each period. That is, if there are I units on hand of age i at the end of a period (after satisfying demand in that period), then only $r_i I$ units will be available at the start of the next period, where $0 \leq r_i \leq 1$. On the surface, this looks like simple exponential decay. However, because decay constants are age dependent, it is in fact much more general. As they note, their model includes fixed life perishability as a special case by defining $r_i = 1$ for $1 \leq i \leq m$, and $r_i = 0$ for $i > m$.

The input to the algorithm is the cost matrix having elements, a_{jt} defined as the holding and marginal production cost of meeting one unit of demand in period t from production in period j, where $1 \leq j \leq t$. It follows that $a_{jt} = c_j \left(\prod_{k=1}^{t-j} r_k \right)^{-1} + \sum_{i=j}^{t-1} h_t \left(\prod_{k=i-j+1}^{t-j} r_k \right)^{-1}$ for $j < t$. The idea behind the first term is the following. If a unit is produced in period j and held until period t, it will have decayed in the intervening periods to $\prod_{k=1}^{t-j} r_k$, which means one will have had to have produced the inverse of this quantity in period j. By applying the decay factor in each of the intervening periods between periods j and t, one determines the holding cost in those intervening periods, thus accounting for the second term. Based on this cost matrix, the authors develop a dynamic programing algorithm for solving the problem. The cost matrix only includes marginal production costs and holding costs. Fixed production costs are treated separately.

While the authors show that it is not necessarily true that ordering only occurs in periods in which starting inventory is zero, they did correctly state that all of the demand in a period is completely satisfied by production in a single prior (or current) period only. This observation is an important feature of their solution algorithm.

Friedman and Hoch's results were extended by Hsu (2000). Hsu defined the following:

z_{it} = the amount of the demand from period t to be satisfied from production in period i.
y_{it} = the amount produced in period i and held at the beginning of period t which excludes the amount z_{it} used to satisfy the demand in period t.
$H_{it}(y_{it})$ = the cost of holding y_{it} units of inventory in period t, which are produced in period i.
α_{it} = the fraction of y_{it} which is lost during period t.

Hsu's development is easier to follow, as he clearly defines the state variables. He generalizes the aging mechanism defined by Friedman and Hoch. Note the variable α_{it} depends on both the age of the inventory and the period in which aging occurs. This allows for different rates of aging based not only on the age of the inventory, but also based on the time of year. For example, one might expect food products to age more quickly in hot weather than in cool weather. This would be reflected by larger values of α_{it} in summer months.

Hsu assumes a more general cost structure than Friedman and Hoch. In particular, both the production and holding cost functions are assumed to be nondecreasing concave functions. He makes the following assumptions:

For $1 \leq i \leq j \leq t \leq n$:

Assumption 1. $\alpha_{it} \geq \alpha_{jt}$
Assumption 2. $H_{it}(y) \geq H_{jt}(y)$ for $y \geq 0$.

Note that by the way the aging mechanism is defined, the age of an item in period t that is produced in period i is the difference $t - i$. It follows that since $j \geq i$, $t - j \leq t - i$. Therefore, Assumption 1 says that older items age deteriorate at least as fast as younger items. Assumption 2 says that as items age, the cost of holding older items is at least as high as the cost of holding younger items. Part of the justification for Assumption 2 is that the holding cost term may also include a disposal cost for perishable items. Note that Hsu's formulation uses exactly the same mechanism for tracking the age of items as that developed by Friedman and Hoch.

The optimization problem is then stated as

$$\text{Minimize } \sum_{t=1}^{n} [C_t(x_t) + \sum_{i=1}^{t} H_{it}(y_{it})]$$

$$\begin{aligned}
&\text{subject to: } x_t - z_{tt} = y_{tt} && 1 \leq t \leq n \\
&(1-\alpha_{i,t-1})y_{i,t-1} - z_{it} = y_{it} && 1 \leq i < t \leq n \\
&\sum_{i=1}^{t} z_{it} = d_t, && 1 \leq t \leq n \\
&x_t, y_{it}, z_{it} \geq 0 && 1 \leq i \leq t \leq n
\end{aligned}$$

He assumes that both the order cost functions $C_t(x_t)$ and the holding cost functions $H_{it}(y_{it})$ are concave nondecreasing functions, generalizing Friedman and Hoch's assumptions that these were linear. Hsu shows that the solution of this generalized version of Friedman and Hoch's model is equivalent to a minimum cost network flow problem on a specially constructed network with flow loss.

Hsu shows that for every t, $1 \leq t \leq n$, there is a unique i, $1 \leq i \leq t$, such that $z^*_{it} = d_t$. This means that at an optimal solution, every demand is satisfied completely by production in a single prior period. Note that the production quantity must be inflated by the decay losses as was noted by Friedman and Hoch. A second structural result (which we do not quote here) is then used to construct the solution algorithm. Hsu notes the result obtained by Friedman and Hoch that the zero inventory property (namely that an optimal solution only produces when starting inventory is zero) does not necessarily hold when perishability is present. In fact, even when perishability is not present, nonstationary costs could result in this property failing to hold. The solution algorithm developed for solving the problem is similar to the one developed by Friedman and Hoch. However, because Hsu allows for nonlinear costs, the cost matrix approach of Friedman and Hoch is no longer possible. In summary, Hsu generalizes Friedman and Hoch's results in two ways: one is allowing for the decay variables to depend on both the age of the inventory and the planning period, and the other is allowing for more general holding and ordering cost functions.

An extension of these results to allow for backorders is considered in Hsu (2003). In this paper, Hsu essentially combines the results for Hsu (2000) and Hsu and Lowe (2001), which did not deal with perishables, but considered the ELS problem with backorders and age-dependent backorder costs. The idea is to allow for backorder costs to depend on both the period in which the backorder occurs and the period in which an item is produced to meet that backorder. We do not review these papers in detail, but note that the earlier Hsu and Lowe paper built a solution algorithm based on properties similar to those discovered by Hsu (2000) when generalizing Friedman and Hoch's model.

An extension of Hsu (2000) was considered by Chu et al. (2005). The model is identical to Hsu (2000), except that the order cost function, $C_t(^*)$ is generalized to the so-called *economies of scale function*.

A cost function, $F(X)$, defined on $[0, \infty]$ is called an *economies of scale function* if:

1. $F(0) = 0$.
2. $F(X)$ is nondecreasing on $[0, \infty]$.
3. The average cost function defined as $\bar{F}(X) = F(X)/X$ for $X>0$, is a nonincreasing function on $(0, \infty)$.

The idea behind the economies of scale function is that it includes many types of discount schedules not captured by the simpler concave nonincreasing function assumed in Hsu (2000). Because there is no requirement that the economies of scale function be continuous or have continuous derivatives, it includes incremental, all units, and carload discount schedules. Except for that, the assumptions and structure of the model are the same as Hsu (2000). The problem is that the important property of an optimal solution under the simpler cost structure that demand in a period is completely filled by production in a single prior period no longer holds. Hence, the algorithm developed in Hsu (2000) no longer holds. It turns out that finding an optimal solution to this problem is NP hard, so the authors consider approximations.

They suggest an approximation which has properties similar to the optimal solution structure in Hsu (2000). Their main result is that this approximation results in a cost that is no higher than $(4\sqrt{2} + 5)/7$ (1.5224) times the optimal solution, and this bound is tight. Unfortunately, this means the cost error can be more than 50%, making the value of this approximation questionable. Note that even though the general problem is NP hard, it does not mean that most reasonable sized problems cannot be solved for an optimal solution by dynamic programing.

Chapter 8
Decaying Inventories

Contents

Decay differs fundamentally from fixed life perishability. Decay means that the amount of loss due to outdating is a function of the amount of on hand inventory. The most common case is exponential decay, which means that a fixed fraction of the inventory is lost each unit of time. Decay has been studied in both continuous review and periodic review settings.

There is a substantial literature on the decay problem. Raafat (1991) lists 70 references, and a more recent review by Goyal and Giri (2001) has 130 references, almost all since Raafat's review. Undoubtedly, the number is in the several hundreds at the time of this publication. One might think that this implies a large number of real-world applications. However, this is not the case. Few products age in this way. Certainly, exponential decay is clearly not appropriate for modeling fixed life inventories. The aging process in these cases is described by the fixed life models treated in earlier chapters of this monograph.

One might conjecture that exponential decay would be a reasonable model of the aging process for fresh produce or other fresh food items. A little reflection quickly reveals this not to be the case, however. For example, suppose a shipment of bananas arrives at a grocery store. When placed on the shelves, they are likely to be either green (unripe) or nearly ripe. At this point, one would expect the loss due to spoilage to be virtually zero. As the batch ages, the rate of spoilage increases, until it eventually reaches 100%, as all of the unsold bananas become overripe. Exponential decay would posit that some fixed percentage, say 10%, of the bananas would spoil each period.

One might argue that if the bananas are replenished on a daily basis, there would always be a mix of old and new bananas in stock, and therefore assuming some

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_8,

fixed fraction spoils each period would be reasonable. However, the fraction of the on hand stock spoiling each period would not be fixed. It varies depending on many factors, including the realization of demand, the age of the bananas when they arrive in stock, the manner in which the bananas are stored, etc.

Volatile liquids, such as alcohol and gasoline, have been suggested as examples of exponential decay. However, these products are virtually always stored in sealed containers, thus making the volatility a nonissue. One might then ask what real situations are accurately described by exponential decay. Radioactive decay is perhaps the one case accurately modeled in this way. Even in this case, it is only the level of radioactivity that is subject to decay, not the quantity of inventory on hand. The interesting application study of Emmons (1968) is perhaps one of the few examples of true exponential decay. Why then, one might ask, is there such a voluminous literature on decaying inventory models? This writer speculates that the reason is that these problems are mathematically tractable, thus affording opportunities for publication. While many of these papers are mathematically interesting, their real-world applicability is suspect at best.

One might justify studying decaying inventories because they provide a reasonable approximation to fixed life inventories. However, this does not appear to be true. Tests which have compared various approximation schemes for ordering fixed life inventories show that exponential decay does not provide a very accurate approximation (Nahmias 1975b). Why, then, is there a chapter devoted to this topic here? Decaying inventories are of theoretical interest, even if they are of limited practical value. And, in the opinion of this writer, this monograph would not be complete if this topic were ignored, especially in light of the very substantial literature on the problem.

8.1 EOQ Models with Decay

The earliest work on exponential decay in the context of inventory management seems to be that of Ghare and Schrader (1963). Their approach was to analyze the differential equations resulting from incorporating continuous exponential decay into the standard EOQ model. Consider a system in which the loss due to decay is assumed to be θ units per unit time. Then, temporarily ignoring losses due to demand, the on hand inventory level function, say $I(t)$, declines according to an exponential function, which can be represented in the form

$$I(t + \mathrm{d}t) = I(t)e^{-\theta \mathrm{d}t}.$$

It follows that the loss due to decay during the interval $(t, t + \mathrm{d}t)$ is

$$I(t) - I(t + \mathrm{d}t) = I(t)(1 - e^{-\theta \mathrm{d}t}).$$

Now, suppose that the demand rate at time t is given by the continuous function $D(t)$. The differential in the decline in the inventory due to both demand and decay is given by

$$\mathrm{d}I = -I(t)(1 - e^{-\theta t}) - D(t)\,\mathrm{d}t.$$

Using the fact that $e^{-\theta \mathrm{d}t} \approx 1 - \theta\,\mathrm{d}t$, one obtains the following differential equation for the decline in on hand inventory due to both demand and decay:

$$\frac{\mathrm{d}I(t)}{\mathrm{d}t} = -\theta I(t) - D(t).$$

One can derive the solution of this differential equation in terms of the integral of the function $D(t)$. When $D(t) = \lambda$, the model reduces to the standard EOQ with decay. The analysis leads to a transcendental equation (which has no explicit algebraic solution). To avoid this problem, the authors suggest in most practical cases of interest it is likely true that the life period, $1/\theta$, is much larger than the cycle time, T. In that case, one can approximate the resulting exponential with the first two terms of the Taylor series expansion. $\left(\text{that is } e^{\theta T} \approx 1 + \theta T + \frac{\theta^2 T^2}{2}\right)$. Doing so gives the following expression for the cycle time, T:

$$T = \sqrt{\frac{K}{\frac{\lambda\theta c}{2} + \lambda hc + \lambda hc\theta T}},$$

where c is the unit purchase cost, K the fixed set up cost, and h the unit holding cost. Note that this is equivalent to a cubic equation in T, which can be seen most easily by squaring both sides of the equation. Ghare and Schrader suggest an iterative scheme for finding the optimal cycle time, T^*. Given the optimal T, say T^*, the optimal order quantity is $Q^* = \lambda T^* + \frac{\lambda\theta T^{*2}}{2}$. We present Ghare and Schrader's model in detail because it does appear to be the first inventory study of exponential decay. The approach is based on several approximations whose accuracy would have to be verified in any particular application.

Ghare and Schrader's model was first extended by Covert and Philip (1973). The analysis and the model were similar to Ghare and Schrader's, except that Covert and Philip assumed that the instantaneous deterioration followed a Weibull distribution rather than an exponential distribution. It is well known that the Weibull distribution plays an important role in reliability theory. It accurately describes the aging of many types of electrical components. One can envision the following scenario where the Covert and Philip model might be applied.

Consider a collection of items, such as automobile batteries, that are stored and issued to the field as required. Suppose that the batteries age in storage according to a Weibull distribution. Assume batteries are ordered in fixed lot sizes only when the entire stockpile is depleted and delivery of fresh batteries is instantaneous. (This assumption is easily relaxed by placing the order prior to the end of the cycle as is done with the simple EOQ model.) The demand rate is a known constant.

Because the shipment of new batteries occurs only when the previous batch has been used up, either by demand or failure, one never mixes items of different ages. Hence, the instantaneous rate of decay is the same for the entire batch. Covert and Philip followed a similar method of analysis to that of Ghare and Schrader. In this case, one writes the differential decline of the inventory, dI, as

$$-\mathrm{d}I = I(t)Z(t)\,\mathrm{d}t + \lambda\,\mathrm{d}t,$$

where the instantaneous decay rate of the inventory is given by

$$Z(t) = \alpha\beta t^{\beta-1}.$$

(note that the original article had several typos in these expressions). This leads to the differential equation of the system

$$\frac{\mathrm{d}I}{\mathrm{d}t} + \alpha\beta t^{\beta-1} I(t) = -\lambda.$$

It is easy to find the optimal order quantity, Q^*, in terms of the cycle time, T. It is

$$Q^* = \int_0^T R\exp(\alpha t^{\beta})\,\mathrm{d}t.$$

The difficult part of the analysis is to find an expression for the cycle time, T. Without presenting all of the details, by using a Taylor series expansion of the exponential, they obtain the following implicit equation for the cycle time, T:

$$c\lambda\sum_{n=1}^{\infty}[\alpha^n n\beta T^{(n\beta-1)}]/[(n\beta+1)n!] + h\lambda\exp(\alpha T^{\beta})/2 - K/T^2 = 0.$$

The authors suggest a simple gradient search to find the optimal value of T.

Cohen (1977) considered an extension of the basic Ghare and Schrader model to include price as well as inventory level as a decision variable. Assume that the demand rate, $D(p)$, is a function of the price p. If one holds the price fixed, Cohen shows that the cycle time, T_p can be approximated by the simple formula $T_p = \sqrt{\dfrac{2K}{D(p)(c\lambda+h)}}$ where K is the fixed set up cost per order, c is the unit cost per item ordered, and h is the holding cost. This expression results from approximating the exponential with the first two terms of the Taylor Series expansion, as previous authors have done. Define the profit function,

$$\pi(T,p) = pD(p) - cD(p) - K/T - (c\lambda + h)D(p)T/2.$$

For a given value of p, the profit function is maximized at T_p. Therefore, the problem reduces finding the solution to

$$\max_{p \geq 0} \pi(T_p, p).$$

Cohen shows that this function achieves its maximum value at $p^* > c$, where p^* could be infinite. One numerical example suggests that the optimal price is relatively insensitive to the decay rate, λ. The model is extended to allow for complete backlogging of excess demand.

8.2 Uncertain Demand

If one assumes fixed lifetime, the memoryless property of the exponential distribution is lost. However, exponential decay is a memoryless process, thus allowing for Markovian analysis. One example of a case where Markovian analysis has been applied is due to Kalpakam and Sapna (1996) who consider a system in which orders are placed on a one-for-one basis (i.e., an $(S-1, S)$ system). They assume that lead times are exponential, demands are generated by a renewal process, and the inventory is subject to exponential decay. While it would seem unlikely that these assumptions would hold in any real world scenario, the model is interesting in that it does lead to a tractable analysis of the system.

The authors show that under their particular assumptions, the joint process $(I,T) = \{I_n, T_n : n = 0, 1, 2, ...\}$ is a Markov Renewal Process, where I is the inventory level and T is the times of demand epochs. Using the theory of Markov renewal processes, they obtain expressions of the various system measures in terms of the steady state distribution of the inventory level. Explicit results are obtained only in the case, where the demand process is assumed to be a memoryless (that is, Poisson) process.

There is a large body of literature on decay models with uncertain demand, but virtually all of these assume zero order lead time or in a few cases, exponential lead time. Few real scenarios would appear to be accurately described by an exponential lead time distribution, and the simultaneous assumptions of continuous demand and zero lead time are even more unrealistic (this was discussed in Chap. 4 in the context of fixed life inventories.)

Periodic review models with zero lead time are useful when lead times are relatively small. When lead times are assumed to be zero, exponential decay is easily incorporated into the standard periodic review models. Consider first the case where there is no order lead time and excess demand is lost. In this case, it is well known that the one period transfer function is

$$s(y,t) = (y-t)^{+}$$

where y is the on hand inventory in the previous period after ordering and t is the realization of demand in the current period. Interpret $s(y,t)$ as the starting stock in the current period. Recall that $x^{+} = \max(x,0)$.

Now, suppose that a fixed fraction of the on hand inventory is lost each period due to decay. Suppose that θ represents the fraction of inventory surviving from one period to the next (i.e., $1-\theta$ is the decay rate). Then, again assuming zero lead time and lost sales, the one period transfer function now becomes

$$s(y,t) = \theta(y-t)^{+}.$$

Clearly, the addition of decay in this context has no effect on the form of an optimal policy (which is well known to be an order up to policy, or S policy). Now, consider the case of full backlogging of demand. This case is slightly more complex, since decay can only be applied to the on hand inventory and not to the backlogged inventory. In this case, we obtain

$$s(y,t) = \begin{cases} \theta(y-t) & \text{if } y \geq t \\ t-y & \text{if } y < t. \end{cases}$$

The optimal policy is still an order up to policy in every period.

If a positive order lead time is present, the problem is much more complex. The problem is that the decay can only be applied to the inventory on hand and not to the inventory on order. Hence, one cannot collapse the state vector into the inventory position (i.e., total stock on hand and on order). The state vector must include each outstanding order as a separate state variable. (The same issue arises in the analysis of the lead time, lost sales problem.)

The only study known to this writer that considers decay in the context of stochastic demand and positive lead times is that of Nahmias and Wang (1979). The authors' approach was to modify the well-known heuristic (Q, R) model developed by Hadley and Whitin (1963) for the no-decay case, and modify it to account for exponential decay. The resulting heuristic was tested against the optimal (Q, R) policy found by simulation for a variety of demand variance to mean ratios and parameter values. The simulated and heuristic (Q, R) values agreed well in most cases, resulting in a worst case performance for the heuristic of 2.77% cost error. Note that the heuristic performed better than the simulated policy in several test cases due to the high variance in the simulated cost making it difficult to reliably search for the optimal (Q, R) in the simulation.

Chapter 9
Queues with Impatient Customers

Consider the following queueing scenario: Customers arrive one-at-a-time to a single server and are served in a first-come–first-served sequence (FIFO). However, customers are impatient in that if they have not completed service by a fixed time, say *m,* they leave the system. There is an obvious analogy with the perishable inventory problem. Identify the queue with the on hand inventory, completion of service with occurrences of demand, and the impatience time with the lifetime of the product. Hence, it is possible that such queueing models could be useful for describing some types of perishable inventory systems.

However, there are also fundamental differences between the queueing scenario and the inventory management scenario. For one, in the queueing scenario, one assumes that arrival rates and service rates are fixed and known. Hence, one does not reorder units in response to demand. The replenishment process (i.e., the arrival process) is outside the control of the user. Even if one treats the arrival rate as a control variable, arrivals still occur according to a random process. Another issue is what happens when the queue is empty. In queueing, service is suspended when the queue is empty, but in the inventory context demands continues to occur when the system is out of stock. For these reasons, this analogy has limited utility in the context of perishable inventory management.

However, one important real world problem that could be accurately described by a queuing model is blood banking. Units of fresh blood "arrive" to a blood bank via donations, which are likely to occur one-at-a-time according to some random process. One might affect the arrival rate by using Bloodmobiles, advertising, company-wide donation programs, etc. Since blood banks rarely stock out, the issue of what happens when the queue is empty is moot.

For the most part, queueing models are descriptive rather than proscriptive. That is, given a set of parameters and assumptions about the system, the goal of queueing analysis is to describe various measures of the system. Inventory models are proscriptive in that a solution involves optimizing the replenishment policy.

There is a very substantial literature on queues with impatient customers. Because these models only have limited applicability to perishable inventory management, we only briefly review this literature. The first study of queues with

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_9,

impatient customers appears to be that of Barrer (1957). Barrer generalized the standard M/M/1 queueing analysis to the case of impatient customers. He considered both the cases, where the customer leaves the system if service has not begun by time m, and where the customer leaves the system if service is not completed by time m. The latter assumption would appear to be most appropriate for describing perishable inventories.

In an M/M/1 queue, arrivals occur according to a simple Poisson process and service times are independent exponential random variables. The traditional method of analysis is birth and death analysis, with the state being the number of customers in the system (i.e., the number in the queue plus the number in service). Birth and death analysis is based on the fact that the number of customers in the system is a Markov process. However, a little reflection reveals the fact that the number of customers in the system is not Markovian when customer impatience is introduced. To determine if a customer leaves the queue before completing service, one must keep track of the elapsed time that customer has been in the queue. Hence, it would appear that Barrer's analysis, which uses birth and death analysis, is flawed in principle as was pointed out by Gnedenko and Kovalenko (1968). However, it does appear that his results were correct.

The proper way to analyze a queueing model with impatience is via the virtual waiting time process, and a closely related process, that of the oldest unit in the queue. Gnedenko and Kovalenko analyze the queue with impatient customers in this fashion assuming multiple servers. In the inventory context, the appropriate assumption is a single server.

The first to use a queueing model in the context of controlling perishable inventories appears to be Nahmias (1982). Assume that customers arrive to the queue at a rate of λ and service occurs at a rate μ. Customers whose service is not completed by time m leave the system. Let the traffic intensity be $\rho = \lambda/\mu$, and define $\alpha = \mu m$. Furthermore, assume that there are two costs incurred. When the system is empty, there is a stockout cost of p per unit per unit time and outdating costs are incurred at θ per unit. Then, Nahmias shows that the expected cost rate as a function of the traffic intensity is

$$G(\rho) = \frac{(\rho - 1)\mu}{\rho e^{\alpha(\rho-1)} - 1}\left(\rho\theta e^{\alpha(\rho-1)} + p\right) \text{ if } \rho \neq 1,$$

and

$$G(\rho) = \frac{\mu}{\alpha + 1}(\theta + p) \text{ if } \rho = 1.$$

The goal of the analysis is to find the value of the traffic intensity that minimizes $G(\rho)$. Note that if demand rate, μ, is known, this is equivalent to optimizing the arrival rate λ. Nahmias was unable to prove that $G(\rho)$ is a convex function. However, numerical tests showed that $G'(\rho)$ appeared to follow sign pattern minus/plus, implying a unique solution to the equation $G'(\rho) = 0$. The solution is easily found by numerical methods.

Graves (1982) was the next to consider the application of queueing models for controlling perishable inventories. Graves considered a different model from Nahmias and a different mode of analysis. He assumes that production occurred at a constant rate of c units per unit time. This would be appropriate, for example, if one is considering a chemical processing plant with continuous production. Furthermore, the demand process was assumed to be a compound process. That is, demand occurrences followed a Poisson process with rate μ, and the size of individual demands follows an exponential distribution with rate γ. Inventory is issued to meet demand on an oldest first basis (FIFO), backorders are not permitted, and inventory that reaches age m before being consumed by demand is outdated and leaves the system.

Graves' approach bypasses the (apparently incorrect) birth and death analysis used by Barrer, and instead focuses on the process $A(t) =$ age of the oldest unit in inventory at time t. Interestingly, the process $A(t)$ is Markovian. It is, in a sense, the dual of the virtual waiting time process (which would be $m - A(t)$). Although Graves does not consider optimizing the production rate, he does derive simple expressions for the expected outdates per unit time, expected shortages per unit time, the expected age of the oldest unit supplied to satisfy demand, and the expected inventory level. He also considers a model with unit demands occurring according to a Poisson process. In this case, only approximate results are obtained, however.

Kaspi and Perry (1983) is the first in a long series of research papers by the second author. These papers are studies of queues with impatient customers, although their titles would indicate otherwise. They are typical of queueing papers in that they do not treat issues of optimization, but only consider various descriptive measures of the system under specific assumptions about the arrival and service processes. The extent to which these models can be applied to perishable inventory control is unclear. The key result in this paper is that the process of deaths (that is, outdating) is a delayed renewal process when the input process is a Poisson process and services are exponentially distributed. We do not review this stream of research here.

Several authors have considered variations of the basic queueing model suggestion specifically by blood banking applications. One such study is due to Goh et al. (1993). They considered a blood banking system with Poisson input and two classes of age category: new and old. New blood (typically 10 days old or newer) is reserved for special applications, such as heart transplants and neonatal procedures. They consider two issuing policies: one where the stock of newer blood is reserved only for the use of these critical procedures. In the second case, new blood may be used to satisfy demand if the stock of older blood is zero. The analysis is based on identifying first and second moments of time between successive outdate epochs. The expressions are complex functions of the various system parameters. Simpler approximations are also considered and tested via simulation.

In a more recent study suggested by the same application, Deniz and Scheller-Wolf (2007) also consider the case, where customers demand items of different ages. As with Goh et al. (1993) only two age categories are allowed. However, this study differs from that of Goh et al. in that replenishment policies are considered.

The authors consider two heuristic ordering policies. They are labeled TIS and NIS. The TIS policy is what we previously referred to as an order-up-to-policy or an S policy. The NIS policy is what we previously referred to as fixed order quantity policy: one orders S units of new stock independent of the inventory position.

A unique aspect of this study is the assumptions made regarding upward and downward substitution. First, suppose there are given fractions of customers denoted $0 \leq \pi_D \leq 1$ and $0 \leq \pi_U \leq 1$ that are willing to accept downward and upward substitution, respectively. Downward substitution means that new product is sold to a customer that requests old stock, and upward substitution is the opposite. The authors examine the following four scenarios: no substitution, downward substitution only, upward substitution only, and both downward and upward substitution. What makes this study interesting and unique is the consideration of four distinct issuing policies. The vast majority of the perishable inventory literature assumes FIFO issuing; and a few studies assume LIFO issuing. The primary focus of this research is to determine under what conditions one of these four issuing rules is preferred given the replenishment policy. In particular, they delineate ten scenarios involving relationships among the costs and substitution parameters resulting in a preferred issuing policy.

Chapter 10
Blood Bank Inventory Control

As noted earlier, many of the models reviewed in this monograph were motivated by the problem of optimally storing and issuing whole blood. However, actual blood banking applications have features too complex to include in the mathematical models. In this chapter, we briefly review the significant papers in the management science literature aimed specifically at blood banking that incorporate these features.

The earliest theoretical work on perishables appeared in the literature in the 1970s. Even earlier than that, researchers began to consider blood banking applications. The field of blood banking is very large, with dedicated journals on the subject (notably, *Transfusion*). This chapter treats only mathematical and computer simulation models aimed at improving inventory management of blood banks.

The first quantitative analysis of inventory management of blood banks is due to Elston and Pickerel (1963, 1965). Their earlier paper considers the distribution blood usage in the hospital. Let Y be a random variable representing the demand for blood of a specific type at a hospital blood bank. They show that under certain assumptions about the frequency of demands for blood, and the size of each request, the distribution of Y follows a negative binomial distribution with parameters p and k, where the parameters depend on the blood type, the day of the week, and possibly other factors. The form of the distribution is

$$P(y) = \frac{(y-k-1)!}{y!(k-1)!} p^y (1-p)^k, \quad y = 0, 1, 2, \ldots$$

They claimed that this distribution also describes input and usage. Note that the negative binomial distribution is the result of assuming that the number of patients requiring transfusions is a Poisson random variable, and the size of each transfusion is a logarithmic random variable. The authors raised several issues here and in Elston and Pickerel (1970) that were later considered in much more detail by other researchers. These included a comparison of outdates when using the freshest

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_10,

(LIFO) verses the oldest (FIFO) blood as well as simulating the operation of the blood bank assuming various S type order up to policies.

Jennings (1973) followed up on the work of Elston and Pickerel with a simulation study of S type policies. Jennings noted that there were several special features of blood banking management not considered in the theoretical models discussed in this monograph. For one, blood is crossmatched and assigned prior to use. Consider a physician making a request for four units of blood prior to a surgery. The correct blood type is removed from the hospital blood bank and tested against the patient's blood to make sure that it is a match. Then, the blood is assigned for the surgery, and is not available for any other use between the time it is assigned and the surgery occurs. Typically, the surgeon requests more blood than he or she expects to use. The unused units are then returned to the unassigned inventory, assuming that they have not outdated.

Jennings notes that dynamic programing models of blood inventory are impractical, owing to the large dimension of the state variable. (At the time, blood had a legal lifetime of 21 days, resulting in a 20 dimensional state vector. Today, it appears that it has been extended under certain circumstances to longer periods, perhaps as much as 35 days.) Jennings assumed an S-type order policy and analyzed the problem via Monte Carlo simulation. He assumed that the hospital blood bank can raise its total inventory of a blood type to a desired level S at the start of each day. (This ignores real issues of supply uncertainty.) Jennings' analysis focused on the trade-off is between shortages and outdates. As the value of S increases, shortages decline, but outdates increase. He represents both shortages and outdates as percentages of the annual number of pints transfused. This obviates the problem of assigning costs to these measures. He constructs a shortage/outdating trade-off curve, where each point on the curve is generated by a different value of S.

The shortage outdating curve can potentially be a useful tool for managers seeking the most effective inventory level. He also develops a family of curves for regional systems consisting two, five, and twenty bank systems. He considers the cases where the system operates using a common inventory policy (which means, the regional system operates as if it were a single hospital) versus a threshold transfer policy, which defines how units are transferred from one hospital to another.

Pegels and Jelmart (1970) present a Markov chain formulation of the blood-banking problem that, at first blush, appears to provide insight into the problem. However, Jennings and Kolesar (1973) note that there are several problems with the approach. For one, Jennings notes that it would be almost impossible to find the transition probabilities in a real-world setting. Furthermore, Kolesar points out that their construction is not even a Markov chain (that is, it does not satisfy the memoryless property required of Markov chains).

Cumming et al. (1976) consider alternatives for collection policies to improve system wide efficiency. They note that requirements for blood do not appear to be seasonal, but do exhibit a fair amount of variation over days of the week. Based on their data, usage rates are highest on Monday, Tuesday, and Wednesday and lowest on weekends and holidays.

They considered several models of issuing policies based on age. The random draw model, for example, assumed that any age of blood was equally likely to be transfused. Since most hospitals follow FIFO issuing to some extent, this is clearly not an accurate descriptor of reality. A better model of age at issue that they discovered is the following. Let

$$p_{jt} = \gamma^{(21-j)} p_{21,t}$$

where p_{jt} is the probability that a unit of blood having a remaining life of j days is transfused on day t, and γ is a constant that specifies the preference for older or fresher blood. In most cases, one would expect that $\gamma < 1$, which implies a type of FIFO policy. Based on actual data, the authors discovered that γ was typically in the range of 1.05–1.10.

Cohen and Pierskalla (1979) follow-up on the work of Jennings (1973). They treat the trade-off of shortages and outdating in terms of costs rather than as a percentage of annual demand. They approached the problem in terms of developing a decision rule that specifies the order-up-to-level S as a function of various key parameters of the system. Parameter values are optimized via linear regression. Using this approach, they obtained the following decision rule:

$$\ln S^* = 1.7967 + 0.7604 \ln(d_{\mathrm{M}}) + 0.1216 \ln(p) - 0.0677 \ln(D)$$

which is a result of taking natural logs of both sides of the multiplicative equation

$$S^* = 6.03(d_{\mathrm{M}})^{0.7604} p^{0.1216} D^{0.0677},$$

where d_{M} is the mean daily demand for a blood type, p is the average transfusion to crossmatch ratio (i.e., the proportion of crossmatched blood actually transfused), and D is the crossmatch release period (i.e., the amount of time that elapses from the point the blood is ordered until it is ready for transfusion or is returned to the inventory).

The idea of a multiplicative decision rule has been used before in other inventory planning models, but Cohen and Pierskalla were the first to apply it in the blood-banking context.

Several researchers considered improving the efficiency of regional blood distribution systems via rotation policies among the hospitals in a region. This approach was first considered by Brodheim and Prastacos (1979a, b). In the second reference, they formulate the problem as a mathematical program which is executed parametrically. The algorithm was successful in providing rotation policies in the Long Island region. The technical details of the algorithm are presented in Prastacos and Brodheim (1980). Kendall and Lee (1980) also suggest a mathematical programing model of regional redistribution and illustrate their method with data from a Midwest regional blood system. Their method does not appear to have been implemented, however.

An improvement in the standard crossmatching policy was considered by Dumas and Rabinowitz (1977). Because crossmatching and assigning blood means that this blood is effectively removed from the system, and often 50% or more of assigned blood comes back to the bank, their concept was that efficiencies could be obtained by double-crossmatching. That is, assigning units of blood to two patients rather than one, thus increasing the probability that a larger portion of the assigned blood would be used. To show why significant wastage occurs, they quote the experience at Mount Sinai hospital in New York City. First, half of the original reservation requests are cancelled on average. Of those requests not cancelled, on average two thirds of the assigned units are transfused, and one third returned to the blood bank. That means that typically only one third of the originally requested units are actually used. Now, suppose that p is the probability that an assigned unit is transfused, and suppose the unit is shared by two patients, each having the same probability of usage. The unit will be used if either or both of the patients require it. The problem can be modeled with a binomial distribution with $n = 2$ and p as the probability of success. The probability of at least one success in two trials is easily seen to be $2p(1 - p) + p^2 = p(2 - p) > p$, which means the probability of usage increases with the double crossmatching. Of course, actually implementing such a policy has many complications (involving reservation cancellations, valid and invalid blood type substitutions, using triple rather than double crossmatching, etc.) We do not go into these details here. It does not appear that anyone has ever attempted to implement a double crossmatching policy, so its effectiveness has not been tested in the field.

Pegels et al. (1977) consider a variety of policy changes that could improve blood banking effectiveness. The four policy alternatives they consider are

1. Utilizing frozen red cells
2. Increasing the legal shelf life of whole blood
3. Rescheduling blood collection operations
4. Improving inventory control

Since they made these suggestions, it appears that the use of frozen red cells has increased substantially. Also, they suggest that increasing the legal shelf life from 21 to 28 days. In some circumstances, blood over 30 days old has been used. In order to evaluate the effect of these policy changes, they used an empirical database consisting of six months of data from a Midwest blood region that collects 50,000 units annually and seeks to maintain 2,000 units in inventory at all times. Using this database, they simulated the operation of the system under the various policy scenarios above. We do not go into their results in detail, but provide only a brief summary of their findings. For one, they found that increasing the use of frozen red blood cells smoothed out the fluctuations in inventory, but increased the operating cost of the system.

A simulation of the system was developed to test the effect of increasing the legal lifetime from 21 to 28 days. Their results were surprising. They found that the average age at transfusion increased from 9.53 to 13.50 days (implying the quality of the transfused blood decreased), but the decrease in units wasted per day improved only minimally (from 19.4 to 18.9). The percent wastage only declined by 0.2%.

They also considered smoothing out the collections process. They noted that the recruitment of donors and actual collections were poorly distributed over time and over geographic locations. They found that changing the schedule of the bloodmobile was effective in smoothing out the variations in the inventory levels.

Finally, they considered increasing the FIFO-ness of transfusions. This means increasing the adherence to using an FIFO issuing policy, and using the older blood first as often as possible. This issue is not so simple, however. Surgeons insist on the use of fresher blood for certain critical procedures, such as organ transplants, and the process of crossmatching and assigning often means that it can be very difficult to use the oldest blood for a procedure.

Virtually all of the studies discussed which attempt to compare shortages and outdating for various values of stocking levels employ computer simulation. As noted, the issue of crossmatching significantly complicates the problem, thus making mathematical modeling difficult. The only study this writer is aware of that considers a purely mathematical model for this tradeoff is Jagannathan and Sen (1991). Suppose that the lifetime of blood is L and the crossmatch release period is d. (This is the time between the point that the blood is crossmatched and tested for compatibility and it is released or made available for either transfusion or returned to inventory.) As a rule, a significantly larger amount of blood is ordered by the surgeon than is expected to allow for unforeseen contingencies. Thus, the proportion of crossmatched blood that is actually transfused is less than one. Typically, these ratios vary from 0.27 to 0.67. This is known as the transfusion-crossmatch ratio, denoted by p. Crossmatched demand is denoted by δ, and is assumed to be normally distributed. One interesting result they obtain is that if the crossmatch demand δ is constant, the mean daily outdating is given by

$$\omega = \frac{p\delta(1-p)^n}{1-(1-p)^n}$$

where $n = L/d$. The authors obtain analytical expressions for several other measures of system performance, and show by a comparison with simulated results that they provide accurate estimates of shortages and outdating for various stocking levels.

Chapter 11
Afterword

The purpose of this monograph has been to provide a reasonably comprehensive summary of the theory of inventory management of perishables. Admittedly, much of the focus has been on my own work. However, I have tried to cover the primary developments in the field as I see them. As the field has grown enormously since I did the formative work on my doctoral dissertation some 40 years ago, I have undoubtedly omitted citing some significant references. I apologize in advance for this. While I expect there to be errors of omission, I hope there are few errors of commission.

My hope in writing this monograph was to create a springboard for aspiring researchers. Given the now large body of work on perishables, it is a daunting task try to familiarize oneself with the literature, which is certainly required before embarking on a research program. While much work has been done, I believe that there are substantial opportunities for additional contributions. I look forward to seeing these contributions in the literature over the coming years.

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5_11,

References

Albright SC (1976) Optimal stock depletion policies with stochastic field lives. Management Science 22:852–857

Arrow KA, Harris TE, Marschak T (1951) Optimal inventory policy. Econometrica 19:38–39

Arrow KA, Karlin S, Scarf HE (eds) (1958) Studies in the mathematical theory of inventory and production. Stanford University Press, Stanford, CA

Barrer DY (1957) Queueing with impatient customers and ordered service. Operations Research 5:650–656

Berk E, Gurler U (2008) Analysis of the (Q, r) inventory model for perishables with positive lead times and lost sales. Operations Research 56:1238–1246

Bomberger EE (1961) Optimal inventory depletion policies. Management Science 7:294–303

Brodheim E, Prastacos G (1979a) A regional blood management system with pre-scheduled deliveries. Transfusion 19: 455–466

Brodheim E, Prastacos G (1979b) The Long Island blood distribution system as a prototype for regional blood management. Interfaces 9:3–20

Brodheim E, Derman C, Prastacos G (1975) On the evaluation of a class of inventory policies for perishable products such as blood. Management Science 21:1320–1325

Chazan D, Gal S (1977) A markovian model for a perishable product inventory. Management Science 23:512–521

Chu LY, Hsu VN, Shen M (2005) An economic lot-sizing problem with perishable inventory and economies of scale costs: approximation solutions and worst case analysis. Naval Research Logistics 52:536–548

Cohen MA (1976) Analysis of single critical order policies in perishable inventory theory. Operations Research 24:726–741

Cohen MA (1977) Joint pricing and ordering policy for exponentially decaying inventory with known demand. Naval Research Logistics Quarterly 24:257–268

Cohen MA, Pierskalla WP (1979) Target inventory levels for a hospital blood bank or a decentralized regional blood banking system. Transfusion 19:444–454

Cooper WL (2001) Pathwise properties and performance bounds for a perishable inventory system. Operations Research 49:4555–466

Covert RP, Philip GC (1973) An EOQ model for items with weibull distribution deterioration. AIIE Transactions 5:323–326

Cumming PD et al. (1976) A collections planning model for regional blood suppliers: description and validation. Management Science 22:962–971

Deniz, Karaesmen, and Scheller-Wolf (2007) Managing inventories of perishable goods: the effect of substitution. Unpublished manuscript, 34 pages

Derman C, Klein M (1958) Inventory depletion management. Management Science 4:450–456

S. Nahmias, *Perishable Inventory Systems*, International Series in Operations Research & Management Science 160, DOI 10.1007/978-1-4419-7999-5,

Derman C, Klein M (1959) Discussion: a note on the optimal depletion of inventory. Management Science 5:210–213

Deuermeyer B (1979) A multi-type production system for perishable inventories. Operations Research 27:935–943

Deuermeyer B (1980) A single period model for a multiproduct perishable inventory system with economic substitution. Naval Research Logistics Quarterly 27:177–185

Dumas MB, Rabinowitz M (1977) Policies for reducing blood wastage in hospital blood banks. Management Science 23:1124–1132

Edgeworth FY (1888) The mathematical theory of banking. Journal of Royal Statistical Society 51:113–27

Eilon S (1961) FIFO and LIFO policies in inventory management. Management Science 7:304–315

Eilon S (1963) Obsolescence of commodities which are subject to deterioration in store. Management Science 9:623–642

Elston RC, Pickerel JC (1963) A statistical approach to ordering and usage policies for a hospital blood bank. Transfusion 3:41–47

Elston RC, Pickerel JC (1965) Guides to inventory levels for a hospital blood bank determined by electronic computer simulation. Transfusion 5:465–470

Elston RC, Pickerel JC (1970). Blood bank inventories. CRC Critical reviews in Clinical Laboratory Sciences 8:527–545

Emmons H (1968) A replenishment model for radioactive nuclide generators. Management Science 11:263–273

Friedman Y, Hoch Y (1978) A dynamic lot size model with inventory deterioration. INFOR 16:183–188

Fries BE (1975) Optimal order policy for a perishable commodity with fixed lifetime. Operations Research 23:46–61

Ghare PM, Schrader GF (1963) A model for an exponentially decaying inventory. The Journal of Industrial Engineering 14:238–243

Gnedenko B, Kovalenko IN (1968) Introduction to queueing theory. Israel Program for Scientific Translation

Goh C, Greenberg BS, Matsuo H (1993) Two stage perishable inventory models. Management Science 39:633–649

Goyal SK, Giri BC (2001) Recent trends in modeling of deteriorating inventory. European Journal of Operational Research 134:1–16

Graves S (1982) The application of queueing theory to continuous perishable inventory systems. Management Science 28:400–406

Greenwood JA (1955) Issue priority. Naval Research Logistics Quarterly 2:251–268

Gurler U, Ozkaya Y (2003) A note on "Continuous review perishable inventory systems: models and heuristics". IIE Transactions 35:321–323

Gurler U, Ozkaya Y (2008) Analysis of the (s, S) policy for perishables with random shelf life. IIE Transactions 40:759–781

Harris FW (1915) Operations and cost. Factory Management Series, Shaw, Chicago

Hadley GJ, Whitin TM (1963) Analysis of inventory systems. Prentice Hall, Englewood Cliffs

Hsu VN (2000) Dynamic economic lot size model with perishable inventory. Management Science 46:1159–1169

Hsu VN (2003) An economic lot size model for perishable products with age dependent inventory and backorder costs. IEE Transactions 35:775–780

Hsu VN, Lowe TJ (2001) Dynamic economic lot size models with period-pair dependent backorder and inventory costs. Operations Research 49:316–321

Jagannathan R, Sen T (1991) Storing crossmatched blood: a perishable inventory model with prior allocation. Management Science 37:251–266

Jennings JB (1973) Blood bank inventory control. Management Science 19:637–645

Jennings JB, Kolesar PJ (1973) Comments on a blood-bank inventory model of Pegels and Jelmert. Operations Research 21:855–858

Kalpakam S, Sapna KP (1996) A lost sales (s − 1, s) inventory system with renewal demand. Naval Research Logistics 43:129–142

Kaspi H, Perry D (1983) Inventory systems of perishable commodities. Advances In Applied Probability 15:674–685

Kaplan R (1970) A dynamic inventory model with random lead times. Management Science 16: 491–507

Kendall KE, Lee SM (1980) Formulating blood rotation policies with multiple objectives. Management Science 26:1145–1157

Lieberman GJ (1958) LIFO vs FIFO in inventory depletion management. Management Science 5:102–105

Liu L, Lian Z (1999) (s, S) continuous review models for products with fixed lifetimes. Operations Research 47:150–158

Muckstadt JA (2005) Analysis and algorithms for service parts supply chains. Springer Series in Operations Research and Financial Engineering, 57 pages

Nahmias S (1972) Optimal and approximate ordering policies for a perishable product subject to stochastic demand. Unpublished doctoral dissertation. Northwestern University, Evanston, Ill. 158 pages

Nahmias S (1974) Inventory depletion management when the field life is random. Management Science 20:1276–1283

Nahmias S (1975a) Optimal ordering for perishable inventory – II. Operations Research 23:735–749

Nahmias S (1975b) A comparison of alternative approximations for ordering perishable inventory. INFOR 13:175–184

Nahmias S (1976) Myopic approximations for the perishable inventory problem. Management Science 22:1002–1008

Nahmias S (1977a) Comparison between two dynamic perishable inventory models. Operations Research 25:168–172

Nahmias S (1977b) Higher order approximations for the perishable inventory problem. Operations Research 25:630–640

Nahmias S (1977c) On ordering perishable inventory when both the demand and lifetime are random. Management Science 24:82–90

Nahmias S (1978) The fixed charge perishable inventory problem. Operations Research 26: 464–481

Nahmias S (1980) Queueing models for controlling perishable inventories. Proceedings of the First International Symposium on Inventories, Budapest

Nahmias S (1981) Managing reparable item inventory systems: a review, in Multilevel Production/Inventory Control Systems: Theory and Practice, edited by L. Schwarz. TIMS Studies in the Management Sciences, North Holland, Vol. 16, pp. 253–277

Nahmias S (2009) Production and Operations Analysis, Sixth Edition. McGraw Hill/Irwin, Burr Ridge, Illinois, 787 pages

Nahmias S, Pierskalla WP (1973) Optimal ordering policies for a product that perishes in two periods subject to stochastic demand. Naval Research Logistics Quarterly 20:207–229

Nahmias S, Pierskalla WP (1976) A two-product perishable/nonperishable inventory problem. SIAM Journal on Applied Mathematics 30:483–500

Nahmias S, Schmidt CP (1986) An application of the theory of weak convergence to the dynamic perishable inventory problem with discrete demand. Mathematics of Operations Research 11:62–69

Nahmias S, Wang SS (1979) A heuristic lot size reorder point model for decaying inventories. Management Science 25:90–97

Nandakumar P, Morton TE (1993) Near myopic heuristics for the fixed-life perishability problem. Management Science 39:1490–1498

Pegels CC, Jelmert AE (1970) An evaluation of blood-inventory policies: a Markov chain application. Operations Research 18:1087–1098

Pegels CC et al. (1977) An analysis of selected blood service policy changes. Medical Care XV:147–157

Perry D, MJM Posner (1998) An (s-1,s) inventory system with fixed shelf life and constant lead times. Operations Research 46: Supp. #3, S65–S83

Pierskalla WP (1967a) Optimal issuing policies in inventory management. Management Science 13:395–412

Pierskalla WP (1967b) Inventory depletion management with stochastic field life functions. Management Science 13:877–886

Prastacos GP, Brodheim E (1980) PBDS: A decision support system for regional blood management. Management Science 26:451–463

Raafat F (1991) Survey of literature on continuously deteriorating inventory models. Journal of the Operational Research Society 42:27–37

Schmidt CP, Nahmias S (1985) (S-1, S) policies for perishable inventory. Management Science 31:719–728

Smith LA (1975) Simultaneous inventory and pricing decisions for perishable commodities with price fluctuation constraints. INFOR 13:82–87

Tekin E, Gurler U, Berk E (2001) Age-based vs. stock level control policies for a perishable inventory system. European J. of Operational Research 134:309–329

Weiss HJ (1980) Optimal ordering policies for continuous review perishable inventory models. Operations Research 28:365–374

Van Zyl GJJ (1964) Inventory control for perishable commodities. Unpublished doctoral dissertation, University of North Carolina, Chapel Hill, NC

Veinott AF Jr (1960) Optimal ordering issuing and disposal of inventory with known demand. Unpublished doctoral dissertation, Columbia University, New York, NY

Veinott AF Jr (1978) Private communication

Veinott (1965) Optimal policy for a multi-product, dynamic, non-stationary inventory problem. Management Science 12:206–222

Wagner HM, Whitin TM (1958) Dynamic version of the economic lot size model. Management Science 5:89–96

Whitin (1957) The Theory of Inventory Management. Princeton University Press, Princeton, NJ

Zehna P (1962) Inventory depletion policies. in Studies in Applied Probability and Management Science, Arrow, Karlin, Scarf (eds.), Stanford University Press, Stanford, CA. 263–287

About the Author

Steven Nahmias is a Professor of Operations and Management Information Systems in the Leavey School of Business and Administration at Santa Clara University. He holds a B.A. in Mathematics and Physics from Queens College, a B.S. in Industrial Engineering from Columbia University, and M.S. and Ph.D. degrees in Operations Research from Northwestern University. Prior to joining the faculty of the Leavey School at Santa Clara in 1979, he served on the faculties of University of Pittsburgh and Stanford University.

Dr. Nahmias has an international reputation for his research on stochastic inventory models, especially those dealing with perishable inventories. He is the author or co-author of more than 50 scientific articles published over the past 40 years. These have appeared in *Operations Research, Management Science, SIAM Journal on Applied Mathematics*, and many others. He has served as an Area Editor for Supply Chain Logistics for *Operations Research*, Associate Editor for both *Management Science* and *Naval Research Logistics*, and was a founding Senior Editor for the *M&SOM* journal. At Santa Clara University, he has served as department chair and Director of the Competitive Manufacturing Institute.

He is a member of Alpha Pi Mu, Beta Gamma Sigma, Omega Rho and Sigma Chi honoraries, and is listed in Who's Who in Engineering, Who's Who in California, American Men and Women of Science, and Who's Who in Frontier Science and Technology. He received First Prize in the 39th Nicholson Student Paper Competition (1971), Second Prize in the TIMS Student Paper Competition (1972), and is the recipient of The University Award for Sustained Excellence in Scholarship from Santa Clara University (1998). In June of 2010, Dr. Nahmias's lifetime contributions were recognized when he was made a Fellow of the Manufacturing and Service Operations Management Society.

He is the sole author of *Production and Operations Analysis* originally published by Richard D. Irwin in 1989. The text is currently in sixth edition which was published by McGraw Hill with a 2009 publication date. This book has been adopted at more than 100 outstanding universities and colleges worldwide including Stanford, Berkeley, Harvard, Penn, Columbia, MIT, and many others. It has been translated into Spanish (two editions), Hebrew, and Chinese. The seventh edition is due out in 2013.

In addition to his academic activities, Dr. Nahmias has served as a consultant to a variety of firms and agencies including Litton Industries, Xerox Corporation, IBM Corporation, LEX Automotive, Department of Transportation of Santa Clara County, Tropicana, and others. He enjoys golf and is currently handicap chairman at Sharon Heights Country Club. For exercise, he swims and bikes. He is also an accomplished jazz trumpeter. At the current time, he performs with the Quintessence quintet, Tuesday Nite Live big band, and the Bob Enos Soundwave Orchestra.

Index