Describing Service in Independent Demand Inventory Systems

RICHARD EHRHARDT
Department Information Systems and Operations Management
The University of North Carolina at Greensboro
Joseph M. Bryan School of Business and Economics
Greensboro, NC 27402-6165
USA

Abstract: - We discuss the notion of service in independent demand inventory systems with random demand, and note several areas where existing pedagogy can be improved. We survey a collection of 23 operations management textbooks and recommend that different treatments be made. Specifically we recommend that the concept of Cycle Service, sometimes referred to as Type I Service should be abandoned as it is typically treated in operations management textbooks. We also recommend making a fundamental change in the way that periodic review systems are depicted. The recommendations are made on the basis of general concepts of the notion of customer service and on the basis of long-standing theoretical results from the field of mathematical inventory theory.

Key-Words: - Operations Management, Inventory Systems, Education, Pedagogy.

1 Introduction
Consider two inventory managers who stock a particular model of television for retail customers. Both managers place a replenishment order whenever stock falls below 5, but the first manager orders one unit each time, while the second always orders a lot size of 20 units. Also assume that both managers face the same demand and replenishment environments. Which manager provides better service, on average, to their customers? It’s quite clear that over the long run the second manager will turn away fewer customers empty-handed than the first. It’s also clear that the second manager will satisfy a higher percentage of the long run quantity of demand.

The standard textbook measure of service, however, would declare that both managers operate their systems at the same service level. The reason for this discrepancy is that the standard measure of service is defined as the probability of satisfying demand only during the order leadtime, when the risk of stocking out is the greatest. It is incapable of distinguishing between higher risk systems that order more frequently and lower risk systems that order less frequently and carry larger amounts of inventory on average.

Why then do most of our operations management text books say that both managers provide the same level of service? Why is it that even top-level professional journals publish articles that rely on the same questionable measure of service? These questions may be impossible to answer definitively, but it is useful to assess the state of operations management pedagogy in this area and to offer constructive criticism.

We confine attention to models with probabilistic demand that is independently and identically distributed across different intervals of time. Replenishment costs have fixed (setup) and linear components, and the lead time may be fixed or random. Inventory holding costs are assumed to accrue at a rate proportional to the amount of stock.

Virtually every OM text prescribes an ordering policy of the (r,Q) form for independent demand systems under continuous review. An (r,Q) policy prescribes an order of size Q (usually the EOQ) whenever the inventory position falls below the reorder point r. While all these assumptions are rather standard, models differ substantially in the way stockouts are treated and how performance is measured. That is the focus of this paper.

We give an overview of service description in Section 2, specifically discussing basic concepts of customer service in inventory systems, and how the method for reviewing inventory levels is involved. We proceed to discuss the various ways that operations management textbooks treat service specification and the appropriate calculation of safety stock in Section 3. Finally, we draw conclusions in Section 4.
2 Issues in Describing Service

The general underlying issue is the description and management of a random environment, something that is becoming increasingly important in modern times. It is essential that we do a good job in teaching the art and science of probabilistic reasoning. Two topics discussed below are especially relevant in this regard.

2.1 Continuous vs. Periodic Review

The issue of defining service is clouded in most operations management texts by the confusingly artificial distinction between systems with continuous review and those with periodic review. The distinction between models and the world they attempt to describe is often lost.

The real world is in fact both periodic and continuous in nature. An inventory system may be “perpetual,” but it will usually operate in an environment that has at least some periodic elements. Examples include hours when businesses are open or closed, daily and weekly patterns of events, shipping schedules and so on. A shipment to a customer may not be considered late (unsatisfied demand) if it arrives within one or two business days. Therefore, it may be more accurate to describe a perpetual system as having daily review periods rather than continuous review. Other examples of natural periodicities can easily be postulated.

Regardless of whether a system is more properly described by a periodic or continuous review, the form of an appropriate replenishment policy is the same. For example, if backlogging of unmet demand is permitted and one can identify a cost of backlogging each unit of unmet demand, the policy that minimizes long run expected costs is usually of the (s,S) form regardless of whether periodic or continuous review is assumed [1] [7]. That is, an order is placed whenever the inventory position (on hand plus on order) y is less than s, and the order size is S-y. The general idea is that in the presence of setup costs, one doesn’t place an order unless the inventory position is sufficiently low to justify the fixed cost of ordering. If it is sufficiently low, then the size of the order should be adjusted to compensate for just how little inventory is available at that time. This result has common sense appeal in addition to being theoretically optimal for the cost-minimization model.

Therefore, when presenting models for managing independent demand inventories, virtually all distinctions between periodic and continuous time should be dropped. What matters more is the lumpiness of demand in determining when positions far below the reorder point can occur. If demands occur only one at a time, then replenishment orders can be placed when the reorder point is exactly reached and a fixed order size will be effective. Demand may occur in lumps either because it happens that way in continuous time or because the review period is long. Then something more like an (s,S) policy can be more effective.

The (s,S) policy form is sometimes referred to as a two-bin policy, meaning that stock can be thought of as being stored in two bins, one of capacity s and the other of capacity S-s. Stock is issued from the second bin until it is empty, at which time the first bin is brought into play and a replenishment order is triggered at the same time. Operations management textbooks sometimes refer to this as a simple policy for low-value items, but that must assume that the method for computing s and S is a crude rule of thumb [8] [11] [13]. Obviously this two-bin approach can be very sophisticated if careful forecasts and advanced computation methods are employed.

2.2 Conceptualizing Service

Describing the nature and degree of customer service that is provided by an inventory management system is inherently challenging because it requires a summary of random events that occur at many points in time. Therefore, the use of a single measure to define service is a gross simplification. Most operations management textbooks fall short in getting this point across.

Service should be defined in a way that reflects the system’s ability to serve customers. The best single measures for defining and prescribing service are probably (1) the fill rate (expected fraction of demand satisfied immediately from stock) and (2) the expected cost of shortages per unit time (if such costs can be readily measured), because both of these measures include the frequency and amount of shortages. The next best measures are probably those that reflect the frequency of stockouts, including (3) the proportion of time that the system is out of stock and (4) the expected number of stockout events per unit of time.

The most common measure of service seen in operations management textbooks is not in any of the above 4 categories. It is sometimes referred to as “Type I service” or “Cycle Service,” meaning the likelihood of stocking out between the time a replenishment order is placed and the time at which it is delivered. This is not merely a less desirable measure of service; it does not truly measure service
at all in the usual meaning of the word “measure.” As illustrated in the introduction, it is possible for one item to have more customers denied satisfaction and a greater proportion of its demand unsatisfied than another even if it is stacked according to a rule that specifies higher cycle service than the other. This will happen whenever the item with higher cycle service has an order quantity that is sufficiently smaller than the item with lower cycle service. This conception of service is essentially misleading unless it is used in conjunction with another performance measure as a way to place a upper bound on risk.

3 A Survey of OM Texts

The list of references includes 23 general operations management textbooks [3] [4] [5] [6] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]. All describe the same basic style of writing. It is not meant to be critical of the textbooks. Some may no longer be in print, while others may not be the most recent edition. Most professionals in the field would agree, though, that it is a large enough sample to be a good representation of current pedagogy.

The following subsections give a summary of how the issues described above are treated in these textbooks. The summary is meant to describe the technical content of the books, and is not meant to imply any judgment on the quality of exposition or style of writing. It is not meant to be critical of individual authors, either. It is merely a description of the state of pedagogy, as the observed patterns are quite general.

3.1 Concepts of Service in OM Texts

The list of references includes 23 general operations management textbooks [3] [4] [5] [6] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]. All describe the same basic model for computing a reorder point that delivers the desired level of service. The service level $\alpha$, is defined as the probability that all demand will be satisfied during the time interval between placing the order and its delivery, and is often referred to as “Type I service” or “Cycle Service.” The model operates in continuous time, and is of the $(r,Q)$ form, that is, an order of fixed size $Q$ is placed whenever the inventory level drops to the reorder point $r$. The value of $Q$ is determined by the EOQ formula, and the reorder point $r$ is set equal to the $\alpha$th percentile of the probability distribution of demand during a replenishment leadtime.

The few notable exceptions are in Martinich [15], Melnyk and Denzler [16] and Nahmias [17]. Martinich describes the concept of Cycle Service but does not develop the model, stating that it is “not a good measure of customer service because it does not consider the frequency or magnitude of stockouts” [15]. Melnyk and Denzler develop both the Cycle Service and Fill Rate models but discount Cycle Service as of more concern to stockroom personnel than to customers [16]. Nahmias criticizes cycle service as overly simple and likely to lead to inconsistency across different items being stocked [17]. His development of independent demand models, described in more detail below, stands out as very rigorous and consistent with the status of mathematical inventory theory. This should not be surprising since he is a leading scholar in the field.

All other cited textbooks develop the standard Cycle Service model, but there are a number of additions and departures from that model as noted in the following subsection.

3.1.1 Other Departures from Cycle Service

A number of textbooks develop the Fill Rate model in addition to the Cycle Service model [14] [15] [16] [25], using tabled values of the normal distribution loss function in a method that was introduced in 1959 by Brown [2]. Nahmias develops two iterative models to minimize the sum of holding and setup costs subject to a constraint on fill rate or the proportion of time the system is out of stock [17]. This leads to order quantities different from the EOQ.

Some textbooks discuss models with a shortage penalty cost, and show simple numerical examples of a cost minimization procedure, using either marginal analysis or a payoff table approach [10] [11] [12] [15]. Only two textbooks make mention of the model that actually minimizes long-run expected costs, the $(s,S)$ model. Starr [24] describes it as an advanced system that “combines the advantages of perpetual and periodic ordering,” but states that the difficulty in computing it makes it appropriate only for large installations such as the U.S military. Nahmias [17] also describes the $(s,S)$ model and cites references that evaluate approximation procedures for computing the policy.

Other textbooks mention that management should consider a tradeoff curve between the service level (usually Cycle Service) and the expected inventory level [5] [6] [21] [22], sometimes suggesting alteration of the order quantity to do so [21] [22].
Some textbooks make brief mention of other measures of service without developing a model [11][13][20][21][22][26]. Starr discusses a novel approach for modifying the Cycle Service model, judging the desirability of the chosen value for $\alpha$ by imputing its equivalent stockout cost, based on the resulting level of average inventory.

### 3.2 Periodic Review Models in OM Texts

While most OM texts make a great distinction between continuous review models and periodic review models, there is actually very little difference between the two. Periodic review models are analyzed in discrete time and the demand distributions tend to be more “lumpy,” but they have the same forms of optimal ordering policies as in the continuous review case (assuming the same set of assumptions about costs, service and the replenishment process). This fact does not come across in most OM texts, as is discussed in Section 2.

A few of the cited textbooks do not develop periodic review models at all [5][16][20][23]. The only cited textbook that treats periodic review systems rigorously is Nahmias [17].

Virtually every operations management textbook in the cited references prescribes an order-up-to model for managing inventory in periodic review systems. That is, whenever inventory is reviewed an order is placed so as to bring inventory up to a fixed value. This value is sometimes called a “base stock level” or “single critical number” in the mathematical inventory literature. This is not necessarily cost efficient when there is a setup cost of ordering, especially when the review period is relatively short and/or the setup cost is relatively large.

Having made those observations, it must be pointed out that several textbooks modify the standard periodic review model so that it will perform more reasonably. The modification is to set the review period at a value that would tend to make the replenishment orders roughly equal to the EOQ on average [10][13][18][21][24][26]. This is done by setting the review period equal to the EOQ divided by the average demand rate. While this may have desirable economic effects, it is not compatible with other important operational reasons for setting the review period at a different value. These might include multi-item joint replenishment incentives, supplier delivery schedules, and internal production scheduling, for example.

### 4 Conclusion

We have pointed out a number of ways in which operations management textbooks could improve their treatment of (1) service specification in independent demand inventory systems and (2) periodic review vs. continuous review systems.

We recommend that cycle service be abandoned in favor of a more robust measure of customer service. The calculation of fill rate in the manner of Brown [2] deserves serious consideration, especially since it is now very easy to implement with spreadsheet functions. A second approach which would be an improvement over cycle service would be to convert cycle service $\alpha$ to a measure that more consistently measures service to the customer, such as the mean time between stockouts $\alpha D/Q$.

Another recommendation that deserves serious consideration at this point in the evolution of our field is to treat (s,S) models in textbooks. Two reasons support this recommendation. First, there are now easily computed approximations for optimal values of $s$ and $S$ [7][27]. Second, although the (s,S) model is based on a shortage penalty cost, optimal policies provide an average probability of being in-stock of $p/(p+h)$, where $p$ and $h$ are the unit costs of shortage and holding. Therefore if a manager specifies a service target of $\beta$ for the average probability of being in stock, one can impute an equivalent penalty cost of $\beta h/(1-\beta)$. The resulting optimal (s,S) policy would then yield the desired service level of $\beta$.

Note that this paper focuses on textbooks that are used primarily in undergraduate programs in business administration. That is not meant to imply that this is the only area where improvements to pedagogy are in order. There are examples in graduate-level texts and professional journal articles where cycle service and other issues discussed in this paper exist. An extension of this survey to those categories of publications would be of interest.

### References:

