

Inventory Theory

Suresh P. Sethi

Jindal School of Management, The University of Texas at Dallas, Richardson, TX, USA

Abstract

This entry is a brief survey of classical inventory models and their extensions in several directions such as world-driven demands, presence of forecast updates, multi-delivery modes and advanced demand information, incomplete inventory information, and decentralized inventory control in the context of supply chain management. Important references are provided. We conclude with suggestions for future research.

Keywords

EOQ model, Newsvendor model, Base stock policy, (s, S) policy, Incomplete information

Introduction

Optimal inventory theory deals with managing stock levels of goods to effectively meet the demand of those goods. Because of the huge amount of capital that is tied up in inventory, its management is critical to the profitability of firms. A systematic analysis of inventory problems began with the development of the classical economic order quantity (EOQ) formula of Ford W. Harris in 1913. A substantial amount of research was reported in 1958 by Kenneth J. Arrow, Samuel Karlin, and Herbert Scarf, and much more has accumulated since then. Books on the topic include Zipkin (2000), Porteus (2002), Axsäter (2006), and Bensoussan (2011).

In this entry, we review single- and multi-period models with deterministic, stochastic, partially observed demand for a single product. In these models, our aim is to decide on the time of the orders and the order quantities. The time between issuing an order and its receipt is called the lead time. For most of this review, we will assume the lead time to be zero, and the reader can consult the referenced books for nonzero lead time extensions and other topics not covered here.

Deterministic Demand

We will describe two classical models: the EOQ model and the dynamic lot size model.

The EOQ Model

This basic and most important deterministic model is concerned with a product that has a constant demand rate D in continuous time over an infinite horizon. No shortages are allowed. The costs consist of a fixed setup/ordering cost K and a holding cost h per unit of average on-hand stock per

unit time. The production/purchase cost per unit time is a sunk cost since there is no choice of a total amount to produce, and hence it can be ignored. Although dynamic, the model can be reduced to a static model by a simple argument of periodicity. Moreover, it is obvious that one should never produce or order except for when the inventory level is zero, and one should order the same lot size Q each time the inventory level reaches zero. Since the average inventory level over time is $Q/2$ and the number of setups is D/Q per unit time, the long-run average cost to be minimized is $KD/Q + hQ/2$. The optimal policy that minimizes this cost, obtained using the first-order condition, is to order the lot size

$$Q = \sqrt{\frac{2KD}{h}} \quad (1)$$

every time the inventory level reaches zero. Harris (1913) introduced the model. Erlenkotter (1990) provides a historical account of the formula, and Beyer and Sethi (1998) provide a mathematically rigorous proof involving quasi-variational inequalities (QVI) that arise in the course of dealing with continuous-time optimization problems involving fixed costs.

The Dynamic Lot Size Model

This is an analogue of the EOQ model when the demand varies over time. Wagner and Whitin (1958) developed it in the discrete-time finite horizon framework. With $D(t)$ denoting the demand in period t and other costs similar to those in the EOQ model, they showed that there exists an optimal policy in which an order will be issued just as the inventory level reaches zero, except for the first order. This policy is called the **zero-inventory policy**. With this in hand, the problem reduces to selecting only the order times. This is accomplished by applying a shortest path algorithm. Moreover, there are forward (recursion) procedures for solving the problem.

An important feature of this model is that in most cases, one can detect a **forecast horizon** which essentially separates earlier periods from later ones. More specifically, T is a forecast horizon if the first order in a T horizon problem remains optimal in any finite horizon problem with horizon longer than T , regardless of the demands beyond the period T . For an extensive bibliography of this literature, see Chand et al. (2002).

Stochastic Demand

We shall discuss three classical models and some of their extensions.

The Single-Period Problem: The Newsvendor Model

The problem of a newsvendor is to decide on an order quantity of newspapers to meet a stochastic demand at a minimum cost. If the realized demand is larger than the ordered quantity, it is lost and there is an opportunity loss of c_u (selling price minus purchase cost) for each paper short. On the other hand, for each paper ordered but not sold, there is an opportunity loss of c_o (purchase cost plus holding cost). The newsvendor conceptualizes the decision by each additional paper as a separate marginal contribution. The first is almost certain to be sold. Each additional paper is

less likely to be sold than the previous one. Thus, each additional paper will be worth somewhat less, and the marginal paper at the optimum should be worth exactly zero. Thus, c_u times the probability of selling the marginal paper minus c_o times the probability of not selling it should equal zero. Now, if F denotes the cumulative probability distribution function of the demand D , then clearly the optimal order quantity Q satisfies $c_o \cdot F(Q) - c_u \cdot (1 - F(Q)) = 0$, which gives us the famous newsvendor formula for the optimal order quantity

$$Q = F^{-1} \left(\frac{c_u}{c_u + c_o} \right), \quad (2)$$

where $c_u/(c_u + c_o)$ is known as the critical fractile.

If p denotes the unit sale price, c the unit cost, and h the holding cost per unit per unit time, then $c_u = p - c$ and $c_o = c + h$, and therefore, the critical fractile can be expressed as $(p - c) / (p + h)$. An extension of the newsvendor formula to allow for a unit cost g of lost goodwill and a unit salvage value s received at the end of the period for each unit not sold is immediate. If we let $\alpha > 0$ denote the periodic **discount** factor, then $c_u = p + g - c$ and $c_o = c + h - \alpha s$ and the critical fractile becomes $(p + g - c) / (p + g + h - \alpha s)$, and therefore,

$$Q = F^{-1} \left(\frac{p + g - c}{p + g + h - \alpha s} \right). \quad (3)$$

The newsvendor model has been used extensively in the context of supply chain management with multiple agents maximizing their individual objectives. In this case, inefficiencies arise due to double marginalization. Then, a question of appropriate contracts that can lead to the first-best solution, or coordinate the supply chain, becomes important. Cachon (2003) surveys this literature.

Multi-period Inventory Models: No Fixed Cost

The newsvendor model is a single-period model, and its multi-period generalization requires that the inventory not sold in a period is carried over to the next period. This results in the multi-period inventory model with lost sales. It is assumed that demand in each period is independent and identically distributed (i.i.d.) with F denoting its cumulative probability distribution function. A rigorous analysis requires the method of **dynamic programming**, and it shows that there is a stock level S_t called *base stock* in period t , that we would ideally like to have at the beginning of period t . Thus, the optimal policy in period t , called **the base stock policy**, is to order

$$Q_t(x) = \begin{cases} S_t - x & \text{if } x < S_t, \\ 0 & \text{if } x \geq S_t. \end{cases} \quad (4)$$

In the special case when the terminal salvage value of an item is exactly equal to its cost c , it is possible to come up with the optimal policy using intuition. Since we do not need to salvage unused items in the multi-period setting, one can argue that an item carried over to the next period is worth its purchase cost c . Therefore, its presence means that the next period will need to order one less and thus save an amount c . In the last period, when there is no next period, our terminal salvage value assumption also guarantees a leftover item's worth to be also c . Thus, we can modify (3) and obtain a stationary base stock level

$$S = F^{-1} \left(\frac{p + g - c}{(p + g - c) + (c + h - \alpha c)} \right) = F^{-1} \left(\frac{p + g - c}{p + g + h - \alpha c} \right) \quad (5)$$

for each period t .

Thus, the elimination of the endgame effect delivers us a *myopic policy*, a policy optimal in the single-period case to be also optimal in the dynamic multi-period setting. A more general concept than the optimality of a myopic policy is that of the forecast horizon mentioned earlier in the context of the dynamic lot size model.

Sometimes, when the demand exceeds the on-hand inventory in the period, the demand is not lost but backlogged. In this case, each unit of backlogged demand is satisfied in the next period, and unit revenue p is recovered, but a unit backlogging cost b is incurred, due to expediting, special handling, delayed receipt of revenue, and loss of goodwill. Thus, $c_u = b - (1 - \alpha)c$, where the second term represents the savings due to postponing the purchase of the backlogged demand unit by one period, and $c_o = c + h - \alpha c$ as in (4). This gives us the base stock level

$$S = F^{-1} \left(\frac{b - (1 - \alpha)c}{b + h} \right), \quad (6)$$

which can be used in (5) to give the optimal policy.

Sometimes it is possible to have multiple delivery modes such as fast, regular, and slow as well as demand forecast updates. Then, at the beginning of each period, on-hand inventory and demand information are updated. At the same time, decisions on how much to order using each of the modes are made. Fast, regular, and slow orders are delivered at the ends of the current, the next, and one beyond the next periods, respectively. In such models, a modified base stock policy is optimal only for the two fastest modes. For details and further generalization, see Sethi et al. (2005).

An important extension includes serial inventory systems where stage 1 receives supplies from an outside source and each downstream stage receives supplies from its immediate upstream stage. Clark and Scarf (1960) introduced the notion of the echelon inventory position at a stage to consist of the stock at that stage plus stock in transit to that stage plus all downstream stock minus the amount backlogged at the final stage. Then, the optimal ordering policy at each stage is given by an echelon base stock policy with respect to the echelon inventory position at that stage. It is known that assembly systems can be reduced to a serial system. Details can be found in Zipkin (2000).

Multi-period Inventory Models: Fixed Cost

When there is a fixed cost of ordering, it is clear that it would not be reasonable to follow the base stock policy when the inventory level is not much below the base stock level. Indeed, Scarf (1960) proved that there are numbers s_t and S_t , $s_t < S_t$, for period t such that the optimal policy in period t is to order

$$Q_t(x) = \begin{cases} S_t - x & \text{if } x \leq s_t \\ 0 & \text{if } x > s_t. \end{cases} \quad (7)$$

Such a policy is famously known as an (s, S) policy.

When the demands are not i.i.d., the model has been extended to Markovian demands. In this case, there is an exogenous Markov process, and the distribution of the demand in each period depends on the state of the Markov process, called the demand state, in that period. It can be shown that the optimal policy in period t is (s_t^i, S_t^i) , where i denotes the demand state in the period. Such a policy is also called a state-dependent (s, S) policy. Further details are available in Beyer et al. (2010). Recent advances in information technology have allowed managers to obtain advance demand information in addition to forecast updates. In such cases, a state-dependent (s, S) policy can be shown to be optimal. For details, refer to Ozer (2011).

The Continuous-Time Model: Fixed Cost

The marriage of the two classical results (1) and (7) is accomplished by Presman and Sethi (2006) in a continuous-time stochastic inventory model involving a demand that is the sum of a constant demand rate and a compound Poisson process. The optimal policies that minimize a discounted cost or the long-run average cost are both of (s, S) type. The (s, S) policy minimizing the long-run average cost reduces to the EOQ formula when the intensity of the compound Poisson process is set to zero. And when the constant demand component vanishes, the model reduces to the continuous-review stochastic inventory model with fixed cost and compound Poisson demand.

Incomplete Inventory Information Models (i3)

A critical assumption in the vast inventory theory literature has been that the level of inventory at any given time is fully observed. The celebrated results (1) and (7) have been obtained under the assumption of full observation. Yet the inventory level is often not fully observed in practice, for a variety of reasons such as replenishment errors, employee theft, customer shoplifting, improper handling and damaging of merchandise, misplaced inventories, uncertain yield, imperfect inventory audits, and incorrect recording of sales. In such an environment of incomplete information, inventories are known to be partially observed and most of the well-known inventory policies including (1) and (7) are not even admissible, let alone optimal. In such cases, Bensoussan et al. (2010) show that the dynamic programming equation can be written in terms of the unnormalized conditional probability of the current inventory level given past observations, referred to as signals, instead of just the inventory level in the full observation case. Furthermore, one can write the evolution of the conditional probability in terms of its current value, the current order, and the current observation. However, there are no longer simple optimal policies except in cases of information delay reported in Bensoussan et al. (2009) where modified base stock and (s, S) policies are shown to be optimal.

Summary and Future Directions

We briefly describe some classical results in inventory theory. These are based on full observation. Some recent work on inventory models under incomplete information is reported. This work leads to a number of new research directions, both theoretical and empirical as reported in Sethi (2010). It would be of much interest to know the industries where the i3 problem is serious enough to

warrant the difficult mathematical analysis required. Furthermore, how are the observed signals related to the inventory level? It is also clear from the reviewed literature that there are no simple optimal policies for most i3 problems, so it would be important to develop efficient computational procedures to obtain optimal solutions or to specify a class of simple implementable policies and optimize within this class. An important benefit of solving i3 problems optimally is the provision of an economic justification for technologies such as RFID that may reduce inaccuracies in inventory observations.

Another area of research would be to study multi-period multi-agent supply chains with a stochastic inventory dynamics. While these can be formulated as dynamic games, there are a number of equilibrium concepts to deal with, depending on the information the agents have. Some of them are time consistent or subgame perfect and some are not. Regardless, there are inefficiencies that arise from these decentralized game settings, and developing contracts for coordinating dynamic supply chains remains a wide open topic of research.

Cross-References

- [Nonlinear Filters](#)
- [Stochastic Dynamic Programming](#)

Bibliography

- Arrow KJ, Karlin S, Scarf H (1958) Studies in the mathematical theory of inventory and production. Stanford University Press, Stanford
- Axsäter S (2006) Inventory control. Springer, New York
- Bensoussan A (2011) Dynamic programming and inventory control. IOS Press, Washington, DC
- Bensoussan A, Cakanyildirim M, Feng Q, Sethi SP (2009) Optimal ordering policies for stochastic inventory problems with observed information delays. *Prod Oper Manag* 18(5):546–559
- Bensoussan A, Cakanyildirim M, Sethi SP (2010) Filtering for discrete-time Markov processes and applications to inventory control with incomplete information. In: Crisan D, Rozovsky B (eds) *Handbook on nonlinear filtering*. Oxford University Press, Oxford, UK, pp 500–525
- Beyer D, Cheng F, Sethi SP, Taksar MI (2010) Markovian demand inventory models. *International series in operations research and management science*. Springer, New York
- Beyer D, Sethi SP (1998) A proof of the EOQ formula using quasi-variational inequalities. *Int J Syst Sci* 29(11):1295–1299
- Cachon GP (2003) Supply chain coordination with contracts. In: de Kok AG, Graves S (eds) *Supply chain management-handbook in OR/MS*, chapter 6. North-Holland, Amsterdam, pp 229–340
- Chand S, Hsu VN, Sethi SP (2002) Forecast, solution and rolling horizons in operations management problems: a classified bibliography. *Manuf Serv Oper Manag* 4(1):25–43
- Clark AJ, Scarf H (1960) Optimal policies for a multi-echelon inventory problem. *Manag Sci* 6(4):475–490
- Erlenkotter D (1990) Ford Whitman Harris and the economic order quantity model. *Oper Res* 38(6):937–946
- Harris FW (1913) How many parts to make at once. *Fact Mag Manag* 10(2):135–136, 152. Reprinted (1990) *Oper Res* 38(6):947–950

- Ozer O (2011) Inventory management: information, coordination, and rationality. In: Kempf KG, Keskinocak P, Uzsoy R (eds) Planning production and inventories in the extended enterprise. Springer, New York
- Porteus EL (2002) Stochastic inventory theory. Stanford University Press, Stanford, CA
- Presman E, Sethi SP (2006) Inventory models with continuous and Poisson demands and discounted and average costs. *Prod Oper Manag* 15(2):279–293
- Scarf H (1960) The optimality of (s, S) policies in dynamic inventory problems. In: Arrow KJ, Karlin S, Suppes P (eds) Mathematical methods in the social sciences. Stanford University Press, Stanford, pp 196–202
- Sethi SP (2010) i3: incomplete information inventory models. *Decis Line* Oct:16–19
- Sethi SP, Yan H, Zhang H (2005) Inventory and supply chain management with forecast updates. Springer, New York
- Wagner HM, Whitin TM (1958) Dynamic version of the economic lot size model. *Manag Sci* 5:89–96
- Zipkin P (2000) Foundations of inventory management. McGraw Hill, New York