i3: Incomplete Information Inventory Models

by Suresh P. Sethi, University of Texas at Dallas

Inventory control is among the most important topics in operations research because of large investments in inventory and their effect on the profitability of the firms. A systematic analysis of inventory problems began with the development of the classical EOQ formula of Ford W. Harris in 1913, and a substantial amount of research was reported in 1958 by Kenneth J. Arrow, Samuel Karlin, and Herbert Scarf. A critical assumption in this vast literature has been that the level of inventory at any given time is fully observed. Some of the most celebrated results such as the optimality of basestock or (s,S) policies have been obtained under the assumption of full observation. Yet the inventory level is often not fully observed in practice, as elaborated below. In such an environment of incomplete information, inventories are known to be partially observed and most of the well-known inventory policies are not even admissible, let alone optimal.

Reasons for partial observation of inventory levels may include replenishment errors, employee theft, customer shoplifting, improper handling of damaged merchandise, misplaced inventories, uncertain yield, imperfect inventory audits, and incorrect recording of sales. Details on some of these causes and their consequences are provided below and further discussion is available.[1-3]

Incorrect recording of sales. Unintentional mistakes happen from time to time in the recording of sales transactions. One example involves customer checkout at a grocery store. If a customer buys two different types of soup, each at the same price, the sales clerk often scans only one type of soup twice. In such cases, the recorded inventory levels of the items involved will differ from their actual levels.

Misplaced inventory. When part of the inventory on hand is misplaced, it is not available to meet the demand until it is found. Often, misplaced inventories are not immediately found, so they remain unobserved by the inventory manager (IM); hence, the total inventory available to meet the demand is partially observed. Misplaced inventory can be quite large and have a significant impact on the bottom line. It was reported that customers of a leading retailer could not find 16 percent of the items in their stores because the items were misplaced.[1] Misplacement of these items reduced the retailer's profit by roughly 25 percent.

Misplacement is quite common when the location of items in storage is altered dynamically. Keeping items in fixed locations may improve record accuracy, but this may also lead to inefficient space utilization; dynamically located items may be stored in more than one location. Recent trends in supply chain management such as cross-docking also cause dynamic locations.

Spoilage. Products can naturally lose their properties while they are held in the inventory. Some examples of items with limited lifetimes are drugs, chemicals, and food products. If the lifetime of a product is limited and not immediately observed, then the actual inventory is less than the recorded inventory, and it is partially observed.

If the lifetime is deterministic, as in the case of drugs, an implementation of RFID (radio frequency identification) tags called SMC (smart medicine cabinet) can be used to track expired drugs. Thus,

Suresh P. Sethi
is Charles & Nancy Davidson Distinguished Professor of Operations Management and director of the Center for Intelligent Supply Networks at The University of Texas at Dallas. He has written seven books and published more than 350 research papers in the fields of manufacturing and operations management, finance and economics, marketing, and optimization theory. He serves on the editorial boards of several journals including Production and Operations Management *and* SIAM Journal on Control and Optimization*. Honors include: SIAM Fellow, POMS Fellow (2005), INFORMS Fellow (2003), AAAS Fellow (2003), IEEE Fellow (2001), IITB Distinguished Alum (2008), SIAM Fellow (2009). Two conferences were organized and two books edited in his honor in 2005-2006.*
sethi@utdallas.edu

SMC can make drug spoilage fully observed. Wal-Mart has started using smart RFID tags to track individual pairs of jeans and underwear for better inventory control. If successful, RFID tags will be rolled out on other products at Wal-Mart's more than 3,750 U.S. stores.[13] Some initial work has already been done on justification for investments in such technologies. Needless to say, such economic analyses require evaluation of the optimal cost under partial observations.[10-11]

As an example of random lifetime, consider the number of batteries in a Sears retail store. Only when these batteries are inspected (say by measuring their voltage, one by one) does the inventory level of fully functioning batteries become known. When spoiled inventory is observed immediately, the associated model, in spite of being a challenge to work with, is fully observed.

In retail stores, customers can damage products, making them unsuitable for sale. Some examples of this are: tearing a package to try on a contained clothing item, erasing software on a computer used for display, spilling drinks on clothes, or scratching a car during a test drive. If such damages are not detected by the IM, the actual inventory is not fully observed.

Product Quality and Yield. When the product quality is low or a production process has a low yield, the actual inventory is not known. Receipts at a warehouse can include products that are defective or non-conforming to quality standards. It is often the case that nonconformance of a product is not immediately observed by the IM. Receipts are usually added to the inventory at the warehouse without full inspection. As a result, the inventory on record may consist of both non-defective products (available to meet customer demand) and defective products (not fit for sale). Since the defective products are not immediately observed, the actual (non-defective) inventory becomes partially observed.

If production lead times are long, an IM may have to place a particular order before observing the yields from previous orders, so that production of the order is completed by a given due date. Thus, incomplete information about inventory can be caused by due dates, long production lead times, and uncertain process yields.

Theft. Items in the inventory can be stolen by thieves who violently break into inventory storage, by warehouse employees who calmly pilfer, or by customers who shoplift. Since break-ins are generally investigated, they are usually observed and therefore not relevant for our study. I focus more on continuous pilferage and shoplifting, because they are not always observed without inventory inspections. Instances of theft at furniture retailers and food wholesalers have been documented.[4] Theft is a major problem resulting in value loss and inaccurate inventory records. Thus, the IM relying on inventory records over-estimates the available inventory until a stockout occurs. In this case, there are shortage costs in addition to costs incurred from reordering, expediting, and receiving replacement items. Typically, costs associated with expediting items to urgently meet backlogged demands are higher than regular costs.

Empirical research gives clear evidence of the presence of inventory record inaccuracy in a number of contexts, including government agencies and utilities. A discrepancy between recorded and actual inventory of $142 million, or the equivalent of 21,000 ocean containers, was reported at the well-known apparel retailer The Limited.[5] DeHoratius and Raman found inaccuracies in 65 percent of nearly 370,000 inventory records observed across 37 retail stores of another publicly held retailer.[2] This retailer had annual sales of roughly ten billion dollars and used automated replenishment systems to manage store inventory. Furthermore, approximately 12 percent of the recorded items audited across 37 stores had no inventory on the shelf, yet the recorded inventory quantity was positive.

The direct financial impact of lost merchandise in the retail industry is substantial. Retailers with inaccurate inventory records may incur additional costs due to the uncertainty of their inventory positions. First, retailers may choose to buffer this uncertainty with additional inventory, or else lose sales due to stockouts. Second, inventory record inaccuracy may undermine decision support tools such as automated replenishment and automated demand forecasting systems that do not account for inventory uncertainties.[1] For example, in the presence of inventory record inaccuracy, automated replenishment systems may order when ordering is unnecessary or fail to order when they should. Raman et al. estimate that the retailer loses 10 percent of its current profit due to "freezing," which means that the retailer has no items on the shelf (and hence, no sales) but a sufficiently positive inventory record (and hence, no orders) resulting in a persistent physical stock-out. They go on to say that the problem is common among other retailers.[1] Investments in inventories can run into trillions of dollars, and so better inventory policies can result in huge savings.

By now it should be clear that the incomplete inventory information (i3) problem is quite common in practice, that policies in current use are neither optimal nor applicable, and that there is a great need in many real-life situations for better inventory control policies in systems with partially observed inventories. DeHoratius et al.[6] identify at least three ways to deal with the issue of inventory record inaccuracy:

1. *Prevention:* Reduce or eliminate the root causes of inventory record inaccuracy through the implementation and execution of process improvement.
2. *Correction:* Identify and correct existing inventory record discrepancies through auditing policies.
3. *Integration:* Use inventory planning and decision tools robust enough to account for the presence of record inaccuracy.

In this article, I focus primarily on the third strategy devoted to finding optimal policies in the face of inaccurate inventory records. The analysis of inventory problems under partial observations has been neglected in the past due to its mathematical difficulty. While working with a finite dimensional state space in the case of full observation, one usually has to deal with an infinite dimensional state space in the partial observation setting. In particular, in full observation cases the inventory level is often the state variable, whereas with partial observations, the system state is typically the conditional distribution of the inventory level given the partial observations. In general, the analysis takes place in the space of probability distributions.

Review and Status of the Literature

As noted, there has not been much serious mathematical analysis of the i3 problem until recently, aside from what was done by DeHoratius et al.[6] I shall, therefore, review primarily recent research that has been done over the last five years. This research focuses mainly on i3 problems arising from the inability to observe the demand and the loss of inventory from theft or spoilage. I classify the existing models in which partial observation of the demand depends on how the sales are observed.

Inventory models with only sales observed. There have been a few studies of partial observations in inventory that are devoted to problems in which the demand is observed fully and satisfied only when it is less than the available inventory, and otherwise, only in the event that it is larger than the inventory is observed and the unsatisfied demand is lost. In other words, unmet demand is censored. It is assumed that leftover inventory in a period is salvaged entirely so that every period starts with zero inventory. This assumption decouples the periods as far as the inventory evolution is concerned. However, the periods are still coupled together by the current estimate of the demand distribution. Consequently, the state of the system is the conditional distribution of the demand characterized by a parameter, which is updated in each period based on the partial observation of the demand at that time. Intuitively, if sales are equal to the inventory level, then the IM knows that the demand exceeded sales and will adjust his beliefs about the demand accordingly. Conversely, if sales are less than the available inventory in a period, it is clear that they are equal to the demand. This intuition is captured via Bayesian updating of the IM's belief about the true demand, represented as the conditional distribution of demand. Thus, there is an evolution equation that maps one period's demand distribution to the next period's demand. This evolution is affected by the choice of the order quantity. It is shown that the myopic order quantity is less than the optimal order quantity.[7] Subsequent works of the authors allow for inventory carry over from one period to the next and incorporates demand learning in an adaptive control framework.

Inventory models with no observation of sales. Bensoussan et al.[3] assume that the IM does not observe sales; rather he observes only whether the inventory level is zero or nonzero in each period. They derive the existence of optimal policies in the resulting infinite-state-space control problem. Their paper is part of a greater effort to build a comprehensive theory of inventory control under partial observations. In related works, they study other inventory models with partial observations. In these papers, the dynamic programming equations are highly nonlinear. Using what is known as the unnormalized probability, the authors are able to obtain a Zakai-type equation for the evolution of the conditional distribution. While this transformation does not remove the infinite dimensionality, the linearity permits the proof of the existence and uniqueness of the solution in a number of important inventory control problems with partial observations. Of course, there remain numerical difficulties due to the infinite dimensionality of the state, and computational approximations still require development. Nevertheless, a sound theory is available.

Inventory models with delayed observation of sales. When partial observability arises from information delays in reporting sales, the current inventory level is not observed by the IM. Instead, he observes the exact inventory level of a prior period. In such models when backorders are allowed, Bensoussan et al. are able to find a sufficient statistics and to show that base-stock and $(s.S)$-type policies are optimal.[8]

Inventory models with observation of sales transactions. Inventory record inaccuracy is a significant problem for retailers using automated inventory management systems. DeHoratius et al. explicitly account for the interaction between inventory level uncertainty and observations of sales transactions in a lost sales environment.[6] They develop a model of a lost sales retail inventory system with record inaccuracies and observed sales, and assume that the retailer has uncertainty around the physical inventory available on the shelf, caused by factors such as theft and spoilage. The state of the system in such a model is represented by the inventory's conditional distribution, which they refer to as the Bayesian inventory record. The evolution of the state is based on the observation of the sales transactions. Intuitively, if visible sales (a censored observation of demand) are zero in a period, then a rational IM will realize that the lack of sales may be due to a physical stockout and adjust his beliefs accordingly. Conversely, if visible sales are positive in a period, then the IM knows that the previous physical inventory position could not have been zero. This intuition can be formalized through Bayesian updating of the IM's belief about the true inventory level, and the optimal policies based on the Bayesian inventory record for both replenishment and audit triggering can be obtained. However, computation of these policies is not a simple task, therefore, the authors seek practical heuristic policies that are easily implemented. Simulation shows that their policies consistently outperform commonly used policies that assume accurate records. Thus, their policies are capable of recouping much of the cost of inventory record inaccuracy, without implementation of inventory tracking technologies like RFID. Their replenishment policies avoid the problem of freezing and their audit policy significantly outperforms the popular zero balance walk audit policy. They also present methods for estimating the necessary model parameters and calibrate their model based on real audit data from a leading retailer. Bensoussan et al. have developed a more general cash register model and focus on existence and uniqueness issues.[9]

Important Research Issues

There are many important empirical and theoretical research topics to pursue in the future.

Empirical research. It would be of much interest to know the industries where the problem of incomplete inventory information, or the i3 problem, is serious enough to warrant the difficult mathematical analysis required. Clearly, if the extent of accuracy in inventory information is not bad, then one could still use the classical inventory control approaches without too much sacrifice in profit.

It is now reasonably well documented that the i3 problem is serious in the retail industry. Could the food industry have spoilage problems significant enough to cause inaccurate inventory records? Which other industries are afflicted with significant inventory inaccuracies?

A related issue worth studying is whether or not the presence of inaccurate inventory records is recognized or simply brushed aside. After all, management may be reluctant to admit its presence for the fear of appearing to be negligent or incompetent.

In the case when the presence of the i3 problem is recognized, what signals are observed that may shed light on the inventory's conditional distribution? Keep in mind that even if the underlying inventory problem is the same and requires the same classical modeling, different signals may give rise to different i3 models worth analyzing.

Finally, how are the observed signals related to the inventory level? The answer to this question may come from historical records or in-depth statistical studies of the processes leading to deviations between recorded and actual inventories. Such studies would enable us to specify the way in which the inventory's conditional probability evolves over time.

Theoretical research. Once we have a serious enough i3 problem, we need to develop a model involving the evolution of the conditional probability. The next step is to develop an optimization methodology for such partially observed problems. It has been more or less settled by Bensoussan et al. that the conditional probability is a sufficient statistic and that an optimal feedback solution exists for such problems.[9] It is also clear from the reviewed literature that there are no simple optimal policies for such problems. There are at least two ways to address this issue. One way is to develop efficient computational procedures which would supply optimal solutions or near optimal solutions. In the later case, it would be important to know the extent of near optimality.

Another, perhaps more interesting way is to specify a class of simple implementable policies and optimize within this class. Of course, it would also be important to study the loss in the value function by restricting the class of optimal policies from the largest possible class of non-anticipative policies. These lines of research are largely unexplored at the moment.

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. In some way, this research provides a three-pronged approach to prevent, correct, and integrate as mentioned in the introduction.

Concluding Remarks

In many real-life problems, inventories and demand are partially observed. For such problems, the classical inventory control approaches are no longer adequate and new approaches need to be developed. I've provided a brief review of the extant literature, which shows the difficult nature of partially observed inventory problems. I've also suggested important empirical and theoretical studies required to address the i3 problems.

Endnotes

1. Raman, A., DeHoratius, N., and Ton, Z. 2001. Execution: The missing link in retail operations, *California Management Review, 43*, pp. 136–152.

2. DeHoratius, N., Raman, A. 2008. Inventory record inaccuracy: An empirical analysis. *Management Science, 54*, 627–641.

3. Bensoussan, A., Cakanyıldırım, M., and Sethi, S.P. 2007. Partially observed inventory systems: The case of zero balance walk. *SIAM Journal of Control and Optimization, 46*, 176-209.

4. Bolger, D. 2002. Managing inventory theft, *Furniture World Magazine*, October.

5. Gentry, C. R. 2005. Theft and terrorism. *Chain Store Age, 81*(7), 66–68.

6. DeHoratius, N., Mersereau, A. J., and Raman, A. 2008. Retail inventory management when records are inaccurate. *Manufacturing & Service Operations Management, 10*, 257-277.

7. Bensoussan, A., Cakanyıldırım, M., and Sethi, S.P. 2007. A multiperiod newsvendor problem with partially observed demand. *Mathematics of Operations Research, 32*, 322-344.

8. Bensoussan, A., Cakanyıldırım, M., Feng, Q., and Sethi, S.P. 2009. Optimal ordering policies for stochastic inventory problems with observed information delays.*Production and Operations Management, 18*, 546-559.

9. Bensoussan, A., Cakanyıldırım, M., and Sethi, S.P. 2008. Filtering for discrete-time Markov processes and applications to inventory control with incomplete information. To appear in *Oxford Handbook of Nonlinear Filtering*. Published by Oxford University Press.

10. Bensoussan, A., Cakanyıldırım, M., and Sethi, S.P. 2007. Economic evaluation of systems that expedite inventory information. *Production and Operations Management, 16*, 360-368.

11. Lee, H., O. Ozer. 2007. Unlocking the value of RFID. *Production and Operations Management, 16*(1), 40–64.

12. Bustillo M., 2010. Wal-Mart Radio Tags to Track Clothing. *Wall Street Journal*, page A1, July 23. ■