Computers & Industrial Engineering 63 (2012) 343–353

Contents lists available at SciVerse ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

A simple and robust batch-ordering inventory policy under incomplete demand knowledge ☆

Liwei Bai [a], Christos Alexopoulos [b,*], Mark E. Ferguson [c,1], Kwok-Leung Tsui [b,2]

[a] Tower Research Capital, LLC, 377 Broadway, 11th Floor, New York, NY 10013, United States
[b] School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, United States
[c] Moore School of Business, University of South Carolina, 1705 College Street, Columbia, SC 29208, United States

A R T I C L E   I N F O

Article history:
Received 20 September 2010
Received in revised form 25 October 2011
Accepted 21 February 2012
Available online 16 March 2012

Keywords:
Inventory inaccuracies
Power Approximation method
Batch–ordering inventory policy

A B S T R A C T

Generally, the derivation of an inventory policy requires the knowledge of the underlying demand distribution. Unfortunately, in many settings demand is not completely observable in a direct way or inventory records may be inaccurate. A variety of factors, including the potential inaccuracy of inventory records, motivate managers to seek replenishment policies where the inventory is reviewed periodically and a fixed quantity Q is ordered once the inventory level is found to be under a certain point r. To apply such a policy, however, firms must derive the values r and Q without a clear understanding of the demand distribution. We develop estimators of the first two moments of the (periodic) demand by means of renewal theoretical concepts and a regression-based method, and use these estimators in conjunction with the Power Approximation (PA) method of Ehrhardt and Mosier (1984) to obtain an (r, Q) replenishment policy. The proposed methodology is robust and easy to code into a spreadsheet application. A series of numerical studies are carried out to evaluate the accuracy and precision of the estimators, and to investigate the impact of the estimation on the optimality of the inventory policies. Our experiments indicate that the proposed (r, Q) policy is very close, with regard to the expected total cost per period, to the (s, S) policy obtained via the PA method when the demand process is fully observable and inventory records are accurate.

1. Introduction

While modern-day inventory theory has been under development for more than 50 years, the application of the theoretical results from this effort to real-world applications remains a challenge. In particular, many common theoretically optimal inventory policies fail to deliver their promised results in practice because the assumptions in the underlying models often do not hold. One key underlying assumption that is often violated in practice is that the demand distribution and the associated parameter values are known with certainty by the policy maker. In reality, very few firms know the true underlying demand distributions for their products. Instead, firms assume a specific demand distribution and replace the unknown parameters with estimates computed from historical data. Unfortunately, the data used to estimate these parameters is often limited due to frequent product introductions and changing customer preferences. Even in the rare occasions where the customer demand distribution is known with certainty, inventory control is still problematic due to frequent inaccuracies in the system-reported inventory positions.

Recent studies have shown that system-reported inventory positions are often inaccurate. According to DeHoratius and Raman (2008), records were inaccurate for 65% of the stock-keeping units (SKUs) at a publicly traded retailer. Kang and Gershwin (2005) found that cycle counts at the SKU level across the stores of a global retailer matched the system-reported inventory count in only 51% of the cases. The best performing store in their study only achieved a 70–75% accuracy level. Kök and Shang (2007) provide an example of a heavy equipment manufacturing firm where audits uncovered inventory inaccuracies of 1.6% of total inventory value at their distribution centers. Various problems arise when firms with system-reported inventory inaccuracies continue to manage their stock with policies developed assuming perfect information about inventory positions and demand distributions. Inventory inaccuracies also affect the demand distribution estimation process, as most common inventory replenishment policies require the estimation (based on historical demand as recorded by the information system) of the first and second moments of the

☆ This manuscript was processed by Area Editor Joseph Geunes.
* Corresponding author. Tel.: +1 404 894 2361.
E-mail addresses: danielblw@gmail.com (L. Bai), christos@isye.gatech.edu (C. Alexopoulos), mark.ferguson@moore.sc.edu (M.E. Ferguson), ktsui@isye.gatech.du (K.-L. Tsui).
[1] Tel.: +1 803 777 5923.
[2] Tel.: +1 404 894 2311.

0360-8352/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cie.2012.02.010

demand distribution. Kök and Shang (2007) estimate the cost penalties at the heavy equipment manufacturer for the misuse of such "perfect-information" replenishment policies at 5%.

The motivation for our problem comes from a corporation that managed the inventory of spare parts for one of the US armed forces. The inventory consists of about 60,000 items, ranging from small and inexpensive to large and expensive. Parts are stored in bins which, in turn, are arranged in bin stock locations. Our focus is on the inventory management policies for the small and inexpensive parts (e.g., screws, bolts, washers, etc.). The firm has determined that it is not cost-efficient to take the time to record the use of these items. Thus, the mechanics at the firm do not record the number of parts they retrieve from the bins and the demand process is unknown.

To manage the small but frequently used parts, the firm currently uses the following "two-bin" policy. Within a bin stock location, each part number occupies two bins that can be sized to hold a predefined number of parts. During a periodic review, if the first bin is found to be empty, then an order is placed for a fixed quantity previously negotiated between the firm and its supplier. The replenishment is used to fill the second bin first, and the remainder is used to fill the first bin. Orders are not delivered instantaneously; the lead-times are independent and identically distributed (i.i.d.) from an unknown distribution that depends on the item and its manufacturer or supplier. When an order is placed, the order quantity, the time the order was placed, and the arrival time of the order are recorded in a database. Nevertheless, the exact inventory level cannot be tracked (the parts are usually small with sometimes hundreds in a bin and cycle counting is cost prohibitive) and the direct demand information is unknown. Adding a cycle count via a less expensive inventory count system, such as a weight-based system, is also not cost effective due to the large number of SKUs. Thus, even a weight-based system requires having someone go through and weigh the bins for all SKUs each time a cycle count is conducted. Any unmet demand is backlogged and assessed a penalty by the customer.

If the demands during successive periods are i.i.d. random variables from a known distribution, the lead-times form an i.i.d. random sequence, the exact inventory levels are known at each review period, and the orders involve a setup cost, an (s, S) policy minimizes the long-run expected undiscounted total cost per period. If only the first two moments of the demand distribution are known, an approximately optimal (s, S) policy can be obtained, e.g., by the method of Ehrhardt (1979). Because of the reasons described above however, the actual inventory level is unknown and there is uncertainty about the true demand distribution – even its first two moments. In addition to an inventory policy that acknowledges these realities, the firm also desires a policy that reflects other common constraints of the industry such as a fixed order quantity and a simple robust policy that does not require dynamic (re-)optimization and can be implemented with relative ease in a comprehensive software package. To meet these requirements, we chose an (r, Q) inventory policy. Such a policy could easily work with the firm's existing two-bin system in the following way. The second bin is set to hold r units while the first bin is sized to hold up to Q units. When a periodic check shows the first bin to be empty, a replenishment order is placed for Q units. When the replenishment arrives, the inventory in the second bin is filled to capacity (up to r units) and the remainder of the replenishment shipment is placed in the first bin.

The main objective of this paper is to provide an (r, Q) inventory policy that minimizes the long-run mean (undiscounted) total cost per period subject to the aforementioned constraints. A direct application of a naïve (r, Q) policy is problematic: the experimental results provided in Table 1 show that for some realistic parameter values, such a policy can result in average system costs that are 100+ times higher than the estimated cost incurred by our method.

To calculate the (r, Q) values under incomplete demand information, we first construct estimators for the first two moments of the demand distribution using well-known results from renewal theory and realizations of periods between orders. We proceed with the derivation of approximately optimal inventory policies similar in spirit to the simple and robust Power Approximation (PA) method first described in Ehrhardt (1979) and revised in Ehrhardt and Mosier (1984). The PA method has been used successfully in a variety of settings and owns its popularity to its simplicity and the surprisingly good fit of the regression model to a variety of demand distributions. The proposed (r, Q) policy approximates the optimal (s, S) policy calculated using the PA method. The performance of the (r, Q) policy is evaluated by means of an experimental grid consisting of 216 design points and similar to Ehrhardt's experimental design. Since the cost of the (r, Q) policy is sensitive to the variance of the demand per period, we develop an alternative regression-based estimator of the variance of the periodic demand using the same grid. The incorporation of this estimator results in substantially improved (r, Q) policies with regard to the mean total cost per period.

The main contribution of this paper is the derivation of a near-optimal (r, Q) policy when the parameters of the demand distribution are unknown and/or the inventory stocking levels are also unknown. The first two moments of the demand distribution are estimated using historical data involving order times and order sizes. The proximity of the (r, Q) policy to optimality is evaluated via an extensive experimental study.

The rest of the paper is organized as follows: in Section 2, we position our research in the context of the relevant literature. Section 3 contains the key assumptions and notation, and reviews the PA method. In Section 4, we discuss the estimation of the first two moments of the demand per period and the conversion of (s, S) policies obtained by the PA method to (r, Q) policies. In Section 5, we evaluate the quality of the moment estimators from Section 4 and the performance of the (r, Q) policy by means of an extensive experimental study. In Section 6, we propose a regression-based estimator for the variance of the demand per period, and use the experimental setting from Section 5 to show that this estimator yields an (r, Q) policy that is close to the optimal (s, S) policy with regard to the mean total cost per period. Section 7 contains conclusions and a discussion of managerial implications. The online supplement to this paper contains tables with detailed experimental results.

2. Review of existing inventory inaccuracy compensation methods

In a world with perfect information, no system-reported inaccuracies, and no restrictions on the periodic order quantity, an (s, S) policy minimizes the long-run expected total cost per period. Scarf (1960) and Iglehart (1963) prove the optimality of (s, S) policies, and Veinott and Wagner (1965) consider methods for computing them when demands per period are i.i.d. and normally distributed random variables. Even under these ideal conditions, (s, S) policies are difficult-to-compute. Thus, several approximations including Ehrhardt's (1979) PA method were developed. Porteus (1985) offers a review of some of the most common approximations. *Our scenario assumes no prior knowledge of the demand distribution type or its moments, and that the total demand in a period is not directly observable*.

When the demand is observable but its distribution is unknown, a commonly used approach fits a statistical model to the observed demand data and uses the fitted model to obtain the desired policy. Typically, limited historical data is available to estimate the parameters of the demand distribution. Jacobs and Wagner (1989) investigate how the choices of statistical estimators