Computers & Operations Research 49 (2014) 18–27

Contents lists available at ScienceDirect

Computers & Operations Research

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

Calculating target inventory levels for constrained production: A fast simulation-based approximation

John M. Betts *

Faculty of Information Technology, Monash University, Clayton, Victoria 3800, Australia

ARTICLE INFO

Available online 26 March 2014

Keywords:
Production
Inventory
Simulation
Optimisation
Approximation

ABSTRACT

An accurate model for the inventory shortfall distribution is necessary to calculate the target level required to give a desired service level in manufacturing systems under stochastic demand and production capacity constraint. Existing methods for modelling the inventory shortfall require that the demand distribution be expressible in functional form and that the coefficient of demand variation be small. When these conditions cannot be met, the only recourse to a practitioner is to set target levels using simulation-based optimisation, which is computationally intensive and time consuming. By contrast, this paper presents a model in which the inventory shortfall is approximated by sampling from a single simulation run of the inventory process. The target level required for a given service level can then be calculated efficiently, to a high degree of reliability, using an iterative search. This new model is thus an efficient alternative to conventional simulation-based optimisation. Because the shortfall distribution is generated by simulating demand directly, the model makes no assumptions about the form of the demand distribution. By requiring no user modelling of the functional form or parameters of the demand distribution, this new method is easily used by inventory managers in practice or implemented in software.

1. Introduction

To provide the required level of customer service when demand is stochastic and production capacity constrained, inventory managers using a base stock policy will set a target level. Key to setting the target inventory level is an accurate model of the inventory shortfall at the end of each production-supply cycle. This shortfall distribution can be modelled in its exact form as a Markov process when the state space is small [1] however this method is intractable when the state space of the inventory shortfall is very large, such as occurs, for example, under highly constrained production. When the state space is large, the shortfall distribution may be approximated by a continuous distribution, however this method is inaccurate when the variability of customer demand is large [2]. Another disadvantage of this method is that parameter estimation requires that the demand distribution to be expressible in a functional form. Thus this method cannot be easily applied when the demand distribution is ambiguous or highly variable, which are situations that practitioners encounter frequently. For these cases, the normal recourse for a practitioner is to use simulation-based optimisation or heuristic techniques to set target levels, either approach being time consuming and inefficient.

To address these limitations, this paper presents a simulation-based model in which the inventory shortfall distribution is approximated directly from the customer demand distribution. To achieve this, a single instance of a discrete-event simulation of the unbounded inventory process operating under the production constraint is run, recording the inventory shortfall. Previous research by others has shown that the inventory shortfall can be approximated by a mass exponential distribution [2,3]. By estimating the parameters of this probability distribution from the simulated shortfall, the service level achieved by any given target level can be computed directly. Using an efficient search technique it is then possible to determine the target inventory level required to produce a given service level to a high degree of accuracy. Thus, the new model allows managers to set target levels efficiently for situations where existing methods cannot be used. Previous research by the author [4] demonstrated the feasibility of this approach under restricted demand and service assumptions. The current research extends this work by examining the performance of the model across a wide variety of demand distributions and more general service assumptions. Extensive

* Tel.: +61 3 99055804.
E-mail address: john.betts@monash.edu

http://dx.doi.org/10.1016/j.cor.2014.03.014
0305-0548/

computational testing shows that the new, simulation-based, model out-performs the most accurate analytical approaches in the research literature.

Computational results show that the simulation-based model is more accurate than known existing approximations, calculating target levels to a high degree of accuracy across a wide variety of demand distributions of varying skewness, kurtosis and volatility. These experiments also illustrate the applicability of the model to realistic, practical, situations where high levels of demand variability and/or production constraint prohibit the use of existing methods. By simulating the shortfall distribution directly from historical demand, the model eliminates the need to estimate the parameters and functional form of the demand distribution. Because no user modelling of demand data is required the model is simpler for practitioner use than existing approaches, and could be used by managers having very little mathematical expertise. The model is also sufficiently reliable and robust to be incorporated into a computer-based inventory management or decision support system. Computational results show speed and accuracy behave predictably as a function of the number of iterations and sampling rate of the simulation run enabling the performance of the model to be user-specified. Finally, the model is computationally efficient compared with other simulation-based techniques for setting inventory target levels by requiring that only a single simulation run be conducted.

The following section provides a background to the current research, outlining existing approaches to modelling the inventory shortfall distribution, including the theory on which the simulation-based model is based. In Section 3, the new model for setting target levels, using a single simulation run of the unbounded inventory process under constraint, is presented. Section 4 presents computational experiments where the performance of the simulation-based model is tested in three ways: firstly by comparison with the most accurate model from the existing literature requiring parameter estimation from the data; secondly the accuracy of the new model is analysed across a wide variety of demand distributions, and finally the trade-off between accuracy and computational performance is analysed. In Section 5, an adaptation of the simulation-based model is briefly discussed for inventory cost minimisation. Section 6 summarises and concludes the paper.

2. Relevant background

The single-item, single-stage, inventory system studied in this paper assumes that production and supply follow the following protocol. At the start of each period, random demand, D_t, is observed. Production, limited to a maximum of C units per period, then takes place, based on the quantity demanded and the number of items currently in inventory. If the maximum production quantity plus the amount currently held in inventory is greater than the demand during the period, then all customer orders are filled, otherwise the unmet demand is back ordered. This model forms the basis of many inventory control systems and has been well-studied, for example, [1,2,5].

The optimal method for managing the single-stage inventory system described here is to use an order-up-to or base stock policy [6,7]. This approach requires that the amount of stock ideally held at the end of each production/distribution cycle, the target level, s, be identified. Production occurs during the period it is demanded, as in the case studied here, or after a finite lead time, and supplied. Because demand is random, shortages will occur in periods where production is not sufficient to cover demand. The shortfall each period, v_t, is determined as $v_t = s - I_t$, where I_t represents the inventory level at the end of each period. v_t is determined by the recursive relationship $v_t = \max(v_{t-1} + D_t - C, 0)$. The condition that $E[D_t] < C$ ensures that an infinite backlog of unmet demand does not build up. Under this condition, the inventory process is ergodic, and hence a steady state distribution exists for v_t, denoted by V [5]. Correctly modelling the shortfall distribution is fundamental in accurately calculating the target level necessary to achieve a required level of service. Two methods for determining V are now discussed.

In the case of discrete demand, the inventory shortfall process can be modelled as a Markov chain, from which the steady state distribution of V can derived as the limiting distribution of the transition matrix [5,8]. Although this approach yields the exact distribution of V, when the state space is large evaluating the transition probabilities becomes computationally intractable. When $E[V]$ is small, the transition matrix can be truncated to assist computation without significant loss of accuracy. in situations of high demand variability or resource utilisation, $E[V]$ will necessarily be large, and the actual shortfalls will vary over a wide range due to causes such as bursts of sustained high demand. For these systems, there is very limited potential to truncate the transition matrix without losing significant accuracy.

One way of addressing the problem of dimension arising in the discrete case when $E[V]$ is large is to assume that the demand distribution is continuous. Under this condition, V is also continuous and can be approximated by a mass exponential distribution having the form $P(V=0) = P_0$, where $0 \le P_0 \le 1$, and the remainder of the distribution being exponentially distributed with mean γ^{-1} [5]. Using $\overline{F}_V$ to denote the complementary cumulative distribution function of V,

$$\overline{F}_V(v) = \begin{cases} \overline{P}_0 e^{-\gamma v}, & v > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Glasserman [3], extending the work of Tayur [9], Prabhu [10] and others, proved that F_V provides a good approximation to V for a variety of demand distributions, and that the degree of approximation improves at critical percentiles approaching 1. When D_t is exponential, V is exact. Glasserman also gives analytical expressions for γ under a variety of demand distributions.

To improve the accuracy of the mass exponential approximation of the shortfall distribution, particularly at low service level percentiles, Roundy and Muckstadt [2] and Muckstadt [5] derive the heuristic approximation for the shortfall distribution as

$$\overline{F}^1_V(v) = \overline{F}_D(v+C) + \overline{P}_1 e^{-\gamma(v+C)} \int_{-\infty}^{v+C} e^{\gamma x} f_D(x)dx, \tag{2}$$

where the improved estimate of P_0 is

$$\overline{P}_1 = \frac{\int_{-\infty}^{C} \overline{F}_D(2C-x) f_D(x)dx}{e^{-\gamma C}\int_{-\infty}^{C} e^{\gamma x} f_D(x)dx - \int_{-\infty}^{C} e^{-\gamma(2C-x)}\left[\int_{-\infty}^{2C-x} e^{\gamma y} f_D(y)dy\right] f_D(x)dx}. \tag{3}$$

Subsequent researchers have proposed alternative approaches for modelling the single item capacitated inventory problem. For example, [11] propose a matrix-analytic technique for modelling the inventory level distribution, which is then approximated by a discrete phase-type distribution; [12] model the relationship between excess capacity, demand variability and service level to determine the optimal safety stock placement in a capacitated supply chain; and [13] show that the computation of an optimal base stock policy is NP-hard and propose a polynomial approximation scheme. For the analogous inventory system under continuous review the (s, S) policy is optimal [14]. Although an efficient solution method exists for the unconstrained system, solution under capacity constraint is computationally expensive [15,16], necessitating the use of heuristic methods, for example