Int. J. Production Economics 118 (2009) 34–42

Contents lists available at ScienceDirect

Int. J. Production Economics

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

Assessing the effects of using demand parameters estimates in inventory control and improving the performance using a correction function

Elleke Janssen *, Leo Strijbosch, Ruud Brekelmans

Department of Econometrics and Operations Research, Tilburg University, PO Box 90153, 5000LE Tilburg, The Netherlands

ARTICLE INFO

Available online 22 August 2008

Keywords:
Unknown demand parameters
Inventory control
Normal distribution
Service level criterion

ABSTRACT

Inventory models need some specification of the distribution of demand in order to find the optimal order-up-to level or reorder point. This distribution is unknown in real life and there are several solutions to overcome this problem. One approach is to assume a distribution, estimate its parameters and replace the unknown demand parameters by these estimates in the theoretically correct model. Earlier research suggests that this approach will lead to underperformance, even if the true demand distribution is indeed the assumed one. This paper directs the cause of the underperformance and quantifies it in case of normally distributed demand. Furthermore the formulae for the order-up-to levels are corrected analytically where possible and otherwise by use of simulation and linear regression. Simulation shows that these corrections improve the attained performance.

1. Introduction

Inventory control involves decisions on what to order, when, and how much. The models dealing with these decisions need information about the distribution of demand during some period, e.g. the demand during lead time or during the review period. Bulinskaya (1990) discriminates between three situations: (a) the type of distribution is known, but its parameters are unspecified; (b) only several first moments of the demand distribution are known; and (c) there is no prior knowledge about the demand. The third situation is of course the most realistic one, and different approaches to deal with situation (c) have been proposed in the literature. These approaches can be categorized into parametric and nonparametric methods. Examples of the first category are assuming a distribution and using Bayesian models; examples of the second category are using order statistics, the bootstrap procedure, and kernel densities.

One of the most widespread approaches to deal with unknown demand is assuming a distribution, estimating its parameters, and replacing the unknown parameters by their estimates in the theoretically correct formulae in which distribution and parameters are supposed to be known. Sani and Kingsman (1997), Artto and Pylkkänen (1999), Strijbosch et al. (2000) and Syntetos and Boylan (2006) use this approach with different inventory models, while Kottas and Lau (1980) provide a short discussion on estimating the parameters needed for their model. Another parametric approach is the Bayesian approach; Azoury and Miller (1984), Azoury (1985) and Karmarkar (1994) are three examples of the Bayesian approach. Also Larson et al. (2001) use this approach, but they introduce a nonparametric form. Other nonparametric approaches involve order statistics, references include Lordahl and Bookbinder (1994) and Liyanage and Shanthikumar (2005), the bootstrap procedure, see e.g. Bookbinder and Lordahl (1989) and Fricker and Goodhart (2000), or using kernel densities, see Strijbosch and Heuts (1992).

* Corresponding author. Tel.: +31 13 466 3244; fax: +31 13 466 3280.
E-mail address: E.Janssen_1@uvt.nl (E. Janssen).

0925-5273/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpe.2008.08.029

Consider the approach of assuming a distribution, and let the assumed distribution be the normal one, since it is often used both in research and in practice; see Zeng and Hayya (1999). Furthermore, Strijbosch and Moors (2006) give references to recent articles that involve the normal distribution. The assumption of normality is made because it yields tractable results and it seems to give quite good approximations if used on data with a low coefficient of variation (see Silver et al., 1998; Zipkin, 2000). However, the normal distribution has two major disadvantages: it can take on negative values and is symmetric, while actual demand is nonnegative and often skewed to the right. This may not impose serious problems if the coefficient of variation is low (Zipkin, 2000) or if the demand during review/lead time consists of many individual and independent demands (Silver et al., 1998). For high values of the coefficient of variation, however, these disadvantages get more important; hence Strijbosch and Moors (2006) suggested two simple modifications of the normal distribution. Tyworth and O'Neill (1997) and Lau and Lau (2003) investigate the (non)robustness of using the normal approximation with an (s, Q) inventory policy.

This paper will not deal with situation (c), but with the less realistic situation (a): the true distribution is known, but its parameters are unspecified, so that the effect of estimating them can be studied. Silver and Rahnama (1986, 1987) have investigated this effect in an (s, Q) inventory policy with a cost criterion. They constructed a function that determined the expected cost of estimating the demand distribution rather than knowing it, and they concluded that this function is not symmetrical: underestimating causes larger costs than overestimating. In the second article they propose a method that deliberately biases the reorder point upwards. Strijbosch et al. (1997) and Strijbosch and Moors (2005) have investigated the same effect for an (R, S) inventory policy with a service level criterion. Both papers concluded that also in this case the order-up-to level needed to be biased upwards. Note that the (s, Q) and (R, S) policies are equivalent (see Silver et al., 1998), so the results of Silver and Rahnama (1986, 1987) apply to the (R, S) policy and the results of Strijbosch et al. (1997) and Strijbosch and Moors (2005) apply to the (s, Q) policy as well.

This paper will focus on two service level criteria within an (R, S) inventory policy with zero lead time, so every R time units the inventory is replenished up to the order-up-to level S and the order is delivered instantaneously. The independent and identically distributed demands during review periods (denoted by X_i) have a normal distribution with mean μ and standard deviation σ, which leads to a coefficient of variation $v = \sigma/\mu$. As mentioned before, the normal distribution could lead to negative demand and in our model this is interpreted as returned goods, so demand is actually net demand (demanded goods minus returned goods). In addition, the goods returned by customers can be sent back to the supplier, thus the inventory level at the start of a review period will always equal S. Furthermore, demand during $t+1$ consecutive review periods is assumed to be stationary, and the first t periodic review demands are used to estimate the mean and standard deviation of the demand in review period $t+1$. The mean and standard deviations are estimated by their sample statistics. We prefer using sample statistics instead of exponential smoothing, since derivations are more tractable, while the conclusions will be similar. In forecasting this method is often referred to as the (simple) moving average.

Section 2 discusses the P_1 service criterion in short. An analytical correction of the order-up-to level is given for the case that only μ is unknown. Section 3 focusses on the P_2 service criterion. First two theoretical situations are considered for illustrative purposes: the cases that μ (μ and σ) is (are) unknown, but σ and v (v) are (is) known. The main Sections 3.3 and 3.4 treat the important case that these three parameters are all unknown. We show by simulation that just plugging in estimates leads to serious underperformance. Besides, we derive a correction function for the safety factor that nearly gives the desired fill rates. The last section concludes this research, and provides directions for further research.

2. P_1 service criterion

This section considers the P_1 service criterion. This criterion states that the fraction of periods in which inventory is sufficient to satisfy demand is at least α. It is common to express the order-up-to level as a function of the mean μ, the standard deviation σ, and a safety factor c_α. Since demand is normally distributed, the order-up-to level is (see e.g. Silver et al., 1998)

$$S(\mu, \sigma, c_\alpha) = \mu + c_\alpha \sigma. \tag{1}$$

The safety factor c_α is $\Phi^{-1}(\alpha)$, where $\Phi(\cdot)$ is the cdf of the standard normal distribution. S without arguments is used to denote the theoretically correct order-up-to level when all parameters are known, so $S = S(\mu, \sigma, c_\alpha)$ in case of a P_1 service criterion.

In practice, the mean and variance are unknown, which means that S is unknown too. The common solution is to replace the parameters μ and σ by its estimates. If we (unrealistically) assume that only μ is unknown and use the sample mean $\hat{\mu}_t$ to estimate it, the resulting order-up-to level is $S(\hat{\mu}_t, \sigma, c_\alpha)$ with $\hat{\mu}_t = (1/t)\sum_{i=1}^{t} X_i$. This order-up-to level, although unbiased, will not meet the service requirements in the long run. Consider Fig. 1. Since $S(\hat{\mu}_t, \sigma, c_\alpha)$ is normally distributed with mean S and variance σ^2/t, it is symmetric and a shift from S to the right is equally probable as a shift of the same size to the left. The shift to the right decreases the probability of having backorders with the darker area in Fig. 1, while the shift of the same size to the left will increase the probability of having backorders with the lighter area. The surface of the lighter area is larger than the surface of the darker area and this implies that in the long run, the achieved service level will fall short of the desired one. This phenomenon is mathematically explained by

$$\begin{aligned} P(X_{t+1} < S(\hat{\mu}_t, \sigma, c_\alpha)) &= P\left(X_{t+1} - \frac{1}{t}\sum_{i=1}^{t} X_i < c_\alpha \sigma\right) \\ &= \Phi\left(\frac{c_\alpha}{\sqrt{1+1/t}}\right) < \Phi(c_\alpha) = \alpha, \end{aligned} \tag{2}$$