Expert Systems with Applications 38 (2011) 7929–7934

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A linear programming formulation for an inventory management decision problem with a service constraint

Gerrit K. Janssens *, Katrien M. Ramaekers

Hasselt University, Campus Diepenbeek, Wetenschapspark 5 Bus 6, B-3590 Diepenbeek, Belgium

ARTICLE INFO

Keywords:
Inventory management
Linear programming
Incomplete information

ABSTRACT

Inventory systems with uncertainty go hand in hand with the determination of a safety stock level. The decision on the safety stock level is based on a performance measure, for example the expected shortage per replenishment period or the probability of a stock-out per replenishment period. The performance measure assumes complete knowledge of the probability distribution during lead time, which might not be available. In case of incomplete information regarding the lead-time distribution of demand, no single figure for the safety stock can de determined in order to satisfy a performance measure. However, an optimisation model may be formulated in order to determine a safety stock level which guarantees the performance measure under the worst case of lead-time demand, of which the distribution is known in an incomplete way. It is shown that this optimisation problem can be formulated as a linear programming problem.

1. Introduction

Some uncertainty in an inventory system (such as lead time, quantity and quality) depends on the suppliers. If the suppliers introduce too much uncertainty, corrective action should be taken. Some uncertainty, however, is attributable to customers, especially demand. If insufficient inventory is hold, a stock-out may occur leading to shortage costs. Shortage costs are usually high in relation to holding costs. Companies are willing to hold additional inventory, above their forecasted needs, to add a margin of safety.

Determination of an inventory replenishment policy, of the quantities to order, of the review period are typical decisions to be taken by logistics managers. Decisions are made through optimisation models taking a performance measure into consideration which might be cost-oriented or service-oriented. Performance measures of the service-oriented type may be expressed relatively as a probability of a stock-out during a certain replenishment period, or may be expressed absolutely in terms of number of units short, which is a direct indication for lost sales. Both performance measures are taken into consideration and special attention will be paid to feasible combinations of company's objectives regarding both performance measures.

For a definition of both measures we refer to Chapter 7 in Silver, Pyke, and Peterson (1998)

The *expected shortage per replenishment cycle* (ESPRC) is defined as (with t the amount of safety stock):

$$\text{ESPRC} = \int_t^{+\infty} (x - t) f(x) dx \tag{1}$$

If ordered per quantity Q the fraction backordered is equal to ESPRC/Q and a performance measure, indicated as $P2$, is defined as

$$P_2 = 1 - \text{ESPRC}/Q \tag{2}$$

The other performance measure is the *probability of a stock-out in a replenishment lead time*, defined as:

$$1 - P_1 = \Pr\{x \geqslant t\} = \int_t^{+\infty} f(x) dx \tag{3}$$

From a production or trading company's point of view, a decision might be formulated to answer the following question: *given a maximum expected number of units short and/or a maximum stock-out probability the company wants to face, what should be the safety inventory at least* (*or at most*)? The question with the 'at most' option might be only of academic nature, as it reflects the most optimistic viewpoint. In human terms, this question would be interpreted as: 'would there exist any probability distribution so that I can still reach my preset performance criteria given a specific safety inventory?'. This type of question is not relevant for a manager facing a real-life situation.

In case the distribution of demand is known, determining the inventory level, given a maximum shortage or maximum

* Corresponding author. Address: Transportation Research Institute, Hasselt University, Wetenschapspark 5 Bus 6, B-3590 Diepenbeek, Belgium. Tel.: +32 11 269119; fax: +32 11 269198.

E-mail addresses: gerrit.janssens@uhasselt.be (G.K. Janssens), katrien.ramaekers@uhasselt.be (K.M. Ramaekers).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.12.009

stock-out probability, reduces to the calculation of the inverse cumulative probability function. The decision problem becomes more difficult if incomplete information exists on the distribution of demand during lead time, for example only the range of demand, or the first moment, or the first and second moments are known. In such a case no single value can be determined but rather an interval.

In classical textbooks not too much attention is paid to the shape of the distribution of the demand during lead time. Mostly, based on the first and second moments, the safety stock level is determined using the normal distribution. When of relevance, one rather should look for a distribution, which is defined only for non-negative values and allows for some skewness. In the literature on inventory control, frequent reference is made to the Gamma distribution.

It is generally known that, given a shape of the demand distribution, the higher the coefficient of variation the more a company needs inventory to reach a given service level. In an investigation on the relevance of the demand shape Bartezzaghi, Verganti, and Zotteri (1999) find out that the shape is very relevant. In extreme cases the impact of different demand shapes on inventories is comparable to the effect of doubling the coefficient of variation. An extensive overview of the distribution types used to model period demand, lead time and/or demand during lead time is given in Vernimmen, Dullaert, Willemé, and Witlox (2008).

This research deals with the case where the demand distribution during lead time is not completely known. This situation is realistic either with products which have been introduced recently to the market or with slow moving products. In both cases not sufficient data are available to decide on the functional form of the demand distribution function. Some but not complete information might exist like the range of the demand, its expected value, its variance and maybe some knowledge about uni-modality of the distribution.

In case incomplete information is available regarding the demand distribution the integrals of the performance measures P_1 and P_2 cannot be evaluated in an analytical manner. This means that also the inverse problem of determining the safety stock level to satisfy the performance measures cannot be obtained analytically. However, the integrals can be approximated by a linear programming formulation with a large set of constraints.

2. Bounds on the performance measures in the case of incomplete information

In this section the ESPRC measure is focused. First, a link is identified with a similar integral formulation which appears in the field of actuarial sciences. Second, some results, which were obtained in actuarial sciences, are transferred to our type of application.

2.1. Towards an analogy in insurance mathematics

In insurance mathematics, an insurance company using the option of re-insurance is confronted with a stop-loss premium. A stop-loss premium limits the risk X of an insurance company to a certain amount t. If the claim size is higher than t the re-insurance company takes over the risk $X-t$. The stop-loss premium is based on the expected value of $X-t$, which in case of a known claim size distribution may be defined as:

$$\int_0^\infty (x-t)_+ dF(x) \tag{4}$$

where $F(x)$ represents the claim size distribution (Goovaerts, De Vylder, & Haezendonck, 1984).

The same formula (4) may be useful in the performance evaluation of inventory management in case of uncertain demand during lead time. When a company holds t units of a specific product in inventory starting a period between order and delivery, any demand less than t is satisfied while any demand X greater than t results in a shortage of $X-t$ units. A lesser number of units short results in a better service to the customer. In this way formula (4) is a measure for customer service in inventory management.

In the following sections lower and upper bounds are obtained for the performance measure under study, given various levels of information about the demand distribution. From a production or trading company's point of view, a decision might be formulated to answer the following question: *given an expected number of units short the company wants to face, what should be the safety inventory at least or at most?*

2.2. The case of known range, mean and variance

Let the size of the demand $X$for a specific product in a finite period have a distribution F with first two moments $\mu_1 = E(X)$ and $\mu_2 = E(X^2)$.

From a mathematical point of view, the problem is to find the following bounds:

$$\sup_{F\in\phi} \int_0^\infty (x-t)_+ dF(x) \tag{5a}$$

and

$$\inf_{F\in\phi} \int_0^\infty (x-t)_+ dF(x) \tag{5b}$$

where $\varnothing$ is the class of all distribution functions F which have moments μ_1 and μ_2, and which have support in R^+. Let further $\sigma^2 = \mu_1 - \mu_2^2$. We assume t to be strictly positive.

For any polynomial $P(x)$ of degree 2 or less, the integral

$$\int_0^\infty P(x)dF(x)$$

only depends on μ_1 and μ_2, so it takes the same value for all distributions in $\varnothing$. There exists some distribution G in $\varnothing$ for which the equality holds:

$$\int_0^\infty P(x)dG(x) = \int_0^\infty (x-t)_+ dG(x). \tag{6}$$

As distribution G a two-point or three-point distribution is used. The equality (6) is attained when $P(x)$ and $(x-t)_+$ are equal in both points of G. The best upper and lower bounds on this term with given moments μ_1 and μ_2 are derived. The method is inspired by papers of Janssen, Haezendonck, and Goovaerts (1986) and by Heijnen and Goovaerts (1989). In the following we assume the known range of the distribution to be a finite interval $[a,b]$.

A probability distribution F is called n-atomic if all its probability mass is concentrated in n points at most. The points are called the atoms of the distributions. The problem (5a) has a 2-atomic solution and (5b) has a 3-atomic solution.

If α, β are two different atoms of the 2-atomic probability distribution F satisfying the first-order moment constraint $\int x dF = \mu_1$, then the corresponding probability masses p_α and p_β are

$$p_\alpha = \frac{\mu_1 - \beta}{\alpha - \beta}, \quad p_\beta = \frac{\mu_1 - \alpha}{\beta - \alpha} \tag{7}$$

If α, β, γ are three different atoms of the 3-atomic probability distribution F satisfying the moment constraints $\int x dF = \mu_1$, $\int x^2 dF = \mu_2$, then the corresponding probability masses p_α, p_β and p_γ are