The Exact Fill Rate in a Periodic Review Base Stock System under Normally Distributed Demand

Edward A. Silver,[a] Diane P. Bischak[a]

[a]*Haskayne School of Business*
University of Calgary
2500 University Drive NW
Calgary, Alberta, Canada T2N 1N4

Published in *Omega: the International Journal of Management Science* 39 (2011), 346-349.

Correspondence to: Diane Bischak (diane.bischak@haskayne.ucalgary.ca)

Abstract: In this paper we consider a periodic review order-up-to-level (or base stock) inventory control system under normally distributed demand. For such circumstances an expression for the exact fill rate (fraction of demand satisfied without backordering) has been available in the literature but has not been widely known, let alone used by practitioners. In this paper we redevelop the expression and contrast our derivation with the earlier published one. The paper has two purposes. First, we hope that the reappearance of the exact result in this journal will lead to its wider adoption. Second, showing two contrasting approaches to obtaining the same result may be useful for both research and pedagogical purposes.

1. Introduction

Periodic review, base stock systems are widely used in practice. In these systems, every R units of time enough stock is ordered to raise the inventory position (on-hand minus backorders plus on-order) to an order-up-to-level (or base stock level) S. In addition, such systems are often employed under a constraint of satisfying a desired fill rate, the fraction of demand met directly from stock (see, for example, Lee and Billington [1]).

Under assumptions of normally distributed demand and a constant lead time, in this paper we develop an exact expression for the fill rate. This is not a new result; what is new is the approach to obtaining it. As the result (and the earlier approach to finding it) is not widely known, most practitioners and academics (the authors included) have been using approximate expressions for the fill rate. Historically there were valid reasons—in particular, computational simplicity—for using such approximations, but under certain conditions the approximations can be appreciably in error. The associated model is still relatively simple to understand and use, which is an important consideration (Ward et al. [2]).

The first objective of the current paper is to achieve broader awareness and adoption of the correct fill rate expression. Second, it is hoped that insights (particularly for research or pedagogical purposes) may be achieved by showing two different methods of obtaining the same result. An early paper by one of the authors (Silver [3]) served this type of purpose.

The authors have been able to find the exact result (or how to obtain it) in only a handful of publications. The original derivation was probably by Hadley and Whitin [4]. Subsequently, de Kok [5] indicated the general approach (but without the result). However, in unpublished material (de Kok [6]) he did derive the end result, which is also shown in a footnote on page 280 of Silver *et al.* [7]. Axsäter [8] indicated how the fill rate can be obtained from limiting conditions of a different type of control system (which is actually part of the derivation used by Hadley and Whitin). The exact result is also used in Chen and Zheng [9]. Sobel [10] has developed a somewhat more complicated result, but the added complexity is due to his not permitting negative values of the normal distribution.

Johnson *et al*. [11] mentioned that the Hadley and Whitin result had been largely overlooked in the literature. They showed it and pointed out that for high enough demand variability (coefficients of variation greater than unity) it could be erroneous because of not properly taking account of the substantial likelihood of negative demands (*i.e.* returns). They went on to develop a considerably more involved expression that handles high coefficients of variation. In the current paper we restrict attention to situations where the chance of a negative demand during a review period is reasonably small. In particular, the results presented in Johnson *et al*. indicate that use of our formula would produce negligible errors (under 0.4%) in the estimated (versus desired) fill rate for coefficients of variation no larger than 0.5 (which for the normal implies a probability of just over 2% of a negative demand in R). Even if the coefficient of variation is as high as 1.0 (which implies a more than 15% chance of a negative demand in R), the errors in the estimated fill rate would still be under 2%.

In the next section we present the new derivation. This is followed by an outline of the original method and two numerical illustrations comparing the results obtained with the exact approach to the results produced with a commonly used approximation. Finally, some brief comments are provided in a concluding section.

2. Derivations

2.1 Our approach

Let R and S denote the review interval (in time units) and the order-up-to-level, respectively. Also let Q represent the order quantity at a particular review. Our approach involves developing EUS (Q_0), the expected units short conditional on a given value of Q, denoted as Q_0. (The approach to this stage parallels that for a continuous-review, order-quantity system in Silver [12].) This is then multiplied by the density function of Q and integrated over all possible values of Q_0 to obtain the unconditional expected units short, EUS. The fill rate is then expressed, as usual, in terms of EUS and other factors. Now we present the details.

In an (R, S) system with a constant lead time L, the magnitude of any shortage in a particular cycle depends on the relationship between S and the total demand during $R+L$. (Note that we do not require any restriction on the relative sizes of R and L.) Consider a cycle that begins at time 0 (see Figure 1). For a given order size Q_0 placed at the end of R, the units short at the end of the subsequent L (*i.e.* just before Q_0 arrives), in turn, depend on the relative sizes of $S - Q_0$ and the total demand in L. We will denote the latter by y and the units short by $US(Q_0, y_0)$. Then

$$US(Q_0, y_0) = \begin{cases} 0 & y_0 \leq S - Q_0 \\ y_0 - (S - Q_0) & S - Q_0 \leq y_0 \leq S \\ Q_0 & S \leq y_0 \end{cases}$$

For y_0 strictly greater than S a shortage larger than Q_0 actually occurs but the portion above Q_0 will not be eliminated by the arrival of the order of size Q_0, hence will be counted in the next cycle. An alternative, intuitive interpretation is that for a batch of Q units, at most Q can arrive late.

[Insert Figure 1 about here]

Taking the expectation with respect to the variable y,

$$EUS(Q_0) = \int_{S-Q_0}^{S} [y_0 - (S - Q_0)]\, f_y(y_0)dy_0 + Q_0 \int_{S}^{\infty} f_y(y_0)dy_0$$

which can be manipulated to obtain

$$EUS(Q_0) = \int_{S-Q_0}^{\infty} [y_0 - (S - Q_0)]\, f_y(y_0)dy_0 - \int_{S}^{\infty} (y_0 - S)\, f_y(y_0)dy_0 \,. \tag{1}$$

With normally distributed demand

$$f_y(y_0) = \frac{1}{\sqrt{2\pi}\sigma_L} \exp\left[\frac{-(y_0 - \mu_L)^2}{2\sigma_L{}^2}\right] \tag{2}$$

where μ_L and σ_L are the mean and standard deviation of the demand in L.

Substituting (2) into (1) and employing the transformation

$$z_0 = \frac{y_0 - \mu_L}{\sigma_L}$$

so that z_0 is a unit normal variable, leads to

$$EUS(Q_0) = \sigma_L \left[G\left(\frac{S - Q_0 - \mu_L}{\sigma_L}\right) - G\left(\frac{S - \mu_L}{\sigma_L}\right)\right] \tag{3}$$

where $G(w) = \int_{w}^{\infty} (z_0 - w)\, f_Z(z_0)dz_0$ is the so-called unit normal loss function.

Next

$$EUS = \int_{Q_0} EUS(Q_0) f_Q(Q_0) dQ_0 \tag{4}$$

Because of the nature of an (R, S) system, Q_0 must represent the total demand during the preceding R. Hence,

$$f_Q(Q_0) = \frac{1}{\sqrt{2\pi}\sigma_R} \exp\left[-\frac{(Q_0 - \mu_R)^2}{2\sigma_R^2} \right] \qquad -\infty < Q_0 < \infty \tag{5}$$

where μ_R and σ_R are the mean and standard deviation of the demand in R. Substituting (3) and (5) into (4) and using

$$u_0 = \frac{Q_0 - \mu_R}{\sigma_R}$$

so that u_0 is again a unit normal variable, results in

$$EUS = \sigma_L \left\{ \int_{-\infty}^{\infty} G\left[\frac{S - (\mu_R + \mu_L) - \sigma_R u_0}{\sigma_L} \right] f_u(u_0) du_0 - G\left(\frac{S - \mu_L}{\sigma_L} \right) \right\} \tag{6}$$

From Silver and Smith [13],

$$\int_{-\infty}^{\infty} G(au_0 + b) f_u(u_0) du_0 = \sqrt{1 + a^2}\, G\left(\frac{b}{\sqrt{1 + a^2}} \right).$$

In the integral of (6) we have

$$a = -\frac{\sigma_R}{\sigma_L} \qquad \text{and} \qquad b = \frac{S - (\mu_R + \mu_L)}{\sigma_L}$$

Thus

$$EUS = \sigma_L \left\{ \sqrt{1+\frac{\sigma_R^2}{\sigma_L^2}}\, G\left[\frac{S-(\mu_R+\mu_L)}{\sigma_L\sqrt{1+\sigma_R^2/\sigma_L^2}}\right] - G\left(\frac{S-\mu_L}{\sigma_L}\right)\right\} \tag{7}$$

Denote the mean and standard deviation of demand per unit time by μ and σ. Moreover, note that $\mu_L = L\mu$, $\mu_R = R\mu$, $\sigma_L = \sqrt{L}\sigma$ and $\sigma_R = \sqrt{R}\sigma$. Also let $CV = \sigma/\mu$ represent the coefficient of variation of the demand in a unit time period. Finally, in the usual fashion let

$$S = \mu_{R+L} + k\sigma_{R+L} = (R+L)\mu + k\sqrt{R+L}\,\sigma \tag{8}$$

where k is the safety factor. Then (7) simplifies to

$$EUS = \sqrt{R+L}\,\sigma G(k) - \sqrt{L}\,\sigma G\left(\frac{R}{\sqrt{L}\,CV} + k\sqrt{\frac{R+L}{L}}\right) \tag{9}$$

Finally, the (R, S) system defines a renewal process where each cycle (every R units of time) starts with the inventory position at S. Thus, the fill rate, denoted by P, is given by the expected demand met from stock in a cycle divided by the expected demand per cycle (Ross [14]), *i.e.*

$$P = \frac{E(Q) - EUS}{E(Q)} = 1 - \frac{EUS}{R\mu} \tag{10}$$

Substituting (9) in (10) leads to

$$P = 1 - \left[\frac{\sqrt{R+L}}{R} CV\, G(k) - \frac{\sqrt{L}}{R} CV\, G\left(\frac{R}{\sqrt{L}\,CV} + k\sqrt{\frac{R+L}{L}}\right)\right] \tag{11}$$

It is seen that the fill rate depends upon four parameters: k, CV, R and L. Note that the parameter R could be eliminated by redefining the unit of time to be R.

2.2 The Hadley and Whitin or de Kok approach

The Hadley and Whitin or de Kok approach clearly leads to the exact fill rate more quickly than our approach. However, it hinges on the following statement, which for some readers may not be particularly intuitive or easy to understand. In a particular cycle (see Figure 1), the *EUS* is the average amount backlogged at the end of the cycle ($R + L$ later) minus the average amount backlogged after the initial lead time (*i.e.*, L later), that is,

EUS = (expected amount that demand in R+L exceeds S)
– (expected amount that demand in L exceeds S).

Denoting the total demand in R+L by x and, as earlier, the demand in L by y,

$$EUS = \int_S^{\infty}(x_0 - S)\, f_x(x_0)dx_0 - \int_S^{\infty}(y_0 - S)\, f_y(y_0)dy_0 \tag{12}$$

Now, x has the normal distribution with mean μ_{R+L} and standard deviation σ_{R+L}. Using the associated density function of x, the density function of y (see equation (2)), defining S as in (8), and again employing unit normal transformations, (12) simplifies to the same *EUS* result as in (9). Hence, the fill rate expression of (11) follows. As an aside, note that (12) can be used for any (known) distributions of x and y.

3. Numerical illustrations including the use of a common approximation

A commonly used approximation (see, for example, Carlson and Miltenburg [15]) results from ignoring the second expression on the right side of (9) or, equivalently, the second integral in (12). Instead of (11) one obtains

$$P = 1 - \frac{\sqrt{R+L}}{R} CV\, G(k)$$

$$\text{or } G(k) = \frac{R(1-P)}{\sqrt{R+L}\, CV}. \tag{13}$$

Thus, specified values of *P*, *R*, *L* and *CV* will imply a value of $G(k)$. A table lookup or rational approximation (Brown [16] or Silver *et al*. [7]) can then be used to find the associated value of *k*. The simplicity (compared with solving (11) for *k*) was very important historically but is no longer a relevant issue. By neglecting the second portion of the right side of (9) the approximation overestimates the *EUS* for any value of *k*. Thus, it selects a *k* value (hence an *S* value) higher than the required value (found by using the correct formula of (11)). As a result, the actual fill rate achieved using the approximate method will, in general, be higher than the target value of *P*.

Table 1 shows two numerical illustrations of the use of both the exact approach (equation (12)) and the approximation of (13). Moreover, the actual fill rates achieved when (13) is used are shown. The second illustration demonstrates that the error (difference between targeted and actual fill rates) can be appreciable. In general, this tends to occur when the term $R/\sqrt{L}CV$, in the second *G* function of (9), is relatively small; in particular, when *L* is much larger than *R* and *CV* is not too small.

4. Conclusion

In this paper we have presented a new derivation of the exact fill rate in an (R, S) system under normally distributed demand and have compared it with a more direct, but perhaps less intuitive, derivation that already existed in the literature. We also presented a commonly used approximate method for selecting the value of the safety factor and pointed out that use of the approximate approach can lead to an achieved service level much higher than the targeted value. It should be mentioned that similar approximate approaches can also be quite inaccurate for other types of systems such as a periodic review, order point, order quantity system (Janssen *et al*. [17]). It is hoped that this paper will help ensure that the exact fill rate expression becomes much more widely known, hence used more in practice. Also, the intermediate result (see equation (9)) of an exact expression for the expected units short per cycle could be used in a cost minimizing model that incorporated a cost per unit short (see Chen and Zheng [9]).

Acknowledgments

The authors wish to specifically thank Ton de Kok for his input, including a critical review of a draft of the paper. Thanks are also extended to David Robb and Eric Johnson for their suggestions. In addition, we acknowledge Joe Blackburn and Sven Axsäter for their helpful comments with respect to some of the literature referenced in this paper. The research leading to this paper was partially supported by the Natural Sciences and Engineering Research Council of Canada under grant number 239153-05.

References

[1] Lee HL, Billington C. Managing supply chain inventory: pitfalls and opportunities. Sloan Management Review 1992; 33:65-73.

[2] Ward SC, Chapman CB, Klein JH. Theoretical versus applied models: The newsboy problem. OMEGA 1991; 19:197-206.

[3] Silver EA. Three ways of obtaining the average cost expression in a problem related to joint replenishment inventory control. Naval Research Logistics Quarterly 1973; 20:241-253.

[4] Hadley G, Whitin TM. Analysis of inventory systems. Englewood Cliffs, NJ: Prentice-Hall; 1963.

[5] de Kok AG. Hierarchical production planning for consumer goods. European Journal of Operational Research 1990; 45:55-69.

[6] de Kok AG. Analysis of one product/one location inventory control models. Unpublished course notes at Technische Universiteit Eindhoven [circa 2002].

[7] Silver EA, Pyke DF, Peterson R. Inventory management and production planning and scheduling, 3rd ed. New York, NY: John Wiley and Sons; 1998.

[8] Axsäter S. Inventory control, 2nd ed. Heidelberg, Germany: Springer Verlag; 2006.

[9] Chen F, Zheng Y.-S. Inventory models with general backorder costs. European Journal of Operational Research 1993; 65:175-186.

[10] Sobel MJ. Fill rates of single-stage and multistage supply systems. Manufacturing and Service Operations Management 2004; 6:41-52.

[11] Johnson ME, Lee HL, Davis T, Hall R. Expressions for item fill rates in periodic inventory systems. Naval Research Logistics 1995; 42:57-80.

[12] Silver EA. A modified formula for calculating customer service under continuous inventory review. AIIE Transactions 1970; 2:241-245.

[13] Silver EA, Smith DJ. Setting individual production rates under significant lead time conditions. INFOR 1981; 19:1-19.

[14] Ross SM. Introduction to probability models, 9th ed. Boston, MA: Academic Press; 2007.

[15] Carlson ML, Miltenburg GJ. Using the service point model to control large groups of items. OMEGA 1988; 16:481-489.

[16] Brown RG. Decision rules for inventory management. New York, NY: Holt, Rinehart and Winston; 1967.

[17] Janssen F, Heuts R, de Kok T. The value of information in an (R, s, Q) inventory model, working paper, Co-operation center Tilburg and Eindhoven Universities, Tilburg University, The Netherlands, 2009.

Table 1. Numerical illustrations of the exact and approximate choice of *k*

Case	Target P	R	L	CV	Exact k	Approximate k	Actual fill rate with approximate k
1	0.9	1	8	0.2	0.598	0.607	0.901
2	0.8	1	24	0.3	0.545	0.740	0.850

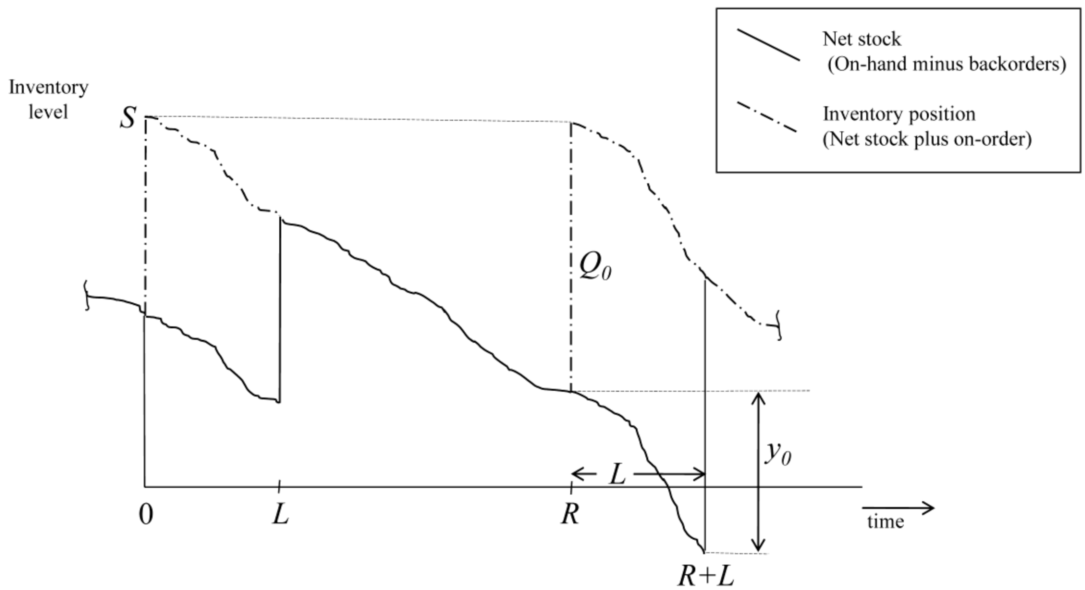

Figure 1 – Sample Behavior of Inventory Through a Cycle