ELSEVIER

On the single item fill rate for a finite horizon

Jiahua Chen[a], Dennis K.J. Lin[b], Douglas J. Thomas[b,*]

[a]*Department of Statistics and Actuarial Science, University of Waterloo, USA*
[b]*Smeal College of Business Administration, The Pennsylvania State University, University Park, PA 16802, USA*

Abstract

We investigate expressions for expected item fill rate in a periodic inventory system. The typical treatment of fill rate found in many operations management texts assumes infinite horizon, independent and stationary demand. For the case when the horizon is finite, we show that the expected value of the actual fill rate is greater than the value given by the infinite horizon expression. The implication of our results is that an inventory manager in a finite horizon situation who uses the infinite horizon expression to set stocking levels will achieve a higher than desired expected fill rate at greater than necessary inventory expense.
© 2002 Elsevier Science B.V. All rights reserved.

1. Introduction

The item fill rate is a pervasive measure of customer service in inventory systems and is defined as the fraction of demand satisfied from on-hand inventory (demand that is not backlogged or lost). Classic inventory expressions for fill rate rely on several assumptions. First, there is an infinite horizon over which demand is stationary. Second, if replenishment lead times are stochastic, the probability of orders crossing is negligible, and third, each replenishment order is sufficient to meet the existing backlog of orders. In this paper, we address the length of horizon effect. Specifically, we investigate an order-up-to inventory system with T periods and deterministic replenishment time from the end of one period to the beginning of the next.

Unmet demand can be lost or backlogged. We show that the infinite horizon fill rate expression underestimates the actual expected fill rate. Furthermore, the single period expected fill rate is always greater than the T period expected fill rate.

We are aware of contracts in practice where rewards (penalties) are doled out to suppliers that exceed (fall short of) predetermined fill rate targets over a finite review period. If such rewards and penalties are settled on a quarterly basis the supplier and customer should be interested in the quarterly fill rate performance rather than the traditional infinite horizon expression. Note that the T in our model is the number of *decision points* in the appropriate horizon. For example, weekly orders for a quarterly performance review would lead to $T = 13$.

It is common practice for inventory managers to use the infinite horizon expression to determine stocking levels. (Many operations management texts describe the procedure for using a standard Normal table or embedded spreadsheet function to determine stock

---

* Corresponding author.
E-mail addresses: jhchen@uwaterloo.ca (J. Chen), dkl5@psu.edu (D.K.J. Lin), dthomas@psu.edu (D.J. Thomas).

levels. See [2] for one example.) One implication of our results is that an inventory manager setting stock levels using this technique in a finite horizon setting will have an expected fill rate greater than that predicted by the infinite horizon expression.

Other authors have investigated underlying assumptions of fill rate expressions. Tyworth and O'Neill [6] investigate the sensitivity of the shape of the lead time demand distribution in a continuous review system. In fact, even if the demand distribution and parameters are obtained precisely by the firm's clairvoyance department, there are still inaccuracies in standard theoretical expressions. Johnson et al. [1] discuss the error in fill rate expressions in periodic inventory systems due to double counting of backorders. This only occurs when the lead time is longer than one period and the replenishment order is not sufficient to cover the existing backlogged demand. This double counting overestimates the number of stockouts, which leads to underestimating fill rate. Computational experiments conducted by Robinson et al. [4] suggests that stochastic lead time models that ignore the possibility of order crossover can significantly overestimate the inventory shortfall and thus underestimate the fill rate.

This paper is organized as follows. In Section 2 we introduce notation. In Section 3 we formally state and prove that the finite horizon fill rate lies between the one period fill rate and the infinite horizon fill rate. Furthermore, we develop two inequalities describing the relationship between T and $T + 1$ period fill rate expressions. Section 4 offers some conclusions and comments on the implication of the results.

2. Notation and preliminaries

In this order-up-to system, with stock level s, s units are on hand at the beginning of the period, demand is observed, if demand exceeds the stock level, some demand is backlogged, and finally a replenishment order is placed to bring next period's beginning stock level up to s.

Let X_t be independent and identically distributed (iid) demand random variables. We denote the number of units satisfied from the shelf in period t by $Y_t \equiv \min(s, X_t)$, the smaller of demand or stock available. Define $\alpha_T(s)$ to be the fraction of demand satisfied

from stock over T periods. The expected fill rate for T periods is

$$E[\alpha_T(s)] \equiv E\left[\frac{Y_1 + \cdots + Y_T}{X_1 + \cdots + X_T}\right].$$

We assume no specific form for demand distributions, however, we do require that the random variables be non-negative. We use the subscript 0 to denote *any* random variable having the appropriate distribution. That is, $X_1, \ldots, X_T$ represent demand random variables in periods $1, \ldots, T$, and X_0 refers to any random variable having the same distribution as $X_1, \ldots, X_T$. When demands are *iid* and $T \rightarrow \infty$, the traditional fill rate expression

$$\lim_{T \to \infty} E[\alpha_T(s)]$$

$$= 1 - \frac{\text{Expected units short in one period}}{\text{Expected demand in one period}}$$

$$= \frac{E[Y_0]}{E[X_0]}$$

follows directly from renewal theory [3, p. 192].

3. Analytic results

Our stated goal in this work is to understand the effect of horizon length on expected fill rate. Here, we establish three results that do not depend on the form of the distribution but do require that demand random variables are non-negative, *iid*.

Our first result establishes that for the same distribution and the same stock level, the expected fill rate over a finite horizon is greater than or equal to infinite horizon fill rate when the stock level is greater than expected demand.

Theorem 1. *For any finite T,*

$$E[\alpha_T(s)] \geq \lim_{R \to \infty} E[\alpha_R(s)].$$

Proof. Without loss of generality, assume $E(X_1) = 1$. We will also assume that X_1 has absolutely continuous distribution with density function $f(x) > 0$ for all $x > 0$. The proof of the general case can be easily obtained from this case. We address the general case at the end of this proof.

For convenience, denote $S_T \equiv \sum_{i=1}^{T} X_i$, and we will work with $E[\alpha_{T+1}]$ instead. Due to the iid structure, we have

$$E[\alpha_{T+1}] = E\left[\frac{(T+1)Y_0}{X_0 + S_T}\right].$$

Since we assumed $E(X_1) = 1$, what we want to show becomes that the function

$$h(s) \equiv E\left[\frac{Y_0}{X_0 + S_T}\right] - \frac{1}{T+1}E(Y_0)$$

is non-negative for all $s \geq 0$.

It is seen that

$$h(s) = E\int_0^s \frac{x}{x+S_T}f(x)\,\mathrm{d}x + E\int_s^\infty \frac{s}{x+S_T}f(x)\,\mathrm{d}s$$
$$- \frac{1}{T+1}\left[\int_0^s xf(x)\,\mathrm{d}x + \int_s^\infty sf(x)\,\mathrm{d}x\right].$$

Hence, the first and second derivatives of h with respect to s are

$$h'(s) = \int_s^\infty E\left(\frac{1}{x+S_T} - \frac{1}{T+1}\right)f(x)\,\mathrm{d}x$$

and

$$h''(s) = f(s)E\left[\frac{1}{T+1} - \frac{1}{s+S_T}\right].$$

Based on these computations, let us list our observations about $h(s)$.

1. $h(0) = h(\infty) = 0$. This is obvious as $E[\alpha_T(0)] = 0$ and $E[\alpha_T(\infty)] = 1$ for all n.
2. $h'(s) \leq 0$ when $s \geq T+1$. Hence, $h(s)$ is a decreasing function when $s \geq T+1$. Since $h(\infty) = 0$, we conclude that $h(s) \geq 0$ when $s \geq T+1$.
3. By Jensen's inequality (If $g(x)$ is a convex function, then $E[g(X)] \geq g(E(X))$ when the relevant expectations exist.)

$$E\left[\frac{1}{X_0 + S_T}\right] \geq \frac{1}{E(X_0 + S_T)} = \frac{1}{T+1}.$$

4. The equality holds only when $P(X_0 = 1) = 1$. This is equivalent to saying that $h'(0) > 0$.

5. $h''(s) = 0$ has at most one solution in s. Note that if $h''(s) = 0$, then

$$E\left[\frac{1}{s+S_T}\right] = \frac{1}{T+1}.$$

6. However, the left-hand side is a strictly decreasing function of s. Let us call this point s^*. Hence, we conclude $h''(s) \leq 0$ when $s \leq s^*$ and $h''(s) \geq 0$ when $s \geq s^*$.
7. When $s^* \leq s$, $h'(s) \leq h'(\infty) = 0$ and hence $h(s)$ is a decreasing function. Therefore, $h(s) \geq h(\infty) = 0$ when $s \geq s^*$.
8. When $s^* \geq s$, $h(s)$ is a concave function. At the same time, $h(0) \geq 0$ and $h(s^*) \geq 0$ from point 5. Therefore, we must have $h(s) \geq 0$ for $0 \leq s \leq s^*$.

In conclusion, we have shown that $h(s) \geq 0$ for all s. If the distribution of X_0 is not absolutely continuous, we consider $X_i^* = X_i + \delta\varepsilon_i$ for $i = 0, 1, \ldots, T$ with $\delta > 0$ being a constant and ε_i's being iid exponential random variables with mean 1, and independent of X_i's. Hence, X_i^* are absolutely continuous with positive density function over $x > 0$ and so the proof applies to X_i^* (the expectation being $1 + \delta$ does not matter). Letting $\delta \to 0$, we obtain the result for the general case. $\square$

Next, we establish that for the same demand distribution and the same stocking level, the single period fill rate is greater than or equal to the T period fill rate. We require the following lemma.

Lemma 1. *Let $f(x_1, x_2, \ldots, x_k)$ and $g(x_1, x_2, \ldots, x_k)$ be two n-variate functions such that both f and g are increasing or decreasing functions concordantly in x_i for fixed $x_1, \ldots, x_{i-1}$ and $x_{i+1}, \ldots, x_n$, $i = 1, 2, \ldots, n$. Let $X_1, \ldots, X_k$ be independent random variables. Then*

$$Cov(f(X_1, \ldots, X_k), g(X_1, \ldots, X_k)) \geq 0.$$

Proof. We first prove the lemma for $k = 1$. Assume that both f and g are increasing functions of x_1, we have

$$\{f(y) - f(z)\}\{g(y) - g(z)\} \geq 0$$

for any y and z. This remains true if both f and g are decreasing functions. Let Y and Z be two independent

random variables with the same distribution as X_1. We have

$$E[\{f(Y) - f(Z)\}\{g(Y) - g(Z)\}] \geq 0.$$

Simplifying the expression, we get

$$2[E\{f(Y)g(Y)\} - E\{f(Y)\}E\{g(Z)\}] \geq 0.$$

That is, $Cov(f(X_1), g(X_1)) \geq 0$.

We assume the lemma is true for $k = 1, 2, \ldots, n-1$, and show that it is also true when $k = n$. The result for $k = 1$ implies

$$E\{f(X_1, X_2, \ldots, X_n)g(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}$$

$$\geq E\{f(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}$$

$$\times E\{g(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}.$$

Hence,

$$Cov(f(X_1, X_2, \ldots, X_n), g(X_1, X_2, \ldots, X_n))$$

$$= E[E\{f(X_1, X_2, \ldots, X_n)$$

$$g(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}]$$

$$- E\{f(X_1, X_2, \ldots, X_n)\}E\{g(X_1, X_2, \ldots, X_n)\}$$

$$\geq E[E\{f(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}$$

$$\times E\{g(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}]$$

$$- E\{f(X_1, X_2, \ldots, X_n)\}E\{g(X_1, X_2, \ldots, X_n)\}$$

$$= Cov(E\{f(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\},$$

$$\times E\{g(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}).$$

However, $E\{f(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}$ and $E\{g(X_1, X_2, \ldots, X_n)|X_2, \ldots, X_n\}$ are obviously increasing functions in each of X_i, $i = 2, 3, \ldots, n$ given others. The induction assumption hence works. This completes the proof.   □

Theorem 2. *For all $T \geq 1$ and $s > 0$, we have $E[\alpha_T(s)] \leq E[\alpha_1(s)]$.*

Proof. Due to the iid structure, we have

$$E[\alpha_T(s)] = E\left[\frac{TY_1}{S_T}\right].$$

Hence, we have

$$E[\alpha_1(s) - \alpha_T(s)] = \sum_{i=2}^{T} E\left[\frac{Y_1}{X_1} \frac{S_T - TX_1}{S_T}\right].$$

Let

$$f(X_1, \ldots, X_T) = \frac{Y_1}{X_1}$$

and

$$g(X_1, \ldots, X_T) = \frac{S_T - TX_1}{S_T}.$$

Note f and g are both decreasing function in X_1, and both increasing function in $X_2, \ldots, X_T$. Hence, by the lemma

$$Cov(f(X_1, \ldots, X_T), g(X_1, \ldots, X_T)) \geq 0.$$

At the same time, it is easy to see that

$$E[g(X_1, \ldots, X_T)] = 0.$$

Hence,

$$E[f(X_1, \ldots, X_T)g(X_1, \ldots, X_T)]$$

$$= Cov(f(X_1, \ldots, X_T), g(X_1, \ldots, X_T)) \geq 0.$$

This completes the proof.   □

These two results establish that for the same distribution and stocking level, the T period fill rate lies below the one period fill rate and above the infinite period fill rate when the stock level is greater than expected demand. One might expect that these expressions are monotonically decreasing with T, $E[\alpha_{T+1}(s)] \leq E[\alpha_T(s)]$. Indeed, we conjecture that this is true although we have been unable to prove it. We can, however, establish two inequalities relating $E[\alpha_{T+1}(s)]$ and $E[\alpha_T(s)]$ stated in Theorem 3.

Theorem 3. *For any $s > 0$,*

(i) $TE[\alpha_{T+1}](s) \leq (T + 1)E[\alpha_T(s)]$;

(ii) $(T+1)E[\alpha_{T+1}](s) \leq TE[\alpha_T(s)] + E(Y_1)E[\bar{X}_T^{-1}]$, *where $\bar{X}_T = T^{-1}S_T$.*

Proof. First, we have

$$E[\alpha_{T+1}(s)] = E\left[\frac{(T+1)Y_1}{S_{T+1}}\right]$$

$$\leqslant E\left[\frac{(T+1)Y_1}{S_T}\right]$$

$$= [1 + T^{-1}]E[\alpha_T(s)].$$

This establishes (i).

Next, we write

$$S_{T+1} = T^{-1}\sum_{j=1}^{T+1} S_{-j},$$

where $S_{-j} = S_{T+1} - X_j$. Hence

$$T(T+1)^{-1}S_{T+1}$$

$$=(T+1)^{-1}\sum_{j=1}^{T+1} S_{-j} \geqslant \left[(T+1)^{-1}\sum_{j=1}^{T+1} S_{-j}^{-1}\right]^{-1}$$

as the arithmetic mean is greater than or equal to the harmonic mean for positive quantities. Therefore,

$$[(T+1)^{-1}S_{T+1}]^{-1} \leqslant n\left[(T+1)^{-1}\sum_{j=1}^{T+1} S_{-j}^{-1}\right].$$

Consequently,

$$E[\alpha_{T+1}(s)] = E[Y_1\{(T+1)^{-1}S_{T+1}\}^{-1}]$$

$$\geqslant (T+1)^{-1}\sum_{j=1}^{T+1} E[TY_1 S_{-j}^{-1}]$$

$$= \frac{T}{T+1} E[TY_1 S_T^{-1}] + \frac{1}{T+1} E[TY_1 S_{-1}^{-1}]$$

$$= [1 - (T+1)^{-1}]E[\alpha_T(s)]$$

$$+ (T+1)^{-1}E(Y_1)E(\bar{X}_T^{-1}).$$

This proves (ii). $\square$

Remark. When T is large, $E(\bar{X}_T^{-1}) \approx [E(X_1)]^{-1}$. Then, inequality (ii) becomes

$$E[\alpha_{T+1}(s)] \leqslant [1 - (T+1)^{-1}]E[\alpha_T(s)]$$

$$+ (T+1)^{-1}E[\alpha_\infty(s)].$$

At the same time, $E[\alpha_T(s)] \geqslant E[\alpha_\infty(s)]$ is true at least for $s > E(X)$. Considering this, our conjecture seems probable, especially for $s > E(X)$. Note that order-up-to levels will rarely be set below the mean.

4. Conclusion

We have investigated the effect that a finite horizon has on the single item expected fill rate. We have shown analytically that the expected fill rate for $T > 1$ periods lies between the expected fill rate with $T = 1$ and the commonly used infinite horizon fill rate expression.

It is important to note here that our results deal with the *expected* fill rate only. While understanding the bias in expected fill rate is important for practitioners, the behavior of other moments of the random variable will be important. The related paper by Thomas [5] further investigates the effect that the finite horizon assumption has on inventory performance.

We started this discussion by noting that several assumptions underlying common fill rate expressions may in fact not hold. Here, we investigated the effect of the horizon length assumption. It is interesting to note that potential double counting of backorders, order crossover (for systems with stochastic lead times) and our finite horizon results presented here all have the same directional effect, namely that the expected fill rate is underestimated.

References

[1] M.E. Johnson, H.L. Lee, T. Davis, R. Hall, Expressions for fill rates in periodic inventory systems, Nav. Res. Logist. 42 (1995) 57–80.

[2] S. Nahmias, Production and Operations Analysis, 2nd Edition, Irwin, Homewood, IL, 1993.

[3] S.I. Resnick, Adventures in Stochastic Processes, Birkhäuser, Basel, 1992.

[4] L. Robinson, J. Bradley, L.J. Thomas, Consequence of order crossover under order-up-to policies, Manuf. Serv. Oper. Manage. 3 (3) (2002) 175–188.

[5] D. Thomas, Measuring item fill rate performance in a finite horizon, Working Paper, Smeal College of Business Administration, Penn State University, 2002.

[6] J.E. Tyworth, L. O'Neil, Robustness of the normal approximation of lead-time demand in a distribution setting, Nav. Res. Logist. 44 (1997) 165–186.