A Partial Backorder Inventory System with Quoted Service Time

Yifan Dou Yongbo Xiao Jian Chen

Research Center for Contemporary Management, Key Research Institute of Humanities and Social Sciences at Universities, School of Economics and Management, Tsinghua University, Beijing 100084, China
Email: xiaoyb@sem.tsinghua.edu.cn

Abstract

In this paper, we introduce a critical time point method of partial backorder system with quoted service time in the traditional (Q,r) model: The retailer sets a critical time point. Excess demand coming before that point will be satisfied with emergency order while other ones will be backlogged until replenishment. Meanwhile, retailer would also announce quoted service time for their customers. Customers are willing to wait for this service time with no backorder cost even they are backlogged.

Keywords: partial backorder, quoted service time, (Q,r) system

1. Introduction

The concept of "backorder" is becoming more and more opaque in these on-line retailers. In old times, when customers step into a bookstore and request for some certain book, it might be disappointing if the bookstore is running out of that one. Nowadays, customers don't know about the exact number of inventory when they click on the websites. There is no more backorder cost during the order and delivery time because customers would always like to wait because of the convenience of on-line purchasing. Due to this new pattern of customers' demand and behavior, retailers are required to reformulate their partial backorder system and to set a more competitive quoted service time, which is HBQ (Hybrid backorder with Quoted service time). There are two time decisions for the retailer: Emergency order time point for internal control; demand comes before that will be filled by emergency order; Quoted service time for the customer. There is no backorder cost for demands come during time period from quoted service time point to the end of leadtime. Two time points come to be the same only when backorder cost is fairly large (which we will show in simulation.) Huang et.al (2007) look into a similar problem in online retailer and reformulate the decision rule of emergency order, which is firstly developed by Axsater (2003). We will extend these previous works into the continuous review (Q,r) system.

2. Brief Literature Review

A vast number of literatures related to our research exist mainly on partial backorder issue. Montgomery et.al (1973) introduces a partial backorder policy that a fraction of unfilled demand is backlogged. Kim and Park (1985) suggest another scenario in which the cost of backorder is assumed to be proportional to length of waiting time.

Moinzadeh (1989) sets customer's waiting time to a constant number, which

is quite similar to us. However, the basic inventory system they adopted is $(S-1, s)$ system. Rabinowitz et.al (1995) analyzes (Q, r) system with a backorder upper bound b. That is, if number of accumulated backorder is more than b, emergency order should be applied to fill the excess demand.

Another topic related to our research is quoted service time, which is widely studied in queuing theory. Analytical work is firstly done by Bertrand (1983). Wein (1991) and many literatures later focus on conditional sojourn time in analyzing DDLT (Due Date Lead Time). Yano (1987) developed a newsboy model to determine the safety lead time. Duenyas and Hopp (1995) also review this problem with method of semi-Markovian decision process. They also connect their work with scheduling problem. Kut.C.So and Song (1998) conclude most findings and categorize methods into 3 levels: quick service, quoted service time, uniform service time. They also develop a model regarding the relationship between service time and capacity and price.

Axsater (2003) introduces a partial backorder system in which unfilled demand can be satisfied with transshipment. In this work, our marginal contribution is to introduce a model with critical time point as Zhang et.al (2003) with quoted service time.

3. The Model

We consider an inventory system with constant lead time L. T is the quoted service time: no backorder cost will be accumulated during first T unit of time for each demand. The interpretation of quoted service time is the response or reaction time which is quite common in web-based retail industry. τ and L-T are two critical cutoff time points in this model. Once the inventory level falls be-

low zero, the demand comes before τ will be filled with emergency orders, others will be backordered. However, the retailers don't promise to deliver service precisely after quoted time T. If demands are not filled after T, time-dependent backorder will then be accumulated until replenishment.

A reasonable T only exists on $[0, L]$. Suppose there is a T which is larger than L, the retailer do not need to hold any inventory and just fill every demand with new order if the fixed order cost is not too large.

We also apply following notations:
e: per unit cost of emergency order;
b: per unit-time cost of each backordered demand.
Q: batch size of the order;
r: reorder point;
K: fixed order cost;
C: per unit variable ordering cost;
h: per unit-time holding cost for each item.
$I(t)$: inventory level of system at time t;
We set time point of reorder as 0; $I(0) = r$;
$T(r)$: time point of rth demand to arrive.
$I(T(r)) = 0$, When Poisson demand is applied, $T(r)$ is Erlang distributed with parameter of r and λ;
$D(t)$: cumulative demands in time period $[0,t]$.
$\bar{T}$: cycle time of inventory system; $\bar{T}$ is time interval between sending out successive 2 orders. For example, if $I(0) = r$, then $I(\bar{T}) = r$(e.g. see figure 1).
I_C: accumulated inventory in a cycle;
E_C: accumulated number of emergency orders in a cycle;
B_C: accumulated time of backorder in a cycle;
B: number of backorders in a cycle, $E(B) = \lambda E(B_C)$;

Briefly we confirm the model again. Demand arriving at time t will be immediately met if $I(t)$ is positive. If not, the demand will be filled with emergency or-

der if $t < \tau$. Otherwise, the customer will be told a quoted service time T. When customers turn back after this quoted service time, they may still be backordered since the order is still outstanding. Time-dependent backorder cost will be charged since then until this demand is finally satisfied with replenishment.

The major difference between our model and Zhang (2003) is the quoted service time T. With this quoted service time, the possible optimal τ will be found only on $[0, L–T)$. (Otherwise, if we got another "optimal" cutoff time τ' on $[L–T, L]$.Then orders between $[L-T, \tau']$ will be

satisfied with emergency orders. This is totally unnecessary because originally these demands will be filled at time point L with no backorder cost.)This analysis is useful in industry, especially for those on-line retailers.

We set the batch size Q and reorder point r to be exogenous. The decision variables are critical time τ, quoted service time T. Total inventory system can be illustrated by figure 1. Dash line indicates those backorders with no backorder cost charged. Note here, since $T(r)$ is Erlang-distributed random variable, it is not necessarily less than τ.
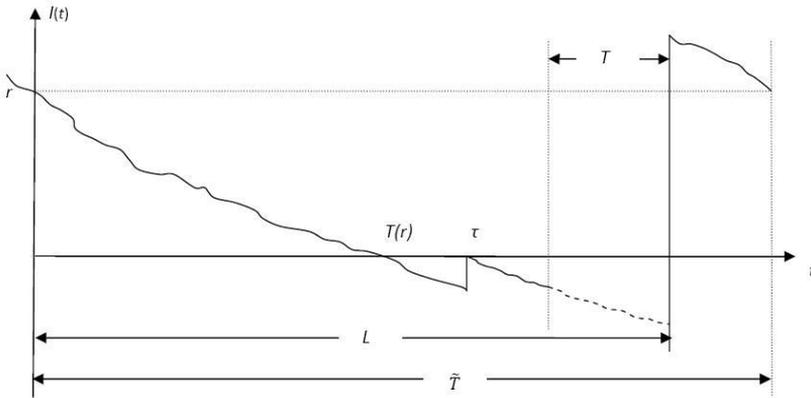

Fig. 1: Inventory Cycle when $T(r) < \tau$.

Expected average cost will be:

$$AC(T,\tau) = \frac{E(C)}{E(T)} \qquad (1)$$

C here stands for total cost for each cycle. C consists of 6 parts:

$$E[C] = K + cQ + hE[I_C] \\ + eE[E_C] + bE[B_C] + IV(T) \qquad (2)$$

We denote

$$p(j : \lambda t) = \frac{(\lambda t)^j e^{-\lambda t}}{j!} \qquad (3)$$

In this way, the Erlang distribution p.d.f of $T(r)$ (see the definition above) is $\lambda p(r - 1 : \lambda t)$; we will have following approximation by integration by parts:

$$\int_0^L p(j : \lambda t)dt \approx \frac{1}{\lambda} P(j+1 : \lambda L) = \frac{1}{\lambda} \sum_{i=j+1}^{\infty} p(i : \lambda L) \qquad (4)$$

3.1. Expected emergency order per cycle: [Ec]

Let T_E denote the time period of emergency orders in a cycle. This time period exists only when r unit of inventory will be consumed within τ unit of time. (e.g.: Figure 1)

$$T_E = \begin{cases} 0 & when \ T(r) \geq \tau \\ \tau - T(r) & when \ T(r) < \tau \end{cases} \quad (5)$$

We can derive the expectation of T_E. Since the p.d.f of $T(r)$ follows Erlang distribution, we have

$$E(T_E) = \int_0^\tau (\tau - t) f(t) dt = \int_0^\tau (\tau - t) \lambda p(r-1 : \lambda t) dt$$

$$= \tau P(r : \lambda \tau) - \frac{r}{\lambda} P(r+1 : \lambda \tau) \quad (6)$$

Since the demand rate is λ, the expected number per cycle is

$$E[E_C] = \lambda \tau P(r : \lambda \tau) - r P(r+1 ; \lambda \tau) \quad (7)$$

3.2. Expected time per cycle: $E[\overline{T}]$

In the traditional model, we set the expected cycle time to be the service time to fill the demand only with regular orders, i.e. $\frac{Q}{\lambda}$. When emergency order is considered, we include the expected time of Emergency order $E[T_E]$.

We have the following definition of expected cycle time:

$$E(\overline{T}) = \frac{Q}{r} + \tau P(r : \lambda \tau) - \frac{r}{\lambda} P(r+1 : \lambda \tau) \quad (8)$$

3.3. Expected backorders per cycle: $E[B_C]$

We denote T_B as the time period of backorders. We can see from figure 1 that existence of T_B is related to $T(r)$.

$$T_B = \begin{cases} 0 & when \ T(r) \geq L-T \\ L-T-T(r) & when \ \tau \leq T(r) < L-T \\ L-T-\tau & when \ T(r) < \tau \end{cases} \quad (9)$$

In this way, we can derive the expected time period for backorders,

$$E(T_B) = \int_\tau^{L-T} (L-T-t) f(t) dt + \int_0^\tau (L-T-\tau) f(t) dt$$

$$= \int_0^{L-T} (L-T) f(t) dt - \int_0^\tau \tau f(t) dt - \int_\tau^{L-T} t f(t) dt$$

$$= (L-T) P[r : \lambda(L-T)] - \tau P(r : \lambda \tau)$$

$$+ \frac{r}{\lambda} P(r+1 : \lambda \tau) - \frac{r}{\lambda} P[r+1 : \lambda(L-T)] \quad (10)$$

We can now get the expected number of backorders $E(B)$ per cycle:

$$E(B) = \lambda E(T_B)$$
$$= \lambda(L-T) P[r : \lambda(L-T)] - \tau \lambda P(r : \lambda \tau)$$
$$+ r P(r+1 : \lambda \tau) - r P[r+1 : \lambda(L-T)] \quad (11)$$

Here we are to compute the cumulative backorders held per cycle, which incurs time-dependent backorder costs. We can get cumulative backorders at time t, $B(t)$, if we replace $L - T$ with t. Since backorders only occur during $[\tau, L - T]$.

$$E(B_c) = \int_\tau^{L-T} B(t) dt = \int_\tau^{L-T} [\lambda t P(r : \lambda t) - \lambda \tau P(r : \lambda \tau) \\ + r P(r+1 : \lambda \tau) - r P(r+1 : \lambda t)] dt$$

$$= \frac{\lambda(L-T)^2}{2} P(r : \lambda(L-T)) + \frac{\lambda \tau^2}{2} P(r : \lambda \tau)$$

$$- \frac{r(r+1)}{2\lambda} P(r+2 : \lambda \tau) + \frac{r(r+1)}{2\lambda} P(r+2 : \lambda(L-T))$$

$$+ r(L-T) P(r+1 : \lambda \tau) - r(L-T) P(r+1 : \lambda(L-T))$$

$$- \lambda \tau(L-T) P(r : \lambda \tau) \quad (12)$$

3.4. Expected inventory per cycle: $E[I_C]$

Obviously cumulative inventory consists of two parts: cumulative inventory during leadtime, I_B , and cumulative inventory after replenishment ,I_A. For I_B, we can easily derive from $I(t)$. Note here that we should include time period $[L - T ; L]$ when computing holding cost:

$$
\begin{aligned}
E(I_B) &= \int_0^L E[I(t)]^+ dt = \int_0^L E[r - D(t)]^+ dt \\
&= \int_0^L [\sum_{x=0}^{r-1} (r-x)p(x:\lambda t)]dt \\
&= rL - \frac{rL^2}{2} + \frac{rL^2}{2}P(r-1:\lambda L) \\
&\quad - rLP(r:\lambda L) + \frac{r(r+1)}{2\lambda}P(r+1:\lambda L)
\end{aligned}
\tag{13}
$$

It will be far more complicated to derive I_A. Because I_A is highly related to the Inventory level at(or just before) the replenish moment $I(L)$. Furthermore, $I(L)$ is determined by the cumulative demands during $[\tau, L]$ and inventory level $I(\tau)$.

Since all demands arrive before time point τ can be filled (either by regular or emergency order), inventory level at time τ will never be negative because no backorder will occur till τ .Distribution of $I(\tau)$ is

$$
\Pr(I(\tau) = j) = \begin{cases} P(r:\lambda\tau) & when \ \ j = 0; \\ p(r-j:\lambda\tau) & when \ \ j = 1,2,3......r. \end{cases}
\tag{14}
$$

Next, we are to derive the probability mass function of $I(L)$ with assumption of only one outstanding order. A lot of literature (Hadley and Whitin 1963, Zhang 2003) thoroughly discussed the importance of this assumption. Here we especially emphasize single outstanding order because only in this way, we can limit the state space of inventory level of $I(L)$ to finite $(Q + r + 1)$ states. We will have(see Appendix of Zhang 2003)

$$
\begin{aligned}
E(I_A) &= \frac{1}{2\lambda}(1 - p(Q - r : \lambda(L-\tau)))P(r:\lambda\tau) \\
&\quad - \sum_{i-1}^{r} p((i+Q-r):\lambda)p(r-i:\lambda\tau) \\
&\quad \times [Q^2 + [2E[I(L)+1]Q \\
&\quad + E[I^2(L)] + E[L] - r(r+1)]]
\end{aligned}
\tag{15}
$$

Now we have the total expected number of inventory $E[I_C]$.

$$
\begin{aligned}
E(I_C) &= rL - \frac{rL^2}{2} + \frac{rL^2}{2}P(r-1:\lambda L) \\
&\quad - rLP(r:\lambda L) + \frac{r(r+1)}{2\lambda}P(r+1:\lambda L) \\
&\quad + \frac{1}{2\lambda}(1 - p(Q - r : \lambda(L-\tau)))P(r:\lambda\tau) \\
&\quad - \sum_{i-1}^{r} p((i+Q-r):\lambda)p(r-i:\lambda\tau) \\
&\quad \times [Q^2 + [2E[I(L)+1]Q + E[I^2(L)] \\
&\quad + E[L] - r(r+1)]]
\end{aligned}
\tag{16}
$$

Now we have all the parts of total cost with decision variable τ and T. τ is bounded by$[0, L - T]$.

The optimizing problem comes to following nonlinear minimizing problem.

$$
\begin{aligned}
&\min_{T,\tau} && AC(T,\tau) \\
&s.t && 0 \le \tau \le L - T \\
&&& 0 \le T \le L
\end{aligned}
\tag{17}
$$

Which is solvable by numerical studies. We will show the numerical results in the extension of this work.

4. Conclusions

In our analysis, we review a classical model of critical point method in PB system and relative literatures in quoted service time. We denote a joint optimization problem for business to manage their demand in a more flexible way. This is especially useful for those online businesses. We set up the average

cost for this model by deriving all the parts of the cost.

Acknowledgements

References

[1]. Axs*a*ter, S. (2003), A new decision rule for lateral transshipments in inventory systems, *Management Science* , 49, 1168-1179.

[2]. Bertrand, J. (1983), The effect of workload dependent due-dates on job shop performance. *Management Science*, 29, 799-816.

[3]. Duenyas, I., and Hopp, W. (1995), Quoted customer lead times, *Management Science* , 41, 43-57.

[4]. Hadley, G. and Whitin, T. (1963), *Analysis of Inventory System*, Prentice Hall. Englewood Cliffs, NJ

[5]. Kim, D. and Park, K. (1985), (Q, r) Inventory model with a mixture of lost sales and time weighted backorders, *Journal of the Operational Research Society*, 36, 231-238.

[6]. Moinzadeh, K. (1989), Operating characteristics for the (S-l. S) inventory systems with partial backorders and constant resupply time, *Management Science* , 35, 472-477.

[7]. Montgomery, D.C., Bazaraa, M.S. and Keswani, A.K. (1973), Inventory models with a mixture of backorders and lost sales, *Naval Research Logistics Quarterly*, 20, 255-263.

[8]. Rabinowitz, G., Mehrez, A., Chu, C. and Patuwo, B.E. (1995). A partial backorder control for continuous review (Q,r) inventory system with Poisson demand and constant leadtime, *Computers and Operations Research* , 22, 689-700.

[9]. Ross, S. (1970), *Applied Probability Models with Optimization Application*, Holden-Day, San Francisco. CA.

[10]. Kut, C.S. and Song, J. (1998), Price, delivery time guarantees and capacity selection. *European Journal of Operational Research,* 111, *28*-49.

[11]. Wein, L. (1991), Due-date setting and priority sequencing in a multi-class M/G/1 queue, *Management Science* , 37, 834-850.

[12]. Yano, C.A.(1987), Setting planned leadtimes in series production systems with tardiness costs. *Management Science,* 33*, 95*-106.

[13]. Zhang, G. et.al (2003), A hybrid inventory system with a time limit on backorders, *IIE transactions* , 35, 679-687.