Supply Chain Management in the Computer Industry

Katta G. Murty

Department of Industrial and Operations Engineering

University of Michigan, Ann Arbor, Michigan 48109–2117, USA

Phone: 734-763-3513; fax: 734-764-3451; e-mail: katta_murty@umich.edu

April 29, 2000

Abstract

An important issue in supply chain management in manufacturing industries is to forecast the demand for each item and to determine when to place orders for it and the order quantities. This issue has been analyzed extensively in the literature using mathematical models of varying degrees of complexity. We present the difficulties encountered in implementing these existing methods in the computer industry, and present a new paradigm that overcomes some of these difficulties.

Key words: Supply contracts, demand distribution and its updating, forecasting demand, shortage and excess penalties, order quantity determination, newsvendor approach

1 Introduction

In manufacturing industries, companies buy items (raw materials, partially finished components, etc.) from their suppliers, put these items through additional processing as necessary, assemble them into their final product, which they then sell to their customers. Traditionally companies maintained inventories of items, but in recent times due to the great success of the *Just-in-Time* philosophy of materials management, all companies are putting a great deal of effort into making the inventories of items as low as possible or necessary. In the computer industry the main driving force behind the effort to keep inventories as small as possible is the rapid rate of technological change and shortened product life cycles, with the result that prices of items fall over time, often every week, and sometimes drastically.

Some of the earliest OR problems to be analyzed mathematically are inventory control problems. We will consider the case where the demand for items is essentially stochastic, which applies to most practical problems in industry. Much of the classical analysis in inventory management literature is based on the distribution of demand being known, typically assumed to be normal. In industry, the reasonableness of the normality assumption is rarely checked, and we realize that this is also not practical as most companies deal with a large number of different items. An additional complication is the fact that often the demand distribution is changing with time.

The normality assumption confers many theoretical advantages, one of them is that the normal distribution is fully characterized by only two parameters, the mean and the standard deviation. So, when the distribution changes, one has to just change the values of these two parameters in the models. In practice, almost always it is only the value of the mean that is changed, the standard deviation is usually assumed to remain the same.

To help keep inventories at minimum possible levels, successful inventory management systems depend heavily on good demand forecasts. The main aim of forecasting methods is to predict future demand, and the output of forecasting is usually presented in the literature as *the forecasted demand quantity for a future period.* Because of this, the purpose of forecasting is often misunderstood to be that of generating a single number, even though sometimes its mean average deviation is also mentioned.

Implicit in the assumption that the demand in a future period is a random variable, is the understanding that we cannot predict its true value in advance; the most that we can do is to derive estimates of its probability distribution, or conditional distribution given the present and past demand values; or important characteristics of these distributions.

Much of the discussion in demand forecasting literature is focussed on procedures for deriving the forecasted demand quantity [2, 6]; in reality what these procedures are doing is estimating the expectation or conditional expectation of future demand. So these procedures are only useful when the assumption is made that the demand follows a distribution like the normal distribution that is completely characterized by a single parameter, the expectation.

The results obtained indeed depend on the demand distribution. So, unless the changes in the demand distribution can be captured reasonably well by changing the value of a single parameter, it is difficult to use the models in classical literature in practice. Also, these models require several other pieces of information such as ordering costs, holding costs, shortage costs, etc. that companies are not able to provide with reasonable accuracy.

In this note we present a simplified paradigm for inventory management in the computer industry that overcomes some of these difficulties, and has been received favorably by some of the companies.

2 Supply Contracts in the Computer Industry

We will use the abbreviation CC for a computer company.

Supply contracts for major components in the computer industry usually have the following features: tentative contracts are drawn for each quarter (of the year) about two weeks before the beginning of the quarter, in which the buyer makes purchasing commitments to the supplier for that quarter, and approximate times during the quarter when orders will be placed with approximate order quantities. The actual order placing dates and the purchase quantities can be modified as time passes and more information about the demand becomes available.

If the actual quantity lifted during the quarter, y, is within a certain percentage (say $\pm\alpha\%$, where α is typically 20) of the commited amount, there is no penalty. If $y <$ 80% of the commited amount, by paying a certain penalty (typically a measure of the suppliers holding cost for the items from one quarter to the next) to the supplier, the CC can agree to pass on the shortfall to the next quarter [1]. During the quarter if the CC realizes that they would actually need $>$ 20% over the committed amount, they may be able to get a portion of this excess amount from the supplier, again by paying some penalty (may be an increase in the purchase price per item, say up to 20%).

Then during the quarter, whenever the CC needs to have the stock of the component replenished, an order is sent. The supplier delivers the order within the lead time for that component. The lead time for delivering an order may depend on the order quantity, but usually for established suppliers with long term relationships and with established order patterns, the lead time can be assumed to be a constant as long as the order qualtity is within the usual range.

CCs usually negotiate deals with their long term suppliers whereby the price of important components decreases from one quarter to the next, particularly for newly introduced components (this may be of the order of 3% of the price or more). For many components, price decreases almost weekly.

For several major components, the transportation (delivery) costs are usually included in their purchase price. For these components then, there is no separate delivery charge. For other components, we found that the transportation costs are very close to being linear in the order quantity. Bulk shipping savings, if any, are negligible. So, for optimizing ordering policies, we ignore transportation costs.

Supply chain management is now-a-days handled using properly designed computer systems. Companies are establishing close relationships with their suppliers including sharing of data by electronic exchange using the internet. Because of these developments, many of the costs such as ordering cost, order processing cost, order handling cost etc. in classical inventory models are either disappearing or becoming insignificant and we ignore them.

Many CCs are now-a-days providing space in their incoming matertials warehouses either directly in their plants or at locations very close to the plants, to their long term suppliers. This space, directly under the control, responsibility, and liability of the supplier, is to be used by them to bring and store the components from their plants, to be delivered to the CC as the demand arises on a *demand-pull* basis. With such an arrangement, usually the CC is required to provide about a week to ten days earlier, the number of units of each component they expect to lift in a week (this information to help the supplier bring adequate stock of components from their manufacturing facilities which may be far away, to the warehouse near the CC). Then every morning the CC puts in an order for the number of units of each component they need for that day. The supplier then delivers that quantity from the local warehouse with a lead time of about 4 hours or less. This type of arrangement is now-a-days called *continuous replenishment.*

3 Seasonality in the Demand Pattern

In the computer industry majority of sales are arranged by sales agents who operate on quarterly sales goals. The sales agents usually work much harder towards the end of the quarter to meet their quarterly goals, so demand for products of the computer industry tends to be higher towards the end of the quarter. As most of the companies are *building to order* now-a-days, weekly production levels and demand for components inherit the same kind of seasonality within the quarter. Roughly, each quarter can be divided into two homogeneous periods; weeks 1 to 8 belong to the *slack period*, and weeks 9 to 13 belong to the *peak period.* As an example, for some components the slack period accounted for about 47% of the total demand, and the peak period about 53%

at a company.

In this case we analyze each of these slack and peak periods separately. The demand distributions for each of these periods are computed and updated separately using data pertaining to that period, and all the management decisions are made using them. Even though the distributions for these two periods are different, the procedures for computing and updating them are the same.

4 Demand Distributions, their Updating and Forecasting

The environment in the computer industry is very competitive with new products replacing the old periodically due to rapid advances in hardware technology. In this dynamic environment, the life cycle of any component or end product is approximately two years or less. Often, this life cycle is partitioned into three periods.

The *growth period* at the beginning of life sees the demand for the item growing from one week to the next due to gradual market penetration, reaching its peak by the end of this period. This is followed by a *stable period* during which the demand for the item stays relatively stable. This is followed by the final *decline period* during which the demand for the item undergoes a steady decline from one week to the next until at the end the item is replaced by a technologically superior one and disappears from the market. See Figure 1.

In recent times, the stable period seems to be getting shorter for many major components. Because of this constant change, models in the literature for supply management based on a single stable demand distribution are not appropriate for applications in the computer industry. We need to use models that frequently and periodically update the demand distributions based on most recent data.

Histograms of daily demand for some components indicate that the demand distributions are not usually symmetric around the mean. Some are skewed to the right of

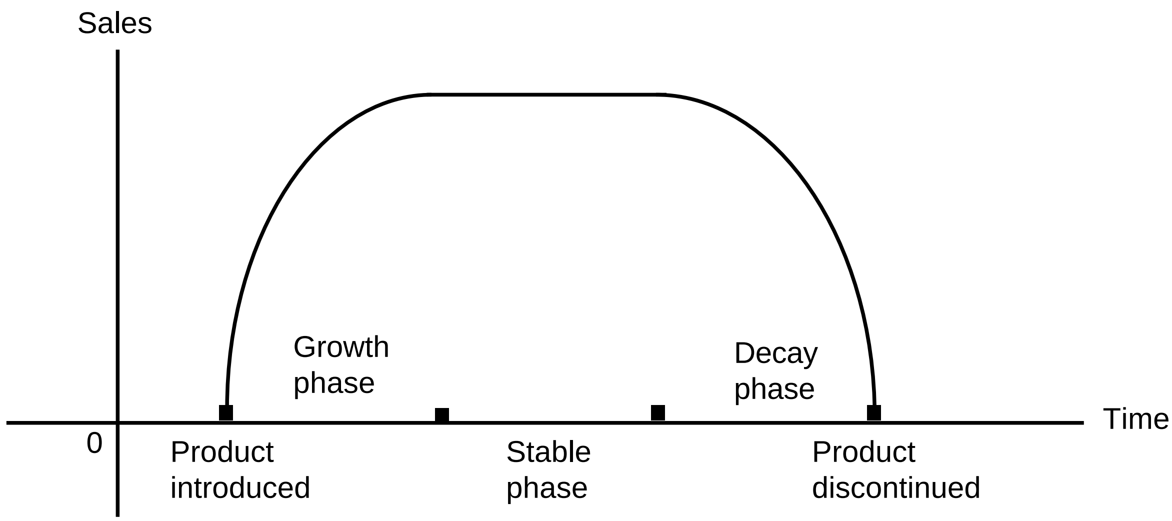

Figure 1: Life cycle of a typical product.

the mean, while others are skewed to the left. Also, as mentioned earlier, approximating the demand distribution by something like the normal or gamma distribution characterized by two or less parameters, only leaves us the freedom to change those one or two parameters when updating the demand distribution. This appears quite inadequate to capture all dynamic changes occuring in the shapes of demand distributions from time to time.

Considering these arguments, it seems that a better strategy would be to approximate the demand distribution by its histogram from past data. We divide the range of variation of the demand into a convenient number of demand intervals (in practice about 10 to 25), and take the relative frequencies among historical data as the probabilities associated with the intervals in the initial distribution. We will call this the *discretized demand distribution*. As an illustration, we give below in Figure 2 this initial

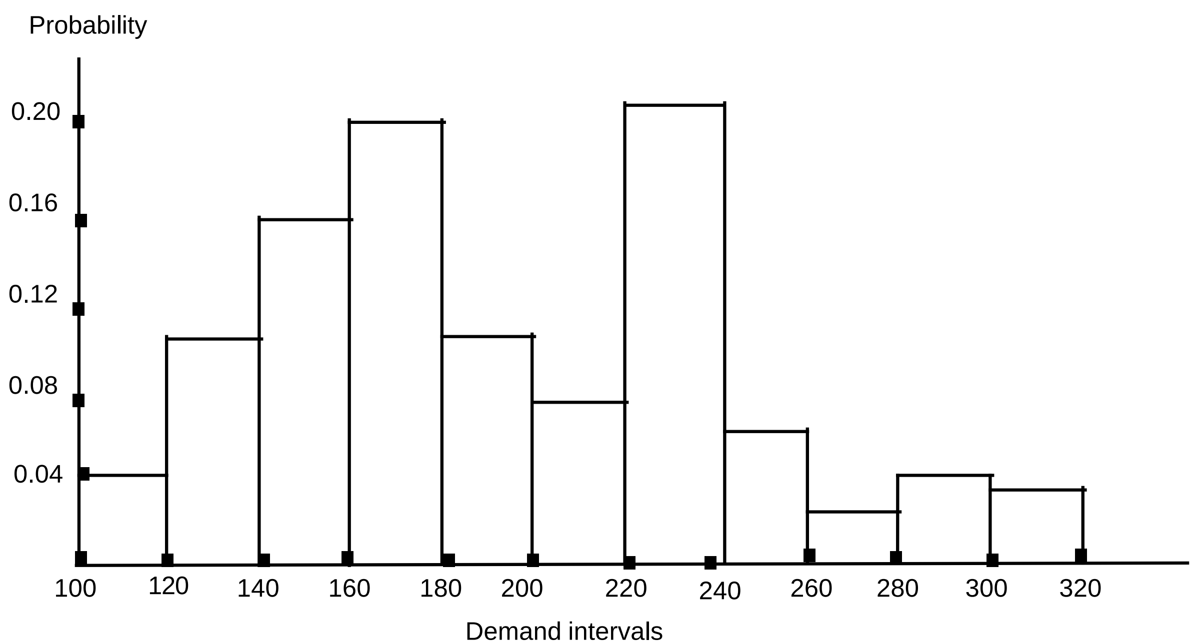

Figure 2: A discretized demand distribution.

distribution for the daily demand of a major component at a plant.

Let $I_1, \ldots, I_n$ be the demand intervals ; and $p_1, \ldots, p_n$ the probabilities associated with them in this discretized demand distribution. In updating the distribution, we have the freedom to change the values of all the p_i, this makes it possible for us to capture any change in the shape of the distribution subject to the discretization used (i.e., division of the range of variation into the demand intervals $I_1, \ldots, I_n$).

Updating the Demand Distribution

If orders are placed daily, it is convenient to have the discretized distribution of daily demand, and to update it daily whenever an order is placed. In this case the period is a day.

Let $I_1, \ldots, I_n$ be the demand intervals, and $p = (p_1, \ldots, p_n)^T$ the probability vector in the present distribution. At the time of updating, take the data on the demand in each period over the most recent k periods (if the period is a day, for example k could be about 50). For $i = 1$ to n, let $f = (f_1, \ldots, f_n)^T$ be the relative frequency vector corresponding to the most recent k periods, i.e., $f_i = r_i/k$ where r_i = number of periods among the most recent k for which the demand was in I_i, $i = 1$ to n.

f represents the estimate of the probability vector corresponding to the most recent demand distribution, but it is based on two few (only k) observations. p is the probability vector corresponding to the demand distribution at the time of previous updating.

Let $x = (x_1, \ldots, x_n)^T$ denote the updated probability vector corresponding to the updated demand distribution. We can take x to be the optimum solution of the following quadratic program:

$$\begin{aligned} \text{Minimize} \quad \beta \sum_{i=1}^{n} (p_i - x_i)^2 \quad &+ \quad (1-\beta) \sum_{i=1}^{n} (f_i - x_i)^2 \\ \text{subject to} \quad \sum_{i=1}^{n} x_i \quad &= \quad 1 \\ x_i \quad &\geq \quad 0 \quad \text{for all } i \end{aligned}$$

where β is a weight between 0 and 1. Typically $\beta = 0.9$ works well, the reason for choosing the weight in the second term in the objective function to be small is because the relative frequency vector f is based on a small number of observations. This is a small convex quadratic program with a unique solution which can be computed efficiently using any of several available algorithms, for example, Lemke's complementary pivot algorithm [5]. Since the quadratic model minimizes the weighted sum of squared forecast errors over all demand intervals, when used at every ordering period, it has the effect of tracking gradual changes in the demand distribution.

If the probabilities associated with demand intervals at either the left or right end are becoming very close to 0, it is an indication of significant increase or decrease in mean demand since the time the original demand intervals were set up. At that time we open new demand intervals at the other end and continue the same updating process.

Demand Forecasts

As mentioned earlier, existing forecasting techniques in the literature concentrate on estimating the expected demand, and hence are only useful when the demand distribution is approximated by something like the normal distribution characterized by the single parameter, the expectation. These classical procedures are not useful to solve inventory management problems with discretized demand distributions.

The procedure discussed above updates the discretized demand distribution at every ordering point using recent demand data. This discretized demand distribution can be combined with the simple classical news-vendor approach, Karlin [4], to yield very convenient inventory management rules for the most common supply management problems in the computer industry. This approach uses the updated demand distribution as the forecast.

5 News-Vendor Models for Periodic Review

As mentioned earlier, the following is a common scheme for replenishing major components in the computer industry today.

(i) Quarterly commitmemts: The CC informs the supplier well before the beginning of each quarter, the aggregate quantity of the component that they will order during that quarter. The purpose of this is to let the supplier know how much of their production capacity to dedicate to the manufacture of the buyer's components.

(ii) Weekly estimates: About ten days before the beginning of a week, the CC provides the supplier a revised estimate of the quantity of the component they would order during that week. This is to let the supplier know the quantity of the component to bring from their manufacturing facilities to the warehouse space near the

CC, and be ready to deliver according to the orders placed each day as described next. Usually the lead time for moving material from the supplier's plant to the warehouse space near the computer company may be several days to a week, and the supplier uses this information to plan those shipments.

(iii) Daily orders: Each workday morning the CC puts an order electronically for the quantity of the component to be delivered that day. This quantity is delivered usually within a lead time of about four hours.

Besides the market induced random fluctuations prominent in the weekly demand, the daily demand is also subject to random fluctuations induced by changes in production schedules on the shop floor. Even though a production schedule for a day is available in the morning of that day, it is subject to disruptions and alterations beyond the control of the company; and the actual daily production and consequently the daily demand for the components would not be known with certainty by the morning of that day. So we maintain the discretized demand distributions for daily demand of components, for determining optimum order quantities. We now develop inventory management rules for the scheme outlined above, using discretized daily demand distributions.

Model to Determine Daily Order Quantity

Let $I_1, \dots, I_n$ be the demand intervals, and $p = (p_1, \dots, p_n)^T$ the probability vector in the present discretized demand distribution for daily demand of a component. Let $u_1, \dots, u_n$ be the midpoints of the various demand intervals. Then $\mu_D = (\sum_{i=1}^n u_i p_i)$ is the expected daily demand for the component, and $\sigma_D = \sqrt{\sum_{i=1}^n p_i (u_i - \mu_D)^2}$ is the standard deviation of the daily demand for this component. Let:

$x_D =$ the daily demand quantity, the random variable.

$d =$ the quantity of this component made available for the process this day, it includes the leftover quantity from previous day, plus the quantity ordered for delivery the morning of this day.

$q =$ the quantity left over unused at the end of this day $= (d - x_D)$.

q will be negative if $x_D > d$. To determine d optimally, we minimize the expectation of a penalty function which contains an excess penalty term that applies when q is large, and a shortage penalty term when q is too small. We use this penalty function as a measure of the costs associated with various values of q, because the actual costs in dollar terms are very hard to assess accurately.

The left over quantity q at the end of this day has to support the production next morning until the delivery ordered next morning arrives at the CC for use on the production line. If q is negative or too small, the CC may be either forced to shut down the production line for lack of this component; or make an unplanned switch in production to some other product that does not need this component, until the next day's order is delivered. Let c_{0D}\$ represent this shortage penalty that the CC would like to associate with this occurrence, it can be taken as the inconvenience penalty of making an unplanned switch in the production schedule. Let q_{0D} denote the safety stock level needed to cover the production until the next morning's order is delivered at the CC (this depends on μ_D, σ_D, the lead time for delivery, and the production schedule for the morning). Then the shortage penalty of c_{0D} is incurred on this day if $q < q_{0D}$.

In the same way, let q_{1D} denote an upper bound on q that the CC aspires to have. The amount of convenient storage space near the production line that can be allocated for this component, and other factors like it go into determining the value of q_{1D}. Also, let c_{1D}\$/unit represent the excess stock penalty that applies for each unit that q is over

q_{1D}.

If y is a real valued variable, we define two new variables associated with it called the positive and negative parts of y to be $y^+ = \text{maximum}\{0, y\}$ and $y^- = \text{maximum}\{0, -y\}$. Also, for any random variable z, we denote its expectation by $E(z)$.

So the total expected penalty associated with d per day is: $f(d) = c_{1D}E(d - x_D - q_{1D})^+ + c_{0D}$ Probability$(d - x_D - q_{0D} < 0)$. For each value of d, this can be computed easily using the discretized distribution of daily demand. To determine the optimal d, we compute the values of $f(d)$ in a sufficient range of values of d around μ_D and take the optimal value to be the one that gives the smallest value for $f(d)$. This is basically the news-vendor approach [4].

Then the optimal quantity to order during that day is: ((optimal d) $-$ (leftover quantity at the end of the previous day))$^+$.

Model to Determine Estimates of Total Orders in a Week

For determining the weekly order estimates, we need the distribution of weekly demand. It is usually difficult to get data over enough number of weeks to estimate the discretized weekly demand distribution with reasonable precision. However, since we are maintaining the discretized distribution of daily demand, it is possible to derive the discretized distribution of weekly demand by either using convolutions, or estimate it using simulation.

Whenever there is a backlog of orders, it is common in the computer industry to work overtime. Because of this, the number of workdays in a week can be 5, 6, or 7; the probabilities of which can be estimated from past data, suppose these are ρ_1, ρ_2, ρ_3.

Select a random variable r which takes values 5, 6, 7 with probabilities ρ_1, ρ_2, ρ_3. Then select r random variates from the discretized distribution of daily demand and add them up, yielding one observation of weekly demand. To estimate the discretized distribution of weekly demand using simulation, draw sufficient number of observations of weekly demand as above and form a histogram of these observations.

Let x_W denote the weekly demand quantity, the random variable; and d' denote the quantity of the component made available for the process in the week, it includes the leftover quantity from previous week, plus the quantity to be ordered this week. Knowing d', the supplier moves d' $-$(leftover quantity from previous week) from their production facility which may be far away, to the warehouse space controlled by them near the CC.

If $x_W > d'$, the supplier may not be able to meet the requirement completely because of the large lead time involved. In this case the CC may not be able to get replenishments until the following week, some of the products scheduled to be manufactured using the component will be delayed and a portion of those customer orders may be lost. Let c_{0W} in \$/unit short reflect the lost profit margin on customer orders lost due to delay. c_{0W} is the shortage penalty per unit short during the week.

If $x_W < d'$, usually there are agreements for them to transfer the excess upto a certain limit, s_1, say, at no cost to the following week ; and the excess over s_1 at a penalty of c_{1w} in \$/unit, to the following week.

Thus the overall expected penalty associated with d' is $f'(d') = c_{1w}E(d' - x_W - s_1)^+$ $+ \, c_{0W}E(x_W - d')^+$. For each value of d' this can be computed easily using the discretized distribution of weekly demand. To determine the optimal d', we compute the values of $f'(d')$ in a sufficient range of values of d' around the expected weekly demand and take the optimal value to be the one that gives the smallest value for $f'(d')$. Then the optimal estimate of new orders for the week is: ((optimal d') $-$ (leftover quantity at the end of the previous week))$^+$.

Model to Determine the Commitment for Total orders in a Quarter

Let x_Q denote the demand during a quarter. Since the probability distribution of demand is expected to change over a quarter, and the life of a component measured in quarters may be quite small, it is not practical to estimate the distribution of x_Q from actual data. However $x_Q = \sum_{i=1}^{13} x_i$ where x_i is the demand during the ith week of the quarter. Knowing the distributions of the weekly demands x_i, we can derive the distribution of x_Q by either using convolutions, or estimate it using simulation.

Let d'' denote the quantity of the component made available for the process in the quarter, it includes the leftover quantity from previous quarter, plus the quantity to be ordered this quarter. Suppose we have the following agreement:

> if $(1-\alpha)d'' \leq x_Q \leq (1+\alpha)d''$, where $0 \leq \alpha \leq 1$ is specified ($\alpha = 0.2$ is common), there is no penalty
>
> if $x_Q > (1+\alpha)d''$, there is a shortage penalty of $c_{0Q}(x_Q - (1+\alpha)d'')$
>
> if $x_Q < (1-\alpha)d''$, there is an excess penalty of $c_{1Q}((1-\alpha)d'' - x_Q)$ and the shortfall $d'' - x_Q$ is transferred to the aggregate order for the next quarter.

So the total penalty associated with d'' is $f''(d'') = c_{1Q}E((1-\alpha)d''-x_Q)^+ + c_{0Q}E(x_Q - (1+\alpha)d'')^+$. We need to determine d'' to minimize $f''(d'')$.

Let μ_Q denote the expected demand during the quarter. If the entire range of variation of x_Q is within the interval $[(1-\alpha)\mu_Q, (1+\alpha)\mu_Q]$ then we can take optimum d'' $= \mu_Q$ since $f''(\mu_Q)$ is then 0.

Otherwise we can determine the optimum d'' with the news-vendor approach using the estimated distribution of x_Q as before.

In this case it may also be reasonable to assume that x_Q, d'' are continuous variables. In this case $g(d'')$ = the derivative of $f''(d'') = c_{1Q}(1-\alpha)$ Prob$(x_Q \leq (1-\alpha)d'')$ $-c_{0Q}(1+\alpha)[1-$ Prob$(x \leq (1+\alpha)d'')]$. $g(d'')$ begins with a negative value at $d'' = 0$, and is monotonically increasing as d'' increases, assuming a positive value when d'' is large. So, $g(d'') = 0$ has a unique solution, which is the continuous value of d'' that minimizes $f''(d'')$. Knowing the distribution of weekly demand, the solution of $g(d'') = 0$ can be estimated easily using simulation.

Then the optimum commitment for the quarter is: ((optimum d'') $-$ (leftover quantity carried over from the previous quarter))$^+$.

Conclusion: We discussed simple procedures for determining order quantities for the most common ordering scheme in the computer industry, while maintaining the discretized distribution of daily demand. For other types of replenishment schemes, procedures for inventory management using diecretized distributions can be developed in a similar way.

Acknowledgement: I am grateful to Cosimo Spera for posing to me the problem of determining optimal order quantities using estimated demand distributions based on data.

6 References

[1] Bassok, Y., A. Bixby, R. Srinivasan, and H. Z. Wiesel, (1997), "Design of Component-Supply Contract with Commitment Revision Flexibility", *IBM J. Res. Develop.*, 41, 6, 693-702.

[2] Box, G.E., and G. M. Jenkins, (1976), *Time Series Analysis Forecasting and Control*, Holden-Day, San Francisco.

[3] Deis, P., (1983), *Production and Inventoty Management in the Technological Age*, Prentice-Hall, Englewood Cliffs, NJ.

[4] Karlin, S., (1958), "Optimal Inventory Policy for the Arrow-Harris-Marschak Dynamic Model", Chapter 9 in *Studies in the Mathematical Theory of Inventory Production*, K. J. Arrow, S. Karlin, and H. Scarf (eds.), Stanford University Press.

[5] Murty, K. G., (1988), *Linear Complementarity, Linear and Nonlinear Programming*, Heldermann Verlag, Berlin.

[6] Nahmias, S., (1989), *Production and Operations Analysis*, Irwin, Boston.