**A Robust Forecasting Technique for
Inventory and Leadtime Management**

by

Uday S. Karmarkar

1990

CMOM 90-03

**CENTER FOR
MANUFACTURING &
OPERATIONS MANAGEMENT**

Working Paper Series

WILLIAM E.
SIMON
GRADUATE SCHOOL
OF BUSINESS ADMINISTRATION
UNIVERSITY OF ROCHESTER
ROCHESTER, NEW YORK 14627

**A ROBUST FORECASTING TECHNIQUE FOR INVENTORY AND
LEADTIME MANAGEMENT**

Uday S. Karmarkar

**Center for Manufacturing and Operations Management
William E. Simon Graduate School of Business Administration
University of Rochester
Rochester, New York 14627**

CMOM Working Paper

January, 1990

A ROBUST FORECASTING TECHNIQUE FOR INVENTORY AND LEADTIME MANAGEMENT

Uday S. Karmarkar

Center for Manufacturing and Operations Management
William E. Simon Graduate School of Business Administration
University of Rochester
Rochester, New York 14627

Abstract

Inventory stocking problems with stochastic demand typically involve an estimate of the location of some fractile of the demand distribution, where the fractile is usually in the 0.8 - .99 range. This fractile is termed the "service level" and is the probability that demands will be satisfied from stock on hand. An analogous procedure must be followed to set safety times or time buffers for delivery and supply lead times. The conventional approach of using the Normal model to estimate this location can sometimes be misleading since it primarily uses information about the center of the distribution to predict behavior in the tail of the distribution.

A new method based on a mixture model is proposed to estimate the location of the appropriate fractile directly. A formal Bayesian approach is derived and heuristic smoothing methods are developed. Simulation is used to evaluate the methods. The estimate obtained is biased in general but robust in the sense that it works well for a variety of distributions. The performance of this method appears to be superior to the conventional Mean - MAD approach. The method is particularly well suited to application in base-stock inventory systems and for safety time determination in supply management when leadtimes are variable.

Introduction

Solutions to inventory management problems often involve setting stocking targets at some particular fractile of the demand distribution. For example, in the classical one-period newsboy problem with excess cost h and shortage cost p , the optimal stock level is at that fractile of the demand distribution determined by the ratio $p/(h + p)$. A similar decision is faced in setting "safety times" when lead times for purchasing or production are variable. Here the time allowed for production or delivery is analogously set at a fractile of the lead time distribution. Hereafter we cast the discussion in terms of the inventory case, recognizing that the results apply to both. We will return to the lead time issue when discussing applications. While the newsboy problem does not usually occur in its pristine form in realistic problems, it often appears as a subproblem in iterative procedures or as an approximation of more exact methods. Furthermore, the concept of a service policy is firmly imbedded in thinking about inventory problems. In some logistics situations it may be natural to specify the service level desired as a priori policy even where the cost parameters of the problem are hard to determine. Suffice it to say that there are many situations of practical importance in inventory management where it is of interest to determine particular fractiles of the demand distribution and that these are generally tail probabilities in the 0.8 - 0.99 range. This paper develops an approach to fractile estimation that is robust and computationally reasonable.

Consider the conventional approach to this problem as used in some inventory management systems for setting order points. A Normal demand distribution is assumed and the statistics collected are usually the Mean and the Mean Absolute Deviation (MAD). The desired fractile location is then estimated as an offset from the mean, using the fact that for a Normal model the population standard deviation is given by 1.25 times the population MAD. The use of the MAD statistic instead of the technically appropriate variance deserves some comment; the reason that is sometimes cited is that computational savings are obtainable by its use. This argument has lost force with the advent of cheap computing power, but as we shall later see, there appear to be other reasons favoring MAD over variance. In either case, the assumption of Normality allows the estimates to be made independently, and methods such as exponential smoothing are used to make some allowance for nonstationarity of the demand process.

The assumption of Normality is likely to be a good one when the coefficient of variation of the demand distribution is small. There are however, many situations where in practice the underlying distribution of demand is likely to be significantly non-Normal. The question arises as to whether the Normal model is useful in such a case. Recalling that the optimal decision in setting order

points consists of the choice of the appropriate fractile P (say) of the distribution of demand; it is clear that any model chosen must be judged on how well this fractile is matched. Since P is usually in the .80 - .99 range, the appropriate criterion for goodness of fit is some measure of how well the model fits the right tail of the empirical demand distribution. Judged on this basis, there are several situations where the Normal model fitted on the usual sample mean, sample variance basis might be a poor choice. In particular, skewed distributions and bimodal distributions will cause difficulties.

Silver and Peterson (1974) have suggested that the double exponential or Laplacian distribution can be used as an alternative to the Normal model. This distribution is found to compare well with the Normal in its performance in inventory control situations and has the added advantage of possessing a closed form expression for its cumulative distribution function. The latter property simplifies computation of fractiles. However, this distribution is also a symmetric, two-parameter distribution which could suffer from similar difficulties when information about the location of the distribution is misleading with respect to tail probabilities. We shall also see that this method can be regarded as a particular case of the approach described here.

A New Approach

Intuitively, the idea behind the method suggested here is to concentrate on estimating the relevant part of the distribution by utilizing the available information on right-tail probabilities. This approach seeks to avoid the pitfall of fitting a model using information about the "center" of the distribution and then making inferences about tail probabilities. A probabilistic model reflecting this idea is presented below. It is by no means the only way of implementing the idea but is certainly one of the simplest. For some fixed x_0 let the distribution of demand over replenishment lead time be described by the density function (shown in Figure 1)

$$f(x) = \begin{cases} p \cdot g(x) & \text{for } 0 \leq x \leq x_0 \\ (1 - p)\lambda e^{-\lambda(x - x_0)} & x > x_0 \end{cases}$$

Figure 1

The Mixture Model

where the $f(x)$ to be a proper density it is required that $\int_{x_0}^{\infty} g(x)dx = 1$. That is to say, the density is a mixture of two distributions, one concentrated on $[0, x_0]$ and the other a translated exponential concentrated on (x_0, ∞) . For our purposes we do not need to specify the form of $g(x)$ so that the parameters of interest are just p and λ . The idea here is that even if the "true" demand distribution is a difficult one to model, the right tail of the density function will often be relatively smooth. The model thus tries to approximate right tail probabilities with a simple functional form while ignoring the general shape of the distribution which is irrelevant in this particular decision problem and may even be misleading.

An immediate question arises about the choice of x_0 . Treating this as a parameter is analytically awkward. In theory any point could be designated as x_0 *a priori*, but in order for the method to be effective in practice, we would like x_0 to be around the 0.8 fractile. This issue is discussed further in later sections.

The probabilistic model is used to derive point estimators of fractile locations using Bayesian methods. Eventually heuristic smoothing methods are to be developed for which the exact Bayesian estimate does not appear to be suitable. Instead a so-called "approximate Bayesian" estimator is developed on which the smoothing method is based. The effectiveness of the methods for a variety of distributions is tested using simulation and analytic methods. In the following, we first describe the estimators qualitatively, and discuss their performance. Analytical derivations are presented in a later section.

Fractile Estimation with Stationary Demand

Suppose that the fractile to be determined is some P which is greater than p . Given the parameters x_0 , p and λ of the mixture model, the value of S , such that $\text{prob}[x \leq S] = P$ is given by

$$S = x_0 + \frac{1}{\lambda} \ln \left[\frac{1 - p}{1 - P} \right] \quad (1)$$

In practice, the parameters must be estimated as demand data becomes available. Initially, we assume that the demand distribution is unchanging (stationary). Data collection is based on maintaining statistics on

- n: the total number of data points
- m: the number of points for which demand is less than or equal to x_0
- r: the number of points for which demand exceeds x_0
- t: the average (conditional) demand when demand exceeds x_0

For the stationary case, a Bayesian analysis of the parameters p and λ can be conducted to give an exact estimator S_1 , based on the unconditional distribution of demand.

$$S_1 = x_0 + t \left[\left(\frac{1 - \bar{p}}{1 - P} \right)^{1/r} - 1 \right]; \quad \bar{p} = \frac{m}{n} \quad (2)$$

An approximate estimator S_2 for the stationary case can also be developed by first making estimates of p and λ and then using these in equation (1).

$$S_2 = x_0 + \left(\frac{1}{\bar{\lambda}} \right) \ln \left(\frac{1 - \bar{p}}{1 - P} \right); \quad \bar{\lambda} = \frac{r}{t} \quad (3)$$

The derivation of these estimators is given in a subsequent section. It can be shown that it is always the case that $S_2 \leq S_1$, though the estimators coincide in the limit.

When these estimators are applied to an arbitrary distribution, they are in general biased. In the case when the underlying demand distribution is Normal, the bias can be determined *a priori* for a given fractile level. For example, when $P = 0.95$, it can be shown that the mixture estimator underestimates the correct value by 0.029σ , where σ is the standard deviation of the Normal distribution.

The validity of the method for other arbitrary distributions is difficult to establish theoretically. Simulation is therefore used to examine the behavior of the estimators under stationary conditions. Four distributions were used to generate data:

- (a) Exponential; $\lambda = 1$
- (b) Normal; $\mu = 5, \sigma = 1$
- (c) Erlang; $\lambda = 1, r = 5$
- (d) NSZ (Normal with a spike at zero); a mixture of the value 0 with probability $\pi = 0.2$ and a Normal component ($\mu = 5, \sigma = 1$) with probability $1 - \pi = 0.8$.

The base value of x_0 for the mixture estimates was set to the theoretical 0.8 fractile in each case and left unadjusted throughout the run. The "target fractile" to be located was 0.95. The results of the test are summarized in Table 1. The sample sizes were between 300 and 500, although in each case the estimates stabilized fairly quickly (well within 100 observations). It should be noted that except in the case of the Normal data generating process, the theoretical 0.95 fractile location is not exactly the same as the empirical fractile location for the data actually generated; however, it is listed in the table so as to sidestep the computational burden of finding the latter value.

In the case of the Normal process, the Normal estimates are closer than the mixture estimates, with the variance based Normal estimate performing better than MAD as is expected. The mixture estimates are slightly low as was suggested in the previous section, but they appear to be quite satisfactory. In each of the other cases, the mixture model outperforms the Normal which underestimates for the positively skewed Exponential and Erlang distributions. In the case of the NSZ distribution the Normal estimators overestimate possible due to the high variance of the process. Parenthetically it is noted that the "approximate" mixture estimate is in all cases lower than the exact, as expected.

Heuristic Methods

In practice the estimation of the unknown parameters may be performed by smoothing methods. This allows for possible nonstationarity of the data-generating process in a heuristic manner. In any period n , given a new data point x_n , the smoothed estimates of p and λ^{-1} are given by

$$\hat{p}_n = \begin{cases} \alpha + (1 - \alpha) \hat{p}_{n-1} & \text{if } x_n \leq x_0 \\ (1 - \alpha) \hat{p}_{n-1} & \text{if } x_n > x_0 \end{cases}$$

TABLE 1**COMPARISON OF ESTIMATES OF 0.95 FRACTILE FOR LARGE SAMPLE SIZES**

Estimator Distribution	Theoretical	Normal (MAD)	Normal (Variance)	"Exact" Mixture	Approx. Mixture
Normal	6.776	6.779	6.776	6.724	6.718
Exponential	2.9	2.54	2.382	2.943	2.932
Erlang	9.15	8.618	8.432	8.928	8.913
NSZ	6.535	7.586	7.514	6.560	6.554

$$\hat{\lambda}_n^{-1} = \begin{cases} \hat{\lambda}_{(n-1)}^{-1} & \text{if } x_n \leq x_0; \\ \beta(x_n - x_0) + (1 - \beta) \hat{\lambda}_{(n-1)}^{-1} & \text{if } x_n > x_0 \end{cases} \quad (4)$$

and where the constants α and β are in the range $(0, 1)$. These smoothed parameter estimates are then used in the "approximate" estimator derived in a subsequent section, and given in equation (3).

The question of initialization of the estimation procedure is quite difficult. A choice of x_0 has to be made and starting values of $\hat{p}_0$ and $\hat{\lambda}_0$ are to be set correspondingly. If a high value of x_0 is selected so that the values of $\bar{p}$ tend to be high, the updating procedure will be slow with respect to λ since very few of the data points obtained will be informative about the tail probabilities beyond x_0 . On the other hand, for x_0 to be chosen relatively small—towards the median of the distribution—is self-defeating since the point of the method is to concentrate attention on the tail of the distribution.

A reasonable starting point may be to set x_0 at about the 0.8 fractile as estimated from initially available data about the demand. The estimate $\hat{p}_0$ is set to 0.8, and $\hat{\lambda}_0$ must be estimated from the data available. If the data are insufficient, they can be fleshed out by fitting say a Normal or some other appropriate distribution to them and then using this model to initialize x_0 , $\hat{p}_0$ and $\hat{\lambda}_0$.

If it subsequently appears that x_0 has been chosen too high or too low—that is, if the estimate of p becomes too high or too low—the value of x_0 can be adjusted heuristically by using a translated exponential model to obtain a new value as follows: At any stage let the estimated value of p be $\hat{p}$ and suppose that the desired location of x_0 is the fractile p' . Then we wish to find x'_0 such that

$$(1 - \hat{p}) \int_{x_0}^{x'} \hat{\lambda} e^{-\lambda(x - x_0)} dx = p' - \hat{p}$$

Simplification yields

$$x'_0 = x_0 - (1/\hat{\lambda}) \ln \left[\frac{1 - p'}{1 - \hat{p}} \right]$$

and the current estimate of λ is left unchanged.

Finally, consider a direct procedure that may be regarded as a special case of these methods. Suppose we choose to set x_0 at a point such that $p = P$. In this case at every stage we set $x'_0 = S$ where S is the calculated target stock level and we set $\hat{p} = P$. We shall refer to this method as a "tracking" approach since we are attempting to track the optimal stock level directly.

The discussion in the previous section indicates that the accuracy of the suggested method in the limiting case is quite satisfactory. However, the behavior of the smoothed estimate in stationary and nonstationary situations remains to be determined and is of greater interest from the point of view of applications such as inventory control. Simulation appears to be the only practical way of studying this behavior.

Simulation Tests

The same four distributions described earlier were used to generate data for testing the heuristic estimators. The estimators tested were:

- (i) Normal with smoothed MAD.
- (ii) Normal with smoothed variance estimates
- (iii) Approximate mixture with smoothed $\hat{p}$ and $\hat{\lambda}$.
- (iv) Tracking estimate.

For the mixture estimate (iii), the base value of x_0 was initialized in each run on the basis of a starting sample of five observations. Thereafter it was adjusted heuristically whenever $\hat{p}$ differed from 0.8 by a pre-specified tolerance level (0.1) in these sets). All smoothing constants were set at 0.01 as trial and error suggested that this gave good results.

In this situation, since the estimates are smoothing heuristics based on incomplete information (finite memory), they do not converge. The criterion of performance should be their stability relative to the data observed and how well they perform in actually locating the target fractile. In the spirit of the inventory problem motivating this investigation, the criterion chosen was the total of excess and shortage costs were chosen so that the ratio of unit shortage cost to unit excess plus shortage was equal to the target fractile. The fractile was set at 0.95.

Two sets of runs were made to test behavior in the stationary and nonstationary cases. In the first instance, 500 sample observations were generated after initialization of the statistics and the costs incurred were summed. The results are reported in Table 2 in terms of ratios of the costs incurred against the cost incurred by the Normal (MAD) estimator.

In the final set, the same smoothing methods were tested against a step change in each of the same four data generating processes. This test was made to gain some indication of the behavior of the estimators in nonstationary situations. The estimates were allowed to stabilize on stationary data; the parameters of the process were then given a step change and the costs incurred were measured over 500 observations. The parameter changes in each of the cases were:

- (i) Exponential, parameter λ ; 1.0 to 0.5.
- (ii) Normal, parameters (μ, σ) ; (5, 1) to (6, 2).
- (iii) Erlang, parameters (λ, r) ; (1, 4) to (0.5, 5).
- (iv) NSZ, parameters (π, μ, σ) ; (0.2, 4, 1) to (0.1, 5, 1).

The results are reported in Table 3, again in terms of cost ratios relative to Normal (MAD).

In both situations, the mixture estimates outperform the Normal estimators in locating the fractile, except for the nonstationary Normal case, where the tracking estimator is poorer than the Normal (MAD) estimate. The smoothed mixture estimator is dominant. The Normal estimator based on smoothed variance statistics generally performed worse than the usual MAD based estimator. This result was something of a surprise. It appears that the use of the MAD statistic gives the Normal model a measure of robustness in dealing with non-Normal data.

TABLE 2

**STATIONARY DEMAND PROCESS: COST RATIOS
AGAINST NORMAL (MAD) ESTIMATOR**

Distribution	Estimation Method	$\left(\frac{\text{Cost of using Estimator}}{\text{Cost with Normal (MAD) Estimator}} \right)$
Exponential	Smoothing	0.924
	Tracking	0.938
	Normal (variance)	1.083
Normal	Smoothing	0.965
	Tracking	0.972
	Normal (variance)	1.014
Erlang	Smoothing	0.927
	Tracking	0.959
	Normal (variance)	1.091
NSZ	Smoothing	0.870
	Tracking	0.878
	Normal (variance)	0.877

TABLE 3

**NONSTATIONARY DEMAND PROCESS: COST RATIOS
AGAINST NORMAL (MAD) ESTIMATOR**

Distribution	Estimation Method	$\left(\frac{\text{Cost of using Estimator}}{\text{Cost with Normal (MAD) Estimator}} \right)$
Exponential	Smoothing	0.917
	Tracking	0.939
	Normal (variance)	1.091
Normal	Smoothing	0.975
	Tracking	1.050
	Normal (variance)	1.098
Erlang	Smoothing	0.886
	Tracking	0.970
	Normal (variance)	1.089
NSZ	Smoothing	0.946
	Tracking	0.982
	Normal (variance)	0.985

A final note: It appeared that the mixture estimators tended to underestimate the target fractile location. Indeed, since the estimator is based on the "approximate" approach, this might have been foreseen. A heuristic adjustment which inflated the estimates slightly resulted in significantly better performance in a preliminary test. This direction deserves more study.

Application to Inventory Control

The methods discussed here are particularly well suited for base-stock inventory systems and for lead time management in purchasing. First considered a base-stock system for inventory control (Kimball, 1988). Here, each stocking location has a base-stock level set for each item. The base stock is chosen so that the total on-hand plus on-order inventory is at a particular fractile of the distribution of demand over the resupply lead time (for example, see Johnson and Montgomery, 1974). The tracking method described earlier can be applied to this problem as follows. The statistics maintained are the proportion of total of demand backordered ($\hat{p}$) and the average of units on backorder ($\hat{\lambda}^{-1}$). The base-stock level is tracked directly and orders are placed on the supplier whenever on-hand plus on-order inventories drop below the base stock. Note that neither the distribution of demand nor resupply lead time need to be monitored continuously.

The smoothing method of the previous section is particularly well suited to the management of purchasing lead times. Here the probability distribution is not of demand but of the supply lead time. The agreed upon delivery date provides a natural choice of x_0 . The planned date of use of the material is set at some time after the delivery date. The use of such active buffers is common in practice. The statistics collected are on-time performance measurement as the fraction of orders arriving before the delivery date ($\hat{p}$) and the average number of days late on late deliveries ($\hat{\lambda}^{-1}$). The safety allowance between the delivery date and the planned use date is adjusted for each new order placed so as to leave the use date at the target fractile value. A version of this method is presently being used by a large equipment manufacturer for the management of lead times on purchased parts. The combined decision of setting safety times and delivery frequency is discussed by Karmarkar and Rummel (1989).

In both of the cases described above, note that the statistics to be collected are natural for the application. The forecasting technique is thus closely integrated with the management process so that software development and implementation are simpler. The technique is also very appealing to management, and can be described without emphasizing the underlying statistical issues.

Analytical Derivations

In this section, we describe the estimation of the desired fractile level S when the parameters p and λ of the mixture are not known with certainty. Exact and approximate Bayesian estimators are derived. The behavior of the estimator, when the underlying distribution is actually Normal, is discussed.

First, suppose that the parameters p and λ for the mixture distribution are known with certainty, and that the optimal fractile value P is such that $p \leq P$. We must choose S such that

$$F(S) = \int_0^{x_0} f(x)dx + \int_{x_0}^S f(x)dx = P.$$

Whence on simplification:

$$S = x_0 + \frac{1}{\lambda} n \left[\frac{1 - p}{1 - P} \right]. \quad (1)$$

Exact Bayesian. Since the two components of the mixture model are concentrated on disjoint intervals, the Bayesian inference problem becomes very simple. We can essentially treat the parameters p and λ independently. The details of the updating procedure are well known and the method is only briefly sketched below using the terminology of Raiffa and Schlaifer (1968).

Suppose that the prior state of knowledge about p is described by a Beta distribution $f_\beta(p|m', n')$. Given a sequence of observations of the demand $x_1, x_2, \dots, x_n$ generated according to the mixture model, we define the imbedded Bernoulli process by the sequence $y_1, y_2, \dots, y_n$ where

$$y_i = \begin{cases} 1 & \text{for } x_i \leq x_0 \\ 0 & \text{for } x_i > x_0 \end{cases}$$

Sufficient statistics for the sample are (m, n) where $m = \sum_{i=1}^n y_i$, for the imbedded process. The

posterior distribution obtained for p in the usual way is also Beta with parameters $m'' = m' + m$ and $n'' = n' + n$. Suppose λ to have a Gamma -1 prior distribution $f_{\gamma 1}(\lambda | r', t')$. To update this distribution we need only consider the subset of observations that lie above x_0 since the remaining observations are not generated according to the translated exponential component of the distribution. Here, the sufficient statistics for the sample are (r, t) where

$$r = \sum_{i=1}^n (1 - y_i)$$

$$t = \sum_{i=1}^n (x_i - x_0) (1 - y_i)$$

Again the posterior density for λ is also the form Gamma -1 with parameters $r'' = r' + r$ and $t'' = t' + t$. Hereafter the double primes on parameters and the detailed parametric notation for densities will be dropped for convenience.

Armed with these updated distributions for the uncertain parameters, we can proceed to find the distribution of demand unconditional on p and λ and then find the order point as the P 'th fractile of this unconditional distribution. Or proceeding more directly, we have

$$\int_0^1 \int_0^{x_0} p \cdot g(x) \cdot f_{\beta}(p) dx dp$$

$$+ \int_0^{\infty} \int_0^1 \int_{x_0}^S (1-p)\lambda e^{-\lambda(x-x_0)} \cdot f_{\beta}(p) \cdot f_{\gamma_1}(\lambda) dx dp d\lambda = P$$

which gives

$$\left[\frac{t}{t + S - x_0} \right]^r = \frac{1 - P}{1 - \bar{p}}; \quad \bar{p} = \frac{m}{n}$$

where r and t represent the updated parameters of the posterior Gamma -1 distribution. Hence we get the estimate

$$S_1 = x_0 + t \left[\left(\frac{1 - \bar{p}}{1 - P} \right)^{1/r} - 1 \right] . \quad (2)$$

Approximate Bayesian. The approach here is to use the updated distributions of p and λ to make point estimates of these quantities and to then use these point estimates as though p and λ were known for certain. If we use the mean as an estimator for both parameters, we have $\bar{p} = m/n$ and $\bar{\lambda} = r/t$. Using these estimates in equation (1):

$$S_2 = x_0 + t \ln \left[\frac{1 - \bar{p}}{1 - P} \right]^{1/r} - 1 . \quad (3)$$

Defining the statistics

$$Y = \left[\frac{1 - \bar{p}}{1 - P} \right]^{1/r}; \quad Z = Y - 1$$

we have $Y \geq 1, Z \geq 0$ for $\bar{p} \leq P$ and $r \geq 1$. We can rewrite the estimates as

$$S_1 = x_0 + tZ$$

$$S_2 = x_0 + t \ln Y = x_0 + t \ln(1 + Z) .$$

Since $Z \geq \ln(1 + Z)$, note that we always have $S_2 \leq S_1$. As the number of observations n increases, r increases approximately as $(1 - p)n$ and hence Z decreases so that agreement between S_1 and S_2 improves, and the estimates coincide in the limit.

Behavior When Demand is Normal

We can conveniently make some observations about the effectiveness of the model when the underlying demand distribution is Normal. Suppose that we require $P = 0.95$ corresponding to a 95% service level, and that x_0 is chosen such that $\Pr[x \leq x_0] = 0.8$; i.e., $p = 0.8$. We also assume that the demand x is distributed according to $N(\mu, \sigma)$ and that p and λ are known with certainty. In the limiting case, given the suggested method of estimation, the parameter λ is given by

$$\lambda^{-1} = E_{x|x > x_0} (x - x_0) = \frac{1}{(1 - p)} \int_{x_0}^{\infty} (x - x_0) \cdot f_N(x|\mu, \sigma) dx.$$

To calculate λ we can use the formula for the right partial expectation of the Normal distribution:

$$E_N^{(r)}(x_0|\mu, \sigma) = \int_{x_0}^{\infty} x \cdot f_N(x|\mu, \sigma) dx = \mu G_N^*(z_0) + \sigma f_N^*(z_0)$$

where the standard Normal variate is defined as $z = (x - \mu)/\sigma$ and G_N^* and f_N^* are the right cumulative distribution function and the density function, respectively, for the standard Normal distribution. We thus have:

$$\lambda^{-1} = \frac{1}{(1 - p)} E_N^{(r)}(x_0|\mu, \sigma) = x_0 \quad . \quad (8)$$

Given that $p = \Pr[x \leq x_0] = 0.8$, we have from tables of the standard Normal distribution that (approximately)

$$z_0 = 0.84 ; G_{N^*}(z_0) = 0.2 ; f_{N^*}(z_0) = 0.28$$

$$x_0 = \mu + 0.84\sigma ; E_N^{(r)}(x_0 | \mu, \sigma) = 0.2\mu + 0.28\sigma .$$

Substituting in equation (8) gives $\lambda^{-1} = 0.56\sigma$ and using the formula derived in equation (1) for the perfect information case, we can compute the estimate $\hat{S}_m$ for S defined by $\Pr[x \leq S] = 0.95 = P$.

$$\hat{S}_m = \mu + 1.616\sigma .$$

From tables it can be determined that the correct value of S is given approximately by

$$S = \mu + 1.645\sigma$$

and the error in the estimate is thus

$$S - \hat{S}_m = 0.029\sigma$$

which depends only on the standard deviation of the Normal distribution as might have been expected.

Summary

This paper has presented a technique for fractile estimation which is particularly appropriate for inventory control applications. The approach proceeds from a Bayesian model and develops computationally realistic and robust heuristic methods for unsupervised estimation of stocking levels and lead times. The statistics to be maintained are natural from a managerial point of view.

The quality of the estimates obtained appears to be superior to that of the usual methods based on a Normal model. It is thought that further research in tuning the heuristic methods will result in improved performance.

REFERENCES

- Johnson, L. A. and D. C. Montgomery (1974), *Operations Research in Production Planning, Scheduling, and Inventory Control*, J. Wiley and Sons.
- Karmarkar, U. S. and J. Rummel (1989), "Delivery Management: Delivery Frequency and Safety Times", CMOM working paper.
- Kimball, G. (1988), "General Principles of Inventory Control", *Journal of Manufacturing and Operations Management*, Vol. 1, Issue 1, pp. 119-130.
- Raiffa, H. and R. Schlaifer (1968), *Applied Statistical Decision Theory*, The M.I.T. Press, Cambridge, MA,
- Silver, E. A. and R. Peterson (1974), "Implementation Considerations in Selecting the Probability Distribution of Demand in a Replenishment Lead Time", presented at the TIMS/ORSA Joint National Meeting, Boston, MA.