NPS 55Ze 9041A

UNITED STATES NAVAL POSTGRADUATE SCHOOL

FORECASTING ERRORS USING MAD

by

Peter W. Zehna

April 1969

Distribution of this Document is Unlimited

32

UNITED STATES NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral Robert W. McNitt, USN
Superintendent

Dr. R. F. Rinehart
Academic Dean

ABSTRACT:

A study of the effects of using mean absolute deviation (MAD) to estimate variability in setting reorder levels for the inventory of a stock item. The method presently employed by NavSup in setting such reorder levels involves exponentially smoothed estimates of the mean and variance of the demand process. Any error involved in setting reorder levels results in a change in the underlying risk which in turn can be translated into costs. Such errors for the method of estimation presently employed are compared with standard maximum likelihood procedures. By simulating several normal systems, the smoothing technique is found to be inferior to classical methods with no reduction in computational difficulties.

This task was supported by the Research and Development Division, Naval Supply Systems Command under NAVSUP RDT&E task area No. TF 015-02-101.

Prepared by:

Peter W. Zehna
Professor

Approved by:

J. R. Borsting
Chairman, Department of
Operations Analysis

Released by:

C. E. Menneken
Dean of
Research Administration

United States Naval Postgraduate School Report #NPS55Ze9041A

UNCLASSIFIED

TABLE OF CONTENTS

1. INTRODUCTION AND SUMMARY

In a previous report [1], the use of MAD (mean absolute deviation) in the analysis of demand variability as proposed by R. G. Brown [3] and presently used by NavSup [4] was discussed. In particular, the use of exponentially smoothed estimates of MAD was criticized on theoretical grounds and a case was made for further study into the probability distributions of the various estimators presently in use. Of special importance, the use of MAD in setting safety levels was mentioned as a problem area in need of further investigation.

The present report concerns itself with the extension of research investigations in the areas just mentioned. Several attempts to derive the necessary probability distribution theory for MAD estimators resulted in disappointing futility for the most part. This was anticipated for, as previously mentioned in [1], absolute deviation as a measure of variability has been abandoned for years on theoretical grounds by Statisticians.

Two minor results that were obtained concern the ratio of MAD, Δ, to standard deviation, σ. The complete result of this ratio has been determined for the Poisson case. As alluded to in [1], the universal use of $\sqrt{\frac{2}{\pi}} \doteq 0.8$ for this ratio is especially bad for low demand Poisson item. The same general criticism can be made in the negative binomial case. However, it was found that for the choice of parameters presently being used by NavSup, the approximation of this ratio by 0.8, is a good one.

Failing to obtain the required distribution theory to study the behavior of MAD estimates, simulation was used to compare smoothed estimates of σ to the usual maximum likelihood estimates mostly for the case of normal demand. Parameter choices that reflect high demand items were selected in order to give MAD the best possible advantage in the comparison with alternate methods of estimation. Such parameter choices were made after consultation with NavSup personnel in order to duplicate situations that actually exist; moreover, large amounts of real data are available for such cases. In every single case studied, the sample variance of the smoothed estimate of MAD and consequent estimate of σ was roughly twice that of the maximum likelihood estimates of σ. The same was true in a slow mover case or two that was examined. This confirms in a general way the observation made in a similar study carried out by Asher and Wallace [2] in which they found MAD to be significantly less efficient (about 20%) compared to the classical minimum variance estimators under the Gauss-Markov assumptions.

Not content merely to confirm this observation, however, the estimates found were used further to see what the actual effects on reorder levels were. This was approached in several ways. First, MAD and maximum likelihood were compared by computing the percentage of time the known theoretical reorder level was overestimated, thereby resulting in too much on hand stock. Secondly, the results were examined to see what percentage of the observations fell within k units of the theoretical order level for $k = 1,2,\cdots$ and various fixed values of risk. Finally, the two methods were compared by setting reorder levels and then computing the actual risk attained by those levels where, for each case again, the theoretical or true risk is known. As seen in Section 4, exponentially smoothed mean and MAD came off second best compared with maximum likelihood estimates of μ and σ in every single case examined.

Recognizing the limitations of simulation, no sweeping claim is here made for proof that exponentially smoothed MAD is an inefficient method of accounting for demand variability. At the same time, one cannot ignore the fact that present methods were uniformly inferior in the situations examined.

In Section 5 recommendations for further study are made. Among these is the suggestion that real data be used from the histories available at FMSO for comparing the results in retrospect with what would have been the case had maximum likelihood procedures been used. The writer wishes to thank Lt. Özden Orneck for his invaluable assistance in constructing and running the computer programs as well as assistance in mathematical derivations. Acknowledgement should also be given to Cdr. Jack E. White of FMSO for his unfailing cooperation in defining the problem areas and supplying parameter values that are realistic in terms of NavSup use. Credit should also be given to Mr. James W. Prichard (SUP 04E) for his continued endorsement and interest in this research area.

2. RATIO OF Δ TO σ

It will be helpful to review some parametric definitions and establish a notation to be used here and in ensuing sections. Let X be a random variable with mean μ and variance σ^2 (standard deviation σ). The mean absolute deviation, MAD for short, will be denoted Δ

and is defined as

$$(2.1) \qquad \Delta = E(|X-\mu|)$$

As remarked in [1], it would be more rational to define MAD as $E(|X-m|)$ where m is the median of X but it is (2.1) that is used by Brown [3] and NavSup [4]. Consequently the same definition is adopted in this report.

Now when X is normally distributed, it is well known that the ratio of Δ to σ is given by $\sqrt{\frac{2}{\pi}}$ or, roughly 0.8. It is somewhat surprising that the ratio is approximately the same for certain other families of probability distributions such as the Exponential, Uniform and Triangular. However, this result is not universally true and is particularly a poor approximation for the Poisson family, a model often used for the so-called slow mover type of inventory item. This fact was demonstrated earlier in [1] by examining the ratio for selected values of the Poisson parameter, λ. That ratio has now been determined for all values of λ and it may be instructive to see the behavior of this ratio in a complete sense.

By $[z]$, we shall mean the greatest integer in z, that is, the largest integer n such that $n \leq z$ (and hence $z < n + 1$). We denote the Poisson (λ) mass function by p so that,

$$p(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \cdots$$

Then we have, recalling that $\mu = \sigma^2 = \lambda$ for this case,

$$\Delta = E(|X-\mu|) = \sum_{x=0}^{\infty} |x-\lambda| p(x) = \sum_{x=0}^{[\lambda]} (\lambda-x) p(x) + \sum_{x=[\lambda]+1}^{\infty} (x-\lambda) p(x)$$

But, letting $F(z) = \sum_{x=0}^{z} p(x)$ define the Poisson distribution function,

$$\sum_{x=0}^{[\lambda]} (\lambda-x)p(x) = \lambda F([\lambda]) - \sum_{x=0}^{[\lambda]} xp(x) = \lambda F([\lambda] + \sum_{x=[\lambda]+1}^{\infty} xp(x) - \lambda$$

$$\sum_{x=[\lambda]+1}^{\infty} (x-\lambda)p(x) = \sum_{x=[\lambda]+1}^{\infty} xp(x) - \lambda + \lambda F([\lambda])$$

Adding these results,

$$(2.2) \quad \Delta = 2\lambda F([\lambda]) - 2\lambda + 2 \sum_{x=[\lambda]+1}^{\infty} xp(x) .$$

But, for $x \geq 1$,

$$xp(x) = xe^{-\lambda}\frac{\lambda^x}{x!} = \lambda e^{-\lambda}\frac{\lambda^{x-1}}{(x-1)!} = \lambda p(x-1) .$$

Hence,

$$\sum_{x=[\lambda]+1}^{\infty} xp(x) = \lambda \sum_{x=[\lambda]+1}^{\infty} p(x-1) = \lambda \sum_{y=[\lambda]}^{\infty} p(y) = \lambda(1-F([\lambda]-1)).$$

Substituting this result in (2.2),

$$\Delta = 2\lambda F([\lambda]) - 2\lambda + 2\lambda - 2\lambda F([\lambda]-1)$$

$$= 2\lambda(F[\lambda]-F([\lambda]-1))$$

$$= 2\lambda\, p([\lambda])$$

Thus,

$$(2.3) \qquad \Delta = 2\lambda e^{-\lambda}\frac{\lambda^{[\lambda]}}{[\lambda]!} \quad \text{for any} \quad \lambda > 0.$$

Since $\sigma = \sqrt{\lambda}$ we have,

$$(2.4) \qquad \frac{\Delta}{\sigma} = 2\sqrt{\lambda}\, e^{-\lambda}\frac{\lambda^{[\lambda]}}{[\lambda]!}$$

The graph of $\frac{\Delta}{\sigma}$ as a function of λ appears in Figure 1.

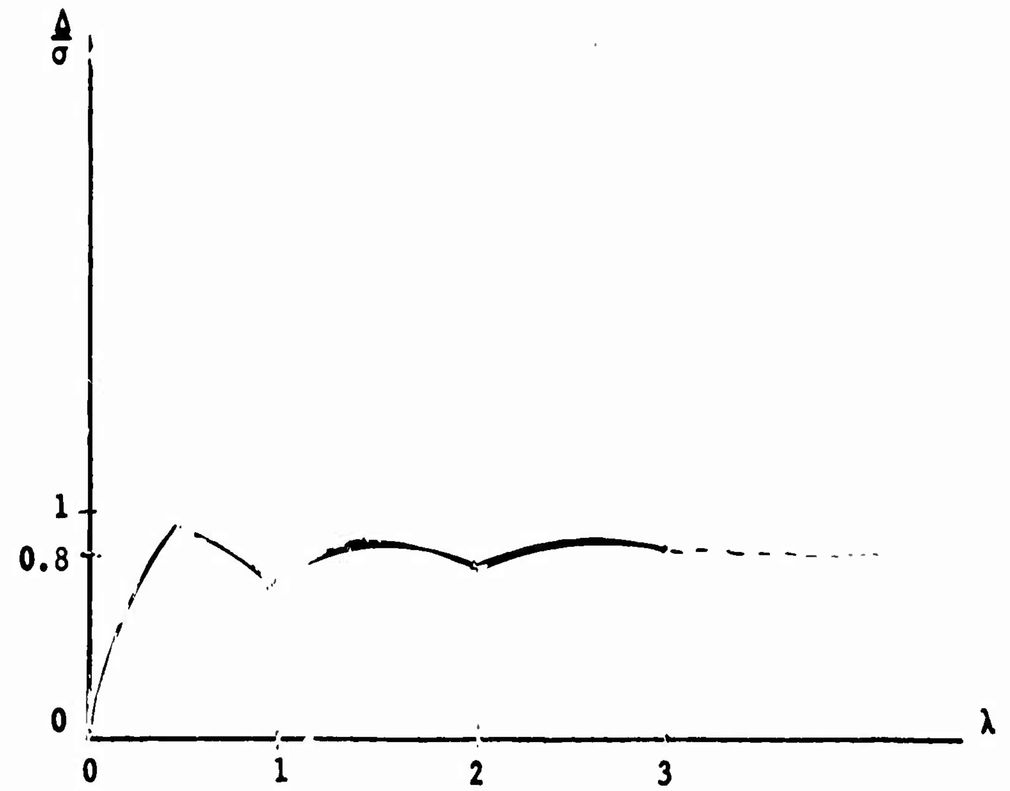

Figure 1. Graph of $\frac{\Delta}{\sigma}$

The analysis that validates the graph is the following.

If $n < \lambda < n + 1$ then $[\lambda] = n$ and, $\frac{\Delta}{\sigma} = \frac{2}{n!} \lambda^{n+\frac{1}{2}} e^{-\lambda}$.

As a function, say f_n, of λ in this open interval, f_n is differentiable with,

$$f_n'(\lambda) = \frac{2}{n!} e^{-\lambda} \lambda^{n-\frac{1}{2}} (n+\tfrac{1}{2}-\lambda)$$

Hence, $f_n'(\lambda) > 0$ if $\lambda < n + \frac{1}{2}$ and $f_n'(\lambda) < 0$ if $\lambda > n + \frac{1}{2}$ so that f_n possesses a relative maximum at $\lambda = n + \frac{1}{2}$ with a maximum value of $f_n(n+\frac{1}{2})$, for $n = 0,1,2,\cdots$.

With the formulas established for f_n and f_{n+1} it is easy to verify that $\frac{\Delta}{\sigma}$ is continuous at each of the integers although not differentiable there. In any case, it is shown in the appendix that the sequence of values of $\frac{\Delta}{\sigma}$ at n,

$$\frac{2}{n!} n^{n+\frac{1}{2}} e^{-n}$$

is monotone increasing. Using Stirlings approximation [7] to $n!$,

$$\sqrt{2\pi}\, n^{n+1}\, e^{-n} < n! < \sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n}\left(1+\frac{1}{12n-1}\right)$$

we see that,

$$\sqrt{\frac{2}{\pi}}\, \frac{1}{1+\frac{1}{12n-1}} < f_n(n) < \sqrt{\frac{2}{\pi}}$$

and $\lim_{n\to\infty} f_n(n) = \sqrt{\frac{2}{\pi}}$. Thus the values of $\frac{\Delta}{\sigma}$ converge in the limit to the factor $\sqrt{\frac{2}{\pi}}$. However, as previously remarked in [1] and as seen visually in Figure 1, approximating $\frac{\Delta}{\sigma}$ by $\sqrt{\frac{2}{\pi}}$ is particularly bad when $\lambda < 1$ which is precisely when the Poisson assumption is of primary interest.

For intermediate movers, the negative binomial distribution is employed as a model for demand. Specifically, from [4], such an assumption is made whenever mean demand is between 2 and 20 units.

The negative binomial distribution is a family of probability mass functions having two parameters r and p where r is a positive integer and $0 < p < 1$. For our purposes the general formula for the mass function is defined by,

$$p(x) = \binom{r+x-1}{x} p^r q^x, \quad x = 0,1,2,\cdots \quad (q=1-p)$$

The mean μ and the variance σ^2 of a random variable having this mass function are given by,

$$\mu = \frac{rq}{p} \quad , \quad \sigma^2 = \frac{rq}{p^2}$$

Proceeding as in the Poisson case,

$$\Delta = E(|X-\mu|) = \sum_{x=0}^{\infty} |x-\mu| p(x) = \sum_{x=0}^{[\mu]} (\mu-x) p(x) + \sum_{x=[\mu]+1}^{\infty} (x-\mu) p(x)$$

But,

$$\sum_{x=0}^{[\mu]} (\mu-x) \binom{r+x-1}{x} p^r q^x = \mu p^r + \sum_{x=1}^{[\mu]} (\mu-x) \binom{r+x-1}{x} p^r q^x$$

and,

$$\sum_{x=[\mu]+1}^{\infty} (x-\mu) \binom{r+x-1}{x} p^r q^x = -\sum_{x=1}^{[\mu]} x \binom{r+x-1}{x} p^r q^x + \mu \sum_{x=0}^{[\mu]} \binom{r+x-1}{x} p^r q^x$$

Adding these results,

$$(2.5) \qquad \Delta = 2[\mu p^r + \mu \sum_{x=1}^{[\mu]} p(x) - \sum_{x=1}^{[\mu]} x p(x)]$$

expressing Δ in terms of finite sums.

Since $\sigma = \frac{\sqrt{rq}}{p}$, it is no problem to compute the ratio $\frac{\Delta}{\sigma}$ for various choices of r and p on the computer.

It is easy to choose values of r and p for which the ratio is very different from 0.8. However, the interest would only be academic for values of r and p such that $\mu \notin [2,20]$. For that range of values, a total of 1,000 cases were considered as follows. Let $\mu \in [2,20]$ be selected and for each such μ, r was allowed to vary over the set $\{2,4,6,\cdots,20\}$. For each such choice, p is then determined from the relation $p = \frac{r}{r+\mu}$ and q is specified by

$q = 1 - p$. Then Δ was computed using (2.5) and finally $\frac{\Delta}{\sigma}$ was evaluated. In each such "group" of ten parameter pairs, it was observed that $\frac{\Delta}{\sigma}$ was invariably monotone increasing although there was no monoticity from group to group. More significantly however, it was observed that $\frac{\Delta}{\sigma}$ varied always between a minimum value of 0.750 and a maximum value of 0.805. For example, if $p = 0.893$ and $r = 20$ whence $\mu = 2.4$, we find that $\frac{\Delta}{\sigma} = 0.805$ although for $p = 0.455$ and $r = 2$, resulting also in $\mu = 2.4$, $\frac{\Delta}{\sigma} = 0.770$. On the other hand, for $p = 0.174$ and $r = 4$, yielding $\mu = 19$, $\frac{\Delta}{\sigma} = 0.781$ so that low values of $\frac{\Delta}{\sigma}$ are not always associated with low values of μ.

In summary, then, we may say that for the negative binomial case and the range of parameter values presently used by NavSup, the approximation $\sqrt{\frac{2}{\pi}}$ for $\frac{\Delta}{\sigma}$ appears to be a safe one.

3. ESTIMATING VARIABILITY WITH MAD

Having established the value of $\frac{\Delta}{\sigma}$ for cases of particular interest in the inventory models, it follows that $\sigma = k\,\Delta$ for some constant k. Consequently, if Δ is estimated directly from the data to obtain an estimate $\tilde{\Delta}$, say, then σ can be estimated by applying the formula $\tilde{\sigma} = k\,\tilde{\Delta}$. In this way, one obtains a measure of the variability of demand through estimates of Δ. How good such a procedure might be clearly depends upon the technique used to estimate Δ in the first place and what the effect is on scaling that estimate by the factor k to obtain $\tilde{\sigma}$.

There are of course many different ways to estimate σ. Following Brown [3], NavSup uses a formula based upon exponentially smoothing certain error forecasts defined as follows. Let $X_0, X_1, \cdots, X_t$ denote independent demands over discrete units of time up to and including t. Suppose that $\hat{X}_{t-1}$ is a forecast of the demand at time t, such forecast being made at time t-1. Then the difference $e_t = X_t - \hat{X}_{t-1}$, represents a forecast error that is dependent upon the method of forecasting used. Again following Brown, suppose that exponential smoothing is used to forecast so that, using smoothing constant α,

$$\hat{X}_{t-1} = \alpha \sum_{k=0}^{t-2} \beta^k X_{t-1-k} + \beta^{t-1} X_0$$

In this report we will only be considering the model where random demand X is normal with constant mean μ and variance σ^2. With this in mind we have, as previously reported in [1],

$$E(e_t) = 0 \quad \text{for all } t$$

and the error terms have time-dependent variances given by,

$$\sigma_e^2(t) = \sigma^2 + \sigma^2 \left(\frac{\alpha + 2\beta^{2t-1}}{2-\alpha}\right)$$

However, if we let $t \to \infty$, we obtain a limiting variance of

$$(3.1) \qquad \sigma_e^2 = \frac{2}{2-\alpha}\sigma^2 \quad \text{or} \quad \sigma = \frac{\sqrt{2-\alpha}}{2}\sigma_e$$

Now for each t, e_t, being a linear combination of independent normal random variables, is again normal and its mean absolute deviation, Δ_e, is related to σ_e by $\sigma_e = \sqrt{\frac{\pi}{2}}\Delta_e$ as previously established. Consequently, from (3.1) we may write, asymptotically,

$$\sigma = \frac{\sqrt{\pi(2-\alpha)}}{2} \Delta_e \quad (3.2)$$

This formula relates demand variability to the mean absolute deviation of error forecasts.

The formula used in [3] and [4] to estimate Δ_e and ultimately σ is based on an exponentially smoothed function of the error forecasts. If

$$\tilde{\Delta}_t = \alpha \sum_{k=0}^{t-1} \beta^k |e_{t-k}|$$

then σ is estimated for demands $X_0, \cdots, X_t$, by

$$\tilde{\sigma} = \frac{\sqrt{\pi(2-\alpha)}}{2} \tilde{\Delta}_t \quad (3.3)$$

As remarked in [1], the worth of this estimation procedure is difficult to assess without some knowledge of the probability distribtuion of $\tilde{\Delta}_t$. Serious efforts notwithstanding, that probability distribution has not been derived to date. The real basic difficulty is that random variables $e_1, e_2, \cdots, e_t$ which compose $\tilde{\Delta}_t$, while normally distributed, are not mutually independent. Even the matter of finding the second moment of $\tilde{\Delta}_t$ has been intractible thus far.

In an effort to discover some indication of the variability involved, a simulation study was undertaken and resulted in mounting evidence that there is reason to believe that the variance of $\tilde{\Delta}_t$ is bounded below by a positive quantity so that even in the limit the actual estimates $\tilde{\sigma}$ based on this quantity will fluctuate about the true value σ.

In turn, this immediately suggests comparing this method of estimation with the classical maximum likelihood estimate $\hat{\sigma}$ given by the formula,

$$\hat{\sigma} = \frac{\sqrt{\sum_{i=1}^{t} (X_i - \bar{X})^2}}{t} \tag{3.4}$$

Now, properties of $\hat{\sigma}$ are well known (see [5] for example), being derived from the so-called Chi-distribution. In fact,

$$E(\hat{\sigma}) = \sqrt{1-\frac{1}{t}}\, \alpha_t \sigma \quad \text{where} \quad \alpha_t = \sqrt{\frac{2}{t-1}}\, \frac{\Gamma(\frac{t}{2})}{\Gamma(\frac{t-1}{2})} \doteq \sqrt{1-\frac{1}{2(t-1)}}$$

and,

$$V(\hat{\sigma}) = \frac{t-1}{t}\, \beta_t^2\, \sigma^2 \;, \quad \text{where} \quad \beta_t^2 = 1 - \alpha_t^2 \doteq \frac{1}{2(t-1)}$$

From these formulas we see that

$$\lim_{t\to\infty} E(\hat{\sigma}) = \sigma$$

$$\lim_{t\to\infty} V(\hat{\sigma}) = 0 \tag{3.5}$$

Hence, if it is true that the variance of $\tilde{\sigma}$ is bounded away from zero, then, even though unbiased in the limit, the fact that $\hat{\sigma}$ has a vanishing variance would make it preferable as an estimating tool. Our simulation results certainly seem to concur that this is the case.

For simulation purposes, several members of the normal family were selected so as to be representative of a wide class of fast moving inventory items. Using random generation samples of size 200 were drawn from these populations and a running account was kept of the estimates $\tilde{\sigma}$ and $\hat{\sigma}$ of the true and known standard deviation σ. This was done by computing the values of (3.3) and (3.4) at each time period t.

Incidentally, it should be observed in this regard, that the running computation of $\hat{\sigma}$ involves no more time nor storage than that of $\tilde{\sigma}$, a feature often cited as one compelling reason for using exponential smoothing. It is true that $\hat{\sigma}^2$ will have to be stored and updated at the next time period and hence one extra operation, that of taking the square root, will be involved. But with modern computers, the time for this extra operation is negligible.

The actual point by point results of the simulation are perhaps not too enlightening. The program that was used has been preserved and is available for further use. Of more interest is a comparison of the two procedures with regard to bias and variance. These have been estimated by computing the sample average and sample standard deviation (S.D.) of each of $\tilde{\sigma}$ and $\hat{\sigma}$ for each parameter choice. While many more cases were examined, the results for parameter choices furnished by NavSup are summarized in Table 1.

Parameter Pairs	Average $\tilde{\sigma}$	Average $\hat{\sigma}$	Bias $\tilde{\sigma}$	Bias $\hat{\sigma}$	S.D. $\tilde{\sigma}$	S.D. $\hat{\sigma}$	M.S.E. $\tilde{\sigma}$	M.S.E. $\hat{\sigma}$
(50,10)	10.27	9.60	0.27	-0.40	2.83	1.33	8.08	1.93
(100,25)	25.67	23.99	0.67	-1.01	7.08	3.31	50.57	11.98
(100,50)	51.33	47.98	1.33	-2.02	14.15	6.63	201.99	48.04
(400,40)	41.07	38.39	1.07	-1.61	11.32	5.30	129.29	30.78
(500,50)	51.38	48.52	1.38	-1.48	14.15	7.40	202.13	56.95
(600,60)	61.60	57.58	1.60	-2.42	16.98	7.95	290.88	69.06
(700,70)	64.99	58.78	-5.01	-11.22	21.40	10.56	483.06	237.40
(800,80)	81.51	76.15	1.51	-3.85	19.18	5.85	370.15	49.04
(900,90)	87.26	86.21	-2.74	-3.79	21.09	5.01	452.30	39.46
(1,000,100)	101.89	95.18	1.89	-4.82	23.97	7.32	578.13	76.81

Table 1. Sample Characteristics

The table clearly brings out the inefficiency of the smoothing technique compared with classical methods of estimation. Both procedures are biased. The smoothing procedure tends to overestimate while the maximum likelihood procedure more conservatively underestimates σ, a fact known from the theory of course. Except for the one extreme case (the pair (700,70)) where the bias for $\hat{\sigma}$ is numerically as large as 11 for some reason that is not consistent with the other results, the bias in both cases is of the same magnitude. There is a marked difference in the variance estimate of the two procedures, however. The variance for $\tilde{\sigma}$ is significantly higher than that for $\hat{\sigma}$ in every case. Indeed, there was never a single normal case studied in which this was not the case.

Since the bias is roughly the same for each procedure this means that the mean squared error (M.S.E.) of each procedure will largely be determined by its variances. Recalling that M.S.E. is given by the variance plus the square of the bias we use the averages and standard deviations of Table 1 to compute estimates of M.S.E. which are recorded in the last two columns. Again the difference in the two procedures are quite striking. The M.S.E. for $\tilde{\sigma}$ reaches as high as 12 times that for $\hat{\sigma}$ in one case and even in the pathological case cited above (the pair (700,70)) where the bias of $\hat{\sigma}$ is unexplainably high, the M.S.E. of $\tilde{\sigma}$ is still twice that of $\hat{\sigma}$. In most of the cases, the ratio is about 4.

A similar analysis was undertaken for Poisson demand with results that are not quite as striking. These are summarized in Table 2 for three typical cases.

Parameter	σ	Average $\tilde{\sigma}$	Average $\hat{\sigma}$	Bias $\tilde{\sigma}$	Bias $\hat{\sigma}$	S.D. $\tilde{\sigma}$	S.D. $\hat{\sigma}$	M.S.E. $\tilde{\sigma}$	M.S.E. $\hat{\sigma}$
.01	0.1	.01	.06	-.09	-.04	.04	.04	.0097	.0032
0.1	0.3	.17	.23	-.13	-.07	.13	.06	.0338	.0085
1.0	1.0	.91	.95	-.09	-.05	.28	.12	.0865	.0169

Table 2. Poisson Demand

Again it should be observed that the M.S.E. for $\tilde{\sigma}$ is at least 3 times that for $\hat{\sigma}$ in each instance examined.

4. EFFECTS ON REORDER LEVELS

Not satisfied to merely summarize the evidence obtained as in the tables of the preceding section, it was deemed advisable to test the effects of the forecasting errors on their use in fixing reorder levels. Again for reporting purposes, the main concentration was spent on the normal case. If demand is normal with mean μ and standard deviation σ, then the (theoretical) reorder level would be set at $\mu + k\sigma$ where k is chosen to satisfy a given risk ρ defined by

$$\rho = P(X > \mu + k\sigma)$$

Thus, the risk, for present purposes, is the probability that demand will exceed the reorder level causing a stockout. For given ρ (or k) the parameter k (or ρ) can be determined from standard normal tables.

Of course μ and σ are unknown so that ρ can never be satisfied exactly. If we were to use μ' and σ' for the reorder level and it should happen that $\mu' + k\sigma' > \mu + k\sigma$ where k has been chosen to satisfy a given risk requirement ρ, then the true risk say $\rho' = P(X > \mu' + k\sigma')$ would be something smaller. This means that simultaneously overestimating μ and σ results in over-stocking, that is, we could stock less to achieve the required risk level. There may also be penalty costs in the way of storage costs to consider for such a situation. On the other hand, if $\mu' + k\sigma' <$ $\mu + k\sigma$ then, while we think we are stocking in such a way as to achieve a given risk ρ, in fact the true risk ρ' would be greater. The resulting shortage "cost" paid for such a position could be disastrous. Since neither position is particularly favorable, and the true parameter values are unknown, it is clear that the most precise estimates of μ and σ are desirable; precise in the sense of minimum fluctuation about the true values.

Trying to determine the true risk incurred when exponentially smoothed estimates $\tilde{\mu}$ and $\tilde{\sigma}$ of μ and σ, respectively, are used theoretically requires the joint probability distribution of X, $\tilde{\mu}$ $\tilde{\sigma}$. If finding the distribution of $\tilde{\sigma}$ alone seems difficult, the task of finding such a joint distribution looms formidable to say the least. Fortunately, in a simulation approach the parameters, and hence the true theoretical reorder level for a given risk ρ, is known. For any procedure used to set actual reorder levels it is then possible to observe the behavior of repeated applications of such a procedure. There are many ways this might be done. One approach is to observe the number of times a reorder level falls within so many units of the theoretical level. Or, for a given percentage P we might ask within how many units of the theoretical reorder levels will P percent of the actual reorder levels be found? We have done a little of both and summarized the findings for various cases in Table 3.

The table is constructed as follows. First a triple of parameters (μ, σ, ρ) is chosen where ρ is the desired risk. From this k is determined from tables and a theoretical reorder level computed. For example, if $\mu = 400$, $\sigma = 40$ and $\rho = .01$ then from $.01 = P(X>400+40k)$ we determine $k = 2.326$ and theoretical reorder level is 493.04. Then, for each parameter triple Q_1 is the percentage (rounded) of 200 reorder levels found to be within 1 unit of the theoretical reorder level first for the level based on exponential smoothing $\tilde{\mu} + k\tilde{\sigma}$ and, second, for a level set using $\bar{X} + k\hat{\sigma}$ where $\hat{\sigma}$ is the maximum likelihood estimate of σ and $\bar{X}$ is the average demand. Next, P_{50} represents the number of units about the theoretical reorder level within which 50% of the computed reorder levels were found. Again, the first column under P_{50} is for smoothed estimates and the second for maximum likelihood estimates. The last column is a similar computation for 95% of the reorder levels.

Parameter Triples	Q_1		P_{50}		P_{95}	
(100,25,.01)	4	32	15	2	35	13
(100,25,.11)	6	48	11	1	24	6
(100,25,.50)	8	19	6	2	17	5
(400,40,.01)	2	22	25	2	57	21
(400,40,.11)	5	38	18	2	38	9
(400,40,.50)	5	10	10	3	27	7
(600,60,.01)	1	15	37	4	85	31
(600,60,.11)	5	26	26	3	58	14
(600,60,.50)	3	6	15	4	40	12
(800,80,.01)	2	15	35	5	103	46
(800,80,.11)	2	16	28	3	69	30
(800,80,.50)	2	23	21	2	56	23
(1000,100,.01)	2	13	44	6	129	58
(1000,100,.11)	1	12	35	4	85	37
(1000,100,.50)	2	18	27	3	70	30

Table 3. Percentiles for Reorder Levels

Many more cases than those reported in the table were examined but the results are omitted for the sake of brevity. The results in these other cases displayed precisely the same pattern however. Indeed, the consistency of the various cases is somewhat startling. Once again there is not one single choice of a mean, a standard deviation or a risk in which the smoothing proceuure does not fall significantly short of maximum likelihood techniques. This was to be expected of course since the variance of the estimators involved determine the amount of flucutation from sample to sample.

One other interesting observation in Table 3 is the monotonic nature of P_{50} and P_{95}. Thus, while it always takes more units of distance (roughly twice as many) to include 95% of the values using smoothing versus maximum likelihood, the comparison at the 50% level is even more striking. In addition to requiring roughly ten times as many units to cover 50% of the observations for smoothing, the number of units for the same percentage using maximum likelihood is comfortably small.

Another way to view the results is to try and estimate the true risk incurred when a given procedure is followed. One way to do this is to imagine that, after 200 observations, the estimates say μ^* and σ^* coincide with the parameters and then compute $P(X>\mu^*+k\sigma^*)$ as a measure of the true risk. With a sample size of 200, the asymptotic result of supposing that estimates and parameters are the same should be fairly sound. Of course the point is that how sound such an estimate is depends upon the precision of the estimate once again. In any event in practice this is essentially what is done and

200 is a figure quite a bit larger than the typical number of periods for which historical data are maintained for making such estimates within NavSup.

In Table 4 our results are summarized as follows. Again, certain triples of parameters μ, σ and ρ are selected. This time, for each pair (μ, σ) down a column, five different risks are associated across a row. For each such combination, an actual reorder level is determined in two ways. First of all, k is determined and then reorder levels fixed at $\tilde{\mu} + k\tilde{\sigma}$ and $\bar{X} + k\hat{\sigma}$. After the 200th observation $\rho' = P(X>\mu^*+k\sigma^*)$ is computed and entered in each cell (suitably rounded), first for $(\mu^*,\sigma^*) = (\tilde{\mu},\tilde{\sigma})$ and second for $(\mu^*,\sigma^*) = (\bar{X},\hat{\sigma})$. For lack of a better name, ρ'' is called the actual risk. In this way, the actual risk incurred using smoothing may be compared with maximum likelihood procedures.

(μ,σ) \ ρ	.01		.05		.11		.25		.50	
(100,25)	.01	.01	.05	.05	.13	.11	.30	.25	.60	.52
(400,40)	.01	.01	.05	.05	.13	.11	.30	.25	.60	.52
(500,50)	.01	.01	.05	.05	.13	.11	.30	.25	.60	.52
(600,60)	.01	.01	.05	.05	.13	.11	.30	.25	.60	.52
(800,80)	.06	.01	.18	.05	.31	.12	.52	.26	.76	.52
(900,90)	.12	.02	.20	.07	.27	.13	.37	.27	.50	.51
(1000,100)	.06	.01	.18	.05	.31	.12	.52	.26	.77	.52

Table 4. Actual Risks

The results are fairly self-explanatory. The results are even fairly consistent in the sense that for most parameter choices the asymptotic results are quite the same. For large means and variances more variation begins to show up. We see that for small values of ρ, both procedures are roughly equivalent. That is to say, after 200 predictions, the actual risk is fairly close to the true risk. But for large or even moderate values of ρ it appears that smoothing consistently produces a larger actual risk than maximum likelihood. Both are larger than the true risk but the maximum likelihood procedure is much closer to the true value and is far more consistent.

Risk	$\tilde{\mu} + k\tilde{\sigma}$	$\tilde{\mu} + k\hat{\sigma}$	$\bar{x} + k\hat{\sigma}$
.01	51	38	44
.02	52	38	43
.03	52	40	41
.04	53	39	40
.05	54	39	39
.06	52	38	37
.07	52	40	37
.08	52	40	34
.09	52	40	35
.10	51	41	34

Table 5. Percentage Overestimates

Finally, supposing that it is far more serious to overestimate the reorder level (resulting in excess stock for a given risk) than to underestimate it, the data were analyzed to see the percentage of time the reorder level was overestimated in 200 trials. For this illustration only one distribution was used, namely normal with mean 50 and standard deviation 10. Again, for various risks (this time no more than 10%) reorder levels are set using $\tilde{\mu} + k\tilde{\sigma}$ and $\bar{x} + k\tilde{\sigma}$. This time we examined a third alternative, a mixture of the two procedures, given by $\tilde{\mu} + k\hat{\sigma}$. The results are summarized in Table 5 and the entries are the percentage of times that the given method overestimates the theoretical reorder level.

It is by this time perhaps not surprising to note that smoothing consistently overestimates the reorder level. Even the mixed procedure, whereby smoothing is used to estimate the mean and σ is estimated by maximum likelihood, is an improvement over a MAD estimate of σ. Compared with straight maximum likelihood, it is better for small values of ρ and then begins to fall off as ρ increases.

5. CONCLUSIONS AND RECOMMENDATIONS

In drawing conclusions it would be well to reiterate the tentative nature of simulation results. Without the exact probability distribution for the exponentially smoothed estimators treated in this report—particularly MAD—no definite statement can be made about the efficiency of this method over others. But the simulation results display a certain consistency that certainly lends support to the claim that maximum likelihood methods would be superior for those purposes to which smoothing is presently employed. Certainly the evidence is sufficient to warrant further examination and comparison with alternatives. In spite of the difficulties involved in the derivation of the probability distributions when exponential smoothing is employed as a forecasting tool, it is our recommendation to continue the theoretical search for those distributions or at least the moments that are involved. In particular it would be helpful to establish that the variance of $\tilde{\sigma}$ is bounded below by a positive quantity as seems to be indicated by the simulations. If this were true then the estimate will always fluctuate about the true σ regardless of how many time periods are considered.

Another reason for continuing research in this area is the fact that even the simulation studies were only carried out for the normal case (and the isolated Poisson cases) with constant mean and variance. This is a strong assumption and entails supposing that no trend is present. It is highly recommended that a similar analysis in the presence of trend be carried out. The trend test and indicator presently used by NavSup as outlined in [4] again employs exponential smoothing as a basic tool for analyzing the data. If the lack of efficiency indicated in this study carries over to that case as well then the consequences should be even more serious. Indeed, there is reason to believe that with the methods presently employed, there can be a 50% chance of rejecting an almost vertical trend, that is a sudden jump in the demand pattern. These matters are certainly topical candidates for further research.

In any or all of the above recommendations, it is suggested that real data be used to analyze, unfortunately perhaps in retrospect, the reorder system had methods such as maximum likelihood been used. Large amounts of such data are available for a wide variety of inventory items and the programs for processing such data have been preserved.

6. APPENDIX

This appendix is devoted to supplying the mathematical details for the monotonic nature of $\frac{\Delta}{\sigma}$ as asserted in Section 3. It should be recalled from that section that $\frac{\Delta}{\sigma}$, expressed as a function of the Poisson parameter λ is given by

$$\frac{\Delta}{\sigma} = 2\sqrt{\lambda}\, e^{-\lambda} \frac{\lambda^{[\lambda]}}{[\lambda]!}$$

In particular, when $\lambda = n$, $\frac{\Delta}{\sigma}$ has a relative minimum with a value u_n given by,

$$u_n = 2\sqrt{n}\, e^{-n} \frac{n^n}{n!}$$

It is conjectured that this sequence of values is monotone increasing. To see this observe that

$$\frac{u_n}{u_{n+1}} = e\left(\frac{n}{n+1}\right)^{n+\frac{1}{2}}$$

and

$$\log \frac{u_n}{u_{n+1}} = (n+\tfrac{1}{2}) \log \frac{n}{n+1} + 1 = -(n+\tfrac{1}{2}) \log(1+\tfrac{1}{n}) + 1 .$$

Recalling the Taylor expansion of $\log(1+x)$ (see [6] for example), we have, for $n > 1$,

$$\log(1+\tfrac{1}{n}) = \frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - R$$

where

$$|R| \leq \frac{(\frac{1}{n})^4}{4} = \frac{1}{4n^4}$$

Then,

$$\log \frac{u_n}{u_{n+1}} = -[1 + \frac{1}{2n} - \frac{1}{2n} - \frac{1}{4n^2} + \frac{1}{3n^2} + \frac{1}{6n^3} + \frac{2n+1}{2} R] + 1$$

$$= -\frac{1}{12n^2} - \frac{1}{6n^3} - \frac{2n+1}{2} R$$

But $R < 0$ for this case so that

$$-\frac{2n+2}{2} R = \frac{2n+2}{2} |R| \leq \frac{1}{4n^3} + \frac{1}{8n^4}$$

Hence,

$$\log \frac{u_n}{u_{n+1}} \leq -\frac{1}{12n^2} - \frac{1}{6n^3} + \frac{1}{4n^3} + \frac{1}{8n^4} = -\frac{(2n^2-2n-3)}{24n^4}$$

Now for $x \geq 2$, the polynomial $2x^2 - 2x - 3$ is positive so that

$$\log \frac{u_n}{u_{n+1}} < 0 \quad \text{and hence} \quad \frac{u_n}{u_{n+1}} < 1$$

In other words,

$$u_n < u_{n+1} \qquad n = 2, 3, \cdots$$

which is what was to be verified.

From [7] (page 64),

$$e^{-(n+\frac{1}{2})} \sqrt{2\pi}\, (n+\tfrac{1}{2})^{n+\frac{1}{2}}\, e^{\frac{1}{24(n+\frac{1}{2})}} < n! < \sqrt{2\pi}\, (n+\tfrac{1}{2})^{n+\frac{1}{2}}\, e^{-(n+\frac{1}{2})}$$

Since $f_n(n+\frac{1}{2}) = \frac{2}{n!}(n+\frac{1}{2})^{n+\frac{1}{2}} e^{-(n+\frac{1}{2})}$ we have

$$\sqrt{\frac{2}{\pi}} < f_n(n+\tfrac{1}{2}) < \sqrt{\frac{2}{\pi}}\, e^{\frac{1}{24(n+\frac{1}{2})}}$$

Thus the sequence of relative maxima of $\frac{\Delta}{\sigma}$ converges to $\sqrt{\frac{2}{\pi}}$ supporting the claim on page 7 that $\frac{\Delta}{\sigma}$ then converges to this limit.

BIBLIOGRAPHY

[1] Zehna, Peter W., Some Remarks on Exponential Smoothing, Naval Postgraduate School, Technical Report No. 72, December 1966.

[2] Ashar, V.G. and T.D. Wallace, "A Sampling Study of Minimum Absolute Deviation Estimators," Operations Research, Vol. 11, No. 5, September-October 1963.

[3] Brown, R.G., Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice-Hall, Inc., 1963.

[4] PAR 1 - Application D, Operations 5, 6, Levels Computation for Consumables and Repairables, FMSO, Mechanicsburg, Pennsylvania, 1964.

[5] Hald, A., Statistical Theory with Engineering Applications, John Wiley & Sons, 1952.

[6] Taylor, Angus E., Advanced Calculus, Ginn and Company, 1955.

[7] Feller, William, An Introduction to Probability Theory and its Applications, Vol. 1, John Wiley & Sons, 1950.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

"FORECASTING ERRORS USING MAD"

4. DESCRIPTIVE NOTES (Type of report and, inclusive dates)

Research Report

5. AUTHOR(S) (First name, middle initial, last name)

Peter W. Zehna

6. REPORT DATE April 1969	7a. TOTAL NO. OF PAGES 33	7b. NO. OF REFS 7

8a. CONTRACT OR GRANT NO. NAVSUP RDT&E No. TF015-02-101 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) NPS 55Ze9041A 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Research and Development Division Naval Supply Systems Command

13. ABSTRACT

A study of the effects of using mean absolute deviation (MAD) to estimate variability in setting reorder levels for the inventory of a stock item. The method presently employed by NavSup in setting such reorder levels involves exponentially smoothed estimates of the mean and variance of the demand process. Any error involved in setting reorder levels results in a change in the underlying risk which in turn can be translated into costs. Such errors for the method of estimation presently employed are compared with standard maximum likelihood procedures. By simulating several normal systems, the smoothing technique is found to be inferior to classical methods with no reduction in computational difficulties.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

UNCLASSIFIED
Security Classification

A-31408

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
INVENTORY THEORY						
EXPONENTIAL SMOOTHING						
FORECASTING						
MEAN ABSOLUTE DEVIATION						
REORDER LEVELS						