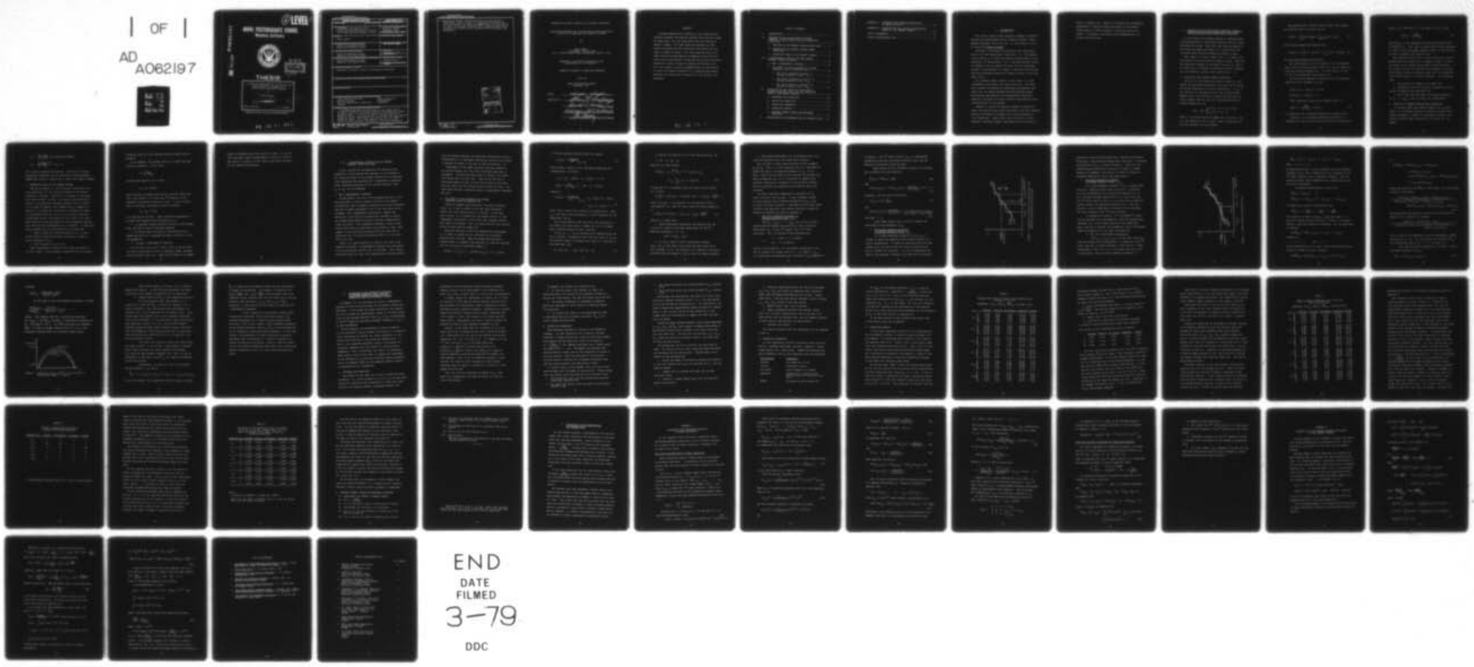
AD-A062 197 NAVAL POSTGRADUATE SCHOOL MONTEREY CALIF F/G 12/2
A DECISION PROCEDURE FOR THE TRANSITION FROM UNCERTAINTY TO RIS--ETC(U)
SEP 78 K SAGDIC

UNCLASSIFIED NL

| OF |

AD
A062197



END
DATE
FILMED
3-79
DDC

1.0 1.1 1.25 1.4 1.6 1.8 2.0 2.2 2.5 2.8 3.2 3.6 4.0 4.5 5.0 5.6 6.3

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

NAVAL POSTGRADUATE SCHOOL

Monterey, California

master's thesis

THESIS

A DECISION PROCEDURE FOR THE TRANSITION FROM UNCERTAINTY TO RISK IN A SINGLE-PERIOD INVENTORY PROBLEM

by

Kadir Sagdic

Sep[redacted]978

58p

Thesis Advisor: G. F. Lindsay

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) A Decision Procedure for the Transition from Uncertainty to Risk in a Single-Period Inventory Problem	5. TYPE OF REPORT & PERIOD COVERED Master's Thesis: September 1978
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Kadir Sagdic	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940	12. REPORT DATE September 1978
	13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Single-period inventory
Newsboy problem
Risk and uncertainty conditions
Minimax
Order statistics
Nonparametric
Estimator

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Minimum expected-cost solutions to the single-period inventory problem (the Newsboy problem) have been well known for many years, for risk cases where the distribution of demand is known. Also well known are minimax cost and Laplace solutions for the uncertainty case where only the range of demand is known. This study explores the use of order-statistic-based quantile estimators as decision procedures

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE UNCLASSIFIED
(Page 1) S/N 0102-014-6601 |

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

CONT → while data gathers during the early decision periods. Simulation, using a variety of demand distributions and unit-cost values, provides recommendations on which period to leave the minimax procedure for a suggested data-based decision rule which will be asymtotic to the optimal risk procedure.

A Decision Procedure for the Transition from Uncertainty to Risk in a Single-Period Inventory Problem

by

Kadir Sagdic
Lieutenant, Turkish Navy
B.S., United States Naval Postgraduate School, 1978

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1978

Author: Kadir Sagdic

Approved by: Glenn F. Lindsay, Thesis Advisor

Donald P. Gaver, Second Reader

M G Sovereign (J K Hartman), Chairman, Department of Operations Research

Dean of Information and Policy Sciences

ABSTRACT

Minimum expected-cost solutions to the single-period inventory problem (the Newsboy problem) have been well known for many years, for risk cases where the distribution of demand is known. Also well known are minimax cost and Laplace solutions for the uncertainty case where only the range of demand is known. This study explores the use of order-statistic-based quantile estimators as decision procedures while data gathers during the early decision periods. Simulation, using a variety of demand distributions and unit-cost values, provides recommendations on which period to leave the minimax procedure for a suggested data-based decision rule which will be asymtotic to the optimal risk procedure.

TABLE OF CONTENTS

I. INTRODUCTION

The classic single period inventory problem is called by various well-known names such as "Newsboy problem," "Newspaperboy problem," and so on. [6]. In this thesis it will be called the Newsboy problem.

Solutions to the Newsboy problem are well known under both "risk" and "uncertainty" conditions. [7]. The purpose of this thesis is to study the Newsboy problem under initial conditions of "Uncertainty," (i.e., when the possible future for demand is not yet known) and then, for the subsequent periods as data gathers, to suggest, structure, and evaluate decision procedures during this transition phase to decision under "Risk."

The inventory model studied in this thesis is a time independent (stationary) and linear-cost type. The study will consider the problem of minimizing the expected cost. Since with the Newsboy problem both the expected profit maximization and expected cost minimization cases have identical solutions, the study should be applicable to both interpretations of the problem.

Chapter II reviews the Newsboy problem for decisions under risk and decisions under uncertainty, and then suggests possible alternative procedures for "from uncertainty to risk" conditions. Among these, Order Statistics (a non-parametric approach) appear promising and are discussed in detail in Chapter III. Chapter IV evaluates the alternative nonparametric transition phase procedures of the Newsboy problem and as a result proposes a transition phase procedure. In Chapter V conclusions and recommendations for further work are given.

II. SOLUTION OF THE SINGLE-PERIOD INVENTORY (NEWSBOY) PROBLEM UNDER RISK AND UNCERTAINTY CONDITIONS

In this section the Newsboy problem is examined under risk and uncertainty, and optimal decision rules for these conditions are given. These two cases form the bounds of the transition phase and permit us to characterize the transition phase decision problem we intend to study.

Here, when it is said, "under risk," it is meant that the probability distribution of demand is known. When the distribution of demand is not known, and data in hand is not sufficient to estimate the possible futures of demand, then the problem is called "under uncertainty." [7].

A. SOLUTION OF THE NEWSBOY PROBLEM UNDER RISK

Let $f_D(x)$ be the probability density function of the demand for continuous populations, and $p_D(x)$ be the probability mass function for discrete populations. If C_{so} is the stock-out or shortage cost per unit of unsatisfied demand, and C_c the carrying or surplus cost per unit of excess supply, then the cost for a particular period of Newsboy problem will be:

$$\text{Cost} \quad C(S) = \begin{cases} C_c(S-X), & \text{if } 0 \leq X \leq S, \text{ and} \\ C_{so}(X-S), & \text{if } X > S, \end{cases}$$

where X is the realization of demand for that period, and S is the inventory level in hand. Here S represents the decision variable for this problem.

The expected cost $E[C(S)]$ which results from keeping an inventory level S in hand, will be:

$$E[C(S)] = \sum_{X=0}^{S} C_c (S-X) p_D(X) + \sum_{X=S+1}^{\infty} C_{so}(X-S) \quad p_D(X) \qquad (1)$$

for discrete demand distributions and,

$$E[C(S)] = C_c \int_0^S (S-x) \quad f_D(x)dx + C_{so} \int_S^{\infty} (x-S) \quad f_D(x)dx, \qquad (2)$$

for continuous demand distributions.

Now, the expected cost for a period is to be minimized and we wish to find an optimal value of S, say S_o, that will yield the minimum expected cost. In other words, (1) and (2) are to be minimized with respect to S.

In the discrete demand distribution case, two necessary conditions for a minimum at S_o are:

$$E[C(S_o + 1)] - E[C(S_o)] \geq 0, \text{ and}$$

$$E[C(S_o - 1)] - E[C(S_o)] \geq 0.$$

These conditions yield the well known result [7]:

$$p_D(S_o - 1) \leq \frac{C_{so}}{C_c + C_{so}} \leq p_D(S_o), \qquad (3)$$

as conditions for a minimum expected-cost solution at S_o.

Similarly, for continuous demand distributions, to minimize (2), the first derivative of $E[C(S)]$ is taken with respect to S, and after it is set equal to zero, yields

$$F_D(S_o) = \frac{C_{so}}{C_c + C_{so}} \tag{4}$$

where $F_D(S_o)$ is the cumulative distribution function (CDF) of demand at S_o. Thus, the S_o values which satisfy equations (3) and (4) are the optimum inventory levels to keep on hand in order to minimize the expected cost.

This solution (under risk) is quite idealistic, since it assumes that the distribution of demand is known. In fact, in reality we are seldom certain about the distribution. In most cases, neither the distribution of demand, nor an adequate amount of data to provide a good basis for estimation of the distribution is initially available. The situation when:

1. Distribution of demand is not known, and

2. Enough data from the system to estimate the distribution of demand is not available, but

3. We do have some idea about the range of demand D, say $0 < D < D_{max}$, is one we shall call "uncertainty."

B. SOLUTION OF NEWSBOY PROBLEM UNDER UNCERTAINTY

It has been shown that whether minimax cost, minimax regret, or the Laplace approach, assuming a uniform distribution of demand is chosen as a principle of choice, the optimal decision under uncertainty is the same [6], [7], namely,

$$S_o = \frac{C_{so} \quad D_{max}}{C_c + C_{so}}, \text{ for continuous demand,}$$

$$S_o < \frac{C_{so}(D_{max} + 1)}{C_c + C_{so}} < S_o + 1,$$

for a discrete demand distribution. The results of these uncertainty procedures are not affected by newly generated demand data unless it causes one to revise the range of demand.

C. TRANSITION PHASE OF THE NEWSBOY PROBLEM

Now the two bounds or limit solutions (uncertainty and risk conditions) for the transition phase of the Newsboy problem are reviewed. Implicitly, to proceed in accordance with these current solutions, one begins under Uncertainty and uses the Uncertainty optimum inventory level S_o (based on the range of demand), until demand data is adequate or complete to estimate the distribution of demand and change to the optimum inventory level S_o for the Risk conditions. As mentioned before, the purpose of this study is to find a good transition phase procedure for the intermediate periods when some data has accumulated and is available. A transition phase procedure may be characterized by two decisions:

1. When to begin the transition phase (i.e., at what period should we leave Uncertainty procedures to begin to use an estimator of S_o).

2. What estimator of S_o to use.

After explaining the nature of the transition phase in the next chapter, three candidate estimators for the optimum inventory level S_o to use during transition phase will be proposed.

As we remember, the optimal choice of S under Risk was based on a quantile. Let us call,

$$p = \frac{C_{so}}{C_c + C_{so}} ,$$

and then from equation (4) we have,

$$S_o = X_p = F_D^{-1}(p) ,$$

which provides a minimum expected-cost solution under Risk. Thus, the problem is to locate the p^{th} quantile of the Cumulative Distribution Function, $F_D(X)$. As more and more data is collected, one might be able to estimate,

$$\hat{S}_o = \hat{X}_p = \hat{F}^{-1}(p) ,$$

as a function of the data. Some alternative approaches to estimate this quantile were considered to be:

1. To consider similar systems for which we have demand data, and then to apply a Bayesian approach.

2. To decide about the form of the distribution of demand and carry out a maximum likelihood procedure to find the parameters.

3. To apply a nonparametric approach.

For the first two cases, if one is able to decide about the form of distribution such as lognormal, gamma, compound-poisson process, beta, etc., then the p^{th} quantile of demand might be estimated quite well subject to data. It is the third approach, that of nonparametric statistics, which is the center of interest in this thesis, and forms the basis for the chapter that follows.

III. A NONPARAMETRIC APPROACH FOR THE NEWSBOY PROBLEM, TRANSITION PHASE

In this section the possibilities for applying a nonparametric or distribution-free approach to the problem of the transition phase of the Newsboy problem will be explored. After explaining the reasoning for a nonparametric approach, three candidate estimators of the optimum inventory level S_o (or X_p) will be proposed.

A. WHY A NONPARAMETRIC APPROACH?

As mentioned at the end of the preceding section, if the decision maker is able to predict the form of the demand distribution, then a parametric approach can be more efficient and have a higher convergence rate to the true value of the estimator than a nonparametric equivalent. Hadley and Whitin [5], give various considerations relating to choice of a demand distribution and about inventory demand predictions, in detail. For a parametric approach, such as would result from fitting a theoretical distribution to data, they suggest to use poisson distribution when the lead time is relatively low, or normal distribution when it is not relatively low, etc.

Since it is quite difficult to specify the form of the demand distribution, a nonparametric approach was considered to be of particular interest for this problem. Also, Hadley and Whitin point out that, for single-period inventory models, like the Newsboy problem, the empirical distribution function (nonparametric) is reasonably applicable, unless we are forced to operate the model at the tail of the distribution. [5].

Performance of the model decision rules was assumed to be critical during, say, the first 50 periods when data is inadequate. The decision maker needs to operate the system well during that critical time. Later on he will have a sufficient amount of data to treat the problem as one under Risk and simply use the optimal decision rules for Risk. For these early periods (transition phase) a nonparametric approach was used.

B. THE EFFECT OF THE ESTIMATORS OF OPTIMUM INVENTORY LEVEL ON EXPECTED COST

Before we suggest some estimators of optimal inventory level, X_p, it may be useful to see how those estimators affect the total expected cost in a period. In the work that follows, it is assumed that the demand process is time independent (stationary), which means that seasonal effects and trends are removed. Also, we assume that for the previous period we know what the demand was.

Before we continue on with the mathematical development, it is useful to expand on the conditions for an optimal expected-cost solution. When the demand distribution is assumed known, we remember from equation (2) that the expected cost for the case of continuous demand is:

$$E[C(S)] = C_c \int_0^S (S - x) \quad f_D(x)dx + C_{so} \int_S^\infty (x - S) \quad f_D(x)dx.$$

To find an optimum inventory level S_o, taking

$$E'[C(S)] = \left.\frac{d\,E[C(S)]}{d\,S}\right|_{S = S_o} , \tag{5}$$

and setting it equal to zero (necessary conditions for minimization), we obtain:

$$- C_{so} + C_{so} \quad F_D(S_o) + C_c \quad F_D(S_o) = 0 \text{ , or}$$

$$F_D(S_o) = \frac{C_{so}}{C_c + C_{so}} = p \text{ , and} \quad S_o = F_D^{-1}(p).$$

Similarly,

$$E''[C(S)] = \left.\frac{d^2\,E[C(S)]}{d\,S^2}\right|_{S = S_o} = C_{so} \quad f_D(S_o) + C_c \quad f_D(S_o)$$

$$= (C_{so} + C_c) \quad f_D(S_o) > 0 \text{ ,} \tag{6}$$

which forms a sufficient condition for minimization. Here, S_o is the value that corresponds to the p^{th} quantile of the population, say X_p.

We are interested in the case where the population is not known, but there exists a sample of size n of demand data, acquired during the last n periods.

Let $X_1, X_2, \ldots, X_n$ be realization of demand during the last n periods. Although we don't know the value of X_p yet, we do have a chance to estimate this value as a function of available data, say:

$$\hat{X}_p = g(X_1, X_2, \ldots, X_n) = \hat{F}_D^{-1}(p) = \hat{S}_o \tag{7}$$

To express the expected cost as the function of $\hat{X}_p$, let

$$\delta = \hat{S}_o - S_o = \hat{X}_p - X_p \ .$$

Equation (2) then becomes:

$$E[C(\hat{S}_o)] = C_c \int_0^{S_o+\delta} (S_o + \delta - x) \quad f_D(x)dx + C_{so}$$

$$+ C_{so} \int_{S_o+\delta}^{\infty} (x - S_o - \partial) \quad f_D(x)dx \qquad (8)$$

If equation (8) is expanded using the Taylor series around S_o, it gives

$$E[C(\hat{S}_o)] = E[C(S_o)] + E'[C(S_o)] \quad E[\delta] + E''[C(S_o)] \quad \frac{E[\delta^2]}{2!} + \ldots$$

Since $E'[C(S_o)] = 0$ by equation (5), and having $E''[C(S_o)]$ from equation (6), then for three terms the Taylor expansion is:

$$E[C(\hat{S}_o)] \cong E[C(S_o)] + (C_{so} + C_c) \quad f_D(S_o) \quad \frac{E[\delta^2]}{2!} \qquad (9)$$

Equation (9) shows that:

1. Under Uncertainty conditions the expected cost for a period is linear in the mean square error of the p^{th} quantile estimators, i.e., in

$$E[\delta^2] = E[(X_p - \hat{X}_p)^2]$$

2. If one is able to find a consistent estimator $\hat{X}_p(n)$, of X_p, then when n gets large the mean square error will converge to zero and $E[C(\hat{S}_o)]$ will converge to $E[C(S_o)]$, providing that the moments of $\hat{X}_p(n)$ exist and behave properly.

3. The optimal performance for an estimator here is to yield the expected total cost under Risk, $E[C(S_o)]$.

Now, we wish to find a good consistent point estimator $\hat{X}_p$ for X_p. Also, it is not required, but it is desirable to obtain one that is an unbiased estimator of X_p. If so, $E[(X_p - \hat{X}_p)^2]$ will be equal to the variance of estimator $\hat{X}_p$. Then, by equation (9), if it is possible to improve this estimator reducing its variance, it will directly reflect and will decrease the expected cost accrued by using that estimator.

To obtain candidate nonparametric estimators of X_p, order statistics will be used. In the following section we shall develop three different extimators of X_p, using an order statistics approach. Additional details about order statistics and empirical distribution functions related to quantile estimation are given in Appendix A.

1. The First Candidate Estimator of Optimum Inventory Level: $X_{(r)}$

Let $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$, be a set of n realizations from an unknown distribution in order statistics form, where $X_{(r)}$ is the r^{th} largest value (r^{th} order statistic). Let r be defined in the following way:

$$r = \begin{cases} np \ , \text{ when } np \text{ is integer} \\ [np + 1], \text{ otherwise,} \end{cases}$$

where p is the quantile, n is the number of periods so far, and [X] denotes the integer part of X. For large n, $X_{(r)}$ is a consistant and unbiased point estimator of X_p (Appendix A).

In Figure 1, the r^{th} order statistic $X_{(r)}$ is represented graphically with the distribution function $F_D(X)$ and the empirical distribution function $G_D(X)$.

Know results of this estimator include the following. For a uniform (0,1) distribution,

$$E[X_{(r)}] = E[U_{(r)}] = \frac{r}{n+1} , \tag{10}$$

and

$$\text{cov}[X_{(r)},X_{(s)}] = \text{cov}[U_{(r)},U_{(s)}] = \frac{r(n-s+1)}{(n+1)^2(n+2)}, \text{ for } r < s \tag{11}$$

Similarly, for any other distribution,

$$E[X_{(r)}] \approx F_D^{-1}(\frac{r}{n+1}) , \tag{12}$$

and

$$\text{cov}[X_{(r)},X_{(s)}] \approx \frac{r(n-s+1)}{(n+1)^2(n+2)} \cdot \frac{1}{f_D(F_D^{-1}(\frac{r}{n+1}))\cdot f_D(F_D^{-1}(\frac{s}{n+1}))} , \tag{13}$$

provided $r < s$.

Later these results will be used to compare the efficiency of the estimators of X_p.

2. <u>The Second Candidate Estimator of Optimum Inventory Level: $\bar{X}_{(r)}$</u>

As a second candidate estimator of X_p, $\bar{X}_{(r)}$ is offered to smooth the empirical distribution function by a linear interpolation. It might be considered that the area between $G_D(X)$ and the actual distribution function $F_D(X)$ may be a measure of deviation. So, here interpolation was done at the midrange, yielding a new empirical distribution

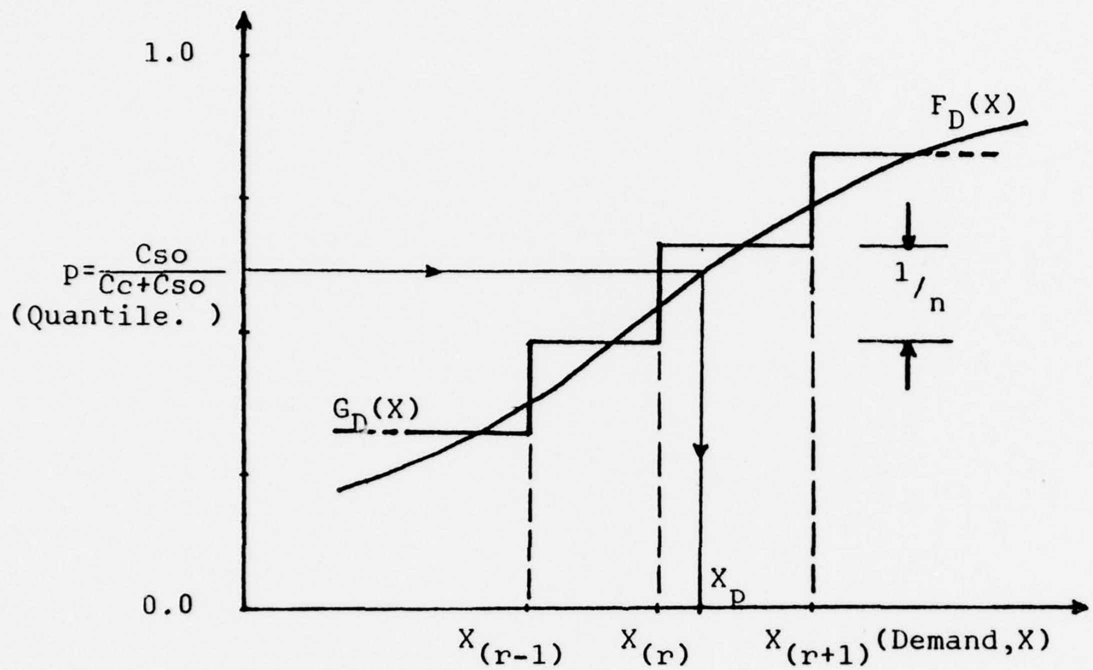


FIGURE 1. Cumulative Distribution Function $F_D(X)$ and Emprical Distribution Function $G_D(X)$.

function $\overline{R}$, which would possibly have a smaller area between $\overline{R}$ and $F_D(X)$, than the area between $G_D(X)$ and $F_D(X)$. A possible realization for $\overline{X}_{(r)}$ is shown in Figure 2. Since the area between those distribution functions is subject to change because of randomness, the results for that case would be evaluated numerically rather than analytically.

3. The Third Candidate Estimator of Optimum Inventory Level: $\overline{\overline{X}}_{(r)}$

As a third candidate estimator of X_p, a pooled order statistic, combining other order statistics, is proposed. For such an approach to be successful by equation (9), it has to reduce the mean square error to a lower value than the r^{th} order statistic does, if it is possible to make this pooled estimator an unbiased estimator of X_p. Then the problem is one of comparing the variance of this estimator with the variance of the r^{th} order statistic.

To examine this subject, it is reasonable to begin with the uniform (0,1) distribution, since the expectations, variances, etc. of its order statistics are simple and exact, and then to discuss the assumptions that are necessary before we generalize it for other distributions.

a. $\overline{\overline{X}}_{(r)}$ for the Uniform Distribution

Let $X_{(r-1)}$, $X_{(r)}$, $X_{(r+1)}$ be the order statistics from the uniform (0,1) distribution. In addition to the r^{th} order statistic, a pooled estimator in the form of a linear combination of the (r-1), (r), (r+1) order statistics will be considered. Here our third candidate estimator is:

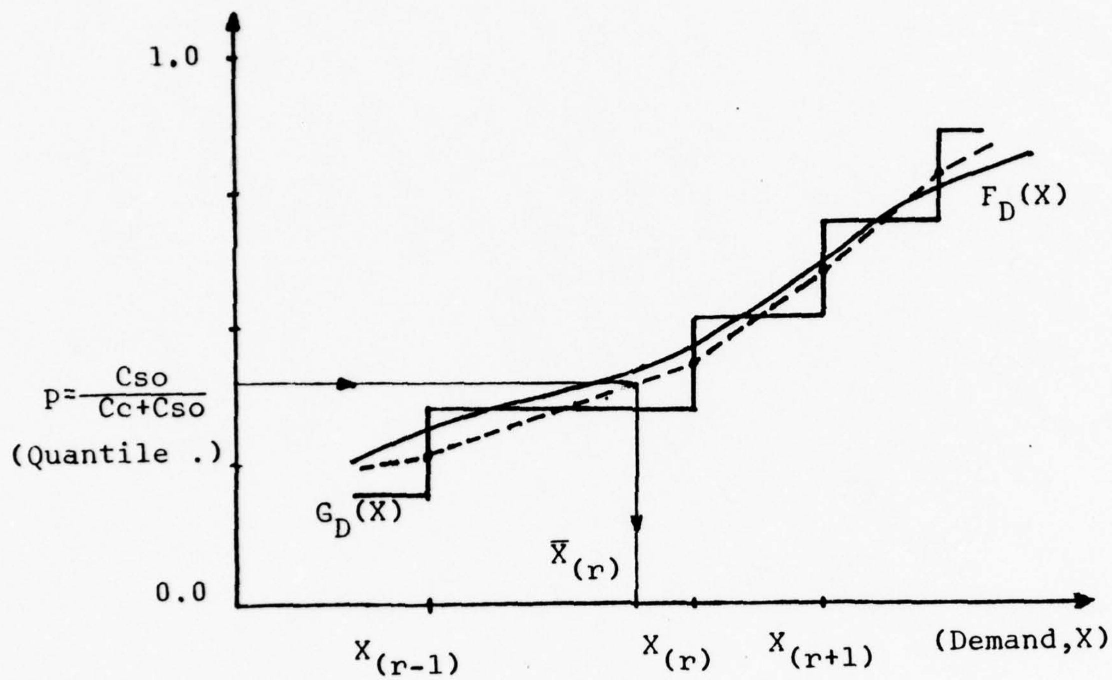


FIGURE 2. Second Candidate Estimator, $\overline{X}_{(r)}$.

$$\bar{\bar{X}}_{(r)} = \hat{X}_p = \alpha \quad X_{(r-1)} + \beta \quad X_{(r)} + \gamma \quad X_{(r+1)} \;,$$

where $\alpha + \beta + \gamma = 1$.

As mentioned earlier, if this estimator can be an unbiased estimator of X_p, then a comparison of the variances of the r^{th} order statistic (First candidate estimator), and this pooled estimator (Third candidate) will indicate which one has better performance.

To maintain unbiasedness, we must have:

$$E[\bar{\bar{X}}_{(r)}] = E[X_{(r)}] \;, \quad \text{or}$$

$$E[\bar{\bar{X}}_{(r)}] = \alpha \; E[X_{(r-1)}] + \beta \; E[X_{(r)}] + \gamma \; E[X_{(r+1)}].$$

If demand has a uniform distribution, then

$$E[\bar{\bar{X}}_{(r)}] = \alpha \; (\frac{r-1}{n+1}) + \beta \; (\frac{r}{n+1}) + \gamma \; (\frac{r+1}{n+1}) \;.$$

This quantity is equal to $E[X_{(r)}] = \frac{r}{n+1}$ only when $\alpha = \gamma$. One way to decide on the values of α, β and γ is to choose values so that the variance is minimized. So, we would like to minimize

$$\text{Var}(\bar{\bar{X}}_{(r)}) = \text{Var}(\alpha \; X_{(r-1)} + \beta \; X_{(r)} + \gamma \; X_{(r+1)}) \;,$$

subject to,

$$\alpha + \beta + \gamma = 1 \qquad \text{and} \qquad \alpha = \gamma$$

By the constraints on α, β and γ, these coefficients can be written in terms of α only, yielding

$$\text{Var}(\bar{\bar{X}}_{(r)}) = \text{Var}\{\alpha(X_{(r-1)} + X_{(r+1)}) + (1-2\alpha) \; X_{(r)}\} \;,$$

or,

$$Var(\bar{\bar{X}}_{(r)}) = \alpha^2(Var(X_{(r-1)}) + Var(X_{(r+1)})$$

$$+ 2\ Cov(X_{(r-1)},X_{(r+1)})$$

$$+ (2\alpha-4\alpha^2)(Cov(X_{(r-1)},X_{(r)}) + Cov(X_{(r)},X_{(r+1)}))$$

$$+ (1 - 4\alpha + 4\alpha^2)Var(X_{(r)}).$$

Taking the derivative of $Var(\bar{\bar{X}}_{(r)})$ with respect to α and setting it equal to zero gives,

$$\alpha = \frac{2\,Var(X_{(r)}) - Cov(X_{(r-1)},X_{(r)}) - Cov(X_{(r)},X_{(r+1)})}{Var(X_{(r-1)}) + Var(X_{(r+1)}) + 2Cov(X_{(r-1)},X_{(r+1)}) - 4[Cov(X_{r-1)},X_{(r)}) + Cov(X_{(r)},X_{(r+1)}) - Var(X_{(r)})]} \qquad (14)$$

Having available the values of variance and covariances by equation (11), equation (14) becomes

$$\alpha = \frac{2r(n-r+1) - (r-1)(n-r+1) - r(n-r)}{(r-1)(n-r+2) + 4r(n-r+1) + (r+1)(n-r) - 4(r-1)(n-r+1) - 4r(n-r) + 2(r-1)(n-r)}$$

the n,r terms in the numerator and in the denominator cancel, resulting in $\alpha = 1/2$. So, to minimize the variance of $\bar{\bar{X}}_{(r)}$, the coefficients have to be $\alpha = \gamma = 1/2$, and $\beta = 0$.

This says that when a population is uniform, a third reasonable estimator of X_p is:

$$\bar{\bar{X}}_{(r)} = 1/2(X_{(r-1)} + X_{(r+1)}), \qquad (15)$$

yielding

$$Var(\bar{\bar{X}}_{(r)}) = \frac{2r(n-r+1) - (n+1)}{2(n+1)^2 (n+2)}$$

If the ratio of the two candidate estimators is taken,

$$\frac{Var(X_{(r)})}{Var(\bar{\bar{X}}_{(r)})} = \frac{2r(n-r+1)}{2r(n-r+1) - (n+1)} > 1 ,$$

always. This implies that for a uniform distribution, $Var\ (X_{(r)}) > Var\ (\bar{\bar{X}}_{(r)})$, and $\bar{\bar{X}}_{(r)}$ is the better estimator of X_p. In Figure 3, for n - 20, these variances are sketched. When r is close to bounds (either close to 1 or n), the difference between those two variances become larger.

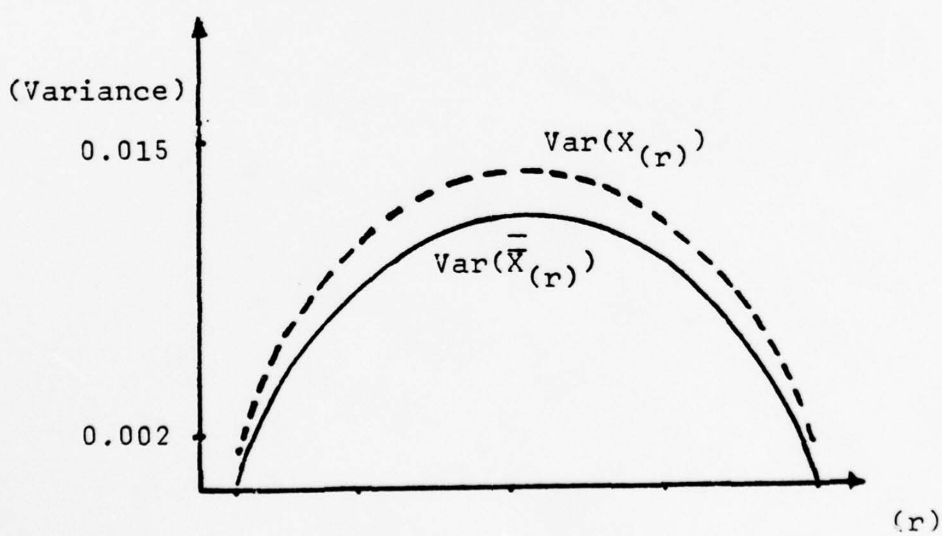


FIGURE 3. Variances of $X_{(r)}$ and $\bar{\bar{X}}_{(r)}$ for Uniform(0,1) Distribution, when n is twenty .

These results apply, of course, only to uniform demand distribution. In the following paragraphs the generalizations to any demand distribution will be discussed.

b. Generalization of $\overline{\overline{X}}_{(r)}$ to any Demand Distribution

For distributions other than uniform, it is difficult to claim that $\overline{\overline{X}}_{(r)}$, as given by equation (15), is exactly unbiased. Also, the variance and covariances, as equation (13), will be functions of the density $f_D(X)$. Even though these two problems exist, some reasonable approximations can be discussed. If we have both a large n and fairly flat densities, since (r-1), (r), (r+1) are the closest neighbors among the other order statistic values, the densities corresponding to these points would tend to equal each other and during the calculation of α, as in equation (14), would essentially cancel.

If we have flat densities (where the three points are nearly linear), or if we have a large number of observations (when the points are close together accordingly), $\overline{\overline{X}}_{(r)}$ would be approximately unbiased, too. Thus, it may be reasonable in some cases to use $\overline{\overline{X}}_{(r)}$ for other distributions besides the uniform.

Incidentally, we would not like to try another pooled estimator of X_p such as

$$\overline{\overline{\overline{X}}}_{(r)} = \alpha_2 \; X_{(r=2)} + \alpha_1 \; X_{(r-1)} + \beta \; X_{(r)} + \gamma_1 \; X_{(r+1)} + \gamma_2 \, X_{(r+2)}$$

For this estimator, the assumptions which were made to permit

$\overline{\overline{X}}_{(r)}$ to apply the distributions other than the uniform may no longer be appropriate. For example, in equation (13), $f_D(F_D^{-1}(\frac{r-2}{n+1}))$ and $f_D(F_D^{-1}(\frac{r+2}{n+1}))$ would probably have quite different values, because these are the densities of the two farthest order statistics, $X_{(r-2)}$ and $X_{(r+2)}$ in $\overline{\overline{\overline{X}}}_{(r)}$. Otherwise, if these densities were known, we would not use a nonparametric approach.

So far, solutions of the Newsboy problem under "Uncertainty" and under "Risk" have been reviewed, and possible procedures for a transition phase between them discussed. Also in this chapter, three candidate estimators of the optimal inventory level S_o (or X_p) were developed for use as alternative transition phase procedures. In the next chapter, the performance of these alternative transition procedures will be evaluated by a computer simulation of Newsboy problem decision making. Finally, from the simulation we will obtain information about what time to leave the minimax (Uncertainty) rule, for each transition phase procedure.

IV. EVALUATION OF THE CANDIDATE ESTIMATORS OF OPTIMUM INVENTORY LEVEL, FOR THE TRANSITION PHASE OF THE NEWSBOY PROBLEM

In Chapter III, the rationale for using a nonparametric approach for the transition phase of the Newsboy problem was discussed. It was shown that under Uncertainty the expected cost for a period was approximately linear in the mean square error of the optimum inventory level S_o (or X_p) corresponding to the p^{th} quantile, and three candidate estimators of X_p or S_o were introduced.

In this chapter, the performance of these estimators will be tested for several demand distributions by simulation. After explaining our reasons for using simulation, we will explain the structure of the simulation program used. In part C, the simulation parameters used in the program are described, together with the reasons for choosing these parameters. Part D illustrates the results of simulation. These results will be used later to compare the performance of the estimators of X_p and the transition phase procedures, and to see whether there were any apparent adverse effects of approximations or assumptions.

A. RATIONALE FOR SIMULATION

The purpose of the study was to find a transition phase procedure (i.e., which estimator of X_p or S_o to use and when) during the transition from Uncertainty to conditions under Risk for the Newsboy problem. As explained before, the performance of an alternative transition phase procedure depends directly on the performance of the candidate estimator of X_p used, in that there is a one-to-one correspondence.

A primary reason for simulation, of course, was to study the question of how long one should continue using the minimax procedure (for Uncertainty) before changing to a transition phase using a quantile estimate. Simulation over several demand distributions should provide insight regarding when (with respect to cost) this change in the Newsboy problem should occur. Additionally, we wish to further examine the question of which of the candidate estimators to use.

However, it was hoped that the estimators of X_p or S_o would come close to having desired order statistics properties, being exact for the case of the uniform distribution (equations (9), (10), (11) and (12)). For example, for the r^{th} order statistic $X_{(r)}$ it has been shown that $F_D(E[X_{(r)}]) = \frac{r}{n+1}$ for a uniform distribution, but $F_D(E[X_{(r)}]) \leq \frac{r}{n+1}$ for convex, and $F_D(E[X_{(r)}]) \geq \frac{r}{n+1}$ for concave cumulative distribution functions of demand. [4]. Unfortunately, there is not an analytical measure to show how much the slack in these inequalities would change depending upon the amount of convexity or concavity of the demand distributions.

Also, for the third candidate estimator of X_p, some additional assumptions were made that might not hold for every distribution.

In summary, our reasons for simulation are:

1. To find the proper time (period) to begin the transition phase (to use one of the candidate estimators), leaving the "Uncertainty" case and the minimax decision rule,

2. To provide information to supplement incomplete analytical knowledge of order statistics regarding population quantiles,

3. To evaluate the impact of the approximations made in developing the third candidate estimator, $\bar{\bar{X}}_{(r)}$, over various demand distributions.

B. DESIGN FOR SIMULATION

The simulation program was written in the FORTRAN IV language. In every simulation run, inventory stocking decisions are made successively using one particular demand distribution and one particular value of the quantile ($P = \frac{C_{so}}{C_c + C_{so}}$), say, demands generated from a normal parent for quantile 0.3 ... etc. The demand distributions and various quantile values used in the simulation are given in the next section. Each run consisted of 50 successive decision periods starting with no demand information other than the range estimate, and replicated 60 times.

As output values of the program, costs for each period were averaged over the number of replications. These average costs per period were calculated for five different cases:

1. The Uncertainty case, with the minimax decision rule being used, denoted by <u>M</u>.

2. The ideal case (Risk), with the quantile value known, denoted by <u>ID</u>.

3. The transition phase case using estimator $X_{(r)}$, denoted by $\underline{R}$.

4. The transition phase case using estimator $\bar{X}_{(r)}$, denoted by $\underline{\bar{\bar{R}}}$.

The minimax (for Uncertainty) and ideal (for Risk) cases would give constant inventory levels, i.e., the same amounts all the way through the simulation run. For Uncertainty, the range of demand was approximately known and would not change unless it was revised, while for Risk conditions it was assumed that the distribution of demand, and thus the quantile, was known.

Transition phase optimum inventory level values generated by $X_{(r)}$, $\bar{X}_{(r)}$ and $\bar{\bar{X}}_{(r)}$ were subject to change from period to period until finally they converge to the value associated with the Ideal case. For the first two periods, $X_{(r)}$, $\bar{X}_{(r)}$, $\bar{\bar{X}}_{(r)}$ would not lead to an inventory level, since they need at least two data points.

The program was run for 50 periods, and replicated 60 times. So, period costs for each transition phase procedure were averaged over 60 replications. Program steps can be ordered in the following way:

1. Define the demand distribution yielding the quantile p, give the stockout cost C_{so}, the carrying cost C_c, and the range of demand.

2. Compute and set minimax and ideal case optimum inventory values.

3. Generate a random demand value from the specified demand distribution.

4. Find the single-period cost for each case (minimax, $X_{(r)}$, $\overline{X}_{(r)}$, $\overline{\overline{X}}_{(r)}$ and Ideal) by comparing their inventory levels and the demand generated for this period. Record these costs. Find the next period inventory levels revising $X_{(r)}$, $\overline{X}_{(r)}$, $\overline{\overline{X}}_{(r)}$ and go to the next period.

5. Repeat steps 3 and 4 for 50 periods.

6. Again, beginning from the first period, repeat step 5 for 60 times (replications), and at the end of 60 replications find the average costs, and the sample variances of costs over 60 replications, for each period.

7. Tabulate the results.

The results obtained from this simulation will be examined in part D.

C. SIMULATION PARAMETERS

In this simulation a variety of possible forms of distributions, including continuous, discrete, symmetric, skewed (long-tailed), etc., were chosen. Demand distributions and their parameters used in this simulation were the following:

Distribution	Parameters
Uniform	the range over [0,30]
Poisson	with mean 10 units
Chi-square	with 10 degrees of freedom
Lognormal	generated by a normal parent having a mean of 3 and a variance of 1
Normal	with mean 35 and variance 100

For D_{max} in the Minimax procedure, $\mu + 3\tau$ is used for tailed distributions. Quantiles $p = \frac{C_{so}}{C_c + C_{so}}$, tested were 0.10, 0.30, 0.50, 0.70, 0.90, 0.95 for each distribution. These were determined so that $C_{so} + C_c$ was constant (in this simulation $C_{so} + C_c = 4.0$ was used). So, for each {quantile, distribution} combination the differences of the expected period costs would come from the performance of the estimators of X_p, as was shown earlier by equation (9).

In the next section we shall present and discuss some of the results from the simulation.

D. SIMULATION RESULTS

In this part, simulation results will be evaluated and, for the Newsboy problem, a transition phase procedure will be proposed. In presenting these results we will identify the quantile value as p. Tables I and II give overall mean costs and standard deviations of cost. Then, Table III will show at what period the transition procedure should begin for a particular distribution, and quantile. Table IV will give relative efficiency of the best transition procedure to the Ideal case (decision under Risk).

On the next page, Table I gives the overall mean values of the cost per period (over 50 periods and 60 replications) for each transition case, using $C_{so} + C_c = 4.00$. As can be seen from Table I, there does not seem a significant difference using R, $\bar{R}$, or $\bar{\bar{R}}$, even though $\bar{R}$ looks quite consistent at the tail ($p = 0.95$). One reason may be the fact that the

TABLE I

Average Costs Over 50 Periods and 60 Replications having $C_c + C_{so} = 4.0$

M=Minimax, $R=X_{(r)}$, $\bar{R}=\bar{X}_{(r)}$, $\bar{\bar{R}}=\bar{\bar{X}}_{(r)}$, ID=Ideal Cases

p		Uniform	Poisson	Chi-Square	Lognormal	Normal
	M	5.49	3.23	3.01	67.00	11.66
	R	6.35	2.38	2.91	14.50	8.09
0.1	$\bar{R}$	6.31	2.38	2.88	14.50	8.09
	$\bar{\bar{R}}$	6.30	2.38	2.87	14.48	8.07
	ID	5.49	2.04	2.56	11.44	6.93
	M	12.68	4.60	5.92	255.96	20.04
	R	13.58	4.55	5.98	36.34	14.25
0.3	$\bar{R}$	13.55	4.52	5.96	36.23	14.18
	$\bar{\bar{R}}$	13.46	4.56	5.94	36.38	14.16
	ID	12.68	4.26	5.62	30.15	13.17
	M	15.08	5.97	11.08	339.86	15.71
	R	16.03	5.35	7.47	51.22	15.80
0.5	$\bar{R}$	15.97	5.33	7.45	51.37	15.78
	$\bar{\bar{R}}$	15.91	5.34	7.45	51.35	15.73
	ID	15.08	4.98	6.98	43.49	14.80
	M	12.75	7.33	11.81	298.67	14.97
	R	13.75	4.74	7.21	57.24	14.24
0.7	$\bar{R}$	13.66	4.74	7.21	57.23	14.19
	$\bar{\bar{R}}$	13.60	4.72	7.20	57.66	14.19
	ID	12.75	4.44	6.67	48.72	13.24
	M	5.39	4.37	6.26	132.42	9.33
	R	6.31	2.68	4.40	42.36	7.76
0.9	$\bar{R}$	6.30	2.67	4.38	42.38	7.76
	$\bar{\bar{R}}$	6.30	2.66	4.36	42.58	7.73
	ID	5.39	2.33	3.71	35.91	6.70
	M	2.83	2.38	3.34	70.42	5.27
	R	3.77	1.75	3.00	31.80	5.04
0.95	$\bar{R}$	3.76	1.75	2.99	31.74	5.02
	$\bar{\bar{R}}$	3.76	1.75	3.00	32.44	5.04
	ID	2.83	1.41	2.27	25.47	3.91

difference between variance of $X_{(r)}$ and variance of $\bar{\bar{X}}_{(r)}$, was not much (Figure 3). But if one wishes to see the difference no matter how much it is, the following normalization by ideal costs may be considered:

1. For a distribution from Table I, for every quantile $\underline{p}$ find $\frac{R_i - ID}{ID}$, and sum this value over all quantiles $\underline{p}$ for every R_i, where R_i is R, $\bar{R}$ or $\bar{\bar{R}}$.

2. Repeat step 1 for every distribution.

3. Pick the R_i's which have the smallest sum among R, $\bar{R}$, $\bar{\bar{R}}$ for every distribution as their best procedure.

The results of this normalization were as follows:

	Uniform	Poisson	Chi-square	Lognormal	Normal
R	0.871	0.767	0.859	1.253	0.839
$\bar{R}$	0.845	0.752	0.831	1.251	0.824
$\bar{\bar{R}}$	0.828	0.755	0.821	1.295	0.816

The best transition procedures were to use $\bar{\bar{R}}$ for uniform, $\bar{R}$ for poisson, $\bar{\bar{R}}$ for chi-square, $\bar{R}$ for lognormal, and $\bar{\bar{R}}$ for normal distributions. To see the effects of the very right tail of a distribution on transition procedures, $p = 0.99$ was used for normal and chi-square distributions. At this quantile the minimax costs were below R, $\bar{R}$, and $\bar{\bar{R}}$ costs. This says minimax can be used at the very tail of a distribution along the transition phase, providing very accurate prediction of the range of demand (say $\pm 5\%$). Otherwise, if we remember from Table I for $p = 0.95$, R, $\bar{R}$ and $\bar{\bar{R}}$ were better than minimax approach.

From Table II (overall standard deviations over 50 periods and 60 replications), it is seen again that there is not much difference of standard deviations, even though $\bar{\bar{R}}$ has a better trend (smaller sample variance of cost) among R, $\bar{R}$ and $\bar{\bar{R}}$.

From these two tables it can be said that there is not much difference of the costs yielded using R, $\bar{R}$ and $\bar{\bar{R}}$, but if an order is necessary it should be the one using $\bar{\bar{R}}$ for p less than 0.95, and $\bar{R}$ at the tail (i.e., when p is greater than 0.95).

During the simulation, as the number of periods increase the estimators converge to X_p or the true S_o, while the minimax and ideal case inventory levels were constant. Since the values of $X_{(r)}$, $\bar{X}_{(r)}$, $\bar{\bar{X}}_{(r)}$ usually change from period to period because they are updated by the most recent demand value, they will generally be associated with a larger variance of cost than the minimix and ideal cases. This is reflected in Table II. Sometimes the minimax procedure had a smaller sample variance of cost (while possibly yielding the largest average cost) than the ideal case. This may be explained as follows:

The ideal solution we are using relates to minimizing expected cost, not the variance of cost. These two minimization points are generally different, depending on the distribution function. If the minimax solution is different from the ideal (expected cost) inventory level and closer to the minimum variance solution, it would yield a smaller variance of cost. Further derivations about the variance

TABLE II

Sample Standard Deviations over 50 Periods and 60 Replications

M=Minimax, $R=X_{(r)}$, $\bar{R}=\bar{X}_{(r)}$, $\bar{\bar{R}}=\bar{\bar{X}}_{(r)}$, ID=Ideal Cases

p		Uniform	Poisson	Chi-Square	Lognormal	Normal
	M	3.15	1.26	1.79	43.46	3.77
	R	6.36	2.69	3.27	27.46	9.89
0.1	$\bar{R}$	6.26	2.65	3.21	27.47	9.70
	$\bar{\bar{R}}$	6.26	2.59	3.18	27.47	9.53
	ID	3.15	1.62	2.26	15.90	6.62
	M	7.46	3.18	4.61	76.10	10.33
	R	9.00	3.60	4.93	57.64	12.17
0.3	$\bar{R}$	8.84	3.53	4.89	57.69	11.97
	$\bar{\bar{R}}$	8.85	3.51	4.86	57.71	11.99
	ID	7.46	3.16	4.42	45.93	10.95
	M	8.82	4.30	6.35	66.95	12.34
	R	10.31	4.14	6.37	84.29	12.73
0.5	$\bar{R}$	10.23	4.12	6.33	83.91	12.68
	$\bar{\bar{R}}$	10.23	4.14	6.44	84.23	12.75
	ID	8.82	3.84	5.80	72.72	12.09
	M	7.30	3.47	4.87	46.79	9.62
	R	9.08	3.95	6.66	99.68	11.65
0.7	$\bar{R}$	9.01	4.04	6.83	99.68	11.65
	$\bar{\bar{R}}$	8.87	4.09	6.76	99.23	11.67
	ID	7.30	3.55	5.66	92.87	10.42
	M	3.16	1.25	1.87	32.45	3.90
	R	6.59	3.28	5.72	99.99	9.10
0.9	$\bar{R}$	6.55	3.33	5.76	99.23	9.09
	$\bar{\bar{R}}$	6.55	3.29	5.69	98.05	9.01
	ID	3.19	2.57	4.10	94.42	6.34
	M	1.66	0.69	1.12	31.66	2.07
	R	6.50	2.91	5.36	94.64	8.50
0.95	$\bar{R}$	6.50	2.95	5.34	92.85	8.47
	$\bar{\bar{R}}$	6.50	2.94	5.23	92.74	8.46
	ID	1.66	1.49	3.15	84.97	4.11

minimization solutions of the Newsboy problem are given in Appendix B.

The simulation results about the period to begin a transition phase are given in Table III. Shown are the first periods in which the average period cost (over 60 replications) using the estimator was below that obtained using the minimax rule. In the simulation it was observed that once costs using an estimator fell below the minimax costs, they stayed below for the remaining time periods. Simulation results showed that, on the average, at the fifth period we were supposed to begin a transition procedure, ceasing to use the minimax procedure. In other words, for the first five periods the large variability of candidate estimators of S_o (or X_p) caused a larger average cost than the minimax procedure yielded. As a very special case, if the distribution of demand is uniform, minimax and Ideal case solutions would match each other and thus the uniform distribution is omitted from Table III. Also by coincidence, if $p = \frac{C_{so}}{C_c + C_{so}}$ is at or close to an intersection of the implied uniform CDF of minimax procedure and the actual distribution function of demand, then the minimax procedure may yield the Ideal case solution. Otherwise, minimax inventory level has always a higher cost than the Ideal case. In Table III the large numbers are those specific values which relate to a quantile P that is close to minimax--actual CDF intersection point. So it takes time for R, $\bar{R}$ and $\bar{\bar{R}}$ to yield costs less than the minimax. No matter how

TABLE III

Periods to Begin Transition Phase
Leaving the Minimax Procedure

Quantile (p)	Poisson	Chi-Square	Lognormal	Normal
0.1	6	6	5	6
0.3	10*	10*	4	4
0.5	6	4	4	10*
0.7	4	4	4	6
0.9	5	5	4	5
0.95	6	5	4	6

* Ideal-Minimax intercept points (or, close to these points).

long it will take to fall below the minimax cost (here, nearly at period ten) from simulation outputs it was observed that after the fifth or sixth period the cost difference between the minimax and the transition procedures was not large. This also can be predicted from the pattern values of Table III. The numbers in Table III are relevant for all candidate estimators, i.e., R, $\bar{R}$ and $\bar{\bar{R}}$ all fell below the minimax cost simultaneously at the same period.

As a measure of efficiency for a transition procedure, the ratio of average cost induced by this procedure to Ideal average cost is used. Table IV gives these ratios for the average costs over second and fifth 10 periods for $\bar{\bar{R}}$ (since $\bar{\bar{R}}$ seems the promising transition procedure among the other ones).

We will examine the data of Table IV for the second 10 period, assuming that the minimax has been left and the transition phase has begun. We also believe that convergence toward the true value of Ideal inventory level S_o for the third, fourth, ... 10 periods will not be worse than the convergence accomplished during the second 10 period.

If the transition procedure (here, $\bar{\bar{R}}$) was operated at the tail (i.e., $p \geq 0.9$), then the efficiency ratio for the second 10 periods was on the average 1.16 and more (worse). As the quantile gets close to 1.0, the minimax procedure was yielding much better results than a transition procedure (assuming the range of demand is predicted quite well).

TABLE IV

The Ratio of the Second and Fifth Ten-Period (11 to 20, 41 to 40) Average Costs to the Ideal Case (Under Risk) Costs, Using $\bar{X}_{(r)}$ Transition Procedure

Quantile (p)		Uniform	Poisson	Chi-Square	Lognormal	Normal
0.1	S	1.063	1.065	1.058	1.033	1.108
	F	1.000	1.023	1.046	1.002	1.008
0.3	S	1.065	1.059	1.050	1.040	1.066
	F	1.007	1.041	1.003	1.004	1.008
0.5	S	1.056	1.069	1.039	1.041	1.077
	F	1.013	1.021	1.013	1.004	1.007
0.7	S	1.073	1.069	1.061	1.057	1.054
	F	1.000	1.001	1.006	1.006	1.004
0.9	S	1.102	1.113	1.128	1.190	1.096
	F	1.008	1.016	1.052	1.026	1.027
0.95	S	1.170	1.193	1.211	1.240	1.183
	F	1.002	1.040	1.054	1.062	1.055

Notes:

S stands for "SECOND," F stands for "FIFTH."

When p is less than 0.9 overall ratio is 1.060 for second and 1.010 for fifth 10 periods.

For the rest of the quantiles where p is less than 0.9, this ratio for the transition procedure was almost constant 1.060. To test the effects of quantiles and the effects of the form of demand distributions on the efficiency ratio (1.060), a two-way Analysis of Variance could be applied to the data of Table IV for quantiles less than 0.9.

Different distributions were significant at a level 0.03, and the different quantiles were significant at a level of 0.01. This result was explained that for $C_{so} + C_c = 4.0$ cost units, whatever the distribution of demand was and for any quantile less than 0.9, the ratio of the cost induced by transition phase procedure ($\overline{\overline{X}}_{(r)}$, here) to the Ideal cost (expected cost under Risk) would be practically constant at 1.060 for second 10 period, and would not exceed this amount for the further periods (for example, at the fifth 10 period this ratio was 1.01).

On the next part, as the summary of this chapter, and using the results obtained in it, a Single-Period inventory (Newsboy) problem transition phase procedure is suggested.

E. PROPOSED NEWSBOY TRANSITION PROCEDURE ALGORITHM

(1) Approximate the bounds of Demand (Range).

(2) Set $p = \dfrac{C_{so}}{C_c + C_{so}}$.

(3) If $p \geq 0.90$, go to step 6, otherwise go to step 4.

(4) Use minimax for the first 5 or 6 periods.

(5) For $6 \leq n \leq 50$, use $\overline{\overline{R}}$ (unless a reason for R, $\overline{R}$), and go to step 10.

(6) If $p \leq 0.95$, go to step 8, otherwise go to step 7.

(7) If step 1 is accurage, may use minimax for all periods and go to step 10, but if it is not accurate, go to step 8.

(8) Use minimax for the first 5 or 6 periods, then go to step 9.

(9) For $6 \leq n \leq 50$, use $\bar{R}$ and go to 10.

(10) Analyze data.

*May fit a probability distribution to the data and begin Decision Under Risk case.

---

*May use $\bar{\bar{R}}$, $\bar{R}$ or R for n > 50 too. Since they converge X_p quite fast, the risk of using them after 50 periods hopefully will be nearly equal to the risk of Ideal case.

V. CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY

For the Newsboy problem, a nonparametric decision procedure that allows transition from Uncertainty to decision under Risk has been developed. Three estimators of the demand corresponding to the p^{th} quantile were obtained, where $p = \frac{C_{so}}{C_c + C_{so}}$. Testing these estimators over several distributions of demand with Newsboy cost criteria, it was found that excluding cases such as drastic skewness, pooling the (r-1)st and (r+1)st order statistics for $\bar{\bar{R}}$ provided essentially the best estimator, where R and $\bar{R}$ also worked well.

It was also found that for the distributions tested and excluding $(\frac{C_{so}}{C_{so} + C_c})$ values very close to 0.0 or 1.0, that the switching Newsboy decision making from minimax solutions to quantile estimator could be made, possibly at the sixth period.

For further work, some parametric studies such as log-normal distribution, or beta and gamma families, depending upon whether the population of demand is finite or not, may be tried. Then, given that the decision maker does not know what the distribution of demand is, which procedure, nonparametric or parametric, gives better response to Uncertainty may be compared. As another recommendation, some work can be afforded to improve nonparametric approaches further.

APPENDIX A

ESTIMATION FOR POPULATION QUANTILES BY ORDER STATISTICS

In this Appendix order statistics, especially those concerning population quantiles, will be briefly reviewed. For a further study of order statistics, references [1], [2] and [3] are cited.

The Joint Distribution of n Order Statistics

Order statistics theory is mainly based on continuous distribution functions. In reference [2] for discrete parents the joint distribution of order statistics is derived as well.

Let $X_1, X_2, \ldots, X_n$ be a random sample from a probability distribution function F_X. If $X_{(i)}, i=1,2,\ldots,n$ is the i^{th} largest value in this sample, then $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ are the order statistics of a random sample $X_1, X_2, \ldots, X_n$. Here, $X_{(r)}$ is called the r^{th} order statistic. Since F_X is a distribution function, by the probability integral transformation $U = F_X(x)$ will have uniform distribution over [0,1], i.e.,

$$F_U(u) = \begin{cases} u, & \text{for } 0 \leq u \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

Letting $F_X(x) = u$ with $F_X(X) \leq u$ if and only if $X \leq x$, the following equations hold:

$$F_U(u) = P(U \leq u) = P\{F_X(X) \leq u\} = P[X \leq F_X^{-1}(u)] = F_X[F_X^{-1}(u)] = u \tag{16}$$

Since $F_X(X)$ is a monotone increasing function and by equation (16) there is a one-to-one correspondence between $F_X(X_{(1)}) < F_X(X_{(2)}) < \ldots < F_X(X_{(n)})$ and $X_{(1)} < X_{(2)} \ldots < X_{(n)}$. Let f_X be the density of F_X, so that

$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n)$ is the joint density of the sample $X_1, X_2, \ldots, X_n$. If each X_i, $i=1,2,\ldots,n$ is independent and identically distributed, then

$$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_X(x_i).$$

Since there exists n! permutation of those sample values,

$$f_{X_{(1)}, X_{(2)}, \ldots, X_{(n)}}(y_i, y_2, \ldots, y_n) = n! \prod_{i=1}^{n} f_X(y_i) \ , \qquad (17)$$

is the joint density of n order statistics.

From the previous references cited:

$$f_{X_{(r)}}(y_r) = \frac{n!}{(r-1)!(n-r)!} [F_X(y_r)]^{r-1} [1-F_X(y_r)]^{n-r} f_X(y_r), \qquad (18)$$

When F_X is a uniform distribution over (0,1), this equation turns to be,

$$f_{X_{(r)}}(n) = \frac{n!}{(r-1)!(n-r)!} [u]^{r-1} [1-u]^{n-r} \ , \qquad (19)$$

and the following sequence of equations are given,

$$E[X_{(r)}{}^k] = \frac{n!}{(r-1)!(n-r)!} \int_0^1 u^{r+k-i} (1-u)^{n-r} du$$

and

$$E[X_{(r)}^k] = \frac{(r+k-1)(r+k-2)....(r+1)r}{(n+k)(n+k-1)....(n+2)(n+1)} , \quad (20)$$

where k is a positive integer. For $k = 1$:

$$E[X_{(r)}] = \frac{r}{n+1} . \quad (21)$$

By equations (20) and (21)

$$Var[X_{(r)}] = E[X_{(r)}^2] - [E(X_{(r)})]^2 = \frac{r(n-r+1)}{(n+2)(n+1)^2} , \quad (22)$$

and,

$$E[X_{(r)} - X_{(s)}] = \frac{r(s+1)}{(n+1)(n+2)} , \quad (23)$$

From equations (6) and (8),

$$Cov[X_{(r)}, X_{(s)}] = E[X_{(r)}X_{(s)}] - E[X_{(r)}]E[X_{(s)}], \text{ and}$$

$$Cov\ [X_{(r)}, X_{(s)}] = \frac{r(n-s+1)}{(n+1)^2\,(n+2)} . \quad (24)$$

Now, we wish a transition from the uniform distribution to a general distribution F_X. Using the relations in equation (16),

$$U(r) = F_X(X_{(r)}) , \quad \text{or} \quad X_{(r)} = F_X^{-1}(U_{(r)}) ,$$

where $U_{(r)}$ is the r^{th} order statistic from Uniform (0,1).

Then, $E[X_{(r)}] = E[F_X^{-1}(U_{(r)})] \approx F_X^{-1}[E(U_{(r)})] = F_X^{-1}(\frac{r}{n+1})$.

(25)

Furthermore, the references given at the beginning of this appendix show that, for continuous distributions having

$n_j = [np_j+1]$, and $0 < p_1 < p_2 \quad \ldots < p_\ell < 1$,

the joint distribution of $X_{(n_1)}, X_{(n_2)}, \ldots X_{(n\ell)}$ asymtotically tends to an ℓ-dimensional normal distribution with means $p_1, p_2, \ldots, p_\ell$, and covariances,

$$Cov[X_{(n_i)}, X_{(n_k)}] = \frac{p_i(1-p_k)}{n\, f(X_{p_i})\, f(X_{p_k})} \text{, provided } i \leq k, \tag{26}$$

So the variance of $X_{(n_i)}$ will be,

$$Var[X_{(n_i)}] = \frac{p_i(1-p_i)}{n\{f(X_{p_i})\}^2}$$

Having $n_i = r$, it can be proven that,

$$U = \frac{X_{(n_i)} - X_{p_i}}{\sqrt{Var(X_{(n_i)})}} = \frac{\sqrt{n}\ (X_{(r)} - X_{p_i})}{\sqrt{p_i(1-p_i)}}\ f_X(X_{p_i}) \approx N(0,1) \tag{27}$$

Equation (27) is quite important for this study. It says that, whatever the distribution function is, the r^{th} order statistic, $X_{(r)}$, is approximately normally distributed around X_p (distribution free property). This also means that asymtotically $X_{(r)}$ is an unbiased estimator of X_p.

Let $G_D(X)$ be the empirical distribution function of order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ having,

$$G_D(X) = \begin{cases} 0 \text{ , if } X < X_{(1)} \\ \frac{j}{n} \text{ , if } X_{(j)} \leq X < X_{(j+1)} \\ 1 \text{ , if } X \geq X_{(n)} \end{cases}$$

In Reference [3] it is said, by the Glivenko-Centelli theorem that, $G_D(X)$ converges uniformly to $F_X(X)$ and $nG_D(X)$ is binomial $(n, F_X(X))$, i.e.,

$$p[nG_D(X)=j] = p[G_D(X) = \frac{j}{n}] = \binom{n}{j}[F_X(X)]^j[1-F_X(X)]^{n-j}, \tag{28}$$

Point and Interval Estimation of Population Quantiles

For point estimation of a population quantile, we assume that X_p is the p^{th} quantile of the distribution function F_X, and $P(X < X_p) = F_X(X_p) = p$, or $F_X^{-1}(p) = X_p$.

Then, by equation (27) the r^{th} order statistic is a consistent and unbiased estimator of the p^{th} quantile, where

$$r = \begin{cases} np & , \text{ if } np \text{ is integer} \\ [np+1] & , \text{ otherwise} \end{cases} , \text{ and } p = \frac{C_{so}}{C_{so}+C_c} .$$

For interval estimation of population quantiles, we are looking for r and s such that,

$$P(X_{(r)} < X_p < X_{(s)}) = \gamma \text{ , where } \gamma \text{ is percent confidence.}$$

Now,

$$P(X_{(r)} < X_p < X_{(s)}) = P(X_{(r)} < X_p) - P(X_{(s)} < X_p), \text{ or}$$

by equation (16)

$$P(X_{(r)} < X_p < X_{(s)}) = P(F_X(X_{(r)}) < p) - P(F_X(X_{(s)}) < p).$$

Since $p = F_X(X_p)$, by equation (28)

$$P(X_{(r)} < X_p < X_{(s)}) = \sum_{i=r}^{n} \binom{n}{i} p^i (1-p)^{n-1} - \sum_{i=s}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

$$= \sum_{r}^{s-1} \binom{n}{i} p^i (1-p)^{n-i} = \gamma \tag{29}$$

To summarize these derivations:

1. For large n the r^{th} order statistic is a consistent, unbiased point estimator of the p^{th} quantile for any distribution.

2. Confidence intervals for the p^{th} quantile as given by equation (29), distributed by the Binomial distribution (n,p).

3. As n gets larger, $X_{(r)}$ converges to X_p and also the empirical distribution function $G_D(X)$ converges to actual cumulative distribution function $F_X(X)$.

APPENDIX B

EXPECTED COST AND VARIANCE MINIMIZATION POINTS OF THE NEWSBOY PROBLEM

In this Appendix we will attempt to explain the difference between minimization of the expected cost, and minimization of the variance of the cost, for the Newsboy problem.

Decision makers' utility functions are involved with both expected cost (or return), and the variance of this cost. If a decision maker is risk prone, the variance of the cost may not affect him, and he just tries to minimize expected cost. On the other hand, if he is risk adverse, then he may give up some of the cost minimization value for a smaller variance (a trade-off between expected cost and variance of cost). As an example [5], let

$U(x) = x - bx^2$, be a utility function. Then,

$$E[U(x)] = E[x] - bE[x^2] = E[x] - b(Var(x) + [E(x)]^2).$$

If we try to optimize this value, we not only deal with expected value of x, but also variance of x.

On the following pages, the variance minimization points of some distributions related with the Newsboy problem are derived.

Let $Var(C) \triangleq E[C^2] - E[C]^2$, and

$$E[C] = C_c \int_0^S (S-x) f_D(x) dx + C_{so} \int_S^\infty (x-S) f_D(n) dx$$

$$E[C^2] = C_c^2 \int_0^S (S-x)^2 f_D(n) dx + C_{so}^2 \int_S^\infty (x-S)^2 f_D(n) dx.$$

To minimize Var(C), with respect to S:

$$\frac{dVar(C)}{d\,S} = 0, \text{ is necessary.}$$

So,

$$\frac{dVar(C)}{d\,S} = \frac{dE[C^2]}{d\,S} - 2E[C] \frac{dE[C]}{d\,S} = 0, \tag{30}$$

Then,

$$\frac{dE[C^2]}{d\,S} = 2C_c^2 \int_0^S (S-x) f_D(x) - 2C_{so}^2 \int_S^\infty (x-S) f_D(x) dx$$

$$- 2E[C] \frac{dE[C]}{dS} = - 2E[C] \{C_c \int_0^S f_C(x) dx - C_{so} \int_S^\infty f_D(x) dx\}$$

$$= - 2E[C]\{C_c F_D(S) - C_{so}(1-F_D(S))\}$$

Since, $\left.\frac{dE[C^2]}{d\,S}\right|_{S=S_o} = 2E[C] \left.\frac{dE[C]}{d\,S}\right|_{S=S_o}$,

then it yields

$$C_c^2\{SF(S) - \int_0^S x\, f_D(x) dx\} - C_{so}^2\{\int_S^\infty x\, f_D(x) dx - S(1-F_D(S))\} =$$

$$(C_c\{S\,F(S) - \int_0^S x\, f_D(x) dx\} + C_{so}\{\int_S^\infty x\, f_D(x) dx - S(1-F_D(S))\})$$

$$(F_D(S)(C_c + C_{so}) - C_{so}). \tag{31}$$

When this is applied to a uniform distribution over $(0, D_{max})$, i.e., $f_D(x) = \frac{1}{D_{max}}$, $0 \leq x \leq D_{max}$ and $F_D(S) = \frac{S}{D_{max}}$, then, the left hand side (LHS) of equation gives,

$$LHS = S^2(C_c^2 - C_{so}^2)\frac{1}{2D_{max}} + C_{so}^2 S - C_{so}^2 \frac{D_{max}}{2},$$

Similarly, right hand side (RHS) of it yields,

$$RHS = ((\frac{C_c+C_{so}}{D_{max}})S - C_{so})(\frac{S^2}{2D_{max}}(C_c + C_{so}) - C_{so}S + \frac{C_{so}D_{max}}{2}).$$

After we equate LHS - RHS and opened them, it was found that,

$$S_o = \frac{C_{so}\ D_{max}}{C_c + C_{so}}, \qquad (32)$$

is the point that minimizes the variance of the cost for the uniform distribution. This point is exactly the same point that minimizes expected cost.

Let us apply the same procedure to a bell-shape, say, Beta ($\alpha = 2$, $\beta = 2$). Then

$$f_D(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} = 6x(1-x) \text{ for } 0 \leq x \leq 1,$$

$$F_D(x) = \int_0^x 6u(1-u)du = 3x^2 - 2x^3, \text{ and}$$

$$1 - F_D(x) = 1 - 3x^2 + 2x^3, \quad \int_x^{1.0} u\ f_D(u)du = \tfrac{1}{2}(1-4x^3 + 3x^4),$$

$$\int_0^x u\ f_D(u)du = \tfrac{1}{2}(4x^3 - 3x^4).$$

Putting those values in equation (2), gives a 5 degree polynomial:

$$(C_c + C_{so})^2 S^5 - \frac{7}{2}(C_c + C_{so})^2 S^4 + 3(C_c + C_{so})^2 S^3 +$$

$$+ \frac{1}{2}(C_c^2 + C_{so}^2 + 5C_c C_{so})S^2 - (4C_{so}^2 + 5C_c C_{so})S + \frac{3}{2}(C_c C_{so} + 2C_{so}^2) = 0 \tag{33}$$

It may be difficult to solve this equation, but it is not difficult to see that it doesn't have the same solution with $\frac{C_{so}}{C_c+C_{so}} = 3S_o^2 - 2S_o^3$, or $3S_o^2 - 2S_o^3 - p = 0$, which is the minimum expected cost solution.

In the exponential $\{\lambda\}$ case,

$f_D(x) = \lambda e^{-\lambda x}$, $F_D(x) = 1-e^{-\lambda x}$, $1-F_D(x) = e^{-\lambda x}$ and

$$\int_0^x u\ f(u)du = \frac{1}{\lambda}[1-e^{-\lambda x}(1+ x)],$$

$$\int_x^\infty u\ f(u)du = \frac{1}{\lambda}[e^{-\lambda x}(1+\lambda x)].$$

Again, inserting these values into equation (2) gives,

$$\frac{F_D(S)}{S} = \frac{C_c\lambda}{C_c+C_{so}}, \tag{34}$$

where $F_D(S) = 1-e^{-\lambda S}$.

If we compare (34) with $F_D(S_o) = \frac{C_{so}}{C_c+C_{so}} = 1-e^{-\lambda S_o}$ or $S_o = -\frac{1}{\lambda}\ln\left(\frac{C_c}{C_c+C_{so}}\right)$, we will see that they are different points. If we examine equation (34) further, it can be shown that $S \geq S_o$, i.e., the variance minimization point S is closer to the tail than the minimum expected cost point S_o.

LIST OF REFERENCES

1. Introduction to Mathematical Statistics, Allen T. Craig amd Robert V. Hogg, The Macmillan Company, 1970.

2. Order Statistics, H. A. David, Wiley, 1970.

3. Nonparametric Statistical Inference, T. D. Gibbons, McGraw-Hill, 1971.

4. Analysis of Inventory Systems, G. Hadley and T. M. Whitin, Prentice-Hall, 1963.

5. Decisions with Multiple Objectives, R. L. Keeney and H. Raiffa, Wiley.

6. The Single Period Inventory Model, T. Masuda, M.S. Thesis, Naval Postgraduate School, Monterey, CA, September 1977.

7. The Analysis of Management Decisions, W. T. Morris and Richard D. Irwin, 1964.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 55 Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
4.	Professor G. F. Lindsay, Code 55 Ls Department of Operations Research Naval Postgraduate School Monterey, California 93940	2
5.	Professor D. P. Gaver, Code 55 Gv Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
6.	Lt. Kadir Sagdic, Turkish Navy Fatih Mah. Cami Sokak No. 5/3 Buyuk Cekmece - Istanbul Turkey	2
7.	Deniz Kuvvetteri Komutanligi Bakanliklar, Ankara Turkey	2
8.	Deniz Harp Okulu Komutanligi Heybeliada, Istanbul Turkey	1
9.	Orta-Dogu Teknik Universitesi Yon-Eylem Arastirmasi Bolumu Ankara Turkey	1