We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Stochastic Control and Improvement of Statistical Decisions in Revenue Optimization Systems

Nicholas A. Nechval and Maris Purgailis

Additional information is available at the end of the chapter

1. Introduction

A large number of problems in production planning and scheduling, location, transportation, finance, and engineering design require that decisions be made in the presence of uncertainty. From the very beginning of the application of optimization to these problems, it was recognized that analysts of natural and technological systems are almost always confronted with uncertainty. Uncertainty, for instance, governs the prices of fuels, the availability of electricity, and the demand for chemicals. A key difficulty in optimization under uncertainty is in dealing with an uncertainty space that is huge and frequently leads to very large-scale optimization models. Decision-making under uncertainty is often further complicated by the presence of integer decision variables to model logical and other discrete decisions in a multi-period or multi-stage setting.

Approaches to optimization under uncertainty have followed a variety of modeling philosophies, including expectation minimization, minimization of deviations from goals, minimization of maximum costs, and optimization over soft constraints. The main approaches to optimization under uncertainty are stochastic programming (recourse models, robust stochastic programming, and probabilistic models), fuzzy programming (flexible and possibilistic programming), and stochastic dynamic programming.

This paper is devoted to improvement of statistical decisions in revenue management systems. Revenue optimization – or revenue management as it is also called – is a relatively new field currently receiving much attention of researchers and practitioners. It focuses on how a firm should set and update pricing and product availability decisions across its various selling channels in order to maximize its profitability. The most familiar example probably comes from the airline industry, where tickets for the same flight may be sold at

many different fares throughout the booking horizon depending on product restrictions as well as the remaining time until departure and the number of unsold seats. Since the tickets for a flight have to be sold before the plane takes off, the product is perishable and cannot be stored for future use. The use of the above strategies has transformed the transportation and hospitality industries, and has become increasingly important in retail, telecommunications, entertainment, financial services, health care and manufacturing. In parallel, pricing and revenue optimization has become a rapidly expanding practice in consulting services, and a growing area of software and IT development, where the revenue optimization systems are tightly integrated in the existing Supply Chain Management solutions.

Most stochastic models, which are used in revenue optimization systems, are developed under the assumptions that the parameter values of the models are known with certainty. When these models are applied to solve real-world problems, the parameters are estimated and then treated as if they were the true values. The risk associated with using estimates rather than the true parameters is called estimation risk and is often ignored. When data are limited and/or unreliable, estimation risk may be significant, and failure to incorporate it into the model design may lead to serious errors. Its explicit consideration is important since decision rules which are optimal in the absence of uncertainty need not even be approximately optimal in the presence of such uncertainty.

In this paper, we consider the cases where it is known that the underlying probability distributions belong to a parameterized family of distributions. However, unlike in the Bayesian approach, we do not assume any prior knowledge on the parameter values. The primary purpose of the paper is to introduce the idea of embedding of sample statistics (say, sufficient statistics or maximum likelihood estimators) in a performance index of revenue optimization problem. In this case, we find the optimal stochastic control policy directly. We demonstrate the fact that the traditional approach, which separates the estimation and the optimization tasks in revenue optimization systems (i.e., when we use the estimates as if they were the true parameters) can often lead to pure results. It will be noted that the optimal statistical decision rules depend on data availability.

For constructing the improved statistical decisions, a new technique of invariant embedding of sample statistics in a performance index is proposed (Nechval et al., 1999; 2004; 2008; 2010a; 2010b; 2010c; 2010d; 2010e; 2011a; 2011b). This technique represents a simple and computationally attractive statistical method based on the constructive use of the invariance principle in mathematical statistics. Unlike the Bayesian approach, an invariant embedding technique is independent of the choice of priors, i.e., subjectivity of investigator is eliminated from the problem. The technique allows one to eliminate unknown parameters from the problem and to find the improved invariant statistical decision rules, which has smaller risk than any of the well-known traditional statistical decision rules.

In order to obtain improved statistical decisions for revenue management under parametric uncertainty, it can be considered the three prediction situations: "new-sample" prediction, "within-sample" prediction, and "new-within-sample" prediction. For the new-sample prediction situation, the data from a past sample are used to make predictions on a future unit or sample of units from the same process or population. For the within-sample prediction

situation, the problem is to predict future events in a sample or process based on early data from that sample or process. For the new-within-sample prediction situation, the problem is to predict future events in a sample or process based on early data from that sample or process as well as on a past data sample from the same process or population. Some mathematical preliminaries for the within-sample prediction situation are given below.

2. Mathematical preliminaries for the within-sample prediction situation

Theorem 1. Let $X_1 \leq ... \leq X_k$ be the first k ordered observations (order statistics) in a sample of size m from a continuous distribution with some probability density function $f_\theta(x)$ and distribution function $F_\theta(x)$, where θ is a parameter (in general, vector). Then the joint probability density function of $X_1 \leq ... \leq X_k$ and the *l*th order statistics X_l ($1 \leq k < l \leq m$) is given by

$$f_\theta(x_1, ..., x_k, x_l) = f_\theta(x_1, ..., x_k)g_\theta(x_l \mid x_k), \tag{1}$$

where

$$f_\theta(x_1, ..., x_k) = \frac{m!}{(m-k)!}\prod_{i=1}^{k} f_\theta(x_i)[1 - F_\theta(x_k)]^{m-k}, \tag{2}$$

$$g_\theta(x_l \mid x_k) = \frac{(m-k)!}{(l-k-1)!(m-l)!}\left[\frac{F_\theta(x_l) - F_\theta(x_k)}{1 - F_\theta(x_k)}\right]^{l-k-1}\left[1 - \frac{F_\theta(x_l) - F_\theta(x_k)}{1 - F_\theta(x_k)}\right]^{m-l}\frac{f_\theta(x_l)}{1 - F_\theta(x_k)}$$

$$= \frac{(m-k)!}{(l-k-1)!(m-l)!}\sum_{j=0}^{l-k-1}\binom{l-k-1}{j}(-1)^j\left[\frac{1 - F_\theta(x_l)}{1 - F_\theta(x_k)}\right]^{m-l+j}\frac{f_\theta(x_l)}{1 - F_\theta(x_k)}$$

$$= \frac{(m-k)!}{(l-k-1)!(m-l)!}\sum_{j=0}^{m-l}\binom{m-l}{j}(-1)^j\left[\frac{F_\theta(x_l) - F_\theta(x_k)}{1 - F_\theta(x_k)}\right]^{l-k-1+j}\frac{f_\theta(x_l)}{1 - F_\theta(x_k)} \tag{3}$$

represents the conditional probability density function of X_l given $X_k = x_k$.

Proof. The joint density of $X_1 \leq ... \leq X_k$ and X_l is given by

$$f_\theta(x_1, ..., x_k, x_l) = \frac{m!}{(l-k-1)!(m-l)!}\prod_{i=1}^{k} f_\theta(x_i)[F_\theta(x_l) - F_\theta(x_k)]^{l-k-1}f_\theta(x_l)[1 - F_\theta(x_l)]^{m-l}$$

$$= f_\theta(x_1, ..., x_k)g_\theta(x_l \mid x_k). \tag{4}$$

It follows from (4) that

$$f_\theta(x_l \mid x_1, ..., x_k) = \frac{f_\theta(x_1, ..., x_k, x_l)}{f_\theta(x_1, ..., x_k)} = g_\theta(x_l \mid x_k), \tag{5}$$

i.e., the conditional distribution of X_l, given $X_i = x_i$ for all $i = 1,\ldots, k$, is the same as the conditional distribution of X_l, given only $X_k = x_k$. This ends the proof. $\square$

Corollary 1.1. The conditional probability distribution function of X_l given $X_k=x_k$ is

$$P_\theta\{X_l \leq x_l \mid X_k = x_k\} = 1 - \frac{(m-k)!}{(l-k-1)!(m-l)!}$$

$$\times \sum_{j=0}^{l-k-1}\binom{l-k-1}{j}\frac{(-1)^j}{m-l+1+j}\left[\frac{1-F_\theta(x_l)}{1-F_\theta(x_k)}\right]^{m-l+1+j}$$

$$= \frac{(m-k)!}{(l-k-1)!(m-l)!}\sum_{j=0}^{m-l}\binom{m-l}{j}\frac{(-1)^j}{l-k+j}\left[\frac{F_\theta(x_l)-F_\theta(x_k)}{1-F_\theta(x_k)}\right]^{l-k+j}. \tag{6}$$

Corollary 1.2. If $l = k + 1$,

$$g_\theta(x_{k+1} \mid x_k) = (m-k)\left[\frac{1-F_\theta(x_{k+1})}{1-F_\theta(x_k)}\right]^{m-k-1}\frac{f_\theta(x_{k+1})}{1-F_\theta(x_k)}$$

$$= (m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}(-1)^j\left[\frac{F_\theta(x_{k+1})-F_\theta(x_k)}{1-F_\theta(x_k)}\right]^j\frac{f_\theta(x_{k+1})}{1-F_\theta(x_k)}, \quad 1 \leq k \leq m-1, \tag{7}$$

and

$$P_\theta\{X_{k+1} \leq x_{k+1} \mid X_k = x_k\} = 1 - \left[\frac{1-F_\theta(x_{k+1})}{1-F_\theta(x_k)}\right]^{m-k}$$

$$= (m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}\frac{(-1)^j}{1+j}\left[\frac{F_\theta(x_{k+1})-F_\theta(x_k)}{1-F_\theta(x_k)}\right]^{1+j}, \quad 1 \leq k \leq m-1. \tag{8}$$

Corollary 1.3. If $l = m$,

$$g_\theta(x_m \mid x_k) = (m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}(-1)^j$$

$$\times \left[\frac{1-F_\theta(x_m)}{1-F_\theta(x_k)}\right]^j\frac{f_\theta(x_m)}{1-F_\theta(x_k)} = (m-k)\left[\frac{F_\theta(x_m)-F_\theta(x_k)}{1-F_\theta(x_k)}\right]^{m-k-1}\frac{f_\theta(x_m)}{1-F_\theta(x_k)}, \quad 1 \leq k \leq m-1, \tag{9}$$

and

$$P_\theta\{X_m \leq x_m \mid X_k = x_k\} = 1 - (m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}\frac{(-1)^j}{1+j}\left[\frac{1-F_\theta(x_m)}{1-F_\theta(x_k)}\right]^{1+j}$$

$$= \left[\frac{F_\theta(x_m) - F_\theta(x_k)}{1 - F_\theta(x_k)} \right]^{m-k}, \quad 1 \le k \le m-1. \tag{10}$$

2.1. Exponential distribution

In order to use the results of Theorem 1, we consider, for illustration, the exponential distribution with the probability density function

$$f_\theta(x) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta} \right), \quad x \ge 0, \ \theta > 0, \tag{11}$$

and the probability distribution function

$$F_\theta(x) = 1 - \exp\left(-\frac{x}{\theta} \right), \quad x \ge 0, \ \theta > 0. \tag{12}$$

Theorem 2. Let $X_1 \le \ldots \le X_k$ be the first k ordered observations (order statistics) in a sample of size m from the exponential distribution (11). Then the conditional probability density function of the lth order statistics X_l ($1 \le k < l \le m$) given $X_k = x_k$ is

$$g_\theta(x_l \mid x_k) = \frac{1}{B(l-k, (m-l+1))} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} (-1)^j \frac{1}{\theta} \exp\left(-\frac{(m-l+1+j)(x_l - x_k)}{\theta} \right)$$

$$= \frac{1}{B(l-k, (m-l+1))} \sum_{j=0}^{m-l} \binom{m-l}{j} (-1)^j \frac{1}{\theta} \left[1 - \exp\left(-\frac{x_l - x_k}{\theta} \right) \right]^{l-k-1+j} \exp\left(\frac{x_l - x_k}{\theta} \right),$$

$$x_l \ge x_k, \tag{13}$$

and the conditional probability distribution function of the lth order statistics X_l given $X_k = x_k$ is

$$P_\theta\{ X_l \le x_l \mid X_k = x_k \} = 1 - \frac{1}{B(l-k, (m-l+1))} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j}$$

$$\times \frac{(-1)^j}{m-l+1+j} \exp\left(-\frac{(m-l+1+j)(x_l - x_k)}{\theta} \right)$$

$$= \frac{1}{B(l-k, (m-l+1))} \sum_{j=0}^{m-l} \binom{m-l}{j} \frac{(-1)^j}{l-k+j} \left[1 - \exp\left(-\frac{x_l - x_k}{\theta} \right) \right]^{l-k+j}. \tag{14}$$

Proof. It follows from (3) and (6), respectively. □

Corollary 2.1. If l = k + 1,

$$g_\theta(x_{k+1} \mid x_k) = (m-k)\frac{1}{\theta}\exp\left(-\frac{(m-k)(x_{k+1}-x_k)}{\theta}\right)$$

$$= (m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}(-1)^j\frac{1}{\theta}\left[1-\exp\left(-\frac{x_{k+1}-x_k}{\theta}\right)\right]^j\exp\left(\frac{x_{k+1}-x_k}{\theta}\right),$$

$$x_{k+1} \geq x_k, \quad 1 \leq k \leq m-1, \tag{15}$$

and

$$P_\theta\left\{X_{k+1} \leq x_{k+1} \mid X_k = x_k\right\} = 1 - \exp\left(-\frac{(m-k)(x_{k+1}-x_k)}{\theta}\right)$$

$$= (m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}\frac{(-1)^j}{1+j}\left[1-\exp\left(-\frac{x_{k+1}-x_k}{\theta}\right)\right]^{1+j}, \quad 1 \leq k \leq m-1. \tag{16}$$

Corollary 2.2. If $l = m$,

$$g_\theta(x_m \mid x_k) = (m-k)$$

$$\times \sum_{j=0}^{m-k-1}\binom{m-k-1}{j}(-1)^j\frac{1}{\theta}\exp\left(-\frac{(1+j)(x_m-x_k)}{\theta}\right)$$

$$= (m-k)\frac{1}{\theta}\left[1-\exp\left(-\frac{x_m-x_k}{\theta}\right)\right]^{m-k-1}\exp\left(-\frac{x_m-x_k}{\theta}\right), \quad x_m \geq x_k, \quad 1 \leq k \leq m-1, \tag{17}$$

and

$$P_\theta\left\{X_m \leq x_m \mid X_k = x_k\right\} = 1 - (m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}\frac{(-1)^j}{1+j}\exp\left(-\frac{(1+j)(x_m-x_k)}{\theta}\right)$$

$$= \left[1-\exp\left(-\frac{x_m-x_k}{\theta}\right)\right]^{m-k}, \quad 1 \leq k \leq m-1. \tag{18}$$

Theorem 3. Let $X_1 \leq \ldots \leq X_k$ be the first k ordered observations (order statistics) in a sample of size m from the exponential distribution (11), where the parameter θ is unknown. Then the predictive probability density function of the lth order statistics X_l ($1 \leq k < l \leq m$) is given by

$$g_{s_k}(x_l \mid x_k) = \frac{k}{B(l-k,(m-l+1))}\sum_{j=0}^{l-k-1}\binom{l-k-1}{j}(-1)^j$$

$$\times \left[1+(m-l+1+j)\frac{x_l-x_k}{s_k}\right]^{-(k+1)}\frac{1}{s_k}, \quad x_l \geq x_k, \tag{19}$$

where

$$S_k = \sum_{i=1}^{k} X_i + (m-k)X_k \tag{20}$$

is the sufficient statistic for θ, and the predictive probability distribution function of the lth order statistics X_l is given by

$$P_{s_k}\{X_l \leq x_l \mid X_k = x_k\} = 1 - \frac{1}{B(l-k,(m-l+1))} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} \frac{(-1)^j}{m-l+1+j}$$

$$\times \left[1 + (m-l+1+j)\frac{x_l - x_k}{s_k}\right]^{-k}. \tag{21}$$

Proof. Using the technique of invariant embedding (Nechval et al., 1999; 2004; 2008; 2010a; 2010b; 2010c; 2010d; 2010e; 2011a; 2011b), we reduce (13) to

$$g_\theta(x_l \mid x_k) = \frac{1}{B(l-k,(m-l+1))} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} (-1)^j$$

$$\times v \exp\left(-\frac{(m-l+1+j)(x_l - x_k)}{s_k} v\right)\frac{1}{s_k} = g_{s_k}(x_l \mid x_k, v), \tag{22}$$

where

$$V = S_k / \theta \tag{23}$$

is the pivotal quantity, the probability density function of which is given by

$$f(v) = \frac{1}{\Gamma(k)} v^{k-1}\exp(-v), \quad v \geq 0. \tag{24}$$

Then

$$g_{s_k}(x_l \mid x_k) = E\{g_{s_k}(x_l \mid x_k, v)\} = \int_0^\infty g_{s_k}(x_l \mid x_k, v)f(v)dv = g_{s_k}(x_l \mid x_k). \tag{25}$$

This ends the proof. □

Corollary 3.1. If $l = k + 1$,

$$g_{s_k}(x_{k+1} \mid x_k) = k(m-k)\left[1 + (m-k)\frac{x_{k+1} - x_k}{s_k}\right]^{-(k+1)}\frac{1}{s_k},$$

$$x_{k+1} \geq x_k, \quad 1 \leq k \leq m-1, \tag{26}$$

and

$$P_{s_k}\left\{X_{k+1} \le x_{k+1} \mid X_k = x_k\right\} = 1 - \left[1 + (m-k)\frac{x_{k+1} - x_k}{s_k}\right]^{-k}, \quad 1 \le k \le m-1. \tag{27}$$

Corollary 3.2. If l = m,

$$g_{s_k}(x_m \mid x_k) = k(m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}(-1)^j\left[1 + (1+j)\frac{x_m - x_k}{s_k}\right]^{-(k+1)}\frac{1}{s_k},$$

$$x_m \ge x_k, \quad 1 \le k \le m-1, \tag{28}$$

and

$$P_{s_k}\left\{X_m \le x_m \mid X_k = x_k\right\} = 1 - (m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}\frac{(-1)^j}{1+j}\left[1 + (1+j)\frac{x_m - x_k}{s_k}\right]^{-k}$$

$$= \sum_{j=0}^{m-k}\binom{m-k}{j}(-1)^j\left[1 + j\frac{x_m - x_k}{s_k}\right]^{-k}, \quad 1 \le k \le m-1. \tag{29}$$

2.2. Cumulative customer demand

The primary purpose of this paper is to introduce the idea of cumulative customer demand in inventory control problems to deal with the order statistics from the underlying distribution. It allows one to use the above results to improve statistical decisions for inventory control problems under parametric uncertainty.

Assumptions. The customer demand at the *i*th period represents a random variable Y_i, $i \in \{1, \ldots, m\}$. It is assumed (for the cumulative customer demand) that the random variables

$$X_1 = Y_1, \ldots, X_k = \sum_{i=}^{k}Y_i, \ldots, X_l = \sum_{i=1}^{l}Y_i, \ldots, X_m = \sum_{i=1}^{m}Y_i \tag{30}$$

represent the order statistics ($X_1 \le \ldots \le X_m$) from the exponential distribution (11).

Inferences. For the above case, we have the following inferences.

Conditional probability density function of Y_{k+1}, $k \in \{1, \ldots, m-1\}$, is given by

$$g_\theta(y_{k+1} \mid k) = \frac{m-k}{\theta}\exp\left(-\frac{(m-k)y_{k+1}}{\theta}\right), \quad y_{k+1} \ge 0; \tag{31}$$

Conditional probability distribution function of Y_{k+1}, $k \in \{1, \ldots, m-1\}$, is given by

$$G_\theta\{y_{k+1} \mid k\} = 1 - \exp\left(-\frac{(m-k)y_{k+1}}{\theta}\right).$$

(32)

Conditional probability density function of $Z_m = \sum_{i=k+1}^{m} Y_i$ is given by

$$g_\theta(z_m \mid k) = (m-k)\frac{1}{\theta}\left[1 - \exp\left(-\frac{z_m}{\theta}\right)\right]^{m-k-1}\exp\left(-\frac{z_m}{\theta}\right), \quad z_m \geq 0, \ 1 \leq k \leq m-1;$$

(33)

Conditional probability distribution function of $Z_m = \sum_{i=k+1}^{m} Y_i$ is given by

$$G_\theta(z_m \mid k) = \left[1 - \exp\left(-\frac{z_m}{\theta}\right)\right]^{m-k}, \quad 1 \leq k \leq m-1.$$

(34)

Predictive probability density function of Y_{k+1}, $k \in \{1, \ldots, m-1\}$, is given by

$$g_{s_k}(y_{k+1} \mid k) = k(m-k)\left[1 + (m-k)\frac{y_{k+1}}{s_k}\right]^{-(k+1)}\frac{1}{s_k}, \quad y_{k+1} \geq 0;$$

(35)

Predictive probability distribution function of Y_{k+1}, $k \in \{1, \ldots, m-1\}$, is given by

$$G_{s_k}(y_{k+1} \mid k) = 1 - \left[1 + (m-k)\frac{y_{k+1}}{s_k}\right]^{-k}.$$

(36)

Predictive probability density function of Z_m is given by

$$g_{s_k}(z_m \mid k) = k(m-k)\sum_{j=0}^{m-k-1}\binom{m-k-1}{j}(-1)^j\left[1 + (1+j)\frac{z_m}{s_k}\right]^{-(k+1)}\frac{1}{s_k},$$

$$z_m \geq 0, \ 1 \leq k \leq m-1;$$

(37)

Predictive probability distribution function of Z_m is given by

$$G_{s_k}(z_m \mid k) = \sum_{j=0}^{m-k}\binom{m-k}{j}(-1)^j\left[1 + j\frac{z_m}{s_k}\right]^{-k}, \quad 1 \leq k \leq m-1.$$

(38)

3. Stochastic inventory control problem

Most of the inventory management literature assumes that demand distributions are specified explicitly. However, in many practical situations, the true demand distributions are not known, and the only information available may be a time-series of historic demand data. When the demand distribution is unknown, one may either use a parametric approach

(where it is assumed that the demand distribution belongs to a parametric family of distributions) or a non-parametric approach (where no assumption regarding the parametric form of the unknown demand distribution is made).

Under the parametric approach, one may choose to estimate the unknown parameters or choose a prior distribution for the unknown parameters and apply the Bayesian approach to incorporating the demand data available. Scarf (1959) and Karlin (1960) consider a Bayesian framework for the unknown demand distribution. Specifically, assuming that the demand distribution belongs to the family of exponential distributions, the demand process is characterized by the prior distribution on the unknown parameter. Further extension of this approach is presented in (Azoury, 1985). Application of the Bayesian approach to the censored demand case is given in (Ding et al., 2002; Lariviere & Porteus, 1999). Parameter estimation is first considered in (Conrad, 1976) and recent developments are reported in (Agrawal & Smith, 1996; Nahmias, 1994). Liyanage & Shanthikumar (2005) propose the concept of operational statistics and apply it to a single period newsvendor inventory control problem.

This section deals with inventory items that are in stock during a single time period. At the end of the period, leftover units, if any, are disposed of, as in fashion items. Two models are considered. The difference between the two models is whether or not a setup cost is incurred for placing an order. The symbols used in the development of the models include:

c = setup cost per order,
c_1 = holding cost per held unit during the period,
c_2 = penalty cost per shortage unit during the period,
$g_\theta (y_{k+1}|k)$ = conditional probability density function of customer demand, Y_{k+1}, during the (k+1)th period,
θ = parameter (in general, vector),
u = order quantity,
q = inventory on hand before an order is placed.

3.1. No-setup model (Newsvendor model)

This model is known in the literature as the *newsvendor* model (the original classical name is the *newsboy* model). It deals with stocking and selling newspapers and periodicals. The assumptions of the model are:

1. Demand occurs instantaneously at the start of the period immediately after the order is received.
2. No setup cost is incurred.

The model determines the optimal value of u that minimizes the sum of the expected holding and shortage costs. Given optimal u (= u^*), the inventory policy calls for ordering $u^* - q$ if $q < u^*$; otherwise, no order is placed.

If $Y_{k+1} \le u$, the quantity $u - Y_{k+1}$ is held during the (k+1)th period. Otherwise, a shortage amount $Y_{k+1} - u$ will result if $Y_{k+1} > u$. Thus, the cost per the (k+1)th period is

$$C(u) = \begin{cases} c_1 \dfrac{u - Y_{k+1}}{\theta} & \text{if } Y_{k+1} \leq u, \\ c_2 \dfrac{Y_{k+1} - u}{\theta} & \text{if } Y_{k+1} > u. \end{cases} \tag{39}$$

The expected cost for the (k+1)th period, $E_\theta\{C(u)\}$, is expressed as

$$E_\theta\{C(u)\} = \frac{1}{\theta}\left(c_1 \int_0^u (u - y_{k+1}) g_\theta(y_{k+1} \mid k) dy_{k+1} + c_2 \int_u^\infty (y_{k+1} - u) g_\theta(y_{k+1} \mid k) dy_{k+1} \right). \tag{40}$$

The function $E_\theta\{C(u)\}$ can be shown to be convex in u, thus having a unique minimum. Taking the first derivative of $E_\theta\{C(u)\}$ with respect to u and equating it to zero, we get

$$\frac{1}{\theta}\left(c_1 \int_0^u g_\theta(y_{k+1} \mid k) dy_{k+1} - c_2 \int_u^\infty g_\theta(y_{k+1} \mid k) dy_{k+1} \right) = 0 \tag{41}$$

or

$$c_1 P_\theta\{Y_{k+1} \leq u\} - c_2 (1 - P_\theta\{Y_{k+1} \leq u\}) = 0 \tag{42}$$

or

$$P_\theta\{Y_{k+1} \leq u\} = \frac{c_2}{c_1 + c_2}. \tag{43}$$

It follows from (31), (32), (40), and (43) that

$$u^* = \frac{\theta}{m - k} \ln\left(1 + \frac{c_2}{c_1} \right) \tag{44}$$

and

$$E_\theta\{C(u^*)\} = \frac{1}{\theta}\left(c_2 E_\theta\{Y_{k+1}\} - (c_1 + c_2) \int_0^{u^*} y_{k+1} g_\theta(y_{k+1} \mid k) dy_{k+1} \right)$$

$$= \frac{c_1}{m - k} \ln\left(1 + \frac{c_2}{c_1} \right). \tag{45}$$

Parametric uncertainty. Consider the case when the parameter θ is unknown. To find the best invariant decision rule u^{BI}, we use the invariant embedding technique (Nechval et al., 1999; 2004; 2008; 2010a; 2010b; 2010c; 2010d; 2010e; 2011a; 2011b) to transform (39) to the form, which is depended only on the pivotal quantities V, V_1, and the ancillary factor η. In statistics, a pivotal quantity or pivot is a function of observations and unobservable

parameters whose probability distribution does not depend on unknown parameters. Note that a pivotal quantity need not be a statistic—the function and its value can depend on parameters of the model, but its distribution must not. If it is a statistic, then it is known as an ancillary statistic.

Transformation of C(u) based on the pivotal quantities V, V_1 is given by

$$C^{(1)}(\eta) = \begin{cases} c_1(\eta V - V_1) & \text{if } V_1 \leq \eta V, \\ c_2(V_1 - \eta V) & \text{if } V_1 > \eta V, \end{cases} \tag{46}$$

where

$$\eta = \frac{u}{S_k}, \tag{47}$$

$$V_1 = \frac{Y_{k+1}}{\theta} \sim g(v_1 \mid k) = (m-k)\exp[-(m-k)v_1], \quad v_1 \geq 0. \tag{48}$$

Then $E\{C^{(1)}(\eta)\}$ is expressed as

$$E\{C^{(1)}(\eta)\} = \int_0^\infty \left(c_1 \int_0^{\eta v} (\eta v - v_1)g(v_1 \mid k)dv_1 + c_2 \int_{\eta v}^\infty (v_1 - \eta v)g(v_1 \mid k)dv_1 \right) f(v)dv. \tag{49}$$

The function $E\{C^{(1)}(\eta)\}$ can be shown to be convex in η, thus having a unique minimum. Taking the first derivative of $E\{C^{(1)}(\eta)\}$ with respect to η and equating it to zero, we get

$$\int_0^\infty v \left(c_1 \int_0^{\eta v} g(v_1 \mid k)dv_1 - c_2 \int_{\eta v}^\infty g(v_1 \mid k)dv_1 \right) f(v)dv = 0 \tag{50}$$

or

$$\frac{\int_0^\infty v P(V_1 \leq \eta v)f(v)dv}{\int_0^\infty v f(v)dv} = \frac{c_2}{c_1 + c_2}. \tag{51}$$

It follows from (47), (49), and (51) that the optimum value of η is given by

$$\eta^* = \frac{1}{m-k}\left[\left(1 + \frac{c_2}{c_1}\right)^{1/(k+1)} - 1 \right], \tag{52}$$

the best invariant decision rule is

$$u^{BI} = \eta^* S_k = \frac{S_k}{m-k}\left[\left(1 + \frac{c_2}{c_1}\right)^{1/(k+1)} - 1 \right], \tag{53}$$

and the expected cost, if we use u^{BI}, is given by

$$E_\theta\{C(u^{BI})\} = \frac{c_1(k+1)}{m-k}\left[\left(1+\frac{c_2}{c_1}\right)^{1/(k+1)} - 1\right] = E\{C^{(1)}(\eta^*)\}. \tag{54}$$

It will be noted that, on the other hand, the invariant embedding technique (Nechval et al., 1999; 2004; 2008; 2010a; 2010b; 2010c; 2010d; 2010e; 2011a; 2011b) allows one to transform equation (40) as follows:

$$E_\theta\{C(u)\} = \frac{1}{\theta}\left(c_1\int_0^u (u-y_{k+1})g_\theta(y_{k+1}|k)dy_{k+1} + c_2\int_u^\infty (y_{k+1}-u)g_\theta(y_{k+1}|k)dy_{k+1}\right)$$

$$= \frac{1}{s_k}\left(c_1\int_0^u (u-y_{k+1})v^2(m-k)\exp\left(-\frac{v(m-k)y_{k+1}}{s_k}\right)\frac{1}{s_k}dy_{k+1}\right.$$

$$\left. +c_2\int_u^\infty (y_{k+1}-u)v^2(m-k)\exp\left(-\frac{v(m-k)y_{k+1}}{s_k}\right)\frac{1}{s_k}dy_{k+1}\right). \tag{55}$$

Then it follows from (55) that

$$E\{E_\theta\{(C(u)\}\} = \int_0^\infty E_\theta\{C(u)\}f(v)dv = E_{s_k}\{C^{(1)}(u)\}, \tag{56}$$

where

$$E_{s_k}\{C^{(1)}(u)\} = \frac{k}{s_k}\left(c_1\int_0^u (u-y_{k+1})g_{s_k}^\bullet(y_{k+1}|k)dy_{k+1} + c_2\int_u^\infty (y_{k+1}-u)g_{s_k}^\bullet(y_{k+1}|k)dy_{k+1}\right) \tag{57}$$

represents the expected prediction cost for the (k+1)th period. It follows from (57) that the cost per the (k+1)th period is reduced to

$$C^{(2)}(u) = \begin{cases} c_1\dfrac{u-Y_{k+1}}{s_k/k} & \text{if } Y_{k+1} \le u, \\ c_2\dfrac{Y_{k+1}-u}{s_k/k} & \text{if } Y_{k+1} > u, \end{cases} \tag{58}$$

and the predictive probability density function of Y_{k+1} (compatible with (40)) is given by

$$g_{s_k}^\bullet(y_{k+1}|k) = (k+1)(m-k)\left[1+(m-k)\frac{y_{k+1}}{s_k}\right]^{-(k+2)}\frac{1}{s_k}, \quad y_{k+1} \ge 0. \tag{59}$$

Minimizing the expected prediction cost for the (k+1)th period,

$$E_{s_k}\{C^{(2)}(u)\} = \frac{k}{s_k}\left(c_1\int_0^u (u - y_{k+1})g^\bullet_{s_k}(y_{k+1}\,|\,k)dy_{k+1} + c_2\int_u^\infty (y_{k+1} - u)g^\bullet_{s_k}(y_{k+1}\,|\,k)dy_{k+1}\right), \quad (60)$$

with respect to u, we obtain u^{BI} immediately, and

$$E_{s_k}\{C^{(2)}(u^{BI})\} = \frac{c_1(k+1)}{m-k}\left[\left(1 + \frac{c_2}{c_1}\right)^{1/(k+1)} - 1\right]. \quad (61)$$

It should be remarked that the cost per the (k+1)th period, $C^{(2)}(u)$, can also be transformed to

$$C^{(3)}(\eta) = \begin{cases} c_1 k\left(\dfrac{u}{s_k} - \dfrac{Y_{k+1}}{s_k}\right) & \text{if } \dfrac{Y_{k+1}}{s_k} \le \dfrac{u}{s_k} \\[3mm] c_2 k\left(\dfrac{Y_{k+1}}{s_k} - \dfrac{u}{s_k}\right) & \text{if } \dfrac{Y_{k+1}}{s_k} > \dfrac{u}{s_k} \end{cases} = \begin{cases} c_1 k(\eta - W) & \text{if } W \le \eta \\ c_2 k(W - \eta) & \text{if } W > \eta, \end{cases} \quad (62)$$

where the probability density function of the ancillary statistic W (compatible with (40)) is given by

$$g^\circ_{s_k}(w\,|\,k) = (k+1)(m-k)\left[1 + (m-k)w\right]^{-(k+2)}, \quad w \ge 0. \quad (63)$$

Then the best invariant decision rule $u^{BI} = \eta^* S_k$, where η^* minimizes

$$E\{C^{(3)}(\eta)\} = k\left(c_1\int_0^\eta (\eta - w)g^\circ(w\,|\,k)dw + c_2\int_\eta^\infty (w - \eta)g^\circ(w\,|\,k)dw\right). \quad (64)$$

Comparison of statistical decision rules. For comparison, consider the maximum likelihood decision rule that may be obtained from (44),

$$u^{ML} = \frac{\hat\theta}{m-k}\ln\left(1 + \frac{c_2}{c_1}\right) = \eta^{ML}_j S_k, \quad (65)$$

where $\hat\theta = S_k/k$ is the maximum likelihood estimator of θ,

$$\eta^{ML} = \frac{1}{m-k}\ln\left(1 + \frac{c_2}{c_1}\right)^{1/k}. \quad (66)$$

Since u^{BI} and u^{ML} belong to the same class,

$$C = \{u : u = \eta S_k\}, \quad (67)$$

it follows from the above that u^{ML} is inadmissible in relation to u^{BI}. If, say, k=1 and $c_2 / c_1 = 100$, we have that

$$\text{Rel.eff.}\{u^{ML}, u^{BI}, q\} = E_\theta\{C(u^{BI})\} / E_\theta\{C(u^{ML})\} = 0.838. \tag{68}$$

Thus, in this case, the use of u^{BI} leads to a reduction in the expected cost of about 16.2 % as compared with u^{ML}. The absolute expected cost will be proportional to θ and may be considerable.

3.2. Setup model (s-S policy)

The present model differs from the one in Section 3.1 in that a setup cost c is incurred. Using the same notation, the total expected cost per the (k+1)th period is

$$E_\theta\{\overline{C}(u)\} = c + E_\theta\{C(u)\}$$

$$= c + \frac{1}{\theta}\left(c_1 \int_0^u (u - y_{k+1})g_\theta(y_{k+1} \mid k)dy_{k+1} + c_2 \int_u^\infty (y_{k+1} - u)g_\theta(y_{k+1} \mid k)dy_{k+1} \right). \tag{69}$$

As shown in Section 3.1, the optimum value u^* must satisfy (43). Because c is constant, the minimum value of $E_\theta\{\overline{C}(u)\}$ must also occur at u^*. In Fig.1,
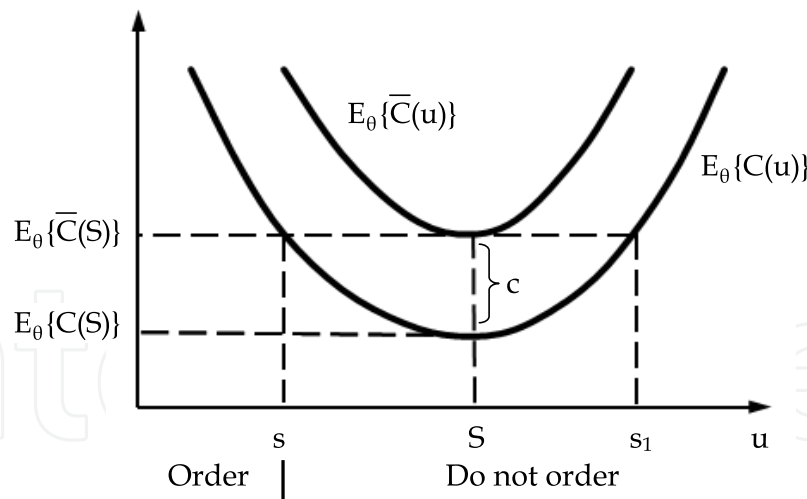
$S = u^*$, and the value of s (< S) is determined from the equation

Figure 1. (s–S) optimal ordering policy in a single-period model with setup cost

$$E_\theta\{C(s)\} = E_\theta\{\overline{C}(S)\} = c + E_\theta\{C(S)\}, \quad s < S. \tag{70}$$

The equation yields another value s_1 (> S), which is discarded.

Assume that q is the amount on hand before an order is placed. How much should be ordered? This question is answered under three conditions: 1) q < s; 2) $s \le q \le S$; 3) q > S.

Case 1 ($q < s$). Because q is already on hand, its equivalent cost is given by $E_\theta\{C(q)\}$. If any additional amount $u - q$ ($u > q$) is ordered, the corresponding cost given u is $E_\theta\{\overline{C}(u)\}$, which includes the setup cost c. From Fig. 1, we have

$$\min_{u>q} E_\theta\{\overline{C}(u)\} = E_\theta\{\overline{C}(S)\} < E_\theta\{C(q)\}. \tag{71}$$

Thus, the optimal inventory policy in this case is to order $S - q$ units.

Case 2 ($s \le q \le S$). From Fig. 1, we have

$$E_\theta\{C(q)\} \le \min_{u>q} E_\theta\{\overline{C}(u)\} = E_\theta\{\overline{C}(S)\}. \tag{72}$$

Thus, it is not advantageous to order in this case and $u^* = q$.

Case 3 ($q > S$). From Fig. 1, we have for $u > q$,

$$E_\theta\{C(q)\} < E_\theta\{\overline{C}(u)\}. \tag{73}$$

This condition indicates that, as in case (2), is not advantageous to place an order – that is, $u^* = q$.

The optimal inventory policy, frequently referred to as the *s – S policy*, is summarized as

$$\begin{array}{l} \text{If } x < S, \text{order } S - x, \\ \text{If } x \ge s, \text{do not order.} \end{array} \tag{74}$$

The optimality of the $s - S$ policy is guaranteed because the associated cost function is convex.

Parametric uncertainty. In the case when the parameter θ is unknown, the total expected prediction cost for the (*k*+1)th period,

$$E_{s_k}\{\overline{C}^{(1)}(u)\} = c + E_{s_k}\{C^{(1)}(u)\}$$

$$= c + \frac{k}{s_k}\left(c_1\int_0^u (u - y_{k+1})g_{s_k}^\bullet(y_{k+1}|k)dy_{k+1} + c_2\int_u^\infty (y_{k+1} - u)g_{s_k}^\bullet(y_{k+1}|k)dy_{k+1}\right), \tag{75}$$

is considered in the same manner as above.

4. Airline revenue management problem

The process of revenue management has become extremely important within the airline industry. It consists of setting fares, setting overbooking limits, and controlling seat inventory to increase revenues. It has allowed the airlines to survive deregulation by allowing them to respond to competitors' deep discount fares on a rational basis.

An airline, typically, offers tickets for many origin–destination itineraries in various fare classes. These fare classes not only include business and economy class, which are settled in separate parts of the plane, but also include fare classes for which the difference in fares is explained by different conditions regarding for example cancellation options or overnight stay arrangements. Therefore the seats on a flight are products, which can be offered to different customer segments for different prices. Since the tickets for a flight have to be sold before the plane takes off, the product is perishable and cannot be stored for future use. The same is true for most other service industries, such as hotels, hospitals and schools.

4.1. Airline seat inventory control

At the heart of airline revenue management lies the airline seat inventory control problem. It is common practice for airlines to sell a pool of identical seats at different prices according to different booking classes to improve revenues in a very competitive market. In other words, airlines sell the same seat at different prices according to different types of travelers (first class, business and economy) and other conditions. The question then arises whether to offer seats at a relatively low price at a given time with a given number of seats remaining or to wait for the possible arrival of a higher paying customer. Assigning seats in the same compartment to different fare classes of passengers in order to improve revenues is a major problem of airline seat inventory control. This problem has been considered in numerous papers. For details, the reader is referred to a review of yield management, as well as perishable asset revenue management, by Weatherford et al. (1993), and a review of relevant mathematical models by Belobaba (1987).

The problem of finding an optimal airline seat inventory control policy for multi-leg flight with multiple fare classes, which allows one to maximize the expected profit of this flight, is one of the most difficult problems of air transport logistics. On the one hand, one must have reasonable assurance that the requirements of customers for reservations will be met under most circumstances. On the other hand, one is confronted with the limitation of the capacity of the cabin, as well as with a host of other less important constraints. The problem is normally solved by the application of judgment based on past experience. The question arises whether or not it is possible to construct a simple mathematical theory of the above problem, which will allow one better to use the available data based upon airline statistics. Two models (dynamic model and static one) of airline data may be considered. In the dynamic model, the problem is formulated as a sequential decision process. In this case, an optimal dynamic reservation policy is used at each stage prior to departure time for multi-leg flights with several classes of passenger service. The essence of determining the optimal dynamic reservation policy is maximization of the expected gain of the flight, which is carried out at each stage prior to departure time using the available data. The term (dynamic reservation policy) is used in this paper to mean a decision rule, based on the available data, for determining whether to accept a given reservation request made at a particular time for some future date. An optimal static reservation policy is based on the static model. The models proposed here contain a simple and natural treatment of the airline reservation process and may be appropriate in practice.

4.2. Expected marginal seat revenue model (EMSR)

Different approaches were developed for solving the airline seat inventory control problem. The most important and widely used model – EMSR – was originally proposed by Littlewood (1972). To explain the basic ideas of the Littlewood model, we will consider the seat allocation problem on a nonstop flight with two airfares. We denote by U the number of seats in the aircraft, by u_L the number of seats reserved for passengers with a lower fare and by p the probability that a passenger who would pay the higher fare cannot find a seat because it was sold to a passenger paying a lower fare. A rise in variable u_L means a rise in probability p. A rise in the number of seats for passengers with the lower fare, i.e. a rise in variable p, decreases the number of seats for passengers with the higher fare, i.e., variable U – u_L decreases. A reduction in variable U – u_L increases the probability that a passenger paying the higher fare cannot find a vacant seat on their desired flight because it has been sold to a passenger with a lower fare. Probability p is given by

$$p = \int_{U-u_L}^{\infty} f_\theta(x)dx, \tag{76}$$

where the $f_\theta(x)$ is the underlying probability density function for the total number of requests X of passengers who would pay the higher fare, θ is the parameter (in general, vector). Littlewood (1972) proposed the following way to determine the reservation level u_L to which reservations are accepted for passengers with the lower fare. We denote by c_1 the revenue from passengers with the lower fare and by c_2 the revenue from passengers with the higher fare. Let u_L seats be reserved for low fare passengers. The revenue per lower fare seat is c_1. Expected revenue from a potential high fare passenger is c_2p. Passengers with lower fares should be accepted until

$$c_1 \geq c_2 p, \tag{77}$$

i.e.,

$$p \leq \frac{c_1}{c_2}. \tag{78}$$

Reservation level u_L is determined by solving the following equality:

$$\Pr\{X > h\} = \gamma, \tag{79}$$

where

$$h = U - u_L, \tag{80}$$

$$\gamma = \frac{c_1}{c_2}. \tag{81}$$

Richter (1982) gave a marginal analysis, which proved that (77) gives an optimal allocation (assuming certain continuity conditions). Optimal policies for more than two classes have been presented independently by Curry (1990), Wollmer (1992), Brumelle & McGill (1993), and Nechval et al. (2006).

Parametric uncertainty. In order to solve (79) under parametric uncertainty, it can be used, for example, the following results.

Theorem 4. Let $X_1 \le \ldots \le X_r$ be the first r ordered past observations from a previous sample of size n from the two-parameter Weibull distribution

$$f(x \mid \beta, \delta) = \frac{\delta}{\beta}\left(\frac{x}{\beta}\right)^{\delta-1} \exp\left[-\left(\frac{x}{\beta}\right)^{\delta}\right] \ (x > 0), \tag{82}$$

where both distribution parameters (β – scale, δ - shape) are positive. Then a lower one-sided conditional γ prediction limit h on the lth order statistic X_l^f in a set of m future ordered observations also from the distribution (82) is given by

$$h = z_h^{1/\hat{\delta}}\hat{\beta}, \tag{83}$$

where z_h satisfies the equation

$$\Pr\{X_l^f > h \mid z^{(r)}\} = \Pr\{X_l^f > z_h^{1/\hat{\delta}}\hat{\beta} \mid z^{(r)}\}$$

$$= \frac{\left[\int_0^{\infty} v_2^{r-2} \prod_{i=1}^{r} z_i^{v_2} \sum_{k=0}^{l-1}\binom{m}{k} \sum_{j=0}^{k}\binom{k}{j}(-1)^j \left((m-k+j)z_h^{v_2} + \sum_{i=1}^{r} z_i^{v_2} + (n-r)z_r^{v_2}\right)^{-r} dv_2\right]}{\int_0^{\infty} v_2^{r-2} \prod_{i=1}^{r} z_i^{v_2}\left(\sum_{i=1}^{r} z_i^{v_2} + (n-r)z_r^{v_2}\right)^{-r} dv_2} = \gamma, \tag{84}$$

$$z^{(r)} = (z_1, z_2, \ldots, z_r), \tag{85}$$

$$Z_i = \left(\frac{X_i}{\hat{\beta}}\right)^{\hat{\delta}}, \ i = 1, \ldots, r, \tag{86}$$

are ancillary statistics, any r–2 of which form a functionally independent set (for notational convenience we include all of $z_1, \ldots, z_r$ in (85); z_{r-1} and z_r can be expressed as function of z_1, $\ldots, z_{r-2}$ only), $\hat{\beta}$ and $\hat{\delta}$ are the maximum likelihood estimators of β and δ based on the first r ordered past observations ($X_1 \le \ldots \le X_r$) from a sample of size n from the two-parameter Weibull distribution, which can be found from solution of

$$\hat{\beta} = \left(\left[\sum_{i=1}^{r} x_i^{\hat{\delta}} + (n-r)x_r^{\hat{\delta}}\right]\Big/ r\right)^{1/\hat{\delta}}, \tag{87}$$

and

$$\hat{\delta} = \left[\left(\sum_{i=1}^{r} x_i^{\hat{\delta}} \ln x_i + (n-r) x_r^{\hat{\delta}} \ln x_r \right) \left(\sum_{i=1}^{r} x_i^{\hat{\delta}} + (n-r) x_r^{\hat{\delta}} \right)^{-1} - \frac{1}{r} \sum_{i=1}^{r} \ln x_i \right]^{-1}, \tag{88}$$

(Observe that an upper one-sided conditional α prediction limit h on the lth order statistic X_l^f from a set of m future ordered observations $X_1^f \leq \ldots \leq X_m^f$ may be obtained from a lower one-sided conditional γ prediction limit by replacing γ by $1-\gamma$.)

Proof. The joint density of $X_1 \leq \ldots \leq X_r$ is given by

$$f(x_1, \ldots, x_r \mid \beta, \delta) = \frac{n!}{(n-r)!} \prod_{i=1}^{r} \frac{\delta}{\beta} \left(\frac{x_i}{\beta} \right)^{\delta-1} \exp\left(-\left(\frac{x_i}{\beta} \right)^{\delta} \right) \exp\left(-(n-r) \left(\frac{x_r}{\beta} \right)^{\delta} \right). \tag{89}$$

Let $\hat{\beta}$, $\hat{\delta}$ be the maximum likelihood estimates of β, δ, respectively, based on $X_1 \leq \ldots \leq X_r$ from a complete sample of size n, and let

$$V_1 = \left(\frac{\hat{\beta}}{\beta} \right)^{\delta}, \quad V_2 = \frac{\hat{\delta}}{\delta}, \quad \text{and} \quad Z_i = \left(\frac{X_i}{\hat{\beta}} \right)^{\hat{\delta}}, \quad i = 1, \ldots, r, \tag{90}$$

Parameters β and δ in (90) are scale and shape parameters, respectively, and it is well known that if $\hat{\beta}$ and $\hat{\delta}$ are estimates of β and δ, possessing certain invariance properties, then V_1 and V_2 are the pivotal quantities whose distributions depend only on n. Most, if not all, proposed estimates of β and δ possess the necessary properties; these include the maximum likelihood estimates and various linear estimates.

Using (90) and the invariant embedding technique (Nechval et al., 1999; 2004; 2008; 2010a; 2010b; 2010c; 2010d; 2010e; 2011a; 2011b), we then find in a straightforward manner, that the joint density of V_1, V_2, conditional on fixed $\mathbf{z} = (z_1, z_2, \ldots, z_r)$, is

$$f(v_1, v_2 \mid z^{(r)})$$

$$= \vartheta^{\bullet}(z^{(r)}) v_2^{r-2} \prod_{i=1}^{r} z_i^{v_2} v_1^{r-1} \exp\left(-v_1 \left[\sum_{i=1}^{r} z_i^{v_2} + (n-r) z_r^{v_2} \right] \right), \quad v_1 \in (0, \infty), \ v_2 \in (0, \infty), \tag{91}$$

where

$$\vartheta^{\bullet}(z^{(r)}) = \left[\int_0^{\infty} \Gamma(r) v_2^{r-2} \prod_{i=1}^{r} z_i^{v_2} \left(\sum_{i=1}^{r} z_i^{v_2} + (n-r) z_r^{v_2} \right)^{-r} dv_2 \right]^{-1}, \tag{92}$$

is the normalizing constant. Writing

$$\Pr\{X_l^f > h \mid \beta, \delta\} = \sum_{k=0}^{l-1} \binom{m}{k} [F(h \mid \beta, \delta)]^k [1 - F(h \mid \beta, \delta)]^{m-k}$$

$$= \sum_{k=0}^{l-1} \binom{m}{k} \left[1 - \exp\left(-\left(\frac{h}{\beta} \right)^{\delta} \right) \right]^{k} \left[\exp\left(-\left(\frac{h}{\beta} \right)^{\delta} \right) \right]^{m-k}$$

$$= \sum_{k=0}^{l-1} \binom{m}{k} \sum_{j=0}^{k} \binom{k}{j} (-1)^{j} \exp[-v_{1}(m-k+j)z_{h}^{v_{2}}]$$

$$= \Pr\{X_{l}^{f} > h \mid v_{1}, v_{2}\}, \tag{93}$$

where

$$z_{h} = \left(\frac{h}{\hat{\beta}} \right)^{\hat{\delta}}, \tag{94}$$

we have from (91) and (93) that

$$\Pr\{X_{l}^{f} > h \mid z^{(r)}\} = \int_{0}^{\infty} \int_{0}^{\infty} \Pr\{X_{l}^{f} > h \mid v_{1}, v_{2}\} f(v_{1}, v_{2} \mid z^{(r)}) dv_{1} dv_{2}. \tag{95}$$

Now v_1 can be integrated out of (95) in a straightforward way to give

$$\Pr\{X_{l}^{f} > h \mid z^{(r)}\}$$

$$= \frac{\left[\int_{0}^{\infty} v_{2}^{r-2} \prod_{i=1}^{r} z_{i}^{v_{2}} \sum_{k=0}^{l-1} \binom{m}{k} \sum_{j=0}^{k} \binom{k}{j} (-1)^{j} \left((m-k+j) \left(\frac{h}{\hat{\beta}} \right)^{\hat{\delta}v_{2}} + \sum_{i=1}^{r} z_{i}^{v_{2}} + (n-r)z_{r}^{v_{2}} \right)^{-r} dv_{2} \right]}{\int_{0}^{\infty} v_{2}^{r-2} \prod_{i=1}^{r} z_{i}^{v_{2}} \left(\sum_{i=1}^{r} z_{i}^{v_{2}} + (n-r)z_{r}^{v_{2}} \right)^{-r} dv_{2}}. \tag{96}$$

This completes the proof. □

Remark 1. If l=m=1, then the result of Theorem 4 can be used to construct the static policy of airline seat inventory control.

Theorem 5. Let $X_1 \le \ldots \le X_k$ be the first k ordered early observations from a sample of size m from the two-parameter Weibull distribution (82). Then a lower one-sided conditional γ prediction limit h on the lth order statistic X_l (l > k) in the same sample is given by

$$h = w_{h}^{1/\hat{\delta}} X_{k}, \tag{97}$$

where w_h satisfies the equation

$$\Pr\{X_{l} > h \mid z^{(k)}\} = \Pr\left\{ \left(X_{l} / X_{k} \right)^{\hat{\delta}} > \left(h / X_{k} \right)^{\hat{\delta}} \mid z^{(k)} \right\} = \Pr\{W > w_{h} \mid z^{(k)}\}$$

$$= \left[\int_0^\infty v_2^{k-2} \prod_{i=1}^k z_i^{v_2} \sum_{j=0}^{l-k-1} \binom{1-k-1}{j} \frac{(-1)^{l-k-1-j}}{m-k-j} \left((m-k-j)(w_h z_k)^{v_2} + j z_k^{v_2} + \sum_{i=1}^k z_i^{v_2} \right)^{-k} dv_2 \right]$$

$$\times \left[\int_0^\infty v_2^{k-2} \prod_{i=1}^k z_i^{v_2} \sum_{j=0}^{l-k-1} \binom{1-k-1}{j} \frac{(-1)^{l-k-1-j}}{m-k-j} \left((m-k) z_k^{v_2} + \sum_{i=1}^k z_i^{v_2} \right)^{-k} dv_2 \right]^{-1} = \gamma, \qquad (98)$$

$$\mathbf{z}^{(k)} = (z_1, \ldots, z_k), \qquad (99)$$

$$Z_i = \left(\frac{Y_i}{\hat{\beta}} \right)^{\hat{\delta}}, \quad i = 1, \ldots, k, \qquad (100)$$

$$w_h = \left(\frac{h}{Y_k} \right)^{\hat{\delta}}, \qquad (101)$$

where $\hat{\beta}$ and $\hat{\delta}$ are the maximum likelihood estimates of β and δ based on the first k ordered past observations $X_1 \le \ldots \le X_k$ from a sample of size m from the two-parameter Weibull distribution (82), which can be found from solution of

$$\hat{\beta} = \left(\left[\sum_{i=1}^k x_i^{\hat{\delta}} + (m-k) x_k^{\hat{\delta}} \right] \middle/ k \right)^{1/\hat{\delta}}, \qquad (102)$$

and

$$\hat{\delta} = \left[\left(\sum_{i=1}^k x_i^{\hat{\delta}} \ln x_i + (m-k) x_k^{\hat{\delta}} \ln x_k \right) \left(\sum_{i=1}^k x_i^{\hat{\delta}} + (m-k) x_k^{\hat{\delta}} \right)^{-1} - \frac{1}{k} \sum_{i=1}^k \ln x_i \right]^{-1}, \qquad (103)$$

(Observe that an upper one-sided conditional α prediction limit h on the lth order statistic X_l based on the first k ordered early-failure observations $X_1 \le \ldots \le X_k$, where $l > k$, from the same sample may be obtained from a lower one-sided conditional γ prediction limit by replacing γ by $1-\gamma$)

Proof. The joint density of $X_1 \le \ldots \le X_k$ and X_l is given by

$$f(x_1, \ldots, x_k, x_l \mid \beta, \delta) = \frac{m!}{(l-k-1)!(m-l)!} \prod_{i=1}^k \frac{\delta}{\beta} \left(\frac{x_i}{\beta} \right)^{\delta-1} \exp\left(-\left(\frac{x_i}{\beta} \right)^\delta \right)$$

$$\times \left[\exp\left(-\left(\frac{x_k}{\beta} \right)^\delta \right) - \exp\left(-\left(\frac{x_l}{\beta} \right)^\delta \right) \right]^{l-k-1} \frac{\delta}{\beta} \left(\frac{x_l}{\beta} \right)^{\delta-1} \exp\left(-\left(\frac{x_l}{\beta} \right)^\delta \right)$$

$$\times \exp\left(-(m-1)\left(\frac{x_1}{\beta}\right)^{\delta}\right).$$ (104)

Let $\hat{\beta}$, $\hat{\delta}$ be the maximum likelihood estimates of β, δ, respectively, based on $X_1 \leq ... \leq X_k$ from a complete sample of size m, and let

$$V_1 = \left(\frac{\hat{\beta}}{\beta}\right)^{\delta}, \quad V_2 = \frac{\delta}{\hat{\delta}}, \quad W = \left(\frac{Y_l}{Y_k}\right)^{\hat{\delta}},$$ (105)

and $Z_i = (Y_i / \hat{\beta})^{\hat{\delta}}$, i=1(1)k. Using the invariant embedding technique (Nechval et al., 1999; 2004; 2008; 2010a; 2010b; 2010c; 2010d; 2010e; 2011a; 2011b), we then find in a straightforward manner, that the joint density of V_1, V_2, W, conditional on fixed $z^{(k)} = (z_1, ..., z_k)$, is

$$f(v_1, v_2, w \mid z^{(k)}) = \vartheta(z^{(k)}) v_2^{k-1} \prod_{i=1}^{k} z_i^{v_2} (wz_k)^{v_2} v_1^k \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} (-1)^{l-k-1-j}$$

$$\times \exp\left[-v_1\left((m-k-j)(wz_k)^{v_2} + jz_k^{v_2} + \sum_{i=1}^{k} z_i^{v_2}\right)\right], \quad v_1 \in (0,\infty), \ v_2 \in (0,\infty), \ w \in (1,\infty), \quad (106)$$

where

$$\vartheta(z^{(k)}) = \left[\int_0^{\infty} v_2^{k-2} \prod_{i=1}^{k} z_i^{v_2} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} \frac{(-1)^{l-k-1-j}\Gamma(k)}{m-k-j}\left((m-k)z_k^{v_2} + \sum_{i=1}^{k} z_i^{v_2}\right)^{-k} dv_2\right]^{-1}$$ (107)

is the normalizing constant. Using (106), we have that

$$\Pr\{Y_l > h \mid z^{(k)}\} = \Pr\left\{\left(\frac{Y_l}{Y_k}\right)^{\hat{\delta}} > \left(\frac{h}{Y_k}\right)^{\hat{\delta}} \Bigg| z^{(k)}\right\}$$

$$= \Pr\{W > w_h \mid z^{(k)}\} = \int_0^{\infty} \int_{w_h}^{\infty} \int_0^{\infty} f(v_1, v_2, w \mid z^{(k)}) dv_1 dw dv_2$$

$$= \left[\int_0^{\infty} v_2^{k-2} \prod_{i=1}^{k} z_i^{v_2} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} \frac{(-1)^{l-k-1-j}}{m-k-j}\left((m-k-j)(w_h z_k)^{v_2} + jz_k^{v_2} + \sum_{i=1}^{k} z_i^{v_2}\right)^{-k} dv_2\right]$$

$$\times \left[\int_0^{\infty} v_2^{k-2} \prod_{i=1}^{k} z_i^{v_2} \sum_{j=0}^{l-k-1} \binom{l-k-1}{j} \frac{(-1)^{l-k-1-j}}{m-k-j}\left((m-k)z_k^{v_2} + \sum_{i=1}^{k} z_i^{v_2}\right)^{-k} dv_2\right]^{-1},$$ (108)

and the proof is complete. □

Remark 2. If l=m, then the result of Theorem 5 can be used to construct the dynamic policy of airline seat inventory control.

5. Conclusions and directions for future research

In this paper, we develop a new frequentist approach to improve predictive statistical decisions for revenue optimization problems under parametric uncertainty of the underlying distributions for the customer demand. Frequentist probability interpretations of the methods considered are clear. Bayesian methods are not considered here. We note, however, that, although subjective Bayesian prediction has a clear personal probability interpretation, it is not generally clear how this should be applied to non-personal prediction or decisions. Objective Bayesian methods, on the other hand, do not have clear probability interpretations in finite samples.

For constructing the improved statistical decisions, a new technique of invariant embedding of sample statistics in a performance index is proposed. This technique represents a simple and computationally attractive statistical method based on the constructive use of the invariance principle in mathematical statistics. The method used is that of the invariant embedding of sample statistics in a performance index in order to form pivotal quantities, which make it possible to eliminate unknown parameters (i.e., parametric uncertainty) from the problem. It is especially efficient when we deal with asymmetric performance indexes and small data samples

More work is needed, however, to obtain improved or optimal decision rules for the problems of unconstrained and constrained optimization under parameter uncertainty when: (i) the observations are from general continuous exponential families of distributions, (ii) the observations are from discrete exponential families of distributions, (iii) some of the observations are from continuous exponential families of distributions and some from discrete exponential families of distributions, (iv) the observations are from multiparametric or multidimensional distributions, (v) the observations are from truncated distributions, (vi) the observations are censored, (vii) the censored observations are from truncated distributions.

Author details

Nicholas A. Nechval and Maris Purgailis
University of Latvia, Latvia

Acknowledgement

6. References

Agrawal, N. & Smith, S.A. (1996). Estimating Negative Binomial Demand for Retail Inventory Management with Unobservable Lost Sales. *Naval Res. Logist.*, Vol. 43, pp. 839 –861

Azoury, K.S. (1985). Bayes Solution to Dynamic Inventory Models under Unknown Demand Distribution. *Management Sci.*, Vol. 31, pp. 1150 –1160

Belobaba, P.P. (1987). Survey Paper: Airline Yield Management – an Overview of Seat Inventory Control. *Transportation Science*, Vol. 21, pp. 63 –73

Brumelle, S.L. & J.I. McGill (1993). Airline Seat Allocation with Multiple Nested Fare Classes. *Operations Research*, Vol. 41, pp. 127–137

Conrad, S.A. (1976). Sales Data and the Estimation of Demand. *Oper. Res. Quart.*, Vol. 27, pp. 123 –127

Curry, R.E. (1990). Optimal Airline Seat Allocation with Fare Classes Nested by Origins and Destination. *Transportation Science*, Vol. 24, pp. 193–204

Ding, X.; Puterman, M.L. & Bisi, A. (2002). The Censored Newsvendor and the Optimal Acquisition of Information. *Oper. Res.*, Vol. 50, pp. 517–527

Karlin, S. (1960). Dynamic Inventory Policy with Varying Stochastic Demands. *Management Sci.*, Vol. 6, pp. 231 –258

Lariviere, M.A. & Porteus, E.L. (1999). Stalking Information: Bayesian Inventory Management with Unobserved Lost Sales. *Management Sci.*, Vol. 45, pp. 346 –363

Littlewood, K. (1972). Forecasting and Control of Passenger Bookings. In: *Proceedings of the Airline Group International Federation of Operational Research Societies*, Vol. 12, pp. 95–117

Liyanage, L.H. & Shanthikumar, J.G. (2005). A Practical Inventory Control Policy Using Operational Statistics. *Oper. Res. Lett.*, Vol. 33, pp. 341–348

Nahmias, S. (1994). Demand Estimation in Lost Sales Inventory Systems. *Naval Res. Logist.*, Vol. 41, 739 –757

Nechval N. A. & Purgailis, M. (2010e). Improved State Estimation of Stochastic Systems via a New Technique of Invariant Embedding. In: *Stochastic Control*, Chris Myers (ed.), pp. 167–193. Publisher: Sciyo, Croatia, India

Nechval, N. A. & Vasermanis, E. K. (2004). *Improved Decisions in Statistics*. SIA "Izglitibas soli", Riga

Nechval, N. A. & Nechval, K. N. (1999). Invariant Embedding Technique and Its Statistical Applications. In: *Conference Volume of Contributed Papers of the 52nd Session of the International Statistical Institute*, Finland, Helsinki: ISI–International Statistical Institute,Availqable from http://www.stat.fi/isi99/procee-dings/arkisto/varasto/nech0902.pdf

Nechval, N. A.; Berzins, G.; Purgailis, M. & Nechval, K.N. (2008). Improved Estimation of State of Stochastic Systems via Invariant Embedding Technique. *WSEAS Transactions on Mathematics*, Vol. 7, pp. 141 –159

Nechval, N. A.; Nechval, K. N. & Purgailis, M. (2011b). Statistical Inferences for Future Outcomes with Applications to Maintenance and Reliability. In: *Lecture Notes in*

Engineering and Computer Science: Proceedings of the World Congress on Engineering (WCE 2011), pp. 865--871. London, U.K.

Nechval, N. A.; Purgailis, M.; Berzins, G.; Cikste, K.; Krasts, J. & Nechval, K. N. (2010a). Invariant Embedding Technique and Its Applications for Improvement or Optimization of Statistical Decisions. In: *Analytical and Stochastic Modeling Techniques and Applications*, Al-Begain, K., Fiems, D., Knottenbelt, W. (eds.),. LNCS, Vol. 6148, pp. 306–320. Springer, Heidelberg

Nechval, N. A.; Purgailis, M.; Cikste, K. & Nechval, K. N. (2010d). Planning Inspections of Fatigued Aircraft Structures via Damage Tolerance Approach. In: *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering (WCE 2010)*, pp. 2470 –2475. London, U.K.

Nechval, N. A.; Purgailis, M.; Cikste, K.; Berzins, G. & Nechval, K. N. (2010c). Optimization of Statistical Decisions via an Invariant Embedding Technique. In: *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering (WCE 2010)*, pp. 1776 –1782. London, U.K.

Nechval, N. A.; Purgailis, M.; Cikste, K.; Berzins, G.; Rozevskis, U. & Nechval, K. N. (2010b). Prediction Model Selection and Spare Parts Ordering Policy for Efficient Support of Maintenance and Repair of Equipment. In: *Analytical and Stochastic Modeling Techniques and Applications*, Al-Begain, K., Fiems, D., Knottenbelt, W. (eds.). LNCS, Vol. 6148, pp. 321–338. Springer, Heidelberg

Nechval, N. A.; Purgailis, M.; Nechval, K. N. & Rozevskis, U. (2011a) Optimization of Prediction Intervals for Order Statistics Based on Censored Data. In: *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering (WCE 2011)*, pp. 63 –69. London, U.K.

Nechval, N.A.; Rozite, K. & Strelchonok, V.F. (2006). Optimal Airline Multi-Leg Flight Seat Inventory Control. In: *Computing Anticipatory Systems*, Daniel M. Dubois (ed.), Melville, New York: AIP (American Institute of Physics) Proceedings, Vol. 839, pp. 591-600

Richter, H. (1982). The Differential Revenue Method to Determine Optimal Seat Allotments by Fare Type. In: *Proceedings of the XXII AGIFORS Symposium*, American Airlines, New York, pp. 339 –362

Scarf, H. (1959). Bayes Solutions of Statistical Inventory Problem. *Ann. Math. Statist*. Vol. 30, pp. 490 –508

Weatherford, L.R.; Bodily, S.E. & P.E. Pferifer (1993). Modeling the Customer Arrival Process and Comparing Decision Rules in Perishable Asset Revenue Management Situations. *Transportation Science*, Vol. 27, pp. 239 –251

Wollmer, R.D. (1992). An Airline Seat Management Model for a Single Leg Route when Lower Fare Classes Book First. *Operations Research*, Vol. 40, pp. 26 –37