CONSTRAINED OPTIMIZATION IN NEWSBOY PROBLEMS UNDER UNCERTAINTY VIA STATISTICAL INFERENCE EQUIVALENCE PRINCIPLE

Nicholas A. Nechval
Department of Mathematical Statistics
University of Latvia
Raina Blvd 19, LV-1050, Riga, Latvia
E-mail: nechval@junik.lv

Konstantin N. Nechval
Department of Applied Mathematics
Transport and Telecommunication Institute
Lomonosov Street 1, LV-1019, Riga, Latvia
E-mail: konstan@tsi.lv

KEYWORDS
Newsboy Problem, Constrained Optimization, Uncertainty, Statistical Inference Equivalence Principle.

ABSTRACT

The aim of the present paper is to show how the statistical inference equivalence principle (the idea of which belongs to the authors) may be employed in the particular case of finding the effective statistical solutions for the multi-product newsboy problems with constraints. To our knowledge, no analytical or efficient numerical method for finding the optimal policies under parameter uncertainty for the multi-product newsboy problems with constraints has been reported in the literature. Using the (equivalent) predictive distributions, this paper represents an extension of analytical results obtained for unconstrained optimization under parameter uncertainty to the case of constrained optimization. An example is given.

INTRODUCTION

The last decade has seen a substantial research focus on the modeling, analysis and optimization of complex stochastic service systems, motivated in large measure by applications in areas such as transport, computer and telecommunication networks. Optimization issues, which broadly focus on making the best use of limited resources, are recognized as of increasing importance. However, stochastic optimization in the context of systems and processes of any complexity is technically very difficult. Most stochastic models to solve the problems of control and optimization of system and processes are developed in the extensive literature under the assumptions that the parameter values of the underlying distributions are known with certainty. In actual practice, such is simply not the case. 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 that are optimal in the absence of uncertainty need not even be approximately optimal in the presence of such uncertainty.

In this paper, we propose a new approach to solve constrained optimization problems under parameter uncertainty. This approach is based on the statistical inference equivalence principle, the idea of which belongs to the authors. It allows one to yield an operational, optimal information-processing rule and may be employed for finding the effective statistical solutions for problems such as multi-product newsboy problem with constraints, allocation of aircraft to routes under uncertainty, airline set inventory control for multi-leg flights, etc.

STATISTICAL INFERENCE EQUIVALENCE PRINCIPLE

In the general formulation of decision theory, we observe a random variable $\mathbf{X}$ (which may be multivariate) with distribution function $F(\mathbf{x};\theta)$ where a parameter θ (in general, vector) is unknown, $\theta\in\Theta$, and if we choose decision d from the set of all possible decisions $\mathcal{D}$, then we suffer a loss $l(d;\theta)$. A "decision rule" is a method of choosing d from $\mathcal{D}$ after observing $\mathbf{x}\in\mathcal{X}$, that is, a function $u(\mathbf{x})=d$. Our average loss (called risk) $E_\theta\{l(u(\mathbf{X});\theta)\}$ is a function of both θ and the decision rule $u(\cdot)$, called the risk function $r(u;\theta)$, and is the criterion by which rules are compared. Thus, the expected loss (gains are negative losses) is a primary consideration in evaluating decisions. We will now define the major quantities just introduced.

Definition 1. A general statistical decision problem is a triplet $(\Theta,\mathcal{D},l)$ and a random variable $\mathbf{X}$. The random variable $\mathbf{X}$ (called the data) has a distribution function $F(\mathbf{x};\theta)$ where θ is unknown but it is known that $\theta\in\Theta$. $\mathcal{X}$ will denote the set of possible values of the random variable $\mathbf{X}$. θ is called the state of nature, while the nonempty set Θ is called the parameter space. The nonempty set $\mathcal{D}$ is called the decision space or action space. Finally, l is called the loss function and to each $\theta\in\Theta$ and $d\in\mathcal{D}$ it assigns a real number $l(d;\theta)$.

Definition 2. For a statistical decision problem $(\Theta,\mathcal{D},l)$, $\mathbf{X}$, a (nonrandomized) decision rule is a function $u(\cdot)$

Proceedings 12th International Conference ASMTA 2005
Khalid Al-Begain, Gunter Bolch, Miklos Telek © ECMS, 2005
ISBN 1-84233-112-4 (Set) / ISBN 1-84233-113-2 (CD)

which to each $\mathbf{x}\in\mathbf{X}$ assigns a member d of $\mathcal{D}$: $u(\mathbf{X})=d$.

Definition 3. The risk function $r(u;\theta)$ of a decision rule $u(\mathbf{X})$ for a statistical decision problem $(\Theta,\mathcal{D},l)$, $\mathbf{X}$ (the expected loss or average loss when θ is the state of nature and a decision is chosen by rule $u(\cdot)$) is $r(u;\theta)=E_\theta\{l(u(\mathbf{X});\theta)\}$.

This paper is concerned with the implications of group theoretic structure for invariant loss functions. Our underlying structure consists of a class of probability models $(\mathcal{X}, \mathcal{A}, \mathcal{P})$, a one-one mapping ψ taking $\mathcal{P}$ onto an index set Θ, a measurable space of actions $(\mathcal{D}, \mathcal{B})$, and a real-valued loss function

$$l(d;\theta)=E_\theta\left\{l^\circ(d;X)\right\} \tag{1}$$

defined on $\Theta \times \mathcal{D}$, where $l^\circ(d;X)$ is a random loss function with a random variable $X\in(0,\infty)$ (or $(-\infty,\infty)$). We assume that a group G of one-one $\mathcal{A}$- measurable transformations acts on $\mathcal{X}$ and that it leaves the class of models $(\mathcal{X}, \mathcal{A}, \mathcal{P})$ invariant. We further assume that homomorphic images $\overline{G}$ and $\widetilde{G}$ of G act on Θ and $\mathcal{D}$, respectively. ($\overline{G}$ may be induced on Θ through ψ; $\widetilde{G}$ may be induced on $\mathcal{D}$ through l). We shall say that l is invariant if for every $(\theta, d) \in \Theta \times \mathcal{D}$

$$l(\tilde{g}d;\bar{g}\theta)=l(d;\theta),\quad g\in G. \tag{2}$$

A loss function, $l(d;\theta)$, can be transformed as follows:

$$l(d;\theta)=l(\tilde{g}_{\hat{\theta}}^{-1}d;\bar{g}_{\hat{\theta}}^{-1}\theta)=l^{\#}(\eta;\mathbf{V}), \tag{3}$$

where $\mathbf{V}=\mathbf{V}(\theta,\hat{\theta})$ is a pivotal quantity whose distribution does not depend on unknown parameter θ; $\eta=\eta(d,\hat{\theta})$ is an ancillary factor; $\hat{\theta}$ is a maximum likelihood estimator of θ (or a sufficient statistic for θ). Then the best invariant decision rule (BIDR) is given by

$$u^{\mathrm{BIDR}}\equiv d=\eta^{-1}(\eta^*,\hat{\theta}), \tag{4}$$

where

$$\eta^*=\arg\inf_\eta E\left\{l^{\#}(\eta;\mathbf{V})\right\} \tag{5}$$

and a risk function

$$r(u^{\mathrm{BIDR}};\theta)=E_\theta\left\{l(u^{\mathrm{BIDR}};\theta)\right\}=E\left\{l^{\#}(\eta^*;\mathbf{V})\right\} \tag{6}$$

does not depend on θ. Consider now a situation described by one of a family of density functions $f(x;\mu,\sigma)$ indexed by the vector parameter $\theta=(\mu,\sigma)$, where μ and $\sigma(>0)$ are respectively parameters of location and scale. For this family, invariant under the group of positive linear transformations: $x\to ax+b$ with $a>0$, we shall assume that there is obtainable from some informative experiment (a random sample of observations $\mathbf{X}=(X_1, \ldots, X_n)$) a sufficient statistic (M,S) for (μ,σ) with density function $h(m,s;\mu,\sigma)$ of the form

$$h(m,s;\mu,\sigma)=\sigma^{-2}h_\bullet[(m-\mu)/\sigma, s/\sigma] \tag{7}$$

such that

$$h(m,s;\mu,\sigma)dmds=h_\bullet(v_1,v_2)dv_1dv_2, \tag{8}$$

where $V_1=(M-\mu)/\sigma$, $V_2=S/\sigma$. We are thus assuming that for the family of density functions an induced invariance holds under the group G of transformations: $m\to am+b$, $s\to as$ $(a>0)$. The family of density functions $f(x;\mu,\sigma)$ satisfying the above conditions is, of course, the limited one of normal, negative exponential, Weibull and gamma, with known index, density functions. The structure of the problem is, however, more clearly seen within the general framework.

Suppose that we deal with a loss function $l^{\dagger}(d,\theta)=E_\theta\{l^\circ(d;X)\}=\omega(\sigma)l(d;\theta)$, where $\omega(\sigma)$ is some function of σ and $\omega(\sigma)=\omega_\bullet(V_2,S)$. In order to obtain an equivalent conditional loss function $l^\bullet(d;m,s)$, which is independent on θ and has the same optimal invariant statistical solution given by (4), i.e.,

$$\arg\min_d l^\bullet(d;M,S)=d^*\equiv u^{\mathrm{BIDR}}, \tag{9}$$

with a risk given by

$$E_\theta\left\{l^\bullet(u^{\mathrm{BIDR}};M,S)\right\}=\omega(\sigma)r(u^{\mathrm{BIDR}};\theta), \tag{10}$$

we define an equivalent predictive conditional probability density function of a random variable X (with a probability density function $f(x;\mu,\sigma)$) as

$$f^\bullet(x;m,s)=\iint_{v_1,v_2} f(x;m,s,v_1,v_2)h_{\bullet\bullet}(v_1,v_2)dv_1dv_2, \tag{11}$$

where

$$f(x;m,s,v_1,v_2)=f(x;\mu,\sigma), \tag{12}$$

$$h_{\bullet\bullet}(v_1,v_2)=\frac{\omega_\bullet^{-1}(v_2,s)h_\bullet(v_1,v_2)}{\iint_{v_1,v_2}\omega_\bullet^{-1}(v_2,s)h_\bullet(v_1,v_2)dv_1dv_2}. \tag{13}$$

Then $l^\bullet(d;m,s)$ is given by

$$l^\bullet(d;m,s)=E_{m,s}\left\{l^\circ(d;X)\right\}=\int_x l^\circ(d;X)f^\bullet(x;m,s)dx. \tag{14}$$

Now the conditional loss function $l^\bullet(d;m,s)$ can be used to obtain efficient frequentist statistical solutions for constrained optimization problems, where the known approaches are unable to do it.

NEWSBOY PROBLEM WITH NO CONSTRAINTS

Preliminaries

The classical newsboy problem is reflective of many real life situations and is often used to aid decision-making in the fashion and sporting industries, both at the manufacturing and retail levels (Gallego and Moon 1993). The newsboy problem can also be used in managing capacity and evaluating advanced booking of orders in service industries such as airlines and hotels (Weatherford and Pfeifer 1994). A partial review of the newsboy problem literature has been recently conducted in a textbook by Silver et al. (1998). Researchers have followed two approaches to solving the newsboy problems. In the first approach, the expected costs of overestimating and underestimating demand are minimized. In the second approach, the expected profit is maximized. Both approaches yield the same results. We use the first approach in stating the newsboy problem. For product *j*, define:

X_j	quantity demanded during the period, a random variable,
$f_j(x_j;\mu_j,\sigma_j)$	the probability density function of X_j,
$\theta_j=(\mu_j,\sigma_j)$	the parameter of $f_j(x_j;\mu_j,\sigma_j)$,
$F_j(x_j;\mu_j,\sigma_j)$	the cumulative distribution function of X_j,
$c_j^{(1)}$	overage (excess) cost per unit,
$c_j^{(2)}$	underage (shortage) cost per unit,
d_j	inventory/order quantity, a decision variable.

The cost per period is

$$l_j^{\circ}(d_j;X_j)=\begin{cases}c_j^{(1)}(d_j-X_j), & \text{if } X_j<d_j,\\ c_j^{(2)}(X_j-d_j), & \text{if } X_j\ge d_j.\end{cases} \quad (15)$$

Complete Information

A standard newsboy formulation (see, e.g., (Nahmias 1996)) is to consider each product j's cost function:

$$l_j^{+}(d_j;\theta_j)=c_j^{(1)}\int_{-\infty}^{d_j}(d_j-x_j)f_j(x_j;\mu_j,\sigma_j)dx_j$$

$$+c_j^{(2)}\int_{d_j}^{\infty}(x_j-d_j)f_j(x_j;\mu_j,\sigma_j)dx_j. \quad (16)$$

Expanding (16) gives

$$l_j^{+}(d_j;\theta_j)=-c_j^{(1)}\int_{-\infty}^{d_j}x_jf_j(x_j;\mu_j,\sigma_j)dx_j$$

$$+c_j^{(2)}\int_{d_j}^{\infty}x_jf_j(x_j;\mu_j,\sigma_j)dx_j$$

$$+(c_j^{(1)}+c_j^{(2)})d_j\left(F_j(d_j;\mu_j,\sigma_j)-\frac{c_j^{(2)}}{c_j^{(1)}+c_j^{(2)}}\right). \quad (17)$$

Let the superscript * denote optimality. Using Leibniz's rule to obtain the first and second derivatives shows that $l_j^{+}(d_j;\theta_j)$ is concave. The sufficient optimality condition is the well-known fractile formula:

$$F_j(d_j^*;\mu_j,\sigma_j)=\frac{c_j^{(2)}}{c_j^{(1)}+c_j^{(2)}}. \quad (18)$$

It follows from (18) that

$$d_j^*=F_j^{-1}\left(\frac{c_j^{(2)}}{c_j^{(1)}+c_j^{(2)}};\mu_j,\sigma_j\right). \quad (19)$$

At optimality, substituting (18) into the last (bracketed) term in Eq. (17) gives

$$(c_j^{(1)}+c_j^{(2)})d_j^*\left(F_j(d_j^*;\mu_j,\sigma_j)-\frac{c^{(2)}}{c^{(1)}+c^{(2)}}\right)=0. \quad (20)$$

Hence (17) reduces to

$$l_j^{+}(d_j^*;\theta_j)=c_j^{(2)}E_{\theta}\{X_j\}-(c_j^{(1)}+c_j^{(2)})\int_0^{d_j^*}x_jf_j(x_j;\mu_j,\sigma_j)dx_j. \quad (21)$$

Parameter Uncertainty

Let us assume that the functional form of the probability density function $f_j(x_j;\mu_j,\sigma_j)$ is specified but its parameter $\theta=(\mu_j,\sigma_j)$ is not specified. Let $\mathbf{X}_j=(X_{j1}, \ldots, X_{jn})$ be a random sample of observations on a continuous random variable X_j. We shall assume that there is obtainable from a random sample of observations $\mathbf{X}_j=(X_{j1}, \ldots, X_{jn})$ a sufficient statistic (M_j,S_j) for $\theta=(\mu_j,\sigma_j)$ with density function of the form (7),

$$h_j(m_j,s_j;\mu_j,\sigma_j)=\sigma_j^{-2}h_{\bullet j}[(m_j-\mu_j)/\sigma_j, s_j/\sigma_j], \quad (22)$$

and with

$$h_j(m_j,s_j;\mu_j,\sigma_j)dm_jds_j=h_{\bullet j}(v_{1j},v_{2j})dv_{1j}dv_{2j}, \quad (23)$$

where $V_{1j}=(M_j-\mu_j)/\sigma_j$, $V_{2j}=S_j/\sigma_j$.

Using an invariant embedding technique (Nechval et al. 2000; 2001), we transform (16) as follows:

$$l_j^{+}(d_j;\theta_j)=\omega_j(\sigma_j)l_j^{\#}(\eta_j;\mathbf{V}_j), \quad (24)$$

where $\omega_j(\sigma_j)=\sigma_j$,

$$l_j^{\#}(\eta_j;\mathbf{V}_j)=c_j^{(1)}\int_{-\infty}^{\eta_jV_{2j}+V_{1j}}(\eta_jV_{2j}+V_{1j}-z_j)f_j(z_j)dz_j$$

$$+c_j^{(2)}\int_{\eta_j V_{2j}+V_{1j}}^{\infty}\left(z_j-\eta_j V_{2j}-V_{1j}\right)f_j(z_j)dz_j, \qquad (25)$$

$Z_j=(X_j-\mu_j)/\sigma_j$ is a pivotal quantity, $f_j(z_j)$ is defined by $f_j(x_j;\mu_j,\sigma_j)$, i.e.,

$$f_j(z_j)dz_j = f_j(x_j;\mu_j,\sigma_j)dx_j\,, \qquad (26)$$

$\mathbf{V}_j=(V_{1j},V_{2j})$ is a pivotal quantity, $\eta_j=(d_j-M_j)/S_j$ is an ancillary factor. It follows from (24) that the risk associated with u_j^{BIDR} (or η_j^*) can be expressed as

$$r_j^+(u_j^{\mathrm{BIDR}};\boldsymbol{\theta}_j)=E_{\boldsymbol{\theta}_j}\left\{l_j^+(u_j^{\mathrm{BIDR}};\boldsymbol{\theta}_j)\right\}$$

$$=\omega_j(\sigma_j)E\left\{l_j^{\#}(\eta_j^*;\mathbf{V}_j)\right\}, \qquad (27)$$

where

$$u_j^{\mathrm{BIDR}}\equiv d_j^*=M_j+\eta_j^*S_j, \qquad (28)$$

$$\eta_j^*=\arg\min_{\eta_j}E\left\{l_j^{\#}(\eta_j;\mathbf{V}_j)\right\}, \qquad (29)$$

$$E\left\{l_j^{\#}(\eta_j;\mathbf{V}_j)\right\}=\iint_{v_{1j},v_{2j}}l_j^{\#}(\eta_j;v_{1j},v_{2j})h_{\bullet j}(v_{1j},v_{2j})dv_{1j}dv_{2j}. \qquad (30)$$

The fact that (30) is independent of θ_j means that an ancillary factor η_j^*, which minimizes (30), is uniformly best invariant. Thus, d_j^* given by (28) is the best invariant decision rule.

Relative Efficiency of Decision Rules

Consider two decision rules based on a sample of observations $\mathbf{X}_j=(X_{j1},\ldots,X_{jn})$, say, $\hat{u}_j\equiv\hat{u}(\mathbf{X}_j)$ and $\tilde{u}_j\equiv\tilde{u}(\mathbf{X}_j)$ having risk function $r_j^+(\hat{u}_j;\boldsymbol{\theta}_j)$ and $r_j^+(\tilde{u}_j;\boldsymbol{\theta}_j)$, respectively. Then the relative efficiency of $\hat{u}_j$ relative to $\tilde{u}_j$ is given by

$$\mathrm{rel.eff.}_{r_j^+}\{\hat{u}_j,\tilde{u}_j;\boldsymbol{\theta}_j\} = r_j^+(\tilde{u}_j;\boldsymbol{\theta}_j)\big/r_j^+(\hat{u}_j;\boldsymbol{\theta}_j). \qquad (31)$$

When $\mathrm{rel.eff.}_{r_j^+}\{\hat{u}_j,\tilde{u}_j;\boldsymbol{\theta}_j^{(0)}\}<1$ for some $\boldsymbol{\theta}_j^{(0)}$, we say that $\tilde{u}_j$ is more efficient than $\hat{u}_j$ at $\boldsymbol{\theta}_j^{(0)}$. If $\mathrm{rel.eff.}_{r_j^+}\{\hat{u}_j,\tilde{u}_j;\boldsymbol{\theta}_j^{(0)}\}\le 1$ for all θ_j with a strict inequality for some $\boldsymbol{\theta}_j^{(0)}$, then $\hat{u}_j$ is inadmissible in relation to $\tilde{u}_j$.

EXAMPLE

Assuming that the demand for product j, X_j, is exponentially distributed with the probability density function,

$$f_j(x_j;\sigma_j)=(1/\sigma_j)\exp(-x_j/\sigma_j)\ \ (x_j>0), \qquad (32)$$

it follows from (16), (19) and (21) that

$$l_j^+(d_j;\sigma_j)=c_j^{(1)}(d_j-\sigma_j)+(c_j^{(1)}+c_j^{(2)})\sigma_j\exp\left(-\frac{d_j}{\sigma_j}\right), \qquad (33)$$

$$d_j^*=\sigma_j\ln\left(1+\frac{c_j^{(2)}}{c_j^{(1)}}\right), \qquad (34)$$

and

$$l_j^+(d_j^*;\sigma_j)=c_j^{(1)}\sigma_j\ln\left(1+\frac{c_j^{(2)}}{c_j^{(1)}}\right), \qquad (35)$$

respectively.

Consider the case when the parameter σ_j is unknown. Let $\mathbf{X}_j=(X_{j1},\ldots,X_{jn})$ be a random sample of observations (each with density function (32)) on a continuous random variable X_j. Then

$$S_j=\sum_{i=1}^{n}X_{ji}, \qquad (36)$$

is a sufficient statistic for σ_j; S_j is distributed with

$$h_j(s_j;\sigma_j)=\frac{1}{\Gamma(n)\sigma_j^n}s_j^{n-1}\exp\left(-\frac{s_j}{\sigma_j}\right)\quad(s_j>0), \quad (37)$$

so that

$$h_{\bullet j}(v_{2j})=\frac{1}{\Gamma(n)}v_{2j}^{n-1}e^{-nv_{2j}}\ (v_{2j}>0). \qquad (38)$$

It follows from (27) and (33) that

$$r_j^+(u_j^{\mathrm{BIDR}};\sigma_j)=E_{\sigma_j}\left\{l_j^+(u_j^{\mathrm{BIDR}};\sigma_j)\right\}$$

$$=\sigma_j\int_0^{\infty}l_j^{\#}(\eta_j^*;v_{2j})h_{\bullet j}(v_{2j})dv_{2j}$$

$$=\sigma_j\left[c_j^{(1)}(n\eta_j^*-1)+\frac{c_j^{(1)}+c_j^{(2)}}{(1+\eta_j^*)^n}\right], \qquad (39)$$

where

$$u_j^{\mathrm{BIDR}}=\eta_j^*S_j, \qquad (40)$$

$$\eta_j^*=\arg\min_{\eta_j}\sigma_j\left[c_j^{(1)}(n\eta_j-1)+\frac{c_j^{(1)}+c_j^{(2)}}{(1+\eta_j)^n}\right]$$

$$=\left[1+\frac{c_j^{(2)}}{c_j^{(1)}}\right]^{1/(n+1)}-1. \qquad (41)$$

For comparison, consider the maximum likelihood decision rule (MLDR) that may be obtained from (24),

$$u_j^{\mathrm{MLDR}} = \widehat{\sigma}_j \ln\left(1+\frac{c_j^{(2)}}{c_j^{(1)}}\right) = \eta_j^{\mathrm{MLDR}} S_j , \qquad (42)$$

where $\widehat{\sigma}_j = S_j/n$ is the maximum likelihood estimator of σ_j. Since u_j^{BIDR} and u_j^{MLDR} belong to the same class

$$\mathscr{C} = \{u_j : u_j = \eta_j S_j\}, \qquad (43)$$

it follows from the above that u_j^{MLDR} is inadmissible in relation to u_j^{BIDR}.

If, say, n=1 and $c_j^{(2)}/c_j^{(1)}$=100, we have that

$$\text{rel.eff.}_{r_j^+}\{(u_j^{\mathrm{MLDR}}, u_j^{\mathrm{BIDR}}; \sigma_j\}$$

$$= r_j^+(u_j^{\mathrm{BIDR}};\sigma_j)\big/ r_j^+(u_j^{\mathrm{MLDR}};\sigma_j)$$

$$= \left(n\eta_j^* - 1 + \frac{1+c_j^{(2)}/c_j^{(1)}}{(1+\eta_j^*)^n}\right)\left(n\eta_j^{\mathrm{MLDR}} - 1 + \frac{1+c_j^{(2)}/c_j^{(1)}}{(1+\eta_j^{\mathrm{MLDR}})^n}\right)^{-1}$$

$$=0.84. \qquad (44)$$

Thus, in this case, the use of u_j^{BIDR} leads to a reduction in the risk of about 16 % as compared with u_j^{MLDR}. The absolute risk will be proportional to σ_j and may be considerable.

In order to obtain an equivalent conditional loss function $l_j^{\bullet}(d_j;s_j)$, which is independent on σ_j and has the same optimal invariant statistical solution given by (40), i.e.,

$$\arg\min_{d_j} l_j^{\bullet}(d_j;S_j) = d_j^* \equiv u_j^{\mathrm{BIDR}}, \qquad (45)$$

with a risk given by

$$E_{\sigma_j}\left\{l_j^{\bullet}(u_j^{\mathrm{BIDR}};S_j)\right\} = r_j^+(u_j^{\mathrm{BIDR}};\sigma_j), \qquad (46)$$

we define (on the basis of (11)) an equivalent predictive conditional probability density function of a random variable X_j (with the probability density function $f_j(x_j;\sigma_j)$) as

$$f_j^{\bullet}(x_j;s_j) = \frac{n+1}{s_j}\left(1+\frac{x_j}{s_j}\right)^{-(n+2)} \quad (x_j>0). \qquad (47)$$

Then $l_j^{\bullet}(d_j;s_j)$ is given by

$$l_j^{\bullet}(d_j;s_j) = E_{s_j}\left\{l_j^{\circ}(d_j;X_j)\right\} = \int_0^{\infty} l_j^{\circ}(d_j;x_j) f_j^{\bullet}(x_j;s_j)dx_j$$

$$= \frac{s_j}{n}\left[c_j^{(1)}\left(n\frac{d_j}{s_j}-1\right) + (c_j^{(1)}+c_j^{(2)})\left(1+\frac{d_j}{s_j}\right)^{-n}\right]. \qquad (48)$$

Now the equivalent conditional loss function $l_j^{\bullet}(d_j;s_j)$ can be used to obtain efficient frequentist statistical solutions for constrained optimization problems, where the known approaches are unable to do it.

NEWSBOY PROBLEM WITH CONSTRAINTS

Complete Information

Define w_j (>0) as product j's per-unit requirement of a constrained resource, and w_Σ as the maximum availability of the resource. The formulation for minimizing the total expected cost of N products subject to one capacity constraint is as follows:

Minimize

$$\sum_{j=1}^{N} l_j^+(d_j;\boldsymbol{\theta}_j) = \sum_{j=1}^{N}\left(c_j^{(1)}\int_0^{d_j}(d_j-x_j)f_j(x_j;\mu_j,\sigma_j)dx_j\right.$$

$$\left. + c_j^{(2)}\int_{d_j}^{\infty}(x_j-d_j)f_j(x_j;\mu_j,\sigma_j)dx_j\right) \qquad (49)$$

Subject to

$$\sum_{j=1}^{N} w_j d_j \le w_\Sigma. \qquad (50)$$

The above problem can be solved as follows:

Compute d_j^* for each product j with Eq. (19) and check whether $\sum_j w_j d_j^*$ exceeds w_Σ. If it does not, the capacity constraint is non-operative, and the optimal order quantity is d_j^*, $\forall j$=1(1)N. Otherwise, the constraint is set to equality and the Lagrange function is introduced as in the following (note that λ is the Lagrange multiplier):

$$L = \sum_{j=1}^{N}\left(c_j^{(1)}\int_0^{d_j}(d_j-x_j)f_j(x_j;\mu_j,\sigma_j)dx_j\right.$$

$$\left. + c_j^{(2)}\int_{d_j}^{\infty}(x_j-d_j)f_j(x_j;\mu_j,\sigma_j)dx_j\right) + \lambda\left(\sum_{j=1}^{N} w_j d_j - w_\Sigma\right).$$

(51)

Again, using Leibniz rule and differentiating, we obtain the optimal solution under capacity constraint

$$d_j^{c*} = F_j^{-1}\left(\frac{c_j^{(2)} - \lambda w_j}{c_j^{(1)} + c_j^{(2)}}; \mu_j, \sigma_j\right), \quad \forall j = 1(1)N, \tag{52}$$

where the value of the Lagrange multiplier λ can be determined by solving the single-variable (λ) non-linear equation

$$\sum_{j=1}^{N} w_j F_j^{-1}\left(\frac{c_j^{(2)} - \lambda w_j}{c_j^{(1)} + c_j^{(2)}}; \mu_j, \sigma_j\right) - w_\Sigma = 0. \tag{53}$$

Parameter Uncertainty

In this case, the problem is as follows:

Minimize the total conditional expected losses

$$\sum_{j=1}^{N} l_j^{\bullet}(d_j; m_j, s_j) = \sum_{j=1}^{N}\left(c_j^{(1)} \int_0^{d_j} (d_j - x_j) f_j^{\bullet}(x_j; m_j, s_j)\, dx_j + c_j^{(2)} \int_{d_j}^{\infty} (x_j - d_j) f_j^{\bullet}(x_j; m_j, s_j)\, dx_j \right) \tag{54}$$

Subject to

$$\sum_{j=1}^{N} w_j d_j \le w_\Sigma. \tag{55}$$

In the same manner as above, we can obtain the optimal statistical solutions under capacity constraint.

CONCLUSION

In this paper, we propose a new approach to solve constrained optimization problems under uncertainty. It is especially efficient when we deal with asymmetric loss functions and small data samples. The results obtained in this paper agree with the simulation results, which confirm the validity of the theoretical predictions of performance of the suggested approach.

ACKNOWLEDGMENTS

This paper is based on research supported, in part, by the Latvian Council of Science and the National Institute of Mathematics and Informatics of Latvia under Grant No. 02.0918 and Grant No. 01.0031. This support is gratefully acknowledged.

The authors thank the three anonymous referees for their helpful suggestions, which improved the presentation of this paper.

REFERENCES

Gallego, G. and I. Moon. 1993. "The Distribution Free Newsboy Problem: Review and Extensions." *The Journal of the Operational Research Society* 44, 825–834.

Nahmias, S. 1996. *Production and Operations Management*. Irwin, Boston.

Nechval, N.A. and K.N. Nechval. 2000. "State Estimation of Stochastic Systems via Invariant Embedding Technique." In *Cybernetics and Systems'2000*, R. Trappl (Ed.). Vienna, Austrian Society for Cybernetic Studies, Vol. 1, 96-101.

Nechval, N.A.; K. N. Nechval; and E.K. Vasermanis. 2001. "Optimization of Interval Estimators via Invariant Embedding Technique." *IJCAS* (*The International Journal of Computing Anticipatory Systems*) 9, 241-255.

Silver, E.A.; D. F. Pyke; and R.P. Peterson. 1998. *Inventory Management and Production Planning and Scheduling*. John Wiley, New York.

Weatherford, L. R. and P.E. Pfeifer. 1994. "The Economic Value of Using Advance Booking of Orders." *Omega* 22, 105–111.

AUTHOR BIOGRAPHIES

NICHOLAS A. NECHVAL received the PhD degree in automatic control and systems engineering from the Riga Civil Aviation Engineers Institute (RCAEI) in June, 1969, and the DSc degree in radio engineering from the Riga Aviation University (RAU) in June, 1993. Dr. Nechval was Professor of Applied Mathematics at the RAU, from 1993 to 1999. At present, he is Professor of Mathematics and Computer Science at the University of Latvia, Riga, Latvia. In 1992, Dr. Nechval was awarded a Silver Medal of the Exhibition Committee (Moscow, Russia) in connection with research on the problem of Prevention of Collisions between Aircraft and Birds. He is the holder of several patents in this field. His interests include mathematics, stochastic processes, pattern recognition, multidimensional statistic detection and estimation, multiresolution stochastic signal analysis, digital radar signal processing, operations research, statistical decision theory, and adaptive control. Professor Nechval is a professional member of the Latvian Statistical Association, the Institute of Mathematical Statistics, and CHAOS asbl (the Institute of Mathematics, based in Liege, Belgium). Dr. Nechval is also a member of the Latvian Association of Professors.

KONSTANTIN N. NECHVAL was born in Riga, Latvia, on March 5, 1975. He received the MS degree from the Aviation University of Riga, Latvia, in 1998. At present, he is a PhD Student in automatic control and systems engineering at the Riga Technical University. His research interests include stochastic processes, pattern recognition, operations research, statistical decision theory, and adaptive control.