

# Solving Operational Statistics via a Bayesian Analysis

Leon Yang Chu<sup>a</sup>, J. George Shanthikumar<sup>b</sup>, Zuo-Jun Max Shen<sup>b,\*</sup>

<sup>a</sup>Department of Information and Operations Management

Marshall School of Business

University of Southern California

Los Angeles, CA 90089, USA

<sup>b</sup>Department of Industrial Engineering and Operations Research

University of California at Berkeley

Berkeley, CA 94720, USA

November 05, 2005; Revised December 05, 2006; April 06, 2007

## Abstract

For the newsvendor inventory control problem with ambiguous demand, Liyanage and Shanthikumar [13] show that integrating parameter estimation and optimization using operational statistics leads to a better solution compared with the traditional approach. However, they leave open the question of finding the optimal operational statistic. We use Bayesian analysis and a non-informative prior to derive the optimal operational statistic.

*Keywords:* Newsvendor model; Model uncertainty, Demand ambiguity, Operational statistics, Estimation and Optimization

\* The authors would like to thank the referee and the area editor for helpful comments on an earlier version that helped improve the presentation of this paper. This work was supported in part by NSF CAREER award DMI-0348209 (for Chu and Shen) and DMI-0500503 (for Shanthikumar). This support is gratefully acknowledged.

\* Corresponding author

*E-mail addresses:* leonyzhu@usc.edu (L. Y. Chu), shanthikumar@ieor.berkeley.edu (J. G. Shanthikumar), shen@ieor.berkeley.edu (Z. J. M. Shen)

# 1 Introduction

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.

Roughly there are two classes of papers in the literature that deal with inventory management problems with unknown demand distribution. The first class assumes that the demand distribution belongs to a parametric family of distributions. Azoury [1], Eppen and Iyer [5], Haksoz and Seshadri [7], Karlin [9], Scarf [15] and many others use Bayesian approach by assuming the demand process is characterized by a prior distribution on the unknown parameter. The second class makes no assumption regarding the parametric form of the unknown demand distribution. Based on the available demand data, either the empirical distribution (e.g., see [13]) or the bootstrapping method (e.g., see [3]) is applied to obtain an inventory control policy. Asymptotic optimality of inventory control policies without the prior knowledge of the demand distribution and their rates of convergence are studied in recent papers by Levi, Roundy and Shmoys [11] and Huh and Rusmevichientong [8].

Another line of research assumes that the decision maker knows the moments of the unknown demand distribution. When the unknown demand distribution is characterized by the first two moments, Scarf [16] derives a robust min-max inventory control policy. Further development and review of this model are given in [6] and [14]. However, one can argue that it is very difficult, if not impossible, for the decision maker to know the true demand parameters (e.g., mean and/or variance).

Liyanage and Shanthikumar [13] propose a new approach – operational statistics – to find a decision rule that maximizes the performance uniformly for all possible values of the unknown demand parameters. No prior knowledge on the parameter value is assumed in their model. They show that operational statistics outperforms the approach that separates estimation and optimization for the newsvendor inventory control problem when the unknown parameter is a scalar.

However, in their analysis, Liyanage and Shanthikumar obtain the optimal operational statistic by restricting the form of operational statistics considered to a small class of functions

that included the standard statistical estimators. Hence, it is not clear whether this is the best one can do in terms of achieving the maximum expected profit. It is also not clear how to apply operational statistics to other general problems. Therefore, one central question in operational statistics is how to find a class of operational statistics (that contains the standard statistical estimators and) within which one can find an operational statistic that maximizes the expected profit uniformly for all possible values of the unknown parameter (see Lim, Shanthikumar and Shen [12] for a formal definition of this to more general problems).

In this paper, we propose a Bayesian analysis to find the optimal operational statistic. The objective of a Bayesian analysis is to find an optimal value for the given sample data over possible values of the unknown parameter, which is quite different from the objective of operational statistics – to find an operational statistic that maximizes the expected profit over all possible samples for each fixed value of the unknown parameter. Nevertheless, by specifying a non-informative prior, we show the relationship between the objective values in operational statistics and in the Bayesian analysis, and show that for each possible sample data, the Bayesian analysis derives the value of the decision variable that is an optimal operational statistic.

The remainder of the paper is organized as follows. In Section 2, we present the formal model for operational statistics for the newsvendor inventory control problem when the unknown parameter is a scalar. In Section 3, we specify the non-informative prior and propose the Bayesian analysis that leads to the optimal mapping. In Section 4, we extend the results in Sections 2 and 3 to the case when both the location and the scale parameters of the demand distribution are unknown. Section 5 illustrates the solution procedure via two examples.

## 2 Model for Operational Statistics

Liyanage and Shanthikumar [13] propose the concept of operational statistics and apply it to a single period newsvendor inventory control problem. In this problem, items are purchased at  $c$  per unit and sold at  $s$  ( $s > c$ ) per unit. Without loss of generality, it can be assumed that the salvage value of the unsold items is zero. The demand  $D$  is a continuous

random variable with a density function  $f_D$ , distribution function  $F_D$  and a survival function  $\bar{F}_D = 1 - F_D$ . We assume that we know the form of  $f_D$  up to a scale parameter  $\theta$ . That is, we assume that  $Z = \frac{D}{\theta}$  has a known density function  $f_Z$  though we do not assume any prior information on  $\theta$ . Following [13], we will explicitly write  $f_D(\cdot|\theta)$  instead of  $f_D$  etc., to highlight its dependence on the parameter  $\theta$ . Note that  $f_D(v|\theta) = \frac{1}{\theta}f_Z(\frac{v}{\theta})$ ,  $F_D(v|\theta) = F_Z(\frac{v}{\theta})$  and  $\bar{F}_D(v|\theta) = \bar{F}_Z(\frac{v}{\theta})$ ,  $v \geq 0$ .

If the order quantity is  $y(\geq 0)$  units, then the profit can be calculated as  $(s \min\{D, y\} - cy)$ , and the expected profit is  $\phi(y, \theta) = E_\theta[s \min\{D, y\} - cy] = s \int_{v=0}^y \bar{F}_D(v|\theta)dv - cy$ . The optimal order quantity  $y^*(\theta) = \bar{F}_D^{inv}(\frac{c}{s}|\theta)$ , where  $\bar{F}_D^{inv}$  is the inverse of  $\bar{F}_D$ .

In operational statistics, we assume no knowledge on the unknown parameter  $\theta$ , and wish to find the optimal order quantity  $y$  based on data  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ . For ease of exposition, we assume that  $\{X_1, X_2, \dots, X_n\}$  is the historic demand data, where  $X_1, X_2, \dots, X_n$  and  $D$  are i.i.d. random variables. Note that the joint density function of  $\mathbf{X}$  is then

$$f_{\mathbf{X}}(\mathbf{x}|\theta) = \prod_{i=1}^n f_D(x_i|\theta) = \frac{1}{\theta^n} \prod_{i=1}^n f_Z(\frac{x_i}{\theta}), \mathbf{x} \in \mathcal{R}_+^n, (i.e., \mathbf{x} \geq \mathbf{0}). \quad (1)$$

The goal is to find an operational statistic  $g(\mathbf{X})$  of the data  $\mathbf{X}$  so that if we set the order quantity  $y = g(\mathbf{X})$  then the expected profit

$$E_\theta[\phi(g(\mathbf{X}), \theta)] = \int_{\mathbf{x} \geq \mathbf{0}} \phi(g(\mathbf{x}), \theta) f_{\mathbf{X}}(\mathbf{x}|\theta) d\mathbf{x} = \int_{\mathbf{x} \geq \mathbf{0}} \phi(g(\mathbf{x}), \theta) \frac{1}{\theta^n} \prod_{i=1}^n f_Z(\frac{x_i}{\theta}) d\mathbf{x} \quad (2)$$

is maximized for all  $\theta$ .

If we did not restrict the class of functions (operational statistics) from which  $g$  is chosen, a single function  $g$  that maximizes  $E_\theta[\phi(g(\mathbf{X}), \theta)]$  for all  $\theta$  may not exist. To see this, for a fixed value of  $\theta$ , say  $\theta_0$ , choosing  $g(\mathbf{x}) = y^*(\theta_0)$  we will maximize  $E_\theta[\phi(g(\mathbf{X}), \theta)]$  for  $\theta = \theta_0$  but not for other values of  $\theta$ . In fact, this function may work very poorly when  $\theta$  is far away from  $\theta_0$ .

Liyanage and Shanthikumar [13] consider the case of exponentially distributed demand with an unknown mean  $\theta$ . For this case, they find the operational statistic  $g$  that maximizes the expected profit  $E_\theta[\phi(g(\mathbf{X}), \theta)]$  for all  $\theta$  by restricting the operational statistic to a class of linear statistics. Specifically they allow  $g$  to be of the form  $g(\mathbf{x}) = \sum_{k=1}^n a_k(x_{[k]} - x_{[k-1]})$  (where  $x_{[k]}$  is the  $k$ -th order statistics of  $\{x_1, \dots, x_n\}$  and  $x_{[0]} = 0$ ). Within this class of

operational statistics, they find that  $g(\mathbf{x}) = ((\frac{s}{c})^{\frac{1}{n+1}} - 1) \sum_{i=1}^n x_i$  maximizes the expected profit over all  $\theta$  for exponentially distributed demand.

We will now consider a much larger class of operational statistics. Later we will show that we can find the optimal operational statistics within this class that maximizes the expected profit over all  $\theta$ . Suppose we decide to order  $g(\mathbf{x})$  units when we observe a demand vector  $\mathbf{x}$ . It is then very intuitive that if we observe a demand vector  $\alpha\mathbf{x}$  that is  $\alpha$  times the earlier demand vector, then we should order  $\alpha g(\mathbf{x})$ . Hence we restrict the operational statistics to the class of degree one homogeneous functions  $\mathcal{H}_1^+$  defined by

$$\mathcal{H}_1^+ = \{g : \mathcal{R}_+^n \rightarrow \mathcal{R}_+; g(\alpha\mathbf{x}) = \alpha g(\mathbf{x}), \alpha \geq 0, \mathbf{x} \geq \mathbf{0}\}.$$

Note that the class of functions considered in [13] is a small subset of the class of degree one homogeneous functions. Next we show that if we restrict the operational statistics to degree one homogeneous functions, then a single statistic will maximize the expected profit  $E_\theta[\phi(g(\mathbf{X}), \theta)]$  for all  $\theta$ . Formally, we have:

**Lemma 1** *If the operational statistics  $g_1$  and  $g_2$  are degree one homogeneous functions, then either  $E_\theta[\phi(g_1(\mathbf{X}), \theta)] \geq E_\theta[\phi(g_2(\mathbf{X}), \theta)]$  for all  $\theta$  or  $E_\theta[\phi(g_1(\mathbf{X}), \theta)] \leq E_\theta[\phi(g_2(\mathbf{X}), \theta)]$  for all  $\theta$ .*

Proof: Substituting  $v = \alpha v$  one finds that

$$\begin{aligned} \phi(\alpha y, \alpha \theta) &= s \int_0^{\alpha y} \bar{F}_D(v|\alpha\theta) dv - c(\alpha y) \\ &= \alpha s \int_0^y \bar{F}_D(\alpha v|\alpha\theta) dv - \alpha c y. \end{aligned}$$

Since  $\bar{F}_D(v|\theta) = \bar{F}_Z(\frac{v}{\theta})$  it is clear that  $\bar{F}_D(\alpha v|\alpha\theta) = \bar{F}_Z(\frac{v}{\theta}) = \bar{F}_D(v|\theta)$ ,  $\alpha > 0$ . Hence from the above equation we see that

$$\phi(\alpha y, \alpha \theta) = \alpha (s \int_0^y \bar{F}_D(v|\theta) dv - c y) = \alpha \phi(y, \theta).$$

That is  $\phi(y, \theta)$  is a degree one homogeneous function. Then substituting  $z_i = \frac{x_i}{\theta}$  into Equation 2 and using the degree one homogeneous property of  $\phi$ , we get

$$E_\theta[\phi(g(\mathbf{X}), \theta)] = \theta \int_{\mathbf{z} \geq \mathbf{0}} \phi(g(\mathbf{z}), 1) \Pi_{i=1}^n f_Z(z_i) dz = \theta E[\phi(g(\mathbf{Z}), 1)].$$

Since  $E[\phi(g(\mathbf{Z}), 1)]$  is independent of  $\theta$ , it is easy to see that

$$E[\phi(g_1(\mathbf{Z}), 1)] \leq E[\phi(g_2(\mathbf{Z}), 1)] \Leftrightarrow E_\theta[\phi(g_1(\mathbf{X}), \theta)] \leq E_\theta[\phi(g_2(\mathbf{X}), \theta)]$$

and vice versa. ■

Lemma 1 shows that the degree one homogeneous function  $g$  that maximizes

$$E[\phi(g(\mathbf{Z}), 1)] = \int_{\mathbf{z} \geq \mathbf{0}} \phi(g(\mathbf{z}), 1) \Pi_{i=1}^n f_Z(z_i) d\mathbf{z}$$

is the optimal operational statistic for our problem. Observe that the above integration over  $\{\mathbf{z} : \mathbf{z} \geq \mathbf{0}\}$  can be converted to the path integral over  $\{\mathbf{z} = r\mathbf{x} : \|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2} = 1; r \geq 0\}$ . It is

$$E[\phi(g(\mathbf{Z}), 1)] = \int_{\mathbf{x}: \|\mathbf{x}\|=1} \left[ \int_{r=0}^{\infty} \phi(g(r\mathbf{x}), 1) r^{n-1} \Pi_{i=1}^n f_Z(rx_i) dr \right] d\mathbf{x}.$$

Observing that  $\phi(g(r\mathbf{x}), 1) = r\phi(g(\mathbf{x}), \frac{1}{r})$ , substituting  $r = \frac{1}{\beta}$  in the above equation, we get

$$E[\phi(g(\mathbf{Z}), 1)] = \int_{\mathbf{x}: \|\mathbf{x}\|=1} \left[ \int_{\beta=0}^{\infty} \phi(g(\mathbf{x}), \beta) \frac{1}{\beta^{n+2}} \Pi_{i=1}^n f_Z\left(\frac{x_i}{\beta}\right) d\beta \right] d\mathbf{x}.$$

The optimal operational statistic can now be found for each  $\mathbf{x}$  by maximizing

$$\int_{\beta=0}^{\infty} \phi(g(\mathbf{x}), \beta) \frac{1}{\beta^{n+2}} \Pi_{i=1}^n f_Z\left(\frac{x_i}{\beta}\right) d\beta. \quad (3)$$

As we will see in the next section, the optimal solution to this can be found using a Bayesian analysis of the inventory problem with a non-informative prior.

### 3 Solving Operational Statistics via a Bayesian Analysis

In this section, we illustrate how to find the optimal degree one homogeneous function  $g$  that maximizes the expected profit via a Bayesian analysis.

To apply the Bayesian analysis, we need to specify two things: the objective function and the prior likelihood of the unknown parameter  $\Theta$ .

- For the objective function, we choose  $\Phi(y, \Theta) = \frac{\phi(y, \Theta)}{\Theta}$ .

Note that the objective function is not a function of data  $\mathbf{x}$  because all the sample information is summarized in the posterior likelihood function of  $\Theta$  and we are only trying to find a single point solution  $y$  for given data  $\mathbf{x}$ . We choose this objective function since it is normalized for the scale effect of the demand.

- For the prior, we choose Jeffrey's non-informative prior  $\pi(\theta) = \frac{1}{\theta}$  (e.g., see [10]). The joint likelihood of data  $\mathbf{x}$  and the parameter under this prior is

$$f_{(\mathbf{x}, \Theta)}(\mathbf{x}, \theta) = \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right).$$

Compare this to the joint density function of  $\mathbf{X}$  in Equation 1.

Intuitively, the ignorance of information about a scale parameter should be expressed as a flat prior on the log-scale. Box and Tiao [4] also motivate Jeffrey's general rule by arguing that under Jeffrey's prior, it is uniform for a parametrization where the likelihood is completely determined except for its location. More importantly, this prior, though not proper, leads to a proper posterior distribution. It is

$$f_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) = \left\{ \frac{1}{\eta(\mathbf{x})} \right\} \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right),$$

where

$$\eta(\mathbf{x}) = \int_{\theta=0}^{\infty} \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta,$$

is integrable and strictly positive for almost all  $\mathbf{x}$ . Integrability follows from the observation:

$$\int_{\mathbf{z} \geq \mathbf{0}} \prod_{i=1}^n f_Z(z_i) d\mathbf{z} = \int_{\mathbf{x}: \|\mathbf{x}\|=1} \int_{\theta=0}^{\infty} \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta d\mathbf{x} = 1.$$

Almost sure positivity follows from the observation:

$$\int_{\mathbf{x} \in \mathcal{S}} \int_{\theta=0}^{\infty} f_{(\mathbf{x}, \Theta)}(\mathbf{x}, \theta) d\theta d\mathbf{x} = 0,$$

for  $\mathcal{S} = \{\mathbf{x} : \int_{\theta=0}^{\infty} \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta = 0; \mathbf{x} \geq \mathbf{0}\}$ . This prior and the posterior ultimately provide the optimal solution for the original operational statistics problem.

The objective function for a given data  $\mathbf{x}$  is then

$$\begin{aligned} E[\Phi(y, \Theta) | \mathbf{X} = \mathbf{x}] &= \left\{ \frac{1}{\eta(\mathbf{x})} \right\} \int_{\theta=0}^{\infty} \Phi(y, \theta) \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta \\ &= \left\{ \frac{1}{\eta(\mathbf{x})} \right\} \int_{\theta=0}^{\infty} \phi(y, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta. \end{aligned}$$

Since  $\eta(\mathbf{x})$  is independent of  $y$  the objective function can be restated as

$$\int_{\theta=0}^{\infty} \phi(y, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta. \quad (4)$$

Therefore, given  $\mathbf{x}$ , we find a (point) solution  $y$  that maximizes  $E[\Phi(y, \Theta)|\mathbf{X} = \mathbf{x}]$ . Let this Bayesian solution be

$$y^* = \arg \max \left\{ \int_{\theta=0}^{\infty} \phi(y, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta : y \geq 0 \right\}$$

and define  $h^*(\mathbf{x}) = y^*$  for each  $\mathbf{x}$ . Comparing Equations 3 and 4 we see that if  $h^*$  is a degree one homogeneous function (as the next Lemma shows), then  $h^*$  is the optimal operational statistic.

**Lemma 2** *If*

$$h^*(\mathbf{x}) = \arg \max \left\{ \int_{\theta=0}^{\infty} \phi(y, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta : y \geq 0 \right\} \quad (5)$$

*then  $h^*$  is a degree one homogeneous function.*

Proof: Observe that

$$h^*(\alpha\mathbf{x}) = \arg \max \left\{ \int_{\theta=0}^{\infty} \phi(y, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{\alpha x_i}{\theta}\right) d\theta : y \geq 0 \right\}.$$

Substituting  $\theta = \alpha\theta$  in the above equation and using the homogeneous order one property of  $\phi$  we get

$$h^*(\alpha\mathbf{x}) = \arg \max \left\{ \frac{1}{\alpha^n} \int_{\theta=0}^{\infty} \phi\left(\frac{y}{\alpha}, \theta\right) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i}{\theta}\right) d\theta : y \geq 0 \right\}. \quad (6)$$

Comparing Equations (5) and (6), it is immediate that

$$h^*(\alpha\mathbf{x}) = \alpha h^*(\mathbf{x})$$

■

The following theorem is now obvious:

**Theorem 1**  *$h^*(x)$  obtained by the Bayesian analysis maximizes  $E_{\theta}[\phi(g(\mathbf{X}), \theta)]$  for all  $\theta$  among all order one homogeneous functions  $g$ .*

In Section 5 we provide two examples to the application of this result. In the next section we will extend this result to demand distributions with unknown location and scale parameters.

## 4 Extension - Demand Distribution with Two Unknown Parameters

In this section, we extend the analysis of Sections 2 and 3 to demand distributions with unknown location and scale parameters. We assume that we know the form of  $f_D$  up to a location parameter  $\tau$  and a scale parameter  $\theta$ . That is, we assume that  $Z = \frac{D-\tau}{\theta}$  has a known density function  $f_Z$  though we do not assume any prior information on  $\tau$  and  $\theta$ . We will explicitly write  $f_D(\cdot|\tau, \theta)$  instead of  $f_D$  etc., to highlight its dependence on the parameters  $\tau$  and  $\theta$ . Note that  $f_D(v|\tau, \theta) = \frac{1}{\theta}f_Z(\frac{v-\tau}{\theta})$ ,  $F_D(v|\tau, \theta) = F_Z(\frac{v-\tau}{\theta})$  and  $\bar{F}_D(v|\tau, \theta) = \bar{F}_Z(\frac{v-\tau}{\theta})$ ,  $v \geq 0$ .

Define the expected profit

$$\phi(y, \tau, \theta) = E_{\tau, \theta}[s \min\{D, y\} - cy] = s \int_{v=0}^y \bar{F}_D(v|\tau, \theta) dv - cy.$$

As before let  $\{X_1, X_2, \dots, X_n\}$  be the historic demand data, where  $X_1, X_2, \dots, X_n$  and  $D$  are i.i.d. random variables. Note that the joint density function of  $\mathbf{X}$  is then

$$f_{\mathbf{X}}(\mathbf{x}|\tau, \theta) = \prod_{i=1}^n f_D(x_i|\tau, \theta) = \frac{1}{\theta^n} \prod_{i=1}^n f_Z(\frac{x_i - \tau}{\theta}), \mathbf{x} \geq \mathbf{0}. \quad (7)$$

The goal is to find an operational statistics  $g(\mathbf{X})$  of the data  $\mathbf{X}$  so that if we set the order quantity  $y = g(\mathbf{X})$  then the expected profit

$$E_{\tau, \theta}[\phi(g(\mathbf{X}), \tau, \theta)] = \int_{\mathbf{x} \geq \mathbf{0}} \phi(g(\mathbf{x}), \tau, \theta) \frac{1}{\theta^n} \prod_{i=1}^n f_Z(\frac{x_i - \tau}{\theta}) d\mathbf{x} \quad (8)$$

is maximized for all  $\tau$  and  $\theta$ .

We will now consider the following class  $\mathcal{H}_1^e$  of operational statistics. This class is chosen based on the observation made in Section 2 and the following observation: Suppose we decide to order  $g(\mathbf{x})$  units when we observe a demand vector  $\mathbf{x}$ . It is then very intuitive that if we observe a demand vector  $\mathbf{x} - \delta \mathbf{e}$  (where  $\mathbf{e} = (1, 1, \dots, 1)'$ ), we should order  $g(\mathbf{x}) - \delta$ . Therefore we set

$$\mathcal{H}_1^e = \{g : \mathcal{R}^n \rightarrow \mathcal{R}; g(\alpha(\mathbf{x} - \delta \mathbf{e})) = \alpha(g(\mathbf{x}) - \delta), \alpha \geq 0, -\infty < \delta < \infty, \mathbf{x} \in \mathcal{R}^n\}.$$

Next we show that if we restrict the operational statistic to  $\mathcal{H}_1^e$ , then a single statistics will maximize the expected profit  $E_{\tau, \theta}[\phi(g(\mathbf{X}), \tau, \theta)]$  for all  $\tau$  and  $\theta$ . Formally, we have:

**Lemma 3** For any two operational statistics  $g_1 \in \mathcal{H}_1^e$  and  $g_2 \in \mathcal{H}_1^e$ , either  $E_{\tau,\theta}[\phi(g_1(\mathbf{X}), \tau, \theta)] \geq E_{\tau,\theta}[\phi(g_2(\mathbf{X}), \tau, \theta)]$  for all  $\tau$  and  $\theta$  or  $E_{\tau,\theta}[\phi(g_1(\mathbf{X}), \tau, \theta)] \leq E_{\tau,\theta}[\phi(g_2(\mathbf{X}), \tau, \theta)]$  for all  $\tau$  and  $\theta$ .

Proof: Substituting  $v = \alpha(v - \delta)$  one finds that

$$\begin{aligned}\phi(\alpha(y - \delta), \alpha(\tau - \delta), \alpha\theta) &= s \int_{v=0}^{\alpha(y-\delta)} \bar{F}_D(v|\alpha(\tau - \delta), \alpha\theta)dv - c(\alpha(y - \delta)) \\ &= \alpha s \int_{v=0}^y \bar{F}_D(\alpha(v - \delta)|\alpha(\tau - \delta), \alpha\theta)dv - \alpha cy - \alpha(s - c)\delta.\end{aligned}$$

Since  $\bar{F}_D(v|\tau, \theta) = \bar{F}_Z(\frac{v-\tau}{\theta})$  it is clear that  $\bar{F}_D(\alpha(v-\delta)|\alpha(\tau-\delta), \alpha\theta) = \bar{F}_Z(\frac{v-\tau}{\theta}) = \bar{F}_D(v|\tau, \theta)$ ,  $\alpha > 0$ ,  $-\infty < \delta < \infty$ . Hence from the above equation we see that

$$\begin{aligned}\phi(\alpha(y - \delta), \alpha(\tau - \delta), \alpha\theta) &= \alpha(s \int_{v=0}^y \bar{F}_D(v|\tau, \theta)dv - cy) - \alpha(s - c)\delta \\ &= \alpha\phi(y, \tau, \theta) - \alpha(s - c)\delta.\end{aligned}$$

Then substituting  $z_i = \frac{x_i - \tau}{\theta}$  into Equation 8 and using the above property of  $\phi$  with  $\delta = \tau$  and  $\alpha = \frac{1}{\theta}$ , we get

$$\begin{aligned}E_{\tau,\theta}[\phi(g(\mathbf{X}), \tau, \theta)] &= \theta \int_{\mathbf{z} \in \mathcal{R}^n} \{\phi(g(\mathbf{z}), 0, 1) + \frac{1}{\theta}(s - c)\tau\} \Pi_{i=1}^n f_Z(z_i) d\mathbf{z} \\ &= \theta E[\phi(g(\mathbf{Z}), 0, 1)] + (s - c)\tau.\end{aligned}$$

Since  $E[\phi(g(\mathbf{Z}), 0, 1)]$  is independent of  $\theta$ , it is easy to see that

$$E[\phi(g_1(\mathbf{Z}), 0, 1)] \leq E[\phi(g_2(\mathbf{Z}), 0, 1)] \Leftrightarrow E_{\tau,\theta}[\phi(g_1(\mathbf{X}), \tau, \theta)] \leq E_{\tau,\theta}[\phi(g_2(\mathbf{X}), \tau, \theta)]$$

and vice versa. ■

Lemma 3 shows that the function  $g \in \mathcal{H}_1^e$  that maximizes

$$E[\phi(g(\mathbf{Z}), 0, 1)] = \int_{\mathbf{z} \in \mathcal{R}^n} \phi(g(\mathbf{z}), 0, 1) \Pi_{i=1}^n f_Z(z_i) d\mathbf{z}$$

is the optimal operational statistics for our problem. Observe that the above integration over  $\{\mathbf{z} : \mathbf{z} \in \mathcal{R}^n\}$  can be converted to the path integral over  $\{\mathbf{z} = r\mathbf{x} - t\mathbf{e} : \|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2} = 1; |\mathbf{x}| = \sum_{i=1}^n x_i = 0; -\infty < t < \infty; r \geq 0\}$ . It is

$$E[\phi(g(\mathbf{Z}), 0, 1)] = \int_{\mathbf{x}: \|\mathbf{x}\|=1; |\mathbf{x}|=0} \left[ \int_{t=-\infty}^{\infty} \left[ \int_{r=0}^{\infty} \phi(g(r\mathbf{x} - t\mathbf{e}), 0, 1) r^{n-2} \Pi_{i=1}^n f_Z(rx_i - t) dr \right] dt \right] d\mathbf{x}.$$

Observing that  $\phi(g(\mathbf{r}\mathbf{x} - t\mathbf{e}), 0, 1) = r\phi(g(\mathbf{x}), \frac{t}{r}, \frac{1}{r}) - \frac{1}{r}(s-c)t$ , substituting  $u = \frac{t}{r}$  and  $\beta = \frac{1}{r}$  in the above equation, we get

$$E[\phi(g(\mathbf{Z}), 0, 1)] = \int_{\mathbf{x}: \|\mathbf{x}\|=1; \|\mathbf{x}\|=0} \left[ \int_{u=-\infty}^{\infty} \left[ \int_{\beta=0}^{\infty} \phi(g(\mathbf{x}), u, \beta) \frac{1}{\beta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i - u}{\beta}\right) d\beta \right] du \right] d\mathbf{x}.$$

The optimal operational statistic can now be found for each  $\mathbf{x}$  by maximizing

$$\int_{u=-\infty}^{\infty} \int_{\beta=0}^{\infty} \phi(g(\mathbf{x}), u, \beta) \frac{1}{\beta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i - u}{\beta}\right) d\beta du. \quad (9)$$

As we will see next, the optimal solution to this can be found using a Bayesian analysis of the inventory problem with a non-informative prior. For this

- we choose the objective function  $\Phi(y, \Gamma, \Theta) = \frac{\phi(y, \Gamma, \Theta) - (s-c)\Gamma}{\Theta}$  which is normalized for the location and scale effect of the demand and
- we choose Jeffrey's non-informative prior  $\pi(\tau, \theta) = \frac{1}{\theta}, -\infty < \tau < \infty; \theta > 0$  (e.g., see [10]). The joint likelihood of data  $\mathbf{x}$  and the parameter under this prior is

$$f_{(\mathbf{x}, \Gamma, \Theta)}(\mathbf{x}, \tau, \theta) = \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right).$$

More importantly, this prior, though not proper, leads to a proper posterior distribution. It is

$$f_{(\Gamma, \Theta)|\mathbf{X}}(\tau, \theta|\mathbf{x}) = \left\{ \frac{1}{\eta(\mathbf{x})} \right\} \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right),$$

where

$$\eta(\mathbf{x}) = \int_{\tau=-\infty}^{\infty} \int_{\theta=0}^{\infty} \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right) d\theta d\tau,$$

is integrable and strictly positive for almost all  $\mathbf{x}$  (can be verified using observations similar to those for the scale parameter case). This prior and the posterior ultimately provides the optimal solution for the original operational statistics problem.

The objective function for a given data  $\mathbf{x}$  is then

$$\begin{aligned} E[\Phi(y, \Gamma, \Theta)|\mathbf{X} = \mathbf{x}] &= \left\{ \frac{1}{\eta(\mathbf{x})} \right\} \int_{\tau=-\infty}^{\infty} \int_{\theta=0}^{\infty} \Phi(y, \tau, \theta) \frac{1}{\theta^{n+1}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right) d\theta d\tau \\ &= \left\{ \frac{1}{\eta(\mathbf{x})} \right\} \int_{\tau=-\infty}^{\infty} \int_{\theta=0}^{\infty} \phi(y, \tau, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right) d\theta d\tau. \end{aligned}$$

Since  $\eta(\mathbf{x})$  is independent of  $y$  the objective function can be restated as

$$\int_{\tau=-\infty}^{\infty} \int_{\theta=0}^{\infty} \phi(y, \tau, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right) d\theta d\tau. \quad (10)$$

Therefore, given  $\mathbf{x}$ , we find a (point) solution  $y$  that maximizes  $E[\Phi(y, \Gamma, \Theta) | \mathbf{X} = \mathbf{x}]$ . Let this Bayesian solution be

$$y^* = \arg \max \left\{ \int_{\tau=-\infty}^{\infty} \int_{\theta=0}^{\infty} \phi(y, \tau, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right) d\theta d\tau : y \geq 0 \right\},$$

and define  $h^*(\mathbf{x}) = y^*$  for each  $\mathbf{x}$ . Comparing Equations 9 and 10 we see that if  $h^* \in \mathcal{H}_1^e$  (as the next Lemma shows), then  $h^*$  is the optimal operational statistic.

**Lemma 4** *If*

$$h^*(\mathbf{x}) = \arg \max \left\{ \int_{\tau=-\infty}^{\infty} \int_{\theta=0}^{\infty} \phi(y, \tau, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right) d\theta d\tau : y \geq 0 \right\} \quad (11)$$

*then*  $h^* \in \mathcal{H}_1^e$ .

Proof: Observe that

$$h^*(\alpha(\mathbf{x} - \delta \mathbf{e})) = \arg \max \left\{ \int_{\tau=-\infty}^{\infty} \int_{\theta=0}^{\infty} \phi(y, \tau, \theta) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{\alpha(x_i - \delta) - \tau}{\theta}\right) d\theta d\tau : y \geq 0 \right\}.$$

Substituting  $\tau = \alpha(\tau - \delta)$  and  $\theta = \alpha\theta$  in the above equation and using the property of  $\phi$  we get

$$h^*(\alpha(\mathbf{x} - \delta \mathbf{e})) = \arg \max \left\{ \frac{1}{\alpha^{n-1}} \int_{\tau=-\infty}^{\infty} \int_{\theta=0}^{\infty} \phi\left(\frac{y - \delta}{\alpha}, \tau, \theta\right) \frac{1}{\theta^{n+2}} \prod_{i=1}^n f_Z\left(\frac{x_i - \tau}{\theta}\right) d\theta d\tau : y \geq 0 \right\}. \quad (12)$$

Comparing Equations (11) and (12), it is immediate that

$$h^*(\alpha(\mathbf{x} - \delta \mathbf{e})) = \alpha(h^*(\mathbf{x}) - \delta).$$

■

The following theorem is now obvious:

**Theorem 2**  *$h^*(x)$  obtained by the Bayesian analysis maximizes  $E_{\tau, \theta}[\phi(g(\mathbf{X}), \tau, \theta)]$  for all  $\tau$  and  $\theta$  among all functions  $g \in \mathcal{H}_1^e$ .*

## 5 Examples

In this section, we solve the newsvendor inventory control problem using Bayesian analysis for two demand distributions: exponential and uniform distributions with unknown scale parameter  $\theta$ .

### 5.1 Exponential Distribution

In this section we consider the case  $f_Z(v) = \exp\{-v\}$ ,  $v \geq 0$ . It is then easy to compute that  $\phi(y, \theta) = s\theta(1 - \exp\{-\frac{y}{\theta}\}) - cy$ . Observing that  $\prod_{i=1}^n f_Z(\frac{x_i}{\theta}) = \exp\{-\frac{\sum_{i=1}^n x_i}{\theta}\}$ ,  $\theta > 0$ , the objective function is then (see Equation 4)

$$\begin{aligned} & \int_{\theta=0}^{\infty} \{s\theta(1 - \exp\{-\frac{y}{\theta}\}) - cy\} \frac{1}{\theta^{n+2}} \{\exp\{-\frac{\sum_{i=1}^n x_i}{\theta}\}\} d\theta \\ &= s \left\{ \frac{(n-1)!}{(\sum_{i=1}^n x_i)^n} - \frac{(n-1)!}{(y + \sum_{i=1}^n x_i)^n} \right\} - cy \left\{ \frac{n!}{(\sum_{i=1}^n x_i)^{n+1}} \right\}. \end{aligned}$$

The optimal value of  $y$  that maximizes the above objective is  $y^* = ((\frac{s}{c})^{\frac{1}{n+1}} - 1) \sum_{i=1}^n x_i$ . Hence the optimal operational statistic is

$$h^*(\mathbf{x}) = ((\frac{s}{c})^{\frac{1}{n+1}} - 1) \sum_{i=1}^n x_i, \mathbf{x} \geq \mathbf{0}.$$

The solution of this example shows that the solution in [13] is in fact optimal among all degree one homogeneous operational statistics.

### 5.2 Uniform Distribution

In this section we consider the case  $f_Z(v) = 1$ ,  $0 \leq v \leq 1$ . It is then easy to compute that  $\phi(y, \theta) = s(y - \frac{y^2}{2\theta}) - cy$ ,  $y \leq \theta$  and  $\phi(y, \theta) = s\frac{\theta}{2} - cy$ ,  $y \geq \theta$ . Observing that  $\prod_{i=1}^n f_Z(\frac{x_i}{\theta}) = 1$ ,  $\theta \geq x_{[n]}$ , (recall that  $x_{[n]} = \max\{x_1, \dots, x_n\}$ ), the objective function is then (see Equation 4)

$$\begin{aligned} & \int_{\theta=x_{[n]}}^{\infty} \{s(y - \frac{y^2}{2\theta}) - cy\} \frac{1}{\theta^{n+2}} d\theta \\ &= s \left\{ \frac{y}{(n+1)x_{[n]}^{n+1}} - \frac{y^2}{2(n+2)x_{[n]}^{n+2}} \right\} - c \left\{ \frac{y}{(n+1)x_{[n]}^{n+1}} \right\}, y \leq x_{[n]} \end{aligned}$$

and

$$\begin{aligned}
& \int_{\theta=x_{[n]}}^y \left\{ s \frac{\theta}{2} - cy \right\} \frac{1}{\theta^{n+2}} d\theta + \int_{\theta=y}^{\infty} \left\{ s \left( y - \frac{y^2}{2\theta} \right) - cy \right\} \frac{1}{\theta^{n+2}} d\theta \\
= & s \left\{ \frac{1}{2nx_{[n]}^n} - \frac{1}{2ny^n} \right\} - c \left\{ \frac{y}{(n+1)x_{[n]}^{n+1}} - \frac{1}{(n+1)y^n} \right\} \\
+ & s \left\{ \frac{1}{(n+1)y^n} - \frac{1}{2(n+2)y^n} \right\} - c \left\{ \frac{1}{(n+1)y^n} \right\}, y \geq x_{[n]}.
\end{aligned}$$

The optimal value of  $y$  that maximizes the above objective is  $y^* = \frac{n+2}{n+1} \left( 1 - \frac{c}{s} \right) x_{[n]}$  if  $\frac{c}{s} \geq \frac{1}{n+2}$  (that corresponds to  $y^* \leq x_{[n]}$ ) and  $y^* = \left( \frac{s}{(n+2)c} \right)^{\frac{1}{n+1}} x_{[n]}$  if  $\frac{c}{s} \leq \frac{1}{n+2}$  (that corresponds to  $y^* \geq x_{[n]}$ ). Hence the optimal operational statistic is

$$h^*(\mathbf{x}) = \frac{n+2}{n+1} \left( 1 - \frac{c}{s} \right) x_{[n]}, \quad n \geq \frac{s}{c} - 2$$

and

$$h^*(\mathbf{x}) = \left( \frac{s}{(n+2)c} \right)^{\frac{1}{n+1}} x_{[n]}, \quad n \leq \frac{s}{c} - 2.$$

## References

- [1] K. S. Azoury, Bayes solution to dynamic inventory models under unknown demand distribution, *Mgt. Sci.*, **31** (1985) 1150–1160.
- [2] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer, New York, 1980.
- [3] J. H. Bookbinder, A. E. Lordahl, Estimation of inventory reorder level using the bootstrap statistical procedure, *IIE Trans.*, **21** (1989) 302–312.
- [4] G. E. P. Box, G. C. Tiao, *Bayesian Inference in Statistical Analysis*, John Wiley & Sons, New York, 1973 (reprinted in 1992).
- [5] G. D. Eppen, A. V. Iyer, Improved Fashion Buying with Bayesian Updates, *Operations Research*, **45** (1997) 805–819.
- [6] G. Gallego, I. Moon, The distribution free newsvendor problem: review and extensions, *J. Opnl. Res. Soc.*, **44** (1993) 825–834.

- [7] C. Hakzoz, S. Seshadri, Monotone forecasts, *Operations Research.*, **52** (2004) 478–486.
- [8] W. T. Huh, P. Rusmevichientong, An Asymptotic Analysis of Inventory Planning with Censored Demand, Technical Report, Department of Industrial Engineering and Operations Research, Columbia University, New York, 2006.
- [9] S. Karlin, Dynamic inventory policy with varying stochastic demands, *Mgt. Sci.*, **6** (1960) 231–258.
- [10] R. Kass, L. Wasserman, Formal rules for selecting prior distributions: A review and annotated bibliography, *Journal of the American Statistical Association*, **91** (1996) 1343–1370.
- [11] R. Levi, R. Roundy, D. B. Shmoys. 2005. Computing Provably Near-Optimal Sample-Based Policies for Stochastic Inventory Control Models, ORIE, Cornell University, Working Paper, 2005.
- [12] A. E. B. Lim, J. G. Shanthikumar, Z. J. M. Shen, Model Uncertainty, Robust Optimization, and Learning, *Tutorials in Operations Research: Models, Methods, and Applications for Innovative Decision Making*, (eds.) M. P. Johnson, B. Norman and N. Secomandi, (2006) 66–94.
- [13] L. H. Liyanage, J. G. Shanthikumar, A practical inventory control policy using operational statistics, *Operations Research Letters*, **33**(4) (2005) 341–348.
- [14] G. Perakis, G. Roels, The Price of Information: Inventory Management with limited information about demand, Working paper, MIT, 2005.
- [15] H. Scarf, A min-max solution of an inventory problem, *Studies in The Mathematical Theory of Inventory and Production*, (Chapter 12) Stanford University Press, Stanford, CA 1958.
- [16] H. Scarf, Bayes Solutions of statistical inventory problem, *Annals of Math. Stats.*, **30** (1959) 490–508.
- [17] M. J. Schervish, *Theory of Statistics*, Springer, New York, 1995.