European Journal of Operational Research 245 (2015) 623–628

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Short Communication

On the relationship between entropy, demand uncertainty, and expected loss

Adam J. Fleischhacker [a,*], Pak-Wing Fok [b]

[a] Department of Business Administration, University of Delaware, Newark, DE 19716, United States
[b] Department of Mathematics, University of Delaware, Newark, DE 19716, United States

ARTICLE INFO

Article history:
Received 9 January 2014
Accepted 9 March 2015
Available online 16 March 2015

Keywords:
Entropy
Measuring demand uncertainty
Stochastic inventory models
Demand variability

ABSTRACT

We analyze the effect of demand uncertainty, as measured by entropy, on expected costs in a stochastic inventory model. Existing models studying demand variability's impact use either stochastic ordering techniques or use variance as a measure of uncertainty. Due to both axiomatic appeal and recent use of entropy in the operations management literature, this paper develops entropy's use as a demand uncertainty measure. Our key contribution is an insightful proof quantifying how costs are non-increasing when entropy is reduced.

1. Background

Entropy is a measure of uncertainty strongly advocated by Jaynes (2003) and originally popularized by Shannon (1948). Let D be a discrete random variable with probability mass function $\mathbf{p} = (p_1, \ldots, p_N)$, then entropy $H(\mathbf{p})$ is defined as

$$H(p_1, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log(p_i). \quad (1)$$

A good introduction to entropy as a measure of uncertainty can be found in Abbas (2006). Maximizing Eq. (1) subject to constraints based on existing knowledge (e.g. the mean, support, moments, etc.) is known as the maximum entropy principle (Jaynes, 1957), and is considered a uniquely correct method for inductive inference (Johnson & Shore, 1983; Shore & Johnson, 1980).

Our work in this paper solidifies the theoretical connection between entropy as a measure of demand uncertainty and expected loss. Despite the concept of entropy being around for over 60 years, the operations management literature has been slow to adopt this uncertainty measure. Maglaras and Eren (2015) comment that "to the best of our knowledge, the operations management and revenue management literatures have not explored the use of ME (maximum entropy) techniques to approximate unknown demand or willingness-to-pay distributions." Recently, however, entropy is being explored. References to entropy within both operations management (see for example Andersson, Jörnsten, Nonås, Sandal, & Ubøe, 2013; Maglaras & Eren, 2015; Perakis & Roels, 2008; Shuiabi, Thomson, & Bhuiyan, 2005) and related contexts are increasing. This includes the contexts of pricing models (Lim & Shanthikumar, 2007), portfolio optimization (Glasserman & Xu, 2013), and discrete optimization (Nakagawa, James, Rego, & Edirisinghe, 2013). Of particular interest, Andersson et al. (2013) numerically demonstrate promising performance characteristics of using entropy-based demand distributions for ordering decisions; our work complements this insight by theoretically connecting entropy and loss in a similar setting.

2. Entropy versus alternative demand uncertainty measures

Existing theoretical connections between uncertainty and supply/demand mismatch costs are often based on a measure of spread or on stochastic ordering techniques. In this work we use entropy to connect uncertainty and expected loss in a way that is intuitively satisfying, facilitates numerical evaluation of uncertainty's effects on mismatch costs, and is also theoretically justified. In addition, it enables the exploration of uncertainty reduction without restricting expected demand to remain constant as uncertainty is increased/decreased. For example, a retailer reducing demand uncertainty by transitioning from hi-lo pricing to everyday low pricing will most certainly see a change in expected demand (see discussion in Lee, Padmanabhan, & Whang, 1997).

Variance is arguably the most preferred measure of demand uncertainty found in the inventory management literature (see for example Kwak & Gavirneni, 2011; Ridder, Laan, & Salomon, 1998; Taylor & Xiao, 2010). While intuitively, variance provides meaning regarding the spread of a distribution and hence, uncertainty; problems arise

* Corresponding author. Tel.: +1 302 831 6966.
E-mail addresses: ajf@udel.edu (A.J. Fleischhacker), pakwing@udel.edu (P.-W. Fok).

http://dx.doi.org/10.1016/j.ejor.2015.03.014
0377-2217/

when one tries to operationalize variance as a measure of uncertainty (Jaynes, 2003, p. 345). In an often cited example, imagine a six-sided die with a known bias such that the expected value of a roll is 4.5 instead of 3.5. How can we assign probabilities to the six outcomes in this clearly under-specified problem? Following Jaynes' logic, we should pick the probability assignment implying maximal uncertainty subject to the problem's constraints. In other words, if every feasible probability assignment were to be evaluated by a numerical measure of uncertainty, then one chooses the assignment that has the largest measurement; to choose otherwise would imply additional information beyond that which is available. Thus, applying variance as an uncertainty measure to the six-sided die problem, assign probability, p_i, by maximizing $\mathrm{Var}(\{p_i\}) = \sum_{i=1}^{6} p_i(i-9/2)^2$ subject to $\sum_{i=1}^{6} p_i = 1$ and $\sum_{i=1}^{6} ip_i = 9/2$. Solving this maximization problem, all probability is placed on the outcomes of one and six with $p_1 = 0.3$ and $p_6 = 0.7$. Unfortunately, this extreme placement of probabilities leads to dissatisfaction with the implied results. Should not some probability remain on outcomes of 2, 3, 4, and 5? Variance as an uncertainty measure leads to these counter-intuitive results and thus, maximal uncertainty and maximal variance are not equal.

More intuitive results are found when entropy is used in place of variance. Mathematically, we maximize $-\sum_{i=1}^{6} p_i \log p_i$ subject to $\sum_{i=1}^{6} p_i = 1$ and $\sum_{i=1}^{6} ip_i = 9/2$. The Lagrange function is

$$\mathcal{L} \equiv -\sum_{i=1}^{6} p_i \log p_i - \lambda_1 \left(\sum_{i=1}^{6} p_i - 1 \right) - \lambda_2 \left(\sum_{i=1}^{6} ip_i - 9/2 \right), \tag{2}$$

where λ_1 and λ_2 are Lagrange multipliers and log is always the natural logarithm in this paper. Maximizing $\mathcal{L}$, we find maximum uncertainty corresponds to $p_1 \approx 5.4\%$, $p_2 \approx 7.9\%$, $p_3 \approx 11.4\%$, $p_4 \approx 16.5\%$, $p_5 \approx 24.0\%$, and $p_6 \approx 34.7\%$. Outcomes of 2–5 are no longer an impossibility and much more in line with what common sense might dictate.

Besides entropy and variance (see Ebrahimi, Maasoumi, & Soofi, 1999, for further comparison of entropy and variance), other alternative uncertainty measures exist. A popular alternative to study the effects of demand variability is to use stochastic ordering criteria as done in Jemaï and Karaesmen (2005), Song (1994), Song, Zhang, Hou, and Wang (2010), and Xu, Chen, and Xu (2010). In all of these works, the authors confirm the intuition that uncertainty generally leads to increased costs. Though, specific cases where larger uncertainty is associated with reduced costs can also be found (Ridder et al., 1998). The previously mentioned works use of stochastic ordering lead to qualitative insight, but do not facilitate numerical connection of uncertainty and mismatch costs. Gerchak and Mossman (1992) enable more numerically driven computations involving uncertainty through use of the mean preserving transformation. Unfortunately, the mean preserving transformation, like the previously mentioned stochastic ordering techniques, assumes uncertainty reduction never leads to changes in expected demand. Common sense requires expected demand changes may indeed be a possible outcome of a demand uncertainty reduction effort; otherwise, why do it? Hence, the extant demand variability literature is largely void of a method of measuring uncertainty that enables us to numerically relate uncertainty and expected mismatch costs, is consistent with common sense, and provides quantifiable evaluation of uncertainty's effect on expected loss.

Our use of entropy and its comparison to variance connotes mathematical risk. However, entropy is a measure of randomness and not a direct measure of risk; entropy's calculation is independent of preferences among potential outcomes of demand. In contrast, risk measures, such as conditional value at risk (CVaR) or expected shortfall (ES) (see Szegö, 2005, Section 5), always depend on the distribution of losses. In this way, a reduction in a risk measure, by definition, quantifies a reduction in the "likelihood of loss or less than expected returns" (see McNeil, Frey, & Embrechts, 2010, chap. 1). Intuitively, a reduction in uncertainty should also lead to a reduction in the likelihood of less than expected returns, but this notion needs to be proved and that is the goal of this work. In addition, this work's focus on uncertainty reduction enables a framework from which to value information, whereas the more common risk-centric focus is better geared toward optimizing ordering decisions (see related discussion in Choi, Ruszczynski, & Zhao, 2011, and references therein).

3. Entropy and expected loss

In this paper, we quantify the relationship between expected supply/demand mismatch costs and entropy in the context of a general stochastic inventory problem. We commence our analysis defining maximal uncertainty and subsequently examine the effects of reducing uncertainty. Consider the loss matrix $\mathbf{A} \in \mathbb{R}^{+N\times N}$ whose elements represent the expected loss associated with N possible ordering decisions and N possible outcomes (e.g. demand realizations). An ordering decision $j \in \{1, 2, \dots, N\}$ is chosen such that the loss, $\mathcal{L}_j(\mathbf{p})$, is minimized

$$\mathcal{L}_{\min}(\mathbf{p}) = \min_{1\le j\le N} \mathbf{q}_j^T \mathbf{A}\mathbf{p}, \tag{3}$$

where demand distribution $\mathbf{p} = (p_1, p_2, \dots, p_N)$ is unknown, and $\mathbf{q}_j$ is the jth unit (column) vector representing an order quantity of j units.

The determination of the optimal order quantity j^* and associated loss $\mathcal{L}_{j^*}(\mathbf{p})$ requires specification of the probability distribution $\mathbf{p}$. Since, we do not know $\mathbf{p}$ and any of an infinite number of N-tuple probabilities $(p_1, p_2, \dots, p_N)$ may be valid representations of $\mathbf{p}$, we find the $\mathbf{p}$ consistent with maximal uncertainty via the principle of maximum entropy: maximize $\sum_{i=1}^{N} p_i \log p_i$, subject to $p_1 + p_2 + \cdots + p_N = 1$, with the unique solution $p_1 = p_2 = \cdots = p_N = 1/N$. The corresponding expected loss is $\mathcal{L}_{\min}(\mathbf{e}/N) = \mathbf{q}_{j^*}^T \mathbf{A}\mathbf{e}/N$ where $\mathbf{e}$ is the vector consisting of all 1s.

With the point of maximal uncertainty (i.e. entropy) defined, we now investigate reducing levels of entropy. To do so, we define the set of all probability distributions consistent with each possible level of uncertainty:

$$S(h) \equiv \left\{ (p_1, \dots, p_N) : H(\mathbf{p}) = h, \right. \\ \left. \sum_{i=1}^{N} p_i = 1, 0 < p_j < 1, j = 1, \dots, N \right\}, \tag{4}$$

where $H(\mathbf{p}) = H(p_1, \dots, p_N) = -\sum_{i=1}^{N} p_i \log p_i$. For a given entropy level h, Eq. (4) yields all N-tuples $(p_1, p_2, \dots, p_N)$ that are consistent with this level of uncertainty. Based on the combinatorial derivation of entropy provided in Niven (2007), all N-tuple's of equal entropy are considered equiprobable. The expected loss subject to the constant entropy constraint $H(p_1, p_2, \dots, p_N) = h$ is therefore an average over all possible N-tuples that lie on S:

$$\mathcal{E}[h] \equiv E[\mathcal{L}_{\min}(\mathbf{p}); S(h)] = \frac{\int_{S(h)} \mathcal{L}_{\min}(\mathbf{p}) ds}{\int_{S(h)} ds}. \tag{5}$$

Note that Eq. (5) is actually an expectation of expected losses.

Some isoentropy surfaces for the $N = 3$ case are shown in Fig. 1. For the general N-decision case, the constant h isoentropy surfaces are $N-2$ dimensional hypershells lying in the $N-1$ dimensional simplex $\sum_{i=1}^{N} p_i = 1$, $0 < p_i < 1$. In Fig. 1(a) we see that for values of h near the maximal value log 3, the isoentropy contour is approximately circular, but for smaller values, they become more triangular. The contours are also geometrically symmetric in p_1, p_2 and p_3, as expected.

In Fig. 1(b), we show that for different values of h, the isoentropy contour can pass through different "loss regions" (separated by the thin red lines) with different optimal order quantities (OQs). Loss region D_2 consists of all 3-tuples (p_1, p_2, p_3) where an OQ of 2 gives the smallest loss. In loss regions D_1 and D_3, OQ of 1 and 3 give the smallest