INVENTORY MANAGEMENT PROBLEM WITH BAYESIAN DYNAMIC PROGRAMMING

Saloni Srivastava*

Dr. R.K. Shrivastava**

ABSTRACT

Considering unobserved lost sales, the application of the envelope theorem is given in a Bayesian inventory management problem. Our intension is to show that the optimal inventory level with unobserved lost sales is greater than the optimal inventory level with observed lost sales. Under the continuous demand distribution, we prove this result and also show that the results can be easily extended to the Markov-modulated demand process.

*Department of Applied Science and Humanities, Sachdeva Institute of Technology, Farah, Mathura.

**Department of Mathematics, Agra College, Agra

1. INTRODUCTION

Decision-making in an uncertain and dynamic environment are mainly found under the business problems. For having optimal decisions, there is a need of using the historical observations to estimate certain system parameters. To improve the estimation and enhance the overall system performance one needs to take active actions to get the information. Bayesian dynamic program can be formulated as a problem concerning joint parameter estimation and optimal control. In the inventory management literature, Scarf(1959, 1960) first used this approach to study optimal inventory policy under a demand distribution with an unknown parameter. On this topic recent work includes Azoury(1984), Lovejoy(1990), Eppen and Iyer(1997), Lariviere and Porteus(1999), Ding et al(2002), Lu et al(2005, 2006, 2008), Bensoussan et al (2007 a,b), Bisi and Dada(2007 a,b), Chen and Plambeck(2008), and Chen (2008). The problem can also be modeled as a partially-observed Markov decision process (POMDP) (e.g., Lovejoy 1993; Treharne and Sox 2002) under the discrete demand. Huh and Rusmevichientong (2009) studied an alternative non-parametric approach to this problem.

In this paper, initial step includes a general Bayesian dynamic program problem. The problem setting includes the un-normalized posterior to linearize the optimality equations. The product of the likelihood function and the prior distribution (results no normalization performed) results the un-normalized posterior.

We show that the original problem is equivalent to the modified problem with the un-normalized prior/posterior in that they share the same optimal solutions. This un-normalized prior/posterior technique was first introduced by Bensoussan et al, (2007 a,b) to study an infinite-horizon inventory planning problem under the perishable inventory assumption.

In this paper, we apply the technique to a more general Bayesian dynamic program problem (which allows for non-perishable inventory) and establish the equivalence result. We prove a generalized envelope theorem for the problems with the un-normalized prior/posterior.

In the microeconomics literature, the classic envelope theorem states that the derivative of a single-period optimal value function with respect to a parameter is equal to the derivative with the objective function fixed the original optimal decision point (Mass-Colell et al, 1995, p.p 964-966). Using the Bayesian dynamic program setting, we show that the derivative of the value function with respect to a parameter is equal to the derivative with the dynamic program objective function fixed at the original optimal control policy on the entire sample path. Here, we want to fix not only the optimal decision for the current period but also the optimal reliant decisions along the sample path also we apply the generalized envelope theorem result to a Bayesian inventory management problem with unobserved lost sales.

Chen and Plambeck (2008), have shown that under the non-perishable inventory assumption, the optimal inventory level with observed lost sales is less than the optimal inventory level with unobserved lost sales and also they establish the result under a general discrete demand distribution. Here, we extending their work by showing that the result also holds under a general continuous demand distribution. With the use of the generalized envelope theorem our proof becomes simpler than the discrete-demand version.

Under the assumption that the random events or demands are independent and identically-distributed, we present our result. Further, we prove that the result can be easily extended to a more general Markov-modulated process, which allows for dependency across different periods. We know that the generalized envelope theorem is useful to know the dynamic pricing and learning about consumer behavior.

2. AN ENVELOPE THEOREM FOR BAYESIAN DYNAMIC PROGRAM

Considering a dynamic program problem with N decision periods and the index of each period is denoted by n, with the initial period indexed by n = 1.

For all n, a random event Z_n occurs with the outcome of the event denoted by z_n. The probability density of Z_n , denoted by $f(z|\phi)$ has an unknown parameter ϕ with $\phi \epsilon \Theta$. The parameter ϕ is subject to a prior distribution $\pi_n(\phi)$. Without loss of generality let us consider that Z_n is independently identically-distributed (IID) means that ϕ is an unknown constant.

Let us assume that x_n be the actions taken in period n. All feasible action belongs to a set defined as : $FA_n = \{x_n \epsilon R | fa_n(x_n , s_n) \geq 0\}$, where s_n is a state variable that determines the feasible action set for period n.

$o_n = o_n(Z_n)$ is the value that observed the random outcome Z_n in each period n.$lh(\phi|o_n)$ is defined as the likelihood function of the unknown parameter ϕ for an given observation o_n.

Let us consider $s_n \; and \; \pi_n$ be the two state variables where s_n determines the feasible action set at the beginning of period n and π_n is the prior distribution of ϕ.

The transition between s_{n-1} and s_n be defined by $s_n = s_n(s_{n-1} ,\; x_{n-1} , o_{n-1})$, which is a function of the previous state s_{n-1}, the action x_{n-1}, and the observation o_{n-1}. The Bayes' rule for the given observation o_{n-1} , the transition from π_{n-1} to π_n is given as:

$$\pi_n(\phi) = \frac{lh(\phi|o_n) \; . \; \pi_{n-1}(\phi)}{\int_{\Theta} lh(\phi|o_n) \; . \; \pi_{n-1}(\phi) \; d\phi} \qquad (1)$$

Define the cost function $cf(x_n, s_n, z_n)$ in period n given the action y_n, the state s_n and the random outcome z_n. The expected cost of a period, denoted by $CF(x_n, s_n, \pi_n)$ is given as:

$CF(x_n, s_n, \pi_n) = \int_{\Theta} E_c \{cf(x_n, s_n, Z_n \mid \phi)\} . \pi_n(\phi) \, d\phi.$

Defining the following sets:

$PM_n : \{o_n | where\ lh(\phi|o_n)\ is\ a\ discrete\ probability\ mass\}$

$CPM_n : \{o_n | where\ lh(\phi|o_n)\ is\ a\ continuous\ probability\ density\}$

For the Bayesian dynamic program the optimality equations are:

$$V_n(s_n, \pi_n) = \min_{a_n(x_n, s_n) \geq 0} \{H_n(x_n, s_n, \pi_n)\}$$

$$= \min_{a_n(x_n, s_n) \geq 0} \left\{ CF(x_n, s_n, \pi_n) + \int_{o_n \in CPM_n} V_{n+1}\left(s_{n+1}, \frac{lh(.|o_n)\,\pi_n}{\int_{\Theta} lh(\phi|o_n)\,\pi_n(\phi)\,d\phi}\right) . \left(\int_{\Theta} lh(\phi|o_n)\,\pi_n(\phi)\,d\phi\right) d\,o_n + \sum_{o_n \in PM_n} V_{n+1}\left(s_{n+1}, \frac{lh(.|o_n)\,\pi_n}{\int_{\Theta} lh(\phi|o_n)\,\pi_n(\phi)\,d\phi}\right) . \left(\int_{\Theta} lh(\phi|o_n)\,\pi_n(\phi)\,d\phi\right) \right\} \quad (2)$$

for $n = 1, \dots \dots, N$, with $V_{N+1}(\cdot\,,\ \cdot) = 0$. The validity of the above optimality equations is given in (Bertsekas and Shreve 1978, Chapter 10). Let $x_n^*(s_n, \pi_n)$ denote the opimal decision point in period n to the above problem (2) and $X_n^*(s_n, \pi_n)$ denote the optimal control policy along the sample path given the initial state s_n and prior distribution π_n. For the smooth progress of the consequent study, let us define the un-normalized posterior $\bar{\pi}_n$ from (1) as:

$$\tilde{\pi}_n(\phi) = lh(\phi|o_{n-1}) . \tilde{\pi}_{n-1}(\phi) \quad (3)$$

For $n = 2, \dots \dots .., N$ with $\tilde{\pi}_1(\phi) = \pi_1(\phi)$. It is easy to show that $\tilde{\pi}_n = \pi_n . \int_{\Theta} \tilde{\pi}_n(\phi)\, d\,\phi$.

On considering the un-normalized posterior let us introduce the modified single-period cost function

$$CF(x_n, s_n, \tilde{\pi}_n) = \int_{\Theta} E_c \{cf(x_n, s_n, Z_n \mid \phi)\} . \tilde{\pi}_n(\phi)\, d\phi \quad (4)$$

And also the modified optimality equations as:

$$\tilde{V}_n(s_n, \tilde{\pi}_n) = \min_{a_n(x_n, s_n) \geq 0} \{\tilde{H}_n(x_n, s_n, \tilde{\pi}_n)\}$$

$$= \min_{a_n(x_n, s_n) \geq 0} \left\{ CF(x_n, s_n, \tilde{\pi}_n) + \int_{o_n \in CPM_n} \tilde{V}_{n+1}(s_{n+1}, lh(.|o_n)\tilde{\pi}_n)\, do_n + \sum_{a_n \in PM_n} \tilde{V}_{n+1}(s_{n+1}, lh(.|o_n)\tilde{\pi}_n) \right\} \quad (5)$$

For $n = 1, \dots \dots \dots ., N$, $also\ \tilde{V}_{N+1}(\cdot\,,\ \cdot) = 0$.

Let us assume that $\tilde{x}_n^*(s_n, \tilde{\pi}_n)$ be the optimal decision point for a period n at the given state s_n and the unnormalized prior $\tilde{\pi}_n$. It is to be noticed that the integration/summation till now is taken without the probability measure directly with respect to o_n .

Consider the following lemma establishes the equivalence between relation (5) and the original problem (2):

Lemma 1: For $\boldsymbol{n = 1, \ldots\ldots, N}$, $\widetilde{H}_{\boldsymbol{n}}(\boldsymbol{x_n}, \boldsymbol{s_n}, \tilde{\pi}_{\boldsymbol{n}}) = \boldsymbol{H_n}(\boldsymbol{x_n}, \boldsymbol{s_n}, \boldsymbol{\pi_n}) \cdot \int_{\ominus} \tilde{\pi}_{\boldsymbol{n}}(\boldsymbol{\phi})\, \boldsymbol{d\phi}$,

$\widetilde{\boldsymbol{V}}_{\boldsymbol{n}}\left(\boldsymbol{s_n}, \tilde{\pi}_{\boldsymbol{n}}\right) = \boldsymbol{V_n}\left(\boldsymbol{s_n}, \boldsymbol{\pi_n}\right) \cdot \int_{\ominus} \tilde{\pi}_{\boldsymbol{n}}(\boldsymbol{\phi})\, \boldsymbol{d\phi}$ **and** $\widetilde{\boldsymbol{x}}_{\boldsymbol{n}}^{*}(\boldsymbol{s_n}, \tilde{\pi}_{\boldsymbol{n}}) = \boldsymbol{x}_{\boldsymbol{n}}^{*}(\boldsymbol{s_n}, \boldsymbol{\pi_n})$**.**

Proof: With the help of backward substitution we prove that

$\widetilde{H}_n(x_n, s_n, \tilde{\pi}_n) = H_n(x_n, s_n, \pi_n) \cdot \int_{\ominus} \tilde{\pi}_n(\phi)\, d\phi$. Using relation (4), we can show that

$CF(\boldsymbol{x_N}, \boldsymbol{s_N}, \tilde{\pi}_{\boldsymbol{N}}) = CF(\boldsymbol{x_N}, \boldsymbol{s_N}, \boldsymbol{\pi_N}) \cdot \int_{\ominus} \tilde{\pi}_n(\phi)\, d\,\phi$. For

$n = N, \widetilde{H}_N(x_N, s_N, \tilde{\pi}_N) = CF(x_N, s_N, \tilde{\pi}_N)$.

Let us assume that the result holds for period n + 1. For period n, we have

$$\widetilde{H}_n(x_n, s_n, \tilde{\pi}_n) = CF(x_n, s_n, \tilde{\pi}_n) +$$

$$= CF(x_n, s_n, \tilde{\pi}_n) + \int_{o_n \in CPM_n} \widetilde{V}_{n+1}\left(s_{n+1}, \frac{lh\,(.\,|o_n)\,\tilde{\pi}_n}{\int_{\ominus} lh(\phi|o_n)\,\tilde{\pi}_n(\phi)\,d\phi}\right).\left(\int_{\ominus} lh(\phi|o_n)\,\tilde{\pi}_n(\phi)\,d\phi\right) d\,o_n$$

$$+ \sum_{o_n \in PM_n} V_{n+1}\left(s_{n+1}, \frac{lh\,(.\,|o_n)\,\tilde{\pi}_{n \int_{o_n \in CPM_n} \overline{V}_{n+1}(s_{n+1}, lh(.|o_n)\overline{\pi}_n)\,do_n + \sum_{a_n \in PM_n} \overline{V}_{n+1}(s_{n+1}, lh(.|o_n)\overline{\pi}_n)}}{\int_{\ominus} lh(\phi|o_n)\,\tilde{\pi}_n(\phi)\,d\phi}\right).\left(\int_{\ominus} lh(\phi|o_n)\,\tilde{\pi}_n(\phi)\,d\phi\right)$$

$$= \left\{CF(x_n, s_n, \pi_n) + \int_{o_n \in CPM_n} V_{n+1}\left(s_{n+1}, \frac{lh\,(.\,|o_n)\,\pi_n}{\int_{\ominus} lh(\phi|o_n)\,\pi_n(\phi)\,d\phi}\right).\left(\int_{\ominus} lh(\phi|o_n)\,\pi_n(\phi)\,d\phi\right) d\,o_n\right.$$

$$\left. + \sum_{o_n \in PM_n} V_{n+1}\left(s_{n+1}, \frac{lh\,(.\,|o_n)\,\pi_n}{\int_{\ominus} lh(\phi|o_n)\,\pi_n(\phi)\,d\phi}\right).\left(\int_{\ominus} lh(\phi|o_n)\,\pi_n(\phi)\,d\phi\right)\right\} \int_{\ominus} \tilde{\pi}_n(\phi)\,d\,\phi$$

$$= H_n(x_n, s_n, \pi_n) \cdot \int_{\ominus} \tilde{\pi}_n(\phi)\,d\,\phi$$

The last equality follows from relation (2), the third equality follows from the fact

$\tilde{\pi}_n = \pi_n \cdot \int_{\ominus} \tilde{\pi}_n(\phi)\, d\,\phi$ and the second equality follows from the induction assumption.

This lemma shows that the problem (5) is an extended version of the original problem (2) and also both problems share the same optimal solution. Therefore, we only concerned about the problem (5).

The unnormalized prior $\tilde{\pi}_n'$ on which a sample path evolving based applying on the control policy $\tilde{X}_n^*(s_n, \tilde{\pi}_n)$, let the expected total cost be $\tilde{G}_n\left(\tilde{X}_n^*(s_n, \tilde{\pi}_n), \tilde{\pi}_n'\right)$. This expected total cost can be recursively defined as:

For $1 \leq n \leq N - 1$,

$$\tilde{G}_n\left(\tilde{X}_n^*(s_n, \tilde{\pi}_n), \tilde{\pi}_n'\right) = CF(\tilde{x}_n^*(s_n, \tilde{\pi}_n), s_n, \tilde{\pi}_n')$$

$$+ \int_{o_n \in CPM_n} \tilde{G}_{n+1}\left(\tilde{X}_{n+1}^*(s_{n+1}, lh(\cdot\,|o_n)\,\tilde{\pi}_n), lh(\cdot\,|o_n)\,\tilde{\pi}_n'\right) d\,o_n$$

$$+\sum_{o_n \epsilon PM_n} \tilde{G}_{n+1}\left(\tilde{X}^*_{n+1}(s_{n+1}, lh(\cdot|o_n)\,\tilde{\pi}_n), lh(\cdot|o_n)\,\tilde{\pi}'_n\right). \quad (6)$$

With $G_{N+1}(\cdot, \cdot) = 0$. Now when $\tilde{\pi}'_n = \tilde{\pi}_n$, then from (5), we get

$\tilde{G}_n\left(\tilde{X}^*_n(s_n, \tilde{\pi}_n), \tilde{\pi}'_n\right) = \tilde{H}_n(\tilde{x}^*_n(s_n, \tilde{\pi}_n), s_n, \tilde{\pi}'_n) = \tilde{V}_n(s_n, \tilde{\pi}_n)$.

Theorem 1: Let $\tilde{\pi}_n = \tilde{\pi}_n(\phi, p)$, where p is a parameter of $\tilde{\pi}_n$. Assume that $\tilde{\pi}_n(\phi, p)$ is differentiable in p. Then $\frac{d\tilde{V}_n(s_n, \tilde{\pi}_n)}{dp} = \tilde{G}_n\left(\tilde{X}^*_n(s_n, \tilde{\pi}_n), \frac{\partial \tilde{\pi}_n}{\partial p}\right)$, where $\tilde{G}_n(\cdot, \cdot)$ is defined in (6).

Proof: Since $\tilde{\pi}_n(\phi, p)$ is differentiable in p, it is clear that $\tilde{V}_n(s_n, \tilde{\pi}_n)$ is also differentiable in p. Through backward induction we can prove the result. Then for n = N we get,

$$\tilde{V}_N(s_N, \tilde{\pi}_N) = \min_{a_N(x_N, s_N) \geq 0} \{CF(x_N, s_N, \tilde{\pi}_N)\}$$

According to the standard envelope theorem, we have

$$\frac{d\tilde{V}_N(s_N, \tilde{\pi}_N)}{dp} = \frac{\partial CF(\tilde{x}^*_N, s_N, \tilde{\pi}_N)}{\partial p} = CF\left(\tilde{x}^*_N, s_N, \frac{\partial \tilde{\pi}_N}{\partial p}\right) = \tilde{G}_N\left(\tilde{\pi}^*_N(s_N, \tilde{\pi}_N), \frac{\partial \tilde{\pi}_N}{\partial p}\right)$$

Let us assume that the result holds for period n + 1. Then for period n, we get

$$\frac{d\tilde{V}_n(s_n, \tilde{\pi}_n)}{dp} = \frac{d\tilde{H}_n(\tilde{x}^*_n, s_n, \tilde{\pi}_n)}{dp} = \frac{\partial \tilde{H}_n(\tilde{x}^*_n, s_n, \tilde{\pi}_n)}{\partial x_n} \cdot \frac{\partial \bar{x}^*_n}{\partial p} + \frac{\partial \tilde{H}_n(\tilde{x}^*_n, s_n, \tilde{\pi}_n)}{\partial p}$$

Also we have $\frac{\partial \tilde{H}_n(\tilde{x}^*_n, s_n, \tilde{\pi}_n)}{\partial x_n} = \lambda^* \cdot \frac{\partial a_n(\tilde{x}^*_n, s_n)}{\partial x_n}$ from the first-order optimality condition, where λ^* is the associated Lagrangian multiplier. Since $\lambda^* \cdot a_n(\tilde{x}^*_n, s_n) \equiv 0$. Taking derivative with respect to p of this equation, we have $\frac{\partial \lambda^*}{\partial p} \cdot a_n(\tilde{x}^*_n, s_n) + \lambda^* \cdot \frac{\partial a_n(\tilde{x}^*_n, s_n)}{\partial x_n} \cdot \frac{\partial \tilde{x}^*_n}{\partial p} = 0$.

For $\lambda^* > 0, a_n(\tilde{x}^*_n, s_n) = 0$. Hence, $\frac{\partial \tilde{H}_n(\tilde{x}^*_n, s_n, \tilde{\pi}_n)}{\partial x_n} \cdot \frac{\partial \tilde{x}^*_n}{\partial p} = 0$ and we get,

$$\frac{d\tilde{V}_n(s_n, \tilde{\pi}_n)}{dp} = \frac{\partial \tilde{H}_n(\tilde{x}^*_n, s_n, \tilde{\pi}_n)}{\partial p}$$

$$= \frac{\partial}{\partial p}\left\{CF(\tilde{x}^*_n, s_n, \tilde{\pi}_n) + \int_{o_n \epsilon CPM_n} \tilde{V}_{n+1}(s_{n+1}, lh(.|o_n)\tilde{\pi}_n)\, do_n + \sum_{a_n \epsilon PM_n} \tilde{V}_{n+1}(s_{n+1}, lh(.|o_n)\tilde{\pi}_n)\right\}$$

$$= CF\left(\tilde{x}^*_n, s_n, \frac{\partial \tilde{\pi}_n}{\partial p}\right) + \int_{o_n \epsilon CPM_n} \frac{d\tilde{V}_{n+1}(s_{n+1}, lh(.|o_n)\tilde{\pi}_n)}{dp}\, do_n + \sum_{a_n \epsilon PM_n} \frac{d\tilde{V}_{n+1}(s_{n+1}, lh(.|o_n)\tilde{\pi}_n)}{dp}$$

$$= CF\left(\tilde{x}^*_n, s_n, \frac{\partial \tilde{\pi}_n}{\partial p}\right) + \int_{o_n \epsilon CPM_n} \tilde{G}_{n+1}\left(\tilde{X}^*_{n+1}(s_{n+1}, lh(\cdot|o_n)\,\tilde{\pi}_n), lh(\cdot|o_n)\frac{\partial \tilde{\pi}_n}{\partial p}\right) do_n$$

$$+\sum_{o_n \epsilon PM_n} \tilde{G}_{n+1}\left(\tilde{X}^*_{n+1}(s_{n+1}, lh(\cdot|o_n)\,\tilde{\pi}_n), lh(\cdot|o_n)\frac{\partial \tilde{\pi}_n}{\partial p}\right)$$

$$= \tilde{G}_n\left(\tilde{X}^*_n(s_n, \tilde{\pi}_n), \frac{\partial \tilde{\pi}_n}{\partial p}\right).$$

Here the second-to-last equality follows from the induction assumption and the last equality follows from the defnition (6).

This completes the induction proof.

The classic envelope theorem states that the derivative of a single-period optimal value function

with respect to a parameter is equal to the derivative with the objective function fixed at the original optimal decision point (Mas-Colell et al. 1995, pp. 964{966).

This theorem is a generalization of the classic envelope theorem in the microeconomics literature.

In theorem 1, under the Bayesian dynamic program problem, the derivative of the value function $\tilde{V}_n(s_n, \tilde{\pi}_n)$ with respect to a parameter p (of the unnormalized prior $\tilde{\pi}_n$) is equal to the expected cost of applying the original optimal control policy $\tilde{X}_n^*(s_n, \tilde{\pi}_n)$ on a sample path evolving according to the derivative of the unnormalized prior $\frac{\partial \tilde{\pi}_n}{\partial p}$. The difference here is that we need to fix not only the optimal decision for the current period but also the optimal contingent decisions along the sample path. Now, we apply Theorem 1 to a Bayesian inventory management problem with unobserved lost sales.

3. BAYESIAN INVENTORY MANAGEMENT WITH UNOBSERVED LOST SALES

Assume a finite-horizon inventory control problem for a non-perishable product with unobserved lost sales and a demand distribution having an unknown parameter. Under a general discrete demand distribution, Chen and Plambeck (2008) have shown that the optimal inventory level with unobserve lost sales is greater than the optimal inventory level with observed lost sales.

We show that this result also holds under a general continuous demand distribution.

The product is stocked and sold for N periods. The demands in each period, denoted by Z_t . The function, denoted by $f(z|\phi)$, has an unknown parameter ϕ,with $\phi \epsilon \ominus$. Also let $\mathrm{F}(z|\phi)$ denote the cumulative distribution function (CDF) and $\bar{F}(z|\phi) = 1 - \mathrm{F}(z|\phi)$ denote the complement CDF. The unknown parameter ϕ is subject to a prior distribution $\pi_n(\phi)$. At the beginning of each period n, an inventory level x_n is chosen to minimize the total inventory holding and stock out penalty costs. Let ζ denote the unit holding cost and p_c the unit stock out penalty cost. The production lead time is assumed to be negligible, so the inventory level is achieved immediately after the decision. The purchase cost of the product is omitted in our formulation as it can be normalized to zero with the standard technique of Heyman and Sobel (1984). The terminal value at the end of the planning horizon is assumed to be zero.

The cost function for the given inventory level x_n and demand z_n is defined as:

$cf(x_n - y_n) = \zeta(x_n - z_n)^+ + q(x_n - z_n)^- = \zeta \cdot max(x_n - z_n\,, 0) + q \cdot max(z_n - x_n\,, 0).$

For a period, expected cost can be written as $CF(x_n\,, \pi_n) = \int_{\Theta} E\{cf(x_n - Z_n)|\phi\} \cdot \pi_n(\phi)d\phi.$

Also $o_n = Z_n$, i.e., when lost sales are observed, the demand is always observed exactly and the likelihood function is $lh(\phi|o_n) = f(o_n|\phi)$.Is to be noted that superscript "l" on the value functions and variables to denote this observed lost sales case. For n = 1,….., N, the unnormalized optimality equations in this case is:

$$\tilde{V}_n^l(s_n\,, \tilde{\pi}_n) = \min_{x_n \geq s_n}\{\tilde{H}_n^l(x_n\,, \tilde{\pi}_n)\}$$

$$= \min_{x_n \geq s_n}\left\{CF(x_n\,, \tilde{\pi}_n) + \int_0^{\infty} \tilde{V}_{n+1}^l((x_n - o_n)^+, f(o_n|\,\cdot)\tilde{\pi}_n)\, d\, o_n\right\}$$

with $\tilde{V}_{n+1}^l(\cdot\,,\ \cdot) = 0$. Let us consider x_n^l be the optimal solution to the (normalized) original problem and $\tilde{x}_n^l$ be the optimal solution to this problem. We know that $x_n^l = \tilde{x}_n^l$ (by lemma 1). The first order derivative of $\tilde{H}_n^l(x_n\,, \pi_n)$ with respect to x_n is:

$$\frac{d\tilde{H}_n^l(x_n\,, \tilde{\pi}_n)}{dx_n} = \frac{dCF(x_n\,, \tilde{\pi}_n)}{dx_n} + \int_0^{x_n} \frac{d}{dx_n}\tilde{V}_{n+1}^l(x_n - o_n\,, f(o_n|\,\cdot)\tilde{\pi}_n)do_n$$

Also $d\tilde{V}_{n+1}^l(0\,,\ f(o_n|\,\cdot)\tilde{\pi}_n)/dx_n = 0$. When the lost sales data observed in period n is given by $l_n = min(Z_n\,,\ x_n)$. And hence the likelihood function is given by

$$lh(\phi|o_n) = \begin{cases} f(o_n|\phi) & if \quad o_n < x_n \\ \bar{F}(x_n|\phi) & if \quad o_n = x_n \end{cases}$$

According to relation (5), for n = 1,…….,N, the unnormalized optimality equations in this case are given by:

$$\tilde{V}_n(s_n\,,\ \tilde{\pi}_n\,) = \min_{x_n \geq s_n}\{\tilde{H}_n(x_n\,,\ \tilde{\pi}_n\,)\}$$

$$= \min_{x_n \geq s_n}\left\{CF(x_n\,,\ \tilde{\pi}_n\,) + \int_0^{x_n} \tilde{V}_{n+1}^l(x_n - o_n\,, f(o_n|\,\cdot)\tilde{\pi}_n)do_n + \tilde{V}_{n+1}(0, \bar{F}(x_n|\,\cdot)\tilde{\pi}_n)\right\}$$

With $\tilde{V}_{N+1}(\cdot\,,\ \cdot) = 0$. Let $\bar{x}_n^*$ be the optimal solution to this problem and x_n^* be the optimal solution to the normalized original problem. By lemma 1, we come to know that $\bar{x}_n^* = x_n^*$. The first-order derivative of $\tilde{H}_n(x_n\,,\ \pi_n\,)$ with respect to x_n is:

$$\frac{d\widetilde{H}_n(x_n, \tilde{\pi}_n)}{dx_n} = \frac{dCF(x_n, \tilde{\pi}_n)}{dx_n} + \int_0^{x_n} \frac{d}{dx_n} \tilde{V}_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n) do_n$$

$$+\tilde{V}_{n+1}(0, f(x_n|\cdot)\tilde{\pi}_n) - \tilde{G}_{n+1}\left(\tilde{X}^*_{n+1}(0, \bar{F}(x_n|\cdot)\tilde{\pi}_n), f(x_n|\cdot)\tilde{\pi}_n\right)$$

here the last term follows from Theorem 1 and the fact that $\partial\tilde{F}(x_n|\phi)/\partial x_n = -f(x_n|\phi)$. We can establish the another the another result by comparing $d\widetilde{H}_n(x_n, \tilde{\pi}_n)/dx_n$ with $d\widetilde{H}^l_n(x_n, \tilde{\pi}_n)/dx_n$ as:

Proposition 1. ***The optimal inventory level with unobserved lost sales is greater than the optimal inventory level with observed lost sales (in a non-perishable inventory system with continuous demand, given the same starting on-hand inventory and prior distribution) i.e.,*** $\boldsymbol{x^*_n \geq x^l_n}$.

Proof: By Lemma 1, we have $\tilde{x}^l_n = x^l_n$ and $\tilde{x}^*_n = x^*_n$ .Therefore, it is sufficient to show that $\widetilde{H}_n(x_n, \tilde{\pi}_n)/dx_n \leq d\widetilde{H}^l_n(x_n, \tilde{\pi}_n)/dx_n$. It is to be noted that

$$\tilde{V}_{n+1}(0, f(x_n|\cdot)\tilde{\pi}_n) = \tilde{G}_{n+1}\left(\tilde{X}^*_{n+1}(0, f(x_n|\cdot)\tilde{\pi}_n), f(x_n|\cdot)\,\tilde{\pi}_n\right)$$

$$\leq \tilde{G}_{n+1}\left(\tilde{X}^*_{n+1}(0, \bar{F}(x_n|\cdot)\tilde{\pi}_n), f(x_n|\cdot)\,\tilde{\pi}_n\right)$$

And we have,

$$\frac{d\widetilde{H}_n(x_n, \tilde{\pi}_n)}{dx_n} \leq \frac{dCF(x_n, \tilde{\pi}_n)}{dx_n} + \int_0^{x_n} \frac{d}{dx_n} \tilde{V}_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n) do_n$$

It is to be noted that:

$$\frac{d\widetilde{H}^l_n(x_n, \tilde{\pi}_n)}{dx_n} = \frac{dCF(x_n, \tilde{\pi}_n)}{dx_n} + \int_0^{x_n} \frac{d}{dx_n} \tilde{V}^l_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n) do_n$$

Therefore it is sufficient to show that

$$\frac{d}{dx_n}\tilde{V}_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n) \leq \frac{d}{dx_n}\tilde{V}^l_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n) \qquad (7)$$

Since we know that

$$\frac{d}{dx_n}\tilde{V}_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n) = max\left(\frac{d}{dx_n}\widetilde{H}_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n), 0\right)$$

$$\frac{d}{dx_n}\tilde{V}^l_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n) = max\left(\frac{d}{dx_n}\widetilde{H}^l_{n+1}(x_n - o_n, f(o_n|\cdot)\tilde{\pi}_n), 0\right)$$

Therefore, by the backward induction assumption relation (7) holds.

As to learn more about the demand information Proposition 1 shows that when lost sales are not observed, one should stock more than what would be optimal if lost sales are observed.

Since the optimal inventory level with observed lost sales is relatively easy to compute, this result can be used to design near-optimal heuristics for the problem with unobserved lost sales.

4. EXTENSION AND CONCLUDING REMARKS

Under the IID assumption we derived Theorem 1 and Proposition 1. But in many real situations, the random events (demands) are often correlated cross different periods and/or following certain seasonality pattern. These situations can be modeled as a Markov-modulated process. We show that our results can be easily extended to this more general case. Let us assume that the unknown parameter ϕ is transitioned between periods according to a *known* stochastic transition function $\tau(\phi|\phi')$. For the given value of ϕ', $\tau(\cdot\,|\phi')$ is a probability distribution). Hence, before observing o_{n-1} , $\pi_{n-1}(\phi)$ is first transitioned to a new distribution by

$$(\tau \; o \; \pi_{n-1})(\phi) = \int_{\Theta} \tau(\phi|\phi')\pi_{n-1}(\phi')\, d\phi'$$

Across different periods, this transition captures the dependency of the random events (demands). Particularly, the dependency is modeled by a Markov-modulated process with ϕ being the modulating state. The posterior $\pi_n$is given by the Bayes' rule, after observing o_{n-1} as:

$$\pi_n(\phi) = \frac{lh(\phi|o_{n-1}) \cdot (\tau \; o \; \pi_{n-1})(\phi)}{\int_{\Theta} lh(\phi|o_{n-1}) \cdot (\tau \; o \; \pi_{n-1})(\phi)\, d\phi}$$

In the IID case, $\tau(\phi|\phi')$ is the Dirac delta function $\delta(\phi - \phi')$; hence, $(\tau \; o \; \pi)(\phi) = \pi(\phi)$ and the above formula reduces to (1). Akin to (3), define the unnormalized prior as:
$\tilde{\pi}_n(\phi) = lh(\phi|o_{n-1}) \cdot (\tau \; o \; \tilde{\pi}_{n-1})(\phi)$, for n = 2,…..,N, with $\tilde{\pi}_1(\phi) = \pi_1(\phi)$. Without loss of generality we can show that $\tilde{\pi}_n = \pi_n \cdot \int_{\Theta} \tilde{\pi}_n(\phi)\, d\,\phi$. Assuming that $\tilde{\pi}_n = \tilde{\pi}_n(\phi\,, p)$ is differentiable in p. We get,

$$\frac{\partial}{\partial p}\left(lh(\phi|o_n) \cdot (\tau \; o \; \tilde{\pi}_n)\right) = \frac{\partial}{\partial p}\left(lh(\phi|o_n) \cdot \int_{\Theta} \tau(\phi|\phi')\tilde{\pi}_n(\phi', p)\, d\phi'\right)$$

$$= lh(\phi|o_n) \cdot \left(\tau \; o \; \frac{\partial \tilde{\pi}_n}{\partial p}\right)(\phi)$$

We can extend Theorem 1 and Proposition 1 to the Markov-modulated process, given this observation following the same proof logic. Besides the application in Bayesian inventory management problems, Theorem 1 is also a useful tool to gain insight on other problems concerning joint parameter estimation and optimal control, such as dynamic pricing and learning about consumer behavior.

5. REFERENCES

[1]. Azoury, K.S. 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Sci.* **31** 1150-1160.

[2]. Bensoussan, A., M. Cakanyildirim, S.P. Sethi. 2007a. A multi-period newsvendor problem with partially observed demand. *Math. of Oper. Res.* **32(2)** 322-344.

[3]. Bensoussan, A., M. Cakanyildirim, S.P. Sethi. 2007b. A note on “The censored newsvendor and the optimal acquisition of information". Forthcoming *Oper. Res.*

[4]. Bertsekas, D., S.E. Shreve. 1978. *Stochastic Optimal Control: The Discrete Time Case.* Academic Press, New York.

[5]. Bisi, A., M. Dada. 2007a. Dynamic learning, pricing, and ordering by a censored newsvendor. *Naval Res. Logist. Quart.* **54** 448-461.

[6]. Bisi, A., M. Dada. 2007b. Optimal polices for the censored newsvendor problem with newsvendor demand distributions. Working Paper, Purdue University.

[7]. Chen, F., J.-S. Song. 2001. Optimal policies for multi-echelon inventory problems with Markov-modulated demand. *Oper. Res.* **49(2)** 226-234.

[8]. Chen, L. 2008. Bounds and heuristics for optimal Bayesian inventory control with unobserved lost sales. Working Paper. Duke University.

[9]. Chen, L., E.L. Plambeck. 2008. Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing & Service Oper. Management.* **10(2)** 236-256.

[10]. Ding, X., M.L. Puterman, A. Bisi. 2002. The censored newsvendor and the optimal acquisition of information. *Oper. Res.* **50(3)** 517-527.

[11]. Eppen,G.D., A.V. Iyer. 1997. Improved fashion buying with Bayesian updates. *Oper. Res.* **45(6)** 805{819.

[12]. Harpaz, G., W.Y. Lee, R.L. Winkler. 1982. Learning, experimentation, and the optimal output decisions of a competitive firm. *Management Sci.* **28(6)** 589-603.

[13]. Heyman, D.P., M.J. Sobel. 1984. *Stochastic Models in Operations Research, Volume II: Stochastic Optimization*. McGraw-Hill Book Company, New York.

[14]. Huh, W.T., P. Rusmevichientong. 2009. A non-parametric asymptotic analysis of inventory planning with censored demand. *Math. of Oper. Res.* **34(1)** 103-123.

[15]. Huh, W.T., R. Levi, P. Rusmevichientong, J. Orlin. 2009. Adaptive data-driven inventory control policies based on Kaplan-Meier estimator. Working Paper. Columbia University. Lariviere, M.A., E.L. Porteus. 1999. Stalking information: Bayesian inventory management with unobserved lost sales. *Management Sci.* **45(3)** 346-363.

[16]. Lovejoy, W.S. 1990. Myopic policies for some inventory models with uncertain demand distributions. *Management Sci.* **36(6)** 724-738.

[17]. Lovejoy, W.S. 1993. Suboptimal policies, with bounds, for parameter adaptive decision processes. *Oper. Res.* **41(3)** 583-599.

[18]. Lu, X., J.-S. Song and K. Zhu. 2005. On “The censored newsvendor and the optimal Acquisition of information". *Oper. Res.* **53(6)** 1024-1026.

[19]. Lu, X., J.S. Song, K. Zhu. 2006. Inventory control with unobservable lost sales and Bayesian updates. Working Paper, Hong Kong University of Science and Technology.

[20]. Lu, X., J.S. Song, K. Zhu. 2008. Analysis of perishable inventory systems with censored data. *Oper. Res.* **56(4)** 1034-1038.

[21]. Mas-Colell, A., M. Whinston, J. Green. 1995. *Microeconomic Theory.* Oxford University Press. Scarf, H.E. 1959. Bayes solution of the statistical inventory problem. *Ann. Math. Statist.* **30** 490-508.

[22]. Scarf, H.E. 1960. Some remarks on Bayes solutions to the inventory problem. *Naval Res. Logist.Quart.* **7** 591-596.

[23]. Treharne, J.T., C.R. Sox. 2002. Adaptive inventory control for nonstationary demand and partial information. *Management Sci.* **48(5)** 607-624.