

SOLVING THE NEWSVENDOR PROBLEM UNDER PARAMETRIC UNCERTAINTY USING SIMULATION

David Fernando Muñoz

Department of Industrial and
 Operations Engineering
 Instituto Tecnológico Autónomo de México
 Rio Hondo #1, Col. Tizapan San Angel
 Mexico City, 01080, MEXICO

David Gonzalo Muñoz

Research and Development
 AOL Advertising
 395 Page Mill Road, Third Floor
 Palo Alto, CA 94404, USA

ABSTRACT

In this paper, we discuss the formulation and solution to the newsvendor problem under a Bayesian framework that allows the incorporation of uncertainty on the parameters of the demand model (introduced by the estimation process of these parameters). We present an application of this model with an analytical solution and we conduct experiments to compare the results under the proposed method and a classical approach. Furthermore, we illustrate the estimation of the optimal order size using stochastic simulation, when the complexity of the model does not allow the finding of a closed form expression for the solution.

1 INTRODUCTION AND LITERATURE REVIEW

Let D represent the demand (during the sales period) of a seasonal item. If $w \geq 0$ denotes the loss for every unsold unit at the end of the period, and $u \geq 0$ denotes the profit for every unit sold during the period, the total profit for an order size of Q units is given by:

$$b(Q) = \begin{cases} uD - w(Q - D), & D < Q, \\ uQ, & D \geq Q. \end{cases} \quad (1)$$

The most common approach (see, for example, Nahmias and Olsen 2015) to find the optimal order size Q_C^* consists in defining a density function $f(y|\theta)$ for the demand D (the analysis is similar for the discrete case), where θ is the parameter vector, and, assuming θ is known, we define the expected profit as

$$B_C(Q|\theta) \stackrel{def}{=} E[b(Q)|\Theta = \theta] = u \int_0^Q y f(y|\theta) dy - w \int_0^Q (Q - y) f(y|\theta) dy + uQ \int_Q^\infty f(y|\theta) dy. \quad (2)$$

$$F_C(Q_C^*|\theta) \stackrel{def}{=} \int_0^{Q_C^*} f(y|\theta) dy = \frac{u}{u + w}. \quad (3)$$

Note that $F_C(y|\theta)$ is the cumulative distribution function (cdf) of the demand given $[\Theta = \theta]$. Furthermore, if $f(x|\theta) > 0$ is continuous in a neighborhood of Q_C^* , condition (3) is sufficient for finding the optimal value Q_C^* . We also mention, for the sake of clarity, that the formulation presented in Nahmias and Olsen (2015) is an equivalent formulation where the authors minimize the expected value of $-b(Q) - uD$, so that (3) follows from Nahmias and Olsen (2015).

In practice, the value of θ is estimated from a data set $x = (x_1, \dots, x_n)$ using, for example, the (maximum likelihood) estimator that maximizes the likelihood function $L(\theta|x)$. The most common approach for finding the optimal order size consists in setting $\theta = \hat{\theta}$ in (3), where $\hat{\theta} = \hat{\theta}(x)$ is a point estimator. While this procedure is found extensively in Operations Management textbooks, it has the downside of assuming that the point estimator equals the parameter. Thus, in this article we discuss a Bayesian approach to the newsvendor problem (i.e., finding the optimal order size) incorporating uncertainty (introduced by the estimation process) in the parameter vector.

Bayesian methods to incorporate parameter uncertainty for inventory management have been proposed since the pioneer work of Scarf (1959), where the author discusses the optimality of a Bayesian updating rule for inventory management. Silver (1965) proposed the incorporation of parameter uncertainty using Bayesian methods and shows how to compute a reorder point by modeling the demand as a multinomial distribution. Bayesian methods have been extensively applied to inventory management in order to propose updating rules for the optimal inventory policy based on new information on the product's demand, see e.g., (Azouri 1985; Eppen and Iyer 1997; Lariviere and Porteus 1999; Chen and Plambeck 2008; Chen 2010; Jain et al. 2015), and the references therein. However, the use of simulation techniques to estimate performance measures for inventory management is not considered in these articles and the related literature.

The following section describes the theory behind the proposed Bayesian approach, which allows the incorporation of parametric uncertainty in the newsvendor problem. Afterwards, in the subsequent section, we illustrate how to estimate the optimal order size using simulation by means of a simple example. This example showcases how to estimate the optimal order size in more complicated problems. In the same section, we present a comparison of the results obtained from applying a classical approach versus the results obtained using the Bayesian approach. Finally, in the last section, we present our conclusions and recommendations.

2 METHODOLOGY

Under a Bayesian framework, the parameter vector is a random variable Θ that has a prior density function $p(\theta)$, thus the posterior density function (given a data set x) is given by

$$p(\theta|x) = \frac{p(\theta)L(x|\theta)}{\int_{S_0} p(\theta)L(x|\theta)d\theta}, \tag{4}$$

where $x \in \mathfrak{R}^d$, $\theta \in S_0$ and $L(x|\theta)$ is the likelihood function. From (4) and following the same notation as in (3), the cdf of the demand (given $[X = x]$) is given by

$$F_B(y|x) \stackrel{def}{=} E[E[F_C(y|\Theta)]|X = x] = \int_{S_0} F_C(y|\theta)p(\theta|x)d\theta, \tag{5}$$

for $y \geq 0$, where $F_C(y|\theta)$ and $p(\theta|x)$ are defined in (3) and (4), respectively. Similarly, from (1) we obtain the expected profit (given $[X = x]$) as

$$B_B(Q|x) \stackrel{def}{=} u \int_0^Q y dF_B(y|x) - w \int_0^Q (Q - y) dF_B(y|x) + uQ \int_Q^\infty dF_B(y|x), \tag{6}$$

where $F_B(y|x)$ is the cdf defined in (5). This shows that $B_B(Q|x)$ has a similar form to $B_C(Q|x)$ defined in (2). Consequently, the optimal order size Q_B^* considering parametric uncertainty satisfies

$$F_B(Q_B^*|x) = \frac{u}{u + w}, \tag{7}$$

where $F_B(y|x)$ is defined in (5). It is worth mentioning that our problem formulation is different from the one proposed in Jain et al. (2015), where the authors propose a dynamic program to solve the newsvendor problem in a multi-period setting in order to show the advantages of using a more efficient updating rule. The main difference is that we consider a single-period expected profit that is explicitly dependent on the available data set x . This formulation allowed us to obtain a simple solution in the form of (7).

It is important to point out that for the case where demand is discrete, taking values $d_1 < d_2 < \dots$, the function $F_B(y|x)$ is not continuous, and equation (7) might not have a solution, in which case we must find the value of d_k that satisfies:

$$P[D \leq d_k | X = x] \leq \frac{u}{u + w} \leq P[D \leq d_{k+1} | X = x], \tag{8}$$

in order to evaluate $B_B(d_k|x)$ and $B_B(d_{k+1}|x)$, where:

$$B_B(Q|x) = u \sum_{j \leq Q} j P[D = j | X = x] - w \sum_{j \leq Q} (Q - j) P[D = j | X = x] + uQ P[D > Q | X = x]. \tag{9}$$

If $B_B(d_k|x) \geq B_B(d_{k+1}|x)$, the optimal order size will be given by $Q_B^* = d_k$, otherwise it will be given by $Q_B^* = d_{k+1}$. Note that in the discrete case (8) is equivalent to (7), in the sense that neither equation considers fixed ordering costs. If there is an initial inventory of Q_0 , we should not order when $Q_0 \geq Q_B^*$, otherwise we should only order $Q_B^* - Q_0$ units if $B_B(Q_B^*|x) - B_B(Q_0|x) > C_0$, where C_0 is the fixed ordering cost.

3 AN ILLUSTRATIVE EXAMPLE

In this section we illustrate the application of the proposed methodology through a similar model to the one presented in Muñoz and Muñoz (2011) for the forecast of an item of intermittent demand. We know that the demand for service parts follows a Poisson process, though there exists uncertainty in the arrival rate Θ_0 , thus, given $[\Theta_0 = \theta_0]$, the times between customers' arrivals are i.i.d. according to the exponential density function:

$$f(y|\theta_0) = \begin{cases} \theta_0 e^{-\theta_0 y}, & y > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $\theta_0 \in S_{00} = (0, \infty)$. Every client can order j units of an item with probability P_j , $j = 1, \dots, q$, $q \geq 2$. Let $\Theta_1 = (P_1, \dots, P_{q-1})$ and $\sum_{j=1}^q P_j = 1$, then $\Theta = (\Theta_0, \Theta_1)$ denotes the vector of parameters and the parameter space is given by $S_0 = S_{00} \otimes S_{01}$, where

$$S_{01} = \left\{ (\rho_1, \dots, \rho_{q-1}) : \sum_{j=1}^{q-1} \rho_j \leq 1; \rho_j \geq 0, j = 1, \dots, q-1 \right\}.$$

The total demand for a period of length T is given by

$$D = \begin{cases} \sum_{i=1}^{N(T)} U_i, & N(T) > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

where $N(s)$ is the number of clients that arrived in the interval $[0, s]$, $s \geq 0$, and U_1, U_2, \dots are the individual item demands (we assume they are conditionally independent with respect to Θ). The information for Θ consists of (i.i.d.) observations $v = (v_1, \dots, v_n)$, $u = (u_1, \dots, u_n)$ of past clients, where v_i is the time between the arrival of client i and previous client $(i-1)$, and u_i is the number of items ordered by client i . The likelihood functions for v and u are given by

$$L(v|\theta_0) = \theta_0^n e^{-\theta_0 \sum_{i=1}^n v_i}, \text{ and } L(u|\theta_1) = \left(1 - \sum_{j=1}^{q-1} \rho_j \right)^{c_q} \prod_{j=1}^{q-1} \rho_j^{c_j}, \tag{11}$$

respectively, where $\theta_1 = (\rho_1, \dots, \rho_{q-1})$, and $c_j = \sum_{i=1}^n I[u_i = j]$ is the number of clients that ordered j items.

From an objective point of view, we can assume a non-informative prior density function for Θ , using Jeffrey's prior density. In the case of the exponential model, Jeffrey's prior density (see, for instance, Bernardo, 2009), is given by $p(\theta_0) = \theta_0^{-1}$, $\theta_0 \in S_{00}$. By taking $\theta = \theta_0$ and $x = v$ in (4), it follows from (11) that

$$p(\theta_0|v) = \frac{\theta_0^{n-1} \binom{n}{\sum_{i=1}^n v_i} e^{-\theta_0 \sum_{i=1}^n v_i}}{(n-1)!}, \tag{12}$$

which corresponds to a $\text{Gamma}(n, \sum_{i=1}^n v_i)$ distribution, where, $\text{Gamma}(\beta_1, \beta_2)$ denotes a gamma distribution with expectation $\beta_1 \beta_2^{-1}$. Similarly, Jeffrey's prior density for the multinomial model (see, for instance, Berger and Bernardo 1992) corresponds to a Dirichlet distribution with density function:

$$p(\theta_1) = \frac{\left(1 - \sum_{j=1}^{q-1} \rho_j\right)^{-1/2} \prod_{j=1}^{q-1} \rho_j^{-1/2}}{B(1/2, \dots, 1/2)},$$

where $B(a_1, \dots, a_q) = \prod_{j=1}^q \Gamma(a_j) / \Gamma(\sum_{j=1}^q a_j)$, for $a_1, \dots, a_q > 0$, thus, it follows from (4) and the multinomial model that

$$p(\theta_1|u) = \frac{\left(1 - \sum_{j=1}^{q-1} \rho_j\right)^{c_q-1/2} \prod_{j=1}^{q-1} \rho_j^{c_j-1/2}}{B(c_1+1/2, \dots, c_q+1/2)}, \tag{13}$$

which corresponds to a Dirichlet($c_1+1/2, \dots, c_q+1/2$) distribution. Let $x_i = (v_i, u_i)$, $i=1, \dots, n$, $x = (x_1, \dots, x_n)$, $\theta = (\theta_0, \theta_1)$, and assuming independence, the posterior density is given by $p(\theta|x) = p(\theta_0|v)p(\theta_1|u)$, where $p(\theta_0|v)$ and $p(\theta_1|u)$ are defined in (12) and (13), respectively.

Note that, in this example, we can obtain a closed form expression for the point estimate of the demand $\mu = E[D|X = x]$, since, from (12) and (13) we have $E[\Theta_0|V = v] = n(\sum_{i=1}^n v_i)^{-1}$ and $E[\Theta_{1j}|U = u] = c^{-1}(c_j + 1/2)$ (where $c = \sum_{j=1}^n (c_j + 1/2) = n + q/2$), and following (10) we have

$$\begin{aligned} \mu &= E[E[D|\Theta, X = x]|X = x] = E[E[N(T)|\Theta]E[U_1|\Theta]|X = x] \\ &= TE\left[\Theta_0 \sum_{j=1}^q j\Theta_{1j}|X = x\right] = TE[\Theta_0|V = v] \sum_{j=1}^q jE[\Theta_{1j}|U = u] \\ &= Tn \left(\sum_{j=1}^n v_i\right)^{-1} (n + q/2)^{-1} \sum_{j=1}^q j(c_j + 1/2), \end{aligned}$$

The above expression allows us to calculate a forecast for the mean μ based on a data set x . In this case, however, it is not easy to obtain a closed form expression for the cdf and the optimal order size. Thus, we can apply the posterior sampling (PS) algorithm described in section 3.1 of Muñoz et al. (2013) in order to calculate, via simulation, the corresponding optimal order size, given a service level $\alpha = u/(u + w)$. It is worth mentioning that the main idea behind the PS algorithm is to generate simulated observations for the demand by first sampling a parameter value θ from the posterior distribution density $p(\theta|x)$, and the sampling a demand observation from the forecasting model for the demand D (conditional on $\Theta = \theta$).

For the case when $q=1$ (i.e., every client orders just one unit), the model is simpler and it is not necessary to turn to simulation in order to find the optimal order size. In this case, we can ignore θ_1 and the values u_i (since they are always equal to 1). Let $x = v$, $\theta = \theta_0$, we have that $P[D = j|\Theta = \theta] = P[N(T) = j|\Theta = \theta] = e^{-\theta T} (\theta T)^j / j!$, and considering (12), it can be proven that

$$P[D = j|X = x] = \binom{n+j-1}{j} \left(\frac{\sum_{i=1}^n x_i}{T + \sum_{i=1}^n x_i} \right)^n \left(\frac{T}{T + \sum_{i=1}^n x_i} \right)^j, \quad (14)$$

for $j=0,1,\dots$, which corresponds to a negative binomial distribution. Using equations (8), (9) and (14) we can determine the optimal order size Q_B^* for this particular case, without resorting to the PS algorithm or simulation.

4 EXPERIMENTAL RESULTS

In order to illustrate the validity of the PS algorithm and, in particular, how it can be applied in order to determine the optimal order size, we will use the example from the previous section that has a closed form expression (14) for the posterior distribution of the demand. First of all, we should point out that we considered the values of $T = 15$, $n = 20$, $\sum_{i=1}^n x_i = 10$, $u = 9$, $w = 1$. With this data, the optimal service level is $\alpha = u/(u + w) = 0.9$. After applying the Bayesian approach described in equations (8) and (9), and the posterior distribution defined in (14), we obtained an optimal order size of $Q_B^* = 41$, then, by following (6), we have an expected profit of $B_B(Q_B^*|x) = 253.38$, and a service level of $F_B(Q_B^*|x) = 0.901$ (slightly higher than 0.9).

With the objective of comparing the results obtained through the classical approach, notice that, from (10), we can find that the cdf $F_C(y|\theta)$ defined in (3) corresponds to a Poisson distribution with mean θT . On the other hand, the maximum likelihood estimator of θ is $\hat{\theta} = 2$, thus, when applying the classical approach with conditions similar to (8) and (9), we obtained $Q_C^* = 37$, reporting an expected profit of $B_C(Q_C^*|x) = 260.05 > B_B(Q_B^*|x)$ from (2), and a service level, from the posterior distribution in (14), of $F_B(Q_C^*|x) = 0.803 < F_B(Q_B^*|x)$. These results suggest that under the classical approach, expected profit is overestimated, and results in a more conservative service level when compared to the Bayesian approach, confirming the intuition that parametric uncertainty proposes a posterior distribution for the demand $F_B(y|x)$ with greater dispersion than the classical approach distribution $F_C(y|x)$. In the following section, we present empirical results that confirm these observations. Subsequently, we will also illustrate how we can estimate the optimal order size when it is not possible (or is extremely complicated) to find a closed form solution.

4.1 Empirical Comparison between the Classical and Bayesian Approaches

In our first experiment, we assumed an arrival rate for clients of $\theta = 2$ and generated $m = 1000$ samples of arrival times, each of size $n = 20$. For every sample, we calculated $\sum_{i=1}^n x_i$ and the optimal order size (under both the classical and Bayesian approaches) with the data from the previous section ($T = 15$, $n = 20$, $u = 9$, $w = 1$). For every sample, we calculated the difference in expected profit between both approaches $B_C(Q_C^*|x) - B_B(Q_B^*|x)$, and the service level for the optimal order size $F_B(Q_C^*|x)$.

Based on Figure 1, notice that the classical approach has overestimated the expected profit in all replications of the experiment. Similarly, based on Figure 2, notice that the classical approach has provided a more conservative service level in every replication of the experiment.

With the objective of showing that the proposed Bayesian approach is consistent with the classical approach, we replicated the previous experiments considering different samples sizes for the time between client arrivals. The results are summarized in Table 1. Notice from the table that, as the sample size increases, the difference in expected profit between both approaches tends to zero and the service level tends to the optimal value 0.9, showing that both approaches coincide as the sample size increases (and as the uncertainty in the parameters becomes negligible).

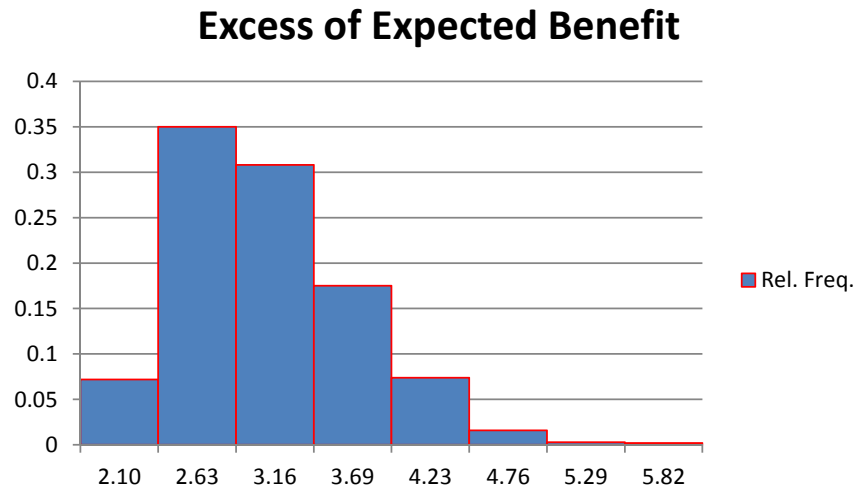


Figure 1: Histogram of the difference $B_C(Q_C^*|x) - B_B(Q_B^*|x)$ based on 1000 replications of the estimation experiments under the classical and Bayesian approaches.

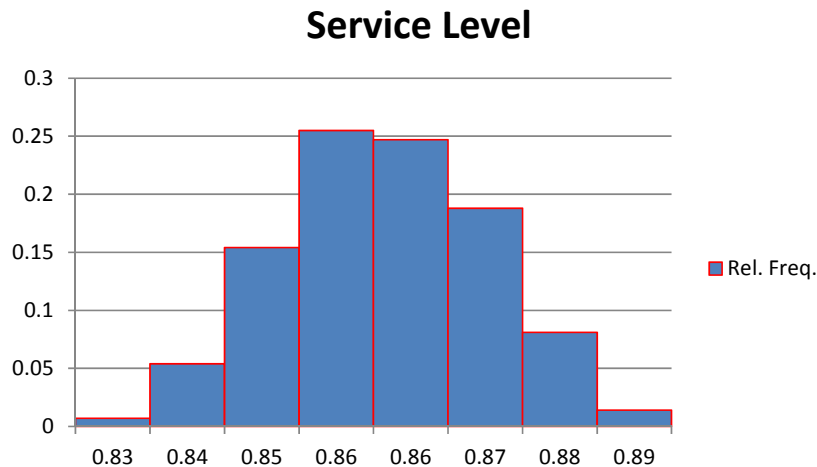


Figure 2: Histogram of the actual service level $F_B(Q_C^*|x)$ for the optimal order size under the classical approach based on 1000 replications of the estimation experiments under the classical and Bayesian approaches.

Table 1: Difference in expected profit and service level under the classical optimal order sizes using different sample sizes.

n	Difference in Expected Profit		Service Level	
	Mean	Std. Dev.	Mean	Std. Dev.
5	25.95	18.02	0.732	0.032
10	13.61	6.13	0.770	0.025
20	7.23	2.16	0.813	0.018
50	3.10	0.59	0.861	0.011
100	1.61	0.22	0.885	0.009
150	1.08	0.12	0.894	0.008
200	0.82	0.08	0.899	0.008
250	0.66	0.06	0.901	0.008
300	0.55	0.05	0.903	0.008

4.2 Estimation of the Optimal Order Size using Simulation

With the objective of illustrating how to calculate the optimal order size when the complexity of the model does not allow the calculation of a closed form expression for the solution, in this section we show the use of the PS algorithm to find the optimal order size using simulation.

Table 2: Results after applying the PS algorithm for $m = 100$ and $m = 1000$.

m	Optimal Order Size		Estimation of the Expected Profit		
			Point	Lower Bound	Upper Bound
100	d_k	42	256.00	240.65	271.35
	d_{k+1}	44	256.20	240.14	272.26
1000	d_k	40	252.37	247.89	256.85
	d_{k+1}	41	252.49	247.90	257.08

In order to apply the PS algorithm in our example, we once again use the data with $T = 15$, $n = 20$, $\sum_{i=1}^n x_i = 10$, $u = 9$, $w = 1$. Using these settings, we know that the optimal order size is $Q_B^* = 41$, with an expected profit of $B_B(Q_B^* | x) = 253.38$. Based on the algorithm described in Figure 2 of Muñoz and Muñoz (2013), the PS algorithm consists in simulating m observations w_1, \dots, w_m of the demand. Each observation w_i is obtained by first simulating the value of the parameter via the posterior distribution $p(\theta | x)$, and then simulating w_i using the forecast model (given the parameter value), which in our case corresponds to model (10).

In the case where the demand allows for a density function, the optimal order size is obtained by setting the service level to $\alpha = u/(u+w)$ and applying a valid method for quantile estimation. Nonetheless, for the discrete case, it is convenient to apply the method described by equations (8) and (9), replacing the cdf $F_B(y|x)$ for the empirical distribution of the observations w_1, \dots, w_m .

From Table 2, notice that for $m = 1000$ observations, the PS algorithm provides an optimal order size of 41, and estimates an expected profit between 247.9 and 257.08, which covers the actual value (253.38). For $m = 100$, the number of observations is insufficient for obtaining an optimal order size (surprisingly, we saw no observation with a value of 43). For values of $m > 1000$, the PS algorithm should still provide an optimal order size of 41, with a better estimate of the expected profit.

5 CONCLUSIONS AND RECOMMENDATIONS

The results obtained by experimenting with the proposed approach show that the classical approach tends to provide more conservative service levels and to overestimate the expected profit when compared to the Bayesian approach. On the other hand, as the number of real data observations increases, the results with both methods tend to coincide.

Based on the obtained results, we recommend applying the proposed Bayesian approach when the number of observations is small since in this case, the uncertainty in the parameters is significant. On the other hand, if we use stochastic simulation in order to estimate the optimal order size, we have to consider a large enough number of simulated observations in order to obtain an adequate precision.

ACKNOWLEDGMENTS

The authors express their gratitude to three anonymous referees for their helpful comments and suggestions that helped improve the quality of this paper. The work of the first author is supported by the Asociación Mexicana de Cultura A.C.

REFERENCES

- Azouri, K. S. 1985. "Bayes Solution to Dynamic Inventory Models Under Unknown Demand Distribution." *Management Science* 31:1150-1160.
- Bernardo, J. M., and A. F. M. Smith. 2009. *Bayesian Theory*, Chichester: Wiley.
- Berger, J. O., and J. M. Bernardo. 1992. "Ordered Group Reference Priors with Application to the Multinomial Problem." *Biometrika* 79:25-37.
- Chen, L., and E. L. Plambeck. 2008. "Dynamic Inventory Management with Learning About the Demand Distribution and Substitution Probability." *Manufacturing Service and Operations Management* 10:236-256.
- Chen, L. 2010. "Bounds and Heuristics for Optimal Bayesian Inventory Control with Unobserved Lost Sales." *Operations Research* 58:396-413.
- Eppen, G. D., and A. V. Iyer. 1997. "Improved Fashion Buying with Bayesian Updates." *Operations Research* 45:805-819.
- Jain, A., N. Rudi, and T. Wang. 2015. "Demand Estimation and Ordering Under Censoring: Stocking-Out Timing is (Almost) All You Need." *Operations Research* 61:134-150.
- Lariviere, M. A., and E. L. Porteus. 1999. "Stalking Information: Bayesian Inventory Management with Unobserved Lost Sales." *Management Science* 45:343-363.
- Nahmias, S., and T. L. Olsen. 2015. *Production and Operations Analysis*. Long Grove: Waveland Press.
- Muñoz, D. F., and D. F. Muñoz. 2011. "Bayesian Forecasting of Spare Parts Using Simulation." In *Service Parts Management: Demand Forecasting and Inventory Control*, edited by N. Altay y L. A. Litteral, 105-123. New York: Springer.

- Muñoz, D. F., D. G. Muñoz, and A. Ramírez-López. 2013. "On the Incorporation of Parameter Uncertainty for Inventory Management Using Simulation." *International Transactions in Operational Research* 20:493-513.
- Scarf, H. 1959. "Bayes Solutions of the Statistical Inventory Problem." *The Annals of Mathematical Statistics* 30:490-508.
- Silver, E. A. 1965. "Bayesian Determination of the Reorder Point of a Slow Moving Item." *Operations Research* 13:989-997.

AUTHOR BIOGRAPHIES

DAVID F. MUÑOZ is Professor and Head of the Department of Industrial and Operations Engineering at the Instituto Tecnológico Autónomo de México. He holds a BS in Statistics from the National Agrarian University of Peru, a M.S. in Mathematics from Catholic University of Peru, and M.S. and Ph.D. degrees in Operations Research from Stanford University. He has been recognized as an "Edelman Laureate" for the Institute for Operations Research and the Management Sciences, for participating in the project "Indeval Develops a New Operating and Settlement System Using Operations Research", winner of the 2010 Edelman Prize. His research interests include input and output analysis for stochastic simulations and the applications of Operations Research. His email address is davidm@itam.mx.

DAVID GONZALO MUNOZ is Director of Yield Analytics at AOL Advertising in Palo Alto, CA. He holds a M.S. in Management Science and Engineering from Stanford University. His research interests include discrete-event simulation, stochastic modeling, big data analytics, and machine learning. His e-mail address is davidmm82@gmail.com.