Int. J. Production Economics 149 (2014) 28–36

Contents lists available at ScienceDirect

Int. J. Production Economics

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

The data-driven newsvendor with censored demand observations

Anna-Lena Sachs*, Stefan Minner

TUM School of Management, Technische Universität München, Arcisstraße 21, Munich 80333, Germany

ARTICLE INFO

Article history:
Received 15 April 2012
Accepted 16 April 2013
Available online 29 April 2013

Keywords:
Inventory
Unobservable lost sales
Newsvendor
Linear programming

ABSTRACT

Motivated by data from a large European retail chain, we tackle the newsvendor problem with censored demand observations by a distribution-free model based on a data-driven approach. The model estimates the optimal inventory levels as a linear function of exogenous variables, e.g., price or temperature. To improve the forecast accuracy, we simultaneously estimate unobservable lost sales, determine the coefficients of the exogenous variables which drive demand, and calculate the optimal order quantity. Since demand exceeding supply cannot be recorded, we use the timing of (hourly) sales occurrences to establish (daily) sales patterns. These sales patterns allow conclusions on the amount of unsatisfied demand and thus the true customer demand. To determine the coefficients of the inventory function, we formulate a Linear Programming model that balances inventory holding and penalty costs based on the censored demand observations. In a numerical study with data generated from the normal and the negative binomial distribution, we compare our model with other parametric and non-parametric estimation approaches. We evaluate the performance in terms of inventory and service level for (non-)price-dependent demands and different censoring levels. We find that the data-driven newsvendor model copes especially well with highly censored data and price-dependent demand. In most settings with price-dependent demand, it achieves similar or higher service levels by holding lower inventories than other benchmark approaches from the literature. Finally, we show that the non-parametric approaches are better than the parametric ones based on real data with several exogenous variables where the true demand distribution is unknown.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

In a lost sales inventory system, demand exceeding supply is usually not recorded. Studies have shown that out-of-stock (OOS) rates amount to 8.3% of SKUs per category worldwide considering non-perishable products (Corsten and Gruen 2003). OOS situations are estimated to occur even more frequently for perishable products, which is due to their short shelf-lives that make full product availability less desirable (ECR Europe, 2003). This leaves the store manager without sufficient knowledge on additional sales that could have been made had the inventory level been higher. Ignoring excess demand and sticking to the same order-up-to level results consecutively in demand misspecification and more lost sales (Nahmias, 1994). Empirical studies of customer reactions to stockouts indicate that customers finding poor availability in a store on a regular basis will not only experience short-term lost sales but will decide not to return to this store in the long-run (Anderson et al., 2006). Retailers often apply rules of thumb to determine the optimal inventory level, underestimating forecast errors and levels which may affect the whole supply chain (Wagner, 2002; Tiwari and Gavirneni, 2007; Hosoda and Disney, 2009).

Consequently, unobservable lost sales estimation is a key factor in inventory planning when it comes to determining optimal order quantities. Existing approaches can be categorized into parametric and non-parametric approaches. Parametric approaches assume some kind of underlying demand distribution. It is often questionable, whether this demand distribution is appropriate in practice. It is thus useful to formulate a non-parametric estimation approach that relies on data readily available to retailers and takes external factors with a strong impact on demand into account. One such factor is price which is usually tracked and linked to the sales quantity. Even though the inventory planning literature has accounted for price-dependent demand in numerous settings (Petruzzi and Dada, 1999, Khouja, 2000), it has not yet been considered in lost sales estimation. If we assume that demand is some function of the sales price, we can incorporate additional information into our approach to better explain demand variations (Fildes et al., 2008).

We suggest a novel approach based on data-driven optimization which overcomes the limitations of existing parametric and

* Corresponding author. Tel.: +49 89 289 28203; fax: +49 89 289 28210.
E-mail addresses: anna.sachs@tum.de (A.-L. Sachs), stefan.minner@tum.de (S. Minner).

0925-5273/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijpe.2013.04.039

non-parametric approaches. There are no prior distribution assumptions and it only requires point of sale (POS) scanner data as typically available in retail stores. Observations including explanatory variables obtained from store-level scanner data are directly incorporated in inventory optimization.

2. Related work

In the following, we will focus on the distribution-free newsvendor model and unobservable lost sales estimation.

The distribution-free newsvendor model goes back to Scarf (1958). The reader is referred to Gallego and Moon (1993) for a review. Bertsimas and Thiele (2005, 2006) solve the inventory planning problem as a Linear Program (LP) that works directly with historical demand observations. This robust, data-driven approach does not require any distributional assumptions. Beutel and Minner (2012) extend this approach to causal forecasting to explain demand variations and directly estimate the optimal inventory level from historical demand data with several explanatory variables.

Lau and Lau (1996) propose a non-parametric model that does not require any prior distributional assumptions based on the product limit method (Kaplan and Meier, 1958). They suggest establishing sales patterns based on the complete demand observations. If a stockout occurs, demand is only partly observed. Based on the level of the demand observations of previous hours with full availability, unobservable lost sales are then estimated for the remainder of the day according to the sales patterns. A Tocher-curve is then fitted to the estimated fractiles using regression. Huh et al. (2011) also use the Kaplan–Meier estimator in a newsvendor-type data-driven inventory model to show that it converges to the optimal inventory policy.

Parametric approaches are based on theoretical demand distributions such as the normal in Nahmias (1994), who suggests to approximate mean demand and standard deviation from sales data given an order-up-to level inventory policy. A Poisson process might be more suitable for discrete and small demands as in Conrad (1976) or the compound Poisson in Springael and Nieuwenhuyse (2005). Agrawal and Smith (1996) emphasize that another important requirement of the distribution is to be capable of capturing the effects of demand variation as present in retailing and they suggest the negative binomial distribution by matching the sample mean and the frequency of facing zero demand to the observed frequency of demand.

Berk et al. (2007) use Bayesian updates for obtaining the parameter values of the negative binomial, Gamma, Poisson and normal distribution for the censored newsvendor problem. They rely on an approximation of the posterior distribution by matching the first two moments, given that one parameter is known (e.g., variance). Lu et al. (2006) consider Bayesian updates in the context of durable goods for a general distribution function and apply their findings to the normal distribution with known variance. Lu et al. (2008) analytically investigate the benefits of overstocking to learn about the true demand in a Bayesian setting. Tan and Karabati (2004) suggest an updating mechanism to achieve a desired service level by iteratively adjusting the inventory level.

3. Data-driven model with unobservable lost sales estimation

A retailer observes historical sales for a newsvendor product. Demand variability can be partially explained by external variables. Therefore, the data-driven newsvendor model determines the optimal order quantity as a function of the external variables (Beutel and Minner, 2012). As an extension of this model, the retailer now only observes censored demands and has to estimate unobservable lost sales before placing an order. Excess demand is lost so that sales data contain incomplete demand information. Any demand occurring after the product is out-of-stock remains unobserved.

Assuming that demand per day follows a similar pattern across all observations, we base our approach to estimate unobserved lost sales on Lau and Lau (1996). We assume that the level of the demand before the stockout occurs also reflects the influence of the external variables on demand. Fig. 1 contains data on average hourly sales quantities of lettuce at one retail store as an example for a product with a short life-cycle, although this is not a newsvendor-product in the strong sense. The level of demand increases with lower prices, but the overall sales pattern is similar for low and high prices. Demand reaches its peak in the morning hours and around noon, then decreases, and in the late afternoon increases again.

3.1. Cost model

The retailer has collected a set of $i=1,\ldots,N$ sales observations and data on external variables X_{ji} per period i. The number of external variables to be considered is denoted by $j=1,\ldots,m$.

If we assume that the product does not stockout, i.e., it is fully available in each period i, Beutel and Minner (2012) determine the inventory level as a linear function of external factors influencing demand for a single-period newsvendor problem. The newsvendor's objective is to minimize total costs consisting of holding costs h for leftover inventory that is salvaged at the end of the day and penalty costs v per unit of unmet demand. The main idea is to fit a linear inventory function that determines the optimal order quantity depending on the values of the external variables X_{ji} to the sample data. The coefficients β_j correspond to these external variables including one coefficient β_0 for $X_{0i}=1$ for the intercept. As a result, we obtain the target inventory level B_i as a product sum of the external factors with their respective coefficients $B_i=\sum_{j=0}^{m}\beta_j X_{ji}$. The coefficients of this linear inventory function are set such that a cost minimization objective or a target service level objective is achieved.

The idea of fitting the coefficients to past demand observations for cost objectives or service level constraints (in-stock probability or fill-rate) is discussed in Beutel and Minner (2012). The models therein assume that the retailer observes complete demand. This assumption may be violated if a retailer cannot track unobservable lost sales as is typically the case with perishable products such as fruits and vegetables.

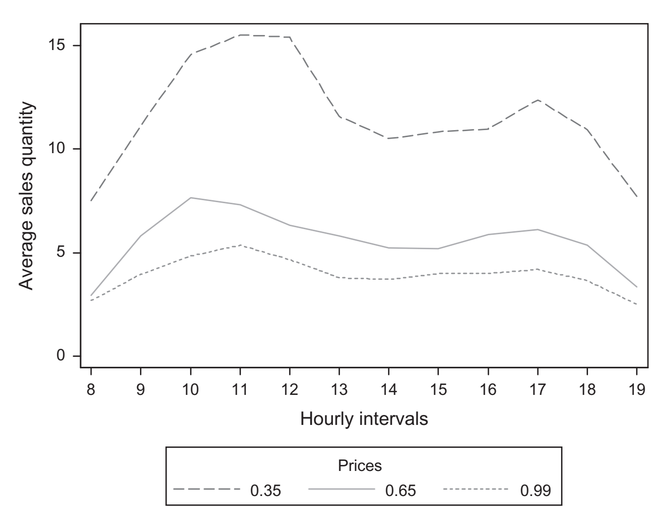

Fig. 1. Sales patterns for different prices.

Download English Version:

https://daneshyari.com/en/article/5080168

Download Persian Version:

https://daneshyari.com/article/5080168

Daneshyari.com