International Journal of Forecasting 28 (2012) 485–496

Contents lists available at SciVerse ScienceDirect

International Journal of Forecasting

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

Forecasting the intermittent demand for slow-moving inventories: A modelling approach

Ralph D. Snyder [a,*], J. Keith Ord [b], Adrian Beaumont [a]

[a] Department of Econometrics and Business Statistics, Monash University, Clayton, VIC 3800, Australia

[b] McDonough School of Business, Georgetown University, Washington, DC 20057, USA

ARTICLE INFO

Keywords:
Croston's method
Exponential smoothing
Hurdle shifted Poisson distribution
Intermittent demand
Inventory control
Prediction likelihood
State space models

ABSTRACT

Organizations with large-scale inventory systems typically have a large proportion of items for which demand is intermittent and low volume. We examine various different approaches to demand forecasting for such products, paying particular attention to the need for inventory planning over a multi-period lead-time when the underlying process may be non-stationary. This emphasis leads to the consideration of prediction distributions for processes with time-dependent parameters. A wide range of possible distributions could be considered, but we focus upon the Poisson (as a widely used benchmark), the negative binomial (as a popular extension of the Poisson), and a hurdle shifted Poisson (which retains Croston's notion of a Bernoulli process for the occurrence of *active* demand periods). We also develop performance measures which are related to the entire prediction distribution, rather than focusing exclusively upon point predictions. The three models are compared using data on the monthly demand for 1046 automobile parts, provided by a US automobile manufacturer. We conclude that inventory planning should be based upon dynamic models using distributions that are more flexible than the traditional Poisson scheme.

1. Introduction

Modern inventory control systems may involve thousands of items, many of which show very low levels of demand. Furthermore, such items may be requested only on an occasional basis. When events corresponding to positive demands occur only sporadically, we refer to the demand as *intermittent*. When the average size of a customer order is large, a continuous distribution is a suitable description, but when it is small, a discrete distribution is more appropriate.

In this paper, our interest focuses upon intermittent demand with low volume. On occasion, such stock keeping units (SKUs) may be of very high value, such as, for example, spare aircraft engines. However, even when the individual units are of low value, it is not unusual for them to represent a large percentage of the number of SKUs, so that they collectively represent an important element in the planning process. Johnston and Boylan (1996a, p. 121) cite an example where the average number of purchases of an item by a customer was 1.32 occasions per year, and "For the slower movers, the average number of purchases was only 1.06 per item [per] customer". Similarly, in the study of car parts discussed in Section 6, out of 2509 series with complete records for 51 months, only 1046 had (a) ten or more months with positive demands, *and* (b) at least some positive demands in the first 15 and the last 15 months.

Demand forecasting for high volume products can be handled successfully using exponential smoothing methods, for which a voluminous body of literature exists; see for example Hyndman, Koehler, Ord, and Snyder (2008)

* Corresponding author. Tel.: +61 3 9905 2366; fax: +61 3 9905 5474.
E-mail addresses: ralph.snyder@monash.edu (R.D. Snyder), ordk@georgetown.edu (J.K. Ord), adrian.beaumont@monash.edu (A. Beaumont).

0169-2070/$ – see front matter © 2011 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ijforecast.2011.03.009

and Ord, Koehler, and Snyder (1997). When volumes are low, the exponential smoothing framework must be based upon a distribution that describes count data, rather than the normal distribution. Further, as was recently emphasized by Syntetos, Nikolopoulos, and Boylan (2010), it is not sufficient to look at point forecasts when making inventory decisions. Instead, they recommend the use of stock control metrics. We accept their viewpoint completely, but since such metrics depend upon the underlying prediction distribution, we have opted to work with such distributions directly. This choice is reinforced by the observation that prediction distributions are applicable to count problems beyond inventory control. Moreover, the information on costs and lead times required when using inventory criteria was not available for the data considered in Section 6.

The remainder of the paper is structured as follows. It begins in Section 2 with a review of the literature on forecasting intermittent demand. The focus here is on models that allow for both non-stationary and stationary features. For example, the demand for spare parts may increase over time as the machines age and then decline as they fail completely or are withdrawn from service. In Section 3, we summarize the different models which will be considered in the empirical analysis and examine how they might be estimated and how they might be used to simulate various prediction distributions. Since our particular focus is on the ability of a model to furnish the entire prediction distribution, not just point forecasts, we examine suitable performance criteria in Section 4. Issues relating to model selection are examined briefly in Section 5. In Section 6 we present an empirical study using data on the monthly demand for 1046 automobile parts. Then, in Section 7, we examine the links between forecasting and management decision making, with an illustration of the use of prediction distributions in inventory management. Finally, various conclusions from our research are summarized briefly in Section 8.

2. Review of the literature on intermittent demand

The classic paper on this topic is that of Croston (1972), with corrections by Rao (1973). Croston's key insight was that:

> When a system is being used for stock replenishment, or batch size ordering, the replenishment will almost certainly be triggered by a demand which has occurred in the most recent interval (Croston, 1972, p. 294).

The net effect of this phenomenon when forecasting the demand for a product which is required only intermittently is that the mean demand is over-estimated and the variance is under-estimated. Thus, an inventory decision based on the application of the usual exponential smoothing formulae will typically produce inappropriate stock levels. Croston then proceeded to develop an alternative approach based upon:

- an exponential smoothing scheme for updating the expected time gap between successive active demand periods;
- an exponential smoothing scheme for updating the expected demand in active periods; and
- an assumption that the time gaps and the demand in active periods are statistically independent.

Since the original paper by Croston, a number of extensions and improvements to the method have been made, notably by Johnston and Boylan (1996a) and Syntetos and Boylan (2005). Syntetos and Boylan (2001) first showed that the original Croston estimators were biased; they then (see Syntetos & Boylan, 2005) developed a new method, which we refer to as the bias-adjusted Croston method, and evaluated its performance in an extensive empirical study. Their out-of-sample comparisons indicate that the new method provides superior point forecasts for "faster intermittent" items; that is, those with relatively short mean times between active demand periods.

Snyder (2002) identified some logical inconsistencies in the original Croston method and examined the use of a time-dependent Bernoulli process. Unlike with Croston's method, distinct smoothing parameters were used for the positive demands and the time gaps. Snyder went on to develop a simulation procedure which provides a numerical determination of the predictive distribution for the lead-time demand. Shenstone and Hyndman (2005) showed that there is no possible model which will lead to the Croston forecast function unless we allow a sample space for active period demands that can take on either negative or positive values.

2.1. Low volume, intermittent demand

There is an extensive body of literature on low count time series models which is potentially applicable to forecasting the demand for slow moving items. Most expositions rely on a Poisson distribution to represent the counts but introduce serial correlation through a changing mean (and variance). Models based on lagged values of the count variable essentially have a single source of randomness (Davis, Dunsmuir, & Wang, 1999; Heinen, 2003; Jung, Kukuk, & Liesenfeld, 2006; Shephard, 1995). By contrast, models which are based upon unobservable components have an additional source of randomness driving the evolution of the mean (Davis, Dunsmuir, & Wang, 2000; Durbin & Koopman, 2001; Harvey & Fernandes, 1989; West & Harrison, 1997; West, Harrison, & Migon, 1985; Zeger, 1988). In addition, there are also several multiple source of error approaches based upon integer-valued autoregressive (INAR) models (Al-Osh & Alzaid, 1987; McCabe & Martin, 2005; McKenzie, 1988).

Single- and dual-source of error models for count data were compared by Feigin, Gould, Martin, and Snyder (2008), who found the dual source of error model to be more flexible, which is not true for Gaussian measurements (Hyndman et al., 2008). However, their analysis was conducted under a stationarity assumption, while, as was noted earlier, demand series are typically non-stationary. The results in Section 6 suggest that non-stationary single source of error models are competitive with other approaches for count time series.

2.2. Evaluation of the Croston method

Willemain, Smart, Shockor, and DeSautels (1994) conducted an extensive simulation study which violated some