Revisiting the Assumption of Geometric Distribution in Modeling Intermittent Demand

Matthew Lindsey
Stephen F. Austin State University, Rusche College of Business
P O Box 13060 SFA Station, Nacogdoches, TX 75965
936-468-1858 lindseymd@sfasu.edu

Robert Pavur
The University of North Texas College of Business
1155 Union Circle #311160, Denton, TX 76203-5017
940-565-3107 pavur@unt.edu

ABSTRACT

The majority of research related to intermittent demand has used a geometric distribution for the interval between demands. Many articles investigate the effect of the distribution of the demand when it occurs on estimates of the population demand per period. However, a gap in the literature occurs in examining the effect of the distribution for the interval between demands. This research will investigate how popular forecasting techniques for intermittent demand perform when certain discrete Beta-Binomial distributions model the distribution of the interval between demands instead of the geometric distribution. Single Exponential Smoothing, Croston's method, and the Syntetos Boylan Approximation (SBA) Bias Corrected Croston Method are investigated with various demand distributions.

INTRODUCTION

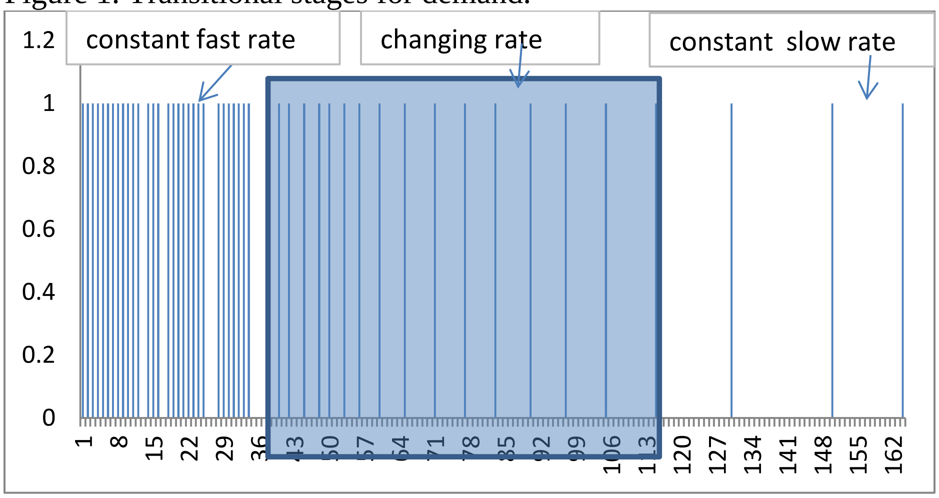
Croston (1972) provided an elegant solution to forecasting intermittent demand by forming two series for the demand size and the time between demands. Since the early work related to intermittent demand, particular attention has been paid to the interval between demands, the size of the demands and the relationship between size and interval (Willemain et al. 1994). Demand classification schemes published in the literature provide guidance on when to use a particular forecasting method.

The slow-moving demand forecasting literature identifies the necessity for classifying demand occurrences to offer direction in selecting the right forecasting methodology and stock control methods (Syntetos, Boylan, and Croston, 2005; Boylan, Syntetos and Karakostas, 2006). By correctly classifying demand, the forecaster can select the best technique for the given situation (Syntetos, et al 2005). When picking the appropriate method, Williams (1984) used the coefficient of variation to classify "lumpiness" in relation to the lead-time as an indicator of "intermittence." This method had shortcomings identified in the research by Syntetos, et al (2005) mainly in that it did not always distinguish demand patterns sufficiently.

Croston (1972) made a few basic assumptions about slow-moving data series. He assumed demand would occur as a Bernoulli process, resulting in independent and identically distributed (IID) demand resulting in a geometric distribution. Demand size was assumed to follow a normal distribution and be IID. A considerable amount of research has studied violations to the demand size assumption and has investigated different distributions for demand size. The majority of the research in this area, however, does not investigate Crostron's assumption of a geometric distribution for the interval between demands. Croston(1972, Segerstedt (1994), Willemain et al, 1994, Bagchi, Havya and Ord (1983) and leven and Segerstedt (2004) demonstrate the Geometric distribution is assumed to be the underlying distribution for time between positive levels of demand. Numerous articles have investigated the demand size and related distributions. Inventory management articles have investigated distributions related to demand lead time. Little investigation has occurred related to the time between demands.

When the demand is transitioning the assumption of a geometric distribution for the intervals between demands is questionable. For example a textbook that has been in print for many years without an update would have a high level of demand for the first few years it is out. As the used book market increases, the demand for the new book will decrease and eventually become a slow mover. As it shifts from regular demand to slow demand the average time between demands is obviously changing. This research will investigate different demand distributions for the time between demands as illustrated in Figure 1. The shaded area shows the area that is of interest to this research. Little research has been done with respect to the distribution of the demand rate

Figure 1. Transitional stages for demand.



Croston's Method

In Croston's (1972) method the smoothing constant like SES it assumes a constant demand average of size μ every p periods, so the mean demand is not μ/p, but

$$y^* = \left(\frac{\alpha}{p}\right)\left\{\frac{p\alpha}{\left[1-(1-\alpha)^p\right]}\right\} \tag{1}$$

Additional estimates using SES are made for the average demand and the time between demands with updates if a demand occurs. Willemain et al. 1994 should be consulted for the methodology.

The variance becomes $V\left(y_t''\right)=\left[\frac{\alpha}{2-\alpha}\right]\left[\frac{(p-1)^2\mu^2}{p^4}+\frac{\sigma^2}{p^2}\right].$ (2)

Optimal results using Croston's (1972) are achieved with alpha values between 0.1 and 0.2.

Bias Correction

Syntetos and Boylan have provided the formulas to compute the bias correction with the true variance from the Beta Binomial distribution. A useful formula in constructing this bias correction is the following, which is based on a Taylor expansion of a ratio.

$$E\left(\frac{x_1}{x_2^{'}}\right) \approx \left[\frac{\Theta_1}{\Theta_2}\right] + \left[\frac{1\partial^2 g}{2\partial_2^2} Var\left(x_2\right)\right] + \ldots \tag{3}$$

If the time between demand interval series is not auto-correlated and the intervals (p_t) are geometrically distributed with a mean of p and homogeneous variance of p(p-1) then:

$$\text{Var}(x_2) = \text{Var}\ (\ p_t'\ ) = \left(\frac{\alpha}{2-\alpha}\right) Var(p_t) = \left(\frac{\alpha}{2-\alpha}\right) p(p-1) \tag{4}$$

Assuming the demand sizes are distributed with a mean m then equation 7 can be transformed to:

$$E\left(\frac{x_1}{x_2^{'}}\right) \approx \left[\frac{\Theta_1}{\Theta_2}\right] + \frac{1}{2}\frac{\alpha}{2-\alpha}\frac{2\Theta_1}{\Theta_2^3}\, p(p-1) \tag{5}$$

Then it follows that:

$$E\left(\frac{z_1'}{p_t^{'}}\right) \approx \left[\frac{\mu}{p}\right] + \frac{\alpha}{2-\alpha}\mu\frac{(p-1)}{p^2} \tag{6}$$

Beta Binomial

The Beta-Binomial Distribution is a family of discrete distributions that arises from probability theory and statistics (Griffiths, 1973). It is a finite set of non-negative integers that occurs when the probability of success from a set number of random or unknown Bernoulli trials. This discrete distribution is the binomial distribution when the probability of success at each trial is not fixed but random and follows the beta distribution. It is generally applied to capture over dispersion in binomial types of distributed data. The Beta-Binomial probability mass function is the following function with parameters α and β.

$$\binom{n}{k}\frac{B\left(k+\alpha, n-k+\beta\right)\alpha}{B(\alpha,\beta)} \tag{7}$$

The Beta-Binomial Cumulative Distribution Function is:

$$1-\frac{B\left(\beta+n=k=1, \alpha+k+1\right)_3 F_2\left(a,b;k\right)}{B(\alpha,\beta)B(n-k,k+2)(n+1)} \tag{8}$$

The mean of the Beta-Binomial Function is:

$$\frac{n\alpha}{(\alpha+\beta)} \tag{9}$$

The variance of the Beta-Binomial Function is:

$$\frac{n\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{10}$$

Table 1 illustrates how the probability mass function (PMF) is calculated for the Beta Binomial distribution. This distribution was selected, as its shape is flexible. Its shape can actually mimic the normal distribution as well as skewed distributions.

Table 1. The computation of the probability for a Beta-Binomial distribution							
n= 9, alpha = 0.028, beta = 0.035							
Actual Time till Demand (t)	Periods with no Demand	Comb(n,x) (A)	Beta(alpha,Beta) (B)	Gamma (n+alpha+Beta) (C)	Gamma (t-1 +alpha) (D)	Gamma (n-t-1+Beta) (E)	Beta-BinPMF (A*D*E)/(C*B)
1	0	1	64.19	46,152.02	35.16	43460	0.516
2	1	9	64.19	46,152.02	0.985	5,408.8	0.016
3	2	36	64.19	46,152.02	1.012	768.85	0.009
4	3	84	64.19	46,152.02	2.053	127.4	0.007
5	4	126	64.19	46,152.02	6.215	25.30	0.007
6	5	126	64.19	46,152.02	25.04	6.27	0.007
7	6	84	64.19	46,152.02	125.9	2.07	0.007
8	7	36	64.19	46,152.02	758.8	1.02	0.009
9	8	9	64.19	46,152.02	5333	0.981	0.016
10	9	1	64.19	46,152.02	42813	28.03	0.405

METHODOLOGY

Simulations were conducted for 600 periods for 5 distributions shown in Table 2, available upon request. Four are discrete Beta-Binomial distributions and the fifth is the geometric distribution. Forecasts were computed with five forecasting methodologies. The first Beta-Binomial distribution has a high probability of demand in the first and last period and very low probability of demand in periods 2 to 9. For the second distribution, the probability of demand is similar in each period and is generally near a 1 in 10 chance of demand occurring. The third set contains probabilities that generally increase from very low to about a 1 in 4 chance of demand and then decrease again. The fourth set is similar to the third but the probabilities increase from 0.018 to a max of 0.182 over 10 periods. For the geometric, the probability of demand slowly decreases.

Simulations were conducted using SES, Croston's (1972) method, the bias corrected SBA Croston Method, SBA Croston with a known probability and Croston with a known variance. The Root Mean Square Error (RMSE) was computed for Croston's method, the bias corrected Croston's method, SES, the bias corrected SBA Croston known probability and the SBA Bias corrected Croston with the Beta Binomial distribution. The research is conducted at two alpha levels. The first being a smoothing constant equal to 0.1 and the second when alpha is 0.3.

RESULTS

This research investigates a non-geometric distribution for the interval between demands for intermittent demand. Simulations are conducted using four discrete Beta Binomial distributions. The average root mean square error is provided in Table 3 (available upon request). The lowest error is in bold.. As might be expected, all techniques developed for intermittent demand yielded lower average errors than forecasts using SES. Generally the lower alpha levels provided the lowest error. The error in the forecast varied more from distribution to distribution than from technique to technique, with the exception of SES. That is all forecasting methodologies using the fourth Beta-Binomial Distribution had less error than any of the other distributions.

CONCLUSIONS

This research provides an initial look at assuming a non-geometric distribution for the interval between demands for intermittent demand. A limited exploration is conducted using four discrete Beta Binomial distributions applied to variations of Croston's method and using the Geometric distribution as a benchmark. An important conclusion is that the SBA Bias correction is designed for the Geometric distributional assumption. A bias correction using the true variance of the distribution for the time between demands often provides slightly better RMSEs. Croston's procedure with no bias correction may actually outperform the bias corrected procedure under the presence of a non-geometric distribution for the length of the time interval between demands.

REFERENCES

Bagchi, U., Hayya, J. C., & Ord, J. K. (1983). The Hermite distribution as a model of demand during lead time for slow-moving items. *Decision Sciences, 14*(4), 447-466.

Boylan, J. E. and Syntetos, A. A., (2007). The accuracy of a modified Croston procedure. *International Journal of Production Economics, 107*(2), 511-517.

Boylan, J. E. , Syntetos, A. A. and Karakostas, G.C. (2006). Classification for forecasting and stock control: a case study. *Journal of the Operational Research Society, 59*, 473-481.

Croston, J. D., (1972). Forecasting and stock control for intermittent demands. *Operational Research Quarterly, 23*(3), 289-303.

Griffiths, D. (1973). Maximum Likelihood Estimation for the Beta-Binomial Distribution and an Application to the Household Distribution of the Total Number of Cases of a Disease. *Biometrics, 29*(4), 637-648. doi:10.2307/2529131

Segerstedt, A. (1994). Inventory control with variation in lead times, especially when demand is intermittent. *International Journal of Production Economics, 35*(1/3), 365-372.

Syntetos, A. A., Boylan, J. E. and Croston, J.D (2005). On the categorization of demand patterns. *Journal of the Operational Research Society, 56*, 495-503.

Teunter, R.H., & Sani, B., (2009). On the bias of Croston's forecasting method, *European Journal of Operational Research, 194*(1), 177-183.

Willemain, T. R., Smart, C. N., Shockor, J. H. and DeSautels, P. A., (1994). Forecasting intermittent demand in manufacturing: A comparative evaluation of Croston's method. *International Journal of Forecasting, 10*(4), 529-538.

Williams, T.M. (1984). Stock control with sporadic and slow-moving demand. Journal of the *Operational Research Society, 35*, 939-948.