

DePaul University

From the Selected Works of Nezh Altay

April, 2012

On the demand distributions of spare parts

A A Syntetos, University of Salford

M Z Babai, King Saud University

Nezh Altay, DePaul University



Available at: https://works.bepress.com/nezih_altay/14/

On the demand distributions of spare parts

A.A. Syntetos^{a*}, M.Z. Babai^{bc} and N. Altay^d

^aUniversity of Salford, Salford, UK; ^bBEM-Bordeaux Management School, Bordeaux, France; ^cKing Saud University, Riyadh, Saudi Arabia; ^dDePaul University, Chicago, USA

(Received 24 May 2010; final version received 31 January 2011)

Spare parts have become ubiquitous in modern societies, and managing their requirements is an important and challenging task with tremendous cost implications for the organisations that are holding relevant inventories. Demand for spare parts arises whenever a component fails or requires replacement, and as such the relevant patterns are different from those associated with ‘typical’ stock keeping units. Such demand patterns are most often intermittent in nature, meaning that demand arrives infrequently and is interspersed by time periods with no demand at all. A number of distributions have been discussed in the literature for representing these patterns, but empirical evidence is lacking. In this paper, we address the issue of demand distributional assumptions for spare-parts management, conducting a detailed empirical investigation on the goodness-of-fit of various distributions and their stock-control implications in terms of inventories held and service levels achieved. This is an important contribution from a methodological perspective, since the validity of demand distributional assumptions (i.e. their goodness-of-fit) is distinguished from their utility (i.e. their real-world implications). Three empirical datasets are used for the purposes of our research that collectively consist of the individual demand histories of approximately 13,000 SKUs from the military sector (UK and USA) and the Electronics Industry (Europe). Our investigation provides evidence in support of certain demand distributions in a real-world context. The natural next steps of research are also discussed, and these should facilitate further developments in this area from an academic perspective.

Keywords: demand distributions; inventory; spare parts management; empirical study

1. Introduction

Parametric approaches to stock control rely upon a lead-time demand distributional assumption and the employment of an appropriate forecasting procedure for estimating the moments of such a distribution (typically mean and variance). For the case of fast demand items, the Normality assumption is typically sufficient. However, Stock Keeping Units (SKUs) often exhibit intermittent or irregular demand patterns that may not be represented by the Normal distribution. This is almost invariably the case for service/spare parts. Such products have become ubiquitous in modern societies, and managing their requirements is an important and challenging task with tremendous cost implications for the organisations that are holding relevant inventories.

Intermittent demand appears at random, with some time periods having no demand at all. Moreover, demand, when it occurs, is not necessarily for a single unit or a constant demand size. In the academic literature, intermittent demand is often referred to as lumpy, sporadic, or erratic demand. A conceptual framework that serves the purpose of distinguishing between such non-normal demand patterns has been discussed by Boylan *et al.* (2007).

As mentioned above, intermittent demand items are typically spare parts; those may be engineering spares (e.g. Mitchell 1962, Hollier 1980, Strijbosch *et al.* 2000) or parts kept at the wholesaling/retailing level (e.g. Sani 1995, Syntetos *et al.* 2009c). Demand for such items arises whenever a component fails or requires replacement as opposed to being generated according to buying behaviours of end-consumers (and the way that demand moves upstream in a supply chain). However, intermittent demand may indeed be characterising any SKU within the range of products offered by all organisations at any level of the supply chain (e.g. Croston 1972, Willemain *et al.* 1994). Such items are typically the so-called ‘slow-movers’, and although they contribute little to the sales of an organisation, they may collectively account for up to 60% of the total stock value (Johnston *et al.* 2003). Consequently, small improvements regarding their management may be translated into substantial organisational savings. Such items are particularly prevalent in the aerospace, automotive, and IT sectors, and they are often

*Corresponding author. Email: a.syntetos@salford.ac.uk

the SKUs at greatest risk of obsolescence. As will be discussed later in this section, the empirical data available for the purposes of our investigation relate to spare parts; although such SKUs constitute the focus of our research, natural extensions with regards to our findings can be made to all products characterised by intermittent demand structures (please refer also to the last section of our paper).

Research in the area of forecasting and stock control for intermittent demand items has developed rapidly in recent years (e.g. Syntetos and Boylan 2005, Chatfield and Hayya 2007, Quintana and Leung 2007, Altay *et al.* 2008, Gutierrez *et al.* 2008) with new results implemented into software products because of their practical importance. Key issues remaining in this area relate to (1) the further development of robust operational definitions of intermittent demand for forecasting and stock-control purposes and (2) a better modelling of the underlying demand characteristics for the purpose of proposing more powerful estimators useful in stock control. Both issues link directly to the hypothesised distribution used for representing the relevant demand patterns. Surprisingly though, not much has been contributed in this area in the academic literature, and this constitutes the main purpose of our work.

1.1 Objectives and structure of the paper

Intermittent demand patterns are built from constituent elements, namely a demand arrival process and a distribution of the demand sizes, when demand occurs. As such, the very nature of such patterns necessitates, conceptually at least, the employment of compound distributions, such as the negative binomial distribution (NBD). Although a number of distributions have been put forward in the literature (please refer also to the second section of this paper) there is still more empirical evidence needed on the goodness-of-fit of these distributions to real data. That is to say, there is a need for more insights into the empirical validity of various distributional assumptions. In addition, a distinction needs to be made between the goodness-of-fit on demand-per-period data and the validity of a distributional assumption for representing lead-time demand. The latter is what matters in a stock-control context, and this distinction has been overlooked in the academic literature. Finally, it is also important to distinguish between the empirical validity and utility of a demand distributional assumption, the latter being related to the stock-control implications of using a particular distribution. Goodness-of-fit tests focus on the entire demand distribution, whereas stock control performance is explicitly dependent upon the fit on the right-hand tail of a distribution. All the above-discussed issues are addressed as part of our work.

In particular, the objective of this paper is threefold: first, we conduct an empirical investigation that enables the analysis of the goodness-of-fit of various two- (or single-) parameter statistical distributions used in the literature in the context of intermittent demand; second, we extend the analysis discussed above to assess the linkage between the validity of various demand distributional assumptions and their utility (i.e. actual implications for stock control in terms of inventory volumes and service levels achieved); third, we discuss the implications of our work for Operational Research (OR) theory and practice, and we provide an agenda for further research in this area. We use three empirical datasets for the purposes of our analysis that collectively constitute the individual demand histories of approximately 13,000 SKUs. Two datasets come from the military sector (Royal Air Force, RAF UK and US Defense Logistics Agency, DLA) and one from the Electronics industry. In all cases, the SKUs are spare/service parts.

At this point, it is important to note that some non-parametric procedures have also been suggested in the literature to forecast intermittent demand requirements (e.g. Willemain *et al.* 2004, Porras and Dekker 2008). Such approaches typically rely upon bootstrapping procedures that permit a re-construction of the empirical distribution of the data, thus making distributional assumptions and estimation of the moments of such distributions redundant. Although it has been claimed that such approaches have an advantage over parametric methodologies, more empirical research is needed to evaluate the conditions under which one approach outperforms the other. In this paper, we will be focusing solely on parametric approaches to inventory management.

The remainder of this paper is organised as follows. In Section 2, an introduction to forecasting and stock control issues in the context of intermittent demand is presented, along with a review of the demand distributions discussed in the literature and/or used by practitioners. In Section 3, we present the datasets used for the purpose of this empirical investigation followed by a discussion on the statistical goodness-of-fit tests that we have considered for analysing the performance of various distributions. The empirical validity and utility of distributional assumptions are then explored in Sections 4 and 5 for demand-per-period data and lead-time demands respectively.

Finally, the conclusions of our research along with its implications and some natural extensions for further work in this area are presented in Section 6.

2. Research background

In this section, a brief review of the literature on issues related to parametric inventory forecasting is presented. First, we establish the need for estimating the mean and variance of demand in an inventory control context, followed by a discussion of various suggestions that have been made in the literature with regards to the hypothesised distribution of demand patterns, focusing on the case of intermittence. The second sub-section culminates with the selection of a number of distributions to be considered in the empirical part of our work.

2.1 Parametric forecasting

Practical parametric approaches to inventory management rely upon estimates of some essential demand distribution parameters. The decision parameters of the inventory systems (such as the Re-Order Point or the Order-Up-To-Level) are then based on these estimates. Different inventory systems require different variables to be forecast. Some of the most cited, for example (R, s, S) policies (Naddor 1975, Ehrhardt and Mosier 1984), require only estimates of the mean and variance of demand. (In such systems, the inventory position is reviewed every R periods, and if the stock level drops to the re-order point s , enough is ordered to bring the inventory position up to the order-up-to-level S .)

In other cases, and depending on the objectives or constraints imposed on the system, such estimates are also necessary, although they do not constitute the 'key' quantities to be determined. We may consider, for example, an (R, S) or an (s, Q) policy operating under a fill-rate constraint – known as P_2 . (In the former case, the inventory position is reviewed periodically, every R periods, and enough is ordered to bring it up to S . In the latter case, there is a continuous review of the inventory position and as soon as that drops to, or below, s an order is placed for a fixed quantity Q .) In those cases we wish to ensure that $x\%$ of demand is satisfied directly off the shelf, and estimates are required for the probabilities of any demands exceeding S or s (for the (R, S) an (s, Q) policy respectively). Such probabilities are typically estimated indirectly, based on the mean demand and variance forecast in conjunction with a hypothesised demand distribution. Similar comments apply when these systems operate under a different service-driven constraint: there is no more than an $x\%$ chance of a stock-out during the replenishment cycle (this service measure is known as P_1). Consequently, we need to estimate the $(100-x)$ th percentile of the demand distribution.

In summary, parametric approaches to forecasting involve estimates of the mean and variance of demand. In addition, a demand distribution needs also to be hypothesised, in the majority of stock-control applications, for the purpose of estimating the quantities of interest. Issues related to the hypothesised demand distribution are addressed in the following sub-section. For a detailed account on issues related to the estimation of mean demand and its variance in an intermittent demand context, interested readers are referred to Syntetos and Boylan (2008) and Syntetos *et al.* (2009b).

2.2 Demand distribution

Intermittent demand patterns are characterised by infrequent demands, often of variable size, occurring at irregular intervals. Consequently, it is preferable to model demand from constituent elements, i.e. the demand size and inter-demand interval. Therefore, compound theoretical distributions (that explicitly take into account the size-interval combination) are typically used in such contexts of application. We first discuss some issues related to modelling demand arrivals and hence inter-demand intervals. We then extend our discussion to compound demand distributions.

If time is treated as a discrete (whole number) variable, demand may be generated based on a Bernoulli process, resulting in a geometric distribution of the inter-demand intervals. When time is treated as a continuous variable, the Poisson demand generation process results in negative exponentially distributed inter-arrival intervals.

There is sound theory in support of both geometric and exponential distribution for representing the time interval between successive demands. There is also empirical evidence in support of both distributions (e.g. Dunsmuir and Snyder 1989, Kwan 1991, Willemain *et al.* 1994, Janssen 1998, Eaves 2002). With Poisson

arrivals of demands and an arbitrary distribution of demand sizes, the resulting distribution of total demand over a fixed lead time is compound Poisson. Inter-demand intervals following the geometric distribution in conjunction with an arbitrary distribution for the sizes results in a compound binomial distribution.

Regarding the compound Poisson distributions, the stuttering Poisson, which is a combination of a Poisson distribution for demand occurrence and a geometric distribution for demand size, has received the attention of many researchers (for example: Gallagher 1969, Ward 1978, Watson 1987). Another possibility is the combination of a Poisson distribution for demand occurrence and a normal distribution for demand sizes (Vereecke and Verstraeten 1994), although the latter assumption has little empirical support. Particularly for lumpy demands, the demand size distribution is heavily skewed to the right, rendering the normality assumption far from appropriate. Quenouille (1949) showed that a Poisson-Logarithmic process yields a negative binomial distribution (NBD). When event arrivals are assumed to be Poisson-distributed, and the order size is not fixed but follows a logarithmic distribution, total demand is then negative binomially distributed over time.

Another possible distribution for representing demand is the gamma distribution. The gamma distribution is the continuous analogue of the NBD and 'although not having a priori support [in terms of an explicit underlying mechanism such as that characterising compound distributions], the gamma is related to a distribution which has its own theoretical justification' (Boylan 1997, p. 168). The gamma covers a wide range of distribution shapes; it is defined for non-negative values only, and it is generally mathematically tractable in its inventory control applications (Burgin and Wild 1967, Burgin 1975, Johnston 1980). Nevertheless if it is assumed that demand is discrete, then the gamma can be only an approximation to the distribution of demand. At this point, it is important to note that the use of both NBD and gamma distribution requires estimation of the mean and variance of demand only. In addition, there is empirical evidence in support of both distributions (especially the former), and therefore they are recommended for practical applications.

If demand occurs as a Bernoulli process, and orders follow the Logarithmic-Poisson distribution (which is not the same as the Poisson-Logarithmic process that yields NBD demand), then the resulting distribution of total demand per period is the log-zero-Poisson (Kwan 1991). The log-zero-Poisson is a three-parameter distribution and requires a rather complicated estimation method. Moreover, it was found by Kwan (1991) to be empirically outperformed by the NBD. Hence, the log-zero-Poisson cannot be recommended for practical applications. One other compound binomial distribution that appeared in the literature is that involving normally distributed demand sizes (Croston 1972, 1974). However, and as discussed above, a normality assumption is unrealistic, and therefore the distribution is not recommended for practical applications.

Despite the inappropriateness of the normal distribution for representing demand sizes, it may in fact constitute a reasonable assumption for lead-time demand itself, when lead times are long (see also Syntetos and Boylan 2008). This is because long lead times permit central limit theorem effects for the sum of demands over the corresponding period, thus making the normality assumption more plausible. In addition, the assumption of normality may prove to be good when the coefficient of variation (CV) of the distribution of demand per period is small. Finally, algorithms based on normality are simple to implement, making the normal distribution a very commonly assumed one among practitioners.

For very slow-moving items, such as those commonly encountered in a military context, for example, the Poisson distribution is known to offer a very good fit, and much of the stock control theory in this area has been developed upon the explicit assumption that demand per period is Poisson distributed (see, for example, Silver *et al.* 1998). In this case, demand is assumed to arrive as a Poisson process couple with unit-sized transactions. In an early work, Friend (1960) also discussed the use of a Poisson distribution for demand occurrence, combined with demands of constant size. Vereecke and Verstraeten (1994) presented an algorithm developed for the implementation of a computerised stock-control system for spare parts in a chemical plant. The demand was assumed to occur as a Poisson process with a package of several pieces being requested at each demand occurrence. The resulting distribution of demand per period was called a 'Package Poisson' distribution. The same distribution has appeared in the literature under the name 'hypothetical SKU' (*h-SKU*) Poisson distribution (Williams 1984), where demand is treated as if it occurs as a multiple of some constant, or 'clumped Poisson' distribution, for multiple item orders for the same SKU of a fixed 'clump size' (Ritchie and Kingsman 1985). The 'Package Poisson' distribution requires, as the Poisson distribution itself, an estimate of the mean demand only.

The review of the literature presented above indicates that it is worth while assessing the empirical validity and utility of the following demand distributional assumptions: (1) Poisson; (2) NBD; (3) Stuttering Poisson; (4) Gamma; and (5) Normal. In the next sections we conduct relevant tests and comment on the plausibility of the relevant assumptions for applications in an intermittent demand context.

3. Empirical investigation

In this section, we first describe the datasets used for the purposes of this empirical investigation, followed by a discussion on statistical goodness-of-fit testing.

3.1 Empirical data

The empirical databases available for the purposes of our research come from the US Defense Logistics Agency (DLA), Royal Air Force (RAF), and Electronics Industry and they consist of the individual monthly demand histories of 4588, 5000, and 3055 SKUs respectively. Some information regarding these datasets is presented in Table 1, followed by detailed descriptive statistics (Mean and Standard Deviation – SD) on the demand data series characteristics for each of the datasets presented in Tables 2–4. (At this point it should be noted that the time series considered have not been tested for stationarity.)

Tables 2–4 indicate that the three datasets considered for the purpose of our empirical investigation reflect a wide range of possible underlying demand patterns. Datasets no. 1 and no. 3 are associated with a somewhat low degree of intermittence; the average demand intervals have a median value of approximately 3 and 2.5 respectively. However, there is a relatively high degree of erraticness; demand sizes are highly variable with a median standard deviation of approximately 5 (6) but a maximum one equal to 1472 (9149) for dataset no. 1 (no. 3). On the other hand, dataset no. 2 consists of average demand intervals that are very large, since their median value is almost nine months, but the demand sizes are associated with a comparatively lower degree of erraticness.

Table 1. Empirical datasets.

No.	Country	Industry	No. of SKUs	Time bucket	History length	Lead-time info
1	USA	Military/DLA	4588	Month	60	No
2	UK	Military/RAF	5000	Month	84	Yes
3	Europe	IT	3055 12,643	Month	48	Constant = 3

Table 2. Dataset No. 1: US Defense Logistics Agency (DLA).

4588 SKUs	Demand intervals		Demand sizes		Demand per period	
	Mean	SD	Mean	SD	Mean	SD
Min	1.000	0.000	1.000	0.000	0.083	0.279
25th percentile	1.967	1.665	2.894	2.314	0.650	1.672
Median	3.278	3.236	5.375	5.142	1.750	3.749
75th percentile	5.600	6.049	11.940	12.435	4.550	9.403
Max	12.000	24.597	1326.875	1472.749	783.917	1219.012

Table 3. Dataset No. 2: Royal Air Force (RAF).

5000 SKUs	Demand intervals		Demand sizes		Demand per period	
	Mean	SD	Mean	SD	Mean	SD
Min	3.824	0.000	1.000	0.000	0.036	0.187
25th percentile	7.273	5.431	1.556	0.815	0.155	0.538
Median	9.000	6.930	3.833	3.062	0.369	1.452
75th percentile	11.571	8.630	11.333	9.315	1.155	4.434
Max	24.000	16.460	668.000	874.420	65.083	275.706

Notes: Dataset No. 2 has been used in the studies conducted by Syntetos *et al.* (2009a) and Teunter *et al.* (2010)

Table 4. Dataset No. 3: electronics.

3055 SKUs	Demand intervals		Demand sizes		Demand per period	
	Mean	SD	Mean	SD	Mean	SD
Min	1.000	0.000	1.000	0.000	0.042	0.245
25th percentile	1.500	1.011	3.462	3.011	0.896	2.215
Median	2.556	2.285	5.900	6.220	2.104	4.501
75th percentile	4.700	4.389	12.122	13.863	6.010	10.480
Max	24.000	32.527	5366.188	9149.349	5366.188	3858.409

Note: Dataset No. 3 has been used in the study conducted by Babai *et al.* (2010).

3.2 Statistical goodness-of-fit tests

Two tests have been mainly used and discussed in the literature for checking statistically significant fit, namely: the chi-square test and the Kolmogorov–Smirnov (K-S) test (see, for example, Harnett and Soni 1991). These tests measure the degree of fit between observed and expected frequencies. Problems often arise with the standard chi-square test through the requirement that data needs to be grouped together in categories to ensure that each category has an expected frequency of at least a minimum of a certain number of observations. Some modifications of this test have also been considered in the literature. A modified chi-square test has been developed for the purpose of testing the goodness-of-fit for intermittent demands (Eaves 2002). This test differs in that boundaries are specified by forming a certain number of categories with similar expected frequencies throughout, rather than combining groups just at the margins. However, the implementation of this test requires the specification of the number of categories to be used. We encountered difficulty in using the standard or modified chi-square test in our research, namely that of deciding how to specify the categories' intervals or the number of categories. On the other hand, the K-S test does not require grouping of the data in any way, so no information is lost; this eliminates the troublesome problem of categories' interval specification.

In an inventory context one could argue that measures based on the entire distribution can be misleading (Boylan and Syntetos 2006). A good overall goodness-of-fit statistic may relate to the chances of low demand values, which can mask poor forecasts of the chances of high-demand values. However, for inventory calculations, attention should be restricted to the upper end of the distribution (say, the 90th or 95th percentiles). The development of modified goodness-of-fit tests for application in inventory control, and even more specifically in an intermittent demand context, is a very important area but not one considered as part of this research. (We return to this issue in the last section of the paper where the natural next steps of research are discussed.) Consequently, we have selected the K-S test for the purpose of assessing goodness of fit.

At this point, it should be noted that the K-S test assumes that the data are continuous, and the standard critical values are exact only if this assumption holds. Several researchers (e.g. Noether 1963, 1967, Walsh 1963, Slakter 1965) have found that the standard K-S test is conservative when applied to data that are discrete. The standard exact critical values provided for the continuous data are larger than the true exact critical values for discrete data. Consequently, the test is less powerful if the data are discrete as in the case of this research; it could result in accepting the null hypothesis at a given significance level while the correct decision would have been not to accept the null hypothesis.

As discussed in the previous section, we are considering five distributions, the fit of which is tested on the demand data related to 12,643 SKUs. First, the distribution of the demand per period has been considered, followed by tests on the lead-time demand. This distinction is a very important one, since the latter distribution may in fact be considerably different from the former and yet is the one required for stock control purposes.

Critical values have been computed based on K-S statistical tables for 1% and 5% significance levels. We consider that:

- there is a 'Strong Fit' if the P -value is less than the critical value for 5%;
- there is 'Good Fit' if the P -value is less than the critical value for 1% but larger than that for 5%;
- there is 'No Fit' if the P -value is larger than the critical value for 1%.

Table 5. Goodness-of-fit results: demand per period.

Dataset No.	No. of SKUs	Distribution	Percentage of SKUs		
			Strong fit	Good fit	No fit
1	4588	Poisson	39.45	5.51	55.04
		NBD	71.19	3.86	24.95
		Stuttering Poisson	84.18	3.64	12.18
		Normal	11.84	14.25	73.91
		Gamma	13.84	3.88	82.28
2	5000	Poisson	59.84	2.94	37.22
		NBD	82.48	2.7	14.82
		Stuttering Poisson	98.64	0.48	0.88
		Normal	12.2	18.12	69.68
		Gamma	19.2	12.32	68.48
3	3055	Poisson	32.64	7.4	59.96
		NBD	73.94	5.31	20.75
		Stuttering Poisson	79.05	4.49	16.46
		Normal	9.92	14.34	75.74
		Gamma	11.69	3.83	84.48

4. Empirical results: demand per period

In Table 5 we present the percentage of SKUs that satisfy the various degrees of goodness-of-fit taken into account in our research, for each of the datasets and statistical distributions considered.

As shown in Table 5, the discrete distributions, i.e. Poisson, NBD and Stuttering Poisson, provide overall a better fit than the continuous ones, i.e. Normal and Gamma. More precisely, and with regards to ‘Strong Fit’, the Stuttering Poisson distribution performs best in all three datasets considered in our research. This is followed by the NBD and then by the Poisson distribution. On the other hand, the Normal distribution is judged to be far from appropriate for intermittent demand items; as will be discussed in the next section, this is partly, at least, due to the experimental structure employed for this part of our analysis that relied upon the distribution of demand per time period rather than the distribution of the lead time demand.

Contrary to our expectations, the Gamma distribution has also been found to perform poorly. This may be explained in terms of the inconsistency between the distribution under concern, which is continuous in nature, and the discreteness of the (demand) data employed in our goodness-of-fit tests. We return to this issue in the last section of the paper where the next steps of our research are discussed in detail.

4.1 Linking the goodness-of-fit to demand characteristics

Johnston and Boylan (1996) offered for the first time an operationalised definition of intermittent demand for forecasting purposes (demand patterns associated with an average inter-demand interval (p) greater than 1.25 forecast revision periods). They compared, on a wide range of theoretically generated demand patterns, the Mean Squared Error performance of Croston’s method that has been specifically designed for intermittent demand estimation, and Single Exponential Smoothing (SES) that is typically used for fast-moving items. They found that the former method performs better than the latter on demand patterns associated with an average inter-demand interval (p) greater than 1.25 forecast revision periods. Their contribution lies on the identification of the average inter-demand interval as a demand classification parameter to determine intermittence rather than the specification of an exact cutoff value. Syntetos *et al.* (2005) took this work forward by developing a demand classification scheme that relies upon both p and the squared coefficient of variation of demand sizes (CV^2). In their work, they compared (assuming fixed lead-times) the theoretical MSE performance of three estimators: Croston, SES, and SBA (Syntetos-Boylan Approximation; Syntetos and Boylan 2005) and they identified regions where one method outperforms all others.

Both schemes serve the purpose of identifying the best performing estimation procedure in each category; however, inventory control issues and demand distributional assumptions were not addressed. Boylan *et al.* (2007)

assessed the stock control implications of the work conducted by Syntetos *et al.* (2005) by means of experimentation on an inventory system developed by a UK-based software manufacturer. The researchers demonstrated, empirically, the insensitivity of the p cutoff value, for demand classification purposes, in the approximate range 1.18–1.86 periods.

In this section, we attempt to explore the potential linkages between demand distributional assumptions and the classification scheme developed by Syntetos *et al.* (2005). By doing so, we aim at providing some further insights into the potential usefulness of such distributions in an intermittent demand context.

In Figure 1 we present for dataset no. 1 and each of the distributions considered the SKUs associated with a ‘Strong Fit’ as a function of the inter-demand intervals (p , expressed as a number of zero demand periods: average interval between successive demand occurrences) and the squared demand coefficient of variation (CV^2 , dimensionless measure) of the demand sizes. (The results for the other two datasets lead to similar insights, and they are presented separately in the Appendix.)

As shown in Figure 1 (and in the Appendix) and theoretically expected, both the Stuttering Poisson and the Negative Binomial distribution perform comparatively better for all the datasets considered. This is true for the SKUs with both high inter-demand intervals (e.g. SKUs with p being up to 12 in dataset no. 1 or SKUs with a p value up to 24 in datasets no. 2 and no. 3) and low demand intervals (e.g. SKUs with p values starting from 1 in datasets no. 1 and no. 3). Moreover, it should be noted that there is a strong fit of NBD and Stuttering Poisson to all the SKUs that are also associated with a strong fit of the Poisson distribution, which is expected since both distributions under concern are compound Poisson ones. The SKUs where there is commonly a strong fit of those three distributions are those characterised by relatively low CV^2 values. The Normal distribution seems to have the narrowest CV^2 range among all tested distributions, showing a good fit roughly between zero and one. This implies that the Normal distribution is not a good fit for erratic data. Furthermore, the Normal distribution performs well

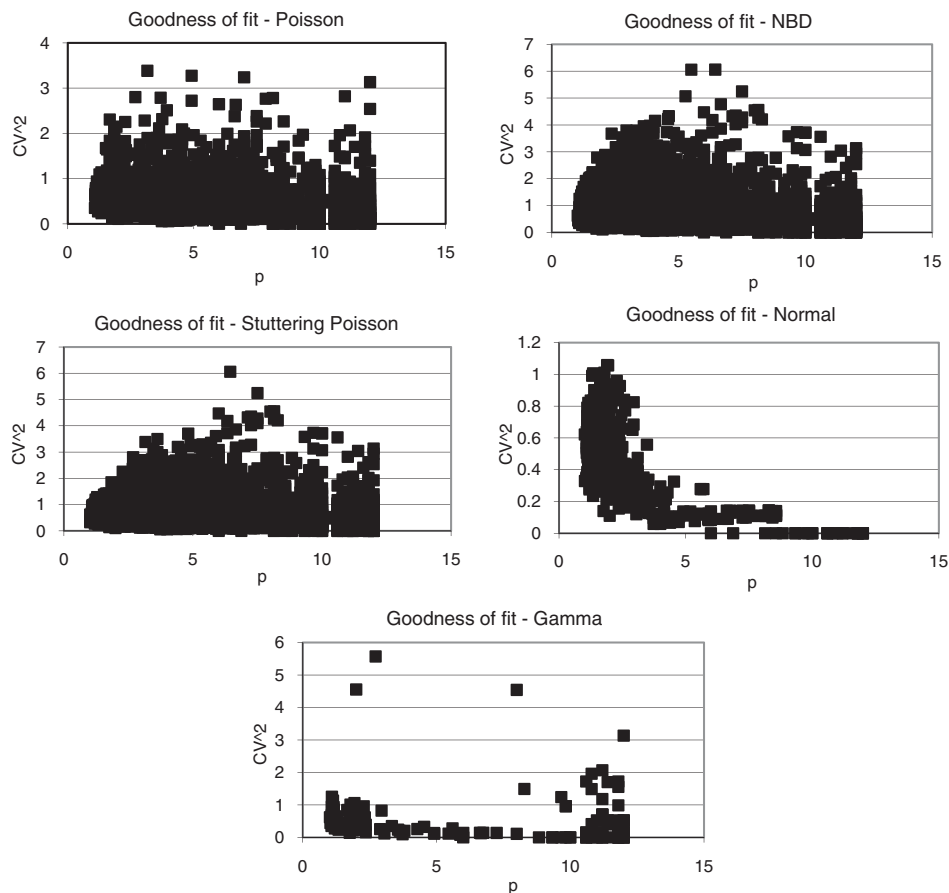


Figure 1. Dataset no. 1: goodness-of-fit results.

for the SKUs with relatively low inter-demand intervals (e.g. SKUs with p values close to 1 in datasets no. 1 and no. 3 and $p = 3.82$ in the dataset no. 2). However, there are also a few SKUs with high inter-demand intervals (p going up to 12 in dataset no. 1 and 24 in datasets no. 2 and no. 3) for which the Normal distribution provides a strong fit. Those latter SKUs have a minimum CV^2 (i.e. $CV^2 = 0$) which can be explained by the fact that their demand is very low (in most of the cases, the demand is equal to one) and can fit to the Normal distribution with low mean (i.e. equivalently high values of p) and variance. Finally, in addition to the SKUs where there is a fit to the Normal distribution (those with low values of p), the Gamma distribution provides also a strong fit to the SKUs with very high values of p (i.e. SKUs with an inter-demand interval going up to 12 periods in dataset no. 1 and 24 periods in datasets no. 2 and no. 3) and high CV^2 values (i.e. SKUs with CV^2 up to 6 in dataset no. 1, $CV^2 = 10$ in the dataset no. 2 and $CV^2 = 8$ in the dataset no. 3). This is also expected since the Gamma distribution is known to be very flexible in terms of its mean and variance, so it can take high values for its p and CV^2 and can be reduced to the Normal distribution for certain parameters of the mean and the variance.

5. Empirical results: lead-time demand

In this section, we are concerned with the goodness-of-fit of the various distributions considered in our work to lead-time demand data. The empirical utility of the relevant demand distributional assumptions is also assessed by means of experimentation with a periodic order-up-to level system. At this point it is important to note that no lead-time information has been made available to us for dataset no. 1, and consequently this has not been considered for this part of our research. All SKUs included in dataset no. 3 are associated with a fixed lead time of three periods (months).

The goodness-of-fit results are presented in Table 6. Although the separation among the tested distributions is not as clear as in the case of demand per period, the results indicate that NBD performs, overall, best with about 62% (weighted average) of 8055 SKUs showing a strong fit. Stuttering Poisson comes in a close second. Despite the general expectation that the Normal distribution would be a good representative of lead-time demand owing to the effect of Central Limit Theorem, our results do not demonstrate dominance by the Normal distribution. On the contrary, the Normal distribution seems to have a split between strong fit and no fit; only about 47% of the 8055 SKUs showing a strong fit for Normal, and about 43% showing no fit. Although the Gamma distribution still compares unfavourably to the compound distributions, it does perform better than the demand-per-period case. On the contrary, the Poisson distribution is associated with an inferior performance when compared with the demand-per-period case; this was theoretically expected since lead-times are overall rather lengthy resulting in high lead-time demand cumulative sizes which cannot be represented by the Poisson distribution. In the next sub-section, we design a simulation study to test the utility of these results.

5.1 Linking the goodness-of-fit to stock control performance

In this sub-section we are concerned with the empirical utility of the demand distributional assumptions discussed thus far in our paper. (The Poisson distribution will not be considered in this part of the analysis, as preliminary

Table 6. Goodness-of-fit results: lead-time demand.

Dataset No.	No. of SKUs	Distribution	Percentage of SKUs		
			Strong fit	Good fit	No fit
2	5000	Poisson	49.08	4.88	46.04
		NBD	68.36	5.00	26.64
		Stuttering Poisson	68.10	6.24	25.66
		Normal	53.58	9.24	37.18
		Gamma	57.42	9.98	32.60
3	3055	Poisson	23.24	6.91	69.85
		NBD	51.29	6.91	41.80
		Stuttering Poisson	50.25	6.38	43.37
		Normal	35.16	10.70	54.11
		Gamma	34.70	12.93	52.34

results indicated a very poor performance as compared with that of all other distributions.) The implications of using these distributions are evaluated through an order-up-to-level (R, S) policy that operates under a Cycle Service Level (CSL) objective (proportion of demand satisfied directly from stock). Mean demand is estimated by using the SBA estimator, and the variance of demand is estimated through the smoothed MSE procedure. By keeping all aspects of the stock-control system fixed and varying only the hypothesised demand distribution, the effects of using such distributions may be evaluated and contrasted with each other.

Three target CSL values have been considered: 90%, 95%, and 99%. Performance is evaluated by means of the average holding volumes and the average backlog volumes (per time period for each SKU) as well as the achieved CSLs. In Tables 7 and 8, we show, for datasets no. 2 and no. 3, respectively, for every demand distribution and for every target CSL, the relevant quantities, averaged across all SKUs.

Overall, the results in Tables 7 and 8 indicate that the Normal distribution is the worst-performing one for higher target CSLs (i.e. $CSL = 95\%$ and 99%). However, for a target CSL equal to 90%, the Normal compares favourably to the other distributions. The Normal distribution may not approximate well the tail of the empirical distribution of data, but performs very satisfactorily otherwise. For high target CSLs, the NBD provides the best results, followed by the Gamma distribution and Stuttering Poisson.

The results presented in the above tables are somewhat difficult to interpret in terms of deriving tangible conclusions on the comparative performance of the various distributions. Higher holding volumes are associated with higher achieved CSLs, making it difficult to appreciate the comparative merits of using one distribution instead of another. To that end, we have also considered the presentation of our results through the use of efficiency curves. In Figures 2 and 3, we plot (for datasets no. 2 and no. 3 respectively) the holding volumes required to support a certain achieved CSL for each of the three target values and each of the distributions considered in our investigation.

Table 7. Stock control-performance: Dataset No. 2.

Distribution	Percentage	Holding (volumes)	Backlog (volumes)	CSL (%)
Normal	90	30.277	0.383	85.9
	95	36.298	0.344	87.4
	99	47.798	0.293	89.7
Gamma	90	23.208	0.461	84.1
	95	33.644	0.365	87.4
	99	72.270	0.250	91.1
Stuttering Poisson	90	23.675	0.460	84.0
	95	34.224	0.364	87.2
	99	74.251	0.249	91.0
NBD	90	23.102	0.464	84.7
	95	33.582	0.365	88.8
	99	72.208	0.248	92.9

Table 8. Stock-control performance: Dataset No. 3.

Distribution	Percentage	Holding (volumes)	Backlog (volumes)	CSL (%)
Normal	90	76.065	3.829	89.3
	95	92.700	3.191	91.1
	99	125.302	2.386	93.5
Gamma	90	64.595	3.873	88.4
	95	87.881	3.001	91.5
	99	161.189	2.017	95.2
Stuttering Poisson	90	66.253	3.872	88.4
	95	89.094	2.997	91.4
	99	161.253	2.012	95.1
NBD	90	64.241	3.891	88.1
	95	87.079	3.017	91.5
	99	159.268	2.042	95.2

In these figures, for a fixed holding volume, a curve that is further from the x -axis shows that the CSL is higher and indicates that the corresponding demand distribution is more efficient.

The efficiency curves show that the NBD and Gamma are the more efficient distributions, followed by the Stuttering Poisson. In fact, for a fixed holding volume, the former two distributions provide the highest CSL. Although the Stuttering Poisson provides a good overall fit (as demonstrated in the beginning of this section) the stock control performance of this distribution is somewhat inferior to that associated with Gamma (that performed less satisfactorily in terms of its goodness-of-fit results). This reinforces potentially earlier comments that we have made with respect to the importance of fitting the tail of the distribution as opposed to the entire distribution of lead-time demand. In addition, the results demonstrate that the Normal distribution is the least efficient one for targets above 90%, and this is in accordance with theoretical expectations.

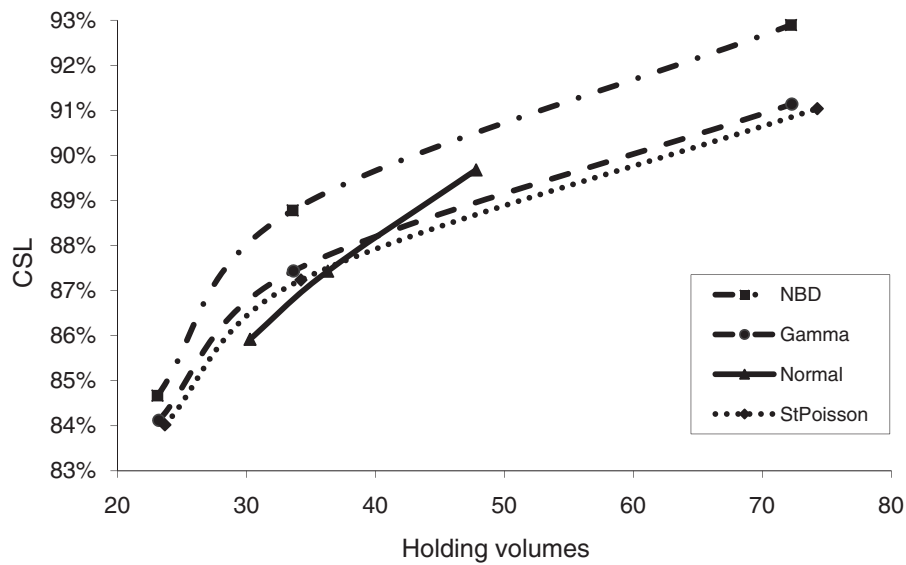


Figure 2. Efficiency curves: CSL against holding volume (dataset no. 2).

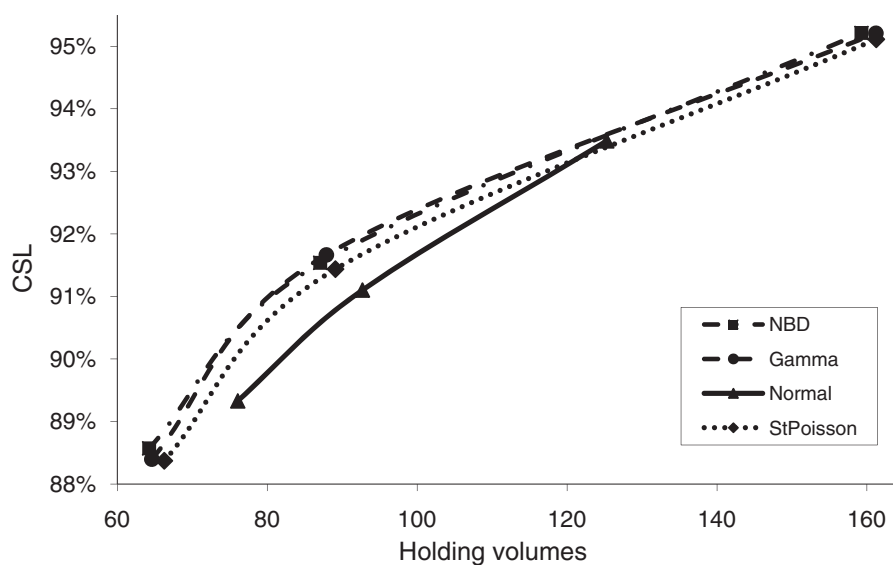


Figure 3. Efficiency curves: CSL against holding volume (dataset no. 3).

In summary, our results for the lead-time demand distribution tests are overall different from the findings emerged from the demand-per-period analysis. The Poisson distribution was found to perform considerably better in the latter case, whereas the opposite is true for the Gamma distribution. The Normal distribution proved to be a weak candidate for the demand-per-period distribution. It was found to perform better for lead-time demand (probably due to Central Limit Theorem effects), although it may still be judged as far from appropriate for implementation in a spare-parts context. The results have also indicated that the goodness-of-fit performance on lead-time demand may not translate to similar insights with regards to the stock-control implications of using the various distributions. Most noticeably, the Gamma distribution was found to compare favourably to the Stuttering Poisson with regards to stock control, whereas the goodness-of-fit tests point to the opposite direction.

6. Conclusions, implications, and further research

Practical parametric approaches to inventory management rely upon an explicit demand distributional assumption. Although the Normal distribution is typically adequate for 'fast' demand items, this is not true for spare parts that are typically characterised by intermittent demand structures. Although a number of distributions have been discussed in the academic literature for representing such patterns, evidence on the empirical validity of the relevant assumptions (i.e. empirical goodness of fit) is lacking. Moreover, the linkage between the validity and the utility of these distributional assumptions, the latter referring to the actual stock control implications of using a particular distribution, has not been explored in the literature. An exception to that is perhaps a PhD thesis conducted at Lancaster University in the early 1990s (Kwan 1991). Unfortunately, no parts of that thesis have ever been published in the academic literature. Finally, an important distinction to be made is that between the performance of a hypothesised distribution when fitting demand-per-period data and that when applied to lead-time demands, the latter being what matters in a real stock control system.

In this paper a number of distributions are first selected as potential candidates for representing intermittent demands, on the basis of: (1) theoretical arguments, (2) intuitive appeal, and (3) empirical support. These distributions are: the Poisson, Negative Binomial Distribution (NBD), Stuttering Poisson, Normal, and Gamma. The empirical goodness-of-fit of these distributions is then assessed for both demand-per-period data and lead-time demands by employing the K-S test. Conclusions are being drawn on the empirical validity of the various distributional assumptions followed by an assessment of their stock control implications (i.e. their empirical utility), under a periodic order-up-to-level system. The empirical databases available for the purposes of our investigation come from the US DLA, RAF, and Electronics Industry, and they consist of the individual monthly demand histories of 4588, 5000, and 3055 SKUs respectively.

When we consider demand-per-period data, the results indicate that both the NBD and Stuttering Poisson provide the most frequent fit. Both these distributions are compound in nature, meaning that they account explicitly for a demand arrival process (Poisson) and a different distribution for the transaction sizes (Log series and Geometric for the NBD and Stuttering Poisson respectively). Despite previous claims, the Gamma distribution does not perform very well, and the same is true for the Normal distribution. This may be attributed to the continuous nature of these distributions (since their fit is tested on discrete observations) but also to the fact that demand per unit time period is considered as opposed to lead time demand. The Poisson distribution provides a 'reasonable' fit and this is theoretically expected for slow moving items.

We have also attempted to link the goodness-of-fit of the various distributions to two key characteristics that have been shown to be collectively sufficient for defining intermittent demands (Syntetos *et al.* 2005). These characteristics are the inter-demand interval and the squared coefficient of variation of the demand sizes. Goodness-of-fit performance was then analysed across all series as a function of those characteristics, and some interesting insights were generated into the comparative performance of the distributions.

Recognising that the lead-time demand distribution is more important from an inventory control perspective, we also tested the goodness-of-fit of the above-mentioned distributions on lead-time demand. NBD and Stuttering Poisson were found to perform best also in this case; however, the remaining distributions were associated with differences in performance. The Poisson distribution was found to perform considerably worse in that case, and this may be explained in terms of the relatively long lead times associated with our empirical dataset. The Gamma distribution though was found to perform better for lead-time demand; similarly, the Normal distribution

proved also to perform better (Central Limit Theorem effects), although it may still be judged as far from appropriate for implementation in a spare-parts context.

Subsequently, simulation experiments using already-proven forecasting (Syntetos–Boylan Approximation) and inventory control methods ((R, S) type periodic review policy) were used to assess the stock-control implications of using the various distributions. Although NBD was indeed confirmed as the best-performing distribution, the results have also indicated that the goodness-of-fit performance on lead-time demand may not translate to similar insights with regards to the stock-control implications of using a particular distribution. Most noticeably, the Gamma distribution was found to compare favourably to the Stuttering Poisson with regards to stock control, whereas the goodness-of-fit tests pointed to the opposite direction.

In summary, the results of our empirical investigation suggest that the NBD performs best in an inventory context. This is followed by the Gamma and Stuttering Poisson distribution, the former found to outperform the latter in terms of their stock-control implications. We expect further academic studies to take note of our findings and build spare-parts management research on these distributions. We also expect these findings to be of interest to practitioners dealing with spare parts and being interested in improving automated procedures towards the most effective management of such items. At this point we should also note that although spare parts constituted the focus of our research, natural extensions with regards to our findings can be made to all products characterised by intermittent demand structures.

Our research has demonstrated that the choice of a demand distribution has considerable effects on stock-control performance in a spare parts context. In addition, the work presented in this paper has revealed a number of interesting themes for further research. Distributional assumptions play a critical role in practical inventory management applications, and further work on the following issues should prove to be valuable from both a theoretical and practitioner perspective:

- The development of modified goodness-of-fit tests for application in inventory control, and even more specifically in an intermittent demand context, is a very important area. In particular, putting more emphasis on the right-hand tail of the distribution seems appropriate for stock-control applications.
- Quantifying the effect that the inconsistency between the discrete nature of demand data and the continuous nature of certain distributions (e.g. Gamma) may have on goodness-of-fit statistics constitutes an interesting research question.
- The inconsistency between the discrete nature of demand observations and the implicit assumption of continuous data employed by various goodness-of-fit tests should be further explored.
- Replication of the analysis conducted in this paper in larger demand datasets coupled with the assessment of the empirical utility of various distributional assumptions under other stock control formulations should help advance knowledge in this area. Also, it would be worth while exploring the sensitivity of our findings under the realistic scenario of variable lead times; see, for example, Babai *et al.* (2009) and, for continuity purposes, also Babai and Dallery (2009).
- Empirical studies are needed to determine the comparative merits of non-parametric approaches (such as bootstrapping) in a spare-parts context. Such approaches do not rely upon any underlying distributional assumption. They rather consider the empirical distribution function and emphasise the direct estimation of the relevant percentiles (say the 95th or the 99th percentile). Although this is a very appealing property, the current empirical knowledge base is not extended enough to suggest how such approaches compare to parametric inventory management in real terms.
- In addition, and given that the inverse of the distribution function is the quantile function, another relevant approach to the issues discussed in this paper relates to *quantile forecasting*. Directly forecasting quantiles avoids the need for assumptions regarding the shape and spread of the underlying distribution (Taylor 2007) and should be given more attention in the inventory literature.
- Finally, assessing Value at Risk (VaR) in inventory management is arguably also an area that has not been given sufficient attention. VaR is increasingly being used in financial management as a natural measure of the risk taken with a given position. In the framework of inventory management it can work as well (Tapiero 2005) and is particularly relevant, since VaR problems involve only one tail of the distribution (of the financial returns). Extensions into the distribution of inventory costs are feasible and intuitively appealing (Luciano *et al.* 2003) and further work into the issue of the efficient computation of VaR would constitute an important addition to the current state of knowledge in the area of inventory control.

Acknowledgements

The research described in this paper has been partly supported by the Royal Society, UK (2007/Round 1, Int. Incoming Short Visits – North America). In addition, we would like to acknowledge the financial support provided by DePaul University that allowed a research visit by Aris Syntetos to DePaul in 2010.

References

- Altay, N., Rudisill, F., and Litteral, L.A., 2008. Adapting Wright's modification of Holt's method to forecasting intermittent demand. *International Journal of Production Economics*, 111, 389–408.
- Babai, M.Z. and Dallery, Y., 2009. Dynamic versus static control policies in single stage production-inventory systems. *International Journal of Production Research*, 47, 415–433.
- Babai, M.Z., et al., 2009. Periodic inventory control systems with non-stationary demand and lead time uncertainty, analysis and empirical investigation. *International Journal of Production Research*, 47, 2461–2483.
- Babai, M.Z., Syntetos, A.A., and Teunter, R., 2010. On the empirical performance of (T, s, S) heuristics. *European Journal of Operational Research*, 202, 466–472.
- Boylan, J.E., 1997. *The centralisation of inventory and the modelling of demand*. Thesis (PhD). University of Warwick.
- Boylan, J.E. and Syntetos, A.A., 2006. Accuracy and accuracy-implication metrics for intermittent demand. *F.O.R.E.S.I.G.H.T: International Journal of Applied Forecasting*, 4, 39–42.
- Boylan, J.E., Syntetos, A.A., and Karakostas, G.C., 2007. Classification for forecasting and stock control: a case study. *Journal of the Operational Research Society*, 59, 473–481.
- Burgin, T.A., 1975. The gamma distribution and inventory control. *Operational Research Quarterly*, 26, 507–525.
- Burgin, T.A. and Wild, A.R., 1967. Stock control experience and usable theory. *Operational Research Quarterly*, 18, 35–52.
- Chatfield, D.C. and Hayya, J.C., 2007. All-zero forecasts for lumpy demand: a factorial study. *International Journal of Production Research*, 45, 935–950.
- Croston, J.D., 1972. Forecasting and stock control for intermittent demands. *Operational Research Quarterly*, 23, 289–304.
- Croston, J.D., 1974. Stock levels for slow-moving items. *Operational Research Quarterly*, 25, 123–130.
- Dunsmuir, W.T.M. and Snyder, R.D., 1989. Control of inventories with intermittent demand. *European Journal of Operational Research*, 40, 16–21.
- Eaves, A., 2002. *The forecasting for the ordering and stock holding of consumable spare parts*. Thesis (PhD). Lancaster University.
- Ehrhardt, R. and Mosier, C., 1984. A revision of the power approximation for computing (s, S) inventory policies. *Management Science*, 30, 618–622.
- Friend, J.K., 1960. Stock control with random opportunities for replenishment. *Operational Research Quarterly*, 11, 130–136.
- Gallagher, D.J., 1969. Two periodic review inventory models with backorders and stuttering Poisson demands. *A.I.I.E. Transactions*, 1, 164–171.
- Gutierrez, R.S., Solis, A.O., and Mukhopadhyay, S., 2008. Lumpy demand forecasting using neural networks. *International Journal of Production Economics*, 111, 409–420.
- Harnett, D.L. and Soni, A.K., 1991. *Statistical methods for business and economics*, 4th ed. New York: Addison-Wesley.
- Hollier, R.H., 1980. The distribution of spare parts. *International Journal of Production Research*, 18, 665–675.
- Janssen, F.B.S.L.P., 1998. *Inventory management systems; control and information issues*. Thesis (PhD). Tilburg University.
- Johnston, F.R., 1980. An interactive stock control system with a strategic management role. *Journal of the Operational Research Society*, 31, 1069–1084.
- Johnston, F.R. and Boylan, J.E., 1996. Forecasting for items with intermittent demand. *Journal of the Operational Research Society*, 47, 113–121.
- Johnston, F.R., Boylan, J.E., and Shale, E.A., 2003. An examination of the size of orders from customers, their characterization and the implications for inventory control of slow moving items. *Journal of the Operational Research Society*, 54, 833–837.
- Kwan, H.W., 1991. *On the demand distributions of slow moving items*. Thesis (PhD). Lancaster University.
- Luciano, E., Peccati, L., and Cifarelli, D.M., 2003. VaR as a risk measure for multiperiod static inventory models. *International Journal of Production Economics*, 81–82, 375–384.
- Mitchell, G.H., 1962. Problems of controlling slow-moving engineering spares. *Operational Research Quarterly*, 13, 23–39.
- Naddor, E., 1975. Optimal and heuristic decisions in single and multi-item inventory systems. *Management Science*, 21, 1234–1249.
- Noether, G.E., 1963. Note on the Kolmogorov statistic in the discrete case. *Metrika*, 7, 115–116.
- Noether, G.E., 1967. *Elements of nonparametric statistics*. New York: Wiley.
- Porras, E.M. and Dekker, R., 2008. An inventory control system for spare parts at a refinery: an empirical comparison of different reorder point methods. *European Journal of Operational Research*, 184, 101–132.
- Quenouille, M.H., 1949. A relation between the logarithmic, Poisson and negative binomial series. *Biometrics*, 5, 162–164.

- Quintana, R. and Leung, M.T., 2007. Adaptive exponential smoothing versus conventional approaches for lumpy demand forecasting: case of production planning for a manufacturing line. *International Journal of Production Research*, 45, 4937–4957.
- Ritchie, E. and Kingsman, B.G., 1985. Setting stock levels for wholesaling: Performance measures and conflict of objectives between supplier and stockist. *European Journal of Operational Research*, 20, 17–24.
- Sani, B., 1995. *Periodic inventory control systems and demand forecasting methods for low demand items*. Unpublished PhD thesis, Lancaster University, U.K.
- Silver, E.A., Pyke, D.F., and Peterson, R., 1998. *Inventory management and production planning and scheduling*. 3rd ed. New York: Wiley.
- Slakter, M.J., 1965. A comparison of the Pearson chi-square and Kolmogorov goodness-of-fit tests with respect to validity. *Journal of the American Statistical Association*, 60, 854–858.
- Strijbosch, L.W.G., Heuts, R.M.J., and van der Schoot, E.H.M., 2000. A combined forecast-inventory control procedure for spare parts. *Journal of the Operational Research Society*, 51, 1184–1192.
- Syntetos, A.A., et al., 2009a. Periodic control of intermittent demand items: theory and empirical analysis. *Journal of the Operational Research Society*, 60, 611–618.
- Syntetos, A.A. and Boylan, J.E., 2005. The accuracy of intermittent demand estimates. *International Journal of Forecasting*, 21, 303–314.
- Syntetos, A.A. and Boylan, J.E., 2008. Smoothing and adjustments of demand forecasts for inventory control. *I.M.A Journal of Management Mathematics*, 19, 175–192.
- Syntetos, A.A., Boylan, J.E., and Croston, J.D., 2005. On the categorisation of demand patterns. *Journal of the Operational Research Society*, 56, 495–503.
- Syntetos, A.A., Boylan, J.E., and Disney, S.M., 2009b. Forecasting for inventory planning: a 50-year review. *Journal of the Operational Research Society*, 60, S149–S160.
- Syntetos, A.A., Keyes, M., and Babai, M.Z., 2009c. Demand categorisation in a European spare parts logistics network. *International Journal of Operations and Production Management*, 29, 292–316.
- Tapiero, C., 2005. Value at risk and inventory control. *European Journal of Operational Research*, 163, 769–775.
- Taylor, J.W., 2007. Forecasting daily supermarket sales using exponentially weighted quantile regression. *European Journal of Operational Research*, 178, 154–167.
- Teunter, R., Syntetos, A.A., and Babai, M.Z., 2010. Determining order-up-to levels under periodic review for compound binomial (intermittent) demand. *European Journal of Operational Research*, 203, 619–624.
- Vereecke, A. and Verstraeten, P., 1994. An inventory management model for an inventory consisting of lumpy items, slow movers and fast movers. *International Journal of Production Economics*, 35, 379–389.
- Walsh, J.E., 1963. Bounded probability properties of Kolmogorov–Smirnov and similar statistics for discrete data. *Annals of the Institute of Statistical Mathematics*, 15, 153–158.
- Ward, J.B., 1978. Determining re-order points when demand is lumpy. *Management Science*, 24, 623–632.
- Watson, R.B., 1987. The effects of demand-forecast fluctuations on customer service and inventory cost when demand is lumpy. *Journal of the Operational Research Society*, 38, 75–82.
- Willemain, T.R., Smart, C.N., and Schwarz, H.F., 2004. A new approach to forecasting intermittent demand for service parts inventories. *International Journal of Forecasting*, 20, 375–387.
- Willemain, T.R., et al., 1994. Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. *International Journal of Forecasting*, 10, 529–538.
- Williams, T.M., 1984. Stock control with sporadic and slow-moving demand. *Journal of the Operational Research Society*, 35, 939–948.

Appendix A. Demand-per-period goodness-of-fit results (datasets no. 2, no. 3)

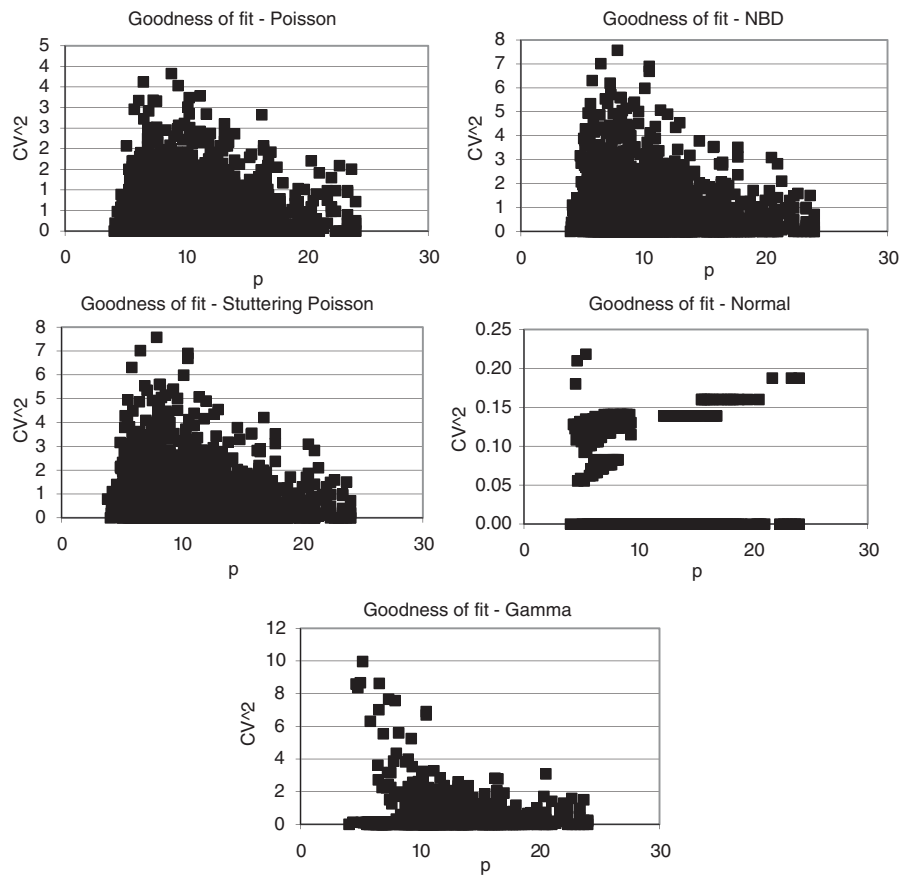


Figure A1. Dataset no. 2: goodness-of-fit results.

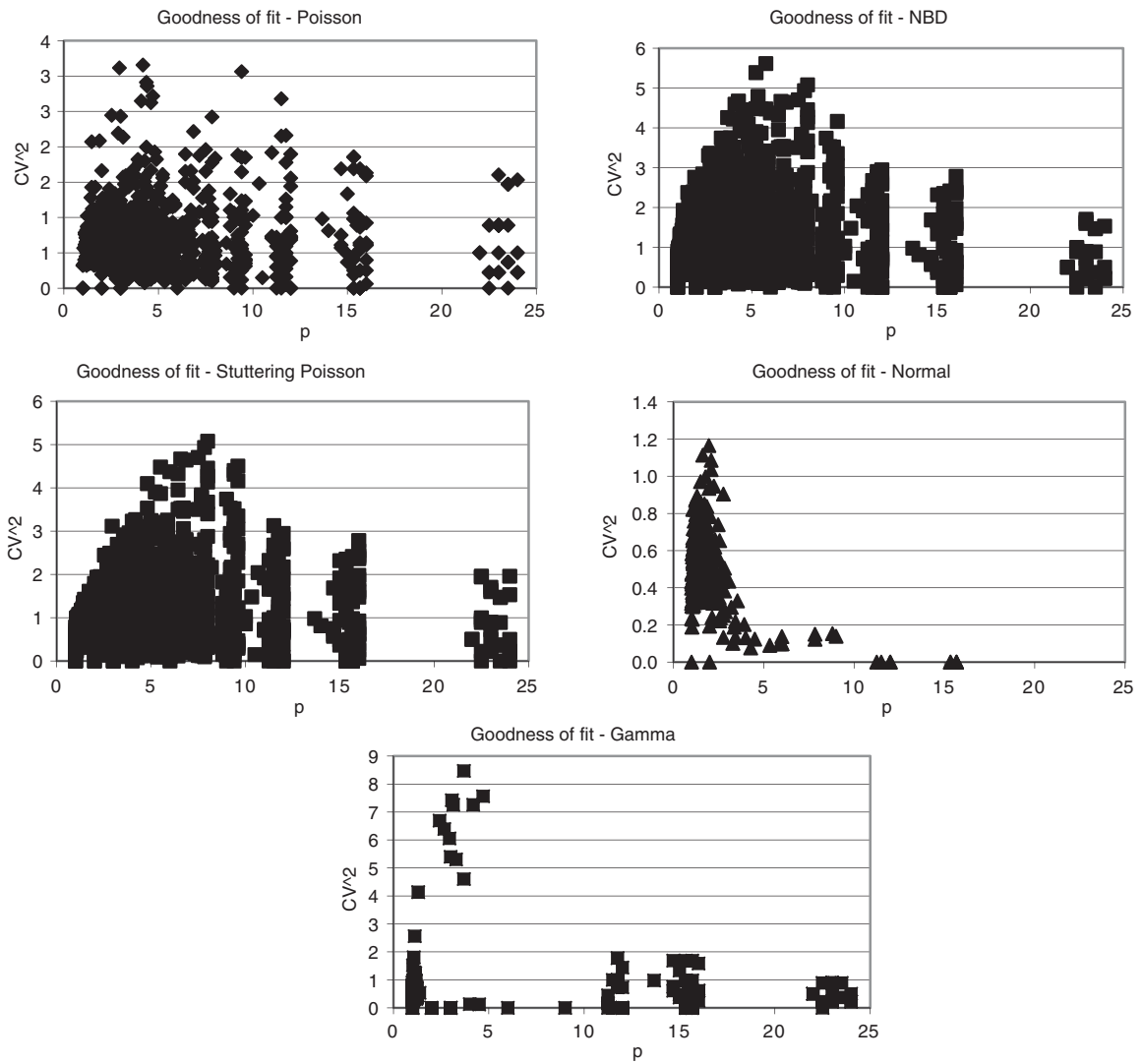


Figure A2. Dataset no. 3: goodness-of-fit results.