



Robust compound Poisson parameter estimation for inventory control[☆]

Dennis Prak^{a,b,*}, Ruud Teunter^a, Mohamed Zied Babai^c, John E. Boylan^d, Aris Syntetos^e

^a Department of Operations, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands

^b Department Industrial Engineering and Business Information Systems, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

^c Kedge Business School, 680 Cours de la Libération, 33405 Talence Cedex, France

^d Department of Management Science, Lancaster University, Lancaster LA1 4YX, United Kingdom

^e PARC Institute of Manufacturing, Logistics and Inventory, Cardiff Business School, Cardiff University, Cardiff CF10 3EU, United Kingdom

ARTICLE INFO

Article history:

Received 12 June 2020

Accepted 4 May 2021

Available online 6 May 2021

Keywords:

Inventory control

Forecasting

Probability

Compound Poisson demand

ABSTRACT

Most companies store demand data periodically and make periodic demand forecasts, whereas many demand processes in inventory control need parameter estimates at the individual customer level. Guidance on estimating the parameters of a continuous-time demand process from period demand data is lacking, in particular for the popular and well-studied compound Poisson class of demand. Whereas the statistics literature typically focuses on asymptotic properties, parameters for inventory control have to be estimated based on a limited number of periodic historical demand observations. We show that the standard Method-of-Moments (MM) estimator – the default choice in applied inventory control research – is severely biased for finite samples. The Maximum Likelihood (ML) estimator – which needs to be obtained by a numerical search – performs better, but both estimators lack robustness to misspecification of the demand size distribution. We propose an intuitive, consistent, closed-form MM alternative that dominates standard MM and ML in terms of estimation accuracy and on-target inventory performance. Its closed form does not depend on the specific demand size distribution, making it robust and easily applicable in large-scale applications with many items. In a case study, we find that the accuracy loss due to storing demand periodically is four times as high under standard MM as under the proposed estimator.

© 2021 The Author(s). Published by Elsevier Ltd.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

The compound Poisson demand process is one of the most general and widely studied demand processes for inventory control. It has a natural business interpretation with customers arriving in continuous time, and their demand sizes following any general compounding distribution, and is therefore a standard choice in textbooks (e.g. [1–3]) and commercial software (e.g. Forecast Pro, Oracle NetSuite, SAP APO, SAS, and Slimstock), especially for intermittent demand patterns. After Galliher et al. [4] studied the special case of Poisson-geometric demand, Feeney and Sherbrooke [5] and Chen et al. [6] derived order-up-to policies under general compound Poisson demand. Archibald and Silver [7] extended this work to general (s, S) policies. Order-up-to policies can be deter-

mined efficiently under compound Poisson demand, as shown by Sherbrooke [8] and Graves [9], who provided approximate procedures, and Axsäter [10], who presented an exact method. Several empirical studies concluded that, especially if products are not fast-moving, compound Poisson distributions provide the best fit (e.g. [11–13]). Therefore, they remain a standard choice in inventory control research (e.g. [14–16]).

In order to exactly characterize inventory policies, a complete specification of the compound Poisson demand process is required. This is most obvious for continuous-review inventory systems. Here, the arrival of a customer immediately triggers an order, and the probability distribution of individual customer demand sizes determines the expected reorder level undershoot. However, also under periodic review all parameters of the compound Poisson demand process are needed for exact calculations [17]. As compound Poisson demand processes operate in continuous time, it would be natural to estimate their parameters from individual transaction data. The average number of customer arrivals in a period is an unbiased estimator for the arrival rate and the parameter(s) of the compounding distribution can be estimated directly from the demand sizes. However, it is common in practice (and a typical assumption in applied inventory control literature [12,13,18]) that

[☆] Area - Supply Chain Management This manuscript was processed by Associate Editor Keskin.

* Corresponding author. Present address: Department Industrial Engineering and Business Information Systems, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

E-mail addresses: d.r.j.prak@utwente.nl (D. Prak), r.h.teunter@rug.nl (R. Teunter), mohamed-zied.babai@kedgebs.com (M.Z. Babai), j.boylan@lancaster.ac.uk (J.E. Boylan), SyntetosA@cardiff.ac.uk (A. Syntetos).

only total demand per period is stored and used as a basis for forecasting. In fact, temporal data aggregation to improve accuracy is frequently suggested (see e.g. [19–21]). As a result, commercial forecasting software is mainly based on period demand. Storing demand per period rather than per individual transaction does lead to an accuracy loss. As we show in this paper, both the choice of the parameter estimator and the period length to which demands are aggregated have a large influence on the magnitude of the error.

Separate forecasting of demand sizes and demand inter-arrival times may be achieved by so-called Size-Interval methods. The most well-known one – twice applying exponential smoothing – was first proposed by Croston [22]. A survey of software packages' forecasting capabilities [23] confirms the popularity of Croston's method. Continuations of this research stream are reviewed by Syntetos et al. [24]. However, period forecasting methods estimate the mean demand size in a period and the mean interval between periods with demand occurrence, rather than the mean demand size of an individual customer and mean interval between individual customer arrivals. Therefore, they cannot be directly used in inventory control, especially not if demand is intermittent [25,26]. Ignoring this mismatch leads even asymptotically to significantly biased estimates and dramatically underachieved service levels, as we will show in Section 2.

Very few authors in the inventory control literature discussed how the unknown parameters of compound Poisson demand processes (or of demand processes in general) can be obtained from period demand data. The authors who did, suggested the use of standard Method-of-Moments (MM) estimators [2,27], which are therefore at present the standard choice in applied work [13,18]. In the wider (statistics) literature, the discussion of parameter estimation centers around (variants of) MM and Maximum Likelihood (ML). Examples are Anscombe [28] and Sichel [29] (logarithmic compounding distribution), Öztürk [30] (exponential compounding distribution), Withers and Nadarajah [31] (gamma compounding distribution), and Balakrishnan et al. [32] (geometric compounding distribution). MM and ML variants are also given for certain compound Poisson distributions in textbooks such as Johnson et al. [33], without a performance comparison. If authors compare properties of different estimators, then the discussion is almost always limited to large-sample properties. In this paper we focus on finite-sample performance instead, which is important in practice. Product life-cycles are shortening, and even if historic demand data are available for multiple years, then not all of them may be indicative of future demand.

Our main contribution is the presentation of an alternative MM estimator that retains Croston's core idea of separately estimating customer arrival rates and demand size means. It bases the customer arrival rate on the fraction of periods without demand, and then uses central moments to estimate the parameters of the (individual) demand size distribution. This estimator has an intuitive, closed-form solution, independent of the specific compounding distribution. It is furthermore robust to misspecification of the compounding distribution, which is an important practical advantage. The idea of estimating distribution parameters based on the 'zero frequency' has sporadically been applied [12,28,34–36]. However, i) a general method to estimate any compound Poisson demand parameters based on this logic has never been given, and ii) a study of such a method's performance is lacking.

Focusing on the exponential compounding distribution in the main body of the paper and discussing the geometric compounding distribution in Appendix B, we compare the finite-sample bias and variance of standard MM, ML, and the proposed alternative MM estimator. Thereafter, we compare the achieved fill rates in a continuous review order-up-to system. We furthermore study all estimators under misspecification of the demand size distribution

and find that both standard MM and ML lack robustness, whereas the proposed method is unaffected by the specific demand size distribution. A case study validates the results and quantifies the accuracy loss due to storing demand per period rather than per transaction.

The remainder of this paper is organized as follows. Section 2 introduces the inventory model, demand model, fill rate service measure, and discusses the misapplication of period Size-Interval methods. Section 3 introduces the consistent standard MM and ML estimator and the proposed estimator. Section 4 provides a comparison of the consistent estimators' biases, whereas Section 5 discusses their variances. Section 6 presents a comparison of achieved fill rates in an order-up-to inventory model, and Section 7 gives robustness results. A comparison of the estimators in a real case study and a quantification of the accuracy loss due to storing aggregated demands per period rather than individual transactions is given in Section 8. Section 9 concludes. Appendix A contains derivations related to Section 4, and Appendix B presents results for the geometric compounding distribution.

2. Demand model, inventory model, and size-interval methods

In this section we specify the general compound Poisson demand model and the compounding distributions on which we base our results. Thereafter we discuss the order-up-to inventory model and the fill rate calculation. Finally, we discuss the consequences of misusing period Size-Interval estimates as compound Poisson demand parameters.

2.1. Demand model

We consider a model where demand follows a compound Poisson process with some general (discrete or continuous) compounding distribution. We have a sample of n period demand realizations x_1, \dots, x_n . That is, for $t = 1, 2, \dots, n$, we observe realizations of the random variable $X_t = \sum_{i=1}^K D_i$, where K follows a Poisson distribution with mean λ and, independently, the individual customer demands D_i are i.i.d. with mean μ_c . The individual customer demands are not observed. We require that $P(D_i \leq 0) = 0$ as a customer can naturally only have positive demand. The set x_1, \dots, x_n of period demand realizations is used to estimate the parameters of the compound Poisson process.

In the remainder of the paper we focus on the exponential compounding distribution which is continuous and therefore more suitable for demonstration, since any fill rate level can be achieved exactly. We refer to Appendix B for additional results based on the geometric compounding distribution. For the exponential distribution, the individual demand sizes D_i have the probability density function

$$f_E(x) = \theta \exp(-\theta x),$$

for $x \geq 0$ and $\theta > 0$, with demand size mean $\mu_c = 1/\theta$.

For the geometric distribution, the individual demand sizes D_i have the (discrete) probability mass function

$$f_G(x) = (1 - \beta)\beta^{x-1},$$

for $x = 1, 2, \dots$ and $0 < \beta < 1$, with demand size mean $\mu_c = 1/(1 - \beta)$.

2.2. Inventory model

We study a continuous review order-up-to inventory control model with a positive fixed lead time of L periods and full back-ordering, under a fill rate constraint. Order-up-to inventory models

are relatively easy to optimize, popular in practice, and theoretically optimal for various cost and service models when order costs are negligible [1,37]. As a service target we use the fill rate, which is the expected fraction of a demand that can be serviced immediately from on-hand inventory.

We follow the computation of the fill rate for compound Poisson demand models by Axsäter [2], and generalize it to handle also continuous compounding distributions. Specifically, denoting an individual demand by D_i and the inventory level by IL , for a given order-up-to level S , the fill rate is the expected fraction of a single order that can be fulfilled from on-hand stock,

$$FR(S) = \frac{E_{IL}[E_{D_i}[\min(\max(IL, 0), D_i)]]}{\mu_c}$$

The distribution of D_i is the selected compounding distribution with mean μ_c , whereas the distribution of IL follows from the observation that the inventory level at time $t + L$ is equal to S minus the lead time demand. Since demand per period is compound Poisson with rate λ and demand size mean μ_c , the demand during the lead time follows a compound Poisson distribution with the same compounding distribution and rate λL . Denoting lead time demand by D_L , we find the relationship $P(IL \leq x) = P(D_L \geq S - x)$. So, $F_{IL}(x) = 1 - F_{D_L}(S - x)$, where F_{IL} denotes the distribution function of the inventory level and F_{D_L} denotes the distribution function of the compound Poisson distribution of the demand process, with arrival rate λL and demand size mean μ_c . The objective is

$$\begin{aligned} &\text{minimize } S \\ &\text{subject to } FR(S) \geq \gamma, \end{aligned}$$

where γ is the target fill rate. In the remainder of this paper, we use $\gamma = 0.95$ (i.e. 95%).

2.3. Misapplication of period size-interval methods

Period Size-Interval methods yield estimates of the time between two periods with positive demand and the size of a non-zero period demand. We demonstrate here the consequences of misapplying these as compound Poisson parameter estimators. We first discuss Croston's method and a 'bias-corrected' version of this method. Thereafter we discuss an intuitive method that uses the average of all non-zero period demands and the average of all times between periods with positive demand, which we refer to as 'Unweighted Averaging'. All discussed methods only update after a period with positive demand.

Croston's method estimates in period n the time between two periods with positive demand (denoted by \hat{p}_n) and the size of a non-zero period demand (denoted by \hat{s}_n) both by exponential smoothing. Formally, given that we have observed n periods, the n th period has a positive demand x_n , and the previous positive demand was q_n periods ago,

$$\begin{aligned} \hat{s}_n &= \alpha_1 x_n + (1 - \alpha_1) \hat{s}_{n-1}, \\ \hat{p}_n &= \alpha_2 q_n + (1 - \alpha_2) \hat{p}_{n-1}, \end{aligned}$$

where $0 < \alpha_1 < 1$ and $0 < \alpha_2 < 1$ are constants.

If Croston's period estimates are wrongly taken as the estimates of the individual demand parameters, then \hat{s}_n is the estimate after n period observations for the mean μ_c of the compounding distribution, whereas $1/\hat{p}_n$ is the estimate for λ . As shown by Syntetos and Boylan [38], using $1/\hat{p}_n$ as an estimate for the arrival rate of a compound Poisson process leads to an inversion bias, as $E[1/\hat{p}_n] \neq 1/E[\hat{p}_n]$. This inversion bias does not diminish as the sample size goes to infinity, as Croston's method uses exponential smoothing and thus does not use all data points, so that the Central Limit Theorem does not hold. Syntetos and Boylan [39] approximately corrected for this bias using a second-order Taylor expansion, and derived the estimator $(1 - \alpha_2/2)/\hat{p}_n$ for λ . The esti-

mator for μ_c does not suffer from the inversion bias and remains unaltered.

It is straightforward to show that, with or without the inversion bias correction, Croston's method yields inconsistent customer demand parameter estimates when applied to period demand. Consider \hat{s}_n , which is unaffected by the inversion bias. Instead of estimating the size of an individual demand, it estimates the size of a positive period demand. Therefore,

$$\lim_{n \rightarrow \infty} E[\hat{s}_n] = E(X_n | X_n > 0) = \frac{\mu_c \sum_{k=1}^{\infty} k \frac{\exp(-\lambda) \lambda^k}{k!}}{1 - \exp(-\lambda)} = \frac{\mu_c \lambda}{1 - \exp(-\lambda)},$$

which is strictly larger than μ_c and smaller than $\mu_c \lambda$ for all $\lambda > 0$ and $\mu_c > 0$. This limit approaches μ_c as $\lambda \rightarrow 0$, and $\mu_c \lambda$ as $\lambda \rightarrow \infty$. The estimator's limit (and thus its bias) is proportional to μ_c .

The asymptotic bias of the arrival rate estimate is more cumbersome to quantify. Instead of estimating the time between two customer arrivals, \hat{p}_n estimates the number of periods between two periods with positive demand. This number of periods (denoted by q) follows a geometric distribution with success probability $1 - \exp(-\lambda)$, mean $1/(1 - \exp(-\lambda))$, and variance $\exp(-\lambda)/(1 - \exp(-\lambda))^2$. Although $E[\hat{p}_n] \rightarrow 1/(1 - \exp(-\lambda))$ as $n \rightarrow \infty$, we can only approximate the limit of $E[1/\hat{p}_n]$. We do so by a second-order Taylor expansion and find, using the mean and variance of the geometric distribution and the fact that \hat{p}_n is an exponential smoothing estimator with asymptotic variance $\text{Var}(q)\alpha_2/(2 - \alpha_2)$, that

$$\begin{aligned} \lim_{n \rightarrow \infty} E[1/\hat{p}_n] &\approx \frac{1}{E[q]} + \frac{\alpha_2}{2 - \alpha_2} \frac{\text{Var}(q)}{E[q]^3} \\ &= \left(1 + \frac{\alpha_2}{2 - \alpha_2} \exp(-\lambda)\right) (1 - \exp(-\lambda)). \end{aligned}$$

As $\alpha_2 \rightarrow 0$, this limit approaches $1 - \exp(-\lambda)$, which is the result without inversion bias. In that scenario, λ is strictly underestimated for all $\lambda > 0$. The approximate limit of the estimator converges to 0 as $\lambda \rightarrow 0$, and to 1 as $\lambda \rightarrow \infty$. Obviously, the bias correction proposed by Syntetos and Boylan [39] scales the estimated arrival rate, and thus its expectation, by $(1 - \alpha_2/2)$. In conclusion, Croston's method generally underestimates the arrival rate, but this is somewhat counteracted by its inversion bias, leading to an overestimation especially for low values of λ and/or high values of α_2 . However, (approximately) correcting for that inversion bias intensifies the underestimation.

While Croston's method and its variants give greater weight to the most recent demand sizes and intervals, this is not a requirement for Size-Interval methods. An alternative approach is to use equal weights, but retain the separate estimation of demand size means and intervals. The Unweighted Averaging method does this by taking as the demand size mean the average of all positive period demand sizes, and as arrival rate the inverse of the average number of periods between two periods with positive demands. Given that we have observed n periods, of which n_p had a positive demand, the n -th period has a positive demand x_n , and we have observed the previous positive demand q_n periods ago, Unweighted Averaging sets

$$\begin{aligned} \hat{s}_n &= \frac{1}{n_p} x_n + \frac{n_p - 1}{n_p} \hat{s}_{n-1}, \\ \hat{p}_n &= \frac{1}{n_p} q_n + \frac{n_p - 1}{n_p} \hat{p}_{n-1}. \end{aligned}$$

It is straightforward to observe that, just as with both Croston variants,

$$\lim_{n \rightarrow \infty} E[\hat{s}_n] = E(X_n | X_n > 0) = \frac{\mu_c \lambda}{1 - \exp(-\lambda)}.$$

Unlike Croston's method, the Unweighted Averaging method is asymptotically unbiased, since it uses all positive demands and

Table 1
Asymptotic Biases and Achieved Fill Rates when Using Period Size-Interval Methods for Compound Poisson Parameter Estimation (Target Fill Rate 95%, Lead Time $L = 2$).

True Parameters			Asymptotic Bias				Achieved Fill Rates (Compounding Distr.)					
							(Exponential)			(Geometric)		
λ	μ_c	α_2	λ CROST	λ SBA	λ UA	μ_c	Standard	SBA	UA	Standard	SBA	UA
1/16	2	0.1	+1.7%	-3.4%	-3.1%	+3.2%	95.5%	95.4%	95.4%	97.2%	97.2%	97.2%
1/16	2	0.3	+13.1%	-3.9%	-3.1%	+3.2%	95.6%	95.4%	95.4%	97.2%	97.2%	97.2%
1/16	2	0.5	+27.3%	-4.5%	-3.1%	+3.2%	95.8%	95.4%	95.4%	97.2%	97.2%	97.2%
1/16	5	0.1	+1.7%	-3.4%	-3.1%	+3.2%	95.5%	95.4%	95.4%	95.7%	95.7%	95.7%
1/16	5	0.3	+13.1%	-3.9%	-3.1%	+3.2%	95.6%	95.4%	95.4%	96.5%	95.7%	95.7%
1/16	5	0.5	+27.3%	-4.5%	-3.1%	+3.2%	95.8%	95.4%	95.4%	96.5%	95.7%	95.7%
1/4	2	0.1	-7.9%	-12.5%	-11.5%	+13.0%	96.5%	96.3%	96.3%	97.0%	97.0%	97.0%
1/4	2	0.3	+0.6%	-14.5%	-11.5%	+13.0%	96.8%	96.2%	96.3%	98.2%	97.0%	97.0%
1/4	2	0.5	+11.4%	-16.4%	-11.5%	+13.0%	97.1%	96.1%	96.3%	98.2%	97.0%	97.0%
1/4	5	0.1	-7.9%	-12.5%	-11.5%	+13.0%	96.5%	96.3%	96.3%	97.0%	96.4%	97.0%
1/4	5	0.3	+0.6%	-14.5%	-11.5%	+13.0%	96.8%	96.2%	96.3%	97.0%	96.4%	97.0%
1/4	5	0.5	+11.4%	-16.4%	-11.5%	+13.0%	97.1%	96.1%	96.3%	97.4%	96.4%	97.0%
1	2	0.1	-35.6%	-38.8%	-36.8%	+58.2%	98.6%	98.4%	98.5%	99.3%	99.0%	99.3%
1	2	0.3	-32.7%	-42.8%	-36.8%	+58.2%	98.7%	98.1%	98.5%	99.3%	99.0%	99.3%
1	2	0.5	-29.0%	-46.8%	-36.8%	+58.2%	98.9%	97.8%	98.5%	99.3%	98.5%	99.3%
1	5	0.1	-35.6%	-38.8%	-36.8%	+58.2%	98.6%	98.4%	98.5%	98.8%	98.6%	98.8%
1	5	0.3	-32.7%	-42.8%	-36.8%	+58.2%	98.7%	98.1%	98.5%	98.9%	98.4%	98.8%
1	5	0.5	-29.0%	-46.8%	-36.8%	+58.2%	98.9%	97.8%	98.5%	99.1%	98.2%	98.8%

inter-arrival times in the data set rather than applying diminishing weights via exponential smoothing. Therefore, the Central Limit Theorem holds and \hat{p}_n converges in probability to $E[\hat{p}_n]$, so that

$$\lim_{n \rightarrow \infty} E[\hat{\lambda}_n] = \lim_{n \rightarrow \infty} E[1/\hat{p}_n] = 1 - \exp(-\lambda).$$

The Unweighted Averaging estimator for λ converges to 0 as $\lambda \rightarrow 0$, and to 1 as $\lambda \rightarrow \infty$. This reveals the flaw in using period demand estimators as individual demand parameters: no distinction can be made between 1 or several demands in a period, leading to an underestimated arrival rate. All individual demands in a period are added and considered as a single customer demand, leading to an overestimated demand size mean.

Table 1 quantifies the (approximate) asymptotic bias resulting from using standard Croston’s method (‘CROST’), the corrected version by Syntetos and Boylan [39] (Syntetos-Boylan Approximation, ‘SBA’), or the Unweighted Averaging method (‘UA’), for several parameter choices. Furthermore, we report the fill rate that would be achieved for both the exponential and geometric compounding distribution, under the true demand parameters, when a target fill rate of 95% is chosen. We use a lead time of 2 periods.

None of the results depend on μ_c , except for the achieved fill rates under the geometric compounding distribution. The reason is that for discrete demand distributions, fill rates cannot be achieved exactly. For low arrival rates and/or high smoothing parameter values, Croston’s inversion bias results in an overall overestimation of λ . The Syntetos-Boylan Approximation achieves an asymptotic underestimation for all scenarios. The Unweighted Averaging method uniformly underestimates λ . The demand size mean μ_c is always overestimated for all methods, and all (absolute) percentage asymptotic biases are increasing in λ .

Interestingly, all achieved fill rates are asymptotically too high, irrespective of whether λ is over- or underestimated. This is because the achieved fill rate depends more strongly on μ_c , as we will further discuss in Section 6. The achieved fill rates are increasing in λ , and for $\lambda = 1$ they are over 97.8% for all compounding distributions. We conclude that misapplying period Size-Interval methods to estimate compound Poisson demand parameters leads to severely overestimated demand size means, significantly over-shot fill rates and correspondingly excessive inventory costs.

3. Consistent estimators

Having discussed the misapplication of period forecasting methods for obtaining compound Poisson demand parameters, we move to consistent estimation methods. We revisit the compound Poisson demand model introduced in the previous section and discuss the standard MM estimator (applied in the inventory control literature), the ML estimator, and the proposed alternative MM estimator - first for general compound Poisson demand processes and then for the specific case of an exponential compounding distribution.

We first give some properties of the Poisson-exponential distribution that are helpful for our analysis [30]. Its probability density function is given by

$$f_{PE}(x) = \exp(-\lambda - x/\mu_c) \sum_{i=1}^{\infty} \left(\frac{(\lambda/\mu_c)^i x^{i-1}}{i!(i-1)!} \right),$$

for $\lambda > 0$, $\mu_c > 0$, and $x > 0$. This distribution has a positive probability mass of $\exp(-\lambda)$ at $x = 0$. The first four central moments are $\mu_1 = \lambda\mu_c$, $\mu_2 = 2\lambda\mu_c^2$, $\mu_3 = 6\lambda\mu_c^3$, and $\mu_4 = (24\lambda + 12\lambda^2)\mu_c^4$. We assume that a sample x_1, \dots, x_n of period demands is available and denote the i th central sample moment of the period demand observations by $\hat{\mu}_i$. In particular, $\hat{\mu}_1$ denotes the sample mean and $\hat{\mu}_2$ the sample variance.

3.1. The standard method-of-moments estimator

Standard MM estimators equate central moments of the compound Poisson demand distribution to their corresponding sample equivalents. That is, if the compounding distribution has P parameters, we solve

$$\begin{aligned} \mu_1 &= \hat{\mu}_1, \\ &\vdots \\ \mu_{P+1} &= \hat{\mu}_{P+1}, \end{aligned}$$

for the unknown arrival rate and the P parameters of the compounding distribution.

An exponential compounding distribution has one parameter, which we take to be its mean. So, there are in total two parameters to be estimated. The standard MM estimators for the arrival

rate and the demand size mean are then

$$\hat{\lambda} = \frac{2\hat{\mu}_1^2}{\hat{\mu}_2} \text{ and } \hat{\mu}_c = \frac{\hat{\mu}_2}{2\hat{\mu}_1}.$$

MM estimators are guaranteed to be consistent, but may still have severe finite-sample biases and variances, as our numerical results will reveal. Furthermore, the second order sample moment (the variance of the compound Poisson distribution) depends on the exact shape of the compounding distribution and not only on its mean. This implies that standard MM is not robust to misspecified demand size distributions.

3.2. The maximum likelihood estimator

The ML estimator of a compound Poisson distribution is found by maximizing the log-likelihood function with respect to the unknown parameters. In most cases, the maximizing arguments cannot be found in closed form, and a numerical search procedure has to be invoked. The log-likelihood function is defined as

$$\mathcal{L}(\theta|x_1, \dots, x_n) = \sum_{t=1}^n \ln(f_{CP}(x_t|\theta)),$$

where f_{CP} denotes the probability density function of the corresponding (full) compound Poisson distribution and θ denotes its vector of parameters, including the arrival rate of the Poisson process and the parameters of the compounding distribution.

For the exponential distribution, we have $f_{CP} = f_{PE}$ and $\theta = (\lambda, \mu_c)$. One can maximize $\mathcal{L}(\lambda, \mu_c|x_1, \dots, x_n)$ to find estimates $\hat{\lambda}$ and $\hat{\mu}_c$, using a non-linear maximization algorithm. Although ML is computationally more demanding than the other methods, it is not prohibitively expensive, as it involves only a two-dimensional search. Moreover, it is guaranteed to provide a consistent estimator that asymptotically has the lowest variance. However, ML offers no guaranteed optimal performance for finite samples.

3.3. The proposed method

We propose an alternative MM estimator and choose its moments such that the intuition of the Size-Interval forecasting literature – separate estimation of the customer arrival rate and the demand size mean – is employed. Also this estimator is guaranteed to be consistent, as it is an MM estimator. We first discuss estimation of the arrival rate λ . Denote by N_0 the (stochastic) number of periods in the sample of size n in which no demand occurred. The corresponding realization of that number of periods is denoted by n_0 . The occurrence of no demand in a period is Bernoulli distributed with probability $\exp(-\lambda)$, and therefore the number of periods without demand out of n total periods follows a binomial distribution with parameters n and $\exp(-\lambda)$. Therefore

$$E[N_0] = n \exp(-\lambda).$$

We equate this sample moment to its expectation to find the estimator for the arrival rate:

$$n \exp(-\hat{\lambda}) = n_0,$$

leading to

$$\hat{\lambda} = -\ln(n_0/n).$$

For each of the parameters of the compounding distribution, we equate (as with standard MM) central moments of the compound Poisson distribution to their sample equivalents. However, the explicit use of the fraction of periods without demand to estimate the arrival rate implies that the highest order of moments used by

this method is always one lower than for the standard MM approach. That is, if the compounding distribution has P unknown parameters, we solve

$$\begin{aligned} \hat{\lambda} &= -\ln(n_0/n), \\ \mu_1 &= \hat{\mu}_1, \\ &\vdots \\ \mu_P &= \hat{\mu}_P. \end{aligned}$$

In the case of an exponential compounding distribution (more generally, in the case of any compounding distribution with a single parameter), the set of equations reduces to

$$\begin{aligned} \hat{\lambda} &= -\ln(n_0/n), \\ \mu_1 &= \hat{\mu}_1. \end{aligned}$$

Using the fact that $\mu_1 = \lambda\mu_c$, we solve for the demand size mean:

$$\hat{\lambda}\hat{\mu}_c = \hat{\mu}_1$$

where $\hat{\mu}_1$ is the sample average of total demand per period. This leads to

$$\hat{\mu}_c = -\frac{\hat{\mu}_1}{\ln(n_0/n)}.$$

To calculate these estimates, one only needs to obtain the fraction of periods without demand and the overall mean demand during all periods. Furthermore, whereas standard MM and ML have to be derived separately for every compounding distribution, this method gives estimates that do not depend on the specific compounding distribution. Irrespective of the compounding distribution, $\hat{\lambda}$ and $\hat{\mu}_c$ have exactly the same closed-form solution. In Section 7 we elaborate on this property, which extends to compounding distributions with an arbitrary number of parameters. Furthermore, we show that this property implies robustness: irrespective of the compounding distribution, the proposed estimators converge to the corresponding true parameters.

4. Bias comparison

In this section we compare the biases of the three estimators discussed in the previous section. Asymptotically, these estimators are all unbiased, but for finite sample sizes this is not the case. As standard MM and the proposed estimator both exist in closed form, we can analytically study their biases, as we will do via a Taylor approximation, and find the conditions under which the proposed estimator is more accurate. Thereafter we perform a numerical study to also compare both estimators with ML.

4.1. Analytical bias comparison of standard MM and the proposed estimator

To analytically compare standard MM and the proposed estimator, we approximate the estimators by a second-order Taylor series, as a function of the sample mean ($\hat{\mu}_1$) and the sample variance ($\hat{\mu}_2$), and then calculate the expected value. For the derivation, we refer to Appendix A. Here we study the results, starting with the arrival rate estimator. For standard MM, we find:

$$E[\hat{\lambda}_{MM}] \approx \lambda + \frac{2\lambda}{n-1} + \frac{2}{n}.$$

For the proposed estimator, we find:

$$E[\hat{\lambda}_{NEW}] \approx \lambda + \frac{\exp(\lambda) - 1}{2n}.$$

Firstly, we observe that both estimators have a positive bias for all values of λ and n . This implies that the true arrival rate is overestimated by both estimators. Secondly, the biases of both estimators only depend on λ and n , and not on μ_c .

Comparing the biases, we find that for $\lambda \leq 2.75$ the new estimator is more accurate than standard MM, for all sample sizes. For $\lambda \geq 3.5$, standard MM is preferred. For $2.75 < \lambda < 3.5$, the new estimator is preferred for smaller samples and standard MM for larger samples. These findings are in line with the expectation that the proposed estimator outperforms standard MM for intermittent demand patterns. Indeed, since $\lambda = 2.75$ corresponds to a 6% probability of zero demand in a period, the new estimator is preferred even for very 'light intermittency'. Put differently, even if only 1 in 17 periods has zero demand, the new estimator still outperforms standard MM. Although both estimators gain accuracy in larger samples, the relative difference remains approximately the same. For $\lambda = 1$ the proposed estimator is approximately 4 times as accurate as standard MM, whereas for $\lambda = 1/4$ this factor has increased to 17.

We continue with the demand size mean estimator. For standard MM, we find:

$$E[\hat{\mu}_{c,MM}] \approx \mu_c \left(1 - \frac{1}{\lambda n}\right).$$

For the proposed estimator, we find:

$$E[\hat{\mu}_{c,NEW}] \approx \mu_c \left(1 - \frac{(\lambda - 2) \exp(\lambda) + \lambda + 2}{2\lambda^2 n}\right).$$

These estimators both have a negative bias for all values of λ , μ_c , and n . The demand size mean μ_c acts solely as a scaling parameter. It follows that for $\lambda < 2$ the new estimator is more accurate than standard MM. The scenario with $\lambda = 2$ corresponds to a 14% probability of zero demand in a period. We see similar dynamics as for the arrival rate estimator. For $\lambda = 1$ the proposed estimator is approximately 7 times as accurate as standard MM, whereas for $\lambda = 1/4$ this factor has increased to 170.

There are many real life settings with moderate to high intermittency. Lengu et al. [12] report that more than 80% of their studied items have an estimated arrival rate below 1, the majority of which are below 1/4. Turrini and Meissner [13] report median demand intervals of 10 and 17 periods, and a maximum of 80 periods. In Section 4.2 we present more insights into the biases of standard MM, ML, and the proposed method by means of a numerical analysis.

4.2. Numerical study of all estimators' biases

In this section we perform a numerical analysis of the biases of standard MM, ML, and the proposed method, under an exponential compounding distribution. We select $\lambda = 1/16$, $1/4$, and 1 , representing high, moderate, and light intermittency, respectively, and present the results for $\mu_c = 2$ only. Experiments with other values and the analysis in Section 4.1 show that μ_c acts only as a scaling parameter. We let the sample size range from $n = 4$ to $n = 200$. Each time we draw 1,000,000 period demand histories from the corresponding compound Poisson distribution. For each draw, we calculate the parameter estimates for λ and μ_c using (i) the standard MM estimator, (ii) the ML estimator, and (iii) the proposed method. We present per parameter combination, per sample size, and per estimator, the average of all obtained estimates.

Two scenarios deserve special attention. The first is the case where only periods without demand are observed. In practice one will not fit any demand distribution then, but in numerical experiments it does occasionally occur. All methods are then undefined and to enable a fair comparison, we set in this case all estimated arrival rate to 0 and do not define demand size mean estimates.

The other extreme scenario is that with no periods without demand. In this unlikely event, we set the proposed method equal to standard MM.

Fig. 1 shows the results, with on the left hand side the estimated arrival rates λ and on the right hand side the estimated demand size means μ_c . Note that in this and all following figures the marker locations have been shifted between estimators to enhance readability. The scenario with $\lambda = 1/16$ (Fig. 1a and b) corresponds to a 94% probability of having zero demand in a period. The standard MM estimator converges the most slowly by far. For $n = 200$, its arrival rate estimate is still 13% too high, whereas its demand size mean estimate is 8% too low. The proposed method and the ML estimator perform almost identically, and are considerably more accurate than standard MM. For $n \geq 50$ the difference with the true parameters is below 1%. As the sample size decreases, the probability of observing no demand at all increases. Especially for low arrival rates, this affects the results.

The scenario with $\lambda = 1/4$ (Fig. 1c and d) corresponds to a 78% probability of having zero demand in a period. Standard MM still shows a significant deviation from the true value for $n = 200$ (5% for λ and 2% for μ_c). The arrival rates converge more slowly than the demand size means, and the proposed estimator actually (slightly) outperforms ML. In Fig. 1e and f, the arrival rate is increased to 1, corresponding to a 37% probability that a period has zero demand. The standard MM estimator still performs much worse than the proposed method and ML, but the gap is smaller than for lower arrival rates. The proposed method now outperforms ML by a larger margin.

Comparing relative accuracies, we find that the ML and proposed arrival rate estimator perform best for low λ , whereas the standard MM arrival rate estimator becomes better when λ increases, although it is still severely outperformed even for $\lambda = 1$. All estimators for the demand size mean are more accurate as λ increases, as fewer periods without demand occur. Directly related to this observation is the question of the optimal period length (e.g. hour, day, week, or month). Different period lengths lead to different levels of intermittency, corresponding to different true values of λ . In Section 6 we study the resulting effect on inventory performance.

An explanation for the very similar performance of ML and the proposed estimator for low arrival rates (here and for the Poisson-geometric distribution in Appendix B) can be found in their definitions. The probability mass of a compound Poisson distribution for periods with no demand reduces to $\exp(-\lambda)$, which is minimized at $\lambda = 0$. For low arrival rates many periods have no demand. The sample log-likelihood function is the sum of the log-densities of all periods and therefore relatively insensitive to (few) nonzero observations. This insensitivity also holds for the proposed arrival rate estimator, as the natural logarithm is relatively flat around 1 (the point where $n_0 = n$). The standard MM arrival rate estimator, however, is more sensitive to (even few) nonzero observations, as it uses the squared mean and the variance of the sample. For higher arrival rates, the demand pattern becomes less intermittent, benefiting standard MM. That is also where the bias of ML differs more from that of the proposed method (e.g. Fig. 1e and f).

The results of this section can be used to validate the accuracy of the analytical bias approximation of Section 4.1. Table 2 presents the comparison for the case $\lambda = 1/4$, $\mu_c = 2$. The proposed estimator's expectation is better approximated by the second-order Taylor series than the standard MM estimator, as would be expected based on their different functional forms. For $n = 10$, the analytical approximation for standard MM differs substantially from the numerical evaluation. There are two reasons for this: 1) the Taylor series approximation is generally more accurate for larger samples, and 2) for small sample sizes the special cases mentioned above occur more frequently. For $n \geq 50$ the difference between the an-

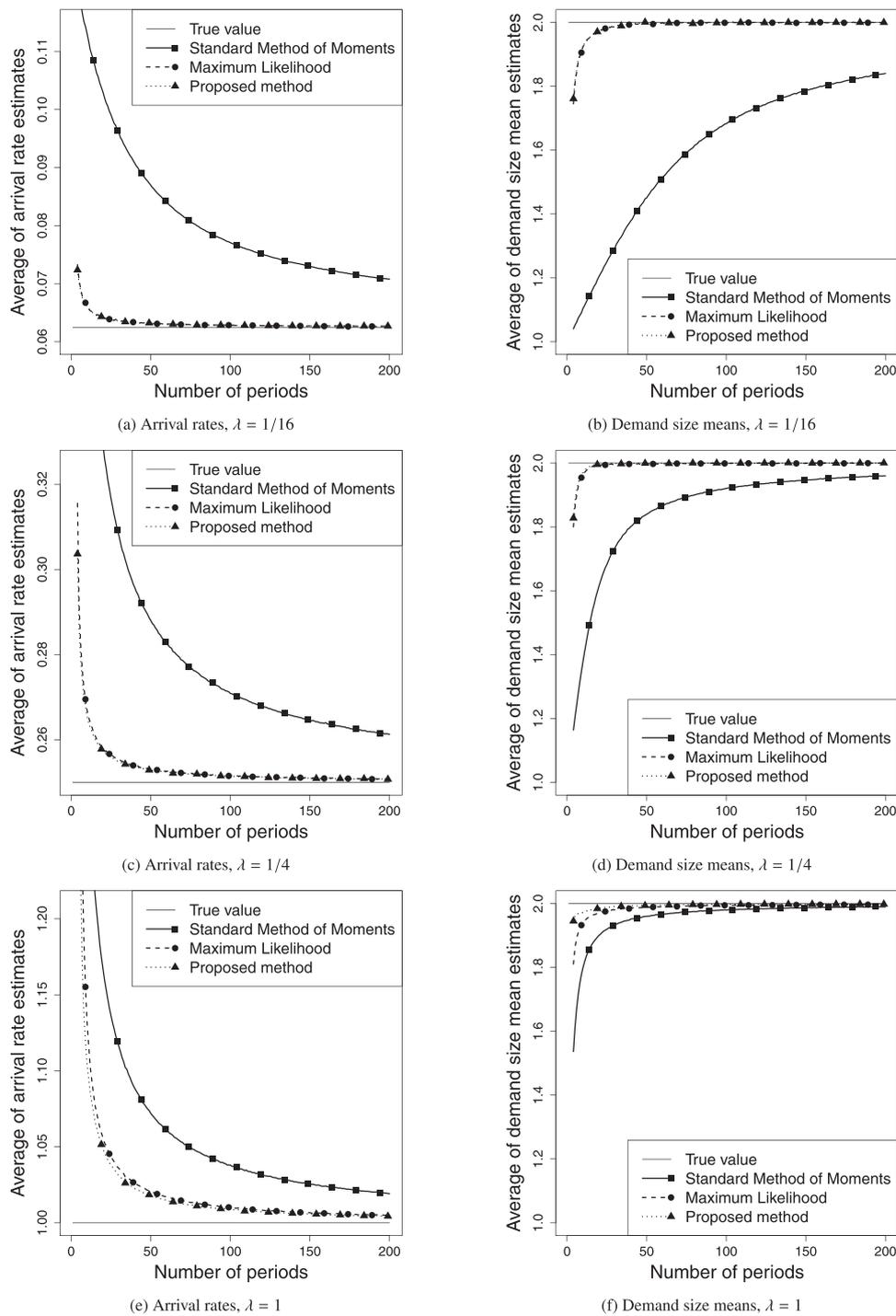


Fig. 1. Sample Averages of Demand Parameter Estimates, $\mu_c = 2$. Vertical Axes Are Truncated for Readability.

alytical and numerical results is small. The proposed estimator is well approximated by the Taylor series for all sample sizes. Furthermore, under both the analytical and numerical evaluation, the methods retain the same ranking.

The proposed method has another important advantage over ML: it has a simple, closed-form solution, whereas optimization of the log-likelihood function for ML involves a two-dimensional numerical search, each time evaluating a function with infinite sums. Although this procedure is not prohibitively expensive, it does require much more computation time than standard MM or the proposed estimator. Table 3 shows the average computation time per

item (for both the arrival rate and demand size mean estimate, excluding the simulation of the demand sample) for the case $\lambda = 1$, $\mu_c = 2$ (executed on a single Intel Xeon 2.5 GHz core). The proposed and standard MM estimator show very similar and fast computation times that mildly increase with the sample size. Contrarily, computation of the ML estimates takes, for $n = 10$, almost 2000 times longer than computing the standard MM or proposed estimates, whereas for $n = 200$ it takes over 14000 times longer. This difference is significant in large-scale applications. Estimating the parameters for 100,000 items each with 200 historical observations

Table 2
Analytical and Numerical Expectation Approximations, $\lambda = 0.25, \mu_c = 2$.

	Approximated expectation			
	$n = 10$	$n = 50$	$n = 100$	$n = 200$
$\hat{\lambda}$, proposed, analytical	0.2642	0.2528	0.2514	0.2507
$\hat{\lambda}$, proposed, numerical	0.2655	0.2529	0.2514	0.2508
$\hat{\lambda}$, standard MM, analytical	0.5056	0.3002	0.2751	0.2625
$\hat{\lambda}$, standard MM, numerical	0.3791	0.2880	0.2709	0.2613
$\hat{\mu}_c$, proposed, analytical	1.9953	1.9991	1.9996	1.9998
$\hat{\mu}_c$, proposed, numerical	1.9701	1.9986	1.9995	1.9997
$\hat{\mu}_c$, standard MM, analytical	1.2000	1.8400	1.9200	1.9600
$\hat{\mu}_c$, standard MM, numerical	1.3784	1.8392	1.9202	1.9596

Table 3
Mean Computation Times per Item for $\lambda = 1, \mu_c = 2$ (Milliseconds).

	Computation time (ms)			
	$n = 10$	$n = 50$	$n = 100$	$n = 200$
Proposed	0.019	0.021	0.023	0.026
Standard MM	0.019	0.021	0.023	0.025
Maximum Likelihood	3.6	10.7	19.7	36.8

takes 61 minutes by ML, and less than 3 seconds by the proposed or standard MM estimator.

5. Variance comparison

In this section we use the same numerical set-up as in the previous section, but report the variances over the draws. Fig. 2 shows the results, with the arrival rates on the left hand side and the demand size means on the right hand side. We firstly observe that the proposed method and ML perform very similarly. This can be explained analogously to their similar biases (see Section 4.2): they differ only in their treatment of non-zero observations, to which both are relatively insensitive compared to the standard MM estimator. Since the variance of ML acts as an asymptotic lower bound and it is practically indistinguishable from the proposed estimator, we conclude that the new estimator shows near-optimal performance.

For the arrival rate estimators, standard MM has a substantially larger variance than the other two methods. As the fraction of non-zero observations increases with λ , so does the difference between standard MM and the other two methods. However, for the demand size mean it can be observed that only from a certain minimum sample size onward, standard MM has a larger variance than the other two methods. This reflects a drawback of our numerical set-up. If only periods without customer arrivals are drawn, the demand size mean is not estimated. The reported averages are therefore biased towards ‘less intermittent’ draws, favoring standard MM. For larger sample sizes (where standard MM does have a significantly larger variance than the other methods) this event virtually never occurs.

6. Inventory control implications

For the order-up-to inventory model described in Section 2.2, we study the service effects when demand parameters are obtained according to the three consistent estimators, again for an exponential compounding distribution. We use a target fill rate of 95% and report the achieved fill rates.

Fig. 3 a shows the results for $\lambda = 1/16, \mu_c = 2$, and $L = 2$. Consistently with the results of Section 4, the ML and the proposed method converge much faster than standard MM. The order-up-to level that is set by any of the three methods is driven by the overestimation of λ (leading to overshooting the order level and fill rate) and the underestimation of μ_c (leading to an undershoot).

For $\lambda = 1/16, \mu_c = 2$, and $L = 2$ the underestimation of μ_c dominates for all three methods, leading to fill rates that converge from below. Although the overestimation of λ is more severe, this parameter affects only the inventory level distribution (via the lead time demand), whereas μ_c affects also the distribution of a single customer’s demand size, which both appear in the fill rate expression. Whereas the proposed method and ML are already within 0.1% from the 95% fill rate target for a sample of 50 observations, standard MM only attains 93.8% even for $n = 200$.

When $\lambda = 1/4$ (see Fig. 3b), the overestimation of λ and underestimation of μ_c by the proposed method and ML almost cancel each other out at the order level calculation, leading to a realized fill rate very close to its target over the entire range of sample sizes. In Fig. 3c ($\lambda = 1$) all methods lead to fill rate overshoots, indicating that the overestimation of λ is dominant. ML performs only marginally better than the proposed method and the differences between the three approaches are smaller than in the previous scenarios. In the last scenario (Fig. 3d), λ is reset to 1/4, but the lead time is quadrupled to 8 periods. Comparing this scenario with that where $\lambda = 1/4$ and $L = 2$, we observe that standard MM actually performs better under a longer lead time. By extending the lead time, more weight is given to the arrival rate, so that the underestimation of μ_c is counteracted.

Summarizing all results, we find that the standard MM estimators show slow convergence of achieved fill rates, especially for highly intermittent demand patterns. ML and the proposed method perform closer to their targets. For low arrival rates, the proposed method, ML, and standard MM all converge to the optimal fill rate from below, whereas for higher arrival rates this pattern shifts to a convergence from above. The analysis in this section shows that – given that demands are stored per period – the best period length for optimal inventory performance is a non-trivial function of the arrival rate, demand size mean, and lead time.

7. Robustness analysis

In real forecasting and inventory control applications one should not only estimate demand parameters, but also select the distributional form. In this section we analyze the robustness of the different methods to misspecification of the demand size distribution. We first analytically demonstrate the robustness of the proposed method and then numerically compare its performance to that of standard MM and ML.

Under the demand model specified in Section 2.1, the proposed estimates for the arrival rate and the demand size mean are $\hat{\lambda} = -\ln(n_0/n)$ and $\hat{\mu}_c = \frac{\hat{\mu}_1}{\hat{\lambda}}$. We exclude, as in Section 4.2, the scenarios where $n_0 = 0$ or $n_0 = n$, as the proposed estimator is then undefined. In all other cases, $\hat{\lambda}$ and $\hat{\mu}_c$ are continuous transformations of n_0/n and $\hat{\mu}_1$. Irrespective of the (discrete or continuous) compounding distribution, the number of periods with zero demand out of n periods in total is binomially distributed with parameters n and $\exp(-\lambda)$ (the probability that no customer arrives in a period). Therefore, $E(n_0/n) = \exp(-\lambda)$. Also, $E(\hat{\mu}_1) = \lambda\mu_c$. Applying the continuous mapping theorem yields that $\hat{\lambda} \rightarrow \lambda$ and $\hat{\mu}_c \rightarrow \mu_c$. In conclusion, the proposed estimators have the same closed form and converge to the true arrival rate and demand size mean for any compounding distribution.

This does not hold for the standard MM estimator. For the parameters of the Poisson-exponential distribution, the standard MM estimator gives $\hat{\lambda} = \frac{2\hat{\mu}_1^2}{\hat{\mu}_2}$ and $\hat{\mu}_c = \frac{\hat{\mu}_2}{2\hat{\mu}_1}$. The second central moment of a Poisson-exponential distribution is $\mu_2 = \lambda(\sigma_c^2 + \mu_c^2)$, where σ_c^2 is the variance of the compounding distribution [40]. Therefore,

$$\hat{\lambda} \rightarrow \frac{2(\lambda\mu_c)^2}{\lambda(\sigma_c^2 + \mu_c^2)}.$$

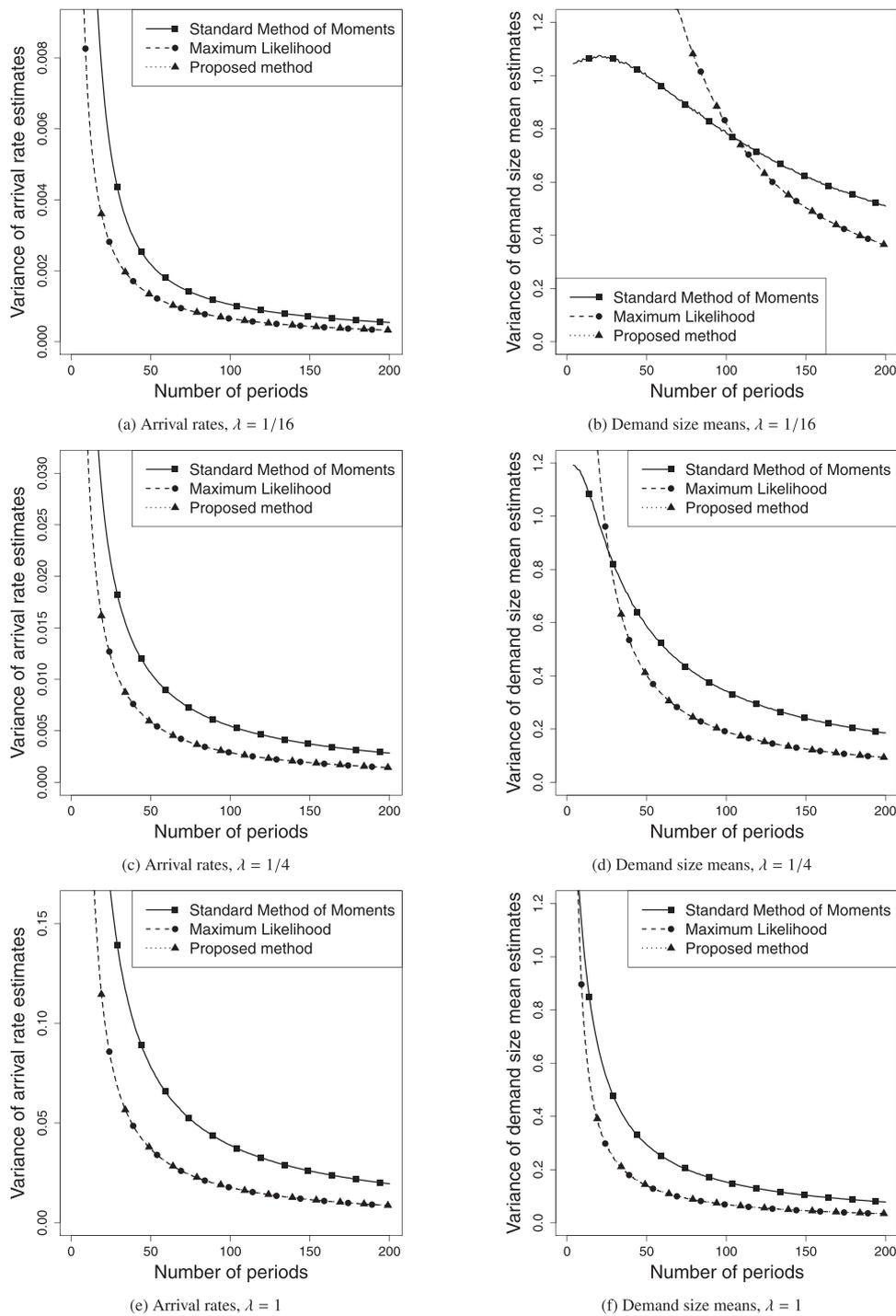


Fig. 2. Sample Variances of Demand Parameter Estimates, $\mu_c = 2$.

Substituting $\mu_c = 1/\lambda$ and $\sigma_c^2 = 1/\lambda^2$ for the exponential distribution, the right hand side simplifies to λ , but if the compounding distribution is not exponential, then this does not hold true generally. Similarly, we can derive that

$$\hat{\mu}_c \rightarrow \frac{\sigma_c^2 + \mu_c^2}{2\mu_c}.$$

For the exponential compounding distribution, the right hand side simplifies to $1/\lambda \equiv \mu_c$, but again this does not hold true generally. We remark that for other compounding distributions, the

asymptotic properties of any estimator can be derived in a similar way. For the asymptotic analysis of the ML estimator we refer to Öztürk [30].

The robustness of the proposed estimator extends to compounding distributions with multiple parameters. The rationale is that it uses P central moments and the probability of zero demand for estimating P parameters of the compounding distribution and the arrival rate. The P th central moment of the full compound Poisson distribution depends only on the first P central moments of the compounding distribution [40]. This is not the case for the standard MM estimator, which always uses $P + 1$ moments for P parameters of the compounding distribution and the arrival rate.

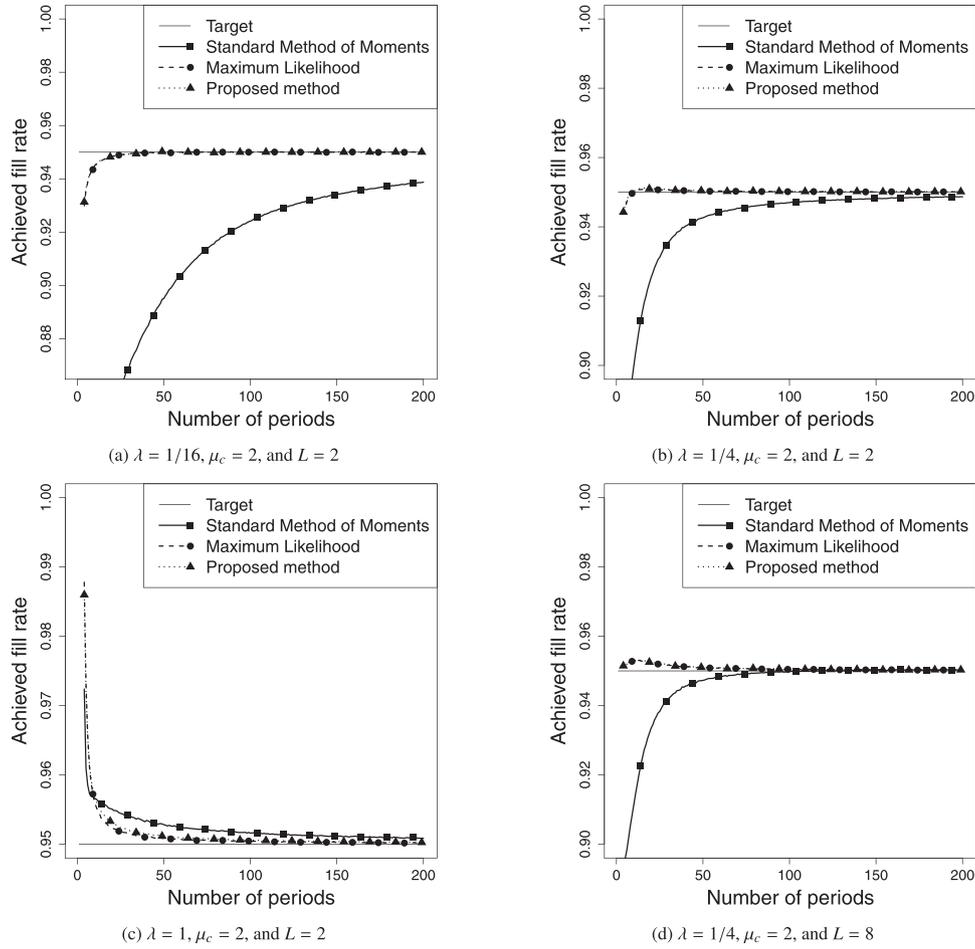


Fig. 3. Achieved Fill Rates (Target Fill Rate 95%). Vertical Axes Are Truncated for Readability.

This makes the MM estimator sensitive to the (next moment of the) compounding distribution and therefore generally inconsistent under distribution misspecification.

We demonstrate this for the case where the compounding distribution has two parameters, which we take to be the mean μ_c and the variance σ_c^2 . The proposed estimator is now the solution to the following three equations

$$\hat{\lambda} = -\ln(n_0/n)$$

$$\hat{\mu}_1 = \hat{\lambda}\mu_c$$

$$\hat{\mu}_2 = \hat{\lambda}(\hat{\sigma}_c^2 + \mu_c^2)$$

The first two equations directly lead to the same estimators for λ and μ_c as for a one-parameter compounding distribution. Substituting these into the last equation gives

$$\hat{\sigma}_c^2 = \frac{\hat{\mu}_2}{\hat{\lambda}} - \frac{\hat{\mu}_1^2}{\hat{\lambda}^2}.$$

This closed form again does not depend on the specific compounding distribution. Since this equation is directly obtained from the general formula of the second moment of a compound Poisson distribution, and as $\hat{\lambda} \rightarrow \lambda$ and $\hat{\mu}_c \rightarrow \mu_c$, the continuous mapping theorem implies that $\hat{\sigma}_c^2 \rightarrow \sigma_c^2$, irrespective of the compounding distribution. The standard MM estimator in this case replaces the first equation by the third central moment equation, which depends on the first three central moments of the compounding distribution [40]. Therefore, it uses the third moment of the compounding distribution to estimate its second moment. If more pa-

rameters (or moments of the compounding distribution) are to be estimated, then the recursive solution procedure of the proposed estimator (and therefore also its robustness) readily extends.

We continue to numerically analyze the achieved accuracy of standard MM, ML, and the proposed estimator under misspecification of the compounding distribution. In line with the remainder of the paper, we assume that the demand sizes are exponentially distributed. However, we use two different distributions to generate demand sizes, namely the (continuous) uniform distribution with bounds 0 and $2\mu_c$, and the (discrete) logarithmic distribution (which was also used by Axsäter [2]) with mean μ_c . The proposed estimators for the arrival rate and the demand size mean have the same closed form for any compounding distribution and converge to the true values by construction. The other estimators, however, do show asymptotic biases under misspecification. Fig. 4 shows the results for the uniform distribution.

The standard MM estimator for λ shows an asymptotic upward bias of 52% in all scenarios, whereas for μ_c it shows an asymptotic downward bias of 33% in all scenarios. This method therefore does not yield reliable estimates at all when the demand distribution is misspecified. ML is more robust than standard MM, but its asymptotic bias is still considerable for larger values of λ . For $\lambda = 1/16$ the error is negligible. For $\lambda = 1/4$ we record for $n = 200$ a 0.7% upward bias in λ and a 0.3% downward bias in μ_c . For $\lambda = 1$ this error has increased to +5% for λ and -4% for μ_c . The proposed method is, in line with our expectation, insensitive to the misspecified demand distribution and shows an unabated accuracy in comparison to the analysis of Section 4.2. Also in the distribution

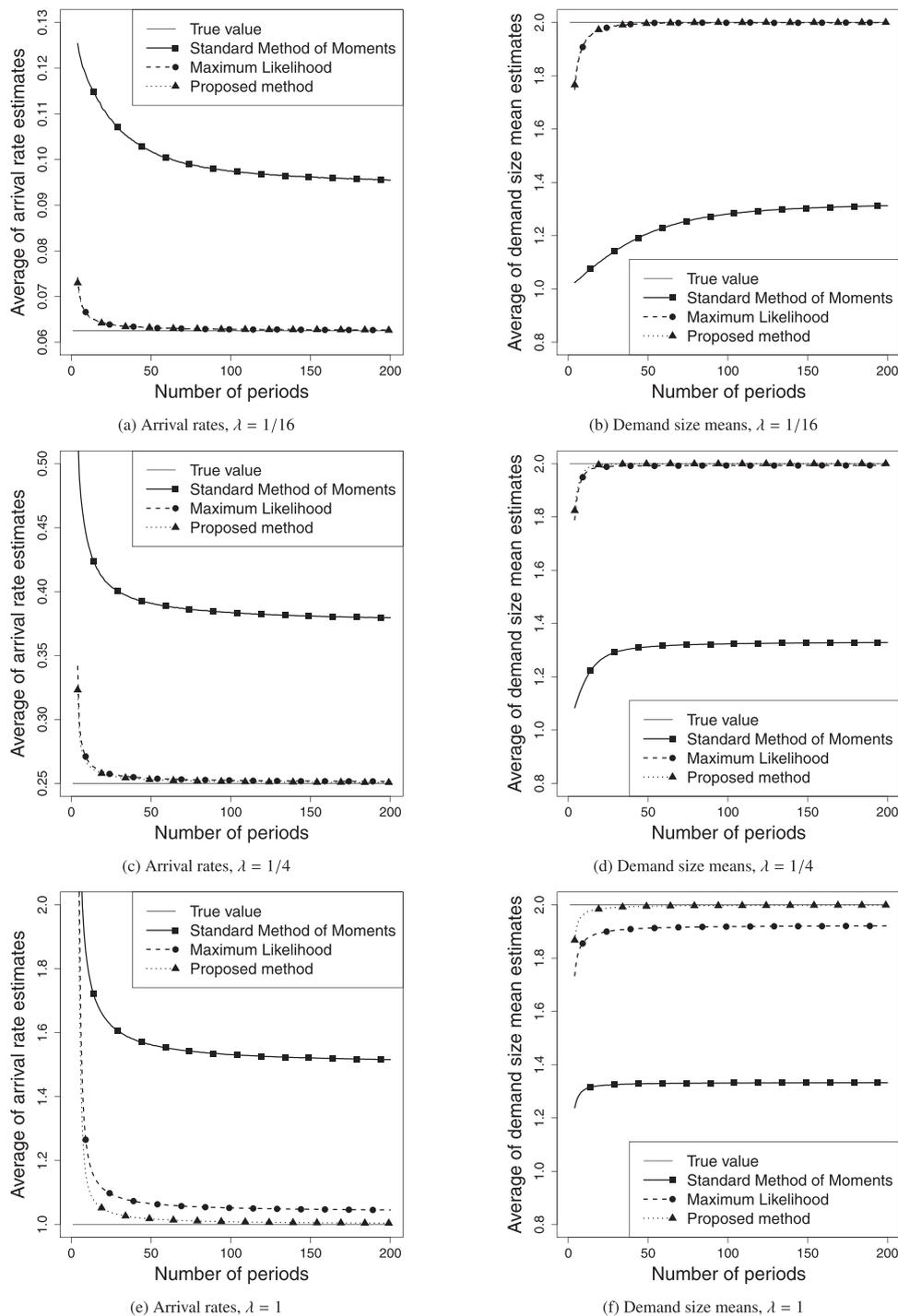


Fig. 4. Averages of Demand Parameter Estimates under Distribution Misspecification, Uniform, $\mu_c = 2$. Vertical Axes Are Truncated for Readability.

misspecification case, the difference between the proposed method and ML is larger if λ is not very small.

Fig. 5 shows the results for the logarithmic distribution. As the logarithmic distribution is discrete, it can happen (especially for small sample sizes) that all observations have the same value and the variance is zero. The standard MM estimator (for the exponential compounding distribution) is then undefined, which is the reason why in Fig. 5e and f the bias of standard MM is not drawn for $n \leq 7$. This issue does not exist for ML and the proposed estimator. The plots are qualitatively similar to those of Fig. 4. As the logarithmic distribution's shape more closely resembles that of the exponential distribution, the asymptotic biases of ML and standard MM

are smaller in this case. Nevertheless, standard MM still shows a large bias in all scenarios, and for $\lambda = 1$ we also see a clear asymptotic bias in ML. Only the proposed estimator is consistent (and has the smallest bias for all sample sizes) in all set-ups.

8. Case study

In this section we compare the different estimators for compound Poisson demand parameters in practice, using a data set of demands for 496 different commercial airline parts. This data set contains demand observations at the individual customer order level, which we aggregate to 125 weekly totals. We restrict our

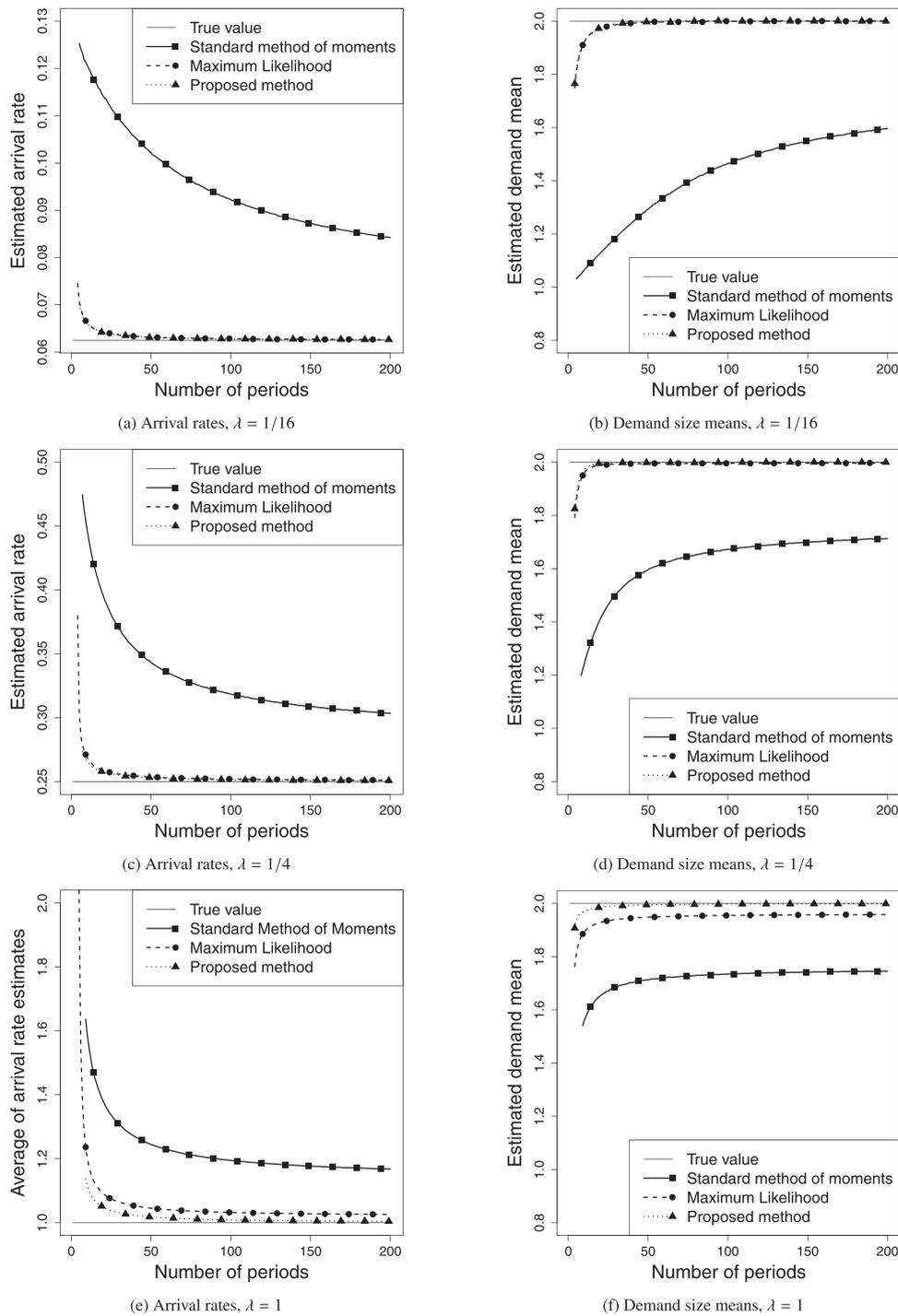


Fig. 5. Averages of Demand Parameter Estimates under Distribution Misspecification, Logarithmic, $\mu_c = 2$. Vertical Axes Are Truncated for Readability.

analysis to the 404 parts that have had at least one demand in the first 20 weeks. In Table 4 we present some descriptive statistics. The estimated arrival rates and demand size means are calculated based on the individually stored demands of the full history per item. Almost all arrival rates are below 1, whereas the demand sizes vary wildly.

Based on the weekly aggregated demand observations, we set up a comparative analysis for the arrival rates and demand size means according to standard MM, ML, and the proposed method. As a benchmark, we furthermore include the ‘theoretically best’ estimates based on the individual orders, which, of course, cannot be exploited using aggregated data. Lengu et al. [12] established that a

compound Poisson process with geometric compounding distribution has a good fit to this data. We adopt this distribution (which is further studied in Appendix B), but in line with the rest of this paper, we also test the exponential compounding distribution. Starting with the first 20 weeks and every time increasing the time horizon by one week, we estimate for every horizon the arrival rate and demand size mean for all items. We report the Mean Absolute Percentage Error (MAPE) between these estimates and the ‘theoretically best’ estimator over the entire sample (for every time horizon, as an average across all items). The proposed method and the ‘theoretically best’ method do not depend on the specific compounding distribution. Furthermore, the ML results do not vary

Table 4
Case Study Data Set Descriptive Statistics.

No. Items	No. Weeks	Estimated Arrival Rate						
		Min	Max	Mean	Median	≤ 0.25	≤ 0.5	≤ 1
404	125	0.04	2.87	0.29	0.19	65%	86%	97%
Estimated (Individual) Demand Size Mean								
Min	Max	Mean	Median	≤ 2	≤ 10	≤ 100		
1.07	802.5	21.42	5.34	6%	74%	94%		

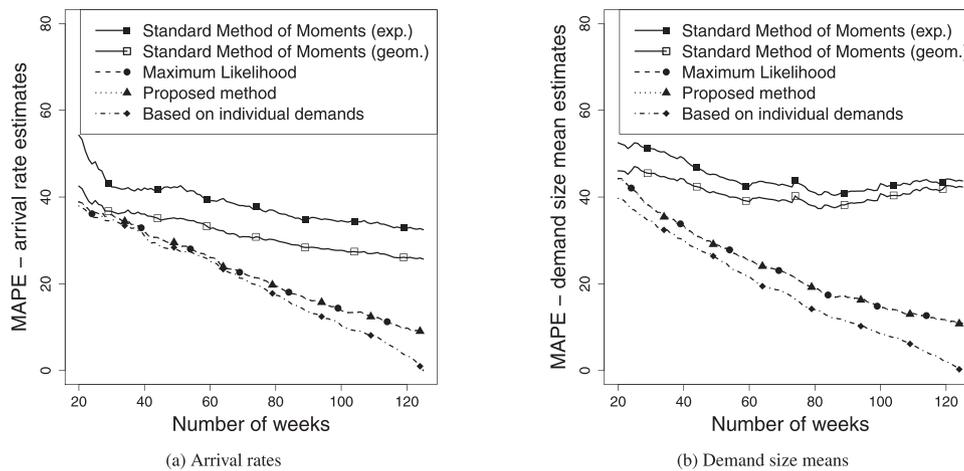


Fig. 6. Mean Absolute Percentage Errors Between the Different Methods and the Theoretically Best Estimator for Case Study Data.

visibly between the exponential and geometric compounding distribution. Therefore, only for standard MM do we report the results under both distributions separately. The results are depicted in Fig. 6.

By construction, the MAPE of the individual demands estimator converges to 0 and the evolution of this estimator throughout the sample benchmarks the other methods. The standard MM estimator clearly performs worse when the exponential distribution is assumed. However, in both cases it performs considerably worse than the other methods. For the arrival rate, the proposed estimator slightly outperforms ML by 0.2% for the full sample. For smaller samples, the differences between the estimators are smaller, but their performance ranking does not change. For the demand size mean, standard MM even deteriorates when the sample size passes 90 weeks. The theoretical results from the preceding sections are confirmed also in this practical case study: standard MM performs by far worst and is most sensitive to the exact specification of the compounding distribution, and the proposed estimator marginally outperforms ML. It is furthermore established that the accuracy loss due to storing only weekly data increases with the sample size. If all 125 weeks are used, then the proposed estimator loses approximately 10% in accuracy compared to using individual transaction data. Standard MM, contrarily, loses approximately 35% or 40%, depending on the compounding distribution.

9. Conclusions

The inventory control literature provides hardly any guidance on obtaining demand parameters from period demand data, although companies typically do store historical demand observations per period and software packages also use period demand as input for their forecasts. Period demand forecasts cannot be directly transformed into compound Poisson parameter estimates, even if the forecasting procedure yields separate estimates for the time between periods with positive demands and the average pe-

riod demand size. Misusing these to obtain demand parameters leads – also asymptotically – to severely biased estimates, overshoot fill rates, and thus excessive inventories and associated costs. It is important to observe that this flaw occurs if any period forecasting procedure is used to estimate individual customer demand parameters. We have addressed this lack of connection between the forecasting and inventory control literature, and performed a comparative study of different compound Poisson parameter estimators. We proposed an alternative MM estimator that uses a non-standard choice of moments, reflecting the intermittent nature of the demand. The fraction of periods without demand is used to estimate the arrival rate, and the estimated arrival rate is combined with further central moments of the (full) compound Poisson distribution to estimate the parameters of the compounding distribution (the distribution of the individual demand sizes).

Our results confirmed that the standard MM estimator (the default option in inventory control) has severe biases for a wide range of intermittent demand patterns and sample sizes, and has a larger variance than the other studied methods. The proposed method's deviation from the true value of both parameters is smaller over the entire range of intermittent demand patterns and sample sizes considered. In several scenarios (particularly also in our case study) the proposed estimator has a lower bias than ML, whereas in the other scenarios it performs almost indistinguishably. A key advantage of the proposed method is that its closed form does not depend on the specific demand size distribution, which implies that it is completely robust to misspecification of that distribution. This property (which neither standard MM, nor ML possesses) is particularly relevant in practice. The closed-form nature of the estimator is useful in large-scale applications where many items are controlled simultaneously and computation speed is an issue. Both practitioners and applied researchers should be aware of 1) the accuracy loss due to storing demand only periodically, and 2) the consequences of their estimator choice on accuracy and inventory performance. It is furthermore important to

realize that this parameter estimation issue exists also for other inventory models and service measures, as well as outside inventory control. Accuracy results carry over directly, but the final effect is of course specific to the decision model.

An important empirical research avenue is to study the effect of bucketing the demand observations into different period lengths. Our study showed that the relationship between the intermittency of the demand pattern and the inventory performance of the different estimators is non-trivial. A further research opportunity lies in the fact that in practice it is sometimes questionable if and for how long the demand process remains stationary. A non-stationary compound Poisson process – which does not have a fixed arrival rate, but a time-varying arrival intensity function – can model this situation. If it is assumed that the intensity function is piece-wise linear, then the process reduces to a stationary compound Poisson process on the corresponding subsets of the data and our method can still be used. For more general intensity functions, however, existing literature is solely based on ML, yielding an opportunity for future research. Finally, an overall connection is lacking in the broader sense between the forecasting and inventory control literature. Forecast accuracy is typically assessed via symmetric loss measures, whereas from an inventory control perspective over- and underestimation carry asymmetric penalties. A study comparing different forecasting methods from this perspective may lead to new insights.

CRedit authorship contribution statement

Dennis Prak: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - original draft, Writing - review & editing, Visualization, Funding acquisition. **Ruud Teunter:** Conceptualization, Methodology, Resources, Writing - review & editing, Funding acquisition. **Mohamed Zied Babai:** Conceptualization, Resources, Writing - review & editing. **John E. Boylan:** Conceptualization, Resources, Writing - review & editing. **Aris Syntetos:** Conceptualization, Resources, Writing - review & editing.

Acknowledgements

Funding: This research was supported by the Dutch Research Council (NWO) [grant nr. 019.191SG.003].

Appendix A. Taylor approximations for biases

In this Appendix we give the derivations of the second-order Taylor approximations leading to the results of Section 4.1. A twice differentiable function $f(x,y)$ can be approximated around the point (a,b) as follows:

$$f(x,y) \approx f(a,b) + \frac{\partial f}{\partial x}(x-a) + \frac{\partial f}{\partial y}(y-b) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x-a)^2 + \frac{\partial^2 f}{\partial x \partial y}(x-a)(y-b) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(y-b)^2.$$

Under the conditions $E[x] = a$ and $E[y] = b$, it follows that

$$E[f(x,y)] \approx E[f(a,b)] + \frac{1}{2} E \left[\frac{\partial^2 f}{\partial x^2}(x-a)^2 \right] + E \left[\frac{\partial^2 f}{\partial x \partial y}(x-a)(y-b) \right] + \frac{1}{2} E \left[\frac{\partial^2 f}{\partial y^2}(y-b)^2 \right].$$

Recall that the first four central moments of the Poisson-exponential distribution are $\mu_1 = \lambda\mu_c$, $\mu_2 = 2\lambda\mu_c^2$, $\mu_3 = 6\lambda\mu_c^3$, and $\mu_4 = (24\lambda + 12\lambda^2)\mu_c^4$.

A1. Standard MM arrival rate estimator

The standard MM arrival rate estimator is given by $\hat{\lambda}_{MM} = \frac{2\hat{\mu}_1^2}{\hat{\mu}_2}$. Using $f(x,y) = 2x^2/y$, $x = \hat{\mu}_1$, $y = \hat{\mu}_2$, $a = \mu_1 = \lambda\mu_c$ and $b = \mu_2 = 2\lambda\mu_c^2$, we have that

$$E[\hat{\lambda}_{MM}] \approx \lambda + \frac{1}{2} E \left[\frac{\partial^2 f}{\partial \hat{\mu}_1^2} (\hat{\mu}_1 - \mu_1)^2 \right] + E \left[\frac{\partial^2 f}{\partial \hat{\mu}_1 \partial \hat{\mu}_2} (\hat{\mu}_1 - \mu_1)(\hat{\mu}_2 - \mu_2) \right] + \frac{1}{2} E \left[\frac{\partial^2 f}{\partial \hat{\mu}_2^2} (\hat{\mu}_2 - \mu_2)^2 \right] = \lambda + \frac{1}{2} \frac{4}{2\lambda\mu_c^2} \text{Var}(\hat{\mu}_1) - \frac{4\mu_c\lambda}{(2\lambda\mu_c^2)^2} \text{Cov}(\hat{\mu}_1, \hat{\mu}_2) + \frac{1}{2} \frac{4(\mu_c\lambda)^2}{(2\lambda\mu_c^2)^3} \text{Var}(\hat{\mu}_2).$$

Let n i.i.d. observations from a population with mean μ_1 and variance μ_2 be given. The sample mean of these observations has variance μ_2/n . Therefore, $\text{Var}(\hat{\mu}_1) = \frac{2\lambda\mu_c^2}{n}$. The covariance of their sample mean and sample variance is μ_3/n [41]. Therefore, $\text{Cov}(\hat{\mu}_1, \hat{\mu}_2) = \frac{6\lambda\mu_c^3}{n}$. The variance of their sample variance is $\frac{\mu_4}{n} - \frac{\mu_2^2(n-3)}{n(n-1)}$ [42]. Therefore, $\text{Var}(\hat{\mu}_2) = \frac{(24\lambda + 12\lambda^2)\mu_c^4 - (2\lambda\mu_c^2)^2(n-3)}{n(n-1)}$. Substituting these terms and simplifying gives the result

$$E[\hat{\lambda}_{MM}] \approx \lambda + \frac{2\lambda}{n-1} + \frac{2}{n}.$$

A2. Proposed arrival rate estimator

The proposed arrival rate estimator is given by $\hat{\lambda}_{NEW} = -\ln(\frac{n_0}{n})$. Using $f(x,y) = -\ln(x)$, $x = N_0/n$, $y = 0$, $a = E[N_0/n] = \exp(-\lambda)$ and $b = 0$, we have that

$$E[\hat{\lambda}_{NEW}] \approx \lambda + \frac{1}{2} E \left[\frac{\partial^2 f}{\partial (N_0/n)^2} (N_0/n - \exp(-\lambda))^2 \right] = \lambda + \frac{1}{2} \frac{1}{\exp(-2\lambda)} \text{Var}(N_0/n).$$

Since N_0 follows a binomial distribution with parameters $\exp(-\lambda)$ and n , we have $\text{Var}(N_0/n) = \frac{\exp(-\lambda)(1-\exp(-\lambda))}{n}$. Substituting and simplifying gives the result

$$E[\hat{\lambda}_{NEW}] \approx \lambda + \frac{\exp(\lambda) - 1}{2n}.$$

A3. Standard MM demand size mean estimator

The standard MM demand size mean estimator is given by $\hat{\mu}_{c,MM} = \frac{\hat{\mu}_2}{2\hat{\mu}_1}$. Using $f(x,y) = y/(2x)$, $x = \hat{\mu}_1$, $y = \hat{\mu}_2$, $a = \mu_1 = \lambda\mu_c$ and $b = \mu_2 = 2\lambda\mu_c^2$, we have that

$$E[\hat{\mu}_{c,MM}] \approx \mu_c + \frac{1}{2} E \left[\frac{\partial^2 f}{\partial \hat{\mu}_1^2} (\hat{\mu}_1 - \mu_1)^2 \right] + E \left[\frac{\partial^2 f}{\partial \hat{\mu}_1 \partial \hat{\mu}_2} (\hat{\mu}_1 - \mu_1)(\hat{\mu}_2 - \mu_2) \right] + \frac{1}{2} E \left[\frac{\partial^2 f}{\partial \hat{\mu}_2^2} (\hat{\mu}_2 - \mu_2)^2 \right] = \mu_c + \frac{1}{2} \frac{2\lambda\mu_c^2}{(\lambda\mu_c)^3} \text{Var}(\hat{\mu}_1) - \frac{1}{4(\lambda\mu_c)^2} \text{Cov}(\hat{\mu}_1, \hat{\mu}_2) = \mu_c \left(1 - \frac{1}{\lambda n} \right).$$

A4. Proposed demand size mean estimator

The proposed demand size mean estimator is given by $\hat{\mu}_{c,NEW} = \frac{-\hat{\mu}_1}{\ln(n_0/n)}$. Using $f(x, y) = -x/(\ln(y))$, $x = \hat{\mu}_1$, $y = N_0/n$, $a = \mu_1 = \lambda\mu_c$ and $b = E[N_0/n] = \exp(-\lambda)$, we have that

$$\begin{aligned} E[\hat{\mu}_{c,NEW}] &\approx \mu_c + \frac{1}{2}E\left[\frac{\partial^2 f}{\partial \hat{\mu}_1^2}(\hat{\mu}_1 - \mu_1)^2\right] \\ &+ E\left[\frac{\partial^2 f}{\partial \hat{\mu}_1 \partial N_0/n}(\hat{\mu}_1 - \mu_1)(N_0/n - \exp(-\lambda))\right] \\ &+ \frac{1}{2}E\left[\frac{\partial^2 f}{\partial (N_0/n)^2}(N_0/n - \exp(-\lambda))^2\right] \\ &= \mu_c + \frac{1}{\exp(-\lambda)\lambda^2} \text{Cov}(\hat{\mu}_1, N_0/n) \\ &+ \frac{1}{2} \frac{\lambda\mu_c(2-\lambda)}{\exp(-2\lambda)\lambda^3} \text{Var}(N_0/n). \end{aligned}$$

From Appendix A.2 we have that $\text{Var}(N_0/n) = \frac{\exp(-\lambda)(1-\exp(-\lambda))}{n}$. Furthermore, we find $\text{Cov}(\hat{\mu}_1, N_0/n) = E(\hat{\mu}_1 N_0/n) - E(\hat{\mu}_1)E(N_0/n)$, where $E(\hat{\mu}_1) = \mu_1$ and $E(N_0/n) = \exp(-\lambda)$. We are left with finding $E(\hat{\mu}_1 N_0/n)$, and first condition on N_0 . Given N_0 periods without demand, $N - N_0$ periods had positive demand. The expected number of customers in a period, given that there was at least one customer, is $\lambda/(1 - \exp(-\lambda))$. As, independently, every customer has a demand with mean μ_c , it follows that

$$\begin{aligned} E(\hat{\mu}_1 N_0/n | N_0) &= \frac{N_0}{n} \frac{n - N_0}{n} \frac{\lambda}{1 - \exp(-\lambda)} \mu_c \\ &= \left(\frac{N_0}{n} - \left(\frac{N_0}{n}\right)^2\right) \frac{\lambda}{1 - \exp(-\lambda)} \mu_c, \end{aligned}$$

so that

$$\begin{aligned} E(\hat{\mu}_1 N_0/n) &= E[E(\hat{\mu}_1 N_0/n | N_0)] \\ &= (E(N_0/n) - E[(N_0/n)^2]) \frac{\lambda}{1 - \exp(-\lambda)} \mu_c. \end{aligned}$$

As $E(N_0/n) = \exp(-\lambda)$ and $\text{Var}(N_0/n) = \frac{\exp(-\lambda)(1-\exp(-\lambda))}{n}$, it follows that

$$E[(N_0/n)^2] = \frac{\exp(-\lambda)(1 - \exp(-\lambda))}{n} + \exp(-2\lambda).$$

Substituting and simplifying gives

$$\begin{aligned} E(\hat{\mu}_1 N_0/n) &= \left(1 - \frac{1}{n}\right)(\exp(-\lambda) - \exp(-2\lambda)) \frac{\lambda\mu_c}{1 - \exp(-\lambda)} \\ &= \left(1 - \frac{1}{n}\right)\lambda \exp(-\lambda)\mu_c. \end{aligned}$$

Putting everything together, we find $\text{Cov}(\hat{\mu}_1, N_0/n) = -\frac{\lambda\mu_c \exp(-\lambda)}{n}$.

Finally, plugging $\text{Cov}(\hat{\mu}_1, N_0/n)$ into the original equation and simplifying, we arrive at

$$E[\hat{\mu}_{c,NEW}] \approx \mu_c \left(1 - \frac{(\lambda - 2) \exp(\lambda) + \lambda + 2}{2\lambda^2 n}\right).$$

Appendix B. Geometric compounding distribution

In this appendix we discuss the application of standard MM, ML, and the proposed method for the geometric compounding distribution and provide numerical results for their biases. The standard MM estimators for a compound Poisson process with geometric compounding distribution are [2,27]:

$$\hat{\lambda} = \frac{2\hat{\mu}_1^2}{\hat{\mu}_1 + \hat{\mu}_2} \text{ and } \hat{\mu}_c = \frac{\hat{\mu}_1 + \hat{\mu}_2}{2\hat{\mu}_1}.$$

For the ML estimators, we use that the probability mass function of the Poisson-geometric distribution is [32]

$$f_{PG}(x) = \exp(-\lambda) \sum_{i=0}^x \frac{1}{i!} \binom{x-1}{i-1} (\lambda/\mu)^i \left(1 - \frac{1}{\mu}\right)^{x-i}$$

and follow the procedure of Section 3.2. The proposed method does not depend on the compounding distribution and stays unaltered:

$$\hat{\lambda} = -\ln(n_0/n) \text{ and } \hat{\mu}_c = -\frac{\hat{\mu}_1}{\ln(n_0/n)}.$$

We perform the same numerical experiment as outlined in Section 4.2, now using the geometric compounding distribution both for generating the individual demands and calculating the estimates. We discuss only the setting with $\lambda = 1/4$ and $\mu_c = 2$ and omit the variance and fill rate analysis, as all results are very similar to those under the exponential distribution. Fig. 7 shows the results. Standard MM is now more accurate than under the exponential distribution, but still leaves a considerable gap with the other methods.

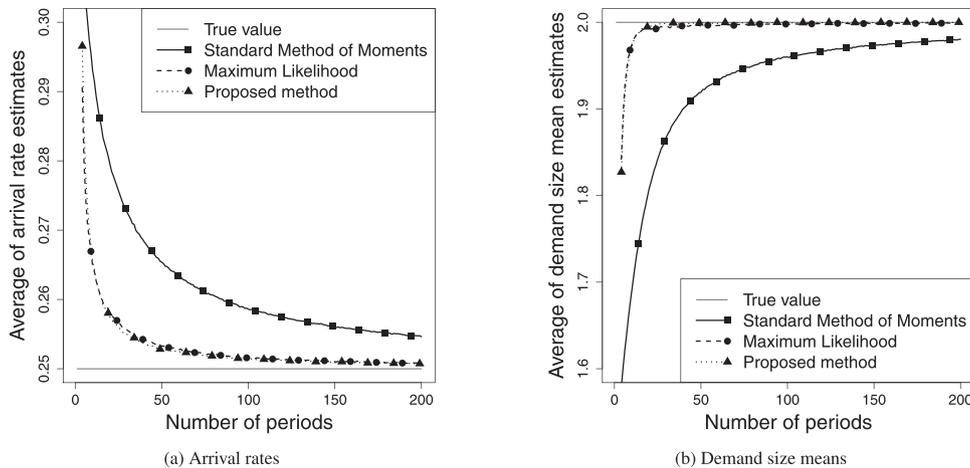


Fig. 7. Averages of Demand Parameter Estimates, Geometric Distribution, $\lambda = 1/4$, $\mu_c = 2$. Vertical Axes Are Truncated for Readability.

References

- [1] Zipkin PH. Foundations of inventory management. 1st ed. New York: McGraw-Hill; 2000.
- [2] Axsäter S. Inventory control. 3rd ed. Basel: Springer International Publishing; 2015.
- [3] Silver EA, Pyke DF, Thomas DJ. Inventory and production management in supply chains. 4th ed. Boca Raton: CRC Press; 2017.
- [4] Galliher HP, Morse PM, Simond M. Dynamics of two classes of continuous-review inventory systems. *Oper Res* 1959;7(3):362–84.
- [5] Feeney GJ, Sherbrooke CC. The $(S-1, S)$ inventory policy under compound Poisson demand. *Manage Sci* 1966;12(5):391–411.
- [6] Chen J, Jackson PL, Muckstadt JA. Exact analysis of a lost sales model under stuttering Poisson demand. *Oper Res* 2011;59(1):249–53.
- [7] Archibald BC, Silver EA. (s, S) policies under continuous review and discrete compound Poisson demand. *Manage Sci* 1978;24(9):899–909.
- [8] Sherbrooke CC. Metric: a multi-echelon technique for recoverable item control. *Oper Res* 1968;16(1):122–41.
- [9] Graves SC. A multi-echelon inventory model for a repairable item with one-for-one replenishment. *Manage Sci* 1985;31(10):1247–56.
- [10] Axsäter S. Simple solution procedures for a class of two-echelon inventory problems. *Oper Res* 1990;38(1):64–9.
- [11] Syntetos AA, Babai MZ, Altay N. On the demand distributions of spare parts. *Int J Prod Res* 2012;50(8):2101–17.
- [12] Lengu D, Syntetos AA, Babai MZ. Spare parts management: linking distributional assumptions to demand classification. *Eur J Oper Res* 2014;235(3):624–35.
- [13] Turrini L, Meissner J. Spare parts inventory management: new evidence from distribution fitting. *Eur J Oper Res* 2019;273(1):118–30.
- [14] van Jaarsveld W, Dollevoet T, Dekker R. Improving spare parts inventory control at a repair shop. *Omega* 2015;57:217–29.
- [15] Chang JA-C, Lu H, Shi JJ. Stockout risk of production-inventory systems with compound Poisson demands. *Omega* 2019;83:181–98.
- [16] Chen J, Dong M, Rong Y, Yang L. Dynamic pricing for deteriorating products with menu cost. *Omega* 2018;75:13–26.
- [17] Rosling K. Inventory cost rate functions with nonlinear shortage costs. *Oper Res* 2002;50(6):1007–17.
- [18] Johansson L, Sonntag DR, Marklund J, Kiesmüller GP. Controlling distribution inventory systems with shipment consolidation and compound Poisson demand. *Eur J Oper Res* 2020;280(1):90–101.
- [19] Babai MZ, Ali MM, Nikolopoulos K. Impact of temporal aggregation on stock control performance of intermittent demand estimators: empirical analysis. *Omega* 2012;40(6):713–21.
- [20] Rostami-Tabar B, Babai MZ, Syntetos AA, Ducq Y. Demand forecasting by temporal aggregation. *Nav Res Logist* 2013;60(6):479–98.
- [21] Kourentzes N, Petropoulos F, Trapero JR. Improving forecasting by estimating time series structural components across multiple frequencies. *Int J Forecasting* 2014;30(2):291–302.
- [22] Croston JD. Forecasting and stock control for intermittent demands. *Oper Res Q* 1972;23(3):289–303.
- [23] Fildes R, Schaer O, Sventunkov I. Forecasting 2018. *ORMS Today* 2018;45(3):44–51.
- [24] Syntetos AA, Babai Z, Boylan JE, Kolassa S, Nikolopoulos K. Supply chain forecasting: theory, practice, their gap and the future. *Eur J Oper Res* 2016;252(1):1–26.
- [25] Bacchetti A, Saccani N. Spare parts classification and demand forecasting for stock control: investigating the gap between research and practice. *Omega* 2012;40(6):722–37.
- [26] Costantino F, Di Gravio G, Patriarca R, Petrella L. Spare parts management for irregular demand items. *Omega* 2018;81:57–66.
- [27] Ward JB. Determining reorder points when demand is lumpy. *Manage Sci* 1978;24(6):623–32.
- [28] Anscombe FJ. Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika* 1950;37(3–4):358–82.
- [29] Sichel HS. The estimation of the parameters of a negative binomial distribution with special reference to psychological data. *Psychometrika* 1951;16(1):107–27.
- [30] Öztürk A. On the study of a probability distribution for precipitation totals. *J Appl Meteorol* 1981;20(12):1499–505.
- [31] Withers C, Nadarajah S. On the compound Poisson-gamma distribution. *Kybernetika* 2011;47(1):15–37.
- [32] Balakrishnan N, Ye Y, Al-Jarallah RA. Likelihood inference for type I bivariate Pólya–Aeppli distribution. *Commun Stat-Simul Comput* 2017;46(8):6436–53.
- [33] Johnson NL, Kemp AW, Kotz S. Univariate discrete distributions. Hoboken: John Wiley & Sons; 2005.
- [34] Evans DA. Experimental evidence concerning contagious distributions in ecology. *Biometrika* 1953;40(1–2):186–211.
- [35] Ehrenberg ASC. The pattern of consumer purchases. *J R Stat Soc Ser C (Applied Statistics)* 1959;8(1):26–41.
- [36] Savani V, Zhigljavsky AA. Efficient estimation of parameters of the negative binomial distribution. *Commun Stat - Theory Methods* 2006;35(5):767–83.
- [37] Song J-S. On the order fill rate in a multi-item, base-stock inventory system. *Oper Res* 1998;46(6):831–45.
- [38] Syntetos AA, Boylan JE. On the bias of intermittent demand estimates. *Int J Prod Econ* 2001;71(1):457–66.
- [39] Syntetos AA, Boylan JE. The accuracy of intermittent demand estimates. *Int J Forecasting* 2005;21(2):303–14.
- [40] Grubbström RW, Tang O. The moments and central moments of a compound distribution. *Eur J Oper Res* 2006;170(1):106–19.
- [41] Zhang L. Sample mean and sample variance: their covariance and their (in)dependence. *Am Stat* 2007;61(2):159–60.
- [42] Mood AM, Graybill FA, Boes DC. Introduction to the theory of statistics. New York: McGraw-Hill; 1974.