

# Order quantity distributions in make-to-order manufacturing: At what level of aggregation do they respect standard assumptions?

Poul Svante Eriksen<sup>1</sup> and Peter Nielsen<sup>2\*</sup>

<sup>1</sup>Department of Mathematical Sciences, Aalborg University, Denmark, <sup>2</sup>Department of Mechanical and Manufacturing Engineering, Aalborg University, Denmark. \*corresponding author [peter@m-tech.aau.dk](mailto:peter@m-tech.aau.dk), Fibigerstraede 16, 9220 Aalborg Oest.

**Abstract.** This paper presents both an analytical and a numerical investigation into the order quantities received by a company in the form of customer orders. A discussion of assumptions regarding the behavior of demand in the form of customer orders from various perspectives within manufacturing planning and control with a special emphasis on the make-to-order environment is presented. A methodological framework for analyzing the behavior of orders and investigate the validity of the assumptions is given. Furthermore, an analytical approach to identify the horizon needed for aggregating orders to achieve a stable demand is developed and tested on data from a real case.

Keywords: Poisson process, aggregate demand, order sizes, planning horizon.

## 1 Introduction

Companies are in the business of satisfying customer demands in the form of customer orders. This paper focuses on the operational characteristics of ordering behavior in the form of demand and discusses the underlying assumptions within a number of fields. These characteristics include volume per planning period (demand rate – often disaggregated to several time horizons and described as a distribution of stochastic variables), accumulation of demand over the planning period, the number of orders per product per planning period and the standard order quantities for each product. The characteristics of the make-to-order manufacturing are depicted by Hendry and Kingsman [9] and Marucheck and McClelland [10]:

- resources are multi-task machinery and flexible workforce – a large number of possible end items within a product line,
- product lead time is vital for customer satisfaction and agreed with customer before production commences,
- capacity planning is based on receipt of customer orders and cannot be planned for in advance,
- demand is volatile and can rarely be predicted – a small volume and high degree of demand variability both in quantity and product mix.

This paper focuses specifically on analyzing the distributions of customer orders, the transformation of these into demand rates and presents an analytical model to calculate the aggregation horizon necessary to ensure a stable planning environment. In this context a stable planning environment is defined as having a demand pattern approximately symmetrically distributed with a given (low) Coefficient of Variance (CV) (variance of distribution divided by mean). The presented model can be used to estimate the planning horizon (i.e. length of planning periods) necessary to ensure a stable demand rate in a given situation.

The remainder of the paper is structured as follows. First, a discussion of assumptions regarding demand in the form of customer orders and their behavior is presented. Second, a methodological framework for analyzing the behavior of orders is given. Third, an analytical approach to identify the horizon needed for aggregating orders to achieve a stable demand is developed. Fourth, a case study application of the presented framework. Finally conclusions and further research are presented.

## 2 Background, implications and approach to analysis

Manufacturing Planning & Control involves coordinating any number of decisions [19]. Several of these are directly related to the customer orders and their characteristics. The demand experienced by companies comes in the form of customer orders that may consist of multiple order lines. Each order (line) will contain the following parameters relevant in this context: desired due date, product identification and quantity. These parameters are then transformed into various forms depending on the usage. Within line balancing customer orders are likewise transformed into demand rates (also termed release or launch rates [2]) for individual products and product mix [13] for a given line. Likewise within lot-sizing demand rates are typically considered deterministic or to stem from a stationary stochastic process [4], despite the fact that demand rates are composed of individual (in)dependent customer orders.

Regardless of application area, the general approach is to aggregate individual customer orders into demand rates for both individual products and product families and use e.g. MRP or MRPII to generate dependent demand and resource requirements [16], [19]. The aggregation can be conducted along a number of dimensions [22] (typically time and products [21]) until a stabilized customer ordering behavior is identified. To facilitate the aggregation and disaggregation of plans a number of assumptions are made regarding the orders handled by the system [1]. These assumptions can be abbreviated to assuming a one-to-one relationship in aggregation and disaggregation [7], [8]. As a result order quantities and the number of orders per period are assumed to be constant. This facilitates calculating time needed for changeovers and the number of changeovers needed per planning period (e.g. Silver et al. [16]). If this is not the case, the disaggregation of aggregate plans may result in suboptimal or even infeasible disaggregate plans. In reality this assumption can be relaxed if the order quantities come from i.i.d. symmetrical distributions (with a low CV; variance compared to the mean of the distribution). A chronological list of order

lines for delivery can e.g. expect to exhibit non-stationary behavior. A typical manifestation of this will be if an order contains order lines for more than one product from the same product family. This becomes critical if the order lines are related in a manner so that their demand will be aggregated within the planning environment. On an individual product level assuming independently distributed order quantities is more reasonable, especially if many customers purchase the same products. For all areas a stabilized behavior is expected as demand is aggregated [6], [7], [17]. The critical aspect is here if the customer ordering behavior allows for a simple linear aggregation along product family or time dimensions? Furthermore, if linear relationships in aggregation are an oversimplification of the issue, what should be done instead? The research presented in this paper addresses the assumptions used to aggregate demand information for various planning purposes. The aim is first to design a method for robust estimation of the behavior of order quantities, i.e. determine whether they can be assumed to be independently distributed from a symmetrical distribution. The second step is to develop a mathematical model that describes the CV for the order quantities, under the assumption that the order quantities can stem from an arbitrary distribution and are dependently distributed. First tests of symmetrical distribution of order quantities are applied. If these are negative, they are considered to stem from a skewed distribution and are therefore also assumed to be non-constant. In order to investigate asymmetry of order quantity distributions, in this paper the following three different statistical tests will be applied:

- The MGG-test presented in Miao et al. [11]
- The CM-test described in Cabilio and Masaro [5]
- The Mira-test developed by Mira [12]

These tests are implemented in the R-package `lawstat` [15]. From the literature review it is known that the individual orders for each product can be considered to be independently distributed, while the aggregate behavior cannot necessarily be assumed to abide by the same behavior. To test for independence the autocorrelation function is used [3]. In the current work only one period lag is considered. Having first established whether the order quantities can be considered to stem from a symmetrical distribution, the next step is to investigate how to stabilize the demand behavior experienced in the form of customer orders. In practice this means determining the horizon needed (i.e. number of orders) to aggregate. The model should therefore be able to determine the number of consecutive orders to aggregate to achieve a distribution with a satisfactory behavior. In this case a symmetrical distribution, using the central limit theorem [20], with a low CV is considered satisfactory. The more aggregate the planning level, the lower CV is desired.

### **3 Analytical modeling of the aggregation horizon**

To formulate the models, the orders for the whole product family are assumed to come from an unbroken time series. We consider a time period of length  $T$ , where the number  $N(T)$  of orders has mean  $\lambda T$  and standard deviation  $\tau\sqrt{T}$ . In case of a Poisson process we have  $\tau^2 = \lambda$ . To describe the volatility of the order quantities the coefficient of variance (CV) is used (e.g. discussed in Nielsen and Eriksen [13] and

used in Tsubone and Furuta [18]). The size  $Q$  of an order is assumed to have mean  $\mu$  and standard deviation  $\sigma$ , i.e.  $Q$  has coefficient of variation  $CV(Q) = \frac{\sigma}{\mu}$ .

Furthermore, the autocorrelation function of successive order quantities is assumed to be zero except for lag one. The autocorrelation of lag one is denoted  $\rho$ . Let  $Q_1, \dots, Q_{N(T)}$  be the successive orders in the period. The aggregated order quantity is  $= \sum_{i=1}^{N(T)} Q_i$ , which has mean conditional on  $N(T)$  given by  $E(AQ|N(T)) = \mu N(T)$  and variance given by:  $Var(AQ|N(T)) = \sigma^2(N(T) + 2\rho(N(T) - 1)) \approx \sigma^2(1 + 2\rho)N(T)$  under the assumption that  $N(T)$  is sufficiently large. For manufacturing environments with a low number of orders, other methods than aggregating orders are used to estimate demand. The mean of  $AQ$  is then given by  $E(AQ) = \mu\lambda T$ , whereas the variance is (approximated by):  $Var(AQ) = E(Var(AQ|N(T))) + Var(E(AQ|N(T))) = \sigma^2(1 + 2\rho)\lambda T + \mu^2\tau^2 T$

This leaves us with a coefficient of variation given by

$$CV(AQ) = \frac{\sqrt{\sigma^2(1 + 2\rho)\lambda + \mu^2\tau^2}}{\mu\lambda\sqrt{T}}$$

Aiming at a time horizon yielding a coefficient of variation given by  $CV_0$ , this is obtained when

$$T = \frac{\sigma^2(1 + 2\rho)\lambda + \mu^2\tau^2}{(\mu\lambda CV_0)^2} = \frac{CV(Q)^2(1 + 2\rho)\lambda + \tau^2}{(\lambda CV_0)^2}$$

Expressed in terms of the mean number of orders we obtain

$$E(N(T)) = \frac{CV(Q)^2(1 + 2\rho)\lambda + \tau^2}{\lambda(CV_0)^2}$$

and in case of a Poisson process this simplifies to:

$$E(N(T)) = \frac{CV(Q)^2(1 + 2\rho) + 1}{(CV_0)^2}$$

The conclusion is that for any given distribution of  $Q$ , for any given company and period it is (given sufficient observations) possible to estimate  $CV(Q)$  and  $\rho$  (only included if significant autocorrelation is present). From this it is then possible to decide upon a desired stability for the system in the form of  $CV_0$  and calculate the number of order ( $E(N(T))$ ) necessary to aggregate over to achieve this stability.

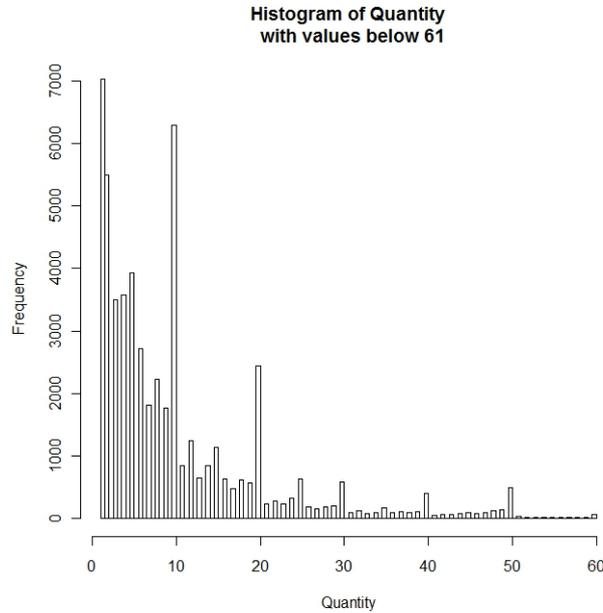
#### 4 A case study application

To illustrate the analysis a test using a set of orders for a product family with variables ItemID, Quantity and DeliveryDate covering a 6 year period was conducted in the statistical analysis tool R [15]. The total number of orders was 54,243 and 2,160 unique products. The data was arranged based on due date, and subsequently sequenced after creation time. This way order lines from the same order would be sequenced to facilitate an investigation of independence on the aggregate set.

First, the distribution of Quantity for the whole product family regardless of ItemID is considered. 99.1% of the order quantities were below 61 and the histogram of these is shown in fig. 1. The overall picture shows an exponential decay with isolated peaks at multiples of 10 and to some extent multiples of 5 and 12. All three tests of symmetry

**Order quantity distributions in make-to-order manufacturing: At what level of aggregation do they respect standard assumptions?** 5

were highly significant with p-values below  $10^{-15}$ . The estimated coefficient of variation was  $CV(Q) = 1.2$ , likewise indicating a very skew distribution.



**Fig. 1.** Histogram of order quantities for the whole data set.

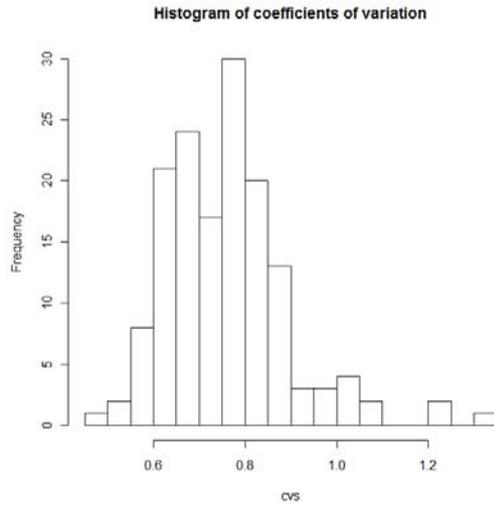
The product specific analyses were limited to the 151 products that had more than 100 orders giving sufficient observations for applying the tests for symmetry. These products only represent 6% of the products, but cover 66% of the total quantity sold. Adopting a 5% significance level, the MGG-test showed that 77.5% of the products had a significant departure from symmetry of the Quantity distribution. The corresponding number for the CM-test was 76.9% and for the Mira-test 67.5%. The conclusion is that on an individual product level there is likewise a skewed distribution of order quantities. Another potential indicator of asymmetry is the CV. Figure 2 shows the histogram of the estimated CV's for the 151 products. Almost all the CV values are above 0.5 indicating significant skewness. For the Quantity time series for the whole product family we have an estimated mean of  $\mu = 10.75$  and an estimated lag one correlation  $\rho = 0.1105$ . As we have no data on arrival times of orders, but rather delivery dates, we assume a Poisson arrival process, i.e. the aggregation level is determined by:

$$E(N(T)) = \frac{CV(Q)^2(1 + 2\rho) + 1}{(CV_0)^2}$$

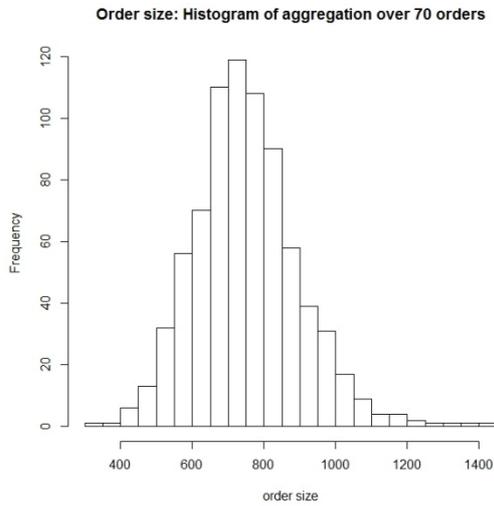
Based on the data estimates, this yields the equation for determining aggregation level in terms of orders given by:

$$E(N(T)) = \frac{2.761}{(CV_0)^2}$$

If e.g. a  $CV_0 = 0.2$  is used, we obtain the integer round up  $E(N(T)) = 70$ . Fig. 3 display the histograms of order quantities after aggregation over 70 orders. The distribution is clearly much closer to symmetry. The skewness is clearly much lower as reflected by the estimated coefficient of variation, which is 0.19.



**Fig. 2.** Histogram of order quantity CV's for 151 products.



**Fig. 3.** Histogram of order quantities, when 31 orders are consecutively aggregated.

The slightly skewed distribution, shown in figure 3, clearly indicates that if an aggregation horizon of 70 orders is used on the aggregate set, the order quantities tend to flow to a symmetrical distribution.

## 5 Conclusions and further research

The analysis and subsequent application of the analytical model gives a number of significant conclusions. First, the behavior of the particular case data set is as expected rather than as assumed in literature. The individual orders are, when the whole data set is considered, neither symmetrically nor independently distributed. Second, the order quantities for the products with sufficient orders are likewise to a large extent not symmetrically distributed. Third, the analytical models allows to calculate the number of orders needed to aggregate to achieve a given CV. Aggregating 70 orders to achieve a CV of approximately 0.2 on the aggregate set gives a distribution that is almost symmetrical. The conclusion is that although the individual customer orders are not behaving as expected, it is possible to aggregate over the order time series so that demand can be considered to stem from a symmetrical distribution. However, this indicates that demand can only be considered to stem from a symmetrical distribution with a sufficiently low CV if orders are aggregated. This is especially challenging in a make-to-order environment where the ability to buffer on time will be significantly influenced by market conditions. On a product family level this may not be a problem (in this example a desired CV of 0.1 on aggregate level would require aggregating 277 orders, 0.5% of the total number of orders, c. 2 weeks of orders). However, on an individual product level with high CV's aggregating even 10 orders indicates the need to aggregate over a very large time period indeed. This supports the assumptions in e.g. Fliedner [7] and Silver et al. [16] regarding the stability of aggregate demand. Future work will focus on simulating the impact of different order quantity distributions on simple manufacturing systems and developing methods to translate the stability of order quantities into a planning horizon for a given manufacturing system.

## 6 References

1. Axäster, S., Aggregation of Product Data for Hierarchical Production Planning, *Operations Research*, 29, pp. 744-756 (1981)
2. Becker, C., Scholl, A., A survey on problems and methods in generalized assembly line balancing, *European Journal of Operational Research*, 168, pp. 694-715 (2006)
3. Box, G. E. P., Jenkins, G., *Time Series Analysis: Forecasting and Control*, Holden-Day (1976)
4. Brander, P., Levén, E., Segerstedt, A., Lot sizes in a capacity constrained facility-a simulation study of stationary stochastic demand, *International Journal of Production Economics*, 93-94 pp. 375-386 (2005)
5. Cabilio, P. and Masaro, J., A simple test of symmetry about an unknown median. *The Canadian Journal of Statistics*, 24, pp. 349-361 (1996)
6. Chen A., Blue J., Performance analysis of demand planning approaches for aggregating, forecasting and disaggregating interrelated demands, *International Journal of Production Economics*, 128, p. 586-602 (2010)
7. Fliedner, G.: Hierarchical forecasting: issues and use guidelines, In: *Industrial Management & Data Systems*, 101, pp. 5-12 (2001)

8. Hax, A. C., and H. C. Meal, Hierarchical Integration of Production Planning and Scheduling, in *Logistics*, edited by M. A. Geisler, vol. 1 of Studies in the Management Sciences, North-Holland/American Elsevier (1975)
9. Hendry, L. C., and Kingsman, B. G., Production planning systems and their applicability to make-to-order companies. *European Journal of Operational Research*, 40, 1-15 (1989)
10. Marucheck, A. S., and McClelland, M. K., Strategic issues in make-to-order manufacturing. *Production and Inventory Management*, 27, 82-96 (1986)
11. Miao, W., Gel, Y. R., and Gastwirth, J. L., A New Test of Symmetry about an Unknown Median. Random Walk, *Sequential Analysis and Related Topics - A Festschrift in Honor of Yuan-Shih Chow*. Eds.: Agnes Hsiung, Cun-Hui Zhang, and Zhiliang Ying, World Scientific Publisher, Singapore (2006)
12. Mira, A., Distribution-free test for symmetry based on Bonferroni's measure. *Journal of Applied Statistics*, 26, pp. 959-972 (1999)
13. Nicholas, J.: *Competitive Manufacturing Management: Continuous Improvement. Lean Production and Customer-focused Quality*. New York: McGrawHill (1998)
14. Nielsen, P. and Eriksen, T., Towards an analysis methodology for identifying root causes of poor delivery performance, *Conference Proceedings of IESM'2011, International Conference on Industrial Engineering and Systems Management* (2011)
15. R-project (webpage) [www.r-project.org](http://www.r-project.org) (2012)
16. Silver, E.A., Pyke, D.F., Peterson, R.: *Inventory Management and Production Planning and Scheduling*, third ed., John Wiley & Sons (1998)
17. Theil, H., *Linear Aggregation of Economic Relations*, first ed., North-Holland Publishing Company (1955)
18. Tsubone H., Furuta H. Replanning timing in hierarchical production planning, *International Journal of Production Economics*, 44, pp. 53-61 (1996)
19. Vollmann, W., Berry D., Whybark T.E., Jacobs F.: *Manufacturing Planning and Control for Supply Chain Management*, McGraw-Hill, Singapore (2005)
20. Walpole, R. E., R. H. Myers, S. L. Myers, and K. Ye, *Probability & Statistics for Engineers & Scientists*, Prentice Hall, New York (2002)
21. Wijngaard, J., On Aggregation in Production Planning, *Engineering Costs and Production Economics*, 6, pp. 259-66 (1982)
22. Zotteri, G., M. Kalchschmidt, and F. Caniato, The impact of aggregation level on forecasting performance, *International Journal of Production Economics*, 93-94, pp. 479-491 (2005)