Identification of the Demand Curve and Forecasts in Subsequent Periods Using the Metropolis-Hastings Algorithm

Łukasz Gołąbek[1], Konrad Gauda[2], Krzysztof Zuk[3], Edward Kozłowski[4]

Abstract:

Purpose: The main purpose of the article is to identify the demand curve and to forecast demand in subsequent periods using the Metropolis-Hastings algorithm.
Design/Methodology/Approach: The Metropolis-Hastings algorithm belonging to the Markov Chain Monte Carlo was used to identify the demand curve and to forecast the demand in subsequent periods. This method consists in generating (drawing) a sample in accordance with the modified distribution and the possibility of rejecting a new sample in case of insufficient improvement of the quality index.
Findings: The results of the conducted research indicate that the presented solution of generating a sample in accordance with the modified distribution and the possibility of rejecting a new sample in the event of insufficient improvement of the quality index is effective in identifying and forecasting the demand.
Practical Implications: The algorithm presented in the article can be used to forecast stays taking into account the product life curve.
Originality/Value: A novelty is the use of the Metropolis-Hastings algorithm to identify the demand curve and the forecast of demand in subsequent periods to determine the strategy of long-term products by analyzing the sales volume of the product.

Keywords: Machine learning, Markov Chain Monte Carlo, Metropolis-Hastings algorithm, forecasting.

JEL codes: C50, C53, O10.

Paper type: Research article.

[1]*Corresponding Author: Netrix Group sp. z o.o. , Lublin, Poland,*
e-mail: lukasz.golabek@netrix.com.pl
[2]*University of Economics and Innovation in Lublin, Lublin, Poland,*
e-mail: konrad.gauda@wsei.lublin.pl
[3]*University of Economics and Innovation in Lublin, Lublin, Poland,*
e-mail: krzysztof.zuk@wsei.lublin.pl
[4]*Faculty of Management, Lublin University of Technology, Lublin, Poland,*
e-mail: e.kozlovski@pollub.pl

1. Introduction

Forecasting future demand is essential to making supply chain decisions. Undoubtedly, historical information about demand can be used to forecast future demand and such analysis affects the functioning of the entire supply chain. Demand forecasts form the basis of all supply chain planning. Consider the division of processes performed in the push / pull supply chain (Chopra and Meindl, 2007). Pull processes are initiated by an order placed by the customer, while push processes are initiated and performed while waiting for the order. When considering these two options for push processes, you should plan your level of activity, be it production, transportation, or any other planned activity.

With pull processes, you need to plan the level of available bandwidth and inventory, but not the actual amount of products to be made. In both cases, the first step to be carried out is to forecast customer demand. In a way, one might be tempted to treat demand forecasting as a kind of art used so as not to leave everything to chance. What the company knows about the previous behavior of its customers is reflected in their future behavior. Demand is influenced by various factors and can be predicted with at least some probability if the firm can identify the relationship between these factors and future demand.

To forecast demand, firms must first identify the factors that affect future demand and then establish a relationship between those factors and future demand. When forecasting demand, enterprises must maintain a balance between objective and subjective factors. In the presented considerations we focus on the methods of quantitative forecasting, however, it should be remembered that companies must take into account human input when preparing the final forecast. The supply chain can experience significant benefits from improving demand forecasting due to qualitative human factors. The company must have knowledge of many factors that are related to the forecasting of demand.

The problem faced by economists, managers, and people responsible for defining long-term strategies is solving the task, where by analyzing the volume of sales/orders for a product in the initial phases, both the demand for this product in subsequent periods and the market absorption (total sales volume on the market) should be predicted. The amount of demand depends on the phases according to the product life cycle (product creation, launch (development), growth, maturity (market saturation), decline) (Mruk and Rutkowski, 1999). The Metropolis-Hastings algorithm (Geyer, 2011; Gamerman and Lopes, 2006; McElreath, 2020) belonging to MCMC (Markov Chain Monte Carlo) was used to identify the demand curve and forecast the demand in subsequent periods.

The sample is generated according to the modified distribution and the possibility of rejecting a new sample in case of insufficient improvement of the quality index.

2. Methodology

2.1 Monte Carlo Sampling with Markov Chains

Monte Carlo sampling with Markov chains (MCMC) (Gamerman and Lopes, 2006; Geyer, 2011) is a class of sampling algorithms using the Monte Carlo method for a certain probability distribution. As a result of sampling, we obtain a Markov chain that has a distribution similar (convergent) to the distribution we are looking for. This technique is usually used to sample complex probability distributions. The greater the number of samples, the more accurately the sample distribution corresponds to the desired distribution. With MCMC, we create sequence as samples of a random variable (s) that can be used to evaluate the objective function against that variable. These chains are stochastic processes that move randomly according to an algorithm. Unlike the Monte Carlo method (Monte Carlo random simulation), the samples used in MCMC are correlated.

The main problem of multivariate systems depends on identification of unknown parameters where the number of observations needed to estimate increases exponentially with the increase in the number of dimensions (Geyer, 2011). Monte Carlo methods using Markov chains are largely immune to the dimensional problem. These methods do not require an analytical solution, nor do they search the entire problem space (set of possible solutions), they rely on iterated sampling. Each consecutive drawn sample focuses (evaluates) more and more precisely around the areas of the distribution with the highest probability.

One of the MCMC algorithms is the Metropolis-Hastings algorithm (Geyer, 2011; Gamerman and Lopes, 2006; McElreath, 2020). This method consists in generating (drawing) a sample in accordance with the modified distribution and the possibility of rejecting a new sample in case of insufficient improvement of the quality index.

2.2 Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm generates a series of samples, where the standard deviation of the samples decreases with increasing number of the sequence. The empirical distribution determined on the basis of the selected samples tends to the distribution of the searched $P(x)$ (Geyer, 2011; Gamerman and Lopes, 2006). The elements of the sequence of samples $\{x_t\}_{t \geq 0}$ are formed iteratively, where the distribution of the next sample depends only on the current value of the sample (according to the Markov property, the future does not depend on the past when the present is known).

In particular, at each step $t \geq 1$ we randomize a sample x' distribution which depends on the last value in the chain (the value obtained in the preceding step), i.e. $g(x'|x_{t-1})$. Then, with a certain probability, the drawn sample is either accepted or rejected. In case of rejection, the value from the preceding step is taken as the

current value of the sample, i.e. $x_t = x_{t-1}$ (and the current value is reused in the next iteration). In case of acceptance we take $x_t = x'$. The acceptance probability is determined by comparing the function value for the drawn sample with the function value for the last element in the chain (compare $f(x')$ with $f(x_t)$).

Algorithm:

1. At the time $t = 0$, we draw the sample x_0 for the apiori distribution $P_0(x)$, which represents the initial state in the Markov chain. We define the transition probability function $g(x|y)$ from state y to state x. Additionally, we define the number of iterations n.
2. Let $t = t + 1$.
3. Randomly select a candidate x' for the transition probability distribution $g(x'|x_{t-1})$.
4. We determine the value of the function $f(x')$ (discussed below in this example for the identification of the product life cycle is selected as a function of the likelihood function).
5. We calculate the acceptance rate $\alpha = \dfrac{f(x')}{f(x_{t-1})}$.
6. Randomly select a number u from the uniform distribution on $[0,1]$.
7. Another element in the Markov chain
8. $x_t = \begin{cases} x', & \text{if } u \le \alpha \\ x_{t-1}, & \text{if } u > \alpha. \end{cases}$
9. If $t < n$, then we go back to point 2, otherwise we stop the algorithm.

2.3 Identification of Demand Size

Product lifecycle is the period in which the product is present on the market. This cycle consists of four phases:

- the launch (of the manufacturer of the goods primarily consist of informing customers about the appearance of the product on the market);
- increase in sales (in the growth phase, the fastest increase in sales occurs, which reduces the unit production costs);
- saturation/maturity (sales growth is weaker, demand mainly from regular users of the product);
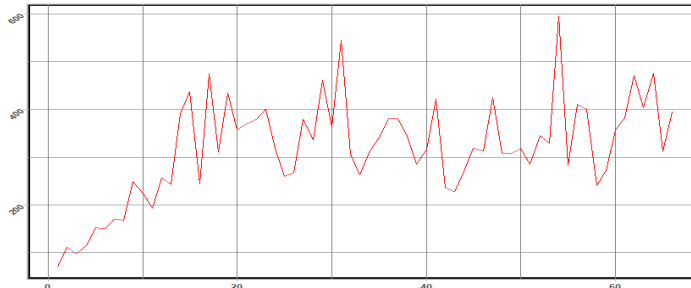- decrease in sales (demand for the product decreases).

The product life cycle curve is a curve that represents the dependence of the quantity of demand on time. The Y axis shows the quantities of demand, and the X axis the moments. Knowing the product life cycle, you can easily make the right decisions by the company at any time (invest in advertising, match production lines to the demand, look for additional recipients, lower the price for the product, and finally withdraw the product from the market). Of course, having the entire history of data, you can determine the curve corresponding to the product life cycle.

Undoubtedly, a significant problem from the economic point of view is to predict a change in demand over time and to make predictions of demand in the future. Therefore, an important problem is the identification of the product life cycle curve based on the information available in the initial phases. The Metropolis-Hastings algorithm will be used to solve such a task.

3. Problem Definition

The problem faced by people responsible for defining long-term strategies is solving the task, where by analyzing the volume of sales/orders for a product in the initial phases, both the demand for this product in subsequent periods and the market absorption (total sales volume on the market) should be predicted. The Metropolis-Hastings algorithm (Gamerman and Lopes, 2006; McElreath, 2020) have been used below to identify the demand curve and forecast the demand in subsequent periods. The size of the demand depends on the phases according to the product life cycle - product creation, introduction to the market (development), growth, maturity (market saturation), decline, e.g., the growth and maturity phases are shown in Figure 1.
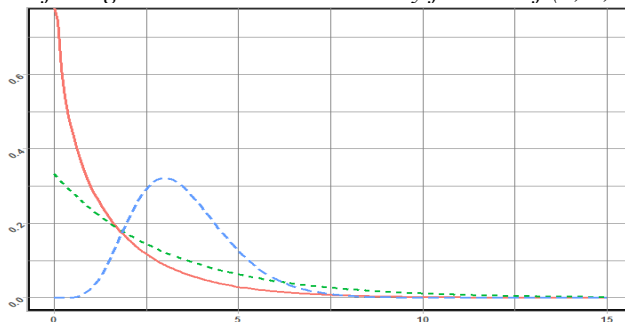
Figure 1. *Demand for the product over time*

Source: Own creation.

Let $(\Omega, \mathcal{F}, P)$ be a probabilistic space and the random variable τ has a gamma distribution (Ross, 1997).

Figure 2. *Graph of the gamma distribution density function f (x, b, c)*

Source: Own creation.

Density function:

$$f(t, b, c) = \begin{cases} \dfrac{t^{b-1}\exp\left(-\dfrac{t}{c}\right)}{c^b \Gamma(b)}, & t \geq 0, \\ 0, & t < 0. \end{cases} \tag{1}$$

and

$$\Gamma(s) = \int_0^{\infty} x^{s-1} \exp(-x)\, dt$$

the parameter $b > 0$ represents the shape parameter, while $c > 0$- the scale parameter.

The distribution of the random variable:

$$F(t) = P(\tau < t) = \int_0^t f(s, b, c)\, ds \tag{2}$$

and $F(\infty) = 1$. Figure 2 shows the values of the gamma distribution density function depending on the magnitude of b and c.

Our task is to estimate the total number of demand for a product $a > 0$ by observing the demand $\{y_i\}_{1 \leq i \leq n}$ in successive periods, where $y_i = \Delta x_i = x_i - x_{i-1}$ and x_i means the demand cumulated up to the moment i and $x_0 = 0$. According to the above notation $\lim_{n \to \infty} x_n = a$.

Let the random variable τ denote the life of the product (e.g. estimating the production for the next 5 years, we do not have a guarantee that during this time the market will not enter a new type of product and sales will decrease drastically). If x_t of the product has been sold by the time t and the total market demand (absorption) for the product is a, then we can assume:

$$P(\tau < t) = \frac{x_t}{a} \tag{3}$$

If we assume $F_a(t) = aF(t)$. Thus, the demand in the period $(t - 1, t]$ is modeled using the equation of state:

$$y_t = af(t, b, c) + \varepsilon_t \tag{4}$$

where $\{\varepsilon_t\}_{t \geq 1}$ is a sequence of independent random variables with the distribution $N(0, \sigma^2)$ for $t > 0$ and the function $f()$ is given by formula (1).

Based on the sequence describing the volume of demand $\{y_i\}_{1 \leq i \leq n}$ to determine the parameters a, b, c in models (4), we use the LSM (Lieberman and Hillier, 1990; Chow, 1995; Wehrens, 2011; Hastie, Tibshirani and Friedman, 2009; Walesiak and Gatnar, 2009). For this, we solve the task:

$$\min_{a,b,c \geq 0} \sum_{t=1}^{n} (y_t - af(t,b,c))^2 \tag{5}$$

the estimated parameter $\hat{a}$ corresponds to the projected total demand for the product, while $\hat{b}$ and $\hat{c}$ are estimators of the gamma distribution (shape and scale respectively).

If we want to predict the volume of demand in the next *k* periods, we estimate as:

$$\sum_{t=n+1}^{n+k} \hat{a} f\left(t, \hat{b}, \hat{c}\right)$$

In order to accurately evaluate the parameters $\hat{a}, \hat{b}, \hat{c}$, the vector of start parameters should be given. We will use the Metriopolis-Hastings algorithm to determine the starting values.

4. Implementation the Metropolis-Hastings Algorithm

From equation (4) we have:

$$E(y_t) = af(t,b,c) \tag{6}$$

$$Var(y_t) = E\left(y_t - E(y_t)\right)^2 = \sigma^2 \tag{7}$$

From (4) - (7) we have that for the set parameters a, b, c, σ the random variable y_t has a normal distribution $N(af(t,b,c), \sigma^2)$. For fixed a, b, c, σ we create a function:

$$L(\{y_t\}_{1 \leq t \leq n} | a, b, c, \sigma) = \prod_{t=1}^{n} \gamma(y_t, af(t,b,c), \sigma) \tag{8}$$

where $\gamma(x, m, \sigma)$ is the density function of the distribution $N(m, \sigma^2)$

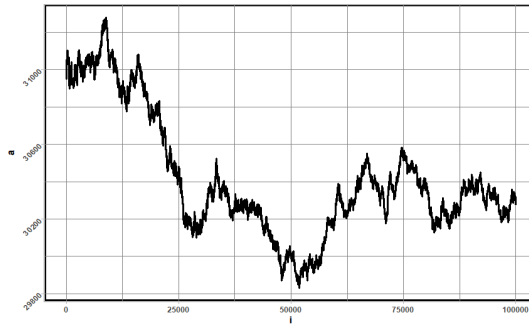$$\gamma(x, m, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(x-m)^2}{2\sigma^2}\right).$$

The likelihood function value (Hastie et al., 2009), (Chow, 1995), (McElreath, 2020) for the sequence $\{y_t\}_{1 \leq t \leq n}$ is defined as follows:

$$L(\{y_t\}_{1 \leq t \leq n}) = L(\{y_t\}_{1 \leq t \leq n} | a, b, c, \sigma) P(a, b, c, \sigma) \tag{9}$$

where $P(a, b, c, \sigma)$ denotes a piori distribution for parameters a, b, c, σ. Below, we assume stochastic independence for the above - mentioned parameters, therefore:

$$P(a, b, c, \sigma) = P_a(a) P_b(b) P_c(c) P_\sigma(\sigma) \tag{10}$$

Figure 3. *Sequence of values for parameter a*

Source: *Own creation.*

From (9) - (10) the logarithm of the likelihood function:

$$\ln L(\{y_t\}_{1 \leq t \leq n}) = -\frac{n}{2} \ln(2\pi) - n \ln(\sigma) + \sum_{t-1}^{n} \frac{(y_t - m)^2}{2\sigma^2}$$
$$+ \ln P_a(a) + \ln P_b(b) + \ln P_c(c) + \ln P_\sigma(\sigma) \tag{11}$$
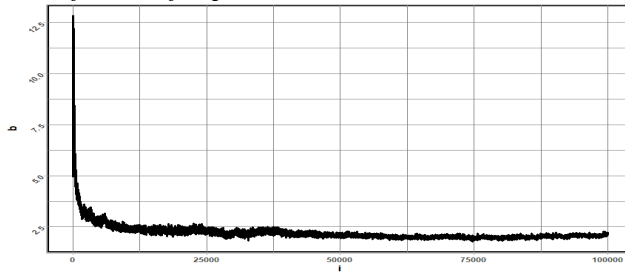
Figures 3 - 6 for the determination of Markov chains using Monte Carlo simulation for the identification of the model parameters (5) are presented below. The last 20,000 elements from each of the sequence were selected as values for the unknown parameters.

Table 1. *Basic statistics for unknown parameters*

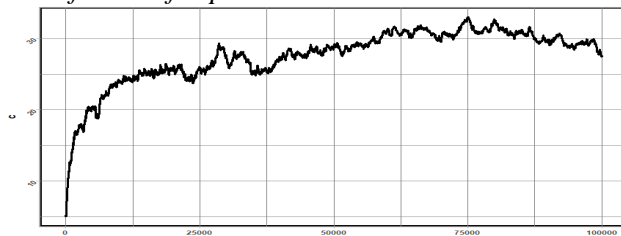|	a	b	c	σ
Average value	30285.38964	2.0270419	29.9951382	75.424009
Standard deviation	73.17689	0.0663256	0.9532362	3.975294
Median	30286.76535	2.0291391	29.8442083	74.191500

Source: *Own creation.*

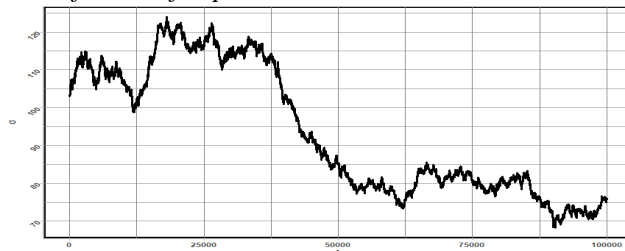**Figure 4.** *Sequence of values for parameter b*

Source: *Own creation.*

Figure 5. *Sequence of values for parameter* c
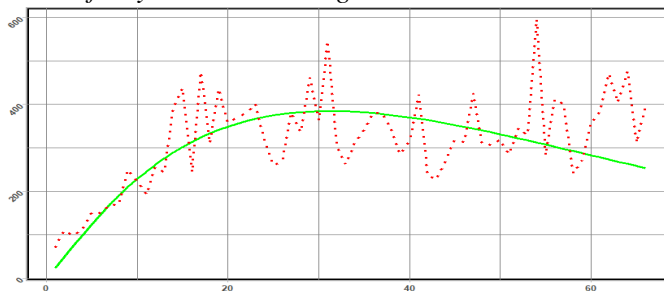

Source: *Own creation.*

Figure 6. *Sequence of values for parameter σ*
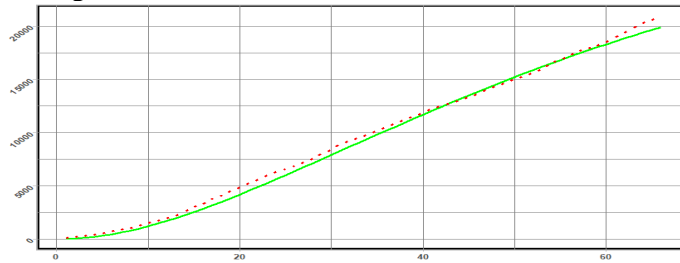

Source: *Own creation.*

The figures below show the fit of the product life cycle curve (demand versus period) and the cumulative demand curve (Figure 7 and Figure 8).

Figure 7. *Product Life Cycle Curve Fitting*

Source: *Own creation.*

Figure 8. Matching total demand

Source: *Own creation.*

5. Conclusions

The article presents the Metropolis-Hastings algorithm of identification for the demand curve and the demand forecast. The method consists in generating a sample in accordance with the modified distribution and the possibility of rejecting a new sample in case of insufficient improvement of the quality index. The volume of demand for the product in subsequent periods as well as the market absorption capacity were estimated as the total number of demand for the product by observing the demand in subsequent periods.

Knowing the product life cycle, it is possible to make appropriate decisions by the company, but having the history of data we can determine the curve corresponding to the product life cycle. A significant problem from the economic point of view is forecasting a change in demand over time and making predictions of demand in the future. The main problem of identification of multidimensional systems i that when the number of dimensions increases, the number of observations needed to estimate the unknown parameters increases exponentially. The applied Monte Carlo method using Markov chains is largely resistant to the problem of dimensionality, does not require an analytical solution and does not search the entire problem space, rely on iterated sampling.

References:

Chopra, S., Meindl, P. 2007. Supply chain management : strategy, planning, and operation. Pearson Prentice Hall.

Chow, G.C. 1995. Ekonometria. Wydaw. Naukowe PWN.

Gamerman, D., Lopes, H.F. 2006. Markov chain monte carlo: Stochastic simulation for bayesian inference. CRC Press.

Geyer, C.J. 2011. Introduction to markov chain monte carlo. Handbook of Markov Chain Monte Carlo. CRC Press.

Hastie, T., Tibshirani, R., Friedman, J. 2009. The elements of statistical learning. Springer-Verlag New York Inc.

Lieberman, F.S.H.G.J., Hillier, F. 1990. Introduction to stochastic models in operations research. USA: McGraw-Hill Publishing Company.

McElreath, R. 2020. Statistical rethinking: A bayesian course with examples in r and stan. CRC press.

Mruk, H., Rutkowski, I.P. 1999. Strategia produktu. Polskie Wydawnictwo Ekonomiczne.

Ross, S. 1997. Introduction to probability models. Academic Press, Sixth Edition.

Walesiak, M., Gatnar, E. 2009. Statystyczna analiza danych z wykorzystaniem programu r. Wydawnictwo Naukowe PWN.

Wehrens, R. 2011. Chemometrics with r. Springer-Verlag GmbH.