DEMAND MODELING FOR THE OPTIMIZATION OF SIZE RANGES

Ivo Malakov, Tzanko Georgiev, Velizar Zaharinov ,
Alexander Tzokev, Velislav Tzenov

TU-Sofia, blvd. Kliment Ohridski 8, Sofia 1000, Bulgaria

Abstract

This article deals with an approach for demand modeling when designing size ranges of technical products. The proposed modeling approach is a stage of the size ranges optimization problem, and by it information can be obtained regarding the requested sizes and their respective quantities (demand). The approach is comprised of four main stages - market segmentation and choice of potential customers, collecting demand data, description and processing of the collected data, and determination of a demand model. The approach is applied for the demand modeling of a particular product – pneumatic modules for linear motion. In the analysis of the market research data the procedures of principle component analysis, factor analysis and regression by principal components are used. The study is carried out in the environment of STATGRAPHICS and SPSS.

Keyword: design; optimization; size ranges; demand modeling; linear motion modules; factor analysis; regression analysis

This Publication has to be referred as: Malakov, I[vo]; Georgiev, T[zanko]; Zaharinov, V[elizar]; Tzokev, A[lexander] & Tzenov, V[elislav] (2016). Demand Modeling for the Optimization of Size Ranges, Proceedings of the 26th DAAAM International Symposium, pp.0435-0444, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-07-5, ISSN 1726-9679, Vienna, Austria
DOI:10.2507/26th.daaam.proceedings.058

1. Introduction

Demand forecasting is an important problem for size ranges optimization, since the choice of the elements included in the size ranges, significantly depend on this forecasting [10]. The collection of great amount of data is required for solving the said problem. By applying appropriate methods and procedures on the collected data, information has to be obtained regarding the customers' demand for technical products with certain values of their main parameters.

The use of conventional methods for demand forecasting [5], [6], [7], [18] is difficult in the case of new products for which there is no information available regarding sales and customer inquiries for previous time periods. Usually the possibility for using information from competitive companies, producing similar products, is also questionable. In this case, the solving of the problem is accompanied by a number of issues, stemming from the high level of uncertainty and the related obscure questions about the necessary and sufficient measure of the calculations' precision when the initial conditions are inexact and incomplete.

In the known publications, related to the optimal design of size ranges, a relatively small amount of attention is paid to the problem of building a demand model and determination of needs. In most of the publications a predefined demand function is used with no information of how it is obtained [2], [9], and in others this function is given in implicit form by assuming even distribution of demand [1], [11], [12], or a limited number of demand distribution models are used [4].

The aim of this article is to propose an approach for demand modeling based on incomplete information regarding the demand for products with certain values of their main parameters - elements of size ranges.

Nomenclature

$N_1, N_2, \dots, N_n$	random variables
$f_1, f_2, \dots, f_k$	random variables
N	vector with elements $N_1, N_2, \dots, N_n$
K	matrix with elements $K_{ij}, (i = 1, n; j = 1, k)$
F	vector with elements $f_1, f_2, \dots, f_k$
$e_1, e_2, \dots, e_n$	residuals
E	vector of the residuals $(e_1, e_2, \dots, e_n)$
G	load capacity
L	stroke length

2. Approach for demand modeling and determination of needs

The demand modeling approach for technical products, which are elements of size ranges, includes the following phases:

- Phase 1. Determination and segmentation of the market.
- Phase 2. Collecting of demand data.
- Phase 3. Processing of the collected data.
- Phase 4. Determination of a demand model.

2.1. Phase 1. Determination and segmentation of the market

In this phase the potential customers for the developed size range are determined, i.e. the market in which it will be offered. As the markets are composed of customers with different requirements for the size range, it is useful to make a grouping of the customers according to different classification characteristics (product requirements, behavioral models, demographic characteristics, etc.) and to determine target market(s). To the problems and tasks that are solved in this phase are dedicated a number of current developments [3], [8], [13].

2.2. Phase 2. Collecting of demand data

For collecting demand data two types of methods can be used: direct and indirect [8]. With direct methods the information is directly obtained from the users and includes market research, sales data including e-trade, customer inquiries, previous experience, market experiments, information systems for consumption monitoring, etc. Indirect

methods include information gathered from specialized literature, annual financial reports, market/financial reports, client's point of view, modeling of customer's choice, etc.

For market research it is necessary to determine what part of the customers will be interviewed or inquired, i.e. to determine a representative sample of potential users, as, in general, it is impossible or unsuitable to obtain information about the needs of all customers in the chosen target market.

The choice of method/s for obtaining the needed information mainly depends on the available resources and time, and also on the experience of the researchers. For every particular case the development of special methods for demand studying is needed. These methods have to take into account the specifics of the solved problem.

2.3. Phase 3. Processing of the collected data

In this phase the set of random variables $N_1, N_2, \dots, N_n$ describing the requests for products with certain values of their main parameters is studied. The probability distributions of the main factors and data (collected requests) are determined. The missing values are registered in the data.

For the purpose of the study it is assumed that the random variables $N_1, N_2, \dots, N_n$ have a covariance or correlation matrix. It is assumed that the mathematical expectations of the set of random quantities are equal to zero. The following hypotheses can be formulated regarding the structure of the set of random variables:

A. Supposed that exist random variables $f_1, f_2, \dots, f_k$, such that fulfill the following condition:

$$N = KF$$

B. Supposed that exist random variables $e_1, e_2, \dots, e_n$, such that the set of independent (with respect to $e_1, e_2, \dots, e_n$) random variables $N_1 - e_1, N_2 - e_2, \dots, N_n - e_n$ is with dimensionality $k < n$, i.e. there exist random variables $f_1, f_2, \dots, f_k$ such that fulfill the following condition:

$$N - E = KF$$

There are two cases subject to hypotheses:

A₁. The random quantities $f_1, f_2, \dots, f_k$ are given and known.

B₁. The random quantities $f_1, f_2, \dots, f_k$ are unknown.

The set of random variables $N_1, N_2, \dots, N_n$ describing the requests for the investigated problem contain missing values. The hypotheses according to the obtained data are shown in Table 1.

	Hypothesis	A	B
Case			
A ₁		Regression	analysis
B ₁		Component analysis	Factor analysis

Table 1. Main hypotheses according to the obtained data

The following comments can be made:

- Principle component analysis is always associated with a linear model and dispersion study;
- Factor analysis allows a linear hypothesis and studies covariance or correlation.

2.4. Phase 4. Determination of a demand model

The determination of demand is carried out in the following sequence:

1. Principle component analysis in the data.
2. Factor analysis.
3. Modelling of the obtained requests by regression models of the principal components.
4. Analysis of the missing values.

On the basis of the regression models, in terms of the obtained requests, the demand for products that are elements of size ranges is forecasted. An analysis of the missing values can also lead to a possible demand forecast.

3. Application of the approach

The proposed approach is applied for demand modelling of linear modules with pneumatic actuation for building positioning systems for the automation of manufacturing processes. The target market is a group of companies in the field of electric equipment manufacturing.

The linear modules with pneumatic actuation are characterized with a number of parameters: stroke length, m; load capacity (maximum effective load, force), kg (N); number of positions; allowable loads and moments; maximum/minimum speed, m/s; acceleration, m/s²; impact energy/maximum impact energy in the end positions, J; accuracy parameters; compressed air consumption/cycle, l (m³); mass of the module, mass of the moving part, kg; reliability indicators; noise level for a given pressure and mass of the moved load, dB; average service life, years; etc.

The load capacity G and stroke length L are chosen as main parameters. The sets of their possible values are determined according to which the corresponding demand for modules is sought:

$$G = \{0,32; 0,4; 0,5; 0,63; 0,8; 1; 1,25; 1,62; 2,5; 3,2; 4; 5; 6,3\}$$

$$L = \{0,02; 0,025; 0,04; 0,05; 0,06; 0,08; 0,1; 0,16; 0,2; 0,25; 0,32; 0,4; 0,45\}$$

For collecting the input information a methodology comprised of five main steps is used:

Step 1. For the totality of preliminary defined objects for automation (parts, assemblies) particular functional schemes are developed. The latter describe manipulation operations performed with the objects in the corresponding workstations to be automated. On that basis, similarly to the ideas of group technology, a generalized functional scheme including all particular schemes is composed.

Step 2. Development of complex system structure for the manipulation system. The structure performs the generalized functional scheme while taking into account the possibilities for concentration and differentiation of the partial functions executed by the manipulation modules.

Step 3. Determination of the possible values for every main parameter of the modules. The values are regulated by norms and standards. If the latter are absent the parameters' values are chosen from preferred numbers.

Step 4. Determination of the functional relationships needed for calculating the values of the chosen main parameters of the modules that are included in the manipulation system's complex structure.

Step 5. For all objects that are going to be manipulated and can be grouped in advance, expedient assemblies for the manipulation system are determined by editing the complex structure. The principle for minimum degrees of freedom has to be observed while determining the particular functional scheme. The values of linear and angular displacements of the manipulation modules are determined in relationship to the manipulated object's characteristics; the characteristics of the technological processes and operations that are automated; the characteristics of the environment – dimensions, shape and location of the main technological equipment's working zone; the number of served workstations, their relative layout, etc. Next a formalized description of the information follows. For every object, data regarding the particular manipulation system are presented in an information card that is suitable for electronic processing.

Step 6. Processing of the input information. After analysis of the objects that are going to be manipulated, and that are defined by the customers, a complex structure of a manipulation system is developed (Fig. 1). For each object are determined the appropriate manipulation system's assembly and the necessary values for the linear and angular displacements of the modules that build it. The data are inputted in information cards that are processed with a software application. The obtained results from the market research regarding demand for linear modules with particular load capacity and stroke length are shown in Fig. 2.

Fig. 1. Complex structure of the manipulation system 1 – module for regional rotation; 2 – module for regional vertical translation; 3 – module for regional horizontal translation; 4 – module for local translation; 5 – module for local rotation; 6 – gripping module

Fig. 2. Real data from market research

Fig. 3. (a) Histogram of the random quantity G without transformation. Exponential Distribution, mean = 0,165769, P-Value = 0,940771; (b) Histogram of the random quantity G with transformation Normal, mean = 0,573212, standard deviation = 1,01171, P-Value=0,9885

For the analysis of the real data are applied procedures of principle component analysis, factor analysis and regression by the principal components [14], [17]. The analysis is done in the environment of STATGRAPHICS and SPSS [15], [16].

First the distributions of the random quantities are studied. Two main factors are considered – stroke length L , load capacity G and the number of requests for each factor pair. As a result it is established that the random quantity G is exponentially distributed with a mean value of 2,75. The histogram is shown in Fig. 3a. If a natural logarithm transformation is used the results shown in Fig. 3b are obtained.

The random quantity L has the following characteristics – without transformation it is exponentially distributed with a mean value of 0,166. The distribution is shown in Fig. 5a. The transformation natural logarithm on the random quantity leads to normal distribution (Fig. 5b).

At first the requests (demand) are considered without the missing values. Birnbaum-Saunders distribution is found for criterion maximum likelihood. The obtained evaluation by χ^2 Chi-Square = 0,81016 with 2 d.f. P-Value = 0,666923. In Fig. 5 the distribution of requests is shown.

Fig. 4. (a) Histogram of the random quantity L without transformation Exponential Distribution, mean = 0,165769, P-Value = 0,940771; (b) Histogram of the random quantity L with transformation Normal, mean = -2,24152, standard deviation = 1,04771, P-Value=0,92943

Fig. 5. Requests' distribution

Component number	Eigenvalue	Percent of variance	Cumulative percentage
1	724,833	73,968	73,968
2	146,677	14,968	88,936
3	72,6975	7,419	96,355
4	20,8567	2,128	98,483
5	6,86321	0,700	99,184
6	5,0628	0,517	99,700
7	1,87	0,191	99,891
8	0,926513	0,095	99,986
9	0,130064	0,013	99,999
10	0,0117145	0,001	100,000

Table 2. Principal components analysis

From the performed analysis so far the following conclusions can be made:

1. The distributions of the two main factors are confirmed with a high P-value.
2. The found transformation leads to normal distribution which is a necessary and sufficient condition for data analysis.
3. The Birnbaum-Saunders distribution has a “hidden” normal structure.

A component analysis of the data is performed under the conditions of hypothesis A and condition B1. Part of them are linearly dependent and are not considered. The following eigenvalues have been found (shown in Table 2).

Threshold constant six is chosen, which allows for revealing of 99,7% of the sample's variance. In practice two components can also be chosen for data analysis, because the first two components reveal 88,936% of the variance (see Table 2).

In Fig 6a is shown a graphical representation of the principal components. The relationship between the components and the variables is shown in Fig. 6b. The obtained results suggest the necessity for studying three factors. The problem for determining the principal components is solved as a first step of factor analysis (Table 3 - factor loading matrix after quartimax rotation).

In this case, three factors have been extracted, since three factors had eigenvalues greater than or equal to threshold constant (30,0). Together they account for 96,3547% of the variability in the original data. Since the principal components method is used, the initial communality estimates have been set to assume that all of the variability in the data is due to common factors. The relationship between the factors and the variables is shown in Fig. 7a.

Hypothesis A and B are described by regression analysis when condition A₁ is satisfied. The principal components are appropriate predictors of the real data.

Fig. 6. (a) Principal components; (b) Share of the variables in the corresponding components

In the role of predictors six principal components can be used. The data are presented in a matrix in which rows is the load capacity, and in which columns is the stroke length. The regression models are formed by a dependent variable which is a chosen column, and independent variables – the principal components.

Depending on the P-value of the coefficients for the individual models the predictors' number is different. The results are shown in Table 3 - characteristics of linear regression model, where C_i is a column from the experimental data matrix (in practice it corresponds to a chosen (fixed) value for the stroke length).

One of the possible models is of the following kind:

$$C_6 = 0,240736 F_1 - 0,0551548 F_2 - 0,154939 F_3 - 0,142337 F_4 + 0,0107204 F_5 - 0,608597 F_6$$

It can be represented with the graphic shown in Fig. 7b.

Factor loading matrix after quartimax rotation				Characteristics of linear regression model				
Column of Data	Factor 1	Factor 2	Factor 3	Variable	Number of predictors	R ² – adjusted [%]	Standard error of the estimate	Durbin-Watson (DW) statistic
C3	6,10851	0,794675	2,23017	C1	5	98,9014	0,0290707	2,11178
C4	4,53872	-0,269085	-0,642947	C2	Linear dependent	-	-	-
C5	3,32156	-0,0892171	0,0915042	C3		98,7442	0,870267	3,45572
C6	6,56271	-0,0195374	-0,984033	C4	6	99,556	0,473327	2,87444
C7	5,4843	-0,0990264	-4,89347	C5	2	93,3557	1,23827	2,89583
C8	5,42788	-0,168117	-1,23314	C6	5	98,9673	0,976763	3,47721

C9	9,98393	1,83438	6,33052	C7	6	99,4197	0,776874	3,35185
C10	15,8413	-2,53395	-2,31961	C8	6	99,3615	0,536957	3,04229
C11	0,0799093	10,9791	0,335603	C9	5	99,8286	0,519949	2,81307
C12	14,1108	1,41282	3,52618	C10	6	99,9955	0,129426	2,57749

Table 3. Factor loading matrix after quartimax rotation and characteristics of linear regression model

Fig. 7. (a) Share of the variables in the individual factors; (b) Approximation of experimental data by linear regression. The confidence limits of the prediction are shown in red

Fig. 8. Forecasted demand by linear regression of principal components

All considered models are without a constant and can be compared. The coefficient R^2 shows, that the linear regression models describe the correlation in the experimental data between 86,9% and 99,99% (Table 3 - characteristics of linear regression model). The forecasted demand is shown in Fig. 8.

The procedure for processing missing values assumes random nature of these values so the missing value is independent of other values. For analysis of the missing values descriptive statistics of one variable is used. For their recovery a regression method is used that is included in the package SPSS.

The analysis process of the missing values can be taken as a possible prediction of the requests of consumption. The obtained results are shown in Fig. 9.

Fig. 9. Linear modules' demand after recovery of the missing values

4. Conclusion

The problems of demand forecasting and synthesis of a demand model for the design of optimal size ranges are discussed. For solving these problems a demand modeling approach for size ranges design of technical products is proposed. The approach is applied for demand modeling of modules for linear motion with pneumatic actuation. An original methodology for collecting of the input data is developed. For the data analysis are applied procedures of principle component analysis, factor analysis and regression by the principal components. Two main factors are considered as main parameters of the pneumatic modules – stroke length L and load capacity G . It is established that the random quantities L and G are exponentially distributed, and when natural logarithm transformation is applied it leads to normal distribution for both. It is also established that the distributions of the two main factors are confirmed with a high P-value. The found transformation on the distributions of L and G leads to normal distribution which is a necessary and sufficient condition for data analysis. Two main cases are studied:

1. Assuming the data is known a demand forecasting is made based on a study of the dispersion's structure. A multitude of regression models have been found on the basis of principal components (Table 3, Fig. 8). It has been established that the principal components are appropriate predictors of the real data.
2. Under the assumption of missing values in the data a procedure is applied for analysing these values, and the latter are recovered by a regression method (Fig. 9). The analysis process of the missing values can be taken as an possible prediction of the requests of consumption.

It is possible with presented approach to obtain the prediction and error estimation of the discussed problem. As a further development to this work neuron networks can be applied for analysis and solving of the problem. Future work will be the integration of the developed approach in a complete methodology for design of optimal size ranges optimization problems.

5 Acknowledgements

The presented scientific research is financed from the Inner competition of TU-Sofia – 2015.

6. References

- [1] G. Voronin et al., Mechanical engineering. Encyclopedia. Standardization and certification in mechanical engineering (Машиностроение. Энциклопедия. Стандартизация и сертификация в машиностроении), (in Russian), vol. 1-5, Moscow, Mashinostroenie, 2002.
- [2] A.I. Dashenko, A.P. Belousov, Design of automated lines, (Проектирование автоматических линий), (in russian), Vysshaya Shkola, Moscow, 1983.
- [3] R. Aydin, C.K. Kwong, P. Ji, H.M.C. Law, Market demand estimation for new product development by using fuzzy modeling and discrete choice analysis, Neurocomputing, 142, Elsevier, 2014, 136–146.

- [4] A. Ginsburg, F. Börjesson, G. Erixon, Size ranges in modular products – a financial approach, in: J. Malmqvist (Ed.), Proceedings of the 7th Workshop on Product Structuring – Product Platform Development, Chalmers University, Göteborg, 2004, pp 141-150.
- [5] W. H. Green, Econometric Analysis, Pearson Education Limited, 2012.
- [6] D.N. Gujarati, Basic Econometrics, New York, McGraw-Hill, 2009.
- [7] D.N. Gujarati, D.C. Porter, Essentials of econometrics. New York, McGraw-Hill, 2010.
- [8] Y. Haik, T. Shahin, Engineering Design Process, Second Edition, Cengage Learning, 2011.
- [9] T. Kipp, D. Krause, Computer aided size range development – data mining vs. optimization, in: M.N. Bergendhal, M. Grimheden, L. Leifer, P. Skogstad, U. Lindemann (Eds.), Proceedings of ICED 09 the 17th International Conference on Engineering Design, Vol. 4, Product and Systems Design, Palo Alto, 2009, pp 179-190.
- [10] I. Malakov, V. Zaharinov, V. Tsenov, Size ranges optimization, Procedia Engineering, 100, B. Katalinic (Ed.) 25th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM 2014, Elsevier, 2015, pp. 791-800.
- [11] D. Mueller, A cost calculation model for the optimal design of size ranges, in: N. Quircke (Ed.), Journal of Engineering Design, Vol. 22, Issue 7, Published by Taylor & Francis , 2011, pp 467-485.
- [12] G. Pahl, W. Beitz, J. Feldhuzen, K.H. Grote, Engineering Design: A Systematic Approach, Third Edition, Springer, 2007.
- [13] M.R. Solomon, G.W. Marshall, E.W. Stuart, Marketing: real people, real choices, 7th edition, Prentice Hall, 2012.
- [14] D.N. Lawley, A.E. Maxwell, Factor Analysis as a Statistical Method, 2nd edition, London, Butterworth, 1971.
- [15] StatPoint Technologies, Inc. STATGRAPHICS, Rev. 7, Operators, Principal Components, Factor Analysis, 2009.
- [16] IBM Corporation SPSS, ver. 20, Estimating Statistics and Imputing Missing Values, 2011.
- [17] T. P. Ryan, Modern Engineering Statistics, Acworth, Georgia, A John Wley & Sons, Inc. Publication, 2007.
- [18] E. Kadric, H. Bajric, M. Pasic, Demand modeling with overlapping time periods, Procedia Engineering, 100, B. Katalinic (Ed.) 25th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM 2014, Elsevier, 2015, pp. 791-800.