Industrial Engineering Methods in Management of Motor Vehicle Breakdowns Behaviour Analysis

Danijel Begović
PhD Student
University of Belgrade
Faculty of Mechanical Engineering

Mirjana Misita
Assistant Professor
University of Belgrade
Faculty of Mechanical Engineering

Dragan D. Milanović
Full Professor
University of Belgrade
Faculty of Mechanical Engineering

Dragan Lj. Milanović
Associate Professor
University of Belgrade
Faculty of Mechanical Engineering

The paper presents the methodology of grouping and processing of the data related to the motor vehicles' breakdowns in the observed sample. In order to perform statistical research with possible motor vehicle behaviour prediction, it is needed to have perennial tracking of selected population features. Using a non-parametric statistical Kolmogorov-Smirnov test, it is possible to identify the distribution affinity of particularly selected vehicle features in terms of breakdowns phenomena. By the introduction of advanced polynomials computation, it is possible to determine the function for prediction of certain vehicle breakdowns in the vehicle maintenance system. The results of the application of the function for prediction of certain vehicle breakdowns in the observed sample of motor vehicles allow monitoring of the behaviour of certain vehicle breakdowns. By the introduction of polynomials concentration, it is possible to overcome the problem of handling with generative polynomial functions with a high degree.

Keywords: *industrial engineering methods, breakdowns analysis, motor vehicle.*

1. INTRODUCTION

The behaviour of any system usually depends on several variables and so is given by one or several functions of one variable. These functions, in turn, provided they are sufficiently continuous [1], can be approximated by polynomials within some range and within some accuracy. As an example, the planning of car malfunction may be given by several polynomials featuring the behaviour of particularly selected features of vehicle systems. Those features represent the vehicle's breakdown classes which are derived and selected by means of the power of sophisticated data base tools and its queries. Polynomial coefficients are obtained by observing the subjected behaviour of the vehicle malfunction over a longer period of time.

Prior to further application of advanced polynomial computations, it is needed to identify whether subjected features behave congenially in statistical sense or not. The basic statistical analysis that should be applied considers the testing of the statistical system hypothesis. This is done by means of the Kolmogorov-Smirnov test. After the respective test is carried out, the hypothesis about the equality of the classes' malfunction distributions is either accepted or not.

2. DATA COLLECTING AND NEW STRATIFICATION

Data were collected and selected according to the following rules:

- the vehicles that are travel related,
- the vehicles to which they relate were purchased in Serbia or Montenegro,
- the vehicles during the repairs were in a basic warranty period,
- the vehicles are repaired in the period from 2004 to 2008.

Proved to be repaired during that time, about 16,000 vehicles with 70,000 malfunctions have been recorded.

The relevant data were recorded in the central data base which was priory projected in order to support the core business. The methodology of data collecting refers to the following:

- Sampling location – in authorized workshops.
- The sort of data – damage codes including all needed information such as affected construction group, potential cause of malfunction, causality of damage, etc.
- The sort of sampling – the data were collected as they arrived, daily and depending on the type and sort of damage. Dynamics of the sampling could not have been predicted due to the process's stochastic nature. The previous sampling features confirmed the hypothesis about the sample without replacement.
- The difficulties in sampling – there have been two main challenges. The first is related to the vehicles with unknown service history and inaccurate service data and the second is exclusion of all works related to the service measures and/or recall campaigns.
- The aim of data collection and processing – the main goal is research of vehicle behaviour during exploitation. The monitoring of failures with exact and detailed belonging information may be used in statistical analysis of the respective failures.

Queries on the database, which had about 70,000 records, gave only usual answers. The engineering experts have to place the following Decision Support System: based on the technical similarities of certain

Received: March 2011, Accepted: October 2011
Correspondence to: Dr Mirjana Misita
Faculty of Mechanical Engineering,
Kraljice Marije 16, 11120 Belgrade 35, Serbia
E-mail: mmisita@mas.bg.ac.rs

components of older and newer types of cars of the same class, we can predict the behaviour of malfunctions in totally new types of cars of the same class that will inherit similar circuits.

Therefore, a particular class of malfunctions was observed in this paper. We choose those classes particularly selected by previous experience to analyze, for example:

- the class of malfunctions related to the turbo compressor of diesel engine, and
- the class of malfunctions related to the fuel injection system.

The formed SQL queries in the database showed that there is particularly a large number of malfunctions which can be divided into two classes:

- MF1: breakdown of diesel engines turbo chargers, and
- MF2: failure of high pressure fuel injectors.

One particular malfunction is determined by unique vehicle damage code. Herewith, one singular damage code might be related to one or more events, where the event is considered as a particular workshop visit. The classes MF1 and MF2 do not depend on the number of cylinders and their arrangement. The data obtained (Table 1) are related for the 4-cylinder diesel engine case study.

Table 1. Table of malfunctions by year

Class	Year	Quarter			
	2004	I	II	III	IV
MF1	33	6	13	9	5
MF2	25	1	3	13	8
	2005	I	II	III	IV
MF1	13	1	2	5	5
MF2	11	2	4	4	1
	2006	I	II	III	IV
MF1	19	6	8	2	3
MF2	13	1	5	5	2
	2007	I	II	III	IV
MF1	5	2	0	2	1
MF2	0	0	0	0	0
	2008	I	II	III	IV
MF1	1	0	0	1	0
MF2	2	1	0	1	0

3. TESTING THE BASIC STATISTICAL HYPOTHESIS

Distribution of the class's malfunctions of MF1 and MF2 is unknown. The basic statistical analysis that should be applied considers the testing of the following statistical system hypothesis:

$$H_0(F_1 = F_2) \tag{1}$$

$$H_1(F_1 \neq F_2)\,. \tag{2}$$

Testing the equality of the distribution of these classes' malfunctions was carried out by the Kolmogorov-Smirnov test. The results of testing are:

$$Q(0.69811) = 0.2722 \tag{3}$$

with the probability of 0.95.

Taking the obtained result for the function Q, the conclusion is that there is not any reason to reject the hypothesis of equality of distributions of Class MF1 and MF2. Considering the importance of appropriate selection of the time series and the effect of the relevant sampling density on it, it is interesting that the critical values in testing are obtained for the year 2008. The respective diagram is given as follows (Fig. 1).

4. APPLICATION OF THE POLYNOMIAL COMPUTATIONS TO THE RESPECTIVE PROBLEM

If X is a discrete random variable taking values on some subset of the non-negative integers $\{0, 1, \ldots\}$, then the probability – generating functions of X is defined as [2]:

$$G_X(t) = E(t^X) = \sum_{x=0}^{\infty} p_x t^x \tag{4}$$

where p_x is probability mass function of X. The probability – generating functions are useful for dealing with the functions of independent random variables. If $X_1, X_2, \ldots, X_n$ is some sequence of independent random variables and [2]:

$$Y = \sum_{i=1}^{n} a_i X_i \tag{5}$$

then the probability – generating function is given by [2]:

$$G_Y(t) = E(t^Y) = E(t^{\sum_{i=1}^{n} a_i X_i}) \tag{6}$$

and taking the multiplicatively feature of mathematical expectation E, we get the probability – generating function given as:

$$G_Y(t) = G_{X_1}(t^{a_1})\, G_{X_2}(t^{a_2}) \ldots G_{X_n}(t^{a_n})\,. \tag{7}$$

A very important fact is that the coefficient with t^{y_i} is the probability of event $\{Y = y_i\}$, where there exist several ways to calculate y_i.

Forming an empirical distribution malfunction class MF1 we get the probability through generating the function of the 19th degrees with the coefficients given in the following table (Table 2):

Table 2. Probability – generating function of MF1 class

Degree	Coef.	Degree	Coef.
0	0.08451	10	0.02817
1	0.18310	11	0.04225
2	0.12676	12	0.02817
3	0.07042	13	0.00000
4	0.01408	14	0.02817
5	0.02817	15	0.01408
6	0.07042	16	0.00000
7	0.07042	17	0.00000
8	0.08451	18	0.01408
9	0.11268	19	0.00000

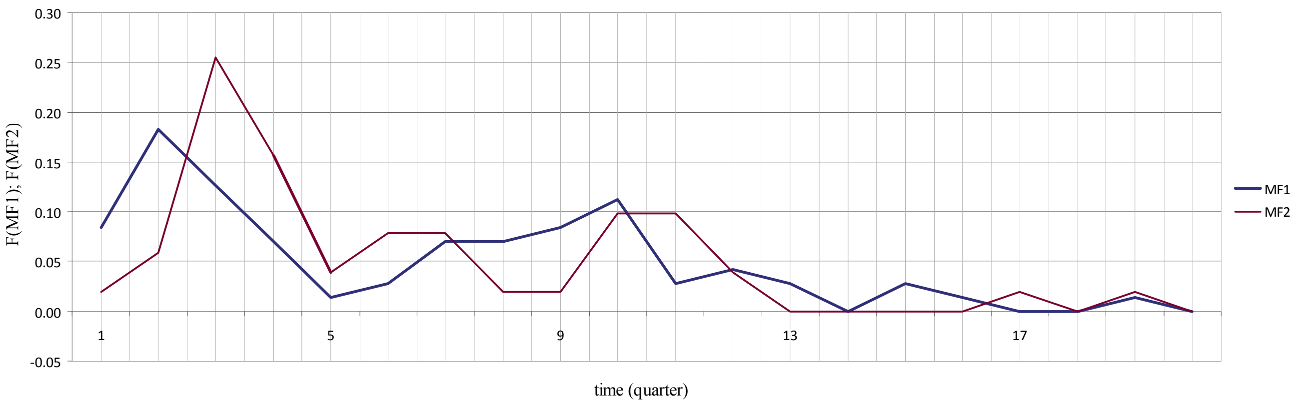

Figure 1. Behaviour of MF1 and MF2 in time

Similarly, for the class of MF2 faults we get the probability – generating function of 19 degrees with the coefficients given in Table 3.

Table 3. Probability – generating function of MF2 class

Degree	Coef.	Degree	Coef.
0	0.01961	10	0.09804
1	0.05882	11	0.03922
2	0.25490	12	0.00000
3	0.15686	13	0.00000
4	0.03922	14	0.00000
5	0.07843	15	0.00000
6	0.07843	16	0.01961
7	0.01961	17	0.00000
8	0.01961	18	0.01961
9	0.09804	19	0.00000

Although the technical systems of each class's failures, MF1 and MF2, are functionally independent, those variables are mutually influential, which is clear to engineering experts, but the reported malfunctions are not considered in pairs. Therefore, we can form the probability – generating function random variables:

$$Y = \mathrm{MF1} + \mathrm{MF2}\,. \tag{8}$$

Variable Y observes the mutual influence of malfunction classes on the overall slowdown of the technical system over time. Random variable Y is a carrier-class percentage share of malfunction that has been formed. Also, the use of the quasi probability – generating function is suitable, because the events at the beginning and end, with a small probability, can be achieved due to the nature of the observed problems.

Applying [3] and [4], referring to the product of the polynomials, after extracting (5), (6), (7) and multiplying:

$$G_Y(t) = G_{\mathrm{MF1}}(t)\,G_{\mathrm{MF2}}(t) = G_{\mathrm{MF1+MF2}}(t) \tag{9}$$

we get the distribution of Y (Table 4).

Based on (8), it can be concluded that generative polynomial function of variable Y, defining mutual influence of the taken malfunction classes, has the degree of 38. An example given in Table 4 shows the volume of work that needs to be done, in order to come to some conclusions related to the random variables MF1 and MF2, which is interpreted in Figure 2.

Table 4. Probability – generating function of *Y* = MF1 + MF2 class

Degree	Coef.	Degree	Coef.
0	0.001657	20	0.021541
1	0.008561	21	0.013808
2	0.034797	22	0.007457
3	0.068766	23	0.007180
4	0.068766	24	0.008285
5	0.053024	25	0.006352
6	0.043634	26	0.003038
7	0.043630	27	0.004419
8	0.049434	28	0.002486
9	0.058271	29	0.001381
10	0.075946	30	0.001105
11	0.090307	31	0.000276
12	0.069594	32	0.000552
13	0.048605	33	0.000276
14	0.040597	34	0.000276
15	0.032312	35	0.000000
16	0.035073	36	0.000276
17	0.037835	37	0.000000
18	0.033416	38	0.000000
19	0.027064		

One of the outcomes is that in the 3rd, 4th and 5th quarters after the beginning of the warranty period malfunctions can be expected, which are observed in both classes. Especially critical are the 10th, 11th and 12th quarters after the beginning of the warranty period due to both high malfunction's probability.

The conclusion after the analysis of the resulting probability values is that the total degree of resulting polynomial function has much higher degree than a particular probability – generative polynomial. Its coefficients are formed by multiplying the probability – generative polynomial of a particular random variable.

Based on [3] authors found much deeper result. If one of the polynomials has a high coefficient and other some concentration at low degrees, the product has high coefficient. In order to challenge this situation, it is proposed to use the concentration of polynomial function. The general idea for introducing the concentration and concentration factor replaces the actual degree of the polynomial by the concentration factor in order to obtain estimates independent of the actual degree of the polynomial.

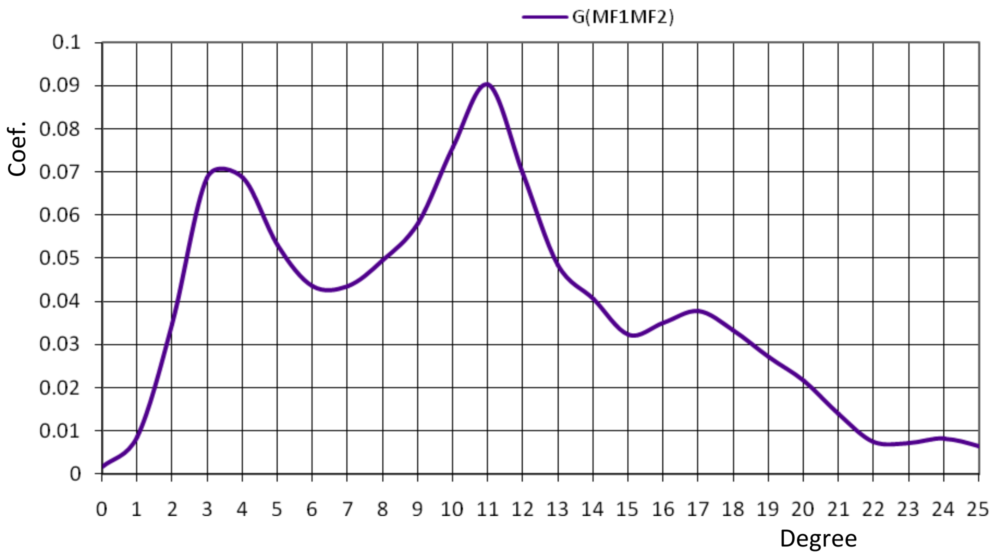

Figure 2. Generative function of the malfunctions classes MF1 and MF2

Since the degree of the resulting polynomial is high, and becomes higher over time, the aim is to reduce the degree of polynomial function without losing respective information it carries. This can be done by the use of quasi probability – generating function (10). Taking the methodology described in [3], the following basic assumption can be made.

If p_x is probability mass function of X, where $X \in \{x_0, x_1, \ldots, x_n\}$ and $x_i \in N_0$ then quasi probability – generating functions of X are defined as:

$$\underline{G}_X(t) = \sum_{i=1}^{n} p_{x_i} t^{x_i} \ \text{ or } \ \overline{G}_X(t) = E(t^X) = \sum_{i=0}^{n-1} p_{x_i} t^{x_i} . \quad (10)$$

Here, the derived generating polynomial function satisfies all conditions [3] about the concentration factor of the polynomial at its total degree, where events $\{X = x_0\}$ or $\{X = x_n\}$ can be omitted. Further calculations are to be directed toward finding the best constant satisfying the respective requirements [3] in order to establish a decision rule: which pair of events is critical in the procedure of prediction and planning of vehicle service system.

5. CONCLUSION

Using the features of the probability – generative polynomial, it is possible to observe the influence of independent random variables having the same unknown distribution. The observation comes down to analyzing the value of the resulting probability – generative polynomial coefficients that were formed by multiplying the probability – generative polynomial of particular random variables. The consequence is that the resulting probability – generative polynomial is of much higher degree than a particular probability – generative polynomial. The respective problem of handling huge polynomial functions can be overcome by the use of polynomials concentration theory.

REFERENCES

[1] Beliakov, G.: How to build aggregation operators from data, International Journal of Intelligent Systems, Vol. 18, No. 8, pp. 903-923, 2003.

[2] Grinstead, C.M. and Snell, J.L.: *Introduction to Probability*, American Mathematical Society, Providence, 1991.

[3] Beauzamy, B. and Enflo, P.: Estimations de produits de polynômes, Journal of Number Theory, Vol. 21, No. 3, pp. 390-412, 1985.

[4] Beauzamy, B., Bombieri, E., Enflo, P. and Montgomery, H.L.: Products of polynomials in many variables, Journal of Number Theory, Vol. 36, No. 2, pp. 219-245, 1990.

NOMENCLATURE

MF1	random variables of class 1 malfunction
MF2	random variables of class 2 malfunction
F1	distribution function of MF1
F2	distribution function of MF2
Q	Kolmogorov-Smirnov's probability function
G	generative probability function

ПРИМЕНА МЕТОДА ИНДУСТРИЈСКОГ ИНЖЕЊЕРСТВА У АНАЛИЗИ УПРАВЉАЊА ПОНАШАЊЕМ ОТКАЗА КОД МОТОРНИХ ВОЗИЛА

Данијел Беговић, Мирјана Мисита, Драган Д. Милановић, Драган Љ. Милановић

У раду је изложена методологија груписања и обраде података о отказима на моторним возилима у посматраном узорку. Применом непараметарског теста Колмогоров-Смирнов, могуће је установити једнакост расподела изабраних врста појава отказа моторног возила. Коришћењем генеративног полинома, могуће је установити утицај независних случајних величина које имају непознату расподелу. Применом производа полинома са комплексним коефицијентима могуће је одредити функцију за предвиђање одређених појава отказа у оквиру система одржавања возила. Резултати примене функције за предвиђање одређених појава отказа на посматраном узорку моторних возила омогућавају праћење понашања одабраних појава отказа на посматраним моторним возилима.