UDC 519.242.7+ 519.233.4

V.N.Fomin, A.V.Dik

Ye.A.Buketov Karaganda State University
(E-mail: vitfomin@mail.ru)

Using one-way analysis of variance in the stochastic determinated design of experiment

The possibility of using one-way dispersion analysis ANOVA in this article is discribed, for the evaluation of statistical significance of the empirical dependensis which was gift by using of stochastic determinated design of experiments (SDDE). It is shown that use of ANOWA can disambiguate the interpretation of some results.

Key words: ANOWA, stochastic determinated design of experiment, regression analysis, analysis of variance, factor experiment.

The mathematical modeling is important tool in the study of complex processes in the chemistry, chemical engineering and other sciences. One of the mathematical models constructing methods it is design of experiments. One of the popular methods of DOE it is the full factor experiment. Mathematically rigorous implementation of the full factor experiment requires a large number of tests. The number of tests which required the full factor plan is rapidly rises with increasing of the number of studied factors and/or levels of their variation. One of the best ways of the optimization factor experiment is the use of the stochastic determinated design of experiments. The method of experiment planning is based on using of the Latin squares, and was first used in the 30s of XX-th century by R.Fisher in agriculture [1]. Subsequently, a method has been developed in the researches of M.Protodjakonov and R.Teder [2]. In the last quarter of the XX century, the stochastic determinated design of experiments came to a modern state through the work of the Kazakhstan scientist V.Malyshev and his followers [3, 4]. Today, the method is widely used in the study of multifactorial dependencies in chemistry, biology, etc. One of the reasons for the limited popularity of the method SDDE is a possibility ambiguous estimate which is obtained with the help of individual dependencies and, as a consequence, the final equation.

Within the SDDE for evaluate the accuracy of approximation individual dependencies and the final equation used coefficient of nonlinear multiple correlation (CNMC) and its significance. These interrelated values characterize only the degree of correlation of the experimental points with the curve fit. Question, is there any influence of the factor on the result of the experiment, within the SDDE has no mathematically definite answer. The «significance» of coefficient non-linear multiple correlation does not mean the same thing what the adopted under other methods of mathematical processing of the experimental data of factor «significance».

For example, consider a mathematical model of the well-studied the decomposition reaction of nitric oxide (V) [5]:

$$2N_2O_5=4NO_2+O_2$$

The dependence between rate constants of the reaction and temperature obtained from the experimental data is presented in Table 1.

Репозиторий КарГУ

Table 1

The dependence between the rate constant for the decomposition N_2O_5 and temperature

Temperature, K	Rate constant, c^{-1}
283	$3.83 \cdot 10^{-6}$
293	$1.71 \cdot 10^{-5}$
303	$6.94 \cdot 10^{-5}$
313	$2.57 \cdot 10^{-4}$
323	$8.78 \cdot 10^{-4}$

The dependence between the rate constant and temperature is determined by the Arrhenius equation:

$$k = Ae^{\frac{-E_a}{RT}}. \quad (1)$$

The dependence between degree of conversion and time is described by the Kazeev-Erofeev equation:

$$\alpha = 1 - e^{kt}. \quad (2)$$

From (1) and (2) is seen that the degree of conversion independent of temperature and time, but does not depend on the concentration of the substance.

To prevent systematic errors used the data is calculated by the equation (2) instead of experimental data. In the calculated data make an randomly error within ±0.5 % or ±2.5 % for the experimental error simulation. Make up the experimental design, taking into account the 4 factors with 3 levels. The first factor is the initial concentration of N_2O_5 and vacated position for fourth factor. The second and third factors are respectively the temperature and time of the process (Table 2).

Table 2

Plan and the results of the calculation of the degree of conversion of N_2O_5, 4 factors, 3 levels

№	$C_{N_2O_5}$, M	T, K	t, c	Vacant position	Definitely	±0.5 %	±2.5 %
1	0.11	283.0	5400.0	1	0.020470	0.02040859	0.020122
2	0.11	303.0	16200.0	2	0.675114	0.677814456	0.689291
3	0.11	323.0	21600.0	3	1.000000	0.999	1.01
4	0.22	283.0	16200.0	3	0.060160	0.06003968	0.06028
5	0.22	303.0	21600.0	1	0.776656	0.779762624	0.780539
6	0.22	323.0	5400.0	2	0.991272	0.987306912	0.984333
7	0.33	283.0	21600.0	2	0.079398	0.079398	0.078048
8	0.33	303.0	5400.0	3	0.312546	0.313171092	0.307858
9	0.33	323.0	16200.0	1	0.999999	0.997999002	0.996999

Let's carry out the mathematical processing of the obtained data methods which is adopted in SDDE. For the averaging the data using the arithmetic mean. To approximate the dependence of the conversion degree from vacant factor using a linear equation, the temperature dependence suppose of the logarithmic, and time dependence is exponential. For the intended degree of conversion, depending on the initial concentration of nitrogen oxide searched the best approximating function. Tables 3–6 show the selections of individual dependences.

Table 3

Dependence of the conversion degree on concentration N_2O_5

№	Result ±0,5 %	Result ±2,5 %	№	Result ±0,5 %	Result ±2,5 %	№	Result ±0,5 %	Result ±2,5 %
1	0.02040859	0.020122	4	0.06003968	0.06028	7	0.079398	0.078048
2	0.677814456	0.689291	5	0.779762624	0.780539	8	0.313171092	0.307858
3	0.999	1.01	6	0.987306912	0.984333	9	0.997999002	0.996999
aver.	0.565741	0.573138	aver.	0.609036	0.608384	aver.	0.463523	0.460968

Репозиторий КарГУ

Table 4

Dependence of the conversion degree on temperature

№	Result ±0,5 %	Result ±2,5 %	№	Result ±0,5 %	Result ±2,5 %	№	Result ±0,5 %	Result ±2,5 %
1	0.02040859	0.020122	2	0.677814456	0.689291	3	0.999	1.01
4	0.06003968	0.06028	5	0.779762624	0.780539	6	0.987306912	0.984333
7	0.079398	0.078048	8	0.313171092	0.307858	9	0.997999002	0.996999
aver.	0.0532821	0.0528167	aver.	0.590249	0.592563	aver.	0.994769	0.997111

Table 5

Dependence of the conversion degree on time

№	Result ±0,5 %	Result ±2,5 %	№	Result ±0,5 %	Result ±2,5 %	№	Result ±0,5 %	Result ±2,5 %
1	0.02040859	0.020122	2	0.677814456	0.689291	3	0.999	1.01
6	0.987306912	0.984333	4	0.06003968	0.06028	5	0.779762624	0.780539
8	0.313171092	0.307858	9	0.997999002	0.996999	7	0.079398	0.078048
aver.	0.440296	0.437438	aver.	0.578618	0.58219	aver.	0.619387	0.622862

Table 6

Dependence of the conversion degree on vacant position

№	Result ±0,5 %	Result ±2,5 %	№	Result ±0,5 %	Result ±2,5 %	№	Result ±0,5 %	Result ±2,5 %
1	0.02040859	0.020122	2	0.677814456	0.689291	3	0.999	1.01
5	0.779762624	0.780539	6	0.987306912	0.984333	4	0.06003968	0.06028
9	0.997999002	0.996999	7	0.079398	0.078048	8	0.313171092	0.307858
aver.	0.59939	0.59922	aver.	0.581506	0.583891	aver.	0.457404	0.459379

The coefficients obtained in approximating functions depend little on the specified accuracy. Graphs obtained with different functions exactly visually indistinguishable and are presented in Figures 1–4.

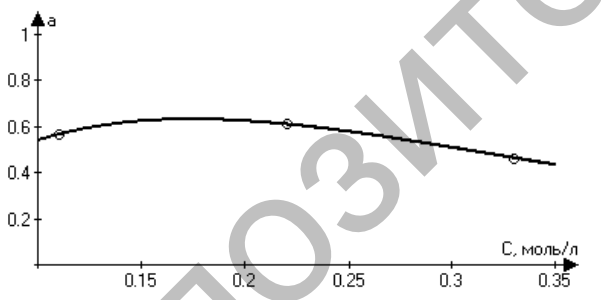

Figure 1. Dependence of the conversion degree on concentration

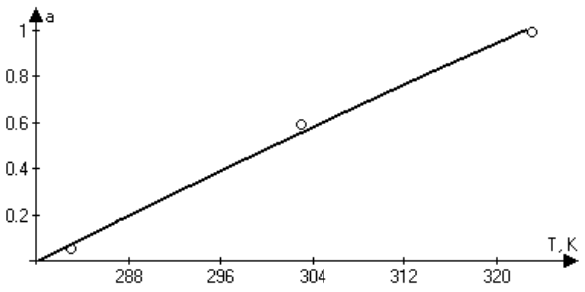

Figure 2. Dependence of the conversion degree on temperature

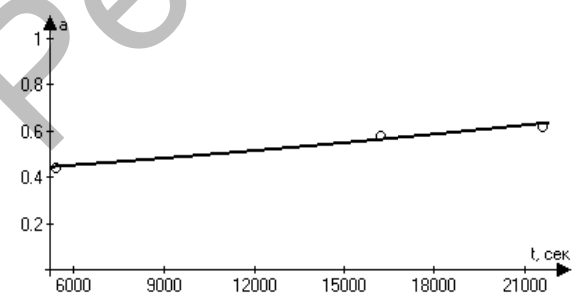

Figure 3. Dependence of the conversion degree on time

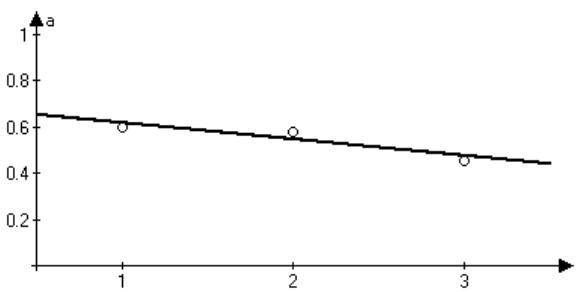

Figure 4. Dependence of the conversion degree on vacant factor

Репозиторий КарГУ

Approximating functions and their corresponding correlation coefficients are given in Table 7.

Table 7

The approximative functions

Dependence of the conversion degree from		«Data» accuracy, %	The approximating function	CNMC	«Significance» of CNMC
Concentration	1	±0,5	$\alpha = 17.3339 \cdot e^{-6.92501 \cdot C} \cdot C^{1.20538}$	1	∞
	2	±2,5	$\alpha = 16.1716 \cdot e^{-6.8472 \cdot C} \cdot C^{1.17215}$	1	∞
Temperature	3	±0,5	$\alpha = -40.1821 + 7.12994 \cdot \ln(T)$	0.996168	130.229687
	4	±2,5	$\alpha = -40.3031 + 7.15136 \cdot \ln(T)$	0.995969	123.788195
Time	5	±0,5	$\alpha = 0.395523 \cdot e^{\tau \cdot 2.16711E-5}$	0.974247	19.161953
	6	±2,5	$\alpha = 0.391578 \cdot e^{\tau \cdot 2.248831E-5}$	0.970676	16.797160
Vacant factor	7	±0,5	$\alpha = 0.6881 - 0.071 \cdot V$	0.828075	2.634733
	8	±2,5	$\alpha = 0.6873 - 0.0699 \cdot V$	0.813624	2.407058

The coefficients in the approximating function, nonlinear multiple correlation and levels of significance are hardly depend on the accuracy of the calculated data. Thus, all subsequent arguments can lead to the same level of accuracy.

From the graph in Figure 2 it is shown that with increasing temperature the degree of conversion rises. Given the high value of the CNMC and its significance, it is possible to make an unambiguous conclusion about the dependence of the conversion degree on the process temperature. In the case of the other three parameters of the data SDDE not clear conclusion. CNMC values in all cases large and have high levels of significance, but it influences whether variation in these factors examined within the degree of decomposition of nitric oxide? It is known that the vacant factor and the initial concentration does not affect the reaction. However, prove it within the SDDE method of calculation is impossible. Exclusion from the final equation «concentration component» reduces of the CNMC and its importance. Inclusion in the equation on vacant factor increases its mathematical adequacy (Table 8). This is contrary to originally defined model.

Table 8

Variants of the final equation

Final equation	CNMC	«Significance» of CNMC
$\alpha = 7.12994 \cdot \ln(T) + 0.395523 \cdot e^{2.16711*10^{-5}\tau} - 40.7282$	0.964132	33.522337
$\alpha = 17.3339 \cdot e^{-6.92501 \cdot C} \cdot C^{1.20538} \cdot 7.12994 \cdot \ln(T) + +0.395523 \cdot e^{2.16711 \cdot 10^{-5}\tau} - 40.7282$	0.975679	45.404003
$\alpha = 17.3339 \cdot e^{-6.92501 \cdot C} \cdot C^{1.20538} \cdot 7.12994 \cdot \ln(T) + +0.395523 \cdot e^{2.16711*10^{-5}\tau} - 0.071 \cdot V - 40.0401$	0.990762	107.746218

This behavior of a mathematical models obtained by SDDE frequently are found in various publications. As a rule, accuracy of the real process final equation increases with the number of levels of varying factors, i.e., with increasing number of experiments.

Without increasing the number of experiments is necessary to use other methods of statistical data processing.

It can be noted that the data grouped by level of factors to output partial dependencies (Tables 3–6) as the same structure can be used in the analysis of variance. The unique problem the analysis of variance is found how a change in the studied factors is influence on the final result. Table 9 shows the results of one-way ANOVA to investigate the dependence.

Репозиторий КарГУ

Table 9

Results of the data processing using method ANOWA

Dependence of degree of conversion from concentration N_2O_5						
Source of variation	SS	df	MS	F	P	F critical
Between groups	0.035555739	2	0.01777787	0.074128044	0.929389808	5.14325285
Within groups	1.438959024	6	0.239826504			
Total	1.474514764	8				
Dependence of degree of conversion from temperature						
Source of variation	SS	df	MS	F	P	F critical
Between groups	1.346675562	2	0.673337781	31.6024086	0.000651696	5.14325285
Within groups	0.127839202	6	0.021306534			
Total	1.474514764	8				
Dependence of degree of conversion from time						
Source of variation	SS	df	MS	F	P	F critical
Between groups	0.056989929	2	0.028494965	0.120611494	0.888473856	5.14325285
Within groups	1.417524834	6	0.236254139			
Total	1.474514764	8				
Dependence of degree of conversion from vacated position						
Source of variation	SS	df	MS	F	P	F critical
Between groups	0.035293533	2	0.017646766	0.07356798	0.92989796	5.14325285
Within groups	1.439221231	6	0.239870205			
Total	1.474514764	8				

The results of analysis of variance showed that on the concentration of vacant factor and results of numerical experiments clearly does not depend ($F<<F_{crit}$, P close to one). There is no dependence of the nitrogen oxide decomposition degree on temperature, according ANOWA ($F>>F_{crit}$, P value is very small).

ANOVA did not answer the question, does the processing time on the degree of it completion. The result of analysis of variance indicates the independence of the nitric oxide decomposition degree on the time of the process. Thus, the calculations carried out by two independent methods are not possible to detect the dependence of the existence of which we are obviously aware. Apparently, the problem is in the unfortunate choice of time interval. May be need simultaneous consideration of the temperature and the process duration.

Summarizing the above, we can conclude that the method ANOWA is a good addition to the method SDDE, and can be used by researchers to confirm or verification of the conclusions.

References

1 *Malyh A.E., Danilov V.I.* The historical development of the theory of Latin squares and some of their applications // Bull. of Perm State University. — 2010. — Vol. 4. — P. 95–104.

2 *Protod'iakonov M.M., Teder R.I.* The rational planning technique of the experiment. — Moscow: Nauka, 1970.

3 *Malyshev V.P.* Mathematical design of metallurgical and chemical experiments. — Alma-Ata: Nauka, 1977. — P. 37.

4 *Belyaev S.V., Malyshev V.P.* Development ways of stochastic determinated design of experiment // Complex processing of mineral resources of Kazakhstan. State. Problem. Solutions. — Almaty, 2008. — Vol. 9. Information technology in mineral complex. — Chapter 8. — P. 599–633.

5 *Fremantle M.* Chemistry in action. In 2 vol. Part 1 / Trans. from Engl. — Moscow: Mir, 1998. — 528 p.

В.Н.Фомин, А.В.Дик

Тәжірибені болжам-детерминирленген жоспарлау шеңберінде бірфакторлы дисперсиялық талдауды пайдалану

Мақалада тәжірибені болжам-детерминирленген жоспарлау (ТБДЖ) әдісімен алынған эмпирикалық тәуелділіктерді статистикалық маңыздылығын бағалауда бірфакторлы дисперсиялық талдау (ANOWA) қолдану мүмкіндігі зерттелді. Тәжірибені болжам-детерминирленген жоспарлауға қосымша ANOWA пайдалану кейбір алынған нәтижелердің интерпретациялану сапалық көрсеткіштерін жоғарылатады.

Репозиторий КарГУ

В.Н.Фомин, А.В.Дик

Использование однофакторного дисперсионного анализа в рамках вероятностно-детерминированного планирования эксперимента

В статье исследована возможность применения однофакторного дисперсионного анализа (ANOWA) для оценки статистической значимости эмпирических зависимостей, полученных с использованием вероятностно-детерминированного планирования эксперимента (ВДПЭ). Показано, что использование ANOWA в дополнение к ВДПЭ может устранить неоднозначность интерпретации некоторых результатов.

Репозиторий КарГУ