Scientific Bulletin. Series F. Biotechnologies, Vol. XIX, 2015
ISSN 2285-1364, CD-ROM ISSN 2285-5521, ISSN Online 2285-1372, ISSN-L 2285-1364

AN APPLICATION OF ANALYSIS OF VARIANCE TO A PROBLEM OF BIOLOGY

Mioara VARGA

University of Agronomic Sciences and Veterinary Medicine of Bucharest,
59 Mărăşti Blvd, District 1, 011464, Bucharest, Romania, Phone: +4021.411.15.90,
Email: vargamioara@yahoo.com

Corresponding author email: vargamioara@yahoo.com

Abstract

In this paper, we discuss about an application of Analysis of Variances from biology. We are interested to determine the effect of adding the NaCN (sodium cyanide) on the uptake in vitro of a particular amino acid by intestinal preparations from certain species of fish. First, we introduce some theoretical notions about this statistical method; we present the most important calculation formulas and the way they are used for comparing the treatment means in case of an experiment with k treatments, b blocks and n replicates per treatment per block. Afterwards, we consider a two-treatment experiment (with and without NaCN) and the results obtained for three replicates from each of the 4 species of fish are expressed as μ mol g^{-1} dry weight per 20 min period. After we calculate the sums of observational data per sub-class, per treatment and grand totals, we determine the elements from the table of analysis of variances (sums of squares, degrees of freedom, the error mean square) and we use the Fisher test for establishing the conclusions. Then we find that the difference between treatment means is 0,49, thus with 95 % confidence probability, we may state that the addition of the NaCN to the medium reduces uptake by 0,68 μ mol g^{-1} dry weight per 20 min period.

Key words: *sums of squares, degrees of freedom, the error mean square, randomized block.*

INTRODUCTION

With certain and many nutritional qualities, fish has an important role in human nutrition. Content rich in amino acids, fats and minerals provide both basic nutritious properties and increasing energy needs.

Numerous studies have proven that all living aquatic shows accumulation and concentration capabilities even in natural environmental conditions, therefore no direct sources of pollution, which continues to justify concern for maintaining a habitat as unpolluted.

With wide application in technological processes, NaCN is a dangerous pollutant of the environment. Discharge of wastewater containing this substance must be achieved through strict observance of the rules, otherwise it may cause, real environmental disasters. Studies show that cyanide reacts with other elements and is broken down into hundreds of compounds containing various concentrations of cyanide in accumulating in fish tissues.

In this context, the work shows the effect of adding the NaCN (sodium cyanide) on the uptake in vitro of a particular amino acid by intestinal preparations from certain species of fish, this topic constituting the subject of a previous article (Varga, 2010), the conclusions being established by applying analysis of variance.

MATERIALS AND METHODS

It is considered an experiment in which are preserved constant all of the factors that influence the character of a statistical population *Y*, except one, which will be variable *X*. The homogeneous population is grouped in *m* homogeneous subpopulations associated with variations $X_i, i=\overline{1,m}$ of the factor X (Parker, 1973).

In the case of a heterogeneous population group, it can be done in *l* sub-populations, homogeneous in relation to the acquisition of *Y* under study. Experimental material will also be heterogeneous consisting of the *l* polls of the subpopulations, each sample consisting of *n* repetitions corresponding to the *m* variants, denoted $x_{ij}, i=\overline{1,m}, \quad j=\overline{1,n}$

If the number of repetitions is different from the variants of the factor *X* $(m \neq n)$ is used the model (drawing) of the randomized blocks $B_1, B_2, ...B_l$. Each repetition is designated by a box which is denoted applied variant factor *X*, ie $X_i, i = \overline{1,m}$ and the response repetition x_{ij}^b , $i = \overline{1,m}, j = \overline{1,n}$, $b = \overline{1,l}$

The results can be organized in the following table:

Table 1.The experimental data

B_1	X_1 $x_{11}^1, x_{12}^1, ...x_{1n}^1$	X_2 $x_{21}^1, x_{22}^1, ...x_{2n}^1$	…	X_m $x_{m1}^1, x_{m2}^1, ...x_{mn}^1$
B_2	X_1 $x_{11}^2, x_{12}^2, ...x_{1n}^2$	X_2 $x_{21}^2, x_{22}^2, ...x_{2n}^2$		X_m $x_{m1}^2, x_{m2}^2, ...x_{mn}^2$
…	…………..	…………..	……	…………
B_l	X_1 $x_{11}^l, x_{12}^l, ...x_{1n}^l$	X_2 $x_{21}^l, x_{22}^l, ...x_{2n}^l$	……	X_m $x_{m1}^l, x_{m2}^l, ...x_{mn}^l$

The responses are placed in a new table to be processed by bi-factorial analysis of variance with *n* repetitions for each cell.

Table 2. The totals/block

Block	Totals/ block
B_1	$s_1 = \sum_{j=1}^{n} x_{1j}^1 + \sum_{j=1}^{n} x_{2j}^1 + ... + \sum_{j=1}^{n} x_{mj}^1$
B_2	$s_2 = \sum_{j=1}^{n} x_{1j}^2 + \sum_{j=1}^{n} x_{2j}^2 + ... + \sum_{j=1}^{n} x_{mj}^2$
…………	…………………………
B_l	$s_l = \sum_{j=1}^{n} x_{1j}^l + \sum_{j=1}^{n} x_{2j}^l + ... + \sum_{j=1}^{n} x_{mj}^l$

If $GT^2 = s_1^2 + s_2^2 + ... s_n^2$ and $C = \frac{GT^2}{lmn}$, it is determined the sums of squares, the degrees of freedom and the mean squares to calculate the experimental value of the Fisher test in order to complete the analysis of variance table.(Varga, 2014)

Table 3.The analysis of variance table

Sources of variation	Sums of-squares	Degrees of freedom *GL*	Mean squares S^2	*F*
X	*SSX*	*m*-1	S_X^2	$F_X = S_X^2 / S_E^2$
B	*SSB*	*l*-1	S_B^2	$F_B = S_B^2 / S_E^2$
Q	*SSQ*	*n*-1	S_Q^2	$F_Q = S_Q^2 / S_E^2$
E	*SSE*	(m-1)(n-1)	S_E^2	
T	*SS*	*mnl*-1		

Where,

$$SSX = \sum_{i=1}^{m} (T_i)^2 / nl - C, \quad T_i = \sum_{b=1}^{l} x_{ij}^b, \quad i = \overline{1,m}$$

$$SSB = \sum_{b=1}^{l} (s_b)^2 / mn - C \quad SS = \sum_{i,j,b} (x_{ij}^b)^2 - C$$

$$SSE = SS - (SSX + SSB) \quad SSQ = (\sum_{b=1}^{l} s_b^2) / n - C$$

$$S_X^2 = SSX / GL_X, \ S_B^2 = SSB / GL_B, \ S_E^2 = SSE / GL_E$$

The values F_X and F_B are compared with the critical values $F_{0,05}$, $F_{0,01}$, $F_{0,001}$ extracted from the Fisher distribution tables for pairs of the corresponding degrees of freedom (GL_X, GL_E) respectively (GL_B, GL_E). If the Fisher test experimental values (last column in Table 3) are lower than the values tabulated $F_{0,05}$, $F_{0,01}$, $F_{0,001}$, then different versions of the *X* factor significantly influences the character *Y*.

It is considered an experiment conducted on four species of fish, to which they apply two types of treatments (with and without NaCN in their environment); for the investigation on the effect of NaCN on the uptake in vitro of a particular amino acid by intestinal preparatio The results (expressed in $\mu \, mol \, g^{-1}$ dry weight per 20 min period) are grouped in the following table (Table 4).

Table 4.The experimental data

Fish	$X_1 =$ Treatment without NaCN	$X_2 =$ Treatment with NaCN
Fish 1	1,54; 1,92; 2,26	1,10; 1,42; 1,04;
Fish 2	1,52; 2,02; 1,91	1,31; 1,15; 1,51
Fish 3	1,00; 1,12; 1,13	0,79; 0,84; 0,86
Fish 4	1,58; 1,78; 1,52	1,24; 0,81; 1,32

RESULTS AND DISCUSSIONS

First, it is calculated the amounts of observational data sub-classes, treatments, and total (Varga, 2014).

1. Sums of squares:

$$C = \frac{GT^2}{lmn} = \frac{32,69^2}{24} = 44,52$$

$$SS = 1,54^2 + 1,92^2 + ...1,32^2 - 44,52 = 3,81$$

$$SSX = \sum_{i=1}^{m} (T_i)^2 / nl - C = \frac{T_1^2 + T_2^2}{12} - C = 1,45$$

$$SSB = \frac{9,28^2 + 9,42^2 + 5,74^2 + 8,25^2}{6} - C = 1,45$$

$$SSE = SS - (SSX + SSB) = 0,91$$

2. Degrees of freedom:
$GL_X = 1$; $GL_B = 3$; $GL_T = 23$; $GL_E = 19$.

3. Mean-squares:
$S_X^2 = 1,45$, $S_B^2 = 0,48$, $S_E^2 = 0,04$

4. The experimental values of Fisher test are:

$$F_X = S_X^2 / S_E^2 = 30,38 \; ; \; F_B = S_B^2 / S_E^2 = 10,1$$

Next, it is completed the table of variances analysis (Table 5)

Table 5.The analysis of variance table for the problem studied

Sources of variation	Sums of-squares	Degrees of freedom	Mean squares	F
X	1,45	1	1,45	$F_X = 30,38$
B	1,45	3	0,48	$F_B = 10,1$
E	0,91	19	0,04	-
T	3,81	23		-

The population is heterogeneous with respect to *Y*, but can be divided into homogeneous blocks (l = 4): B_1(Fish 1), B_2(Fish 2), B_3(Fish 3) B_4(Fish 4).
There are m = 2 variants of *X* factor (NaCN Without Treatment, Treatment with NaCN) and n = 3 repetitions / treatment / block.

5. From the tables of Fisher distribution for (1; 19) GL, the critical values are obtained: $F_{0.05}$ = 4,38, $F_{0.01}$ = 8,18 and $F_{0.001}$ = 15,08; for (3,19) GL, $F_{0.05}$ = 3,13, $F_{0.01}$ = 5,74 şi $F_{0.001}$ = 8,28.

6. By comparing the values $F_X = 30,38 > F_{0,001}$ $F_B = 10,1 > F_{0,001}$ we obtain the conclusion that the treatments with and without Na CN and block variation leads to very significant differences.
The contributions of the variation of *X*-factor and *B* to the variation of *Y* are:

$$A_X = \frac{SSX}{SS} = \frac{1,45}{3,81} = 38\%,$$

$$A_B = \frac{SSB}{SS} = \frac{1,45}{3,81} = 38\%,$$

$$A_X = 100\% - A_X - A_B = 24\%$$

Because the experiment was performed considering the many replicas/block, the variation of the averages is studied assuming that blocks the effects of treatments gathered. Again is applied Fisher test, first there are been calculated the sizes of variance analysis table, necessary to determine the experimental value of the test (Samboan and Bad, 1986).

$$SSQ = 5,72^2 + 5,45^2 + ...3,37^2 / 3 - 44,52 = 3,06$$

$$GL_Q = n - 1 = 3, \quad S_Q^2 = SSQ / GL_Q = 0,05,$$

$$SSE = SS - (SSQ + SSX + SSB) = 0,74$$

$$GL_E = GL_T - (GL_Q + GL_X + GL_B) = 16$$

$$S_E^2 = SSE / GL_E = 0,04 \quad F_Q = S_Q^2 / S_E^2 = 1,25$$

For (3,16) GL and the confidence probability of 95%, the tabular value is 3.24.
Such, it follows that the interaction is not significant, so it can make a quantitative expression of the difference between the

treatments, which is applicable to all blocks (Feller, 1996).

The difference between treatments averages is

$$D = {(19,30 - 13,39)}/{12} = 0,49$$

If considered

$S_E^2 = 0,04$, $GL_E = 16$, $\alpha = 5\%$, $t_{\alpha/2} = 2,12$,

The confidence interval for D

$$\left(0,49 - 0,1; 0,49 + 0,1\right)$$

CONCLUSIONS

In conclusion, we may state that the addition of the NaCN to the medium of fish, reduced uptake by $D = 0,49 \pm 0,1$ $\mu\, mol\, g^{-1}$ dry weight per 20 min period.

REFERENCES

Feller W. 1996. An Introduction to Probability Theory and its Applications, John Wiley & Sons, Inc., New York, London, 40-55

Parker R.E. 1973. Introductory Statistics for Biology, London, 67-75.

G. Samboan G., Bad M.D., 1986. Elements of Mathematics for biologists, Bucharest. Univ. Ed., Bucharest, 289-313.

Varga M., 2014. Elements of Mathematics and Statistics with applications in biotechnology, Printech Ed., Bucharest, 199-204.

Varga M., 2010. Some results on effect of NaCN on the uptake in vitro of a particular amino acid by intestinal preparations from a certain species of fish,Proceedings of the International Symposium on New Researches in Biotechnology, Bucharest, 128-133.

FOOD
BIOTECHNOLOGY