Romanian Biotechnological Letters *Vol. 15, No. 6, 2010*

ORIGINAL PAPER

An application of variance analysis for establishing the effect of the saturated fats and of the manganese ions concentrations from the food ration on the cholesterol levels.

Received for publication, July 10, 2010
Accepted, November 8, 2010

VARGA MIOARA*, POPA CIPRIAN, TAMBA-BEREHOIU RADIANA*****
**Faculty of Biotechnology Bucharest (vargamioara@yahoo.com)*
***FARINSAN S.A. Giurgiu (cipnpopa@yahoo.com)*
**** Faculty of Biotechnology Bucharest (radianatamba@yahoo.com)*

Abstract

One of the most used tools in biology, for examining the variation within the whole group of experimental values determined for a sample from a statistical population, is represented by the analysis of variance. Some theoretical aspects of this method are presented, in order to describe the influence of two factors upon one character of a statistical population with normal distribution. The previous defined concepts were applied for solving a problem, aiming to establish the effect of the saturated fats concentrations and of the of manganese ions concentration in the food ration on cholesterol levels.

Keywords: *sum of squares of deviations, degrees of freedom, variance*

Introduction

In many problems of biology and agriculture, it is interesting to find out the influence of certain factors on a character of a statistical population that can be experimentally determined. Analysis of variance is a statistical technique applied to solve such situations, by the achievement of a study on the simultaneous activity of several variability factors and pointing out in the same time their interactions. This method was introduced in connection with the analysis of data and with the designing of agricultural field experiments. Some terms used to describe the analysis reflect this origin, for example “treatment”, “plot” and “block” as noted by R.E.Parker, 1973 [3].

In the present paper a control model for two factors of variability acting on the same character was applied, which was supposed to be normally distributed. After introducing the necessary theoretical elements for solving such problems, (based on the mentioned references) it was chosen the studying of the effect of the saturated fats concentrations and of the concentration of Mn ions in the food ration on cholesterol level, as was first discussed by G. Samboan, 1986 [4].

Material and Method

A brief characterization of variance analysis method is following, as previously described by several authors [1], [2], [5].

It was supposed that upon the character of a statistical population, which in terms of probabilities is a random variable normally distributed with mean μ and dispersion σ^2, two variability factors are acting simultaneously, their influence being evidenced by the mean of the X character. It was considered that the factor A has $A_1, A_2, ..., A_m$ - m versions and the factor

B has $B_1, B_2, ..., B_n$ - n versions. The experiment gives one observation for each couple $\left(A_i, B_j\right)$, $i = \overline{1,m}$, $j = \overline{1,n}$

The observational data can be arranged in the table below:

Table 1. The observational data for each couple (A, B)

B / A	A_1	A_2	...	A_m
B_1	x_{11}	x_{21}	...	x_{m1}
B_2	x_{12}	x_{22}	...	x_{m2}
...	...			...
B_n	x_{1n}	x_{2n}	...	x_{mn}

It was intended to establish how the factors A and B act on the character X by applying Fisher test. This involves the calculation of a value based on experimental observations, which is compared with a theoretical value corresponding to a particular confidence probability. If the experimental value is greater than the tabular one, then the influence of the factor on the studied character is significant.

The values involved in the problem of bifactorial variance analysis for finding the experimental value, noted $F_{\exp}$, can be arranged in a table, as follows:

Table 2. The table of variance analysis

Sources of variation	Sum of square	Degrees of freedom	Variances
A	SPA_A	GL_A	S_A^2
B	SPA_B	GL_B	S_B^2
R	SPA_R	GL_R	S_R^2
T	SPA_T	GL_T	

They have the following meanings:

1. The sum of squares of deviations (SPA)

- The sum of squares of deviations due to factor A:

$$SPA_A = m\sum_{i=1}^{m}\left(\overline{X_i} - \overline{X}\right)^2 . \tag{1}$$

- The sum of squares of deviations due to factor B:

$$SPA_B = n\sum_{j=1}^{n}\left(\overline{Y_j} - \overline{X}\right)^2 \tag{2}$$

- The total sum of squares of deviations:

$$SPA_T = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij} - \overline{X}\right)^2 . \tag{3}$$

- The sum of squares of deviations of error (residual):

$$SPA_R = SPA_T - SPA_A - SPA_B . \tag{4}$$

2. The degrees of freedom (*GL*)
- Degrees of freedom of factor *A*:

$$GL_A = m-1. \tag{5}$$

- Degrees of freedom of factor *B*:

$$GL_B = n-1. \tag{6}$$

-

Total degrees of freedom:

$$GL_T = mn-1. \tag{7}$$

- Degrees of freedom of error

$$GL_E = GL_T - GL_A - GL_B = (m-1)(n-1). \tag{8}$$

3. The variances (S^2)
- Variance of *A* factor:

$$S_A^2 = SPA_A / GL_A . \tag{9}$$

- Variance of *B* factor:

$$S_B^2 = SPA_B / GL_B . \tag{10}$$

- Residual variance:

$$S_R^2 = SPA_R / GL_R . \tag{11}$$

The influence of the factor *A* is highlighted by applying Fisher test to the variable $F_{\exp A} = S_A^2 / S_R^2$ and the influence of the factor *B* by using the appropriate variable $F_{\exp B} = S_B^2 / S_R^2$. In order to establish the connection strength between a factor *A* and the character *X*, the correlation index is calculated by formula:

$$I_C(A) = \sqrt{SPA_A / SPA_T} \tag{12}$$

If this index is close to value 1, then the correlation between factor *A* and character *X* is very significant. Similarly, for studying the correlation between a factor *B* and the character *X*, the correlation index is calculated:

$$I_C(B) = \sqrt{SPA_B / SPA_T} \tag{13}$$

Results and Discussions

The problem which was intended to be studied was to find the effect of the concentration of saturated fats and of the concentration of Mn ions from the food rations on the cholesterol levels.

This problem fits the model control of two factors of variability. We denoted by *X* = the cholesterol level, expressed as mg per 100 ml serum, representing the character studied, the analysis being applied on a sample of $n = 12$ bovines.

Two factors of variability were considered, with $m_1 = 3$, respectively $m_2 = 4$ variants each:

A= the concentration of saturated fats in the ration:

$A_1 = 35\%$ of the caloric value of the ration;

$A_2 = 30\%$ of the caloric value of the ration;

$A_3 = 25\%$ of the caloric value of the ration;

B = the concentration of Mn ions in the salt mixture:

$B_1 = 0.25\%$ mg to 100 grams of food,

$B_2 = 0.5\%$ mg to 100 grams of food,

$B_3 = 0.75\%$ mg to 100 grams of food,

$B_4 = 1\%$ mg to 100 grams of food

The obtained values are presented below:

Table 3. The table of values obtained

B/A	A_1	A_2	A_3	$\overline{Y_j}$
B_1	510	500	490	500
B_2	520	510	510	513.3
B_3	535	530	520	528.3
B_4	545	540	530	538.3
$\overline{X_i}$	527.5	520	512.5	

The selection mean, calculated by $\overline{X} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}}{n}$ was $\overline{X} = 520$. By reporting the working initial data to this mean, the following values are obtained:

Table 4. The table of differences between the obtained values and the mean

B/A	A_1	A_2	A_3	Total
B_1	-10	-20	-30	-60
B_2	0	-10	-10	-20
B_3	15	10	0	25
B_4	25	20	10	55

In order to determine the influence of factors A and B on cholesterol, we calculate the magnitudes needed to complete the table of dispersion analysis were calculated, such as: the sum of square of deviations, the degrees of freedom, the variances, the experimental values and the table values of the Fisher test:

$$SPA_A = 3\sum_{i=1}^{3}\left(\overline{X_i} - \overline{X}\right)^2 = 337.5$$

$$SPA_B = 4\sum_{j=1}^{4}\left(\overline{Y_j} - \overline{X}\right)^2 = 3394.68$$

$$SPA_T = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij} - \overline{X}\right)^2 = 3850.$$

Table 5. The table of analysis of variance

Sources of variation	Sum of square	Degrees of freedom	Variances	F
A	$SPA_A = 337.5$	$GL_A = 2$	$S_A^2 = 337{,}5 / 2 = 168.75$	$F_{\exp A} = S_A^2 / S_R^2 = 7.16$
B	$SPA_B = 3394.68$	$GL_B = 3$	$S_B^2 = 3394{,}68 / 3 = 1131.56$	$F_{\exp B} = S_B^2 / S_R^2 = 48.02$
T	$SPA_T = 3850$	$GL_T = 11$		
R	$SPA_R = 117.82$	$GL_R = 5$	$S_R^2 = 117{,}82 / 5 = 23.564$	

By applying the (12) and (13) formulas to calculate the correlation indices, it was obtained:

$$I_C(A) = \sqrt{SPA_A / SPA_T} = 0.29, \quad I_C(B) = \sqrt{SPA_B / SPA_T} = 0.93 .$$

Conclusions

The application of described statistical method for the existing experimental data, has conducted to the following result: if the risk is considered $\alpha = 5\%$, the superior limit of Fisher test is 6.94. The experimental values $F_{\exp A} = 7.16$ and $F_{\exp B} = 48.02$ are higher than the table value, so we may conclude that both factor *A* (the concentration of saturated fats in food ration) and the factor *B* (concentration of Mn ions) significantly influence the cholesterol level. Since the correlation index $I_C(B) = 0.93$ is very close to value 1, we can say that the correlation between the concentration of Mn ions in the food ration and the cholesterol levels is very significant.

References

1. E. ALMANN, J. RHODES, *Mathematical Models in Biology*, London, 2003, pp. 35-57.
2. W. FELLER, *An Introduction to Probability Theory and its Applications*, John Wiley & Sons, Inc., New York, London, 1996.
3. R.E. PARKER, *Introductory Statistics for Biology,* London, 1973, pp.67-75.
4. G. SAMBOAN, M.D. BAD, *Elements of Mathematics for biologists*, Bucharest. Univ. Ed., Bucharest, 1986, pp.289-313.
5. M. VARGA, *Elements of Mathematics and Statistics with applications in biotechnology* Printech Ed., Bucharest, 2009, pp.199-204.