PROGRAM SUPPORT FOR SOLUTION TO SCIENTIFIC TASKS IN MEAT PROCESSING INDUSTRY

Andrey B. Lisitsyn[1], Marina A. Nikitina[1], Egor B. Sus[1]

[1] FGBNU V.M. Gorbatov VNIIMP, Moscow, Russia

Abstract – This paper presents the statistical support subsystem for biological research (SSS Bio). The area of statistical methods application in biology is significant as many ecological, genetic, cytological, microbiological and radiobiological events are mass in their nature. It is not a single individual or a cell, not a single α-particle, not a single bacterium, but multitudes, i.e., aggregates of cells, α-particles, bacteria, individuals of a species, mixtures and so on that take part in these events. Occurrence of events in these aggregates can be assessed by probabilities, and their analysis requires use of statistical methods. Statistical methods are also essential in performing experiments because only with their support it is possible to establish whether a difference in an observed event between experimental and control plots and animal groups depends on a tested factor or it is random.

The developed software SSS Bio makes it possible to process statistical data and obtain a stochastic dependence (a mathematical model). SSS Bio realizes the processing algorithms for an active experiment (complete factorial, fractional factorial and orthogonal central composite) and passive experiment (one-way analysis of variance, linear correlation test, rank correlation test with use of Spearman's and Kendall's coefficients, linear and multiple regression analysis).

Key Words – stochastic dependences, active and passive experiments, prediction of results.

I. INTRODUCTION

Prospects of meat industry development to a large extent depend on efficiency of scientific research in the area of applied biotechnology and incorporation of the results into the industry. In solving the tasks of scientific research management it is necessary to assess their prospects, evaluate the contribution of individual developments in solution to set problems and develop a strategy for implementation of the obtained results. In connection with the complexity and the depth of the above mentioned problems, the specialists with knowledge on the specific technologies of meat raw material processing who are able to use the instruments of computing techniques are required.

The methods of the mathematical statistics broaden the possibilities of scientific prediction and expedient decision-making in the conditions of uncertainty, when a researcher in principle cannot know the whole complex conditions of performing an experiment.

One of the most important factors that facilitates further development of various areas of natural sciences including meat processing branch is implementation of information technologies and, as a consequence, mathematics into these areas.

Many years ago Russian scientist Pavlov [1] said: "The whole life from the simplest to the most complex organisms, including, of course, humans, is a long range of environment equilibrations with their increasing complication up to the highest degree. Time will come, although the remote, when the mathematical analysis based on the natural scientific one, will embrace by the magnificent equations all these equilibrations eventually incorporating itself into them."

The modern information technologies based on application of a wide range of mathematical models and methods, and realized in a form of programs for personal computers can be successfully used both in conducting scientific research works associated with establishing regularities, predicting results of technological processes, and in conditions of the real production when choosing optimal ways of processing valuable biological raw material and rational use of production capacities.

II. MATERIALS AND METHODS

SSS Bio was developed in the programming environment Embarcadero Delphi 2010. The program was developed and tested on a personal computer with 64-bit operating system Microsoft Windows 7. When designing SSS Bio, the external

module Matrices.pas was used. This module was realized by JKCoders Team, has the actual version 1.0 and is used for realization of TMatrix class, which allows performing the main actions on matrices.

The work with SSS Bio program is carried out after starting the executing file SSSBio.exe (Fig.1).

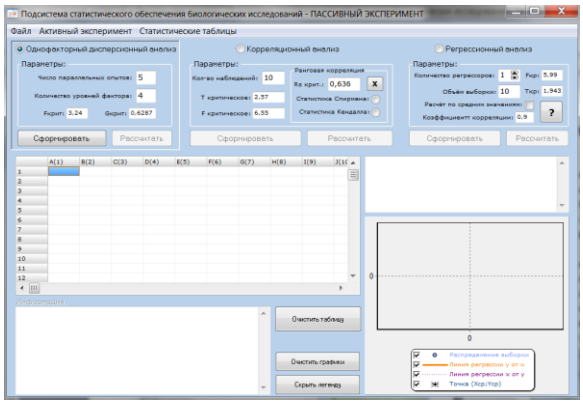

Figure 1. Program interface upon starting

In order to create mathematical models in the form of the stochastic dependences, the statistic information about an object under investigation is required. There are various methods for collection and accumulation of the necessary initial information. The main sources of information are experimental investigations. A distinction is made between passive and active experiments [2-5].

A passive experiment is characterized by a large number of experiments with successive variation of each test factors. The coefficients of the regression model are determined by the method of least squares and depend on each other. The presence of insignificant coefficients requires recalculation of all other coefficients.

An active experiment is carried out according to a preliminarily drawn plan. Simultaneous changing of all factors is stipulated, which makes it possible to establish the strength of relationships between them. In an active experiment, the number of trials is significantly reduced. Coefficients of the obtained model are independent of each other. Presence of insignificant coefficients does not require recalculation of all others.

In both cases, the target mathematical model is the response function *y*, which links the parameter of optimization that is characterized by the experimental results with the variables that are varied when performing trials.

The user interface of SSS Bio program consists of two main forms (active (Fig. 2) and passive (Fig.3) experiments) and one auxiliary for creating statistical tables (Fisher's F-test, Student t-test, Cochran's G test)

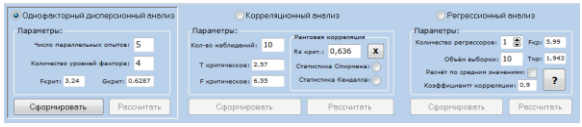

Fig. 2 Menu of the initial settings of the input parameters of a passive experiment

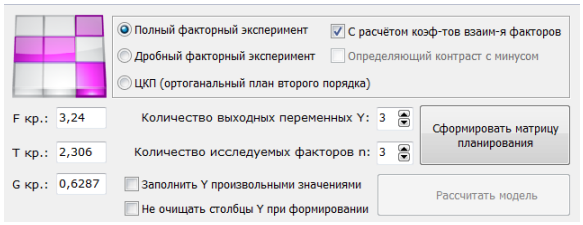

Fig. 3 Settings of an active experiment

To improve a perception of the results, the text fields are embedded into SSS Bio program, into which the statistical data, some explanations related to computations as well as recommendations (for example, in case of an inadequate model) are displayed.

Another important feature of SSS Bio program is the possibility to build regression charts and visual presentation of a sample arrangement. This function can often be useful directly for researchers as they can visually assess correlation and regression relationships.

Tabular results of a conducted analysis can be saved both in a new file in the format *.xls or *.xlsm and in a previously opened file. To this end, it is necessary to click *File* in the Main Menu and select *Save* or *Save as*.

III. RESULTS AND DISCUSSION

As an example of the system work, we examine the regression analysis; the initial data presented in Table 1 were taken from the book [6].

Table 1 The data on the live weight and chest circumference in cows

	X	Y	X*X	XY
1	225	145	50625,00	32625,00
2	275	156	75625,00	42900,00
3	325	160	105625,00	52000,00
4	375	166	140625,00	62250,00
5	425	170	180625,00	72250,00
6	475	175	225625,00	83125,00
7	525	182	275625,00	95550,00
8	575	182	330625,00	104650,00
Сумма	3200,00	1336,00	1385000,00	545350,00

Table 1 demonstrates the columns of the initial data as well as the columns x^2 and xy. The last row shows the total values.

Fig. 4 presents the graph of the analyzed data. In addition, the dotted line shows the inverse relationship. This regression line is actual only if both traits have random variation (as in the present example), i.e., the regression is two-way.

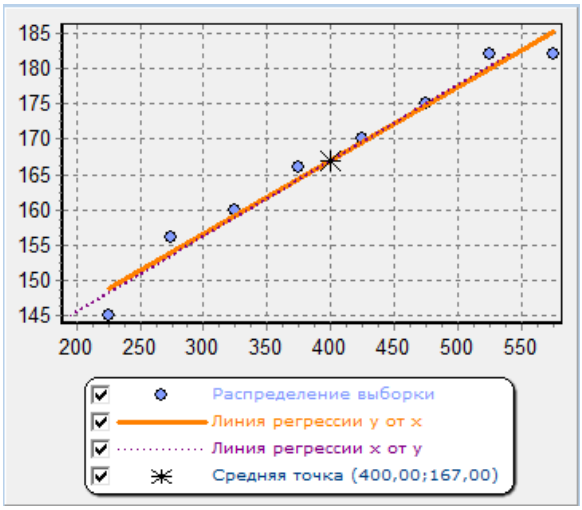

Figure 4. Charts of regression lines and sample distribution

Analyzing data from Fig. 4, it is possible to make a conclusion without additional calculations about the close linear correlational relationship of the traits.

Table 2 presents the results of the other calculations of the regression analysis.

As a result of performing the regression analysis, SSS Bio program presented the comprehensive quantity of statistical data.

Table 2 The results of the regression analysis

TextField *Information*	Additional TextField
Coeff. a = 125.29 Coeff. b = 0.10 Correlaion coefficient : r=0.9846 Determination coefficient: R^2=0.969 Regression equation: Y = 125.2857 + 0.1043 * X Student test t_b = 13.7820 Student test t_a = 39.7933 Confidential interval for α and β: 0.090 < β < 0.119 119.168 < α < 131.403 The analyzed relationship is statistically significant! F = 189.9446 The model is adequate!	Sampling variance X = 13125.00 Sampling variance Y = 147.25 Standard deviation: S(x)=114.564 Standard deviation: S(y)=12.135 Standard error of regression Sy= 2.452 Standard deviation of random variables a and b: Sa = 3.1484 Sb = 0.0076

Thus, the following regression equation was obtained: $y = 127 + 0{,}10x$. Substituting different values of x into the regression equation, we obtain the theoretical values of the chest circumference, which correspond to the values of live weight.

As an example of building linear models of the first-order design, we used the data from the textbook [5].

Perform a complete factorial experiment for detection of the actual acidity of an environment (pH) in meat cans from beef after primary mechanical processing and ageing.

Varying factors and the range of their change: x1 – duration of the primary mechanical processing (5-40 min.); x2 – duration of the primary ageing (1-48 hours); x3 – the quantity of the added brine (10-25% of the volume). Controlled value: y – pH.

The number of replicate trials q=3.

The results of the calculation are presented in Fig.5; the intermediate results are given below.

	X1*X2	X1*X3	X2*X3	X1..*X3	Y1	Y2	Y3	Y
1	1	1	1	-1	6,752	6,63	6,48	6,62
2	-1	-1	1	1	5,79	5,69	5,91	5,80
3	-1	1	-1	1	6,38	6,39	6,33	6,37
4	1	-1	-1	-1	6,36	6,38	6,51	6,42
5	1	-1	-1	1	6,36	6,31	6,48	6,38
6	-1	1	-1	-1	6,55	6,35	6,25	6,38
7	-1	-1	1	-1	5,83	5,77	5,73	5,78
8	1	1	1	1	6,21	6,6	6,63	6,48
b	0,20	0,18	-0,11	-0,02				

Figure 5. The design matrix

The results of the calculation:

Calculation of the coefficients of regression equation:
b[0] = 6.2780
b[1] = -0.0088
b[2] = -0.0180
b[3] = -0.0222
b[12] = 0.1972
b[13] = 0.1847
b[23] = -0.1095
b[123] = 0.0213
Row-wise dispersion calculation:
1 row: 0.019
2 row: 0.012
3 row: 0.001
4 row: 0.007
5 row: 0.008
6 row: 0.023
7 row: 0.003
8 row: 0.055
Maximum dispersion: 0.055
Cochran's test G = 0.433
Dispersion is uniform – the experiment is reproducible!
The experiment error: 0.016

Assessment of significance of regression coefficients:
T[0]=957.405 significant – the corresponding factor is significant!
T[1]=13.379 significant - the corresponding factor is NOT significant!
T[2]=27.264 significant - the corresponding factor is NOT significant!
T[3]=33.575 significant - the corresponding factor is NOT significant!
T[12]=298.640 significant – the corresponding factor is significant!
T[13]=279.707 significant – the corresponding factor is significant!
T[23]=165.855 significant – the corresponding factor is significant!
T[123]=32.313 significant - the corresponding factor is NOT significant!
Fisher's test F = 0.0000
THE MODEL IS ADEQUATE !

As a result, we obtained the mathematical model $y = 6.278 + 0.1972x_1+0.1847x_2-0.1095x_3$.

Analysis of the regression equation obtained by the results of the complete factorial experiment for the single-stage mechanical processing and ageing revealed that the interaction of the factors practically did not influenced the controlled pH value.

IV. CONCLUSION

The wide use of statistical methods in biological research is connected with the really existing uncertainty in solving practical tasks. This uncertainty is determined by a number of reasons: the errors of the instruments of measurement and control, impossibility to account all affecting factors and so on. The statistical methods do not eliminate this uncertainty; however, they allow its quantitative assessment, which make it possible to obtain the best solutions at all stages of research.

REFERENCES

1. Pavlov, I.P. (1951). The Collected Works. Moscow: Publishing House of the Academy of Sciences of the USSR.
2. Akhnovarova, S.L., Kafarov, V.V. (1985). Experiment optimization in chemistry and chemical technology. Moscow: Higher School.
3. Kafarov, V.V. (1985) Cybernetic methods in chemistry and chemical technology. Moscow: Chemistry.
4. Ivashkin, Yu.A., Protopopov, I.I., Borodin, A.V. et al. (1987). Modelling of production processes and production in meat and dairy industry. Moscow: Agropormizdat.
5. Borodin, A.V., Nikitina M.A. (2008). Information technologies of production in biological raw material processing. Moscow: LLC “Polysuvenir”.
6. Rokitsky, P.F. (1967). Biological statistics. Minsk: Higher school.