DEVELOPMENT OF REGRESSION MODEL OF PROTEINS ATTACKABILITY PROCESS IN MEAT FOOD (IN VITRO)

Alisher S. Pulatov1,*, Marina A. Nikitina2
1 Namangan Engineering-Construction Institute, Namangan, Republic of Uzbekistan
2 V. M. Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences, Moscow, Russia

Keywords: attackability, regression model, meat product, matrix of pair correlation coefficients

Abstract
In the presented article the authors consider the issues of development of regression model for process of food digestion by proteolytic enzymes in human body. The authors use correlation analysis. They analyze the main nutritional values and physical and chemical properties of meat products, the modes of heat treatment of semi-finished lamb products. The essential parameters and features are determined to find the dependence between the factor values and efficient values of the basic raw material, which affect the quality of the technological processes and, in general, the finished product. The regression model equation is mathematically calculated by methods of solving K. Gauss linear equations. The standard deviations of parameters are calculated, the initial data are normalized; the matrices of the pair correlation coefficients, lower and upper limits of their values are compiled. Equations of the mathematical regression model of meat proteins attackability by proteolytic enzymes — in vitro (pepsin, trypsin) are developed. It is proved that the obtained equation represents a regression model of the process of meat food proteins attackability by enzymes (pepsin, trypsin and chymotrypsin), depending on the determined 3 essential factors (weight of a meat piece, duration of frying, collagen content in lamb meat). Also this equation reflects the process of lamb digestibility in a digestive tract of a human body.

Introduction
The rate of proteins digestion in the gastrointestinal tract, or attackability of the proteins in the composition of meat dish by proteolytic enzymes is one of the important factors that determine the biological value of food products [1–5]. Nowadays the theory and methods of correlation analysis can be successfully applied to study mathematical problems, relations between phenomena and features in various fields of science, technology and national economy. Correction relation is considered as established when a row of function values correspond to the same value of the argument. The features that characterize this relation are divided into factor and efficient. The features that affect a certain result are called factor features. And the features that respond to factor features are called efficient [6].

Correlation-regression analysis is often used in scientific research. When processing the big amounts of statistical data, correlation analysis quantifies the strength of link between two or more quantitative variables. Regression or correlation analysis describes the link between variables, while the correlation provides a numerical way to measure the level or strength of link between two variables [7,8].

Here are the examples of some researches that cover the application of correlation-regression analysis in various fields of science. The research [9] assessed the capacity of an interactive dual-energy X-ray absorptiometer (DEXA) installed on the slaughter line of a meat processing plant to determine the composition of lamb carcasses. 607 lamb carcasses from 7 slaughter groups were scanned by DEXA device and later were scanned by computed tomography device to determine the ratio of fat, lean meat and bone in carcasses. The results of those test across whole data range showed high accuracy of body fat percentage forecasting by computed tomography device, with coefficient of determination (R2) = 0.89, compared to 0.69 for lean meat and 0.68 for bones in carcass, that showed less accuracy. Accuracy in the seven groups was also high in comparison with the mean values of bias 0.66, 0.83 and 0.51.

The researches [10] carried out by the Danish Research Institute of Meat are conducted to develop the forecasting models for the shelf life of meat products. Expiration dates of chilled meat were simulated for beef cuts (850 samples), pork cuts (1500 samples) and chicken (1080 samples), minced beef and minced pork (680 samples of each type of minced meat) and for bacon (1080 samples). In this case, the samples of meat and meat products were packed in modified atmosphere into vacuum bags in combination with microwave treatment at various storage temperatures. Research showed that forecasting models serve as tools able to assess the importance of temperature and packaging changes for the shelf life of various meat products. In the research [11] volatiles were studied during roasting beef at a temperature of 180 °C. 70 volatile substances were identified, including non-aromatic, homocyclic and heterocyclic compounds. A significant positive regression model was constructed to forecast the storage of toluene, benzol acetaldehyde, 2-formylfuran, pyrazine, 2,6-dimethylpyrazine, 2,3-dimethylpyrazine, 2-acetylthiazole, and

FOR CITATION: Pulatov, A.S., Nikitina, M.A. (2021). Development of regression model of proteins attackability process in meat food (in vitro). Theory and practice of meat processing, 6(3), 236-241. https://doi.org/10.21323/2414-438X-2021-6-3-236-241

236
2-formyl-3-methylthiophene. For calculation of aging time a linear and logarithmic regression model was chosen.

In a research [12] the authors studied the influence of the carcass parts weights (thigh, breast, wing, back, stomach, heart) on whole carcass weight of white turkeys (Big-6). Data were analyzed with the help of regression analysis based on ridge regression and factor analysis. Both regression models were found to be suitable for turkey carcass weights forecasting. However, the ridge regression method was preferred, as it showed higher R² value and explains carcass weight in a better way.

In the study [13] three varieties of wheat (PKB Talas, BG Merkur and PKB Lepoklas) harvested in 2009 and 2010 were studied. The correlations between the morphological and yield parameters of plants were studied: the number of shoots, the number of spikelets on a wheat head, the number of grains per wheat head, the weight of 1,000 grains and mass of grain per wheat head. Taking into account the parameters of all three varieties, high and positive correlations were found between the number of grains per wheat head and the weight of grain per wheat head (> 0.78), the number of spikelets per wheat head and the number of grains per wheat head (> 0.79), as well as the number of spikelets per wheat head and grain weight per wheat head (> 0.73). Regression analysis was conducted only as an addition to the correlation formula, and was presented in the form of charts showing the link between the studied dependent and independent characteristics.

Researchers in the study [14] constructed a quadratic polynomial that explains the link of three variables (fermentation temperature, X₁; amount of inoculation, X₂; and concentration of solid substrate, X₃) on yield of monacolin K.

The goals of the study [15] were to define: (1) the effect of fertilizers, the environment and their interactions on the thousand grains weight (TGW), hectolitre weight (HW) and grain yield (GY) of winter triticale variety and (2) the correlation between these characteristics in various environments. Negative and significant correlation was found between GY and TGW (minus 0.392) in 2015, positive highly significant correlation was found in 2013 (0.648) and 2014 (0.493).

Recently, researches have appeared which study the effects of meat consumption of health of population in the modern world, including issues related to the consumption of saturated fat [16]. Regression analysis was used to determine the effect of stearic acid on total cholesterol (TC) level in blood plasma and low-density lipoprotein (LDL) cholesterol when people of various ages consume fatty foods is described in researches [17,18,19].

With an increase in standards of living and well-being of peoples of Central Asia, and in particular, the Republic of Uzbekistan, in the diet of population the consumers' demand for natural food products and a wide range of lamb meat dishes cooked in traditional ways increases. In accordance with modern concepts of nutritional science, food must have an attractive appearance and high taste, and also food shall be biologically complete, i.e. the food shall contain all necessary essential amino acids and other important food components in optimal proportions. Those food components are well digested in a human body by digestive enzymes [20,21].

The basis of enzymatic hydrolysis of meat proteins according to method of A. A. Pokrovsky and I. D. Erta-nov [5] is the conditions when the availability of attacked peptide bonds in a meat is determined not only by its physical and chemical parameter of proteins, but also by characteristics related to the structure and chemical composition of the basic food product. It is important and necessary to conduct scientific research in this sphere using mathematical methods of analysis, as they prove the effectiveness of the approaches, the possibility of determining the calculated parameters of the biological and nutritional value of food products and comparing them with the FAO data.

The purpose of our study was to build a regression model of the most important processes of food digestion by proteolytic enzymes (pepsin, trypsin), using correlation analysis and expressing the analytical dependence of the efficient characteristic (Y-process of meat proteins attackability) on factor characteristics $x_i, i = 1, n$.

To achieve this goal the following tasks were solved: based on 20-fold physical and chemical analyzes to obtain the necessary digital parameters of the analyzed meat product (meat) and based on results of obtained data, calculate their position and dispersion characteristics, determine the influence of significant factors — (modes of heat treatment, quantitative content of imperfect proteins) for efficient features (quality of semi-finished meat products).

**Objects and methods**

To conduct the scientific research we selected and prepared sample materials, as well as chose experimental methods for testing. The object of research and study was: meat productivity, quality of meat and raw fat of fat-tailed sheep of “Jaydara” breed, popular in the foothill regions of the Republic of Uzbekistan. Experimental studies analyzed 14 parameters of meat product quality, including the parameter fully covered in this article.

The process of hydrolysis (digestion of proteins in vitro) was run in the Department of Food Technology laboratory on a special 3-cells device that provides continuous mixing and dialysis of the samples, and the products of their breakdown were analyzed by micro-methods that allow simultaneous analysis of a significant number of samples under study [22].

The meat product containing about 150 mg of protein (N2 x 6.25 or x 5.75) is placed in the inner vessel of the specified device, then 15 ml of 0.02 N solution of HCl with pH 1.2 is added into the vessel. In order to comply with the isoionic process, it is also necessary to add 60 ml of
the same solution to the outer vessel. The test samples are incubated in the thermostat at a temperature of 37 °C.

The index of the proteins attackability of meat semi-finished products by proteolytic enzymes (pepsin, trypsin) was estimated by build-up of hydrolysis products in individual samples. The calculations were made according to the corresponding formulas.

For statistical data processing and construction of 1) matrix of pair correlation coefficients; 2) the regression equations, we used “Data Analysis” in a spreadsheet processor MS Excel.

Results and discussion

In order to conduct the mathematical analysis, preliminarily we studied the process of proteins attackability in lamb meat food cooked in various modes and duration of heat treatment, taking into account the unequal content of connective tissue proteins in meat (collagen and elastin), as well as the content of pure hydroxyproline in the samples under study [21].

As a result of physical and chemical analyzes data were obtained, which were subsequently presented as initial data, where:

- $X_1$ — is weight of meat pieces (weight, g);
- $X_2$ — deep fry temperature (°C);
- $X_3$ — duration of frying (min);
- $X_4$ — collagen content in meat (%);
- $X_5$ — elastin content in meat (%);
- $X_6$ — oxyproline content in meat (mg%),
- $Y$ — degree of attackability (mg of hydrolyzed protein).

The layout of the initial data table is compiled on the basis of laboratory analysis values and is presented below in Table 1.

Based on the initial data, the parameters of the multiple regression equation were calculated as follows:

- The arithmetic mean value of each parameter is determined according to the formula (1):
  \[ \bar{X}_k = \frac{\sum_{i=1}^{n} x_{ik}}{n} \]  
  (1)

where $k$ — is number of factors (in our case $k = 1, 2, \ldots, 6$); $\bar{X}_k$ — arithmetic mean of the $k$th factor; $x_{ik}$ — value of the $i$th measurement of the $k$th factor ($i = 1, 2, \ldots, n$); $n$ — number of trials.

The obtained values are presented in the last line of the Table 1 and are highlighted in orange.

- The standard deviation of each parameter was determined by the formula (2):
  \[ C_{sr} = \sqrt{\frac{\sum_{i=1}^{n} (x_{ir} - \bar{X}_r)^2}{n-1}} \]  
  (2)

where $C_{sr}$ — is the mean-square deviation of the $r$th factor.

Further, the initial data were normalized (Table 1) according to the following formula:

\[ x^{*} = \frac{x_{ki} - \bar{X}_k}{C_{sr}} \]  
(3)

where $x^{*}$ — is rationing the value of the $k$th factor.

After normalizing the data the following results were obtained (Table 2).
— The pair correlation coefficients can be calculated according to the formula (4) or using the spreadsheet processor “Data Analysis” in MS Excel.

\[ R_{ks} = \frac{1}{n} \sum \frac{x^*_k \cdot x^*_s}{s} \]  

(4)

where \( R_{ks} \) — is coefficient of correlation dependence between factors \( k \) and \( s \).

The results of calculation are presented in the form of a matrix of pair correlations (Table 3).

**Table 3. Matrix of pair correlation coefficients**

|     | Y  | X1 | X2 | X3 | X4 | X5 | X6 |
|-----|----|----|----|----|----|----|----|
| Y   | 1  |    |    |    |    |    |    |
| X1  | -0.38 | 1 |    |    |    |    |    |
| X2  | -0.33 | 0.39 | 1 |    |    |    |    |
| X3  | 0.94 | -0.27 | -0.45 | 1 |    |    |    |
| X4  | -0.97 | 0.41 | 0.36 | -0.93 | 1 |    |    |
| X5  | -0.33 | 0.22 | 0.29 | 0.30 | 0.35 | 1 |    |
| X6  | -0.50 | 0.28 | 0.13 | -0.34 | 0.57 | 0.64 | 1 |

Analyzing the data presented in the Table 3, it is obvious that the coefficients of pair correlation between Y and factors X1, X2, X3 feature rather low values (all coefficients in absolute value are below 0.4), and it means that the link is quite weak. The moderate link is observed between Y and X4 (correlation ratio is minus 0.50). Strong link is defined between Y and X5, X6 (correlation coefficients to modulo exceed 0.9). In addition to the strong link is observed between variables X1 and X6 (correlation ratio is minus 0.93), moderate link is determined between variables X2 and X5 (correlation ratio is minus 0.96), X3 and X5 (correlation ratio is minus 0.93), and X2 and X6 (correlation ratio is minus 0.97). It means that these coefficients can be collinear.

After finding the pair correlation coefficients, the partial correlation coefficients were determined by the following formula:

\[ r_{yxi \mid xj} = \frac{r_{yxi} - r_{yxj} \cdot r_{xj \mid xi}}{\sqrt{(1 - r_{yxi}^2)(1 - r_{xj \mid xi}^2)}} \]  

(5)

As result the following values are obtained:

| Yx1 | -0.290 | Rx1x3 | -0.368 | Rx1x4 | 0.075 | Rx1x5 | -0.334 | Rx1x6 | -0.288 |
|-----|--------|-------|--------|-------|-------|-------|--------|-------|--------|

Correlation ration is low everywhere

| Rx2x1 | -0.209 | Rx2x3 | 0.309 | Rx2x4 | 0.107 | Rx2x5 | -0.255 | Rx2x6 | -0.303 |
|-------|--------|-------|--------|-------|-------|-------|--------|-------|--------|

Correlation ratio is low everywhere

| Rx3x1 | 0.938 | Rx3x2 | 0.937 | Rx3x4 | 0.369 | Rx3x5 | 0.932 | Rx3x6 | 0.943 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|

Correlation ratio is high (to modulo ober 0.9), except for Rx3x4. In this case the correlation is low.

| Rx4x1 | -0.969 | Rx4x2 | -0.971 | Rx4x3 | -0.789 | Rx4x5 | -0.970 | Rx4x6 | -0.967 |
|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|

Correlation ratio is high everywhere

| Rx5x1 | -0.275 | Rx5x2 | -0.260 | Rx5x3 | -0.161 | Rx5x4 | 0.039 | Rx5x6 | -0.015 |
|-------|--------|-------|--------|-------|--------|-------|-------|-------|--------|

Correlation ratio is low everywhere

| Rx6x1 | -0.448 | Rx6x2 | -0.492 | Rx6x3 | -0.557 | Rx6x4 | 0.292 | Rx6x5 | -0.403 |
|-------|--------|-------|--------|-------|--------|-------|-------|-------|--------|

Correlation ratio is low for Rx6x4, moderate for others.

To establish a connection between all factors and the resulting characteristic, the multiple regression coefficient is found:

\[ R_{yx1 \mid x2} = \frac{\sum \beta_i \cdot r_{yxi}}{\sqrt{\sum \beta_i^2}} \]  

(6)

As a result of the calculations, we found that the coefficient of multiple regression is 0.98, which proves a strong correlation between the entire set of factors and the result. The unadjusted coefficient of multiple determination shows that 96% of the variation in the result are explained by the variation of the factors presented in the equation. The adjusted coefficient of multiple determination defines the correlation ratio, taking into account the degrees of freedom of the total and residual variances and is calculated by the following formula:

\[ R^2_{corr} = 1 - (1 - R^2) \cdot \frac{n-1}{n-m-1} \]  

(7)

The adjusted multiple regression coefficient is 0.94, i.e. practically equal to 1, i.e. the regression equation explains the variation of attackability.

To build a linear multiple regression model, “Data Analysis” in MS Excel (Figure 1) was used.

---

**Figure 1. Regression analysis data**
We came to the conclusion that it is feasible to include the regression equation, as \( F_{\text{table}}(51.64) > F_{\text{obs}}(0.00) \) with probability \( 1 - \alpha = 0.95 \).

Using the particular of F-test of Fisher, we assessed the feasibility of including the factors \( x_i \) in the multiple regression equation after the other factors:

\[
F_{x_i} = \frac{R^2(x_{i+1}, x_n)}{1 - R^2} \cdot (n - m - 1) \quad (9)
\]

where \( R^2(x_{i+1}, x_n) = \sum \beta_i^2 r_{iX} \).

We came to conclusion that it is feasible to include \( x_1, x_2, x_3 \) into the regression equation.

To switch from standardized values to natural values, the following formula is used:

\[
A_i = \frac{C_y}{C_{xi}} \cdot b_i \quad (10)
\]

Where \( A_i \) is the natural coefficient of the equation; \( C_i \) is the standard deviation of the factor; \( C_{xi} \) the standard deviation of the parameter.

Thus the regression equation in natural values looks as follows:

\[
y = 89.41 - 0.033x_1 + 0.516x_2 - 26.792x_3 \quad (11)
\]

This equation represents a regression model of the process of meat proteins attackability by enzymes (pepsin, trypsin and chymotrypsin), depending on the defined 3 essential factors (mass of a piece of meat; duration of frying; collagen content in meat) and reflects the process of digestibility of meat products in a human digestive tract.

It can be seen from the equation that while the mass of a meat piece increases by 1 unit, its attackability also increases by 0.516 units. While the collagen content in meat increases, the protein attackability also increases by 0.516 units. While the collagen content in meat increases, attackability decreases by 26.792 units.

**Conclusion**

As a final summary of the research it is possible to conclude that the analyzes, mathematical calculations, statistical analysis of reliability of the obtained regression equation prove the possibility of successful application of correlation-regression analysis for the calculations of biological and nutritional value assessment — i.e. the processes of protein breakdown in various finished lamb culinary products by proteolytic enzymes (in vitro).

**REFERENCES**

1. Anwar, A. (1962). Evaluation of proteins by in vitro pancreatic digestion. *Poultry Science*, 41(4), 1120–1123. https://doi.org/10.3382/ps.0411120
2. Akesson, W.R., Stahmann, M.A. (1964). A pepsin pancreatic digest index of protein quality evaluation. *The Journal of Nutrition*, 83(3), 257–261. https://doi.org/10.1093/jn/83.3.257
3. Denton, A.E., Elvehjem, C.A. (1953). Enzymatic liberation of amino acids from different proteins. *The Journal of Nutrition*, 49(2), 221–229. https://doi.org/10.1093/jn/49.2.221
4. Evans, R.J., Butts, H.A. (1949). Inactivation of amino acids by autoclaving. *Science*, 109(2840), 569–571. https://doi.org/10.1126/science.109.2840.569
5. Pokrovskiy, A.A., Erganov, I.D. (1965). The attackability of food proteins by proteolytic enzymes in vitro. *Problems of Nutrition*, 3, 38–44. (In Russian)
6. Pandey, S. (2020). Principles of correlation and regression analysis. *Journal of the practice of cardiovascular sciences*, 6(1), 7–11. https://doi.org/10.4103/jpcs.jpcs_2_20
7. Draper, N.R., Smith, H. (1966). Applied regression analysis, 2nd ed.: New York. John Wiley, 1996.
8. Snedecor, G.W., Cochran, W.G. (1967) Statistical methods, 6th ed.: Ames, Iowa State Univ. Press, 1967.
9. Gardner, G.E., Starling, S., Charnley, J., Hocking-Edwards, J., Peterse, J., Williams, A. (2018). Calibration of an on-line dual energy X-ray absorptiometer for estimating carcass composition in lamb at abattoir chain-speed. *Meat Science*, 144(10), 91–99. https://doi.org/10.1016/j.meatsci.2018.06.020
10. Terngren, M.A., Darré, M., Gunvig, A., Bardsenshtein, A. (2018). Case studies of packaging and processing solutions to improve meat quality and safety. *Meat Science*, 144(10), 149–158. https://doi.org/10.1016/j.meatsci.2018.06.018
11. Watanabe, A., Kamada, G., Imanari, M., Shiba, N., Yonai, M., Muramoto, T. (2015). Effect of aging on volatile compounds in cooked beef. *Meat Science*, 107(9), 12–19. https://doi.org/10.1016/j.meatsci.2015.04.084
12. Sogut, B., Kurbal, O.F., Demirulus, H., Inci, H. (2010). Growth performance of Big-6 broiler white turkeys in the different rearing conditions. *Journal of animal and veterinary advancement*, 9(9), 1334–1337. https://doi.org/10.3923/java.2010.1334.1337
13. Djuric, N., Prodanovic, S., Brankovic, G., Djekic, V., Cvijanovic, G., Zilic, S. et al. (2018). Correlation-Regression Analysis of Morphological-Production Characteristics of Wheat Varieties. *Romanian biotechnological letters*, 23(2), 13457–13465
14. Ding, Y., Pu, L., Kan, J. (2017). Hypolipidemic effects of lipid-lowering granulated tea preparation from Monascus-fermented grains (adlay and barley bran) mixed with lotus leaves on Sprague-Dawley rats fed a high-fat diet. *Journal of Functional Foods*, 32(5), 80–89. https://doi.org/10.1016/j.jff.2017.02.025
15. Rajičić, V., Popović, V., Perišić, V., Biberđić, M., Jovović, Z., Gudžić, N. et al. (2020). Impact of nitrogen and phosphorus on grain yield in winter triticale grown on degraded vertisol. *Agronomy*, 10(6), article article 757. https://doi.org/10.3390/agronomy10060757
16. Mann, N.J. (2018). A brief history of meat in the human diet and current health implications. *Meat Science*, 144(10), 169–179. https://doi.org/10.1016/j.meatsci.2018.06.008
17. Derr, J., Kris-Etherton, P.M., Pearson, T.A., Seligson, F.H. (1993). The role of fatty acid saturation on plasma lipids, lipoproteins, and apolipoproteins: II. The plasma total and low-density lipoprotein cholesterol response of individual fatty acids. *Metabolism*, 42(1), 130–134. https://doi.org/10.1016/0026-0495(93)90183-0
18. Yu, S., Derr, J., Etherton, T.D., Kris-Etherton, P.M. (1995). Plasma cholesterol-predictive equations with special emphasis on trans and saturated fatty acids. An analysis from designed controlled studies. *Lipids*, 36(8), 783–791. https://doi.org/10.1007/s11745–001–0785–6
20. Rskediev, B.A., Uzakov, Ya. M., Butsik, V.A., Hegay, G.P. (2008). Complex dressing of mutton. *Meat Industry*, 9, 67–69. (In Russian)
21. Uzakov, Ya. M. (2012). Research of the biological and nutritional value assessment — i.e. the processes of protein breakdown in various finished lamb culinary products by proteolytic enzymes (in vitro).
AUTHOR INFORMATION

Alischer S. Pulatov — candidate of technical sciences, associate professor, Department of Life Safety and Ecology, Namangan Engineering-Construction Institute. Namangan city, 160103. I. Karimov Avenue, 12, Republic of Uzbekistan.
Tel.: +998–69–234–29–63, E-mail: alisherpulatov1950@gmail.com
ORCID: https://orcid.org/0000–0002–1462–2180
* corresponding author

Marina A. Nikitina — candidate of technical sciences, docent, leading scientific worker, the Head of the Direction of Information Technologies of the Center of Economic and Analytical Research and Information Technologies, V. M. Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences. 26, Talalikhina str., 109316, Moscow, Russia.
Tel: +7–495–676–95–11 extension 297, E-mail: m.nikitina@fncps.ru
ORCID: https://orcid.org/0000–0002–8313–4105

All authors bear responsibility for the work and presented data.
All authors made an equal contribution to the work.
The authors were equally involved in writing the manuscript and bear the equal responsibility for plagiarism.
The authors declare no conflict of interest.