MATHEMATICAL MODELING BY USING A C++ SOFTWARE

Ioan Milosan*

*Faculty of Materials Science and Engineering, Transilvania University of Brasov, Romania

Abstract: The aim of the study is to achieve a easy calculation of the linear regression (mathematical modeling) using of a C++ software. This application focused on the fitting and checking of linear regression models, using small and large data sets, with computers. Were constructed logical steps of the procedure and software based on their achievements C++ was made. The performance of regression analysis methods in practice depends on the form of the data generating process, and how it relates to the regression approach being used. It was used some statistical criteria as: Cochran, Student and Fischer criteria. After solving statistical analysis of the linear regression models, in the end there was obtained an applied statistical analysis of the linear regression model through the use of C++ software.

Keywords: analytical models; computer aided software engineering, heat treatment, input variables, mechanical properties.

MSC2010: 15A06, 49K05, 68N17, 97P30, 97Q60, 97R30.

1. INTRODUCTION

Mathematical modeling aims to describe the different aspects of the real world, their interaction, and their dynamics through mathematics [1].

Today, mathematical modeling has an important key role in industrial fields and especially in complex metallurgical processes. Linear regression is a statistical technique that is used to learn more about the relationship between a dependent variable and one or more independent variables [2, 3].

Linear regression equations with one or more variables were obtained under the assumption that the formula \( \hat{Y} = f(X_1, X_2, \ldots, X_n) \) was known in relationships physical analysis of the problem. There are many situations in industrial processes in which this assumption not confirmed; therefore equations obtained by regression analysis are subject to a statistical analysis of the regression equation to determine whether or not consistent with experimental data [4, 5, 6, 7].

The aim of the study is to achieve a easy calculation of the linear regression (mathematical modeling) using of a C++ software for all industrial processes and in particularly for the statistical analysis of the regression equation applied to the mechanical results of a special S.G. cast iron.

2. METHOD OF CALCULATING

Solving statistical analysis of the linear regression models is done in the following steps:
(1) Enter experimental data

(2) Linear regression model building, \( y = b_0 + b_1(x_i - \bar{x}) \);

(3) Calculating the \( b_0 \) and \( b_1 \) coefficients from the regression of the mathematical equation, with the relationships:
\[
b_0 = \frac{1}{K} \sum_{i=1}^{K} \bar{y}_i; \\
h = \sum_{i=1}^{K} x_i \bar{y}_i - \bar{x} \cdot \sum_{i=1}^{K} \bar{y}_i
\]

(4) With coefficients \( b_0 \) and \( b_1 \) can calculate the mathematical model linear regression model building,
\[
\tilde{y} = b_0 + b_1(x_i - \bar{x})
\]

(5) Determining the experimental dispersion, \( s_i^2 \), with the relationships:
\[
s_i^2 = \frac{\sum_{i=1}^{n} (y_{iy} - \bar{y}_i)^2}{v_i} \\
v_i = n_i - 1
\]

(6) The value of the Cochran criterion for the parallel determinations \( (y_1, y_2, \text{ and } y_3) \), is calculated:
\[
G_c = \frac{S_{ \text{max}}^2}{\sum_{i=1}^{K} S_i^2}
\]

(7) Verifying the homogeneity of the dispersions with the help of the Cochran criteria:
\[
G_f = G_{\alpha;v;k}
\]

A

B
(8.a) If \(G_C < G_T\), the dispersions are homogeneous.

(8.b) If \(G_C > G_T\), the dispersions are not homogeneous, *calculation stops*.

(9) Will determine the reproducibility dispersion \(S_0^2\), with the relationship:

\[
S_0^2 = \frac{1}{k} \sum_{i=1}^{K} S_i^2
\]

(10) The value of the Student criteria is calculated:

\[
t_{b_0} = \frac{|b_0|}{\text{Sb}_0}; \quad t_{b_1} = \frac{|b_1|}{\text{Sb}_1}; \quad \text{Sb}_0 = \sqrt{Sb_0^2}; \quad \text{Sb}_1 = \sqrt{Sb_1^2}; \quad \text{Sb}_0^2 = \frac{S^2}{N};
\]

\[
S^2 = \frac{1}{N - l} \sum_{i=1}^{N} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2; \quad N = k \cdot n_i; \quad \text{Sb}_1^2 = \frac{S^2}{\sum_{i=1}^{K} n_i(x_i - \bar{x})^2}
\]

(11) Verifying the coefficient signification \(b_0\) and \(b_1\) with the help of the critical value of the Student criteria, \(t_T\):

\[
t_T = t_{\alpha, \nu}
\]
Fig. 1. Steps for calculating the statistical analysis of the linear regression models.

(12.a) If \( t_{b0} > t_T \) and \( t_{b1} > t_T \), then both \( b_0 \) and \( b_1 \) coefficients are significant and may be part of the regression equation.

(12.b) If \( t_{b0} < t_T \) and \( t_{b1} < t_T \), then both \( b_0 \) and \( b_1 \) coefficients are not significant and can be neglected, calculation stops.

(13) The value of the Fischer criteria is calculated:

\[
F_c = \frac{S_{conc}^2}{S_0^2}; \quad S_{conc}^2 = \frac{\sum_{i=1}^{K} h_i (\bar{y}_i - \bar{y})^2}{k - l}
\]

(14) Verifying the connection between the regression equation and the experimental data with the help of the critical value of the Fischer criteria, with the relationships:

\[
F_T = F_{a;1;\nu;2}
\]

(15.a) If \( F_c < F_T \), then the regression equation consistent with experimental data (the regression equation adequately describes the experimental data and mathematical model is linear).

(15.b) If \( F_c > F_T \) then the regression equation is not consistent with experimental data (mathematical model is not linear).

where:
- \( \tilde{y} \) is the mathematical model;
- \( x_i \) is the process variable;
- \( \bar{x} \) is the average of process variables;
- \( b_0 \) is the slope coefficient from the regression of the mathematical equation, indicating the magnitude and direction of that relation;
- \( b_1 \) is the intercept coefficient from the regression of the mathematical equation, indicating the status of the dependent variable when the independent variable is absent;
- \( b_{0} \) is the slope coefficient from the regression of the mathematical equation, indicating the magnitude and direction of that relation;
- \( b_{1} \) is the intercept coefficient from the regression of the mathematical equation, indicating the status of the dependent variable when the independent variable is absent;
- \( K \) is the number of experimental points;
- \( \bar{y}_i \) is the average of process performances;
- \( s_i^2 \) is the experimental dispersion.
3. EXPERIMENTAL RESEARCHES

The results were obtained after performing some KCU properties tests on samples pieces in the case of an austempered ductile iron. The studied cast iron has the following chemical composition (% in weight): 3.75% C; 2.14% Si; 0.4% Mn; 0.012%P; 0.003%S; 0.05%Mg; 0.40% Ni, 0.42%Cu. This cast iron was made in an induction furnace.

The parameters of the heat treatment done were the following: the austenizing temperature, \( t_A = 900[^\circ C] \), the maintained time at austenizing temperature was, \( \tau_A = 30 \) [min]; the temperature at isothermal level, \( t_{iz} = 300[^\circ C] \); the maintained time at the isothermal level, \( \tau_{iz} = 10, 20, 30, 40, 50 \) and \( 60 \) [min]. The experimental samples were performed at isothermal maintenance in salt-bath, being the cooling after the isothermal maintenance was done in air.
After the heat treating there were used 18 samples for each determining the impact strength (KCU).

The values of experimental points, the process variable, the average of process variable and the process performances (the values of impact strength, KCU) are presented in table 1. The process variable (xi) was the maintaining time at isothermal heat treatment and for each six samples it was made three parallel determinations (y1, y2 and y3).

Table 1. The values of experimental points, the process variable, the average of process variable and the process performances (the values of impact strength)

<table>
<thead>
<tr>
<th>Experimental points</th>
<th>Process variable</th>
<th>Average of process variable</th>
<th>Process performances, KCU [ J / cm²]</th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td>i</td>
<td>y1</td>
<td>y2</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>32</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>41</td>
<td>43</td>
</tr>
</tbody>
</table>

For easy calculation of the statistical analysis regression equation will work spreadsheet, considering the above six steps work, which is presented in table 2.

Table 2. Tabular presentation of data taken into account

<table>
<thead>
<tr>
<th>k</th>
<th>i</th>
<th>x_i</th>
<th>( \bar{y}_i )</th>
<th>( (x_i - \bar{y}_i)^2 )</th>
<th>( (y_{ij} - \bar{y}_i)^2 )</th>
<th>( S_i^2 )</th>
<th>( (y_{ij} - \bar{y}_i)^2 + (\bar{y}_i - \bar{y})^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23</td>
<td>230</td>
<td>625</td>
<td>2.0000</td>
<td>1.0000</td>
<td>3.699</td>
<td>23.1746</td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>540</td>
<td>225</td>
<td>2.0000</td>
<td>1.0000</td>
<td>4.375</td>
<td>27.0603</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>930</td>
<td>25</td>
<td>2.0000</td>
<td>1.0000</td>
<td>5.051</td>
<td>30.9460</td>
</tr>
<tr>
<td>4</td>
<td>35.3333</td>
<td>1413.3333</td>
<td>25</td>
<td>0.6667</td>
<td>0.3333</td>
<td>5.727</td>
<td>34.8317</td>
</tr>
<tr>
<td>5</td>
<td>38.6667</td>
<td>1933.3333</td>
<td>225</td>
<td>0.6667</td>
<td>0.3333</td>
<td>6.403</td>
<td>38.7175</td>
</tr>
<tr>
<td>6</td>
<td>42.3333</td>
<td>2540</td>
<td>625</td>
<td>2.6667</td>
<td>1.3333</td>
<td>7.079</td>
<td>42.6032</td>
</tr>
<tr>
<td>Σ</td>
<td>197.3333</td>
<td>7586.6666</td>
<td>1750</td>
<td>9.3334</td>
<td>4.6666</td>
<td>32.334</td>
<td>197.3333</td>
</tr>
</tbody>
</table>

Solving statistical analysis of the linear regression models in this case is done using the 15 steps above.

4. CALCULATING THE STATISTICAL ANALYSIS OF THE REGRESSION EQUATION WITH C++ SOFTWARE

Solving statistical analysis of the linear regression models with C++ software is done in the following steps:

1. Input data into the program, actually presented in figure 2.
2. Output data into the program, actually presented in figure 3.

(3) Running Windows Commander data to determine all values in accordance with the working steps of the method, presented in figure 4.
Solving statistical analysis of the linear regression models in this case is done in the following six steps, considering the data presented in figure 2 and calculated with the C++ software:

1. Linear regression model building,
   \[ \hat{y} = b_0 + b_1(x_i - \bar{x}) \]

2. After calculating the \( b_0 \) and \( b_1 \) coefficients from the regression of the mathematical equation, there has been obtained the following results: \( b_0 = 32.8889 \); \( b_1 = 0.388571 \), and the mathematical model is:
   \[ \hat{y} = 19.288904 - 0.388571 \cdot x_i \]

3. After calculating the experimental dispersion, \( S_i^2 \) there has been obtained the following results: \( S_i^2 = 5 \).

4. After calculating the homogeneity of the dispersions with the help of the Cochran criteria, there has been obtained the following results: \( G_c = 0.266667 \) and \( G_c < G_T \) respectively.\( 0.426667 < 0.6161 \). Because \( G_c < G_T \) the dispersions are homogeneous and the dispersion value was determined: \( S_0^2 = 0.833333 \).

5. After verifying the coefficient signification \( b_0 \) and \( b_1 \) with the help of the Student criteria, there has been obtained the following results: \( t_{b_0} > t_T \) result 167.587 > 2.12 and \( t_{b_1} > t_T \), result 33.8146 > 2.12, and then both \( b_0 \) and \( b_1 \) coefficients are significant and may be part of the regression equation.

6. After verifying the connection between the regression equation and the experimental data with the help of the Fischer criteria, there has been obtained the following results: \( F_C < F_T \) result 0.327619 < 4.53, therefore the regression equation consistent with experimental data (the regression equation adequately describes the experimental data and mathematical model is linear).

5. CONCLUSIONS

(a) The regression analysis is one of the most widely used statistical tools to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships.

(b) The performance of regression analysis methods in practice depends on the form of the data generating process, and how it relates to the regression approach being used.

(c) After verifying the connection between the regression equation and the experimental data with the help of the Fischer criteria, it is noted that the regression equation consistent with experimental data (the regression equation adequately describes the experimental data and mathematical model is linear).

(d) By using C++ software we obtained more accurate results and the application time was reduced by several hours (for the classical calculation) to 2-3 minutes.
REFERENCES


