ANNALS of the ORADEA UNIVERSITY.
Fascicle of Management and Technological Engineering, Volume VII (XVII), 2008

MATHCAD PROGRAM FOR ESTIMATING THE WEIBULL DISTRIBUTION PARAMETERS

Cristin-Olimpiu MORARIU[1], Victoria TOMA[2], Vladimir MĂRĂSCU-KLEIN[2]
[1]University Transilvania Braşov, Department of Manufacturing Engineering, [2]Department of Economic Engineering and Production Systems
e-mail: c.morariu@unitbv.ro, vtoma@unitbv.ro, klein@unitbv.ro

Keywords: matching moments method, least squares method, maximum likelihood method

Abstract: In this article a MathCAD program is presented, which achieves the punctual estimation also with the confidence interval of the three-parametric Weibull distribution parameters.

The parameter estimating methods are:

- the modified moments method;
- the correlation coefficient method for estimating the location parameters of the Weibull distribution;
- the least squares method of parameter estimation;
- the maximum likelihood method.

Finally, the paper presents a case study, which uses sampling values representing working times between failures.

1. INTRODUCTION

For the main sub-ensemblies of the production equipment and for their wearing component parts [1], it is indicated to determine the time evolution of the failure rate, for conceiving an adequate maintenance policy.

Because the Weibull distribution provides a good conformity to the experimental results, its application field is a very large one, covering random, ageing and wearing processes. Consequently, for the statistical modeling of the working times between accidental failures of the production equipment [4] the theoretic Weibull model is adopted. The estimation of specific parameters is achieved by the following methods: the modified moments method, the correlation coefficient method for estimating the location parameter of the Weibull distribution, the least squares method, the maximum likelihood method and the moments method.

2. NOTATIONS AND TERMINOLOGY

Within this article, the following notations and terminology has been used:

F(t)	- Weibull distribution probability;	*1-α*	- confidence level;
β	- shape parameter;	*Ccor*	- correlation coefficient;
γ	- position parameter;	Fs(β,η)	– Fischer information
η	- scale parameter;	Γ(x)	– Gamma function;
N	- sample volume;	f(x)	– probability density function;
α	- significance level;	*Δ(*β,η)	– logarithmic likelihood function.

3. PROGRAM DESCRIPTION

The MathCAD program successively estimates the Weibull distribution parameters, by using: the modified moments method, the correlation coefficient method for estimating the location parameter, the least squares method, the maximum likelihood method and the moments method.

3.1. Modified moments method

The estimation of the values of the three-parametric Weibull distribution parameters ($\hat{\beta};\hat{\eta};\hat{\gamma}$) by the modified moments method, represents the solutions of the following equations system, obtained by equaling theoretical moments to sampling moments [2]:

$$\begin{cases} \dfrac{\Gamma\left(1+\dfrac{2}{\hat{\beta}}\right)-\Gamma^2\left(1+\dfrac{1}{\hat{\beta}}\right)}{\left[\left(1-n^{\frac{-1}{\hat{\beta}}}\right)\cdot\Gamma\left(1+\dfrac{1}{\hat{\beta}}\right)\right]^2}=\dfrac{\dfrac{1}{n-1}\cdot\sum\limits_{i=1}^{n}\left(t_i-\bar{t}\right)^2}{\left(t_1-\bar{t}\right)^2} \\ \hat{\gamma}=\dfrac{1}{n}\cdot\sum\limits_{i=1}^{n}t_i-\hat{\eta}\cdot\Gamma(1+1/\hat{\beta}) \\ \hat{\eta}=\sqrt{\dfrac{\dfrac{1}{n-1}\cdot\sum\limits_{i=1}^{n}\left(t_i-\bar{t}\right)^2}{\Gamma\left(1+\dfrac{2}{\hat{\beta}}\right)-\Gamma^2\left(1+\dfrac{1}{\hat{\beta}}\right)}} \end{cases} \tag{1}$$

3.2. The correlation coefficient method

In the case of the Weibull three-parametric distribution, featured by the existence of the position parameter γ, the expression of the correlation coefficient, depending on this parameter, is given by the following relation [2]:

$$\mathrm{Ccor}(\gamma):=\frac{\sum\limits_i\left(\ln\left(T_i-\gamma\right)\cdot\ln\left(\ln\left(\dfrac{1}{1-Fn_i}\right)\right)\right)-\dfrac{\left(\sum\limits_i\ln\left(T_i-\gamma\right)\right)\cdot\sum\limits_i\ln\left(\ln\left(\dfrac{1}{1-Fn_i}\right)\right)}{n}}{\left[\left[\sum\limits_i\ln\left(T_i-\gamma\right)^2-\dfrac{\left(\sum\limits_i\ln\left(T_i-\gamma\right)\right)^2}{n}\right]\cdot\left[\sum\limits_i\ln\left(\ln\left(\dfrac{1}{1-Fn_i}\right)\right)^2-\dfrac{\left(\sum\limits_i\ln\left(\ln\left(\dfrac{1}{1-Fn_i}\right)\right)\right)^2}{n}\right]\right]^{\frac{1}{2}}} \tag{2}$$

It may be seen that the correlation coefficient is a function of the position parameter γ, allowing using equation (2) for estimating parameter γ under the conditions of maximizing the value of $C_{cor}(\gamma)$.

The bond between the three-parametric and two-parametric shapes of the Weibull model is represented by the position parameter *(γ) which* determines the distribution's position on the x-axis [2]. By determining its value by the correlation coefficient method and making the following substitution:

$$T_{b_i} = T_i - \gamma \tag{3}$$

the initial sample becomes a sample modeled by the two-parametric Weibull distribution, which parameters *β, η* , shall be determined by applying the specific estimation methods for this model.

3.3. The least squares method

The least squares method consists in [2], [3] determination of the line parameters passing among the points containing statistical remarks of the working times, in such a manner, that the squares sum of the existing deviations between the line so traced and the multitude of point, is the least.

Smoothing by a double looking up the logarithm of the expression of the Weibull distribution probability, under knowing *γ*, the relation:

$$\ln\cdot\ln\frac{1}{1-F(T)} = \beta\cdot\ln(T_b) - \beta\cdot\ln\eta \tag{4}$$

is obtained

If this relation is noted :

$$y = \ln\cdot\ln\frac{1}{1-F(T)}\ ;\ x = \ln T_b\ ;\ A = -\beta\cdot\ln\eta\ ;\ B = \beta \tag{5}$$

equation :

$$y = A + B\cdot x \tag{6}$$

is obtained, representing a linear relation between y and x being a " line " in the *xOy* coordinate system, ehere :

$$x_i = \ln T_{b\,i} \quad (\text{ i = 1, 2........n }) \tag{7}$$

Let be: $T_1, T_2 \ldots T_n$, times of good working coming from the observation of the working during a time *T* of a production equipment.

We note by δ_i *(i = 1, 2,.....n)* the upright distance from points $M_i(T_{bi}, y_i)$ to the median line $y = B\cdot x + A$, and it resspults:

$$y_i = A + B\cdot\ln(T_{b\,i}) \pm \delta_i \tag{8}$$

The determination of parameters *A* and *B* by the least squares method implies the minimization of the equation value:

$$\delta = \sum_{i=1}^{n}\delta_i^2 = \sum_{i=1}^{n}\left(A + Bx_i - y_i\right)^2 = \min \tag{9}$$

Thus, we obtain:

$$\begin{cases} \hat{A} = \dfrac{\sum\limits_{i=1}^{n} y_i \cdot \sum\limits_{i=1}^{n} x_i^2 - \sum\limits_{i=1}^{n} x_i \cdot \sum\limits_{i=1}^{n} x_i \cdot y_i}{k \cdot \sum\limits_{i=1}^{n} x_i^2 - \left(\sum\limits_{i=1}^{n} x_i\right)^2} \\ \hat{B} = \dfrac{k \cdot \sum\limits_{i=1}^{n} x_i \cdot y_i - \sum\limits_{i=1}^{n} x_i \cdot \sum\limits_{i=1}^{n} y_i}{k \cdot \sum\limits_{i=1}^{n} x_i^2 - \left(\sum\limits_{i=1}^{n} x_i\right)^2} \end{cases} \tag{10}$$

β and η, parameters estimation is obtained from systems (10) and (5), being given the relations:

$$\hat{\beta} = \hat{B}\,;\ \hat{\eta} = e^{-\frac{A}{B}} \tag{11}$$

3.4. Maximum likelihood method

For estimating parameters *β, η*, if *γ* was determined by the correlation coefficient method, the maximization of the likelihood function conduces to system:

$$\begin{cases} \dfrac{1}{\hat{\beta}} + \dfrac{1}{n}\sum\limits_{i=1}^{n} \ln T_{bi} - \dfrac{\sum\limits_{i=1}^{n} t_i^{\hat{\beta}} \ln t_{bi}}{\sum\limits_{i=1}^{n} t_{bi}^{\hat{\beta}}} = 0 \\ \hat{\eta} = \left[\dfrac{1}{n}\sum\limits_{i=1}^{n} T_{bi}^{\hat{\beta}}\right]^{\frac{1}{\beta}} \end{cases} \tag{12}$$

Solving the first equation of the system (12) needs to use iterative calculating procedures, starting from an approximate initial solution. Within the MathCAD program values of the parameter of *β* shape were taken into calculation, obtained by the modified moments method, respectively the least squares method and the initial solution *β = 0,6* has been adopted, after which it has to come back to the equation system (12) obtaining a high estimating precision, of the 10^{-6} order.

Calculation of the confidence intervals is achieved by means of the Delta method, which presumes a calculation of the Fisher information.

$$Fs(\beta,\eta) := \begin{bmatrix} -\dfrac{d^2}{d\beta^2}\Delta(\beta,\eta) & -\dfrac{\partial}{\partial\beta}\left(\dfrac{\partial}{\partial\eta}\Delta(\beta,\eta)\right) \\ -\dfrac{\partial}{\partial\eta}\left(\dfrac{\partial}{\partial\beta}\Delta(\beta,\eta)\right) & -\dfrac{d^2}{d\eta^2}\Delta(\beta,\eta) \end{bmatrix} \tag{13}$$

In formula (13) the dispersion of the shape and scale parameter is represented by the first element (0,0) and respective (1,1) of the matrix.

The confidence intervals for *β, η* are determined on grounds of the dispersion, by using asymptotic features of the maximum likelihood estimations. If the number *N* of estimation is very high, tending to infinity, than estimators have a normal distribution.

3.5. The moments method

The application of the moments method for determining the estimations, presumes to equal the first *r* moments of the theoretic distribution to the first *r* empiric moment, according to the number of parameters defining the statistical distribution. An equations system supplying as solutions, the estimations of the parameters looked for, is obtained.

In the case of the two-parametric Weibull distribution, the variation coefficient (*Cv*) depends only on parameter *β*. Equalizing the expression of theoretic variation coefficient to its sampling value, equation:

$$\left(\frac{\Gamma\left(\frac{2}{\hat{\beta}+1}\right)}{\Gamma^2\left(\frac{1}{\hat{\beta}+1}\right)}-1\right)^{\frac{1}{2}}=\frac{\frac{1}{n}\cdot\sum_{1}^{n}t_i}{\sqrt{\frac{1}{n-1}\cdot\sum_{i=1}^{n}\left(t_i-\bar{t}\right)^2}} \tag{14}$$

is obtained; by solving it, the estimated values of the shape parameters *β* are obtained.

The estimated value of the scale parameter is obtained by equalizing of the average sampling $\bar{t}$ to the theoretical average value, relation (15):

$$\hat{\eta}=\frac{\bar{t}}{\Gamma\left(\frac{1}{\hat{\beta}}+1\right)} \tag{15}$$

The logical scheme of the MathCAD program for estimating the Weibull distribution parameters is represented in figure 1.

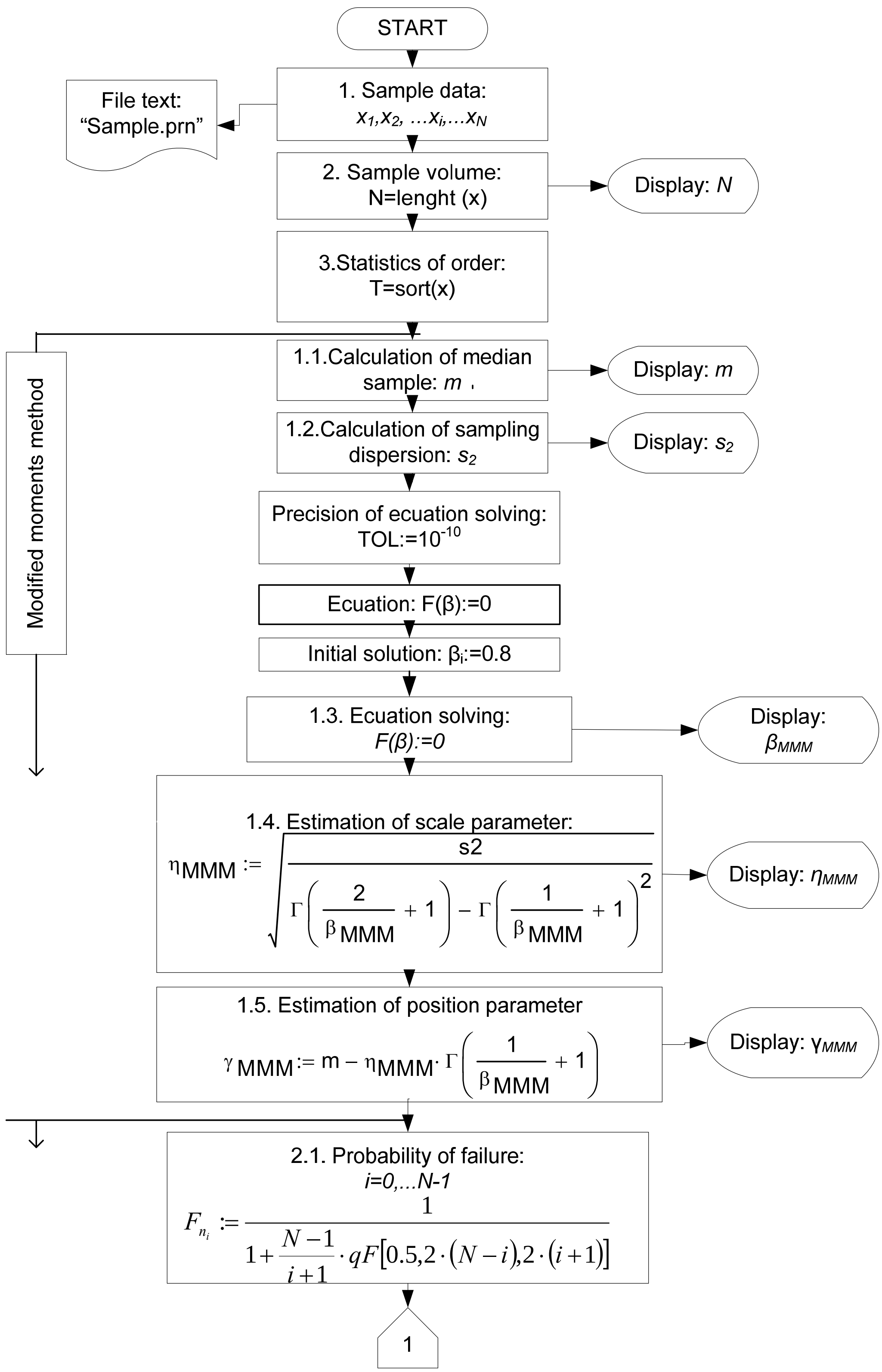
4. CASE STUDY

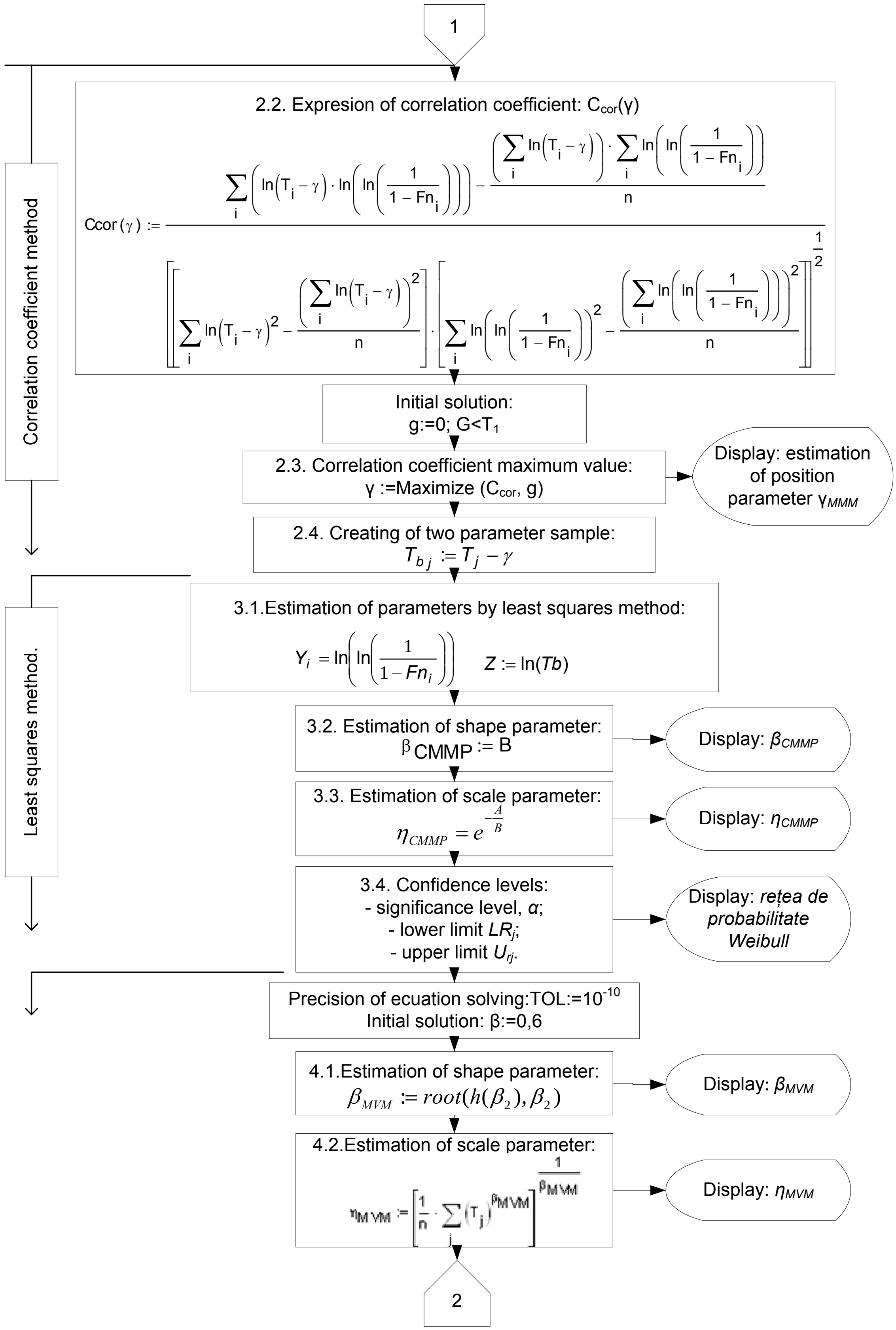
The data used within this case study were collected during a time of 1 year, respectively January - December 2004. During this period, the working of a *Slicer SM-H* -type production equipment specialized in processing wood into thin plates having thicknesses between 0,1-0,2 mm has been followed. The equipment has worked in continuous running, 24 hours of 24, seven days from seven.

The equipment is provided by automate recorders of the staying times in repair. A software written in Visual Basic for Application allowed the fast determination of the times of good working. As a result of the application of the randomness checking tests, identification and elimination of outlier data, the following experimental data, in minutes, of the times of working without failure, have been obtained:

1400	9964	11440	3897	2782	7200	702	1426	500	9039	7167	22568
10027	2693	4310	2518	10080	4280	1355	6396	1144	1419	2854	2606
2673	10039	12955	8222	5382	2789	3895	2437	11339	5390	7189	5287
4187	2688	4308	2026	8084	4076	5728	4091	5326	2591	2587	4278

4222	4206	1427	4247	2821	7101	11155	8502	2769	2405	5474	2840
4221	2537	2839	2786	4276	4212	1455	4246	5747	2781	19684	7180
2750	5746	6869	4268	7128	14353	5502					




1
Correlation coefficient method
2.2. Expresion of correlation coefficient: $C_{cor}(\gamma)$
$$Ccor(\gamma) := \frac{\sum_i \left(\ln(T_i-\gamma)\cdot \ln\left(\ln\left(\frac{1}{1-Fn_i}\right)\right)\right) - \frac{\left(\sum_i \ln(T_i-\gamma)\right)\cdot \sum_i \ln\left(\ln\left(\frac{1}{1-Fn_i}\right)\right)}{n}}{\left[\left[\sum_i \ln(T_i-\gamma)^2 - \frac{\left(\sum_i \ln(T_i-\gamma)\right)^2}{n}\right]\cdot\left[\sum_i \ln\left(\ln\left(\frac{1}{1-Fn_i}\right)\right)^2 - \frac{\left(\sum_i \ln\left(\ln\left(\frac{1}{1-Fn_i}\right)\right)\right)^2}{n}\right]\right]^{\frac{1}{2}}}$$
Initial solution:
g:=0; G<T_1
2.3. Correlation coefficient maximum value:
γ :=Maximize (C_{cor}, g)
Display: estimation of position parameter γ_{MMM}
2.4. Creating of two parameter sample:
$T_{b\,j} := T_j - \gamma$
Least squares method.
3.1.Estimation of parameters by least squares method:
$Y_i = \ln\left(\ln\left(\frac{1}{1-Fn_i}\right)\right)$ $Z := \ln(Tb)$
3.2. Estimation of shape parameter:
$\beta_{CMMP} := B$
Display: β_{CMMP}
3.3. Estimation of scale parameter:
$\eta_{CMMP} = e^{-\frac{A}{B}}$
Display: η_{CMMP}
3.4. Confidence levels:
- significance level, α;
- lower limit LR_j;
- upper limit U_{rj}.
Display: *rețea de probabilitate Weibull*
Precision of ecuation solving:TOL:=10^{-10}
Initial solution: β:=0,6
4.1.Estimation of shape parameter:
$\beta_{MVM} := root(h(\beta_2), \beta_2)$
Display: β_{MVM}
4.2.Estimation of scale parameter:
$\eta_{MVM} := \left[\frac{1}{n}\cdot\sum_j (T_j)^{\beta_{MVM}}\right]^{\frac{1}{\beta_{MVM}}}$
Display: η_{MVM}
2

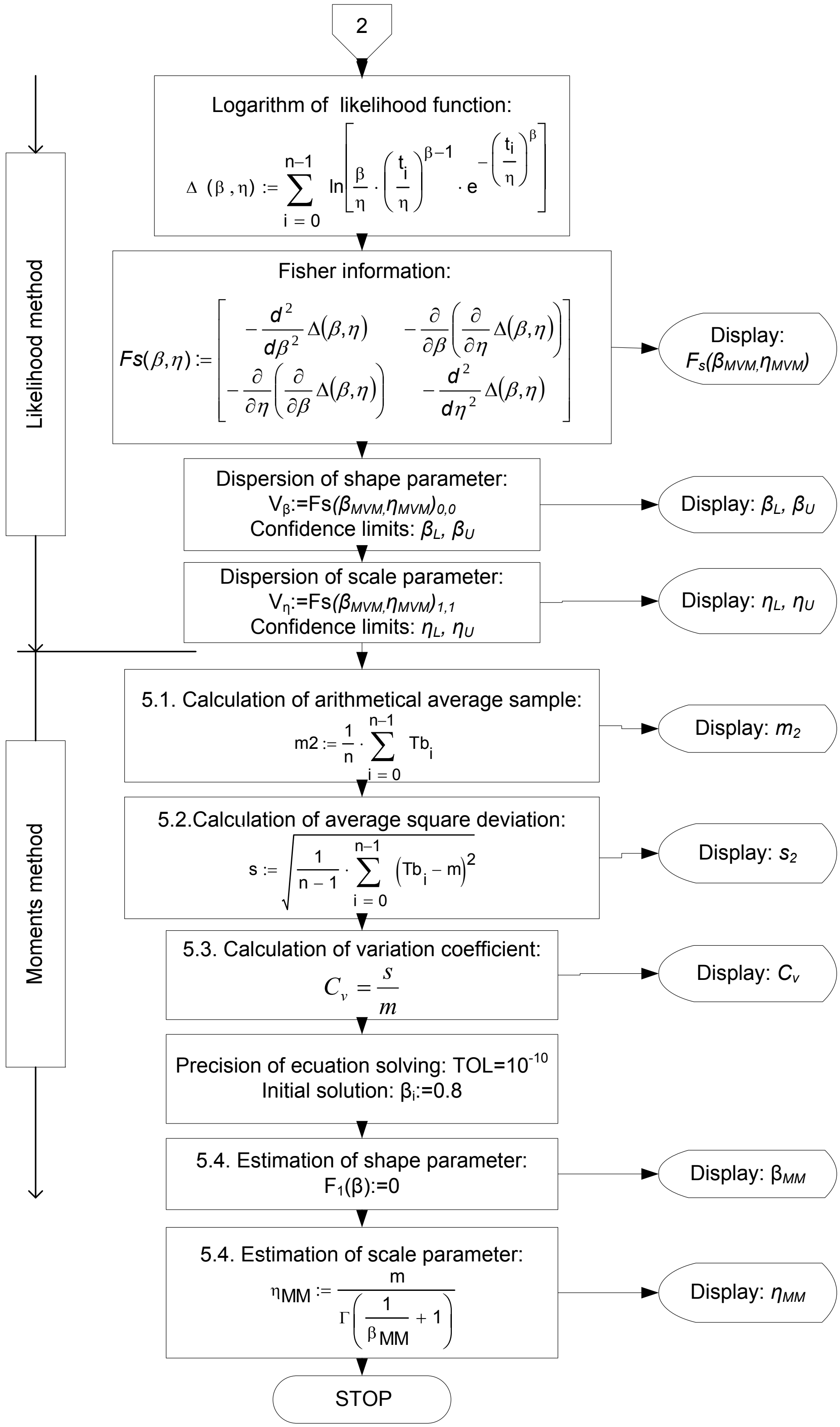


Fig.2.The logical scheme of the MathCAD program

For a correct reading of the observed data in the MathCAD program, their organization in a file ASCII called *Date250408omogenizate*.txt has been necessary.

The calculation program supplies estimations of the three-parametric Weibull distribution parameters, by using the presented methods. By the modified moments method the estimated values of the parameters are: $\beta_{MMM,}=1.198015$; $\eta_{MMM,}= 4994.90383$; $\gamma_{MMM}= 579.49128$.

The values of the correlation coefficient *Ccor* $(\gamma) \geq 0,90$ [2],[6] obtained by the correlation coefficient method (fig.2) show a good correlation between the adopted theoretic model and the observed experimental values, allowing the estimation of the position parameter γ which has a value of 306.32015.

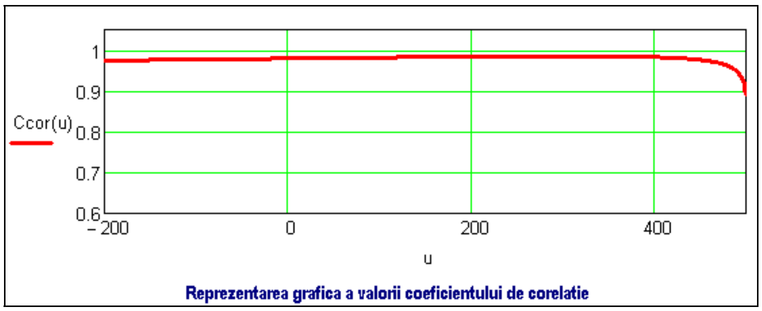


Fig. 2. Variation of the correlation coefficient according to the position parameter

The least squares method supplies the following values of the parameters: $\beta_{CMMP,}=1.48392$; $\eta_{CMMP,}= 5450.07$; $\gamma_{CMMP}=$ 306.32015. The program allows seeing experimental data and confidence intervals on a Weibull-type probability network (fig.3.):

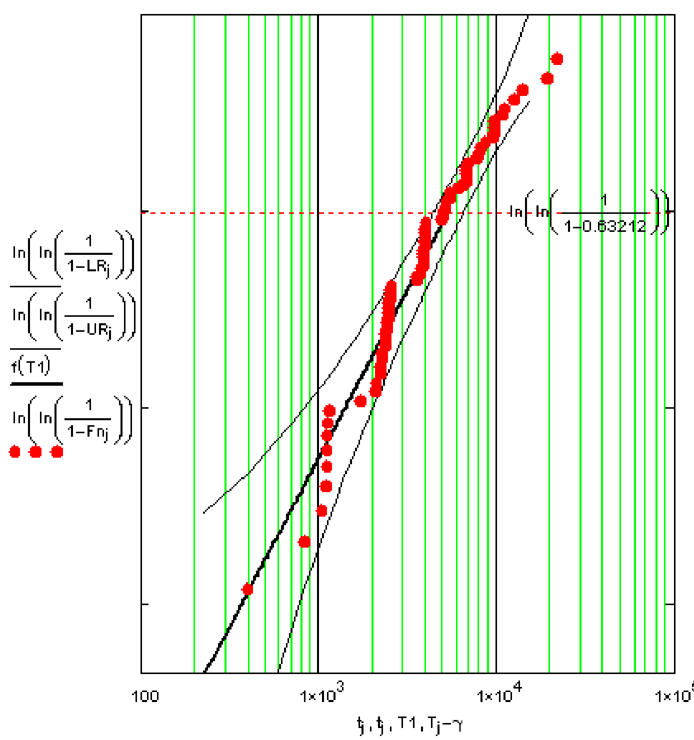


Fig.3. Weibull probability network

By applying the maximum likelihood method the following values have been obtained: $\beta_{MVM,}$=*1.38032;* $\eta_{MVM,}$= *5474.36;* γ_{MVM}= 306.32015*.* For a significance level of 5%, the confidence interval for the shape parameter is $\beta \in [1.17318;1.62403]$ and for the scale parameter $\eta \in [4.62335 \cdot 10^3;6.48202 \cdot 10^3]$.

The application of the moments method leds to the following results: β_{MM} =*1.350416;* $\eta_{CMMP,}$= *5758.27635;* γ_{CMMP}= 306.32015.

5. CONCLUSIONS

Between the estimated values of the Weibull distribution parameters obtained with different estimation methods, there appear insignificant differences, due to the features of the estimators used for.

The calculated values of the correlation coefficient *(Ccor(*γ*) = 0.98384*) confirms the hypothesis concerning the good concordance between the Weibul distribution model and the times between failures of the production equipment.

Due to their features, maximum likelihood estimators are unanimously considered as being the bests.

REFERENCES

1. Jalobeanu, C., Raşa, I.- *Mathcad – Probleme de calcul numeric şi statistic*, Imprimeria Ardealul Cluj-Napoca, 1995
2. Kececioglu, D. B. - *Reliability Engineering HandBook.* Vol. I şi II, Inc. Englewood Cliffs, PTR Prentice Hall, New-Jersey, 1993
3. Martinescu, I., Popescu, I.: *Fiabilitate*, Editura Gryfon, Braşov, 1995
4. Mărăscu-Klein, V., Toma, V.: *Managementul mentenanței,* Editura Universității Transilvania din Braşov, 2007
5. Monchy, F.: *Maintenance – Methodes et organisation*, Edition Dunod, 2003
6. Morariu, C.O.,: *Optimizarea încercărilor de fiabilitate a rulmenților,* Teză de doctorat, Universitatea Transilvania din Braşov, 1999