Annals of the Academy of Romanian Scientists
Online Edition Series on Engineering Sciences
ISSN 2066 – 8570 Volume **7**, Number **1**/**2015**

INPUT MODELLING USING STATISTICAL DISTRIBUTIONS AND ARENA SOFTWARE

Elena Iuliana GINGU (BOTEANU)[1],
Miron ZAPCIU[2]

Rezumat. *Lucrarea prezintă o metodă de alegere corectă a distribuţiilor probabile pentru timpul de defecţiune într-un sistem flexibil de fabricaţie. Mai multe distribuţii bine-cunoscute oferă adesea o aproximare bună în practică. Distribuţiile continue frecvent utilizate sunt: Uniformă, Triunghiulară, Beta, Normala, Lognormala, Weibull şi Exponenţială. În acest articol este studiat modul de utilizare a Input Analyzer în limbajul de simulare Arena pentru a potrivi distribuţiile de probabilitate a datelor, sau pentru a evalua cât mai bine o anumită distribuţie. Obiectivul a fost de a furniza selecţia distribuţiilor statistice cele mai adecvate şi pentru a estima valorile parametrilor timpilor de defecţiune pentru fiecare maşină a unei linii de fabricaţie reală.*

Abstract. *The paper presents a method of choosing properly the probability distributions for failure time in a flexible manufacturing system. Several well-known distributions often provide good approximation in practice. The commonly used continuous distributions are: Uniform, Triangular, Beta, Normal, Lognormal, Weibull, and Exponential. In this article is studied how to use the Input Analyzer in the simulation language Arena to fit probability distributions to data, or to evaluate how well a particular distribution. The objective was to provide the selection of the most appropriate statistical distributions and to estimate parameter values of failure times for each machine of a real manufacturing line.*

Keywords: **S**tatistical distribution, Random numbers, Failures times, Input Analyzer

1. Introduction

Nowadays, the increase of competitiveness and the quick changes of the market conditions constrain many enterprises to optimize the production or to reorganize.In order to remain competitive, many enterprises use innovative methods, as discrete event simulation. This method solves problems in a virtual way and by the aid of the dedicated simulation software [1].

The discrete event simulation software allows us to easily test different scenarios with little effort, and embed it in the simulation of our own algorithms.

In the simulation of discrete event, typical input data are the statistical distribution, as a distribution of time to failure of a component.

[1]PhD Student, Faculty of Engineering and Management of Technological Systems, University Politehnica of Bucharest, (e-mail: iuliana_boteanu@yahoo.com).
[2]Professor, Faculty of Engineering and Management of Technological Systems, University Politehnica of Bucharest; Academy of Romanian Scientists, Splaiul Independenței 54, 050094, Bucharest, Romania (e-mail: miron.zapciu@upb.ro).

One of the purposes of a simulation study will be to optimize the availability of resources, such as machines, operators, buffers, etc. This means that you have to reduce the number of failures and down-times of machines. When a machine fails, it has to be repaired, which leads to delays in the production and costs money.

2. Modelling Random Processes and Statistical Distributions

Random numbers are a necessary basic piece in the simulation of almost all discrete systems. Many computer programs use random number generators to create a stream of random numbers. Usually, there are deterministic processes and random processes. For a deterministic process, we can predict the result of the process; if we do not know all factors that characterize a process, its result is random [2]. For example, when we simulate a production processes:

- The occurrence of pauses is a deterministic process.
- The Failures of machines are random processes.

The usefulness of statistical distributions to model activities is generally for unpredictable or uncertain events, as: inter-arrival time, demands for a product, failure time, etc.

Those variables are modelled as random variables with specified probability distribution then, statistical procedures for estimating the parameters of distribution are introduced and finally it is testing the validity of the model.

The analysis of the data necessary for stochastic simulation can be used with Arena software. The Input Analyzer from Arena provides three numerical measures of the quality of fit of a distribution to the data to help you decide.

1. Mean square error: The simplest to understand. Smallest is better.
2. Chi-square: Corresponding p-values fall between 0 and 1. Larger *p*-values indicate better fits.
3. Kolmogorov-Smirnov: Larger p-values indicate better fits.

Fig. 1. Statistical distribution.

The commonly used continuous distributions are: Normal, Exponential, Beta, Lognormal, Weibull, Uniform, Triangular (Fig. 1).

The graphical representation of the main statistical distributions are shown in Figure 2.

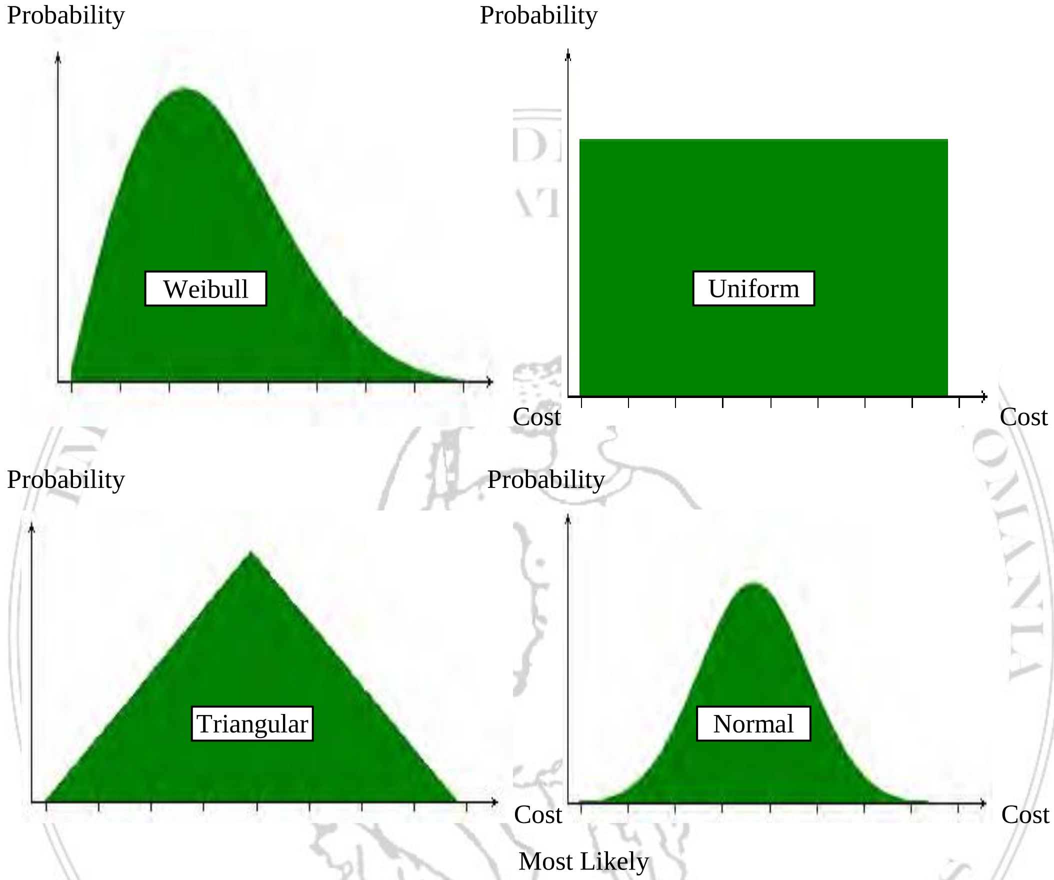


Fig. 2. The graphical representation of statistical distributions.

3. Modelling failures using statistical distributions – case study

The approach to modelling input data is to fit a probability distribution from a standard family of continuous distributions based on sample data.

A random number that can take all values within a limited or an unlimited interval of numbers is continuous. For example: he mean time between failures and the mean time to repair of a machine are continuous random numbers.

To describe a continuous random number, we cannot specify the probability for an individual number to occur. We have to specify an interval, the probability with which this random number is located between two given values.

The machines are parts of the Flow Production Systems that process the products. They are the active parts of the system as they initiate the movement of products in the system.

The time of a machine is operational and therefore, if it is not in a state of failure we call it the ‘uptime’.

The remaining time is ‘downtime’, this being in fact, the time when the machine is under repair [3].

A schedule downtime can be used to describe the effective working time of a machine and the time of repair or maintenance [4].

In our study we want to examine the input data who define the Machines: Mean Time between Failures and Mean Time to Repair.

If we note MTTR - Mean Time to Repair and MTBF - Mean Time between Failures, then the formulas are:

MTBF = (Total up time) / (number of breakdowns) (1)

MTTR = (Total down time) / (number of breakdowns) (2)

Where,

- "Mean Time" means, statistically, the average time;
- Up time means the moment at which a machine began operating (initially or after a repair);
- Down time means the moment at which a machine failed after operating since the previous uptime-moment.

Production planning specialists do observe the availability of resources, the MTTR and the MTBF. The availability is calculated by the formula:

$$AV = \frac{100 \cdot MTBF}{MTBF + MTTR} \quad (3)$$

An availability of 100% has an MTTR of 0, as the machine is available and does not have to be repaired.

The case study is related to a manufacturing line of a car headrest support work piece (Fig.3).

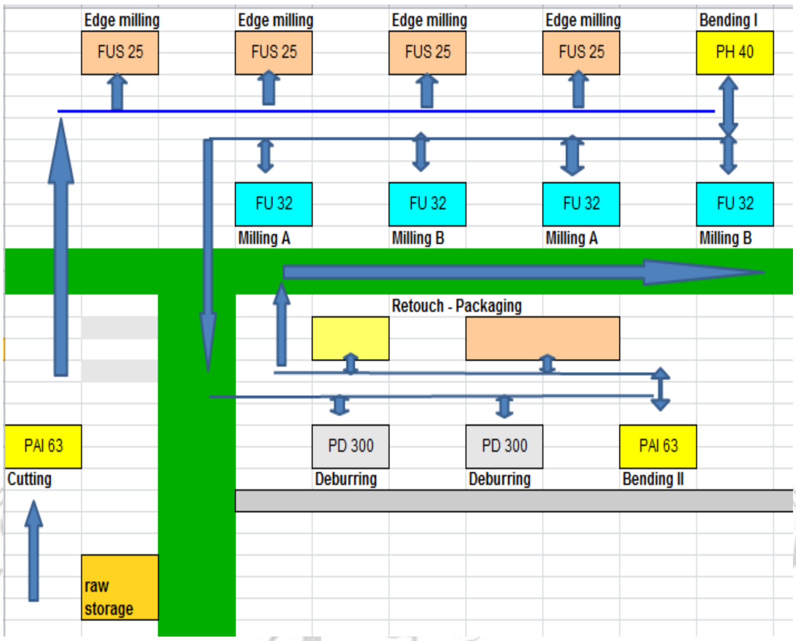


Fig. 3. Layout of a manufacturing line case study.

Oftentimes the Mean Time between Failures (MTBF) and the Mean Time To Repair (MTTR) are stochastically distributed. Because of the qualitative properties of a system with failures, we know the type of distribution for the MTBF and for the MTTR. What we do not know are the parameters of the distribution [5, 6].

To develop the data input we need to respect four steps:

1. Collection of data from the real system (Fig.4)

T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	Total down time (min)	MTTR	MTBF	Frequency	Operation availability (min/mo)
9	6	6	0	1	1	0	15	12	2	52	0,44	211,08	118	24960
9	6	6	0	1	2	0	15	24	2	65	0,55	210,97		
9	7	7	0	1	3	0	15	25	2	69	0,58	210,94		
9	7	8	1	1	3	10	16	25	2	82	0,69	210,83		
9	7	8	1	1	4	10	16	26	2	84	0,71	210,81		
9	7	8	2	1	4	10	16	26	2	85	0,72	210,81		

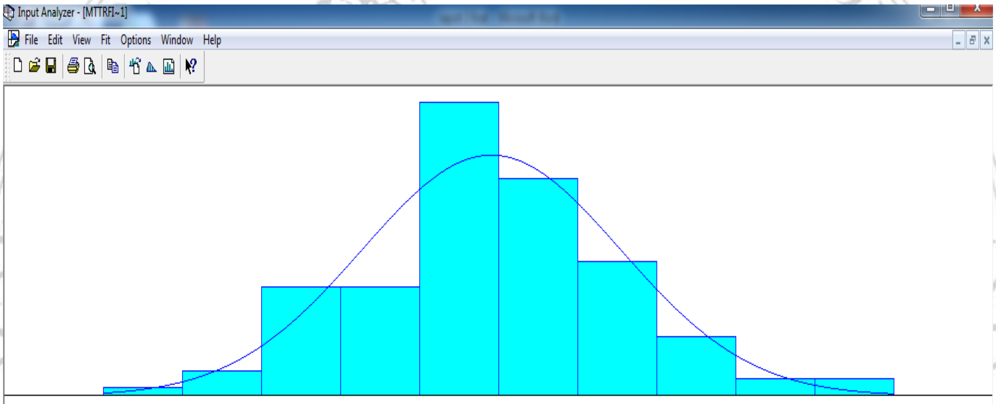
Fig. 4. Database for the delay times

2. Identification of distributions of probabilities to represent the input process: The Lognormal distribution, the Normal distribution, the Erlang distribution and the Negative Exponential distribution are especially suited for modelling failures.

3. Selection of parameters to determine a specific distribution: The collected data are entered into the program's Arena Input Analyser to fit the statistical distributions.

4. Evaluation of distribution and its testing: After a family of distributions has been selected the next step will be parameter estimation. The last step proposes testing methods Chi-Square distributions, Anderson-Darling and Kolmogorov-Smirnov [7].

By means of the data collected it was established Mean Time to Repair MTTR and Mean Time between Failures MTBF. After entering data into the Input Analyzer, the most appropriate distribution for MTTR is Normal (Mean = 1.37, Standard Deviation = 0.378) (Fig.5).




```
          Distribution Summary

Distribution:    Normal
Expression:      NORM(1.37, 0.378)
Square Error:    0.008669

Chi Square Test
  Number of intervals    = 6
  Degrees of freedom     = 3
  Test Statistic         = 5.6
  Corresponding p-value  = 0.146

Kolmogorov-Smirnov Test
  Test Statistic         = 0.061
  Corresponding p-value  > 0.15

          Data Summary

Number of Data Points    = 118
Min Data Value           = 0.44
Max Data Value           = 2.35
Sample Mean              = 1.37
Sample Std Dev           = 0.38

          Histogram Summary

Histogram Range          = 0.24 to 2.55
Number of Intervals      = 10
```

Fig. 5. Statistical distribution for MTTR

For MTBF, the most appropriate distribution is Normal distribution with the parameters: Mean = 210, Standard Deviation = 0.377 (Fig. 6).

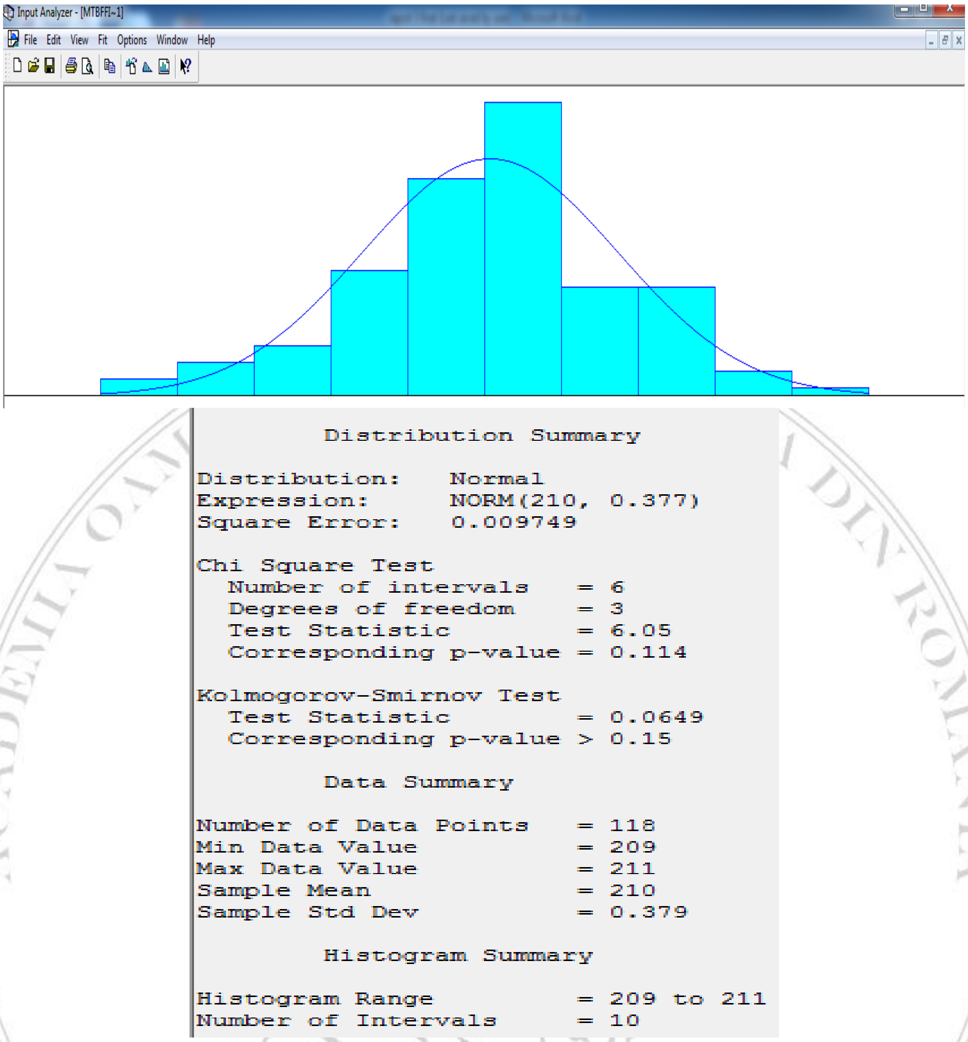


Fig. 6. Statistical distribution for MTBF.

With these results, we can establish the repair process and the failure distribution for each machine in a flow production system. The estimation of failures time could be very useful to industrial practitioners.

Conclusions

In this article a method of choice of the statistical distributions and their parameters for a manufacturing line has been presented. The model is realized by using Arena software. The research was focused on the correct selection of the stochastic distribution of failure time. To fit this distribution we need to study modelling random processes. Failures of machines or the rejects that are produces are random processes.

Input Analyzer from Arena software help us to establish the properly distribution for failure times. In our study the appropriate distribution is Normal with the parameters Mean and Standard Deviation. The input data of the Machines: Mean Time between Failures and Mean Time to Repair were examined. The results of the research showed that the failure time estimation model gave us satisfactory calculation of the failure time. The method to determine failures time estimation could be very useful to industrial practitioners.

Acknowledgment

The work has been funded by the Sectorial Operational Program Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/134398.

R E F E R E N C E S

[1] http://www.ipa.fraunhofer.de/Factory_Planning_and_Production_Optimization

[2] J. Banks , J. Carson , B. Nelson, D. Nicol, *Discrete-Event System Simulation*, Prentice-Hall International Series in Industrial and Systems Engineering, ISBN 0-13-088702-1.

[3] S. Sankar Sana, *Preventive maintenance and optimal buffer inventory for products sold with warranty in an imperfect production system*, International Journal of Production Research, Vol. 50, No. 23, 2012, pp. 6763–6774.

[4] I. Gingu (Boteanu), M. Zapciu, *Modeling and simulation of manufacturing system layout using Delmia Quest software*; Proceedings of the Academy of Romanian Scientists PRODUCTICA, Scientific Session; 9 May 2014, ISSN 2067-2160 Volume 6, Number 1/2014; 205-214.

[5] K. Hee Han, S. Moon Bae, S. Hyun Cho, G. Lee, *Parameter-Driven Rapid Virtual Prototyping of Flexible Manufacturing System*, International Journal of Mathematics and computers in simulation, 387- 396, Issue 4, Volume 6, 2012

[6] S. Taja, G Pfeilc, M. Roumayad, *A spread-sheet model for efficient production and scheduling of a manufacturing line/cell*, International Journal of Production Research Vol. 50, No. 4, 2012, 1141–1154.

[7] R. Popa (Scurtu), O. Lupescu, I. Popa, *Determining the safety coefficient in order to optimize the spare parts stocks*, International Journal of Modern Manufacturing Technologies ISSN 2067–3604, Vol. IV, No. 2 / 2012.