25th ABCM International Congress of Mechanical Engineering
October 20-25, 2019, Uberlândia, MG, Brazil

COMPARATIVE ANALYSES OF TWO PROGRAMING TOOLS BASED ON RELIABILITY

Isaíres Kayla Farias de Luna Barbosa [1]
Pedro Henrique Vasconcelos Pereira[2]
UFRPE/UACSA [1] UFRPE/UACSA[2]
isalunaluna123@gmail.com[1] pedrovasconcelos056@gmail.com[2]

Adriano da silva Marques
UFRPE/UACSA/ Depto. de Engenharia Mecânica
adriano.marques@ufrpe.br

Abstract. *Day to day in industry brings within it a routine of maintenance, in which it is necessary to know how the machines works, when and where they will fail. To analyze these phenomena the engineers use modeling functions that enable the understanding of the work environment. This article was written based on the comprehension of reliability of systems and programing language and aims to compare didactics and the development of a program in excel and with python language, by modeling a hazard function, a reliability function and a density function. In this work are presented results from a computational code developed in python language and the comparison with the results presented in (ASSIS, 2018). Excel is software with well-defined tools, however the Python language has in itself a versatility that enables the student to explore high-level libraries for free. The realization of the work in a program language facilitate the work in other languages that are widely diffused on academia, as the C.*

Keywords: *Reliability, Python, distributions, programing, probability.*

1. INTRODUCTION

According to Fogliatto and Ribeiro (2011) the reliability can be defined in terms of probability of surviving of a given product over time Silva (et al, 2017) affirms that the reliability can be represents by some important functions such as the reliability function, R(t), and hazard function, h(t). This is the most commonly used function in the reliability study, and its unit of measure is failure per unit of time. In order to study the behavior of the risk function there are models called probability distributions, "probability distributions are mathematical models that derive from a certain number of failures in a certain range of performance" (Szidarovszky, 2018). The distributions are useful in many areas, such as equipment maintenance and knowledge of the diametrical distribution of forest stand (Binoto et al 2017). Among the most well-known distributions are: exponential, weibull, gamma and lognormal, which are widely used due to their relative simplicity and fit with minimal possible error.

To study the behavior of these distributions in the classroom or academic work, one usually chooses software already known as Excel. This is a widely used and used spreadsheet editing tool, the biggest reasons so many people use this tool "its intuitive cell-based structure and simple interface that is easy even for new users" (Tarek and Tolga, 2001). Another great benefit of programming with this platform is the ability to work with complex equations and get fast results even with large amounts of data. However, some functionality may be compromised due to errors commonly found in spreadsheets, such as typos when there is a lot of data to add, and the fact that information is not always displayed in cells makes it even harder to find errors. . Another downside to using Excel is that it is not intuitive, spreadsheets created for companies and companies are only useful when their creator is present, because all knowledge of how it works is with the author.

According to Menezes (2014) programming is a task that requires time and dedication, where it is not enough to just study and make examples, our mind has to work freely. Likewise Menezes (2014) affirms that if there is no satisfactory program, one must create something that meets our expectations. And the choice of Python as the basis for the data analysis was through this justification. Python is a simple, clear and powerful language, and can be used in system administration to develop large projects (Menezes, 2014). In addition to being free software, it can be used in any operating system, be it Linux, Microsoft Windows or Mac OS X. According to Borges (2010): "Python has a clear and concise syntax that favors the readability of the source code, making the language more productive. "Other programming languages involve numerous markups that make the program difficult to read

This paper uses exponential distribution modeling from an interactive classroom activity that simulated machines on production lines and their time to failure. In the data adjustment was used for comparative effects the results presented in an Excel table in Assis (2018) work and a computational code, developed by the authors, developed in Python language. In both it is possible to observe results linked to the distribution model and its graphs, which allowed the comparison of both data and interactivity between the programmer / program.

It is noteworthy that the purpose of this article is not to underestimate a highly powerful tool like Excel, but to open the doors of methodologies used in the classroom, precisely in the Maintenance-Centered Reliability discipline, to take advantage of other required disciplines during student's training.

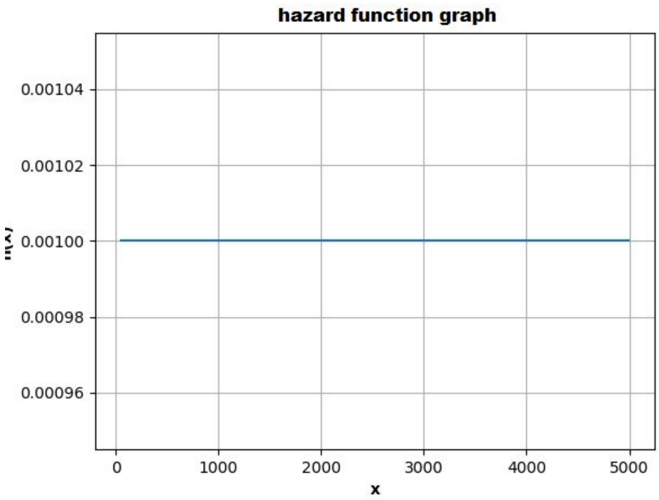
2. THEORETICAL FOUNDATION

This paper discusses reliability analysis in two software programs and we will divide this section into three parts: introduction to reliability distributions, the use and the functionality of Python and Excel..

2.1 Introduction to reliability distributions.

The discipline of systems reliability exposes the theory of failure analysis in components and processes through the study of distribution models: exponential, weibull, lognormal, among others. Each of these distributions are accompanied by various functions that describe a variable for the purpose of studying it. Some of the main functions can be defined, according to Fogliatto and Ribeiro (2011), as: R (t), reliability function; MTTF (t), mean time to failure, represents the probability of a unit succeeding in the given time interval; h (t), risk function, can be considered the most widespread reliability measure in practice, such function can be interpreted as the amount of risk associated with a unit at time t.

The main characteristic of exponential distribution is a constant risk function, ie, units with the same reliability R (t) are assumed for any t, regardless of their age or time of use. The mathematical simplicity of the expressions derived from the exponential was what spread its use in the area, even if sometimes inadequate. Units that exhibit wear or fatigue are adequately modeled by exponential only during their lifetime, which is when the occurrence of failures is relatively constant over time.



graph 1: hazard function. Source: Elaborated by the authors.

Risk function equation, h (t), is equal to a constant (λ). A constant risk indicates that failure usually happens during the "life" of a product.

$$h(t) = \lambda \qquad (1)$$

We will not dwell upon distribution by distribution, as this is not our goal. The article was conceived during the elaboration of a work in the Reliability discipline, where the team used interactive methods to explain the use of exponential distribution with practical example, but is restricted to analyze the operation of the program developed in python.

2.2 Excel Usage and Functionality

The spreadsheet is seen as “[...] a two-dimensional arrangement of cells. Each cell, intersecting a row with a column, can contain a label, a value, or an expression ”according to Maxim and Verhey (1995, cited COXFORD and SHULTE 1999, p. 205). These spreadsheets are the result of software evolution and we cannot deny that Excel is an extremely developed tool, a leader in the engineering sector. According to the office package website, Microsoft Excel is the industry's leading spreadsheet program, a powerful data visualization and analysis tool, used at home by the high school student, the bakery owner, and large companies. Thus, Excel can be an effective, easy and convenient tool for teaching. Its interface is simple and clear, highlighting the following components: cells (small squares where we add the information), Workbook (within the same file you can create multiple folders), toolbar (where you access all available functions) , as can be seen from Figure 1

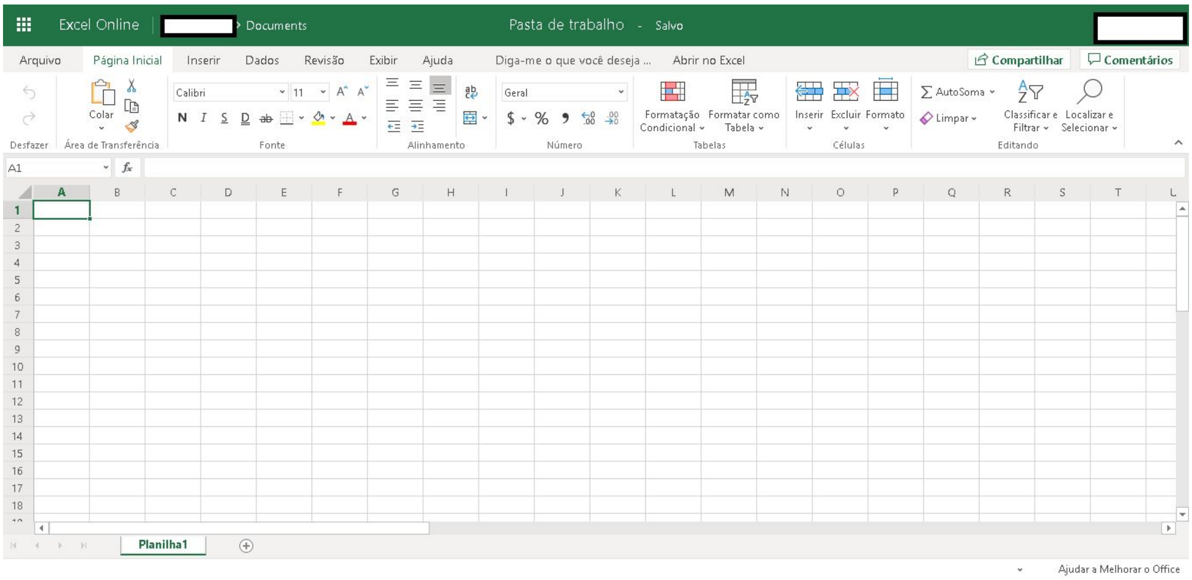


Figure 1. Excel Online Interface in Office 365 / Student Version

Excel is also used for assessment and validation of learning methods. Spreadsheets are used in didactics for teaching mathematics, where mathematical activities related to the study of mathematical functions, graphing (bar, line, histograms, etc.) are developed.

The point against the software used is that we have to pay for the full office package, and most students don't have the financial structure to get the software. Despite the "student version" output, the software becomes quite limited. Moreover, as it is a program with a general purpose proposal, there is a lack of functionality for more specific applications, such as the analysis of the thermodynamic properties of water (Giuma, 2019). Such gaps are filled with the addition of extensions called add-ins.

2.2 Python Usage and Functionality

According to data taken from the Python Brazil website, Python is a programming language that offers several aspects, such as:

- Web systems building with Django, Flask and Pyramid.
- Data Analysis, Artificial Intelligence and Machine Learning with Numpy, Pandas and Matplotlib.
- Building apps with Kivy and Pybee
- Building desktop systems with Tkinter and WxPython.

We can use various IDE's like Notepad ++, Sublime, Spyder and Pycharm.

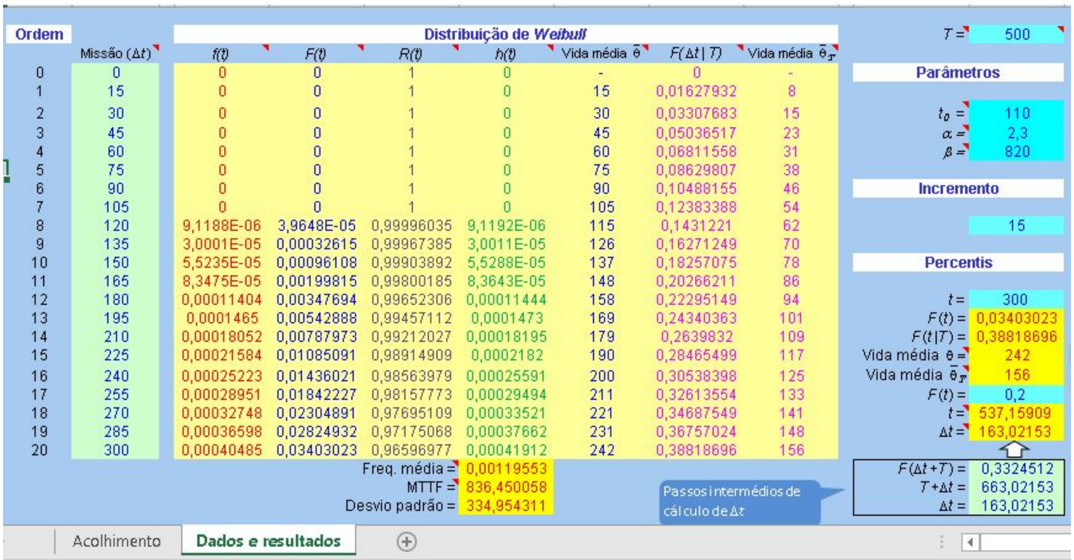
As mentioned earlier Python is concise, with a high level of abstraction, simplicity, conciseness and versatility, simple input and output, readable code, and a community that encourages the development of many modules and packages for a wide variety of applications. According to Chien-An et al (2015) Python should be taught as the first programming language because it has simpler syntax and high-level data structures that make writing code easier for students, and academic work proves that python is a Great programming language for didactic use.

According to a study by Oswaldo and Liliana (2019) at the University of Valle, Colombia, the numbers confirm many failures in programming courses with computer and engineering students. They were totally affected by the high level of disapproval, the students who enrolled in the courses with false prejudices and a prior knowledge of the level of the existing difficulty. Students were subjected to programming language such as C and Java which, for those who are not used to, offer complex code. The students formed very bad opinions about the importance of the programming course in their academic formation, considered as boring and disposable, however programming is an important tool that solves a wide range of engineering and computing problems. To solidify the article, the researchers compared the performance of the eight years prior to the application of the methodology, and the years after the application of the methodology. The pillars of the methodology are the use of python as the first programming language, project-oriented and problem-based learning, multimedia resources and assessment rubrics. The application of this methodology yielded partial results with positive impact.

The example given above does not only focus on the programming language used, however it shows that much of the advance was basically due to the change in the language used. Because C and Java have complex syntax (classes, methods, types, and complex syntax can be a challenge for beginning programmers and can make the learning process difficult), students spend more time trying to learn the syntax rather than understanding the programming logic. High-level scripting languages are simpler, more secure, and more flexible than low-level ones.

3. METHODOLOGY

As already mentioned during the reading, in this work there is a comparative study between two tools used in the academic environment. We evaluate the advantages and disadvantages of replacing paid software with free software, employing the knowledge gained in programming disciplines. The research results are obtained by evaluating the development of the parameters under which the tools are submitted and the characteristic curves obtained. For this we use a spreadsheet for academic use entitled “weibull distribution” which is available in Assisi (2018). This worksheet has column spaces on the right that must be filled in with the conditions under which our reliability function will work, some cells are in yellow, as they represent information that should not be changed as some additional data. reliability functions. In the center is a table showing data obtained from several of the major reliability functions. Note that the exponential curve is actually a special case of the weibull formula, where the shape parameter (α) is equal to 1. The spreadsheet requires some data, such as the minimum guaranteed life (t_0), shape parameter (α) and and scale parameter (β), which were obtained with the dynamics made in room. With this data in place, the spreadsheet performs the calculations necessary to obtain the main reliability functions, with the lifetime varying at user-defined intervals.



Ordem		Distribuição de *Weibull*						
	Missão (Δt)	$f(t)$	$F(t)$	$R(t)$	$h(t)$	Vida média $\bar{\theta}$	$F(\Delta t \mid T)$	Vida média $\bar{\theta}_T$
0	0	0	0	1	0	-	0	-
1	15	0	0	1	0	15	0,01627932	8
2	30	0	0	1	0	30	0,03307683	15
3	45	0	0	1	0	45	0,05036517	23
4	60	0	0	1	0	60	0,06811558	31
5	75	0	0	1	0	75	0,08629807	38
6	90	0	0	1	0	90	0,10488155	46
7	105	0	0	1	0	105	0,12383388	54
8	120	9,1188E-06	3,9648E-05	0,99996035	9,1192E-06	115	0,1431221	62
9	135	3,0001E-05	0,00032615	0,99967385	3,0011E-05	126	0,16271249	70
10	150	5,5235E-05	0,00096108	0,99903892	5,5288E-05	137	0,18257075	78
11	165	8,3475E-05	0,00199815	0,99800185	8,3643E-05	148	0,20266211	86
12	180	0,00011404	0,00347694	0,99652306	0,00011444	158	0,22295149	94
13	195	0,0001465	0,00542888	0,99457112	0,0001473	169	0,24340363	101
14	210	0,00018052	0,00787973	0,99212027	0,00018195	179	0,2639832	109
15	225	0,00021584	0,01085091	0,98914909	0,0002182	190	0,28465499	117
16	240	0,00025223	0,01436021	0,98563979	0,00025591	200	0,30538398	125
17	255	0,00028951	0,01842227	0,98157773	0,00029494	211	0,32613554	133
18	270	0,00032748	0,02304891	0,97695109	0,00033521	221	0,34687549	141
19	285	0,00036598	0,02824932	0,97175068	0,00037662	231	0,36757024	148
20	300	0,00040485	0,03403023	0,96596977	0,00041912	242	0,38818696	156

Freq. média = 0,00119553
MTTF = 836,450058
Desvio padrão = 334,954311

T = 500

Parâmetros

t_0 = 110
α = 2,3
β = 820

Incremento

15

Percentis

t = 300
$F(t)$ = 0,03403023
$F(t|T)$ = 0,38818696
Vida média θ = 242
Vida média $\bar{\theta}_T$ = 156
$F(t)$ = 0,2
t = 537,15909
Δt = 163,02153

$F(\Delta t + T)$ = 0,3324512
$T + \Delta t$ = 663,02153
Δt = 163,02153

Passos intermédios de cálculo de Δt

Figura 2. Planilha original da distribuição de Weibull. Disponível em: http://www.rassis.com/manutencao.html.
Source: Elaborated by the authors.

For Python programming, Python version 3.7 was used and the IDE (integrated development environment) was Pycharm, whose version is 2018.2.4. Obtaining the desired results Python programming was achieved with the model of Figure 2. To realize the dynamics, we first divided the reliability class into small groups that represented stages of a manufacturing process, and were given gummy candy of a specific color, that we call standard, and other ones with various colors that are different from the standard ones, along with bags to collect them. Standard color gummy candy meant a machine success and color gummy candy meant a mistake. The intervals between a success and an error were noted by the groups and recorded in a text file (.txt), as seen in Figure XXX. .

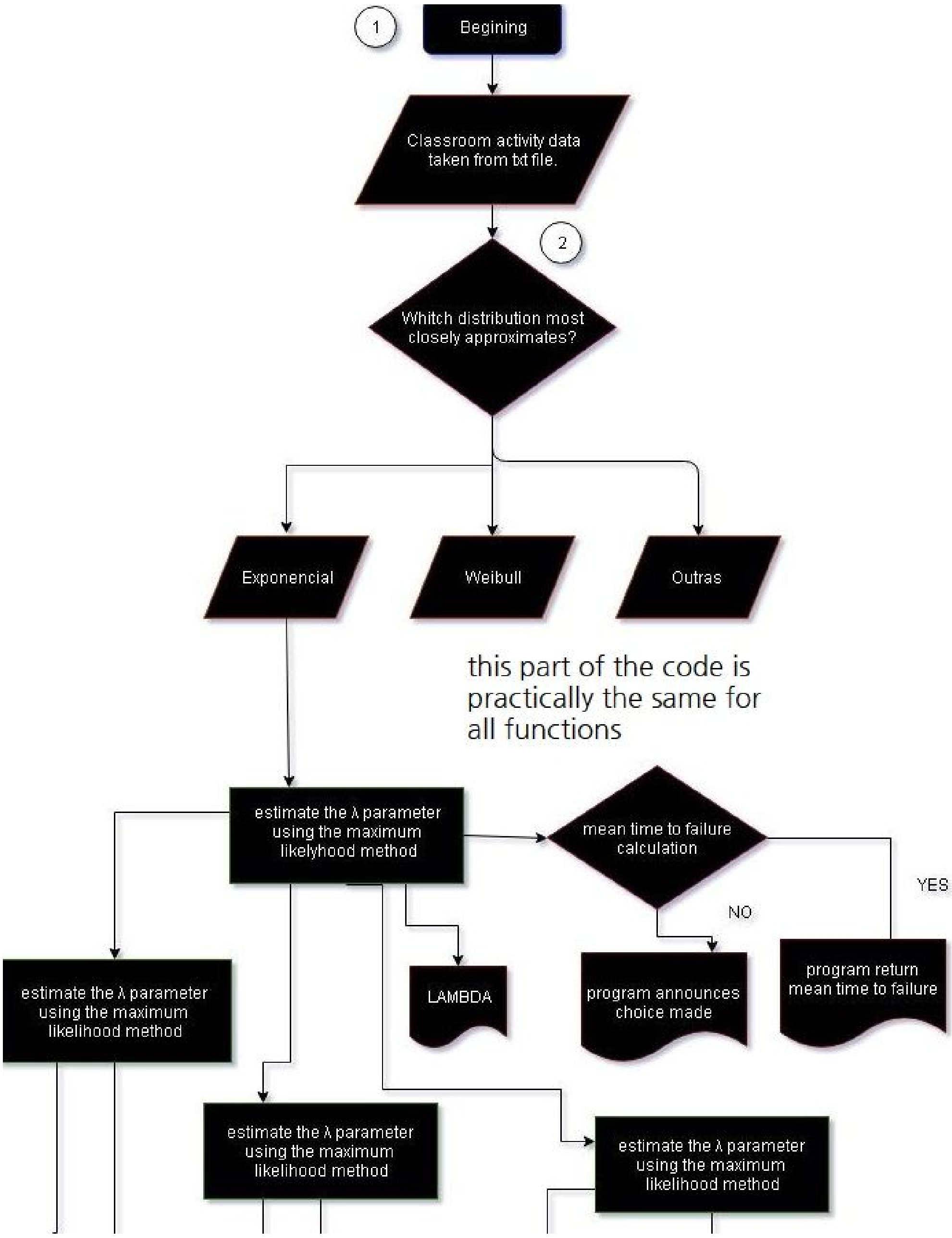
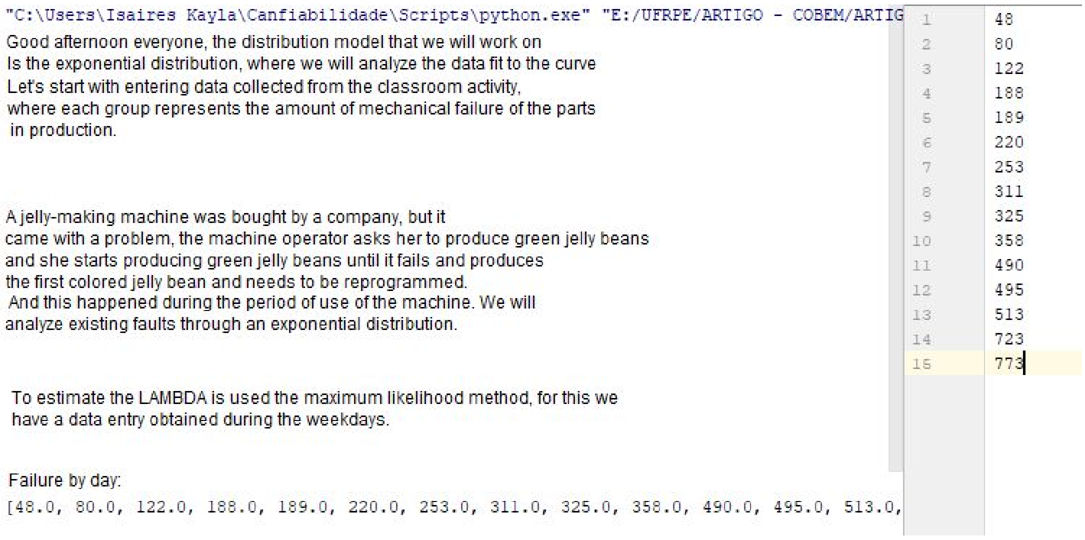

1
Begining
Classroom activity data taken from txt file.
2
Whitch distribution most closely approximates?
Exponencial
Weibull
Outras
this part of the code is practically the same for all functions
estimate the λ parameter using the maximum likelyhood method
mean time to failure calculation
YES
NO
LAMBDA
program announces choice made
program return mean time to failure
estimate the λ parameter using the maximum likelihood method
estimate the λ parameter using the maximum likelihood method
estimate the λ parameter using the maximum likelihood method

Figure 3 . Partial fluxogram of python program. Drawn on website: https://www.draw.io .
Source: Elaborated by the authors.



```
"C:\Users\Isaires Kayla\Canfiabilidade\Scripts\python.exe" "E:/UFRPE/ARTIGO - COBEM/ARTIG
Good afternoon everyone, the distribution model that we will work on
Is the exponential distribution, where we will analyze the data fit to the curve
Let's start with entering data collected from the classroom activity,
where each group represents the amount of mechanical failure of the parts
 in production.


A jelly-making machine was bought by a company, but it
came with a problem, the machine operator asks her to produce green jelly beans
and she starts producing green jelly beans until it fails and produces
the first colored jelly bean and needs to be reprogrammed.
 And this happened during the period of use of the machine. We will
analyze existing faults through an exponential distribution.


 To estimate the LAMBDA is used the maximum likelihood method, for this we
 have a data entry obtained during the weekdays.


Failure by day:
[48.0, 80.0, 122.0, 188.0, 189.0, 220.0, 253.0, 311.0, 325.0, 358.0, 490.0, 495.0, 513.0,
```

```
48
80
122
188
189
220
253
311
325
358
490
495
513
723
773
```

Figure 4. Fraction of the program showing (right) txt file input data and program interactivity (left). Source: Elaborated by the authors.

The Python program collected the data from the text file, and made it into lists whose language could manipulate its data. It is emphasized that the code was fully worked to communicate with the operator, not only in decision making but in the advancement of each phase. The data referred to each function were also added in txt file, and the graphics were customized to the most didactic form possible. To obtain the modeling of the function the maximum likelihood method was used. The most used methods to estimate parameters according to Fogliatto (2011) are: the moments, least squares and maximum likelihood method. These methods need to be non-biased, where it does not underestimate or overestimate the actual value of the parameter, need to be consistent (those that converge quickly to a result), efficient and sufficient.

Also according to Fogliatto (2011), one of the best methods for obtaining point parameter estimators is the maximum likelihood method, where an estimator will be given by the parameter value that maximizes the likelihood function. The maximum likelihood method was developed in 1920s by a famous British statistician, Sir RA Fisher, Montgomery & Runger (2006), and to use this method the population distribution must be known or assumed (in our case, the exponential).

Due to the fact that Python and Excel have different natures, the first one is a programming language and the second one is software, throughout the text we will call them work tools.

4. RESULTS

The program presents us as a result, after entering the data found in the txt file, the graphs of probability density, reliability and hazard function. Analyzing the reliability functions presented by the program in Python and Excel, we can see that both demonstrate the expected characteristics of an inverted exponential function, decreasing the probability of operation as the cycles pass. The two graphs start at just over 0.7 on the R (x) axis and cross the 70% reliability line around 5 cycles, and follow the same inverted exponential function until around 30 cycles, when they approach 0 ,1.

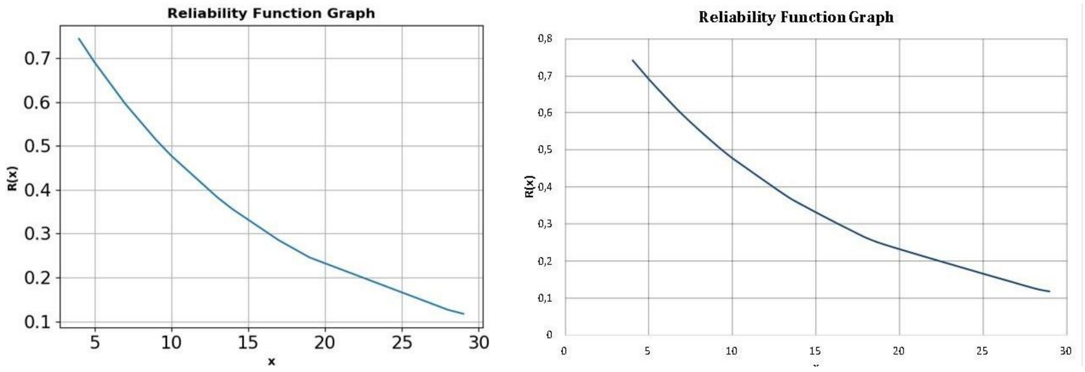


Figure 5. graphics of the reliability function drawn respectively on Excel and Python. Source: Elaborated by the authors.

As for the numerical results, there is a great similarity that once again makes it possible to use Python. Observing the reliability function, which according to the exponential distribution is a constant represented by h (x), the results differ by the last decimal place, as seen in Table 01. In the Python program the team made use of the method of calculation provided by Fogliatto and Ribeiro (2011) and the table obtained uses a similar device, differentiated only by the calculation of the density function. We then attribute this ninth decimal point to the calculation difference and the rounding off of the work tools.

Function\Tool:	EXCEL	PYTHON
Hazard function	0,073891626	0,073891625
Reliability function data associated to the first obtained point	0,744109928	0,744109928

Table 1. Comparative data withdraw from the work tools.

The points associated with the reliability density function can be seen in the following table. Note that the points are exactly the same:

Python data	Excel data
0.054983492	0,054983492
0.054983492	0,054983492
0.051067147	0,051067147
0.051067147	0,051067147
0.044051444	0,044051444

Table 2 Reliability density function data

Other relevant features involve the relationship between the person who manipulates the tool and the tool itself. In excel, the data is stored in easy-to-manipulate columns within the spreadsheet itself, and Python organizes it into txt files, as mentioned earlier. These files, besides being easily manipulated in Pycharm, can be accessed for any other utility quickly. Regarding charts in Excel, they belong to the same folder, but in different tabs. In Pycharm these charts can also be viewed on different tabs, although they open automatically. The layout of the data must suit the needs of the operator, he can choose between the facilities that Python or Excel have.

5. CONCLUSION

In this paper we use a program made by the authors to analyze the exponential distribution by estimating its reliability graphs and risk and reliability functions in a room dynamics that simulated a production line. The python language was chosen for its practicality and ease of manipulation, allowing the emergence of improvements in didactics and assimilation of content obtained by students, compared to existing programs.

With the program in place, it was possible to combine the production of realistic function graphs with the ease of user input more satisfactorily than was found using the Excel platform. The graphical results are quite coincident in format and approximation of data. Its popularity in the academic environment and its simplicity of use have led Excel to today's engineering-leading status, but the power to have a freer language that can make a program that matches Our expectations without any loss in the accuracy of results can meet the need to use software to simplify the data collection and processing work in engineering.

It was observed that this study opens possibilities to get out of repeatability, using a unique software, and enjoy the diversity of information and stimuli that students absorb every day to create programs that solve a wide range of problems that are found in engineering and in computing.

6. ACKNOWLEDGEMENTS

We thank God for all the blessings and blessings granted and those who benefit from our work: family, patience and understanding, friends, appropriate skills and studies, and the university with the pedagogical support for technical research. . Our article partners Fernando and Wildson and our advisor teacher, Adriano Marques.

7. REFERENCES

Fogliatto, F. S. Andibeiro, J. L. D., 2011. Confiabilidade e manutenção industrial. Elsevier, Rio de Janeiro, 1nd edição.

Silva, E.C., Ferreira, G.S, Santos, M.M.M., Andrade, P.C.R., Rosa, R.M.G., 2017. "Análise de Dados de Falha de um Transmissor de Fibra ". *Revista Thema*, Vol. 14, No. 4, pp. 259 - 266.

Szidarovszky, M. ; Szidarovszky, F.; Honecker, Sharon L., 2018. "A General Formula for Expected Number of Failures". In: *Reliability And Maintainability Symposium*.

Binoti, D.H.B., Leite, H.G., Silva, M.L.M., 2017. "Sistema Computacional para Ajuste de Funções Densidade de Probabilidade ". *Floram, Floresta e Ambiente,*

Menezes, N.N.C., 2014. Introdução a Programação em Python. *Novatec*, São Paulo, 2nd edição.

Borges, L.E., 2010. Python para Desenvolvedores. 2nd edição. < *http://ark4n.wordpress.com/python/*>

RELIASOFT CORPORATION. Life data analysis reference. Tucson, Arizona: [s. n.], 2015. Disponível em: *http://www.reliawiki.org/index.php/Attribution_Example.* Acesso em: 6 fev. 2019.

ASSIS, Rui Manuel Assis. Distribuição Weibull. [S. l.], 17 jul. 2018. Disponível em: *www.rassis.com/manutencao.html.* Acesso em: 29 out. 2018

GIUMA,Fellah. (2019). Excel Spreadsheet as a Tool for Simulating the Performance of Steam Power Plants. *Spreadsheet in Education.* vol. 2. 1-19.

Hegazy, Tarek & Ersahin, Tolga. (2001). Simplified Spreadsheet Solutions. I: Subcontractor Information System. *Journal of Construction Engineering and Management-asce - J CONSTR ENG MANAGE-ASCE*. 127.

Solarte, Oswaldo & Machuca, Liliana. (2019). Fostering Motivation and Improving Student Performance in an Introductory Programming Course: An Integrated Teaching Approach. *Revista EIA*. 16. 65.

Montoya-Araque, Exneyder A. & Suarez-Burgoa, Ludger. (2018). Application software in Python 3 for calculating the second-order tensors Jelínek's statistics in anisotropy of magnetic susceptibility data - Software de aplicación en Python 3 para el cálculo de la estadística de tensores de segundo orden de Jelinek en datos de anisotropía de susceptibilidad magnética. *Boletín de ciencias de la tierra*. 44. 49-58.

Miños Fayad, Alejandro. (2016). Uso didáctico de estrategias inductivas en un curso introductorio de programación estructurada. *TED: Tecné, Episteme y Didaxis*.

Hegazy, Tarek & Ersahin, Tolga. (2001). Simplified Spreadsheet Solutions. I: Subcontractor Information System. *Journal of Construction Engineering and Management-asce - J CONSTR ENG MANAGE-ASCE*. 127.

Kuka, S., & Karamani, B. (2013). Using Excel and VBA for Excel to Learn Numerical Methods. *1st International Symposium on Computing in Informatics and Mathematics (ISCIM 2011).*

8. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper