A simulation of data censored rigth type I with weibull distribution

Daniel Gaspar
Department of Mechanical Engineering
ESTGV, Institute Polytechnic of Viseu
Viseu, Portugal
dangas05@gmail.com

Luis Andrande Ferreira
Department of Mechanical Engineering
FEUP, University of Porto
Porto, Portugal
lferr@fe.up.pt

***Abstract*—In the maintenance and reliability field, there are frequent analyses with data being censored. In reliability research, many articles do simulation, but few explain how they do it. the loss of information resulting from the unavailable exact failure times will impact negatively the efficiency of reliability analysis. This paper presents four different algorithms to generate random data with a different number of censored values. The four algorithms are compared, and tree parameters are used to select the best one. The Weibull distribution is used to generate the random numbers because it is one of the most used in reliability studies. The results of the algorithm chosen are very relevant; with a sample of $n = 50$ and a number of cycles of simulations $m = 1000$, the standard deviation is higher when the shape factor of Weibull distribution is $\beta = 0.5$ and slowly decreases until the shape factor equals 5. The percentage error (PE), one of the indicators selected, is much higher when the percentage of censored data is $c = 5\%$, then goes down when the shape factor increases. There is a different behaviour when censored data is $C = 20\%$ and the percentage error (PE) is higher when shape factor is $\beta = 1.5$.**

This article presents an algorithm that it considers the best for simulating right-censored type-I data. The algorithm has excellent accuracy, random data i.i.d and excellent computational performance.

***Index Terms*—Data censored, Reliability, Algorithm simulation, Weibull distribution**

I. Introduction

In the survival analysis and reliability field, there are several situations in which equipment, components, and units are lost or taken from the study while they are still working. The data censored may occur in control situations, as in life-testing and preassigned time or in actual operations, and to make a predictive analysis of failures on time, with systems with huge numbers of sensors and monitoring lots of parameters; in this case, using reliability models containing censored data is fundamental.

Genschel and Meeker (2010) refer that, in practice, life test data are almost always time-censored or type I because the study defines the time at which the test will end [1]. Balakrishnan et al. (2000) have more details about when the progressive censoring schemes take place [2]. Several methods and techniques have been proposed for analyzing different types of reliability data over the past decades. Most of them refer to complete data. However, the evaluation of highly censored reliability data has not been widely studied. Nelson (1985) presented an excellent work on this topic [3]. In the beginning, few of the studies used simulation tools, but over time the use of simulation in the reliability field increased, most of them to estimation parameters.

Olteanu and Freeman (2010) conducted a simulation study that compared the performance of maximum likelihood (ML) and median-rank regression (MRR) methods in estimating Weibull parameters for highly censored reliability data [4]. In addition to the well-known large-sample optimal properties associated with ML estimators, experience, including many simulation studies, has shown that ML estimators are generally hard to beat consistently even in small samples [1], [5]. Recently, the estimation of parameters from different lifetime distributions based on progressive type-II censored samples are studied by several authors, including [6], [7], [8], [9].

This article is concerned with the analysis of the simulation of censored reliability data. It is true that the loss of information resulting from the unavailable exact failure times will impact negatively the efficiency of reliability analysis. Many articles use the percentage of data censored (% C) to compare and analyse the model and study simulations, like in [10], [11], and [12]. The use and application of data censored in the field of reliability can be seen in [13], [14]. The type of distribution used in this study is typically used in the reliability field. The significant contributions brought forth by this paper are: (i) understand and develop a systematic method to build an accurate simulation model in the presence of data censored, (ii) give more accuracy and precision to the simulation process in the reliability field, and (iii) in addition, our proposed algorithm can find an accurate solution within a a relatively short time of the simulation.

II. The right data censored

The data is considered complete when the exact time of each system failure is known. In many cases, the data contain uncertainties, i.e., the exact moment when the failure occurred is not known. The data containing such uncertainty as to be when the event occurred are regarded as incomplete or partial. Incomplete data can be classified as censored or truncated [15].

Censoring, from the theoretical point of view, may not be the most efficient way to conduct an experience, but due to time, cost or practical things, it's so frequent that researchers had to find ways to deal with it.

Characterizing the censoring mechanisms is essential to analyze the data and the phenomena in the study. Such a report can be based on several elements, the status of the entity observed, the span of the study, the dynamic of the system in the study, and the times of start and finish of the observations. Censoring mechanisms can also be characterized based on when and how the time to finish the study is defined. One of the most common types of censored data that may arise in real cases is type-I right censored data.

In type-I right censored data, all units of a system are observed up to the date of completion of the study. For this censorship scheme, the time each unit is under observation is fixed, while the number of units that fail (uncensored observations) is random. In this type of censoring, the stopping time (t_c) is defined or pre-established, and the number of failures observed during the analysis period is random. Putting an end to the experiment and stopping monitoring all the entities at some pre-specified time t_c, independent of the event of interest. The Weibull distribution is the most popular statistical distribution used in reliability engineering [16]. It can be used to fit many life distributions, and it has a significant advantage in the reliability field by changing the parameters to adjust perfectly to the reliability data.

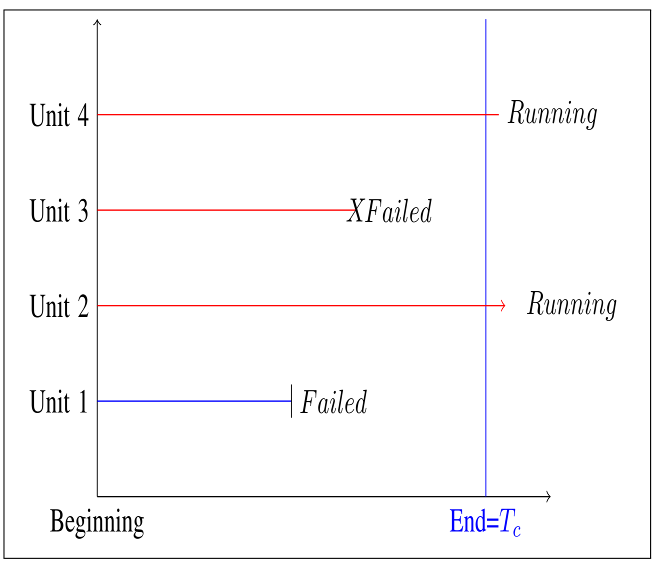

Fig. 1. Fixed type I right censored

The type I censoring occurs when the experiments are run only for a fixed duration t_c; the lifetimes are known for those individuals whose lifetimes are $t_i \leq t_c$, as it's possible to see in fig. 1.

The difference between type I and type II is that in type I censoring, the number of observed lifetimes is a random variable, and in type II the number of observed events is fixed.

III. The algorithms to generate random data censored (right type I)

Burton et al.(2006) proposed to generate a random non-informative right censoring with a specific proportion of censored observations in a similar manner to the uncensored survival times by assuming an exponential distribution for the censoring times but can be Weibull or uniform [17]. For [17] it's by iteration that the parameters of censoring distribution will be achieved [17]. Halabi and Singh (2004) in another way, provide formulas for determining parameters for standard survival and censoring distribution [18]. The censoring mechanism can also be extended to incorporate dependent, informative censoring.

A fundamental part of any simulation is the algorithm used to generate the random numbers. The function of *R* software to generate random number generation is the "Mersenne-Twister", from Matsumoto and Nishimura (1998). A twisted GFSR with period 219937^{-1} and equidistribution in 623 consecutive dimensions (over the whole period). The "seed" is a 624-dimensional set of 32-bit integers plus a current position in that set.

In the algorithms, it's essential to define how the results will be stored after each simulation, to avoid the risk to repeat the simulations. The estimate of interest will be t_c - time censoring for each sample. The number of samples - *n* will be 50 and 1000. The routine is made *m* times (in this case 1000) and it's important to calculate the mean, as a measure of the true estimate of interest:

$$mu = \frac{\sum_{i=1}^{m} t_{c_i}}{m} \tag{1}$$

The results of simulations can measure the uncertainty in the estimate of the parameter t_c which represents the percentage of %C data censored. The empirical standard deviation (SE) σ is calculated as the standard deviation of the estimates of interest from all simulations (in this case $m = 1000$).

$$\sigma = \sqrt{\frac{\sum_{i=1}^{m}(t_{c_i} - \mu)^2}{m}} \tag{2}$$

The average of the estimated within the study simulation, σ could be used. Increasing the number of simulations will reduce the SE - σ of the simulation process, i.e. $\sigma(t_c)/\sqrt{m}$, but this will be computationally expensive and therefore variance reduction techniques could be used.

After the simulation has been performed, it's necessary to define the criteria for evaluating the results obtained from the different scenarios or statistical approach, in this study more precisely is the change of parameters of each distribution being studied.

The comparison of the results with the true values of the simulation provides a measure of the performance and associated precision of the model and the algorithm in the study. Some examples of performance measures that are often used include assessment of bias, accuracy, time of simulation, etc...

The estimates of simulations are the main reason and hence the average of estimates overall simulation is used to calculate accuracy measures. When analyzing different scenarios or models, there is a trade-off between the amount of bias and the dispersion or variability. Some authors argue that having less bias is more crucial than producing a valid estimate of sampling variance. However, models, scenarios, or methods that result in a biased estimated with little variability may be considered not so accuracy or conversely if exist an unbiased estimate with large variability.

To evaluate the performance of statistical methods and algorithms with different distribution parameters we use MSE and the PE - percentage error associated with the estimated of each t_c time censoring. The PE associated with the estimate t_c with a true value of T_C is computed by the following relation:

$$PE_{T_v} = \xi = \frac{\left|t_c - T_{c(exact)}\right|}{T_{c(exact)}} x100 \quad (3)$$

A. The first algorithm

The first algorithm initialized with the percentage of data censored ($\%C$), the parameters of distribution (β, α) and the number of simulation cycles (m); the number of samples - n is not defined. The time censoring t_c is selected taking into account the parameters from the Weibull distribution and the required value defined by the experience. The algorithm uses the cycle/loop using the *do-while* function until get the target, and only follows to the next step after the number of censored data is the same value required;

Step 1 Define initial parameters $(\%C, \beta, \alpha, m\ )$
Step 2 Select T_c and verify if is scaled
Step 3 Generate the vector Y that represent t_i random times from distribution model
Step 4 Compare $\frac{n_{cens}}{n_{total}} \leq C$
Step 5 Repeat *n* times from step 3 to 4
Step 6 Repeat the procedure *m* times - number of simulations
Step 7 Collect and analyse results

The flow chart that resumes this first proposal model can be seen in fig. 2.

The model can be used in a situation when don't know exactly the number of censor data required or the time of censoring it's not controlled or easy to control. The solution has a great time-consuming and resources in computation point of view.

B. The second algorithm

The second algorithm, defined in the beginning all parameters except the censoring time t_c. In this case, the censoring time is a result of the cycle and the random generation and their calculation depend on the %C censoring data. In this algorithm, the value of %C is exact but can happen, sometimes, not having enough random censor data or the inverse... It's a simulation model that can be used in some applications where the censoring times are not important and the time of computation must be optimized. In this case the number *n* of sample is define but the t_c is not controlled.

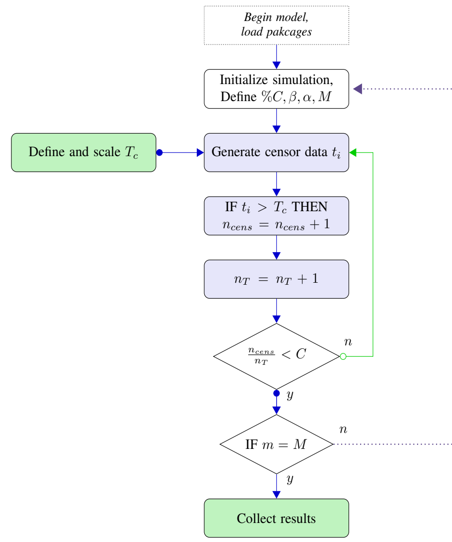

Fig. 2. Flow chart of model/algorithm 1

Step 1 Define initial parameters $(\%C, \beta, \alpha, m, n)$
Step 2 Generate the vector Y that represent t_i random n times from distribution model
Step 3 Find t_c that represent $\frac{n_{T_c}}{n_T} < C$
Step 4 Repeat the procedure *m* times - number of simulations
Step 5 Collect and analyse results

The flow chart that resumes this second model can be seen in fig. 3.

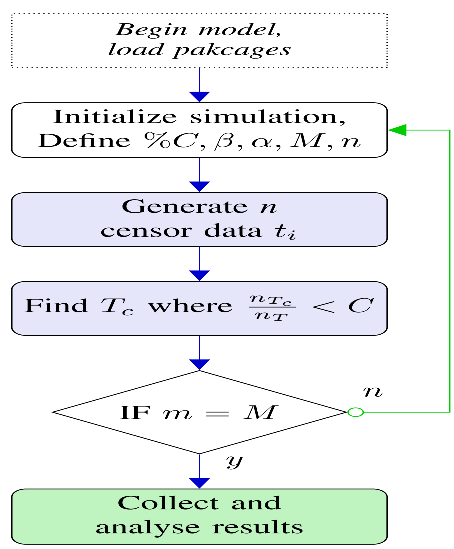

Fig. 3. Flow chart of model/algorithm 2

The model can be used in a situation when don't know exactly the time of censoring, but know the number of samples.

This model can be used when the number of samples and percentage of data censored is important. It's a model that has the great disadvantage of not controlling the time censoring. Normally this can be fit better when simulating the right type II data censored.

C. The third algorithm

The third algorithm begins to define all the parameters required for simulation, the percentage of data censored ($\%C$), the parameters of distribution (β, α), the number of simulation cycles (m) and the number of samples - n. After that, the model generates a random vector Y of t_i with the dimension of the samples n of the Weibull distribution and generates another random vector X from a binomial distribution with *(0,1)* and the number of zero's is equal of the percentage of data censored - % C. The model presents a very practical solution, but with some loss of accuracy and it's easier to have some bias in the output of the model. It needs to define all parameters and in the initial step take out all the values t_i that exceed t_c. The algorithm filter the values t_i that exceed t_c, and for that reason it requires one more or two steps, and can take the algorithm no so fast.

Step 1 Define initial parameters $(\%C, \beta, \alpha, m, n, t_c)$
Step 2 Generate the random vector Y with n t_i and repeat until all $t_i < t_c$
Step 3 Generate the random Binomial *(0,1)* vector X with number correspond of percentage of % C of zero's
Step 4 Multiply the two vectors (*O's* represents the t_i censored)
Step 5 Repeat the procedure m times - number of simulations
Step 6 Collect and analyse results

The flow chart that resumes the third proposal model can be seen in fig. 4.

This model uses the random vectorization data to be faster and less time-consuming for resource computation. It's a model that has one step in the first generating random time censoring, which could be necessary to repeat several times.

D. The four algorithm

The four models optimize the simulation and technically give very good results. Begin to initialize all parameters, but exist one step before running the model: - the calculation of the time of censoring t_c for each percentage of data censored, and that varies and is different from each statistical distribution.

The probability of a value - random number generate - falling between a region $(x, +\infty)$ is:

$$P(x_1 > X) = \int_{x_1}^{\infty} f(x)dx \tag{4}$$

Which can see as the same as the definition of the function $R(t)$ - reliability

$$R(x) = \int_{x_1}^{\infty} f(x)dx \tag{5}$$

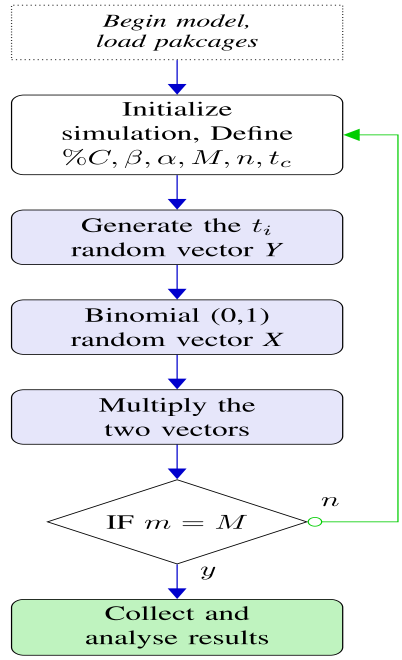

Fig. 4. Flow chart of model/algorithm 3

All values random generates that fall in that region $(x, +\infty)$ are the censored data; It's easy to achieve this relation between the reliability and the C% percentage of data censored with the expression:

$$R(t_c) = \int_{t_c}^{\infty} f(t)dt = C \tag{6}$$

The density function is given by

$$f(x, \eta, \beta) = \frac{\beta}{\eta^{-\beta}} t^{\beta-1} e^{-(\frac{t}{\eta})^{\beta}} \; with \; t \in \mathbf{R}^{+} \tag{7}$$

And the corresponding reliability function is

$$R(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}} \tag{8}$$

The shape parameter β is non-dimensional and reflects the type of failure mode, such as infant mortality ($\beta < 1$), random or exponential ($\beta = 1$), or wear-out ($\beta > 1$). To have %C of data censored, it's the same to equal the expression of reliability:

$$R(t_c) = \int_{t_c}^{\infty} f(x)dx = e^{-\left(\frac{t}{\eta}\right)^{\beta}} = C \tag{9}$$

and resolve the equation in order of t_c, results:

$$t_c = \eta * (-\log(C))^{\frac{1}{\beta}} \tag{10}$$

that gives the time censoring with the %C percentage of data censored required.

Generically is to do the inverse function of *pdf* function, calculate the time censoring t_C, and put this value in the algorithm of simulation with this value. This procedure reduces the time-consuming computation and with a large sample is very precise and comes closest to the percentage of censored data defined or theoretical.

The algorithm to generate random data censor have the follow steps:

Step 1 Define initial parameters $(\%C, \beta, \alpha, m, n)$
Step 2 Calculate t_c (with parameters of distribution and number of data censor %C)
Step 3 Calculate the order of i^th number that begins the censored data of a sample (censoring-order i_c).
Step 4 Generate the vector Y that represent t_i random times from distribution model
Step 5 Order the vector Y
Step 6 Find the time for censoring-order - $Y(i_c)$
Step 7 Repeat M times from step 3 to 6 (save to T_{cens})
Step 8 From vector T_{cens} calculate the mean and standard deviation
Step 9 From step 8 calculate the error $\xi = |T_c - \mu|$

The flow chart that resumes the four models can be seen in fig. 5.

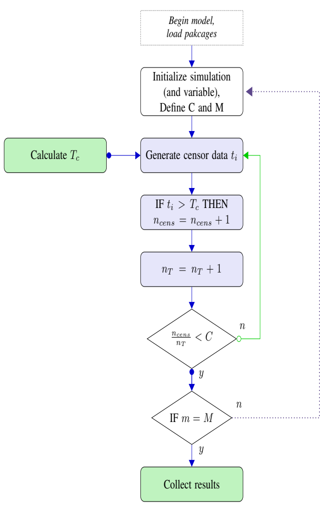

Fig. 5. Flow chart of model/algorithm 4

In conclusion, define first the %C of censored data and then calculate the time censoring t_c to produce the random number generating to which the result has the probability calculated.

IV. The Simulation Study

To illustrate and compare the results from the algorithm, two random sample of size of *n=50* and *1000*, to take care of medium and large data sets. The scale parameter was chosen to be **1** and the shape parameter or standard parameter, depending on the statistical distribution and vary between *0.5, 1, 1.5, 3* and *5*. These were replicated **1000 times - m**, the number of cycles simulations. For each model, it calculates the censorship time which has a reliability of 5%, 10%, 20%, or 30%. In a great number of reliability studies, it's used the same or similar, range of values from %C - percentages of censorship data.

One way to compare the performance and quality of algorithms is to compare against some criteria. For the first criteria the accuracy indicator is used, that tells whether the algorithm is very close or far from the values that are the original or true ones. The second criteria was to identify whether the values of each sample could be considered i.i.D, using for this purpose the test batteries on the randomization of values. Finally, the criteria of cycle time, and computational performance. This evaluation was based on empirical knowledge and work developed by [19]. The Table 1 summarizes an evaluation made of each of the models with three parameters of comparison. The scale used for each of the parameters was from 0 to 5. In the end, the calculation is made to identify the best model. In this case, can be seen in table one that model four clearly stands out and is undoubtedly the best.

TABLE I
Assesment and Compare Models

Model.	Accuracy	i.i.d	Time cycle	Total
1	2	3	2	12
2	3	3	2	18
3	4	3	3	36
4	4	4	4	**64**

A. *Results from model four with Weibull distribution*

The study from Weibull distribution performed an analysis for the shape factor β with a range from *0.5, 1, 1.5, 2, 3 and 5*, which are very illustrative of the shape factor β variation; the scale factor used is $\alpha = 1$. The simulation for each shape factor and the following percentage of censored data is 5%, 10%, 20%, and 30%. The resume of the study is in two tables that summarize the analysis. The first table is the simulation made with sample $n = 50$ and the second is with sample $n = 1000$.

TABLE II
Simulation right type I , Weibull $(\beta, C\%)$, $\eta = 1, n = 50$

	5			10			**20**			30		
	μ	σ	ξ	μ	σ	ξ	μ	σ	ξ	μ	σ	ξ
$\beta_{0.5}$	7.67	2.87	14.5	5.18	1.96	2.3	2.60	0.88	0.5	1.47	0.51	1.3
β_1	2.64	0.51	11.8	2.23	0.42	3.2	1.58	0.28	1.6	1.22	0.22	0.9
$\beta_{1.5}$	1.88	0.29	9.5	1.73	0.20	0.6	1.33	0.15	3.1	1.11	0.15	1.8
β_2	1.61	0.17	6.9	1.51	0.13	0.8	1.24	0.11	2.4	1.09	0.09	0.8
β_3	1.38	0.08	4.0	1.29	0.08	2.6	1.16	0.08	1.0	1.06	0.06	0.5
β_5	1.20	0.05	3.3	1.17	0.04	0.9	1.09	0.04	0.7	1.03	0.03	0.6

The results of the Weibull distribution are very interesting. With a sample of $n = 50$ and a cycle of simulations $m = 1000$, the standard deviation is higher when $\beta = 0.5$ and slowly decreases until $\beta = 5$. The percentage error is much higher when the C=5% and then goes down when the shape factor increases. There is different behaviour in C=20%; in this case, the standard deviation and the *PE - percentage error* is higher when $\beta = 1.5$; this could have an explanation because of the transition of shape from exponential to standard shape. The simulation to a sampling number of $n = 1000$ doesn't have the same behaviour, probably could be some phenomena with the random generation number in these particular distribution parameters.

TABLE IV
SIMULATION RIGHT TYPE I, WEIBULL $(\beta, C\%)$, $\eta = 1$, $n = 1000$

	5			10			**20**			30		
	μ	σ	ξ	μ	σ	ξ	μ	σ	ξ	μ	σ	ξ
$\beta_{0.5}$	8.93	0.81	0.5	5.27	0.44	0.7	2.59	0.21	0.2	1.46	0.12	0.4
β_1	2.99	0.14	0.2	2.30	0.10	0.2	1.61	0.06	0.1	1.20	0.05	0.2
$\beta_{1.5}$	2.08	0.06	0.1	1.74	0.05	0.2	1.37	0.04	0.0	1.13	0.03	0.1
β_2	1.73	0.04	0.3	1.52	0.03	0.1	1.27	0.02	0.2	1.10	0.02	0.2
β_3	1.44	0.02	0.2	1.32	0.02	0.1	1.17	0.02	0.1	1.06	0.01	0.0
β_5	1.24	0.01	0.1	1.18	0.01	0.0	1.10	0.01	0.1	1.04	0.01	0.1

Table IV shows a simulation of a sample of 1000 and in this case, the standard deviation and PE are smaller than in the case of the number of samples is 50. To all simulations, it can be noted that the error is less than 1%, which is very small, and even the dispersion itself is minimal, as can be seen by the table IV and table II. As the shape factor increases, there is a slight decrease in dispersion and error. Finally, can conclude that for Weibull distribution, this algorithm can be used. Still, it needs to have cautious and choose a higher number of sampling to give more accuracy to the simulation study. In fig. 6 and fig. 7, it can see the dispersion, the bias, and the mean of simulation graphically and compare with the true value.

V. CONCLUSIONS AND OUTLOOK

Survival testing and reliability studies are usually focused on estimating an unknown cumulative distribution function (CDF). In simulation studies, it's normal to use computational power to test particular hypotheses and assess the validity and accuracy of various statistical methods or procedures concerning a known truth. These procedures and algorithms provide an empirical estimation of the sampling distribution of the parameters of interest.

In fig. 8, which summarizes in a condensed and graphic form the results of the tables IV and II, it can be seen that there is not an equal pattern for all shape factors, but rather a tendency to approach the value as the percentage of censored data increases. That is, the more censored data there is in the sample, the less bias and less error, relative to the true value, the sample has. The behavior of factors $\beta = 0.5$ and $\beta = 1.5$ are very similar but with different scales. Undoubtedly, the curve with the worst behavior, that is, the greatest deviations, is with $\beta = 0.5$. And on the opposite side, the curve with the best performance and lowest error is when β is very high, in this case $\beta = 5$.

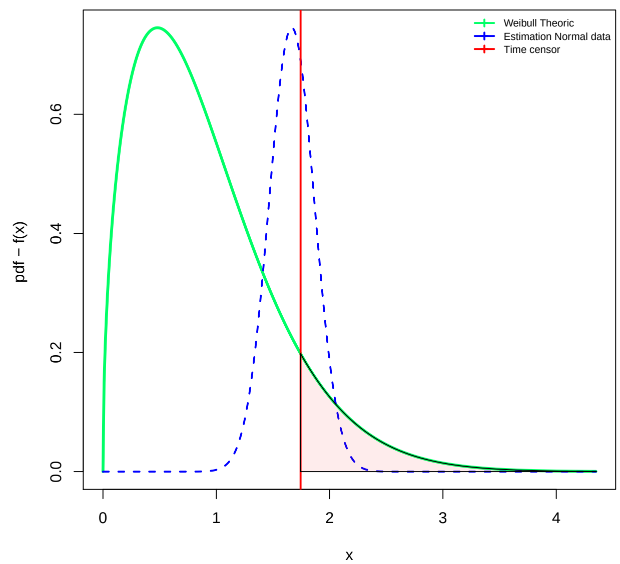

Fig. 6. $\beta_{1.5}$,$C = 10\%$ and n=50

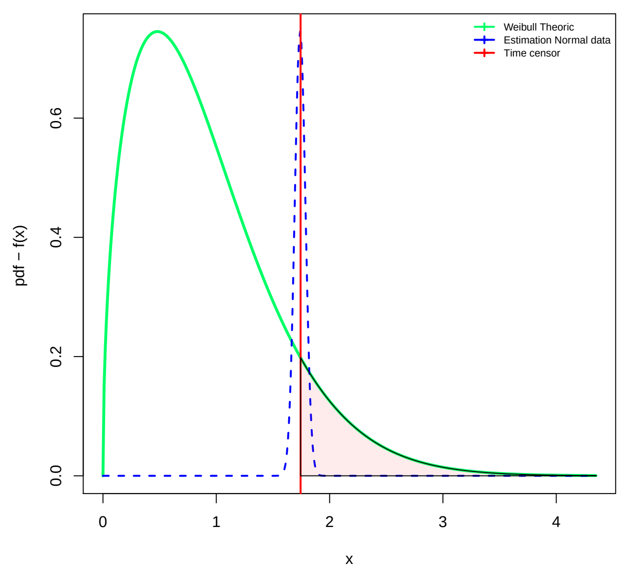

Fig. 7. $\beta_{1.5}$, $C = 10\%$ and n=1000

The study confirms under the physically motivated assumption that the distribution of the generalized deviations does not depend on changes in specific parameters (e.g. the scale parameter in the distribution in the study). Based on the experiences and intuitions, a value of σ in the neighbourhood and below 1 tends to make the deviation distribution close to

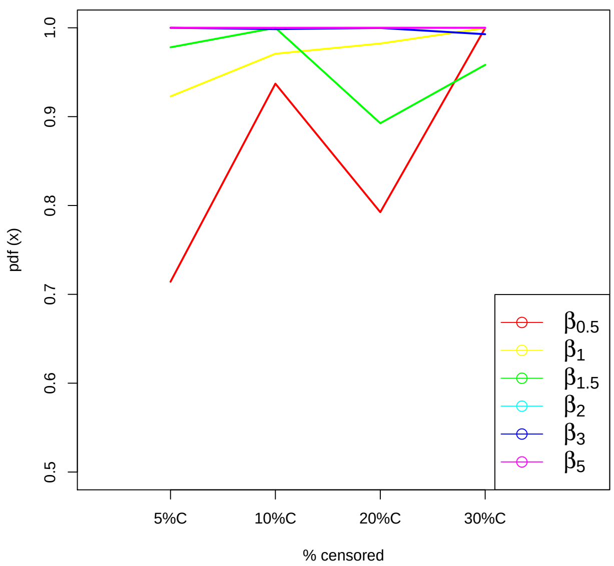

Fig. 8. Weibull desviation changing %C and n=1000

i.i.d. $N(\mu, \sigma^2)$ over a wide range of testing conditions.

This paper pretends to help the development of the best procedures to generate a sample of data with a particular characteristic (right censored and type I) and needs to be random (i.i.d) to be used to simulate in the reliability field.

In conclusion, the work pretends to generate more discussion and attention to the algorithms that simulate data censored and give some tools and results to make the simulations and the studies more accurate and optimized.

The next steps for this work would be to continue the study with the same algorithm for the other statistical distributions, namely the exponential, Gamma, Log-normal and Normal distributions. Another important step would be to verify whether the chosen algorithm was well adapted to other types of censored data, as would be the case with type II censored data. Finally, as this work was developed in a specific software language, in this case the R software, it would also be interesting to verify the performance of algorithms in other languages, such as python or C++.

References

[1] U. Genschel and W. Q. Meeker, "A comparison of maximum likelihood and median-rank regression for weibull estimation," *Quality Engineering*, vol. 22, no. 4, pp. 236–255, 2010.

[2] N. Balakrishnan, N. Balakrishnan, and R. Aggarwala, *Progressive censoring: theory, methods, and applications*. Springer Science & Business Media, 2000.

[3] W. Nelson, "Weibull analysis of reliability data with few or no failures," *Journal of Quality Technology*, vol. 17, no. 3, pp. 140–146, 1985.

[4] D. Olteanu and L. Freeman, "The evaluation of median-rank regression and maximum likelihood estimation techniques for a two-parameter weibull distribution," *Quality Engineering*, vol. 22, no. 4, pp. 256–272, 2010.

[5] A. Somboonsavatdee, V. N. Nair, and A. Sen, "Graphical estimators from probability plots with right-censored data," *Technometrics*, vol. 49, no. 4, pp. 420–429, 2007.

[6] N. Balakrishnan and N. Kannan, "Ch. 14. point and interval estimation for parameters of the logistic distribution based on progressively type-ii censored samples," *Handbook of statistics*, vol. 20, pp. 431–456, 2001.

[7] M. A. Mousa and Z. Jaheen, "Statistical inference for the burr model based on progressively censored data," *Computers & Mathematics with Applications*, vol. 43, no. 10-11, pp. 1441–1449, 2002.

[8] A. Childs and N. Balakrishnan, "Conditional inference procedures for the laplace distribution when the observed samples are progressively censored," *Metrika*, vol. 52, no. 3, pp. 253–265, 2000.

[9] A. A. Soliman, "Estimation of parameters of life from progressively censored data using burr-xii model," *IEEE Transactions on Reliability*, vol. 54, no. 1, pp. 34–42, 2005.

[10] N. Balakrishnan and D. Mitra, "Left truncated and right censored weibull data and likelihood inference with an illustration," *Computational Statistics & Data Analysis*, vol. 56, no. 12, pp. 4011–4025, 2012.

[11] A. Birolini, *Reliability Engineering: Theory and Practice*. Springer Berlin Heidelberg, 2017. [Online]. Available: https://books.google.pt/books?id=SXckDwAAQBAJ

[12] S. Ross, *Simulation*, ser. Knovel Library. Elsevier Science, 2012. [Online]. Available: https://books.google.pt/books?id=1Dwsyyty3P8C

[13] P. Wang and D. W. Coit, "Repairable systems reliability trend tests and evaluation," in *Proceedings of the Annual Reliability and Maintainability Symposium*. Citeseer, 2005, pp. 416–421.

[14] R. Horst, *The Weibull Distribution A Handbook*, 2009.

[15] I. Gijbels, "Censored data," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 2, pp. 178–188, 2010.

[16] D. P. Murthy, M. Xie, and R. Jiang, *Weibull models*. John Wiley & Sons, 2004, vol. 505.

[17] A. Burton, D. G. Altman, P. Royston, and R. L. Holder, "The design of simulation studies in medical statistics," *Statistics in medicine*, vol. 25, no. 24, pp. 4279–4292, 2006.

[18] S. Halabi and B. Singh, "Sample size determination for comparing several survival curves with unequal allocations," *Statistics in Medicine*, vol. 23, no. 11, pp. 1793–1815, 2004.

[19] D. Gaspar, "The advanced reliability and maintainability analysis using simulation tools."