

A “FRAMEWORK” TO MANAGE REPAIRABLE SUPPLIES USING R-STUDIO SOFTWARE.

OSCAR DE ALMEIDA MACHADO

Academia Militar das Agulhas Negras
Rod. Presidente Dutra Km 306, Paraíso, Resende-RJ, 27534-970.
Instituto Tecnológico de Aeronáutica
Praça Marechal Eduardo Gomes, 50 - Vila das Acácias, São José dos Campos - SP, 12228-900.
Universidade Federal de São Paulo
Avenida Cesare Mansueto Giulio Lattes, 1201 - Jardim Coqueiro, São José dos Campos - SP,
12247-014
oscarmachado2014@gmail.com

ANDRÉ LUIZ GOULART MATOS

Academia Militar das Agulhas Negras
Rod. Presidente Dutra Km 306, Paraíso, Resende-RJ, 27534-970.
aluizgoulartmatos@gmail.com

MAURI APARECIDO DE OLIVEIRA

Instituto Tecnológico de Aeronáutica
Praça Marechal Eduardo Gomes, 50 - Vila das Acácias, São José dos Campos - SP, 12228-900.
mauri@ita.br

ABSTRACT

This paper aims to propose a “*framework*” to calculate the ideal turnover quantity of repairable supplies, knowing the number of items removed from a given set of equipment and the number of items that return from the repair center in serviceable conditions. The proposed “*framework*” was developed using the R-Studio platform. To test the effectiveness of this proposal, data from the literature were utilized, leading to the observation that there was no prevalence in the best choice of sampling distribution, although the Negative Binomial distribution proved to be more suitable in 5 out of the 15 cases analyzed. Finally, it was concluded that the present “*framework*” has a simple implementation, requiring only minimal user data collection and very few financial resources, since R-Studio is a free software.

KEYWORDS. Repairable Supplies, “*framework*”, R-Studio.

Paper topics: Logistics and Transportation, Statistics and Probabilistic Models.

1. Introduction

In today's business environment, managers are increasingly concerned with saving resources associated with the maintenance and storage of spare parts to ensure the continuity of company operations. In this context, the efficient management of the repairable supply chain takes on a prominent role, since such items have high acquisition, storage, transportation, and maintenance costs, demanding greater attention from managers to these materials.

Repairable supplies are an efficient way to ensure the operational capability of the company. For example, suppose a helicopter air taxi company has a fleet of aircraft. It is known that various items on the aircraft undergo periodic inspections and unplanned maintenance (failures). If an aircraft engine experiences a failure or requires intervention, it can be removed to a repair center. To avoid leaving the aircraft out of operation, resulting in lost profits, a spare engine would be needed to quickly get the aircraft back in operation. As the company has a fleet, it will likely need more than one spare engine. Thus, the question arises: how many engines should we have in circulation to ensure the desired operational capability?

Calculating the number of items that should be in circulation becomes even more complex when we understand that we are in a VUCA world, which is an acronym for Volatility, Uncertainty, Complexity, and Ambiguity. According to [Stiehn 2002], this term was coined by the U.S. Army War College to describe the challenges of military leadership on the battlefield.

Repairable items fit perfectly into this VUCA concept. They are volatile, as they are mostly manufactured and maintained in other countries, making the supply chain subject to frequent changes due to strategic and customs policies. They present uncertainty because demand is highly variable, as is their repair time, which are the two main parameters needed to calculate the number of items in circulation. They are complex, requiring high skills for manufacturing, maintenance, and storage. Finally, they exhibit a certain ambiguity, as the emergence of new technologies renders items obsolete, making it difficult to obtain maintenance parts and necessitating modernization.

To address the question above and effectively deal with the adversities of the VUCA world, this work proposes a framework, which according to [Nguyen et al. 2024] is a methodological structure designed to assist modelers in decision-making. This framework encompasses the necessary phases to calculate the quantity of repairable supplies in circulation, based on the historical series of demand (items removed) and repair time (items returned from the repair center).

For this purpose, R-Studio software was used, a powerful tool for statistical analysis and data visualization, offering extensive resources for handling large volumes of data. Additionally, the tool, developed by the company Posit, has the advantage of being free, eliminating costs related to software licenses.

The stages that comprise the framework will be presented, along with a practical example and the code used at each stage. Following this presentation, a study conducted with literature data will be discussed, with observations regarding the application of the framework. A brief conclusion will highlight that this framework, being entirely developed in R-Studio, incurs no costs for the company, making it a strong alternative to paid statistical packages available in the market.

It is important to note that although repairable supplies are widely used in the aviation sector, they can also be found in other areas, such as maritime and land fleet management, military equipment, medical-hospital equipment, and even in radio communication industries and various industrial machines in general.

2. Literature Review

Repairable items began to gain prominence in the academic literature in the late 1960s with the works of [Schrady 1967] and [Sherbrooke 1968]. In fact, [Teunter 2004] states that Schrady was the first to study the subject in 1967. While the former developed a deterministic model, the latter used a probabilistic model based on the premise that the failure of repairable items followed a Poisson distribution.

In the following decade, the work of [Nahamias and Rivera 1979] emerged, refining Schrady's deterministic model. In the 1990s, the works of [Mabini et al. 1992] and [Richter 1996] appeared. The former extended Schrady's original model to multiple items, and the latter developed inventory control policies using the unit purchase lot, where repairable items are acquired in batches of only one unit due to their high cost.

Since the 2000s, there has been an exponential growth in works in this area. [Koh et al. 2002] and [Teunter 2004] expanded the models of Schrady and Nahamias and Rivera. Noteworthy among others are the works of [Wong et al. 2005], [Konstantaras et al. 2010], [Basten et al. 2012], [Tracht et al. 2013] and [Arts 2017], which present probabilistic models instead of deterministic ones.

Some of these works were published in the European Journal of Operational Research, gaining worldwide recognition, and highlighting to the scientific community the need to use stochastic methods to effectively manage a repairable supply chain.

In Brazil, the subject is very little studied. Nationally, the works of [Igawa 2006], [Machado 2010], [Machado and Ferreira 2010], and [Machado 2017] are noteworthy, the latter being a book exclusively dedicated to the topic written in Portuguese.

3. Method

To efficiently address the uncertainties of demand and repair time of materials, while simultaneously reducing the aggregated costs of the supply chain, the following framework is proposed to calculate the optimal circulation quantity. For this, we will use R-Studio software, which uses the R language, as a reference. The proposed framework is developed in five phases described below:

3.1 Phase 1 – Data Collection

This phase is crucial for the success of the model development, as the data used must be reliable and have the property of being constantly audited. Essentially, the historical series from the organization's database is used. Data on the monthly demand, which is characterized by the number of items removed from a system and sent to a repair center for a specific repairable item should be collected, as well as the number of items that returned from repair during the same month and the number of new items that entered the inventory and consequently the repair cycle. The time horizon to be considered must be at least 20 months.

A time horizon of 20 months was considered because, in statistical terms, the sample size impacts the reliability of predictions. A sample size of more than 20 observations, as noted in general Statistics textbooks, such as [Devore, 2011], [Cochran, 1977] and [Montgomery and Runger, 2014], is more suitable for the study at hand and provides greater reliability to it.

To ensure better data reliability, the data should be audited according to the documentation relevant to the item's entry into inventory, such as invoices, movement guides, etc. This auditing should be conducted regularly, at intervals defined by organizational policy.

3.2 Phase 2 – Combining Demand and Repaired Item Data into a Single Sample Distribution

In this phase, the two sample distributions mentioned in the previous phase are combined using formula (1) to form a third sample distribution of data, called the shortage, as described in [Machado and Ferreira, 2010]:

$$f_n = f_{n-1} + D_n - Tr_n, \text{ and for } n=1 \quad f_1 = D_1 - Tr_1, \quad (1)$$

where:

f_n : Number of items' shortage in n-th month.

D_n : Demand in the n-th month.

Tr_n : Number of items that return in n-th month.

It is important to note that this third sample distribution, which represents the shortage variable, depends only on the value of this variable in the previous month, the demand for the item in the current month, and the number of items returning from repair in the month.

As a premise of the model, it is considered that if an item is condemned, a new item is immediately acquired to replace the discarded component. Therefore, in the quantity (Tr_n), the new items delivered as replacements for scrapped items must also be included.

3.3 Phase 3 – Determine the Best Sample Distribution for the Shortage Variable Data

In this phase, the "fitdistrplus", "logspline", "nortest", and "DescTools" packages in R-Studio were used to perform the steps described here. The tools in the "fitdistrplus" package are detailed in [Delignette-Muller and Dutran 2018]. For the "DescTools" package, it is suggested to consult the package documentation available in [Signorell et al. 2024]. This phase is necessary to find the sample distribution that best fits the data obtained in the previous phase. To do so, the following steps are followed:

3.3.1 – Create the Histogram with Density Line for the Data

In this step, a histogram with density lines is created to visually identify a sample distribution that may fit. For example, if the density line has a "bell" shape, the normal distribution may be a good fit. However, this step is quite subjective, making it necessary to perform the following steps. An example of R code that performs this step and its respective output is:

```
item1 <- c(0,0,1,1,1,1,1,1,2,3,4,4,4,4,5,6,4,6,7,11,11,10,13,13)
hist(item1,main="Item SC7282",prob=TRUE)
lines(density(item1))
```

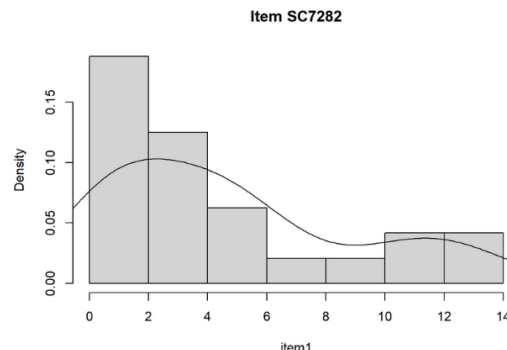


Fig 1. Example of a histogram with the density line. Note that the histogram, in this case, does not provide a good indication of which sample distribution to use. Source: Authors.

3.3.2 – Plot the Cullen and Frey Graph

The Cullen and Frey graph is a two-dimensional plot, where the x-axis represents the square of skewness, and the y-axis represents the kurtosis of the presented data. The blue point on the graph corresponds to the observed data, and it is sufficient to observe which sample distributions are closest to this point. Its practical application was first presented by [Cullen and Frey, 1999].

It is possible to plot the graph for both continuous distributions (Normal, Gamma, Weibull, logistic, etc.) and discrete distributions (Poisson and Negative Binomial). Figures 2 and 3 show examples of the Cullen and Frey graph for continuous and discrete distributions, along with the code that generates such a graph.

```
library(fitdistrplus)
library(logspline)
```

```
descdist(item1,discrete=FALSE)
# Produces the Cullen and Frey graph for continuous distributions. If
discrete=TRUE, such a plot will be created for discrete distributions.
```

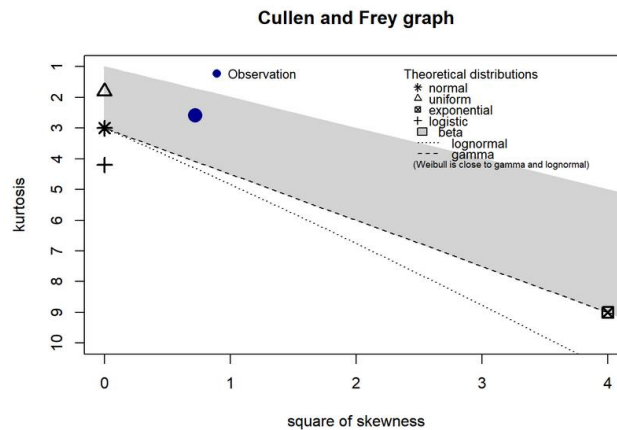


Fig 2. Cullen and Frey Graph for continuous distributions. Source: Authors.

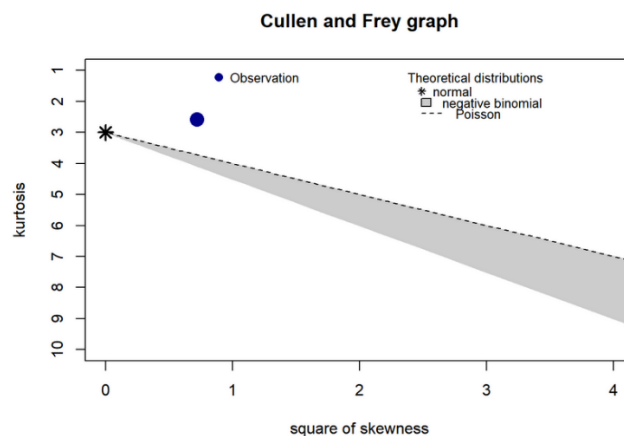


Fig 3. Cullen and Frey graph for discrete distributions. Source: Authors.

Note that in Fig 2, in the gray area, the distribution approaches a beta distribution. On the dashed line, a gamma distribution, and on the dotted line, a lognormal distribution. In the triangle, there is a uniform distribution, at the asterisk, a normal distribution, at the cross a logistic distribution, and at the square with the cross inside, an exponential distribution. The Weibull distribution corresponds to the region between the gamma and lognormal lines.

In Fig 3, in the gray area, the distribution approximates a negative binomial distribution, and on the dashed line, a Poisson distribution.

Constructing the Cullen and Frey graph is important for a pre-selection of distributions to be studied. Analyzing both graphs, it is observed that the most suitable continuous distributions in this example would be the Weibull, Normal, Gamma, and Lognormal. For discrete distributions, both Poisson and negative binomial are good fitting options. Note that, although the observed data point is in the gray area, indicating an excellent fit for the beta distribution, it is not recommended because in this type of distribution, the variable's value ranges between 0 and 1, which is not the case for the observed values.

3.3.3 – Plot the Fitted Curve with GGPlot

For the distributions selected in the previous step, curve fitting is performed using the "fitdist" function and GGPlot in R. In the example in question, GGPlot did not provide a

satisfactory fit for the Weibull, gamma, and lognormal distributions. Therefore, only the normal, Poisson, and negative binomial distributions remain, whose code is presented below, along with their respective outputs.

```
fit.normal <- fitdist(item1, "norm")
fit.poisson <- fitdist(item1, "pois")
fit.binomialneg <- fitdist(item1, "nbinom")
plot(fit.normal)
plot(fit.poisson)
plot(fit.binomialneg)
```

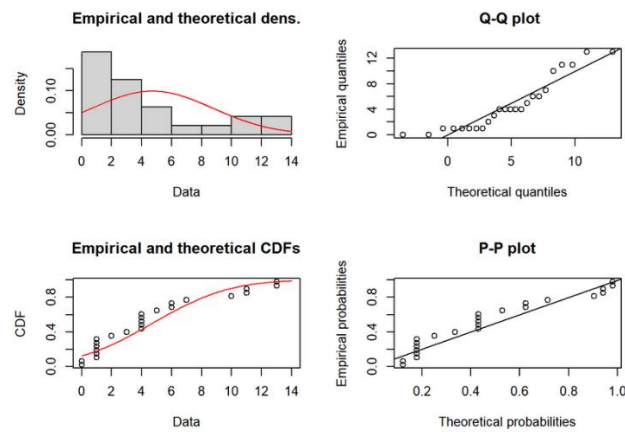


Fig 4. The fit of data for the Normal curve. Source: Authors.

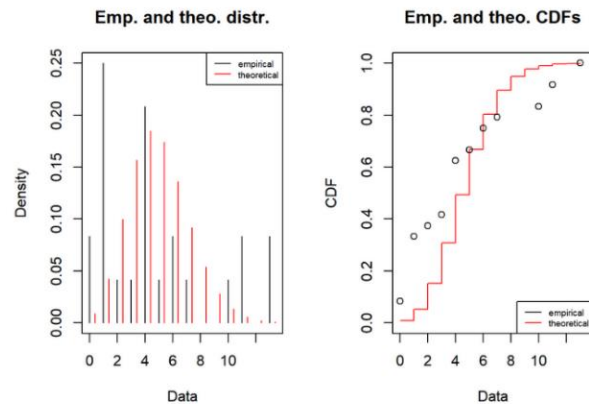


Fig 5. The fit of data for the Poisson's curve. Source: Authors.

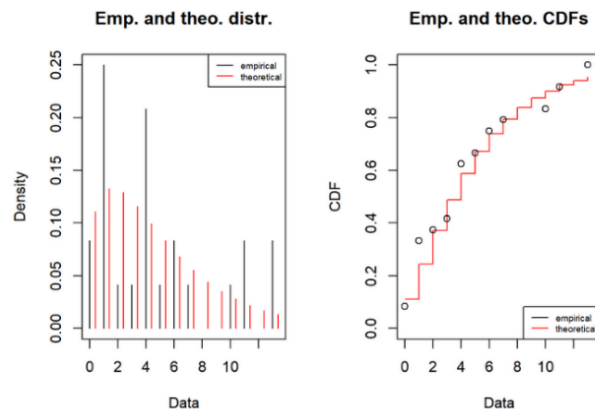


Fig 6. The fit of data for a Binomial Negative curve. Source: Authors.

In the example, by analyzing the plots produced by GGPlot, the observed data are closer to a negative binomial distribution.

3.3.4 – Determine the Akaike Information Criteria (AIC).

Akaike Information Criteria (AIC) are indicators introduced by the Japanese statistician Hirotugu Akaike (1927-2009), which serve as an alternative to non-parametric hypothesis tests. [Akaike 1974] mentions that "the practical utility of hypothesis testing in model construction should be considered quite limited.

The procedure based on AIC provides a versatile model identification procedure." For this reason, it was chosen to introduce AIC in the present framework.

Furthermore, according to [Akaike 1974], the model with the lowest AIC value is the most suitable for the problem. Therefore, the sample distribution with the lowest AIC is selected. To do so, the following code is used in R-Studio, which provided the following results for the cited example:

```
fit.normal$aic
[1] 139.2046
fit.poisson$aic
[1] 157.9699
fit.binomialneg$aic
[1] 129.7104
```

3.3.5 – Create the Fitting Model for the Chosen Curve.

In this step, we aim to find the parameters that best fit the chosen sample distribution. To do so, we use the code below to fit to a negative binomial distribution, which had the lowest AIC presented. As observed from the generated output, the best fit will be a negative binomial distribution with the parameters mean (μ) equal to 4.708716 and size (size) equal to 1.608986.

```
mle_item1 <- fitdist(item1,"nbinom",method="mle")
summary(mle_item1)
```

The output given by R-Studio for this code is:

```
Fitting of the distribution 'nbinom' by maximum likelihood
Parameters :
estimate Std. Error
size 1.608986 0.6545419
mu 4.708716 0.8777681
Loglikelihood: -62.85519 AIC: 129.7104 BIC: 132.0665
Correlation matrix:
size      mu
size 1.0000000000 -0.0001320458
mu -0.0001320458 1.0000000000
```

3.3.6 – Perform Hypothesis Testing for the Chosen Curve.

At this point, the most common hypothesis tests in the literature should be conducted, namely the Kolmogorov-Smirnov and Anderson-Darling tests. If the chosen distribution is normal, Shapiro-Wilk and Lilliefors tests for normality can be performed to provide further support for the chosen distribution. It is important to note that the Kolmogorov-Smirnov test should not be used when the data have many ties, i.e., if they repeat throughout the series. In the example at hand, if we perform the mentioned test, we will encounter an error.

It is important to emphasize that this step is more of a practical confirmation of the adequacy of the chosen sample distribution to the observed data, since AIC serves as an alternative to the aforementioned non-parametric hypothesis tests. The code below loads the "DescTools"

library, which contains the "AndersonDarlingTest" function, used to perform the Anderson-Darling test for any sample distribution.

```
library(DescTools)
AndersonDarlingTest(item1, null = "pnbinom", size=1.608986, mu=4.708716, log=FALSE)
```

The output given by R-Studio for the code above is:

```
Anderson-Darling test of goodness-of-fit
Null hypothesis: negative binomial distribution
with parameters size = 1.609, mu = 4.709
data: item1
An = 0.79555, p-value = 0.4828
```

As the p-value provided for the test was 0.4828, a value greater than 0.05, i.e., 5%, it can be said that the null hypothesis cannot be rejected, indicating that the data subjected to the test are from a negative binomial distribution, with a significance level of 5%.

With the completion of the Hypothesis Testing indicating the non-rejection of the hypothesis that the data fits the chosen sample distribution (null hypothesis), the 3rd phase is concluded. If the null hypothesis is rejected, it is necessary to return to the determination of the Akaike criteria and choose the second distribution with the lowest AIC, and so on.

3.4 – 4th Phase – Determine the Ideal Probability of Shortage for the Model.

In this phase, a model must be created to minimize the costs involved in the Supply Chain, the most important of which, according to [Taha 2008], are the order cost, inventory holding cost, and shortage cost. By solving such a model, a closed formula is obtained to determine the ideal shortage probability (P) to minimize the system cost.

At this point, the model proposed by [Machado 2017] (eq.2) can be used, which provides the ideal shortage percentage as a function of the following parameters:

$$P = \frac{hy}{s(1-r)d+hy}, \text{ where} \quad (2)$$

h: Holding cost expressed in units/month.

y: Size of ideal purchase lot.

s: Shortage cost expressed in units/month.

r: Repair rate, where $0 \leq r \leq 1$;

d: Item demand expressed in units/month.

[Machado 2017] also mentions that such a model can be simplified if we realize that repairable items, due to their high cost and manufacturing difficulty, make it impractical to acquire them in larger batches than the single-item purchase lot, in which case $y = 1$. This conclusion is also found in other works such as [Sherbrooke, 1968] and [Richter, 1996]. Therefore, it is suggested, without significant loss, to use the model below (eq. 3) to determine the ideal shortage probability (P):

$$P = \frac{h}{s(1-r)d+h}. \quad (3)$$

Applying the formula above, we can determine the ideal probability of shortage, finishing this current phase.

3.5 – 5th Phase – Determine the Ideal Stock Quantity to Maintain in Rotation to Meet the Desired Shortage Percentage.

At this point, the value related to the ideal shortage probability is determined according to the chosen sample distribution in the third phase. To illustrate, we can follow the example used thus far, where the chosen distribution was a negative binomial, with mean (μ) equal to 4.708716 and size equal to 1.608986. Let's also suppose that the ideal shortage probability obtained with the application of the model from the previous phase was about 15%. In this case, we choose to determine the cumulative distribution for 85% ($1 - P$, where P is the ideal shortage probability). For the mentioned example, we used the following code in R-Studio:

```
ceiling(qnbinom(0.85, size = 1.608986, mu = 4.708716, lower.tail = TRUE))
```

After this command, we obtain the response that the ideal quantity of items in turnover will be 9 (nine) items for the example in question, which completes the proposed framework. It should be noted that each distribution has a specific function in R-Studio. For a Normal distribution, one can use "qnorm", for Poisson "qpois", and so on. The "ceiling" function was used here because the rotation quantity must be an integer value.

4. Results and Discussions

To provide a comparison of the framework proposed here developed in R-Studio with well-established software in the statistical field, the data contained in [Machado and Ferreira, 2010] was used, which used the EasyFit software to perform hypothesis tests and choose the sample distribution using four distributions (Normal, Weibull, Poisson, and Gamma). For this purpose, the shortage probability considered in that work was kept at 5%. The results were obtained in Table 1.

From the analysis of Table 1, it can be concluded that in 9 out of the 15 items analyzed, there was an increase in the rotation quantity obtained. This is because [Machado and Ferreira, 2010] used, when the data adhered to more than one distribution, the distribution that provided the lowest rotation quantity. In the present framework, the Akaike Information Criterion (AIC) was used as the criterion for choosing the distribution. Therefore, the chosen distribution was not the one that provided the lowest rotation quantity, but the one with the lowest AIC.

Item Description	Machado and Ferreira (2010)		Observed in the framework	
	Turnover	Distribution	Turnover	Distribution
1 RT ARC 182	27	Normal	30	Binomial Negative
2 ANTENNA MECHANISM	19	Normal	19	Normal
3 RADARSCOPE	18	Normal	19	Weibull
4 TORQUE INDICATOR HM-1	20	Normal	22	Weibull
5 TORQUE INDICATOR HA-1	19	Normal	23	Binomial Negative
6 DELTA NG INDICATOR	25	Gamma	25	Gamma
7 TORQUE TRANSMITTER	26	Normal	29	Binomial Negative
8 T4 INDICATOR	4	Weibull	5	Gamma
9 ENGINE VIBRATION SET	9	Poisson	11	Weibull
10 STARTER GENERATOR	30	Normal	30	Normal
11 LEFT SERVO COMMAND	7	Normal	6	Poisson
12 CDU FDR	12	Poisson	13	Normal
13 MDAU FDR	10	Normal	10	Normal
14 REAR SERVO COMMAND	12	Normal	13	Binomial Negative
15 RIGHT SERVO COMMAND	7	Normal	7	Binomial Negative

Table 1. Comparison between [Machado and Ferreira 2010] and the framework. Source: Authors

Indeed, there are no reasons to believe that the best curve fit depends on the outcome, but rather on a widely used criterion in Statistics, such as the Akaike Information Criterion (AIC). For

this reason, the present framework provides greater accuracy in calculating the rotation of repairable supplies, despite yielding a higher quantity of items.

Regarding the distribution that best fits the cases in question, it is noteworthy that no distribution prevailed over another, differing from the work of [Machado and Ferreira 2010], where the Normal distribution showed a certain prevalence among others. In the proposed framework, the Negative Binomial distribution appeared in 5 out of the 15 items studied, while the Normal distribution appeared in 4, the Weibull in 3, the Gamma in 2 of the studied items, and finally, the Poisson distribution appeared in only 1 item. The Lognormal and exponential distributions were not efficient in shaping the problem.

Regarding the calculation of the AIC, the following values were observed for the different types of sample distribution:

	Distribution								
	Uniform	Normal	Lognormal	Weibull	Gamma	Bin. Neg.	Poisson	Min.AIC	Choose
1	154.504	170.704	NA	NA	NA	170.290	223.720	154.504	Bin. Neg.
2	145.333	150.147	NA	NA	NA	153.860	172.136	145.333	Normal
3	137.084	149.227	151.492	146.706	147.879	147.548	164.202	137.084	Weibull
4	139.994	155.576	161.638	154.357	156.305	155.954	186.015	139.994	Weibull
5	152.370	162.823	NA	NA	NA	147.686	247.858	147.686	Bin. Neg.
6	154.504	162.823	158.248	158.082	157.416	157.885	188.474	154.504	Gamma
7	160.389	172.196	NA	NA	NA	168.283	242.533	160.389	Bin. Neg.
8	81.253	66.450	62.755	67.402	62.566	78.895	76.895	62.566	Gamma
9	103.813	120.193	120.249	117.946	118.819	119.221	118.614	103.813	Weibull
10	160.389	168.429	185.035	171.028	176.382	174.246	207.319	160.389	Normal
11	97.404	102.448	NA	NA	NA	100.183	98.478	97.404	Poisson
12	114.524	120.726	135.109	121.991	128.112	123.096	121.096	114.524	Normal
13	103.813	122.628	NA	NA	NA	124.790	133.348	103.813	Normal
14	127.118	139.205	NA	NA	NA	129.710	157.970	127.118	Bin. Neg.
15	97.404	107.974	NA	NA	NA	104.257	106.472	97.404	Bin. Neg.

Table 2. AIC values obtained for sample distributions. Source: Authors.

NA: indicates that the data did not conform to the distribution (GGPlot failed to fit).

From the analysis of Table 2, it can be observed that the continuous uniform distribution obtained the lowest AIC in 13 out of the 15 items studied; however, in all cases, the conducted hypothesis test (Anderson-Darling) provided a very low p-value, less than 0.0001, leading to the rejection of the null hypothesis, where the data would come from a continuous uniform distribution. Therefore, it was chosen to model the data output by the distribution with the second lowest AIC.

One notable occurrence in 8 out of the 15 studied items was that GGPlot failed to fit the Gamma, Weibull, and Lognormal distributions. In the other 7 items, GGPlot managed to perform the fittings, with the Weibull distribution being chosen in three of these cases and the Gamma in two. This fact highlights that data analysis in GGPlot helps filter distributions for the next phase with relative effectiveness.

Another noteworthy observation concerns the data collected in item 8. When fitting by the GGPlot graph, it was observed that none of the sample distributions fitted satisfactorily, a fact ratified by the hypothesis test. All null hypotheses at a significance level of 5% were rejected, with the highest observed p-value being 0.04848 for the Gamma distribution, very close to 5% (0.05). At this point, the significance level was reduced to 1%, and the Gamma distribution was used to model the problem. If desired, we could choose to create a specific sample distribution for this item.

To ensure the reproducibility of the experiment described in this work, the files used in qmd format and their respective outputs in .html format were made available at the following link:

https://drive.google.com/drive/folders/1DbnwD9u36-5_NKg437eUr9As3dnPT6yv?hl=pt-BR

5. Conclusion

Based on the above, it can be observed that as there is no prevalence of one sample distribution over another, there arises the need to always identify the sample distribution that best fits the problem at hand.

The incorporation of the Cullen and Frey graph into the framework proved to be quite efficient, as it provides an indication, albeit quite incipient, about the most likely distributions to be suitable for the data. The creation of fitting models in GGPlot with the "fitdistrplus" package also proved to be quite efficient, as it allows for a graphical analysis of the best-fitting distribution. In this phase, in some datasets, the Gamma, Weibull, and Lognormal distributions were discarded.

The Akaike Information Criterion (AIC) proves to be quite efficient in breaking the tie between sample distributions that have not yet been discarded. At this point, the uniform distribution, although it has the lowest AIC, did not fit the data and was rejected in the subsequent hypothesis test.

The hypothesis test cannot be dispensed with, as in the case of the uniform distribution, which has a low AIC but was rejected in the hypothesis test. For other types of distributions, the Anderson-Darling test was only a confirmation of the results obtained in the AIC determination step.

Regarding the proposed framework, the feasibility of implementing it in a free software like R-Studio is evident. Such implementation is relatively simple since the source code is available in the provided files, and R-Studio offers the possibility of extracting data from spreadsheets, making it an excellent option for implementing the method described here. Another great advantage of using R-Studio is its high reproducibility potential, i.e., the ease of redoing the steps followed and confirming the results obtained, as well as expanding the research to other datasets.

As opportunities for improvement and future studies, it is suggested to study more thoroughly the fact that the Anderson-Darling test rejects the continuous uniform distribution even though it has the lowest AIC. Another possibility for improvement would be to automate the proposed framework so that the user enters the initial data from the 1st phase and after a click, receives the proposal of items to compose the turnover quantity. To do so, the interpretation of the Cullen and Frey graph and the graphs of the "fitdistrplus" package should be automated.

In conclusion, it can be inferred that the present framework is an efficient way to organize data to obtain the ideal turnover quantity, given the demands and the items returning from repair in the month, requiring few data from users and consuming few financial resources, presenting great applicability in the business world.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6): 716–723.
- Arts, J. (2017). A multi-item approach to repairable stocking and expediting in a fluctuating demand environment. *European Journal of Operational Research*, 256(1): 102–115.
- Basten, R.J.L. Van der Heijden, M.C. & Schutten, J.M.J. (2012). Joint optimization of level of repair analysis and spare parts stocks, *European Journal of Operational Research*, vol. 222(3):474–483.
- Cochran, W. G. (1977). *Sampling Techniques* (3rd Edition). John Wiley & Sons.
- Cullen, A. & Frey, H. (1999). Probabilistic Techniques in Exposure Assessment. Plenum Publishing Co, 1st edition.
- Delignette-Muller, M.L. & Dutran, C. (2018). fitdistrplus: An R Package for Fitting Distributions.
- Devore, J. L. (2011). *Probability and Statistics for Engineering and the Sciences* (8th Edition). Cengage Learning.

Igawa, M. (2006). Otimização do Suprimento de Itens Reparáveis para Frotas de Aeronaves. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal Fluminense, Niterói.

Koh, S.G., Hwang, H., Sohn, K.I. & Ko, C.S. (2002). An optimal ordering and recovery policy for reusable items. *Computers and Industrial Engineering*, 43: 59-73.

Konstantaras, I., Skouri, K. & Jaber, M. (2010). Lot sizing for a recoverable product with inspection and sorting. *Computers and Industrial Engineering*, 58:452-462.

Mabini, M.C., Pintelon, L.M. & Gelders, L.F. (1992). EOQ type formulations for controlling repairable inventories. *International Journal of Production Economics*, 28:21-33.

Machado, O. A. (2010). O Cálculo da Quantidade de Giro de Suprimentos Reparáveis para Frotas de Asa Rotativa. Monografia. Universidade Estadual Paulista, Guaratinguetá.

Machado, O. A. (2017). Gerenciamento da Cadeia de Suprimentos Reparáveis: Um enfoque prático. Ed Cia do Livro, Timburi-SP.

Machado, O. A & Ferreira, U. R. (2010). O Cálculo da Quantidade de Giro de Suprimentos Reparáveis para Frotas de Asa Rotativa. *Simpósio de Pesquisa Operacional e Logística da Marinha*.

Montgomery, D. C., & Runger, G. C. (2014). *Applied Statistics and Probability for Engineers* (6th Edition). John Wiley & Sons.

Nahmias, N. & Rivera, H. (1979). A deterministic model for a repairable item inventory system with a finite repair rate. *International Journal of Production Research*, 17(3), 215–221.

Nguyen, L. K. N, Howick, S. & Megiddo, I. (2024). A framework for conceptualizing hybrid system dynamics and agent-based simulation models. *European Journal of Operational Research*, 315: 1153-1166.

Richter, K. (1996). The EOQ repair and waste disposal model with variable setup numbers. *European Journal of Operational Research*, 95:313-324.

Sherbrooke, C. C. (1968). Optimal Inventory for Modelling of Systems: Multi-Echelon Techniques. Boston: Kluwers Academic Publishers.

Signorell A. et al. (2024). Package ‘DescTools’ - Reference Manual. Disponível em: <https://cran.r-project.org/web/packages/DescTools/DescTools.pdf> acesso em 12 mar 24 00:38:00.

Stiehm, J. H. (2002). The U.S. Army War College: Military education in a democracy. Philadelphia, PA: Temple University Press.

Taha, H.A. (2008). Operations Research: an introduction. 8th edition. Pearson Prentice Hall.

Teunter, R. H. (2004). Lot-sizing for inventory systems with product recovery. *Computers and Industrial Engineering*, 46:431-441.

Tracht, K. Von der Hagen, F. & Schneider, D. (2013). Applied repairable-item inventory modeling in the aviation industry. *Procedia Cirp*, 11:334-339.

Wong H., Cattrysse, D. & Oudheusden, D. V. (2005). Inventory pooling of repairable spare parts with non-zero lateral transshipment time and delayed lateral transshipments. *European Journal of Operational Research*. 165:207–218.