v. 4, n. 23, 2024

Journal of
Engineering
Research

Acceptance date: 27/09/2024

DETERMINATION OF TIME DISTRIBUTION USING NONPARAMETRIC STATISTICS

Leticia del Pilar de la Torre González
Professor at the Tecnológico Nacional de México/ I.T. Chihuahua, México

Rita Luna Gándara
Professor at the Tecnológico Nacional de México/ I.T. Chihuahua, Mexico

Maríana Elizabeth Carreon Quiñonez
Student At Tecnológico Nacional de México/ I.T. Chihuahua, Mexico

Johana Anahí Cuevas Duarte
Student At Tecnológico Nacional de México/ I.T. Chihuahua, Mexico

Francisco Alfredo Marrufo Carmona
Student At Tecnológico Nacional de México/ I.T. Chihuahua, Mexico

INTRODUCTION

In the present investigation on the distribution of the time that a person takes at an ATM to withdraw cash, several distributions were observed, such as the Normal Distribution, the Weibull Distribution, the Exponential Distribution, among others. In these different distributions, tests for the determination of the distribution such as Anderson Darling, Kolmogorov Smirnov and Chi-square were applied. In these different distributions covered in the work, a great variety of results and numbers can be observed between one and the other, as for example in the Chi-Square test that accepted 2 of 3 distributions, but in the Anderson Darling test it rejected the 3 distributions due to the different characteristics between each of the tests, since some tests are more sensitive than others in certain aspects and should be used depending on their reliability. It can be seen that throughout the process Excel was used as the main tool, as well as Minitab software as an alternative tool. The Box-Cox method is used to normalize the data, which can be used in parametric statistics for decision making. The time taken was 35 data and was in an ATM to withdraw cash and will allow to appreciate the distribution of the same, performing the appropriate statistical tests to determine its parameters.

THEORETICAL FRAMEWORK

To better understand this topic, it is necessary to know more about the different distributions, the different types of tests, as well as the different methods that are used when it comes to

PROBABILITY DISTRIBUTION

A probability distribution is one that makes it possible to establish the full range of outcomes likely to occur in a given experiment. That is, it describes the probability that an event will occur in the future.

Any probability distribution is generated by a random (because it can take different values) variable x (because the value it takes is completely random).

NORMAL DISTRIBUTION

Minitab (2021) defines the following concept of Normal Distribution:

> "The normal distribution is a continuous distribution that is specified by the mean (μ) and standard deviation (σ). The mean is the peak or center of the bell-shaped curve. The standard deviation determines the dispersion of the distribution."

This is a very important concept within the methods for determining the type of distribution since we are talking about one of the most important and most common distributions, since many statistical analyses assume that the data come from normally distributed populations.

Importance of the Normal Distribution

J. Lejarza & I. Lejarza (2007) tell us that:

> "On the one hand, a large number of real phenomena can be modeled with this distribution (such is the case of the quantitative characteristics of almost all large populations). On the other hand, many of the frequently used distributions tend to approximate the normal distribution under certain conditions; and, finally, by virtue of the Central Limit Theorem, all those variables that can be considered to be caused by a large number of small effects (such as may be observation errors) tend to be distributed with a normal distribution."

As we mentioned before, the Normal distribution is the most common distribution and therefore it is also the most important and the one we encounter most in everyday life.

Characteristics of the Normal Distribution

The Society of Anesthesiology of Chile (2017) proposes the following characteristics of the Normal Distribution:

"1. It is bell-shaped.

2. It is symmetrical.

3. It reaches its maximum at μ (the mean).

4. The mean is also the mode and median.

5. It is asymptotic to the abscissa axis and, since it never touches it, any value of X between-infinity and +infinity is theoretically possible.

6. The relative position on the abscissa axis is determined by μ (more to the right the greater it is) and its greater or lesser flattening or width is determined by σ (the standard deviation), being more flattened the greater its magnitude. This characteristic is called kurtosis (from the Greek, curved): narrow or leptokurtic (literally, narrow curve), medium or mesokurtic and widened or platykurtic (literally, wide curve). The height of the curve is of no practical importance or use."

These characteristics are easy to digest and to observe within the Normal Distribution since we are talking about a symmetric distribution where the mean is the peak of the bell and the standard deviation is the one found in the dispersion of the bell.

WEIBULL DISTRIBUTION

Minitab (2022) defines the concept of Weibull distribution:

> "The Weibull distribution is a versatile distribution that can be used to model a wide range of applications in engineering, medical research, quality control, finance and climatology. For example, the distribution is frequently used with reliability analysis to model time-before-failure data. The Weibull distribution is also used to model asymmetric process data in capacity analysis."

This distribution is very useful because it is very versatile and allows us to apply it in several areas, in order to model and interpret the data before the failure time, also to verify if the data are symmetrical and if not, to see what the problem is and thus control it.

Importance of the Weibull Distribution

Ministerio de Trabajo y Asuntos Sociales España in collaboration with the Centro Nacional de Condiciones de Trabajo (1994) define the significance of the Weibull distribution as:

> "The Weibull distribution allows us to study what is the distribution of failures of a key safety component that we intend to monitor and that through our failure log we observe that failures vary over time and within what is considered normal usage time."

The Weibull distribution is very important because it allows us to study the failures in our times and thus control them and have a normal distribution.

Characteristics of the Weibull Distribution

Anly Quiroga and Paula Andrea Riano through the Pubs page (2016).

> "The Weibull distribution function is a statistical model that represents the probability of failure after a time F(t) as a function of elapsed time, or of an analogous variable, or it can be interpreted to mean that R(t) is the probability that the components of an ensemble survive to time t (reliability)."

As we could observe this distribution is very helpful to see where is the fault in our times, as well as to see the probabilities of our function.

EXPONENTIAL DISTRIBUTION

Minitab (2021) defines the exponential distribution as:

> "The exponential distribution has a great practical utility since we can consider it as a suitable model for the probability distribution of the waiting time between two events following a Poisson process. In fact, the exponential distribution can be derived from an experimental Poisson process with the same characteristics as those we stated when studying the Poisson distribution, but taking as a random variable, in this case, the time it takes to produce a done"

We believe that this concept explains in a very clear way what this distribution is and we see that it is very helpful when we want to calculate a probability of time, it tells us that it is very useful because we can consider it suitable for the probability of times following a Poisson process.

Importance of the Exponential Distribution

100functions (2018) says that the significance of the exponential distribution is:

> "Thanks to the knowledge of the format of exponential functions, it is easier to graph equations or certain formulas. It is possible to set up a table of values to find the graphing curve with a higher degree of accuracy."

We can see that it is important because thanks to this it is easier to plot equations or certain formulas, as well as to change values to find the curve more accurately.

Characteristics of the Exponential Distribution

Permatics (2014) mentions the following characteristics of the exponential distribution:

- "The domain of an exponential function is R.
- They are continuous functions.
- If a > 1 the function is increasing.
- If 0 < a < 1 the function is decreasing.
- They are always concave.
- The x-axis is a horizontal asymptote."

We can see that these characteristics are very easy to understand since it tells us what criteria must be met to be within this distribution, whether it is continuous or the x-axis is a horizontal asymptote.

CHI-SQUARE TEST

Minitab (2022) defines Chi-Square as:

> "The chi-square distribution is a continuous distribution that is specified by the degrees of freedom and the non-centrality parameter. The distribution is positively skewed, but the skewness decreases with increasing degrees of freedom."

We can see that this concept is of great importance since the Chi-Square is a continuous distribution specified by the degrees of freedom and its parameter, it has an asymmetric distribution, but this is modified as the degrees of freedom increase.

Significance of Chi-Square

QuestionPro (2022) says the following about the importance of Chi-Square.

> "The basic idea of the test is that you compare the actual data values with what you would expect if the null hypothesis were true."

We can understand that the importance of Chi-Square is to help us compare in an efficient way the real data with what we expect to fulfill our hypotheses.

Chi-Square Characteristics

Pubs (2016) mentions the following characteristics:

> "The distribution has a single parameter, k, called the degrees of freedom of the random variable.

The continuous random variable X has a chi-square distribution, with k degrees of freedom, if its density function is given by:

$$f(x;k) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2} & \text{para } x \geq 0, \\ 0 & \text{para } x < 0 \end{cases}$$

where k is a positive integer."

These characteristics are very specific and we can see that we need a parameter, the variable continues called x, in order to apply these values in the corresponding formula.

ANDERSON DARLING

Jensen & Alexander (2016) mention that:

> "The Anderson-Darling goodness-of-fit statistic measures the area between the fitted line - based on the normal distribution - and the empirical distribution function - which is based on the data points."

The Anderson-Darling test is designed to detect deviations from normality, to confirm or disconfirm that the data are from a normal distribution.

Importance of Anderson Darling

Guisande & Barreiro (2006) tell us that:

> "The Anderson Darling statistic can be used to test whether data satisfy the assumption of normality for a t-test. It can also be defined as that nonparametric statistic that is used to test whether a set of sample data comes from a population with a specific continuous probability distribution, usually a normal distribution."

This test is important because we can use it to check if the data satisfies the normality assumption, since it is not parametric, it is used to test if it is a set of data that complies with a continuous distribution.

Characteristics of Anderson Darling

Business Process Simulation (2011) mentions the following characteristics:

> "Its purpose is to corroborate whether a sample of random variables comes from a population with a specific probability distribution.
>
> It is a modification of the kolmogorov-smirnov test.
>
> This is very sensitive at the extremes of the distribution, so it should be used with great care in distributions with a bounded lower limit, and is not reliable for discrete-type distributions."

These characteristics are easy to understand as they tell us specifically what requirements this test shows such as that it modifies the Kolmogorov-Smirnov test, and realize that it is very sensitive at the extremes of the distribution and because of this it is not reliable for discrete type distributions.

KOLMOGOROV SMIRNOV

García R.; González J. and Jornet J.M. mention that:

> "The Kolmogorov-Smirnov test for a sample is a "goodness-of-fit" procedure, which allows measuring the degree of agreement existing between the distribution of a data set and a specific theoretical distribution. "

Minitab (2022) defines the Kolmogorov - Smirnov test as:

> "A goodness-of-fit test that is used to determine whether the input variable follows a specific probability distribution.
>
> This tells us that the Kolmogorov-Smirnov test is intended to show us whether the data come from a specific distribution."

This concept tells us that this test allows us to measure the degree of agreement between a set of data and its distribution, so its objective is to show us that they come from a specific distribution.

Significance of Kolmogorov Smirnov

Psychology and Mind (2019) states that:

> "It is important to verify whether or not the scores we have obtained from the sample follow a normal distribution. That is, it allows us to measure the degree of agreement existing between the distribution of a set of data and a specific theoretical distribution."

This test is of great importance because it helps to verify whether or not the sample data follow a normal distribution, with this we can observe if there is a distribution among the data set and see if the hypotheses are true or not.

Kolmogorov Smirnov Characteristics

Psychology and Mind (2019) mentions the following characteristics:

> "The Kolmogorov-Smirnov test is a type of nonparametric test. Nonparametric (also called free distribution) tests are used in inferential statistics, and have the following characteristics:

- Hypotheses on goodness-of-fit, independence and...
- The level of measurement of the variables is low (ordinal).
- They do not have excessive restrictions.
- They are applicable to small samples.
- They are robust."

The characteristics of this test are very clear, since it is a non-parametric test with some restrictions, such as hypothesis testing, robustness and application in small samples.

BOX-COX METHOD

Statologos (2021) defines this method as:

> "A Box Cox transformation is a transformation of non-normal dependent variables into a normal form. Normality is an important assumption for many statistical techniques; if your data are not normal, applying a Box-Cox means you can run a larger number of tests."

It can be understood that the Box-Cox transformation is mainly used to normalize data and thus be able to apply the corresponding tests.

DEVELOPMENT

The time it takes for people to withdraw cash from an ATM was taken, with due respect to the user.

A random sample of 35 times was taken from the beginning of the operation until the moment the cash was withdrawn from the ATM, taking care not to lose any data.

All of this happened in a single particular teller to avoid adding variations to the times and prevent outliers.

The 35 data were taken with an IPhone X stopwatch. The data in seconds are shown below.

57.62	61.85	51.84	51.27	61.81	54.77	54.19	45.42	51.76	57.84	60.76	45.48
62.27	50.74	45.93	61.45	49.33	48.28	54.93	47.32	45.78	51.54	48.63	48.24
61.42	51.15	58.46	50.91	54.35	52.42	52.69	52.32	52.47	47.52	52.13	

Journal of Engineering Research ISSN 2764-1317 DOI https://doi.org/10.22533/at.ed.3174232418099

RESULTS

NORMAL DISTRIBUTION WITH X^2

With the following data, the X-test² will be performed to test whether the data fit a normal distribution, with a NC=95%. The data represent time in seconds.

Times:

57.62, 61.85, 51.84, 51.27, 61.81, 54.77, 54.19, 45.42, 51.76, 57.84, 60.76, 45.48,

62.27, 50.74, 45.93, 61.45, 49.33, 48.28, 54.93, 47.32, 45.78, 51.54, 48.63, 48.24,

61.42, 51.15, 58.46, 50.91, 54.35, 52.42, 52.69, 52.32, 52.47, 47.52, 52.13

As a first step, the parameters necessary for the normality test are estimated and the histogram is made.

Maximum value = 62.27

Minimum value = 45.42

Range = 62.27 - 45.42 = 16.85

k = $\sqrt{35}$ = 5.91607978 ≈ 6

h = 16.85 / 6 = 2.81

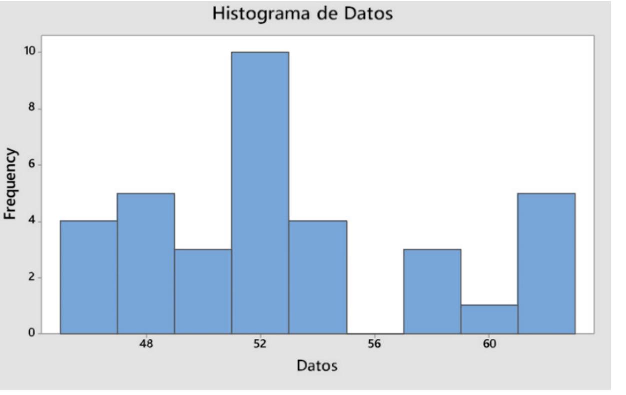
Estimated Mean = 53.27

Standard Deviation = 4.76

FREQUENCY DISTRIBUTION

LRI	Classes		LRS	Frequency
41.415	45.42	48.23	48.235	6
48.235	48.24	51.05	51.055	6
51.055	51.06	53.87	53.875	10
53.875	53.88	56.68	56.685	4
56.685	56.69	59.50	59.505	3
59.505	59.51	62.32	62.325	6
				35

histogram with MINITAB



Frequency histogram

When looking at the histogram, the data does not have a smooth curve or better known as the Gaussian bell, even so, the histogram is not good evidence of normality, so the Chi-Square test for goodness of fit is performed.

The probabilities of each Class are calculated using the Probability- Distributions application.

To calculate the expected frequencies (Ei), each probability is multiplied by the total number of data, in this case 35.

Probability	Cumulative probability	Ei	Grouped	Grouped Oi
0.145	0.145	5.075	5.075	6
0.1758	0.3208	6.153	6.153	6
0.2297	0.5505	8.0395	8.0395	10
0.2129	0.7634	7.4515	7.4515	4
0.1414	0.9048	4.949	8.281	9
0.0952	1	3.332		
1		35	35	35

The expected frequency values are grouped, because the test requires the expected values to be greater than 5, so the observed frequency values are also grouped.

Hypothesis

H0; Data fit a normal distribution with μ=53.27 and σ=04.76 H1; Data DO NOT fit a normal distribution with μ=53.27 and σ=04.76

Calculation of test statistic X^2

$$X^2 = \frac{(6-5.075)^2}{5.075} + \frac{(6-6.153)^2}{6.153} + \frac{(10-8.0395)^2}{8.0395} + \frac{(4-7.4515)^2}{7.4515} + \frac{(9-8.281)^2}{8.281} = 2.3116$$

To calculate the value of X^2 it is necessary to obtain the degrees of freedom so that with the significance level the chi-square value is obtained.

V = k-1- m

Where:

k = Number of classes

m = Number of calculated parameters.

Therefore:

V = 6-1-2 = 3

Then:

$$X^2_{(0.05,\ 3)} = 7.8147$$

Conclusion

Since 2.3116<7.8147, H0 is NOT rejected and it is concluded that the data fit a normal distribution with α= 0.05.

The data are entered into the MINITAB software in order to find the distribution to which the data best fit.

Prueba de bondad del ajuste

Distribución	AD	P	LRT P
Normal	0.783	0.038	
Transformación Box-Cox	0.466	0.238	
Lognormal	0.623	0.096	
Lognormal de 3 parámetros	0.478	*	0.274
Exponencial	13.214	<0.003	
Exponencial de 2 parámetros	1.548	0.018	0.000
Weibull	1.240	<0.010	
Weibull de 3 parámetros	0.544	0.171	0.002
Valor extremo más pequeño	1.490	<0.010	
Valor extremo por máximos	0.471	0.238	
Gamma	0.698	0.072	
Gamma de 3 parámetros	0.482	*	0.128
Logística	0.727	0.033	
Loglogística	0.604	0.077	
Loglogística de 3 parámetros	0.462	*	0.253
Transformación de Johnson	0.456	0.252	

Minitab uses the Anderson Darling test, and it can be seen that the data do not fit a normal distribution. The P value for this test is 0.038 and since it does not fit a normal distribution, MINITAB automatically transforms the values with the BOX-COX method.

Next, we have the Anderson-Darling test to conclude whether the data fit a normal distribution or not.

Normal Distribution with Anderson Darling

The variance and standard deviation are obtained from the data:

Mean = 52.9968571

Variance= 26.7005575

Standard deviation = 5.16725822

To start this test, the following table is made, where the data and the operations necessary to find the sum of the last column, which is necessary for the formula of A^2 n, are found.

C0	C1	C2	C3	C4	C5	C6	C7	C8
i	Yi	Yn+1-i	2i-1	ADP(Yi)	1-PEA(Yn+1-i)	ln C4	ln C5	C3 * (C6+C7)
1	45.42	62.27	1	0.07128048	0.036359016	-2.64113271	-3.31431308	-5.95544579
2	45.48	61.85	3	0.07287492	0.043327269	-2.61901077	-3.13897307	-17.2739515
3	45.78	61.81	5	0.08125925	0.044043683	-2.51011062	-3.12257334	-28.1634198
4	45.93	61.45	7	0.08571518	0.05092975	-2.45672535	-2.97730805	-38.0382338
5	47.32	61.42	9	0.13596676	0.051540279	-1.99534484	-2.96539166	-44.6466284
6	47.52	60.76	11	0.14459152	0.066500562	-1.93384262	-2.71054488	-51.0882626
7	48.24	58.46	13	0.17863575	0.145196143	-1.72240644	-1.92966974	-47.4769903
8	48.28	57.84	15	0.18066451	0.174308514	-1.71111349	-1.74692848	-51.8706296
9	48.63	57.62	17	0.19902711	0.185473898	-1.61431421	-1.68484112	-56.0856406
10	49.33	54.93	19	0.23896586	0.354159808	-1.43143457	-1.03800703	-46.9193904
11	50.74	54.77	21	0.33114232	0.36574294	-1.10520702	-1.00582454	-44.3316627
12	50.91	54.35	23	0.34315725	0.396711361	-1.06956648	-0.92454631	-45.8645943
13	51.15	54.19	25	0.36039046	0.408694521	-1.02056722	-0.89478729	-47.8838628

14	51.27	52.69	27	0.36911728	0.523677232	-0.99664085	-0.64687975	-44.3750562
15	51.54	52.47	29	0.38899473	0.540606056	-0.94418948	-0.61506444	-45.2183637
16	51.76	52.42	31	0.40541156	0.544444378	-0.90285254	-0.60798949	-46.8361031
17	51.84	52.32	33	0.41142448	0.552108237	-0.88812979	-0.59401117	-48.9106518
18	52.13	52.13	35	0.43338621	0.566613795	-0.83612602	-0.56807735	-49.1471178
19	52.32	51.84	37	0.44789176	0.588575517	-0.80320368	-0.53005004	-49.3303875
20	52.42	51.76	39	0.45555562	0.594588445	-0.78623746	-0.5198858	-50.9388072
21	52.47	51.54	41	0.45939394	0.611005268	-0.77784717	-0.4926497	-52.0903716
22	52.69	51.27	43	0.47632277	0.630882717	-0.74165957	-0.4606353	-51.6986796
23	54.19	51.15	45	0.59130548	0.639609538	-0.52542251	-0.44689739	-43.7543954
24	54.35	50.91	47	0.60328864	0.65684275	-0.50535952	-0.42031063	-43.5064975
25	54.77	50.74	49	0.63425706	0.668857677	-0.45530095	-0.40218398	-42.0167616
26	54.93	49.33	51	0.64584019	0.761034137	-0.43720319	-0.27307706	-36.2242928
27	57.62	48.63	53	0.8145261	0.800972886	-0.2051488	-0.22192818	-22.6350803
28	57.84	48.28	55	0.82569149	0.819335488	-0.19153408	-0.19926165	-21.493765
29	58.46	48.24	57	0.85480386	0.821364245	-0.15688324	-0.19678861	-20.1592955
30	60.76	47.52	59	0.93349944	0.855408481	-0.06881492	-0.15617617	-13.2744741
31	61.42	47.32	61	0.94845972	0.86403324	-0.05291596	-0.14614404	-12.1426597
32	61.45	45.93	63	0.94907025	0.914284821	-0.05227246	-0.08961314	-8.93879239
33	61.81	45.78	65	0.95595632	0.91874075	-0.04504306	-0.0847513	-8.43663322
34	61.85	45.48	67	0.95667273	0.927125083	-0.04429392	-0.07566679	-8.03736755
35	62.27	45.42	69	0.96364098	0.928719516	-0.03703648	-0.07394851	-7.65796378
								-1252.42223

$\sum$ (2i-1)[lnPEA(Yi) + ln(1-PEA(Yn+1-i))]= -1252.42223

Hypothesis

H0; The data conform to a normal distribution H1; The data do not conform to a normal distribution

Calculation of Anderson Darling test statistic

Now A^2 n is calculated using the result of the previous table following the formula: A^2 n = -[n + 1/n * $\sum$ (2i-1)[lnPEA(Yi) + ln(1-PEA(Yn+1-i))]]]

$$\mathbf{A^2n = -(35 + \frac{1}{35}(-1252.42223)) = 0.78349228}$$

It is necessary to adjust the test statistic to the normal

$$\boldsymbol{NORMAL = 0.78349228\,(1 + \frac{4}{35} - \frac{25}{35^2}) = 0.857044616}$$

The Anderson-Darling table is used in the normal distribution with an α= 0.05 to obtain the value of A^2 n0.05.

$$\mathbf{A^2n}_{0\cdot 005} = 0.751$$

Conclusion

As 0.85704 > 0.751 H0 is rejected and the data are said not to fit a normal distribution with μ = 52.99685 and σ = 5.167258.

The Anderson Darling test has the advantage of detecting discrepancies at the extremes of the distributions and also that in this test the critical values are calculated for each of the data. Unlike Chi-square where data are grouped and it is not so sensitive at the extremes of the distribution.

The data will be transformed to normalize them with the Box-Cox method, so that they can be used for subsequent studies using parametric statistics.

BOX-COX

This method consists of finding the most appropriate lambda value. To do so, it is necessary to evaluate each lambda from -5 to 5, using the Minitab program. The results of this evaluation yielded a λ = -2. Therefore, it was decided to only do -2 to 2 to compare the lambdas.

```
transformación de Box-Cox: λ = -2

Función de Transformación de Johnson:
-9.93833 + 3.60751 × Ln( X - 36.4154 )


Prueba de bondad del ajuste

Distribución                    AD       P    LRT P
Normal                          0.783    0.038
Transformación Box-Cox          0.466    0.238
```

To start with, it is necessary to obtain the geometric mean of the data. This is found with the following formula:

$$\dot{y} = \sqrt[n]{\prod_{i=1}^{n} yi} \quad \text{like} \quad n=35 \quad \dot{y} = \sqrt[35]{\prod_{i=1}^{35} yi} = 52.756$$

The geometric mean is equal to 52.756, so each lambda is evaluated by substituting it into the following equations:

$$y^2 = \frac{y^{-\lambda}-1}{\lambda * \dot{y}^{\lambda-1}} \text{ Si } \lambda \neq 0$$

$$y^2 = \dot{y} ln y \text{ Si } \lambda = 0$$

Having the results of each lambda ready and already selected the lambda which was λ= -2, we proceed to elevate the data to the -2 to normalize the data.

Data	-2	-1.5	-1	-0.5	0	0.5	1	1.5	2	Raised to -2
45.42	73379.54485	13432.823	2721.91866	652.654672	201.314392	83.3748674	44.42	28.0041508	19.5425771	0.000484737
45.48	73379.63869	13432.9101	2721.9995	652.729706	201.384037	83.4395103	44.48	28.0598415	19.594268	0.000483458
45.78	73380.10234	13433.3413	2722.40053	653.102661	201.730889	83.7620866	44.78	28.3388454	19.8537455	0.000477143
45.93	73380.33077	13433.5543	2722.59907	653.287766	201.903463	83.9229784	44.93	28.478691	19.9841241	0.000474031
47.32	73382.34527	13435.4479	2724.37906	654.960992	203.476359	85.4015875	46.32	29.7854162	21.2125862	0.000446592
47.52	73382.62067	13435.709	2724.62661	655.195684	203.698865	85.6125404	46.52	29.9750321	21.3923573	0.00044284
48.24	73383.58391	13436.6266	2725.50077	656.028452	204.492203	86.3683187	47.24	30.6609508	22.0458109	0.00042972
48.28	73383.63616	13436.6766	2725.54857	656.07417	204.53593	86.4101404	47.28	30.6992084	22.082402	0.000429008
48.63	73384.08789	13437.1095	2725.96347	656.471793	204.916998	86.7753443	47.63	31.0346378	22.4038678	0.000422855
49.33	73384.96268	13437.9524	2726.7756	657.254306	205.670976	87.5018305	48.33	31.7091201	23.0537654	0.000410939
50.74	73386.61611	13439.5624	2728.34344	658.781063	207.157753	88.9496982	49.74	33.0822656	24.3910418	0.000388418
50.91	73386.80623	13439.749	2728.5266	658.960843	207.334212	89.1228975	49.91	33.2491253	24.5548194	0.000385828
51.15	73387.07142	13440.0098	2728.78311	659.213123	207.58233	89.3669227	50.15	33.4851666	24.7869674	0.000382216
51.27	73387.20262	13440.1391	2728.91047	659.338598	207.705953	89.4887206	50.27	33.6033953	24.9034508	0.000380429
51.54	73387.49448	13440.4272	2729.19485	659.619313	207.983049	89.7622459	50.54	33.8699156	25.1665365	0.000376453
51.76	73387.72892	13440.6592	2729.42437	659.846417	208.207761	89.984589	50.76	34.0875973	25.3819243	0.00037326
51.84	73387.81343	13440.7429	2729.50735	659.928642	208.289237	90.0653238	50.84	34.1668691	25.4604746	0.000372109
52.13	73388.11653	13441.0438	2729.80602	660.225119	208.583539	90.3574667	51.13	34.4547421	25.7462364	0.00036798
52.32	73388.31238	13441.2386	2729.9999	660.418025	208.775471	90.5484303	51.32	34.6437835	25.9343241	0.000365312
52.42	73388.41461	13441.3405	2730.10138	660.519132	208.876208	90.6487982	51.42	34.743417	26.0335924	0.00036392
52.47	73388.46551	13441.3912	2730.15198	660.569578	208.926505	90.6989462	51.47	34.7932694	26.0832976	0.000363227
52.69	73388.68772	13441.6131	2730.37345	660.790684	209.147241	90.9193142	51.69	35.0129021	26.3025637	0.0003602
54.19	73390.13143	13443.066	2731.83559	662.262159	210.62814	92.4097204	53.19	36.5225824	27.8220117	0.000340535
54.35	73390.27841	13443.215	2731.98679	662.415511	210.783676	92.5674729	53.35	36.684862	27.9866034	0.000338533
54.77	73390.65812	13443.6012	2732.37948	662.814857	211.189791	92.9804717	53.77	37.1119823	28.4209654	0.000333361
54.93	73390.80049	13443.7463	2732.52749	662.965783	211.343683	93.1373879	53.93	37.2751268	28.5873161	0.000331422
57.62	73393.01929	13446.0372	2734.89294	665.408324	213.865951	95.7421105	56.62	40.0533284	31.456748	0.000301199
57.84	73393.18718	13446.2129	2735.07666	665.600513	214.066997	95.9524196	56.84	40.283466	31.6974903	0.000298912
58.46	73393.65018	13446.699	2735.58699	666.136288	214.629491	96.5429664	57.46	40.9343897	32.3808818	0.000292605
60.76	73395.24573	13448.3946	2737.38916	668.051682	216.665288	98.7068027	59.76	43.3791976	34.9796952	0.000270872

61.42	73395.67081	13448.8521	2737.88138	668.58135	217.235255	99.3201358	60.42	44.089417	35.7439571	0.000265082
61.81	73395.91562	13449.1166	2738.1673	668.890341	217.569182	99.6810117	60.81	44.5108919	36.1994475	0.000261747
61.85	73395.94047	13449.1435	2738.19642	668.921867	217.603311	99.7179601	60.85	44.5541955	36.2463274	0.000261409
62.27	73396.19848	13449.4234	2738.49993	669.251054	217.960347	100.1052	61.27	45.0097276	36.7403982	0.000257895
61.45	73395.68981	13448.8726	2737.9035	668.605223	217.261017	99.3479361	60.45	44.1217908	35.7788924	0.000264823

Once the Box-Cox method has been completed, new **normalized data** are available:

0.000484737	0.000483458	0.000477143	0.000474031
0.000446592	0.00044284	0.00042972	0.000429008
0.000422855	0.000410939	0.000388418	0.000385828
0.000382216	0.000380429	0.000376453	0.00037326
0.000372109	0.00036798	0.000365312	0.00036392
0.000363227	0.0003602	0.000340535	0.000338533
0.000333361	0.000331422	0.000301199	0.000298912
0.000292605	0.000270872	0.000265082	0.000261747
0.000261409	0.000257895	0.000264823	

With these new data, the Anderson-Darling test is performed to conclude whether the data fit a normal distribution.

NORMAL DISTRIBUTION WITH ANDERSON-DARLING (NORMALIZED DATA)

The mean, variance and standard deviation of the new data are calculated.

Mean = 0.00036569

Variance = 4.7038E.09

Standard Deviation = 6.8584E-05

C0	C1	C2	C3	C4	C5	C6	C7	C8
i	Yi	Yn+1-i	2i-1	ADP(Yi)	1-PEA(Yn+1-i)	ln C4	ln C5	C3 * (C6+C7)
1	0.00025789	0.00048474	1	0.05801095	0.041298993	-2.84712348	-3.18691715	-6.03404064
2	0.00026141	0.00048346	3	0.06419897	0.042973992	-2.74576812	-3.14716019	-17.6787849
3	0.00026175	0.00047714	5	0.064821	0.052072075	-2.73612562	-2.95512646	-28.4562604
4	0.00026482	0.00047403	7	0.07069054	0.057085569	-2.64944357	-2.86320392	-38.5885324
5	0.00026508	0.00044659	9	0.07120238	0.11907383	-2.64222907	-2.12801156	-42.9321657
6	0.00027087	0.00044284	11	0.08341362	0.130308639	-2.48394363	-2.03784949	-49.7397244
7	0.00029261	0.00042972	13	0.14330527	0.175247741	-1.94277817	-1.74155464	-47.8963266
8	0.00029891	0.00042901	15	0.16511965	0.177938244	-1.80108494	-1.72631873	-52.911055
9	0.0003012	0.00042285	17	0.17353512	0.20227158	-1.75137526	-1.59814403	-56.9418279
10	0.00033142	0.00041094	19	0.30867054	0.254691812	-1.17548077	-1.36770104	-48.3204545
11	0.00033336	0.00038842	21	0.31869629	0.370164372	-1.14351669	-0.99380812	-44.8838212
12	0.00033853	0.00038583	23	0.34607662	0.384509401	-1.06109508	-0.95578704	-46.3882886
13	0.00034053	0.00038222	25	0.35690512	0.404781158	-1.0302853	-0.90440871	-48.3673501
14	0.0003602	0.00038043	27	0.46811224	0.414909505	-0.75904718	-0.87969484	-44.2460347
15	0.00036323	0.00037645	29	0.48568811	0.437634457	-0.7221886	-0.82637129	-44.9082368
16	0.00036392	0.00037326	31	0.48971868	0.456042713	-0.71392417	-0.78516881	-46.4718823
17	0.00036531	0.00037211	33	0.49781727	0.462703918	-0.6975222	-0.77066792	-48.4502739
18	0.00036798	0.00036798	35	0.51333284	0.486667158	-0.66683083	-0.72017484	-48.5451985
19	0.00037211	0.00036531	37	0.53729608	0.502182733	-0.62120597	-0.68879122	-48.469896
20	0.00037326	0.00036392	39	0.54395729	0.510281318	-0.60888455	-0.6727931	-49.9854285
21	0.00037645	0.00036323	41	0.56236554	0.514311886	-0.57560321	-0.66492542	-50.8616736
22	0.00038043	0.0003602	43	0.5850905	0.53188776	-0.53598875	-0.63132279	-50.1943962
23	0.00038222	0.00034053	45	0.59521884	0.643094877	-0.51882614	-0.44146301	-43.2130118
24	0.00038583	0.00033853	47	0.6154906	0.653923377	-0.48533561	-0.42476509	-42.774733
25	0.00038842	0.00033336	49	0.62983563	0.681303708	-0.4622964	-0.3837471	-41.4561315
26	0.00041094	0.00033142	51	0.74530819	0.691329456	-0.29395747	-0.36913879	-33.8179092
27	0.00042285	0.0003012	53	0.79772842	0.826464877	-0.22598707	-0.19059786	-22.079001

28	0.00042901	0.00029891	55	0.82206176	0.834880353	-0.19593976	-0.18046685	-20.7023636
29	0.00042972	0.00029261	57	0.82475226	0.85669473	-0.19267223	-0.15467363	-19.7987141
30	0.00044284	0.00027087	59	0.86969136	0.916586376	-0.13961689	-0.08709897	-13.3762356
31	0.00044659	0.00026508	61	0.88092617	0.928797623	-0.12678146	-0.07386441	-12.2393979
32	0.00047403	0.00026482	63	0.94291443	0.929309463	-0.05877974	-0.07331348	-8.32187305
33	0.00047714	0.00026175	65	0.94792792	0.935178998	-0.05347681	-0.06701733	-7.83211877
34	0.00048346	0.00026141	67	0.95702601	0.935801031	-0.04392471	-0.0663524	-7.38856638
35	0.00048474	0.00025789	69	0.95870101	0.941989049	-0.04217603	-0.05976163	-7.03369841
								-1241.30541

$\sum (2i-1)[\ln PEA(Yi) + \ln(1-PEA(Yn+1-i))] = -1241.30541$

Hypothesis

H0; The data fit a normal distribution H1; The data do not fit a normal distribution

Calculation of Anderson Darling test statistic

: $A^2 n = -[n + 1/n * \sum (2i-1)[\ln PEA(Yi) + \ln(1-PEA(_{Yn+1-i}))]]]$

$$^2 An = -(35 + .\frac{1}{35}(-1241.30541)) = 0.465868779$$

In this case we are also looking to see if the data follow a normal distribution and therefore it is necessary to adjust the test statistic to the normal with the following formula: $A^2 n^*(1+\frac{4}{n} -)\frac{25}{n2}$

$$NORMAL = 0.465868779\,(1 + \frac{4}{35} - \frac{25}{35^2}) = 0.509603399$$

Table value:

$$A^2 n_{0\cdot 005} = 0.751$$

Conclusion

As 0.509603399 < 0.751 H0 is not rejected and the data are said to fit a normal distribution with $\mu = 0.00036569$ and $\sigma = 6.8584E-05$.

With a Pvalue = 0.238

It can be observed that when normalizing the data using the Box-Cox method, the Anderson Darling test does not reject H0, so with the normalized data Anderson Darling tells us that the data fit a normal distribution.

Continuing with the investigation, the X^2 and Anderson- Darling test will be performed with the original data to conclude if the data fit a Weibull distribution.

WEIBULL DISTRIBUTION WITH X2

To start this test, we entered the data into the Minitab program to calculate and output the shape and scale parameters.

Estimaciones ML de los parámetros de distribuciónS

Distribución	Ubicación	Forma	Escala	Valor umbral
Normal*	52.99686		5.16726	
Transformación de Box-Cox*	0.00037		0.00007	
Lognormal*	3.96569		0.09629	
Lognormal de 3 parámetros	2.74742		0.31545	36.61186
Exponencial			52.99686	
Exponencial de 2 parámetros			7.79970	45.19715
Weibull		10.80469	55.38251	
Weibull de 3 parámetros		1.59765	9.18014	44.72993
Valor extremo más pequeño	55.63599		5.10624	
Valor extremo por máximos	50.56878		4.24461	
Gamma		110.33430	0.48033	
Gamma de 3 parámetros		3.60430	2.83045	42.79476
Logística	52.65086		2.98791	
Loglogística	3.96130		0.05595	
Loglogística de 3 parámetros	2.59685		0.21727	38.76498
Transformación de Johnson*	0.02043		1.13955	

$\alpha = 10.80469$

$\beta = 55.38251$

Orderly times:

45.42, 45.48, 45.78, 45.93, 47.32, 47.52, 48.24, 48.28, 48.63, 49.33, 50.74, 50.91, 51.15, 51.27, 51.54, 51.76, 51.84, 52.13, 52.32, 52.42, 52.47, 52.69, 54.19, 54.35, 54.77, 54.93, 57.62, 57.84, 58.46, 60.76, 61.42, 61.45, 61.81, 61.85, 62.27

We order the data from smallest to largest, in order to make the detailed diagram and sheet and then make the table of classes and frequencies.

Maximum value = 62.27

Minimum value = 45.42

Range = 62.27 - 45.42 = 16.85

k = $\sqrt{35} = 5.91607978 \approx 6$

h = 16.85 / 6 = 2.81

Stem	Sheet
45	42, 48, 78,93
46	
47	32, 52
48	24,28, 63
49	33
50	74, 91
51	15, 27, 54,76,84
52	13,32,42,47,69
53	
54	19,35,77,93
55	
56	
57	62,84
58	46
59	
60	76
61	42,45,81,85
62	27

With the help of the stem-and-leaf diagram, the table of classes and frequencies is made and the expected frequencies are calculated, which are the product of the Weibull probability by the total number of data (35).

LRI	Classes		LRS	Frequency	Expected		
41.415	45.42	48.23	48.235	6	7.04361	7.04361	6
48.235	48.24	51.05	51.055	6	4.848375	11.29359	16
51.055	51.06	53.87	53.875	10	6.445215		
53.875	53.88	56.68	56.685	4	7.33509	7.33509	4
56.685	56.69	59.50	59.505	3	5.340265	9.32771	9
59.505	59.51	62.32	62.325	6	3.987445		
				35	35	35	35

The expected frequencies that are less than 5 are grouped and simultaneously the observed frequencies are grouped.

Hypothesis

Ho; The data fit a Weibull distribution

H1; The data do not fit a Weibull distribution

Calculation of test statistic X^2

$$X^2 = \frac{(6-7.04361)^2}{7.04361} + \frac{(16-11.29359)^2}{11.29359} + \frac{(4-7.33509)^2}{7.33509} + \frac{(9-9.32771)^2}{9.32771} = 3.64384027$$

We know that our value of degrees of libertarianism is:

V = k-1- m

Where:

k = Number of classes

m = Number of calculated parameters, in this case α and β. Therefore:

V = 4-1-2 = 1

X^2(0.05; 1) = 3.8414

Conclusion

As 3.6430<3.8414, H0 is not rejected and is said to follow a Weibull distribution with α = 10.80469 β = 55.38251

As we observe X^2 tells us that the data does fit a De Weibull distribution, however, Minitab once again tells us with Anderson-Darling that the distribution does not fit Weibull with a Pvalue = 0.010.

Prueba de bondad del ajuste

Distribución	AD	P	LRT P
Normal	0.783	0.038	
Transformación Box-Cox	0.466	0.238	
Lognormal	0.623	0.096	
Lognormal de 3 parámetros	0.478	*	0.274
Exponencial	13.214	<0.003	
Exponencial de 2 parámetros	1.548	0.018	0.000
Weibull	1.240	<0.010	
Weibull de 3 parámetros	0.544	0.171	0.002
Valor extremo más pequeño	1.490	<0.010	
Valor extremo por máximos	0.471	0.238	

An Anderson-Darling test is then performed for comparison with the X test.[2.]

Weibull Distribution with Anderson Darling

$\alpha = 10.80469$ $\beta = 55.38251$

C0	C1	C2	C3	C4	C5	C6	C7	C8
i	Yi	Yn+1-i	2i-1	ADP(Yi)	1-PEA(Yn+1-i)	ln C4	ln C5	C3 * (C6+C7)
1	45.42	62.27	1	0.1107159	0.02877217	-2.2007878	-3.5483467	-5.7491345
2	45.48	61.85	3	0.11221367	0.03695177	-2.18735042	-3.29814173	-16.4564765
3	45.78	61.81	5	0.11995907	0.037810504	-2.12060466	-3.27516832	-26.9788649
4	45.93	61.45	7	0.12399558	0.046196641	-2.08750935	-3.07484819	-36.1365027
5	47.32	61.42	9	0.16698187	0.046950213	-1.78987005	-3.05866753	-43.6368382
6	47.52	60.76	11	0.17404762	0.065764549	-1.74842631	-2.72167435	-49.1711073
7	48.24	58.46	13	0.20144681	0.166335969	-1.60222991	-1.79374563	-44.147682
8	48.28	57.84	15	0.20306113	0.20218708	-1.59424824	-1.59856187	-47.8921516
9	48.63	57.62	17	0.21760948	0.215654495	-1.52505319	-1.53407771	-52.0052254
10	49.33	54.93	19	0.24902141	0.400447928	-1.39021639	-0.91517154	-43.8023706
11	50.74	54.77	21	0.32180014	0.411981368	-1.1338246	-0.88677715	-42.4326368
12	50.91	54.35	23	0.33142282	0.442195099	-1.10436031	-0.81600409	-44.1683812
13	51.15	54.19	25	0.34530826	0.453654122	-1.06331774	-0.79042022	-46.343449
14	51.27	52.69	27	0.3523808	0.557868892	-1.04304286	-0.5836313	-43.9202024
15	51.54	52.47	29	0.36860332	0.572447626	-0.99803424	-0.55783403	-45.1201797
16	51.76	52.42	31	0.38213009	0.575729531	-0.96199419	-0.55211729	-46.9374558
17	51.84	52.32	33	0.38711567	0.582257196	-0.94903174	-0.54084301	-49.1658668
18	52.13	52.13	35	0.40547686	0.594523142	-0.90269148	-0.51999564	-49.794049
19	52.32	51.84	37	0.4177428	0.61288433	-0.87288934	-0.48957906	-50.4113306
20	52.42	51.76	39	0.42427047	0.617869913	-0.85738413	-0.48147734	-52.2155973
21	52.47	51.54	41	0.42755237	0.631396683	-0.84967849	-0.45982096	-53.6894771
22	52.69	51.27	43	0.44213111	0.647619196	-0.81614882	-0.43445242	-53.775853
23	54.19	51.15	45	0.54634588	0.654691735	-0.60450303	-0.42359079	-46.2642216
24	54.35	50.91	47	0.5578049	0.668577177	-0.58374602	-0.40260344	-46.3584246
25	54.77	50.74	49	0.58801863	0.678199857	-0.53099664	-0.38831326	-45.0461853
26	54.93	49.33	51	0.59955207	0.750978586	-0.51157245	-0.28637814	-40.6954801
27	57.62	48.63	53	0.7843455	0.782390519	-0.24290566	-0.24540128	-25.8802678
28	57.84	48.28	55	0.79781292	0.796938875	-0.22588115	-0.2269773	-24.9072144
29	58.46	48.24	57	0.83366403	0.798553191	-0.1819248	-0.2249537	-23.1920743
30	60.76	47.52	59	0.93423545	0.825952375	-0.06802678	-0.19121816	-15.295452
31	61.42	47.32	61	0.95304979	0.833018132	-0.04808813	-0.18269987	-14.0780683
32	61.45	45.93	63	0.95380336	0.876004419	-0.04729775	-0.13238414	-11.3199594
33	61.81	45.78	65	0.9621895	0.880040928	-0.03854387	-0.12778686	-10.8114975
34	61.85	45.48	67	0.96304823	0.887786326	-0.03765179	-0.11902419	-10.4972903
35	62.27	45.42	69	0.97122783	0.889284098	-0.0291942	-0.11733852	-10.1107582
								-1268.40773

Journal of Engineering Research ISSN 2764-1317 DOI https://doi.org/10.22533/at.ed.3174232418099

Hypothesis

Ho; The data fit a Weibull distribution

H1; The data do not fit a Weibull distribution

Calculation of Anderson Darling test statistic

$$\mathbf{A^2n} = -\left(\mathbf{35} + \frac{\mathbf{1}}{\mathbf{35}}(-\mathbf{1268.40773})\right) = \mathbf{1.240220746}$$

Recall that the test statistic should be fitted to the Weibull distribution with the following formula: $A^2n = (1+\frac{1}{5\sqrt{n}})$

$$\boldsymbol{WEIBULL} = \mathbf{1.240220746}$$

$$\left(\mathbf{1} + \frac{\mathbf{1}}{\mathbf{5\sqrt{35}}}\right) = \mathbf{1.28214786}$$

$\mathbf{A^2n}_{\,0.005} = 0.757$

Conclusion

As 1.282147 > 0.757 H0 is rejected and the data is said not to fit a Weibull distribution with $\alpha = 10.80469$ and $\beta = 55.38251$ and a NC = 95%.

It is seen again that Anderson-Darling says the opposite of X^2 , this is because Anderson-Darling is a modification to the Kolmogorov-Smirnov test and this modification makes it more effective in detecting discrepancies at the extremes of the distribution.

It is then concluded that the data do not fit the Weibull distribution.

ANDERSON-DARLING WEIBULL DISTRIBUTION (NORMALIZED DATA)

Since the data do not approximate a Weibull distribution, we proceed to perform the Anderson-Darling test, but this time with the normalized Box-Cox data.

0.000484737	0.000483458	0.000477143	0.000474031
0.000446592	0.00044284	0.00042972	0.000429008
0.000422855	0.000410939	0.000388418	0.000385828
0.000382216	0.000380429	0.000376453	0.00037326
0.000372109	0.00036798	0.000365312	0.00036392
0.000363227	0.0003602	0.000340535	0.000338533
0.000333361	0.000331422	0.000301199	0.000298912
0.000292605	0.000270872	0.000265082	0.000261747
0.000261409	0.000257895	0.000264823	

We enter the data into Minitab to output the shape and scale parameters.

$\alpha = 6.07347$ $\beta = 0.00039$

Estimaciones ML de los parámetros de distribución

Distribución	Ubicación	Forma	Escala	Valor umbral
Normal*	0.00037		0.00007	
Transformación de Box-Cox*	0.00037		0.00007	
Lognormal*	-7.93139		0.19257	
Lognormal de 3 parámetros	-3.03209		0.00140	-0.04785
Exponencial			0.00037	
Exponencial de 2 parámetros			0.00011	0.00025
Weibull		6.07347	0.00039	
Weibull de 3 parámetros		791.69210	0.05010	-0.04970
Valor extremo más pequeño	0.00040		0.00006	
Valor extremo por máximos	0.00033		0.00006	
Gamma		28.48238	0.00001	
Gamma de 3 parámetros		3.84334	0.00004	0.00022
Logística	0.00037		0.00004	
Loglogística	-7.92260		0.11191	
Loglogística de 3 parámetros	4.15888		0.00000	-63.99974

* Escala: Estimación de ML ajustado

C0	C1	C2	C3	C4	C5	C6	C7	C8
i	Yi	Yn+1-i	2i-1	ADP(Yi)	1-PEA(Yn+1-i)	ln C4	ln C5	C3 * (C6+C7)
1	0.00025789	0.00048474	1	0.07790683	0.023608418	-2.5522416	-3.74615195	-6.29839355
2	0.00026141	0.00048346	3	0.08429263	0.02505807	-2.47346085	-3.68655936	-18.4800606
3	0.00026175	0.00047714	5	0.08492856	0.03325356	-2.46594481	-3.40359346	-29.3476913
4	0.00026482	0.00047403	7	0.09088106	0.037968297	-2.39820362	-3.27100377	-39.6844517
5	0.00026508	0.00044659	9	0.09139624	0.102572208	-2.39255092	-2.27718826	-42.0276526
6	0.00027087	0.00044284	11	0.10352699	0.114926462	-2.26792294	-2.16346281	-48.7452433
7	0.00029261	0.00042972	13	0.16023908	0.164924407	-1.83108831	-1.80226805	-47.2336326
8	0.00029891	0.00042901	15	0.18027463	0.167929001	-1.71327385	-1.784214	-52.4623178
9	0.0003012	0.00042285	17	0.18795867	0.195081462	-1.67153317	-1.63433805	-56.1998108
10	0.00033142	0.00041094	19	0.31074089	0.25312381	-1.16879587	-1.37387654	-48.3107758
11	0.00033336	0.00038842	21	0.31993062	0.37696162	-1.13965113	-0.9756119	-44.4205237

12	0.00033853	0.00038583	23	0.34515218	0.391892401	-1.06376985	-0.93676796	-46.0123696
13	0.00034053	0.00038222	25	0.35518267	0.412816662	-1.03512306	-0.8847517	-47.9968691
14	0.0003602	0.00038043	27	0.46047959	0.423194541	-0.77548674	-0.8599233	-44.1560711
15	0.00036323	0.00037645	29	0.47756128	0.44629593	-0.7390628	-0.80677303	-44.8292389
16	0.00036392	0.00037326	31	0.48149727	0.464825749	-0.73085471	-0.76609268	-46.4053691
17	0.00036531	0.00037211	33	0.48942738	0.471491157	-0.71451918	-0.75185493	-48.3903457
18	0.00036798	0.00036798	35	0.50470184	0.495298162	-0.68378744	-0.70259535	-48.5233978
19	0.00037211	0.00036531	37	0.52850884	0.510572619	-0.63769574	-0.6722224	-48.4669713
20	0.00037326	0.00036392	39	0.53517425	0.518502728	-0.62516288	-0.65680999	-49.996942
21	0.00037645	0.00036323	41	0.55370407	0.522438723	-0.5911249	-0.64924758	-50.8552718
22	0.00038043	0.0003602	43	0.57680546	0.539520409	-0.55025023	-0.61707467	-50.1949704
23	0.00038222	0.00034053	45	0.58718334	0.644817332	-0.53241818	-0.43878821	-43.7042874
24	0.00038583	0.00033853	47	0.6081076	0.654847816	-0.49740344	-0.42335241	-43.2755251
25	0.00038842	0.00033336	49	0.62303838	0.680069384	-0.47314716	-0.38556045	-42.0766728
26	0.00041094	0.00033142	51	0.74687619	0.689259112	-0.29185585	-0.37213801	-33.8636869
27	0.00042285	0.0003012	53	0.80491854	0.812041328	-0.2170142	-0.20820404	-22.536567
28	0.00042901	0.00029891	55	0.832071	0.819725367	-0.18383751	-0.19878591	-21.0442881
29	0.00042972	0.00029261	57	0.83507559	0.839760917	-0.18023303	-0.17463805	-20.2276515
30	0.00044284	0.00027087	59	0.88507354	0.896473012	-0.12208454	-0.10928709	-13.6509264
31	0.00044659	0.00026508	61	0.89742779	0.908603758	-0.10822262	-0.09584619	-12.4481972
32	0.00047403	0.00026482	63	0.9620317	0.909118936	-0.03870787	-0.09527935	-8.44119509
33	0.00047714	0.00026175	65	0.96674644	0.915071437	-0.03381903	-0.08875314	-7.96719134
34	0.00048346	0.00026141	67	0.97494193	0.91570737	-0.02537737	-0.08805843	-7.60019849
35	0.00048474	0.00025789	69	0.97639158	0.922093166	-0.02389156	-0.08110901	-7.24503964
								-1243.1198

∑ (2i-1)[lnPEA(Yi) + ln(1-PEA(Yn+1-i))]= -1243.1198

Hypothesis

Ho; The data fit a Weibull distribution
H1; The data do not fit a Weibull distribution

Calculation of Anderson Darling test statistic

$$\mathbf{A^2n = -(35 + \frac{1}{35}(-1243.1198)) = 0.517708507}$$

We fit our test statistic to the Weibull distribution.

$$\boldsymbol{WEIBULL} = \mathbf{0.517708507\,(1 + \frac{1}{5\sqrt{35}}) = 0.535210249}$$

$$\mathbf{A^2n}_{0.005} = 0.757$$

Conclusion

Since 0.53521025 < 0.757 H0 is NOT rejected and the data are said to fit a Weibull distribution with α = 6.07347 and β = 0.00039.

Pvalue = 0.218

With normalized data the Anderson-Darling test says that the data do fit a Weibull distribution.

EXPONENTIAL DISTRIBUTION WITH KOLMOGOROV-SMIRNOV (NORMALIZED DATA)

Continuing with the search for the distribution to which the data best fit, we proceed to perform the Kolmogorov-Smirnov test to conclude whether the data fit the exponential distribution.

Journal of Engineering Research ISSN 2764-1317 DOI https://doi.org/10.22533/at.ed.3174232418099

The mean of the data (β) is calculated.

$\beta = 0.00036569$

We proceed to find the value of λ which is equal to $\frac{1}{\beta}$

$\lambda = 2734.57347$

Then we made the following table of frequencies and probabilities.

X	Frequency	Poi	POAi	PEAi	|POAi-PEAi|
0.00025789	1	0.0286	0.0286	0.50600593	**0.4774345**
0.00026141	1	0.0286	0.0571	0.51073069	0.453587835
0.00026175	1	0.0286	0.0857	0.51118331	0.425469021
0.00026482	1	0.0286	0.1143	0.51527755	0.400991838
0.00026508	1	0.0286	0.1429	0.51562042	0.372763282
0.00027087	1	0.0286	0.1714	0.52322948	0.351800907
0.00029261	1	0.0286	0.2000	0.5507389	0.350738897
0.00029891	1	0.0286	0.2286	0.5584204	0.32984897
0.0003012	1	0.0286	0.2571	0.56117331	0.304030454
0.00033142	1	0.0286	0.2857	0.59598234	0.310268058
0.00033336	1	0.0286	0.3143	0.59811913	0.283833415
0.00033853	1	0.0286	0.3429	0.60376315	0.260906004
0.00034053	1	0.0286	0.3714	0.60592651	0.234497934
0.0003602	1	0.0286	0.4000	0.62655821	0.226558205
0.00036323	1	0.0286	0.4286	0.62963651	0.201065078
0.00036392	1	0.0286	0.4571	0.63033795	0.173195095
0.00036531	1	0.0286	0.4857	0.63174287	0.146028584
0.00036798	1	0.0286	0.5143	0.63441963	0.120133911
0.00037211	1	0.0286	0.5429	0.63852378	0.095666638
0.00037326	1	0.0286	0.5714	0.63965988	0.068231311
0.00037645	1	0.0286	0.6000	0.64279283	0.042792829
0.00038043	1	0.0286	0.6286	0.64665503	0.018083602
0.00038222	1	0.0286	0.6571	0.6483776	0.00876526
0.00038583	1	0.0286	0.6857	0.65183374	0.033880541
0.00038842	1	0.0286	0.7143	0.65429065	0.059995063
0.00041094	1	0.0286	0.7429	0.6749395	0.067917648
0.00042285	1	0.0286	0.7714	0.68536059	0.086067977
0.00042901	1	0.0286	0.8000	0.69061045	0.109389547
0.00042972	1	0.0286	0.8286	0.69121204	0.137359385
0.00044284	1	0.0286	0.8571	0.70209461	0.155048245
0.00044659	1	0.0286	0.8857	0.70513495	0.180579338
0.00047403	1	0.0286	0.9143	0.72645077	0.187834941
0.00047714	1	0.0286	0.9429	0.72876839	0.214088749
0.00048346	1	0.0286	0.9714	0.73341242	0.238016155
0.00048474	1	0.0286	1.0000	0.73434256	0.265657439

Finishing the table, select the highest value of the last column and this will be the value of D0, then:

Table value:

$$D_0 = 0.4774345$$

$$D_{0.05,35} = 0.230$$

Hypothesis

H0; The data fit an Exponential distribution with λ = 2734.57347.

H1; The data do NOT fit an Exponential distribution with λ = 2734.57347.

Conclusion

As 0.4774345 > 0.230 H0 is rejected and the data is said not to fit an Exponential distribution with λ=2734.57347 and α=0.05.

EXPONENTIAL DISTRIBUTION WITH ANDERSON DARLING (NORMALIZED DATA)

Now the Anderson-Darling test will be performed to compare the Kolmogorov-Smirnov response. One has:

β = 0.00036569

λ = 2734.57347

So we proceed to perform the table to obtain ∑ (2i-1) [lnPEA(Yi) + ln(1- PEA(Yn+1-i))] and substitute in the formula of A^2 n.

C0	C1	C2	C3	C4	C5	C6	C7	C8
i	Yi	Yn+1-i	2i-1	ADP(Yi)	1-PEA(Yn+1-i)	ln C4	ln C5	C3 * (C6+C7)
1	0.00025789	0.00048474	1	0.50600593	0.26565744	-0.68120689	-1.32554763	-2.00675452
2	0.00026141	0.00048346	3	0.51073069	0.26658758	-0.67191285	-1.32205245	-5.98189588
3	0.00026175	0.00047714	5	0.51118331	0.27123161	-0.67102703	-1.30478219	-9.87904609
4	0.00026482	0.00047403	7	0.51527755	0.27354923	-0.66304959	-1.29627369	-13.7152629
5	0.00026508	0.00044659	9	0.51562042	0.29486505	-0.6623844	-1.22123748	-16.9525969
6	0.00027087	0.00044284	11	0.52322948	0.29790539	-0.64773514	-1.21097933	-20.4458592
7	0.00029261	0.00042972	13	0.5507389	0.30878796	-0.59649445	-1.17510046	-23.0307339
8	0.00029891	0.00042901	15	0.5584204	0.30938955	-0.5826432	-1.17315413	-26.3369599
9	0.0003012	0.00042285	17	0.56117331	0.31463941	-0.57772549	-1.15632804	-29.47891
10	0.00033142	0.00041094	19	0.59598234	0.3250605	-0.51754424	-1.12374395	-31.1844755
11	0.00033336	0.00038842	21	0.59811913	0.34570935	-0.51396533	-1.06215689	-33.0985666
12	0.00033853	0.00038583	23	0.60376315	0.34816626	-0.5045733	-1.05507517	-35.8719147
13	0.00034053	0.00038222	25	0.60592651	0.3516224	-0.50099658	-1.0451974	-38.6548494
14	0.0003602	0.00038043	27	0.62655821	0.35334497	-0.4675136	-1.04031045	-40.7112494
15	0.00036323	0.00037645	29	0.62963651	0.35720717	-0.4626126	-1.02943936	-43.2695067
16	0.00036392	0.00037326	31	0.63033795	0.36034012	-0.46149917	-1.02070692	-45.9483889
17	0.00036531	0.00037211	33	0.63174287	0.36147622	-0.45927282	-1.01755902	-48.7354508
18	0.00036798	0.00036798	35	0.63441963	0.36558037	-0.45504467	-1.00626912	-51.1459828
19	0.00037211	0.00036531	37	0.63852378	0.36825713	-0.44859636	-0.99897386	-53.5600982
20	0.00037326	0.00036392	39	0.63965988	0.36966205	-0.44681868	-0.99516607	-56.2374053
21	0.00037645	0.00036323	41	0.64279283	0.37036349	-0.4419328	-0.99327034	-58.8433288
22	0.00038043	0.0003602	43	0.64665503	0.37344179	-0.43594231	-0.98499312	-61.1002236
23	0.00038222	0.00034053	45	0.6483776	0.39407349	-0.43328204	-0.93121785	-61.4024953
24	0.00038583	0.00033853	47	0.65183374	0.39623685	-0.42796574	-0.92574313	-63.6243171
25	0.00038842	0.00033336	49	0.65429065	0.40188087	-0.42420361	-0.91159957	-65.4543559
26	0.00041094	0.00033142	51	0.6749395	0.40401766	-0.39313223	-0.9062967	-66.2708753
27	0.00042285	0.0003012	53	0.68536059	0.43882669	-0.37781016	-0.82365073	-63.6774274
28	0.00042901	0.00029891	55	0.69061045	0.4415796	-0.37017936	-0.81739698	-65.3166984

29	0.00042972	0.00029261	57	0.69121204	0.4492611	-0.36930864	-0.80015104	-66.6592016
30	0.00044284	0.00027087	59	0.70209461	0.47677052	-0.35368711	-0.74071999	-64.5700188
31	0.00044659	0.00026508	61	0.70513495	0.48437958	-0.34936608	-0.72488643	-65.5294032
32	0.00047403	0.00026482	63	0.72645077	0.48472245	-0.31958456	-0.72417882	-65.757093
33	0.00047714	0.00026175	65	0.72876839	0.48881669	-0.3163993	-0.71576772	-67.0908564
34	0.00048346	0.00026141	67	0.73341242	0.48926931	-0.31004709	-0.71484221	-68.6675832
35	0.00048474	0.00025789	69	0.73434256	0.49399407	-0.30877965	-0.70523176	-69.9667878
								-1600.17657

∑ (2i-1)[lnPEA(Yi) + ln(1-PEA(Yn+1-i))]= -1600.17657

Hypothesis

Ho; The data fit a Weibull distribution H1; The data do not fit a Weibull distribution

CALCULATION OF ANDERSON DARLING TEST STATISTIC

$$\mathbf{A^2n = -\left(35 + \frac{1}{35}(-1600.17657)\right) = 10.7193307}$$

We fit the test statistic to the exponential distribution with the formula: A^2 n (1 +)3

A^2 nα with the Anderson-Darling table with an α=0.05.

$\mathbf{A^2n}_{0.005} = 1.326$

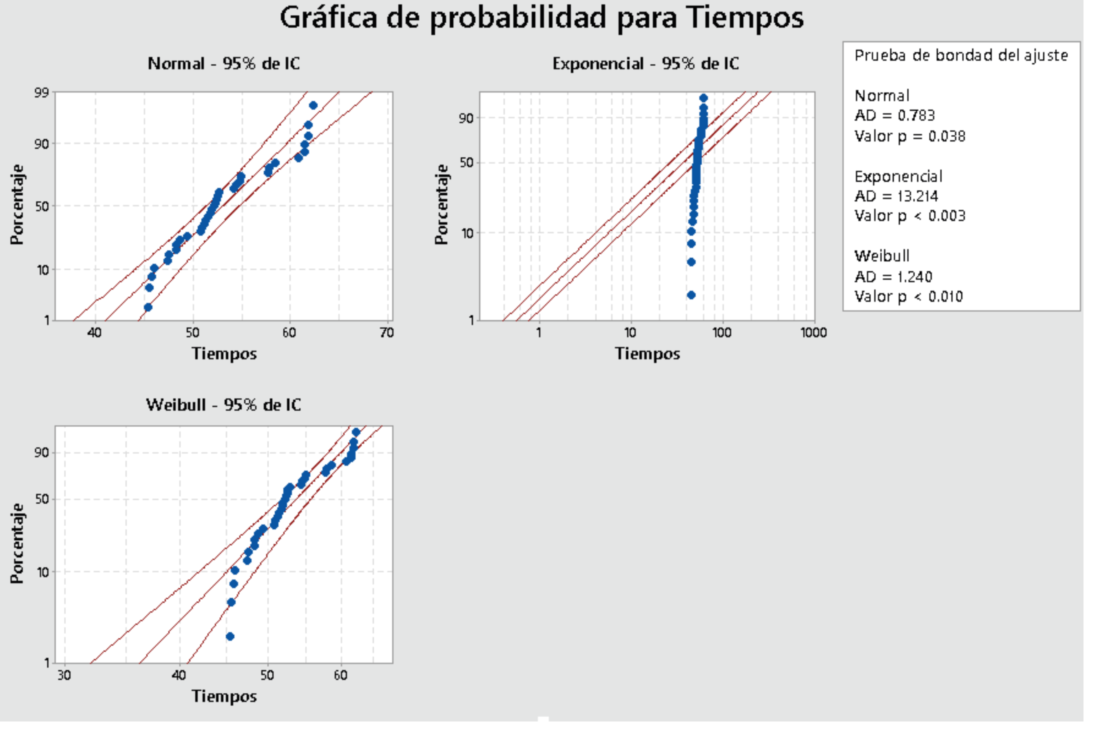
Conclusion

As 10.9030906 > 1.326 H0 is rejected and the data is said not to fit an Exponential distribution with λ=2734.57347.

The graphs obtained from the Minitab program before and after normalizing the data with the Box-Cox method are shown below.

GENERAL CONCLUSION

Probability plot of original data

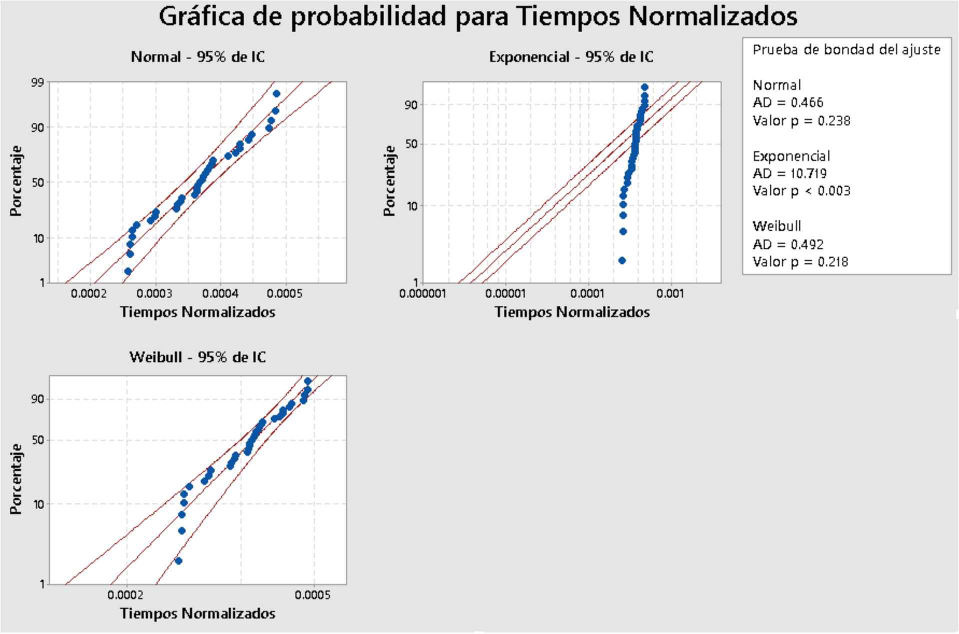


As can be seen the p-value of the Normal, Exponential and Weibull distributions is less than the value of α= 0.05 so the data does not fit any of these distributions.

1. Normal: Pvalue = 0.038 < α= 0.05
2. Weibull: Pvalue = 0.010 < α= 0.05
3. Exponential: Pvalue = 0.003 < α= 0.05

Plot of normalized data with

BOX-COX METHOD



When the data were normalized and Normal, Exponential and Weibull tests were performed the P value increased, so:

1. The data do fit a normal distribution with a Pvalue = 0.238 which is greater than the value of α= 0.05.

2. The data do fit a Weibull distribution with Pvalue = 0.218 which is greater than the value of α= 0.05.

3. The data do not fit an exponential distribution because their value

of Pvalue = 0.003 is well below the value of α= 0.05.

CONCLUSIONS

For data collection it is essential to have the right measuring instrument, with a reliable resolution and to avoid adding any kind of noise, such as changing the person to obtain the data.

The planning of the registration and data collection plays a very important role in this work since any distraction influences the numbers obtained, affecting the conclusions.

Finally, it is recommended to start by testing the data with the distributions to be performed by means of a Chi-Square test for goodness of fit, and once all the distributions have been tested, it is recommended to analyze the results to subsequently perform an Anderson-Darling or Kolmogorov Smirnov test. It is necessary to have an auxiliary tool, such as MINITAB software, to check the conclusions.

After this investigation, it is recommended to make several iterations in order to conclude on the reliability of the non-parametric tests.

REFERENCES

Distribución de Weibull - Minitab. (s. f.). (C) Minitab, LLC. All rights Reserved. 2022. https://support.minitab.com/es-mx/minitab/21/help-and-how-to/probability-distributions-random-data-and-resampling-analyses/supporting-topics/distributions/weibull-distribution/importancia de la distribución weibull - Google Zoeken. (s. f.). https://www.google.com/search?q=importancia+de+la+distribuci%C3%B3n+weibul l RPubs - Weibull. (2016, 29 marzo https://rpubs.com/anlykiro/weibull

Jensen, W., & Alexander, M. (2016). Statistics for Engineering and the Sciences. Journal of Quality Technology, 48(3), 297–299. https://www.tandfonline.com/doi/abs/10.1080/00224065.2016.11918168

Guisande, C., & Barreiro, A. (2006). Tratamiento de datos. Ediciones Díaz de Santos. https://books.google.com.ec/books?id=AhNx24025ZoC&printsec=frontcov er&hl=es#v=onepage&q&f=false

Distribución normal - Minitab. (s. f.-a). (C) Minitab, LLC. All rights Reserved. 2021. https://support.minitab.com/es-mx/minitab/20/help-and-how-to/probability-distributions-random-data-and-resampling-analyses/supporting-topics/distributions/normal-distribution/

LA DISTRIBUCIÓN NORMAL. (2017, 2 diciembre). Revista Chilena de Anestesia. https://revistachilenadeanestesia.cl/la-distribucion-normal/

Lejarza, J., & Lejarza, I. (2007). Distribución normal. Proyecto CEACES– Universitat de Vàlencia. https://www.uv.es/ceaces/pdf/normal.pdf

García Bellido, R, González Such, J. & Jornet Meliá, J.M. (2010). SPSS: PRUEBAS NO PARAMÉTRICAS. En Grupo de Innovación Educativa Universitat de Valencia. Recuperado 17 de octubre de 2022, de https://www.uv.es/innomide/spss/SPSS/SPSS_0802A.pdf

Kolmogorov-Smirnov prueba - Minitab Engage. (s. f.). (C) Minitab, LLC. All rights Reserved. 2022. Recuperado 19 de octubre de 2022, de https://support.minitab.com/es-mx/engage/glossary/k/kolmogorov-smirnov-test/

Mitjana, L. R. & Mitjana, L. R. (2019, 28 mayo). Prueba de Kolmogórov-Smirnov: qué es y cómo se usa en estadística. Recuperado 21 de octubre de 2022, de https://psicologiaymente.com/miscelanea/prueba-kolmogorov-smirnov

Distribución exponencial - Minitab. (s. f.). (C) Minitab, LLC. All rights Reserved. 2021. Recuperado 21 de octubre de 2022, de https://support.minitab.com/es-mx/minitab/20/help-and-how-to/probability- distributions-random-data-and-resampling-analyses/supporting- topics/distributions/exponential-distribution/

Benites, L. (2021, 5 octubre). Transformación de Box Cox: definición, ejemplos. Statologos: El sitio web para que aprendas estadística en Stata, R y Phyton. Recuperado 21 de octubre de 2022, de https://statologos.com/transformacion-de-caja-cox/

RPubs - DistribuciÃ³n chi cuadrado. (2016, 31 marzo). Recuperado 21 de octubre de 2022, de https://rpubs.com/angelicalopez/chicuadrado

Distribución de chi-cuadrada - Minitab. (s. f.). (C) Minitab, LLC. All rights Reserved. 2022. Recuperado 21 de octubre de 2022, de https://support.minitab.com/es-mx/minitab/21/help-and-how-to/probability- distributions-random-data-and-resampling-analyses/supporting- topics/distributions/chi-square-distribution/

Narvaez, M. (2022, 4 mayo). Prueba de chi-cuadrado: ¿Qué es y cómo se realiza? QuestionPro. Recuperado 21 de octubre de 2022, de https://www.questionpro.com/blog/es/prueba-de-chi-cuadrado-de-pearson/

PRUEBA DE ANDERSON DARLING. (2011, 14 marzo). Simulación de Procesos Empresariales. Recuperado 21 de octubre de 2022, de https://simulaciondeprocesosempresariales.wordpress.com/2011/02/28/prue ba-de-anderson-darling/

Mitjana, L. R. & Mitjana, L. R. (2019b, mayo 28). Prueba de Kolmogórov-Smirnov: qué es y cómo se usa en estadística. Recuperado 21 de octubre de 2022, de https://psicologiaymente.com/miscelanea/prueba-kolmogorov-smirnov