Civil Engineering and Architecture 11(3): 1306-1323, 2023 http://www.hrpub.org
DOI: 10.13189/cea.2023.110316

The Modification of Chi-Square Tests for the Identification of Rainfall and River Flow Data Distribution

Heri Sulistiyono[1,*], I Wayan Yasa[1], Ery Setiawan[1], Hendra Ahyadi[2], Supardi[2], Husnulhuda Bajsair[2]

[1]Department of Civil Engineering, University of Mataram, 62 Majapahit Street, Mataram, 83125, Indonesia
[2]River Basin Organization for Nusa Tenggara I, 01, Ahmad Yani Street, Gerimax, Mataram City, West Nusa Tenggara, 83239, Indonesia

Received November 16, 2022; Revised January 4, 2023; Accepted January 28, 2023

Cite This Paper in the Following Citation Styles

(a): *[1] Heri Sulistiyono, I Wayan Yasa, Ery Setiawan, Hendra Ahyadi, Supardi, Husnulhuda Bajsair , "The Modification of Chi-Square Tests for the Identification of Rainfall and River Flow Data Distribution," Civil Engineering and Architecture, Vol. 11, No. 3, pp. 1306 - 1323, 2023. DOI: 10.13189/cea.2023.110316.*

(b): *Heri Sulistiyono, I Wayan Yasa, Ery Setiawan, Hendra Ahyadi, Supardi, Husnulhuda Bajsair (2023). The Modification of Chi-Square Tests for the Identification of Rainfall and River Flow Data Distribution. Civil Engineering and Architecture, 11(3), 1306 - 1323. DOI: 10.13189/cea.2023.110316.*

Abstract The growing population in the community has led to an increase in the need for community infrastructure. Civil engineers have to maintain the safety of the community in the design of urban areas' service infrastructure. The infrastructure must resist the load caused by extreme events, such as rainstorms and floods. Therefore, civil engineers must design the infrastructure based on the precise data parameters. Engineers obtain the precise parameters of a return period through frequency analysis. The precise parameters will produce an acceptable data distribution. Civil engineers can use the Chi-Square method to test the fitness of the data distribution type. However, the original way to get the Expected Frequency is complicated because it uses the integral solving method. This weakness causes the engineers to linger to test the distribution suitability. This article proposes a modification to ease obtaining the expected frequency in the Chi-Square test. This article demonstrates the proposal using rainfall and river flow data around the globe. The demonstration results show that the proposal is easy to implement. The method accurately identifies the type of rainfall and river flow data distribution. Among the seven stations, five groups of data follow a lognormal distribution; one group of data follows a normal distribution. One other group of data follows an exponential distribution.

Keywords Rainstorms, Data Distribution, Return Period, Chi-Square Test

1. Introduction

Design discharge is essential information in water construction design [1-3]. Researchers and civil engineers use design discharges to obtain the dimensions of water structures. Unrealistic design discharge can lead to the failure or the over-budget of construction [4]. Researchers and engineers derive discharge designs based on an acceptable type of discharge data distribution in the frequency analysis [5-8]. However, there are difficulties in recognizing the distribution. So far, the well-known methods for estimating the parameters of distribution data are the Method of Moments, the Method of Maximum Likelihood Estimation, the Method of Probability Weighted Moment, and the Richardson Method [9-11]. Those methods have some weaknesses [12], including: (1) an iterative procedure for maximum likelihood estimation is based only on a complete sample of various population parameters; (2) the solution is only based on a linear system; (3) the method cannot find the precise parameter if the system has more than one peak (4), the method cannot find the parameter if the distribution members are infinite, (5) the solution requires complex mathematical solving skills, such as integral approach [13-16].

To anticipate the weaknesses above, engineers utilize the statistical parameters of distribution to understand the type of distribution [17, 18]. The distribution parameters are the skewness coefficient and the kurtosis coefficient. However, engineers have to compare the frequency of the observed data to the expected frequency to ensure the type of distribution [19, 20]. The expected frequency in the Chi-Square Test is the theoretical frequency that we expect to occur in the data according to the type of distribution. Generally, researchers and engineers calculate the expected frequency based on the equation of the data distribution curve or the Probability Density Function (PDF) plot. Researchers and engineers can draw PDF plots using a mathematical solution approach to the theoretical distribution function [13-16]. Fig. 1 shows the PDFs of several distribution functions.

Statisticians and mathematicians use an integration approach to calculate the area under the PDF curve as the estimate of expected frequency [13-16]. However, this integration approach is complicated; consequently, the chi-square test has become unpopular among engineers.

This paper proposes a modification of Chi-Square Tests to identify the data distribution. This technique helps engineers obtain the type of rainfall and river flow data distribution without having a problem.

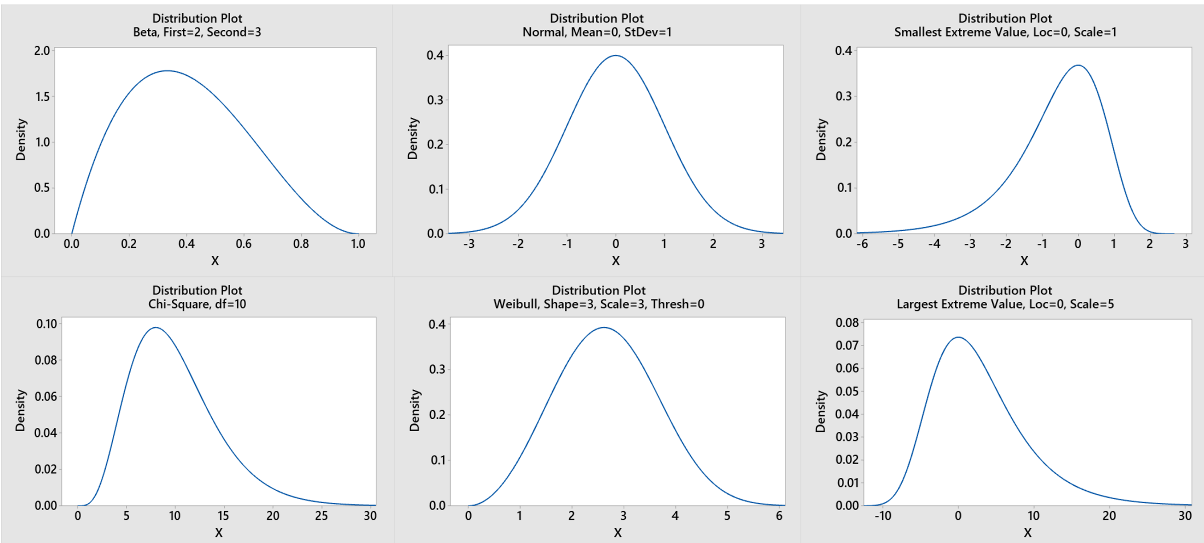

Figure 1. The Plots of Probability Density Function

2. Materials and Methods

2.1. Chi-Square

Pearson [19] developed the method of Chi-Square for the goodness of fit test. The Chi-Square tests the similarity between observed and theoretical data distribution based on the sum of squares of the difference in the data classes. Pearson gave the equation (1) below [21, 22].

$$\chi^2 = \Sigma\,(OF - EF)^2 / \Sigma\,(EF) \quad (1)$$

where: χ is Pearson's cumulative test statistic, OF is the number of observation Frequencies, EF is the expected (theoretical) frequency.

A calculated Pearson's Chi-square smaller than the critical Chi-square from the table indicates the observed data follow the expected distribution. Table 1 presents critical Chi-square. Fig. 2 shows the technique of using the Chi-square method for the identification of rainfall and river flow data distribution types.

Fig. 2 shows the Chi-square test technique to identify the distribution type of rainfall and river flow data. This technique contains three parts of analysis. The first part is to prepare data into classes according to the number of data. The second part is to obtain the expected frequency and the modification. Finally, the third part examines the acceptability of the data distribution.

Table 1. Percentage Points of the Chi-Square Distribution

Degree of Freedom	Probability of a Larger Value of χ^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.00	0.00	0.01	0.10	0.45	1.32	2.71	3.84	6.63
2	0.02	0.10	0.21	0.57	1.38	2.77	4.61	5.99	9.21
3	0.11	0.35	0.58	1.21	2.36	4.11	6.25	7.81	11.34
4	0.29	0.71	1.06	1.92	3.35	5.39	7.78	9.49	13.34
5	0.55	1.14	1.61	2.67	4.35	6.63	9.24	11.07	15.09
6	0.87	1.63	2.20	3.45	5.34	7.84	10.64	12.59	16.81
7	1.23	2.16	2.83	4.25	6.34	9.04	12.02	14.07	18.48
8	1.64	2.73	3.49	5.07	7.34	10.22	13.36	15.51	20.09
9	2.08	3.32	4.16	5.89	8.34	11.39	14.68	16.92	21.67
10	2.55	3.32	4.16	5.89	8.34	11.39	14.68	16.92	21.67

(Source: Devore, 1995)

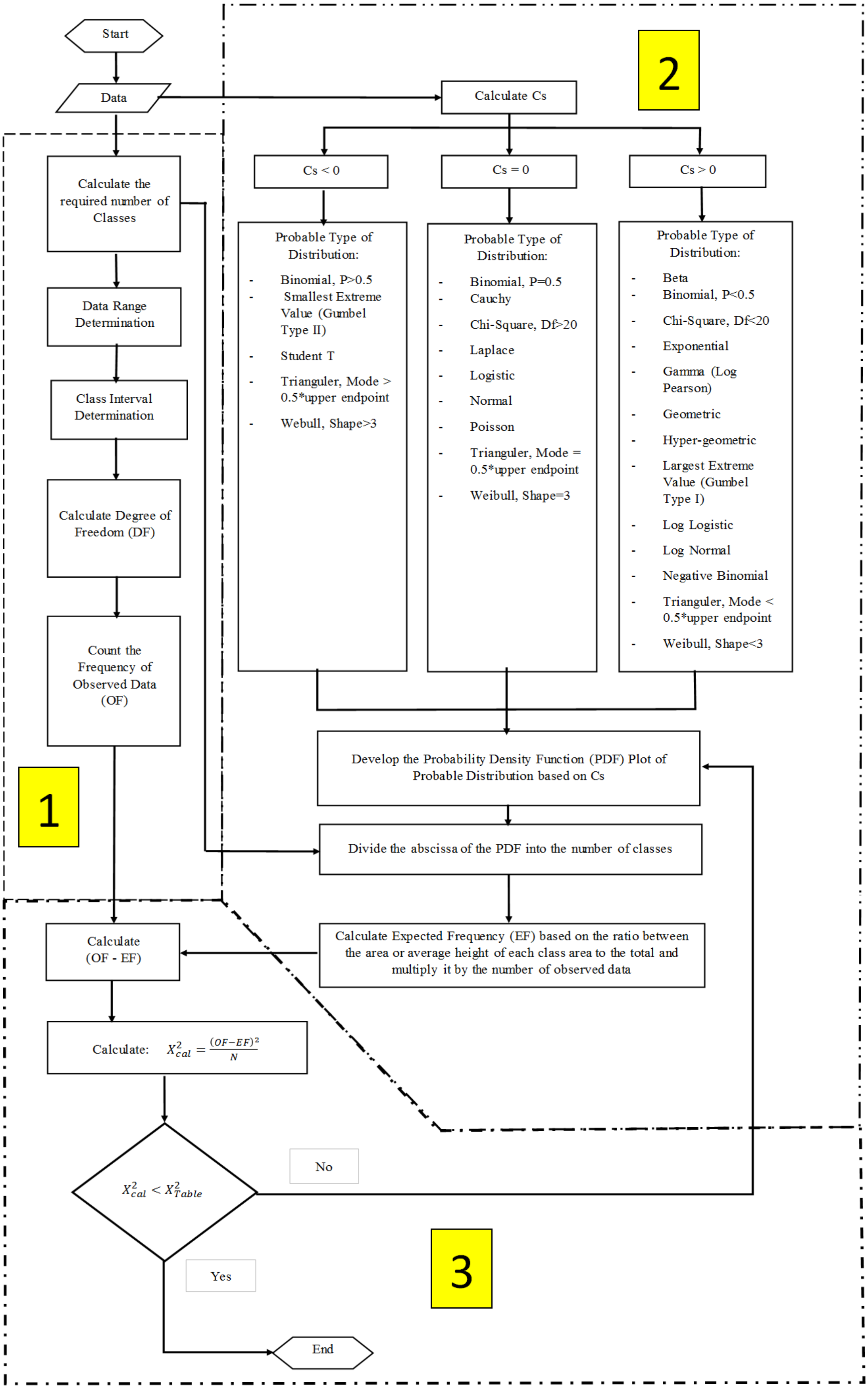

Figure 2. The Proposed Process Diagram

2.2. Chi-Square Calculations

2.2.1. The Number of Class

The Chi-Square calculation needs to group the data based on an interval to count the frequency of occurrence. The data grouped into classes used equation (2) [23, 24, 20].

$$k = 1 + 3.322 (\log10 (n)) \quad (2)$$

where k is the number of class, n is the number of data

2.2.2. The Range of Data

The range of data shows how the data spreads from the smallest to the largest. Equation (3) obtains the range

$$R = H - L \quad (3)$$

where R is the range of data, H is the largest value of data, L is the smallest value of data.

2.2.3. The Interval of Class

Equation (4) obtains the interval of class

$$I = R / k \quad (4)$$

where I is the interval of class, R is the range of data, k is the number of class.

2.2.4. The Degree of Freedom

The degree of freedom (DF) is the number of independent variables in the data set, which still independently maintain fixed parameters [25,26]. The equation of the degree of freedom for the Chi-Square calculation is

$$DF = k - 1 \quad (5)$$

where DF is the Degree of Freedom, k is the number of classes.

2.2.5. The Coefficient of Skewness

Some researchers use the Coefficient of Skewness (Cs) to identify whether the data follow a symmetrical distribution. Symmetrical distributions have Cs equal to zero. Otherwise, the distribution is asymmetrical. The two types of asymmetrical distributions are positive and negative asymmetrical distributions. Pearson gave the equation (6) to calculate the Coefficient of Skewness (Cs) [27-30].

$$Cs = [n/(n-1)(n-2)]\ \Sigma\ [(xi-\mu)/\sigma)^3] \quad (6)$$

where Cs is the Coefficient of Skewness, n is the number of data, xi is the individual data, μ is the average of data, and σ is the standard deviation of data.

2.2.6. The Modification of the Expected Frequency Calculation

The modification is replacing the original method of calculating the area of a PDF curve based on an integration [19, 13-16] with the technique of measuring the height of the curve. The solution becomes simpler. The modification contains five following steps:

Step 1: Get the PDF curve based on the Coefficient of Skewness obtained using the equation (6),

Step 2: Divide the abscissa of the curve according to the number of classes,

Step 3: Obtain the height of each class,

Step 4: Sum all of the heights of the class,

Step 5: Determine the Expected Frequency of each class by the height line of each class divided by the total height, then multiplied by the number of data.

The following section will explain every step of the modification calculations.

2.3. Case Study

This section presents several case studies to demonstrate the application of the proposed technique. Table 2 presents the list of stations used to demonstrate the proposed modification.

Table 2. The list of stations

No	Name of Station	Country	Type of Data	Start Year	End Year	# of data
1	Semongkat	Indonesia	Rainfall	1997	2021	25
2	Abaurrea Alta	Spain	Rainfall	1961	2020	60
3	Nelspruit	South Africa	Rainfall	1961	2015	55
4	Oxford	UK	Rainfall	1853	2022	170
5	Ouse	UK	River Flow	1886	2021	136
6	Nicholson	Canada	River Flow	1911	2012	102
7	Brewarrina	Australia	River Flow	1892	2021	130

Table 2 shows representative rainfall and river basin stations from around the world. The number of years varies from 25 to 170 years to represent the availability of short and long data records. The use of short and long data records in the demonstration ensured the acceptance of the proposed modified technique on the various availability of data.

Fig. 3 shows the Graphs of Data from the 7 stations.

Figure 3. The Graphs of Data

This paper will explain the Chi-Square calculations of Semongkat data in detail, while the other station's data calculations are presented in Table 7 to Table 12. Table 3 presents the Historical Rainfall Data from the Semongkat Station, Sumbawa – Indonesia.

Table 3. The Maximum Daily Rainfall Data from the Semongkat Station

No	Year	Xi	No	Year	Xi
1	1997	145.2	14	2010	86.2
2	1998	92	15	2011	62.3
3	1999	162.3	16	2012	122.3
4	2000	96.6	17	2013	100
5	2001	115.3	18	2014	113.5
6	2002	146.2	19	2015	117.5
7	2003	98.7	20	2016	92
8	2004	145.7	21	2017	145
9	2005	86.4	22	2018	67.5
10	2006	76.8	23	2019	78.3
11	2007	87.8	24	2020	221.1
12	2008	82.4	25	2021	56.4
13	2009	189.3			

2.3.1. The Demonstration of the Proposed Modification to obtain the Expected Frequency

Step 1: Getting the PDF of based on the Coefficient of Skewness obtained using the equation (6),

$$Cs = [25/(25-1)(25-2)]\ (23.40537)] = 1.060026$$

The PDF of data distribution that has Cs > 0 is shown in Fig. 4.

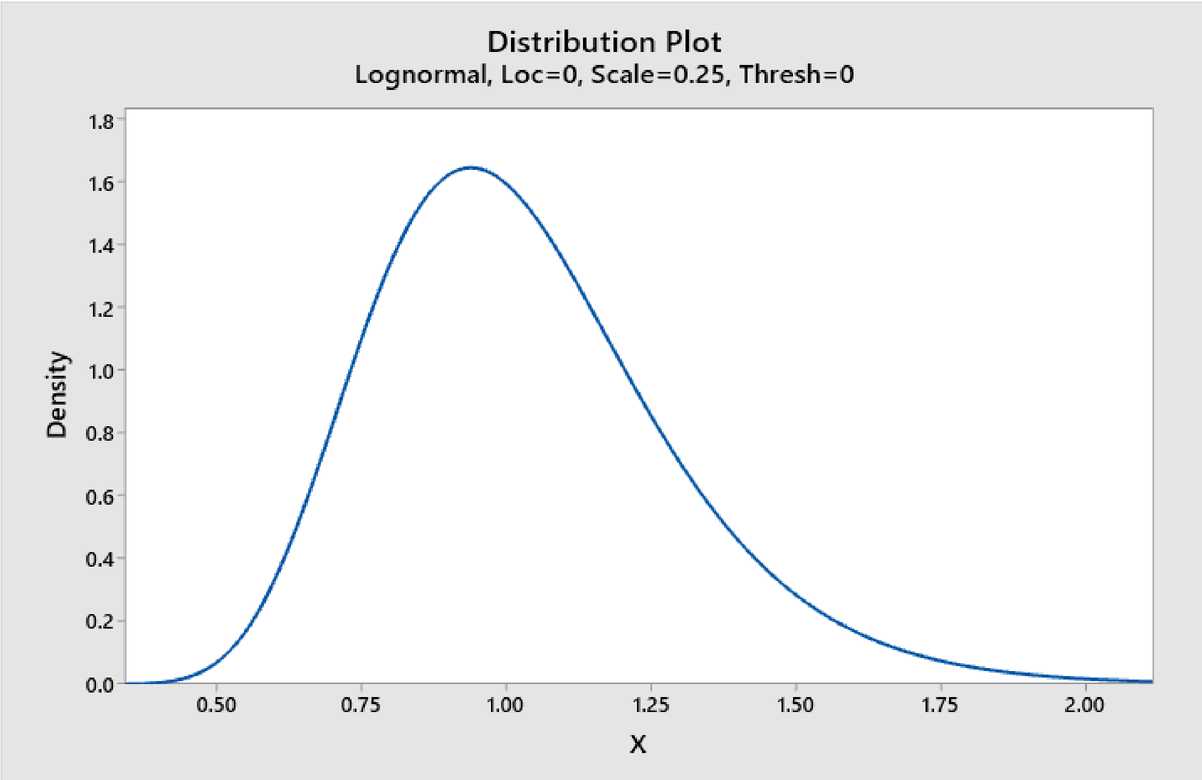

Figure 4. A PDF Plot of Lognormal Distribution

Step 2: The Division of the abscissa as the classes

Using Equation (2), the number of class (k) is

$$k = 1 + 3.322 (\log10 (25)) = 5.643957 \approx 6 \text{ classes}$$

Based on the calculation of k above, the abscissa of the PDF has to be divided into 6 parts to indicate 6 classes. Fig. 5 shows the abscissa of the PDF Plot is divided into six parts.

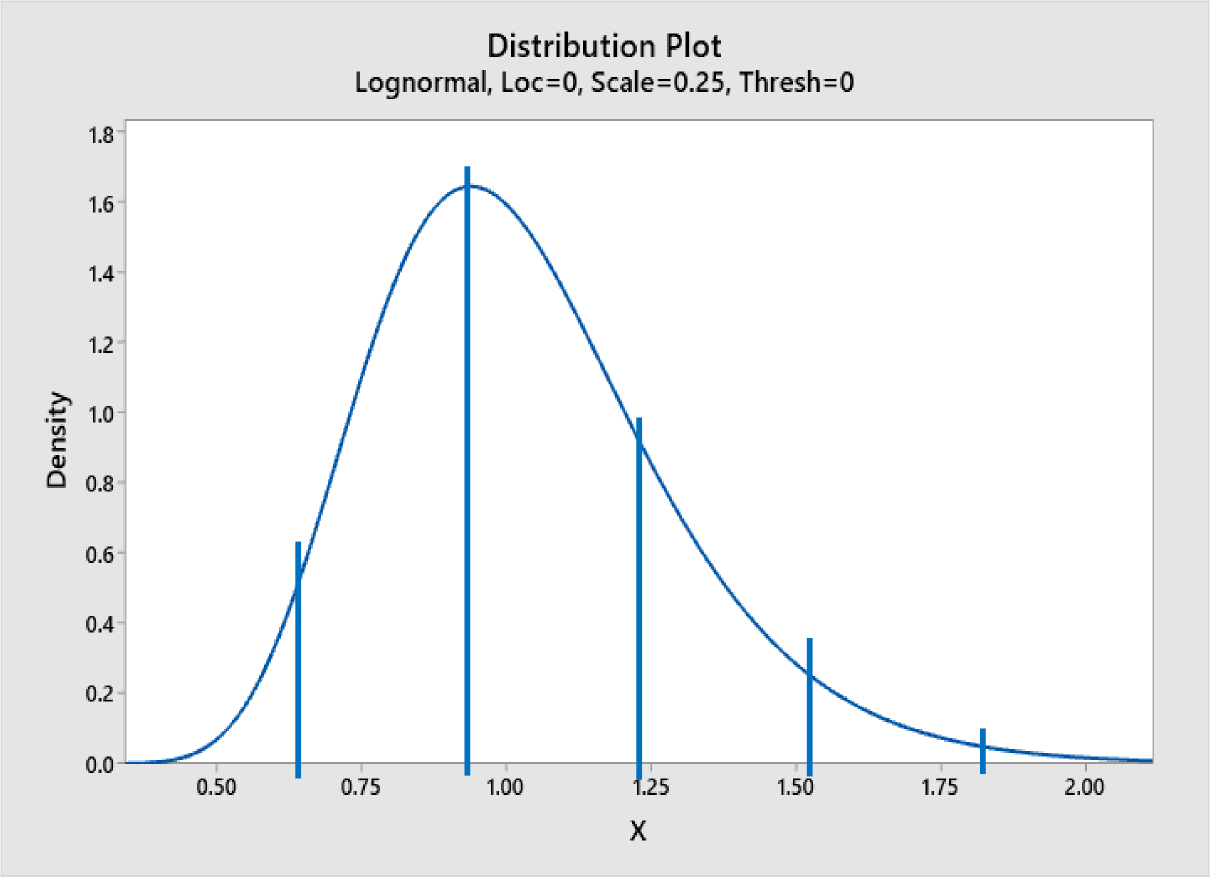

Figure 5. The six parts of PDF Plot

Using Equations (3), the range of data (R) is

$$R = 221.1 - 56.4 = 164.7$$

Using Equation (4), the interval of class is

$$I = 164.7 / 6 = 27.45$$

The first upper bond is the sum of the smallest value of data and the class interval, as shown below:

The First Upper Bond = 56.4 + 27.45 = 83.85

The first lower bond is zero. The succeeding lower bonds are the previous upper bonds.

The second lower bond = The First Upper Bond = 83.85, etc.

The succeeding upper bonds are the sum of the lower bond and the class interval.

The Second Upper Bond = 83.85 + 27.45 = 111.3, etc.

Table 4 shows the classes with the class interval.

Table 4. The Classes with the class interval

No	Lower Bond	Upper Bond	Classes
1	0	83.85	$0 < X \leq 83.85$
2	83.85	111.3	$83.85 < X \leq 111.3$
3	111.3	138.75	$111.3 < X \leq 138.75$
4	138.75	166.2	$138.75 < X \leq 166.2$
5	166.2	193.65	$166.2 < X \leq 193.65$
6	193.65	221.1	$193.65 < X \leq 221.1$

Step 3: Obtaining the high line of each class. The red lines in Fig. 6 represent the high lines

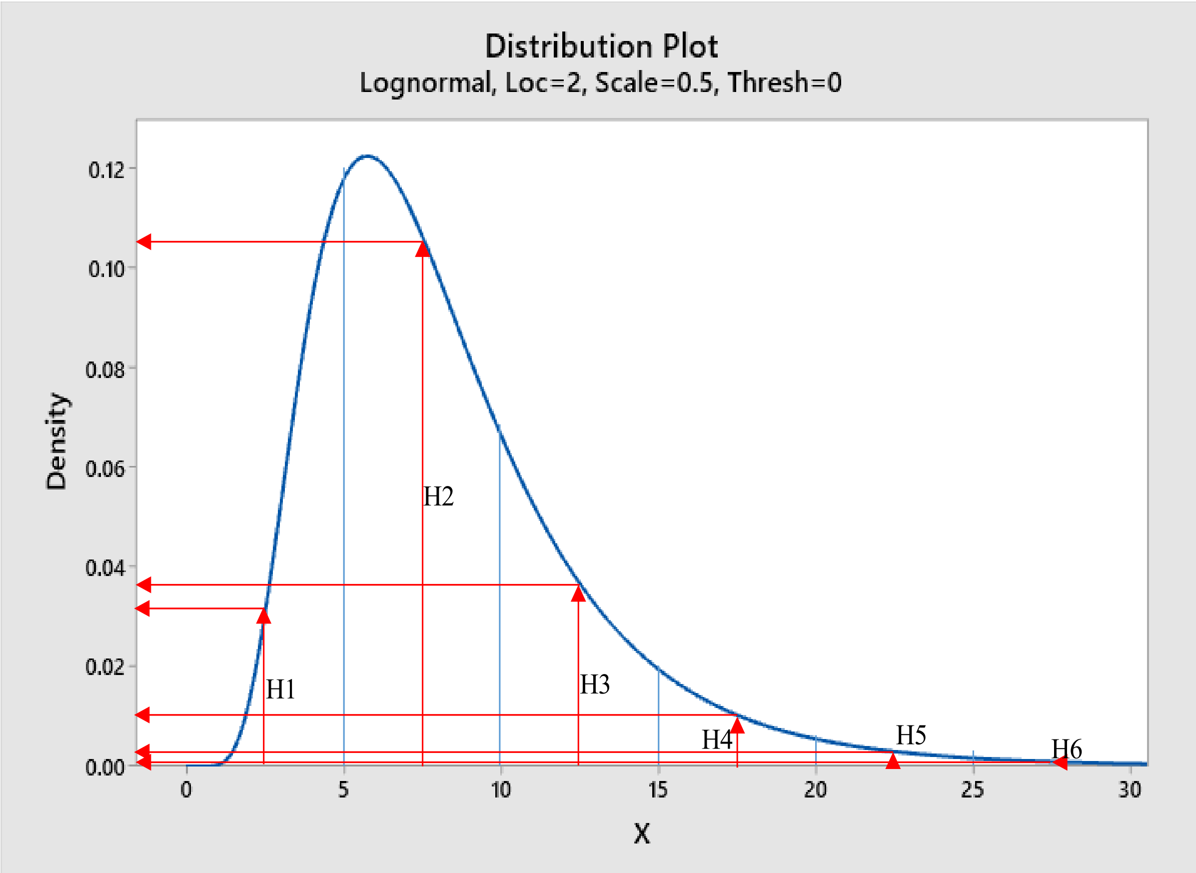

Figure 6. The height lines in the PDF data from Semongkat

Fig. 6 shows H1 is 0.031, H2 is 0.105, H3 is 0.039, H4 is 0.010, H5 is 0.001, and H6 is 0.000.

Step 4: The total of all high lines.

The sum of high lines is 0.186

Step 5: The Expected Frequency. Table 5 shows the calculation of the Expected Frequency

Table 5. The Calculation of Expected Frequency

Classes	Calculation of EF	EF
$0 < X \leq 83.85$	=(0.031/0.186)*25 = 4.16	≈ 4
$83.85 < X \leq 111.3$	=(0.105/0.186)*25 =14.11	≈ 14
$111.3 < X \leq 138.75$	=(0.039/0.186)*25 =5.24	≈ 5
$138.75 < X \leq 166.2$	=(0.010/0.186)*25 =1.34	≈ 2
$166.2 < X \leq 193.65$	=(0.001/0.186)*25 =0.13	≈ 0
$193.65 < X \leq 221.1$	=(0.000/0.186)*25 =0.00	≈ 0
Total		25

The next chi-square calculation is in Table 6.

Table 6. The Chi-Square Calculation

No	Class	OF	EF	OF-OF	$(OF-EF)^2$ /EF
1	0 < X ≤83.85	6	4	2	1
2	83.85 < X ≤ 111.3	8	14	-6	2.57
3	111.3 < X ≤138.75	4	5	-1	0.2
4	138.75 < X ≤ 166.2	5	2	3	4.5
5	166.2 < X ≤193.65	1	0	1	0
6	193.65 < X ≤ 221.1	1	0	1	0
	Total	25	25		8.27

Table 6 shows that the calculated Chi-Square is 8.27. The calculated Chi-Square has to be smaller than the critical Chi-Square.

2.3.2. Degree of Freedom

Using Equation (5), the Degree of Freedom (DF) is

$$DF = k - 1 = 6 - 1 = 5$$

Table 1 shows the critical Chi-Square value based on the degree of freedom of 5 with a significant error of 5% is 11.70. So, the calculated Chi-Square is smaller than the critical Chi-Square. The calculation shows that the rainfall data from the Semongkat follows a Lognormal Distribution.

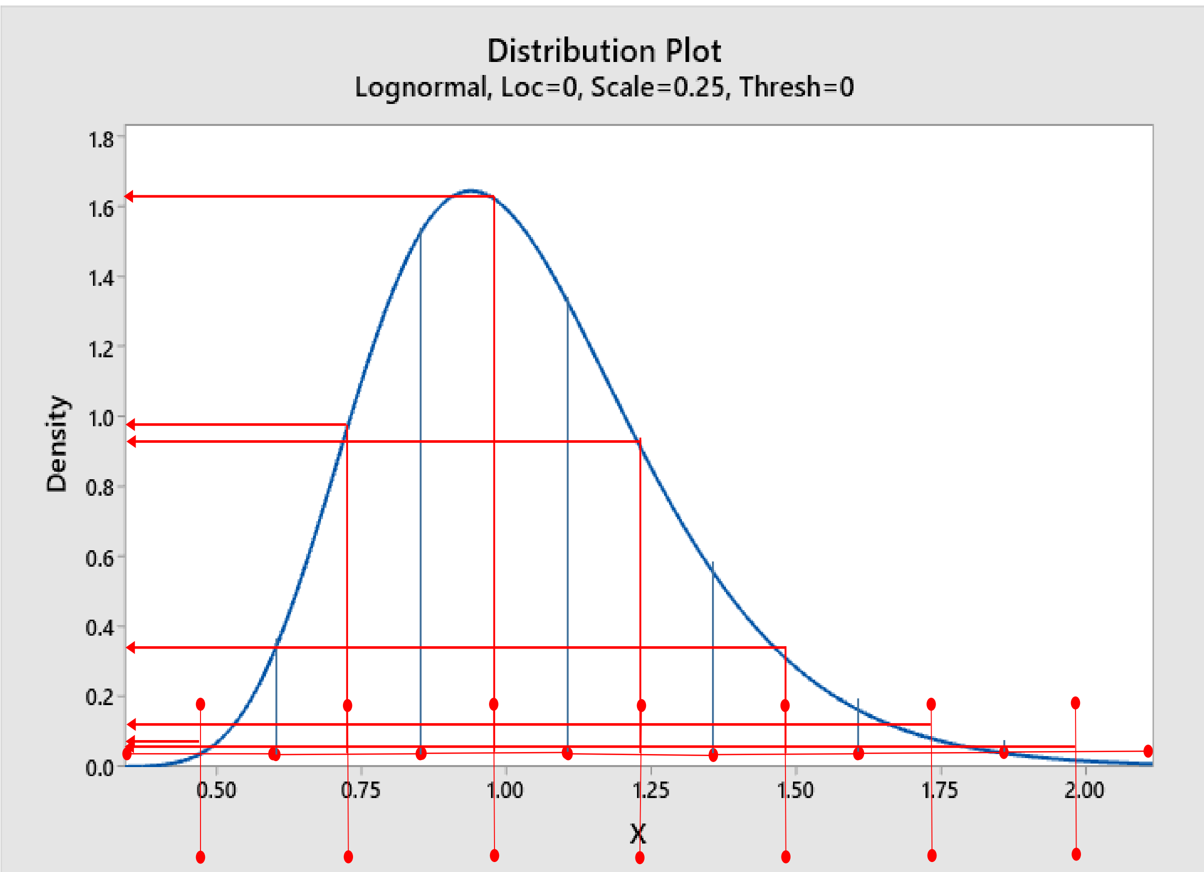

Figure 7. The height lines in the PDF data from the Abaurrea Alta

The rainfall data from Abaurrea Alta has a coefficient of skewness of 0.85. The predicted distribution is lognormal. Fig.

7 shows the PDF of a lognormal distribution. The high lines of the PDF are 0.005, 0.95, 1.62, 0.9, 0.3, 0.1, and 0.02 for H1 to H7, respectively. The sum of the high lines is 3.94. Table 7 shows the Chi-Square calculation of data from the Abaurrea Alta.

Table 7. The Chi-Square Calculation of Data From the Abaurrea Alta

No	Class	OF	EF	$(OF-EF)^2/EF$
1	$0 < X \leq 43.6$	4	0	0
2	$43.6 < X \leq 55.8$	16	15	0.07
3	$55.8 < X \leq 68.1$	22	25	0.36
4	$68.1 < X \leq 80.3$	9	14	1.79
5	$80.3 < X \leq 92.5$	4	5	0.2
6	$92.5 < X \leq 104.8$	4	1	9
7	$104.8 < X \leq 117$	1	0	0
	Total	60	60	11.41

Table 7 shows that the calculated Chi-Square is 11.41. It is smaller than the critical Chi-Square of 12.59 from the degree of freedom of 6. Therefore, the lognormal distribution was accepted for the analysis of the rainfall data from Abaurrea Alta.

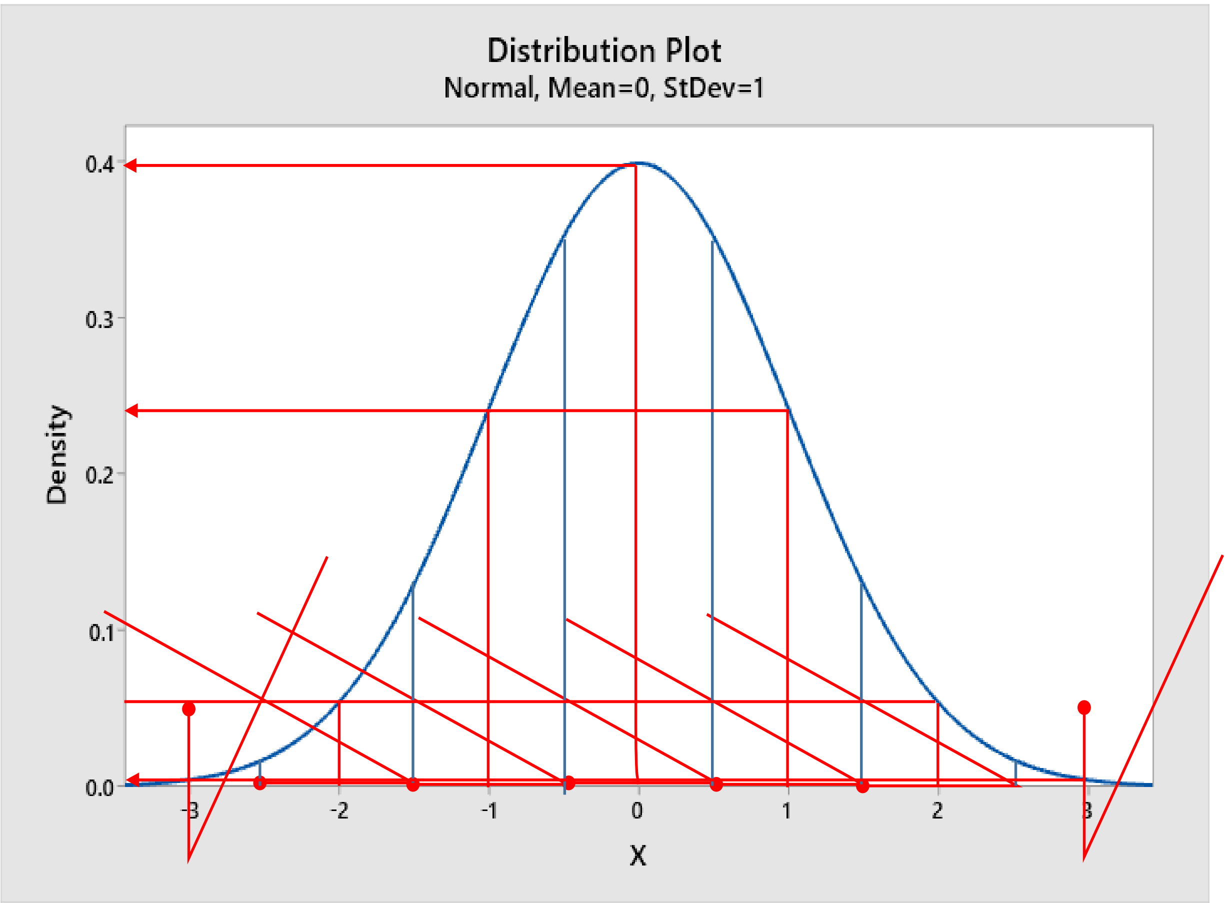

Figure 8. The height lines in the PDF data from the Nelspruit

The rainfall data from Nelspruit has a coefficient of skewness of 0.09. The predicted distribution is normal. Fig. 8 shows the PDF of a normal distribution. The high lines of the PDF are 0.0, 0.05, 0.25, 0.4, 0.25, 0.05, and 0.0 for H1 to H7, respectively. The sum of the high lines is 1.0. Table 8 shows the Chi-Square calculation of data from the Nelspruit.

Table 8. The Chi-Square Calculation of Data From the Nelspruit

No	Class	OF	EF	$(OF-EF)^2$ /EF
1	0 < X ≤18	2	0	0
2	18 < X ≤36	2	2	0
3	36 < X ≤ 54	13	14	0.071
4	54 < X ≤72	19	23	0.696
5	72 < X ≤ 90	12	14	0.286
6	90 < X ≤108	4	2	2
7	108 < X ≤ 126	3	0	0
	Total	55	55	3.05

Table 8 shows that the calculated Chi-Square is 3.05. It is smaller than the critical Chi-Square of 12.59 from the degree of freedom of 6. Therefore, the normal distribution is accepted for the analysis of the rainfall data from the Nelspruit.

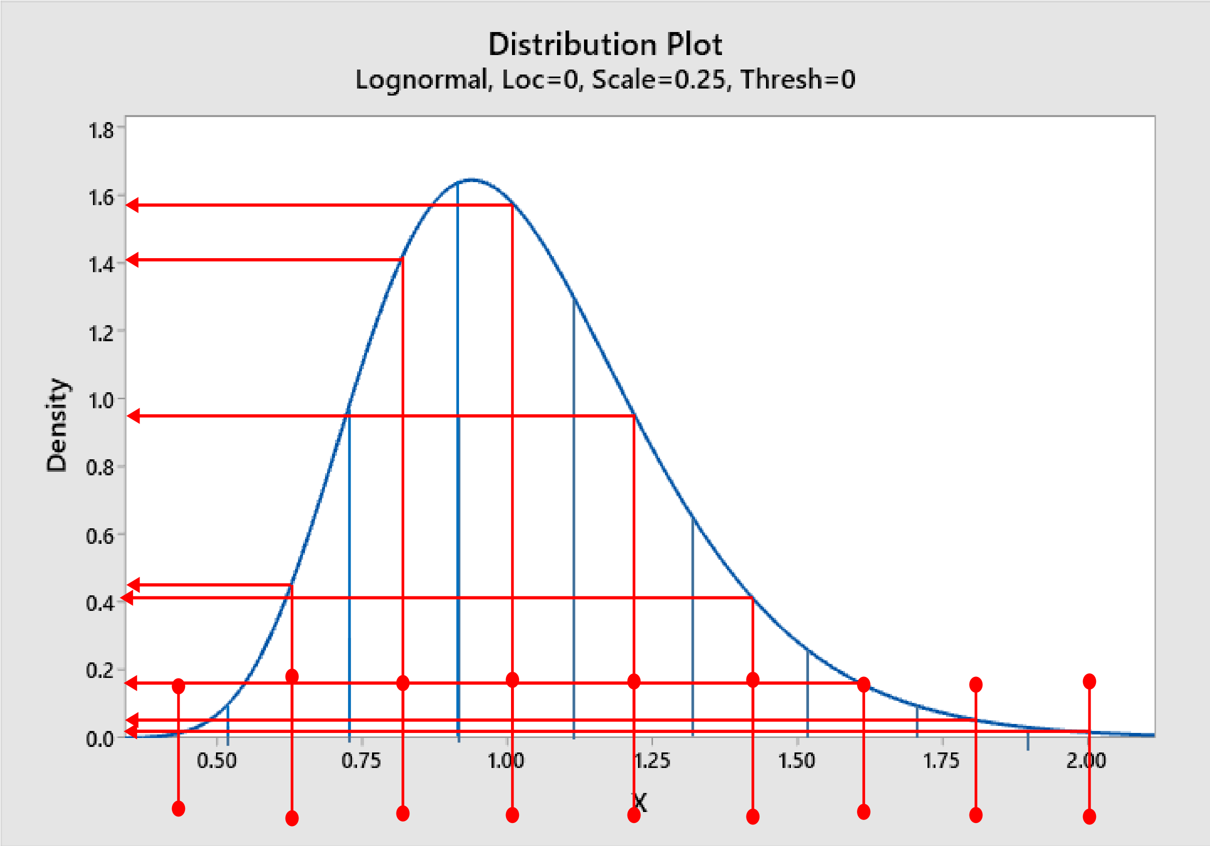

Figure 9. The height lines in the PDF data from the Oxford

The rainfall data from Oxford has a coefficient of skewness of 0.30. The predicted distribution is lognormal. Fig. 9 shows the PDF of a lognormal distribution. The high lines of the PDF are 0.0, 0.45, 1.4, 1.52, 0.96, 0.4, 0.16, 0.05, and 0.0 for H1 to H9, respectively. The sum of high lines is 4.94. Table 9 shows the Chi-Square calculation of data from the Oxford.

Table 9. The Chi-Square Calculation of Data From the Oxford

No	Class	OF	EF	$(OF-EF)^2/EF$
1	0 < X ≤71.92	7	0	0
2	71.92 < X ≤87.04	14	15	0.07
3	87.04 < X ≤ 102.17	39	48	1.69
4	102.17 < X ≤117.29	38	52	3.77
5	117.29 < X ≤ 132.41	35	33	0.12
6	132.41 < X ≤147.53	22	14	4.57
7	147.53 < X ≤ 162.66	10	6	2.67
8	162.66 < X ≤177.78	3	2	0.5
9	177.78 < X ≤192.9	2	0	0
	Total	170	170	13.38

Table 9 shows that the calculated Chi-Square is 13.38. It is smaller than the critical Chi-Square of 16.92 from the degree of freedom of 8. Therefore, the lognormal distribution is accepted for the analysis of the rainfall data from the Oxford.

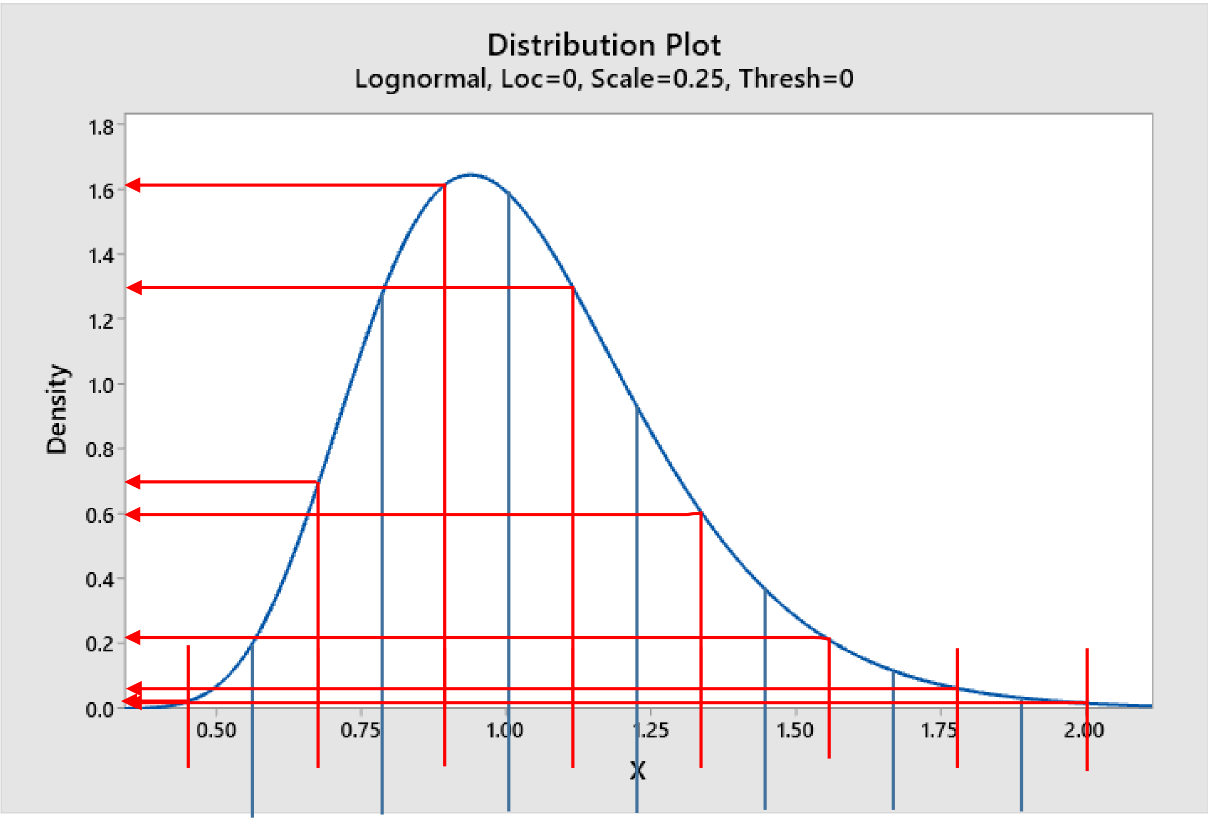

Figure 10. The height lines in the PDF data from Ouse

The river flow data from Ouse has a coefficient of skewness of 0.59. The predicted distribution is lognormal. Fig. 10 shows the PDF of a lognormal distribution. The high lines of the PDF are 0.0, 0.7, 1.6, 1.3, 0.6, 0.2, 0.05, and 0.0 for H1 to H8, respectively. The sum of high lines is 4.45. Table 10 shows the Chi-Square calculation of data from the Ouse.

Table 10. The Chi-Square Calculation of Data From Ouse

No	Class	OF	EF	(OF-EF)2 /EF
1	0 < X ≤215.5	6	0	0
2	215.5 < X ≤268	23	21	0.19
3	268< X ≤ 320.5	37	49	2.94
4	320.5 < X ≤373	32	40	1.6
5	373 < X ≤ 425.5	21	18	0.5
6	425.5 < X ≤478	8	6	0.67
7	478 < X ≤ 530.5	6	2	8
8	530.5 < X ≤583	3	0	0
	Total	136	136	13.90

Table 10 shows that the calculated Chi-Square is 13.90. It is smaller than the critical Chi-Square of 15.51 from the degree of freedom of 7. Therefore, the Lognormal distribution is accepted for the analysis of the river flow data from the Ouse.

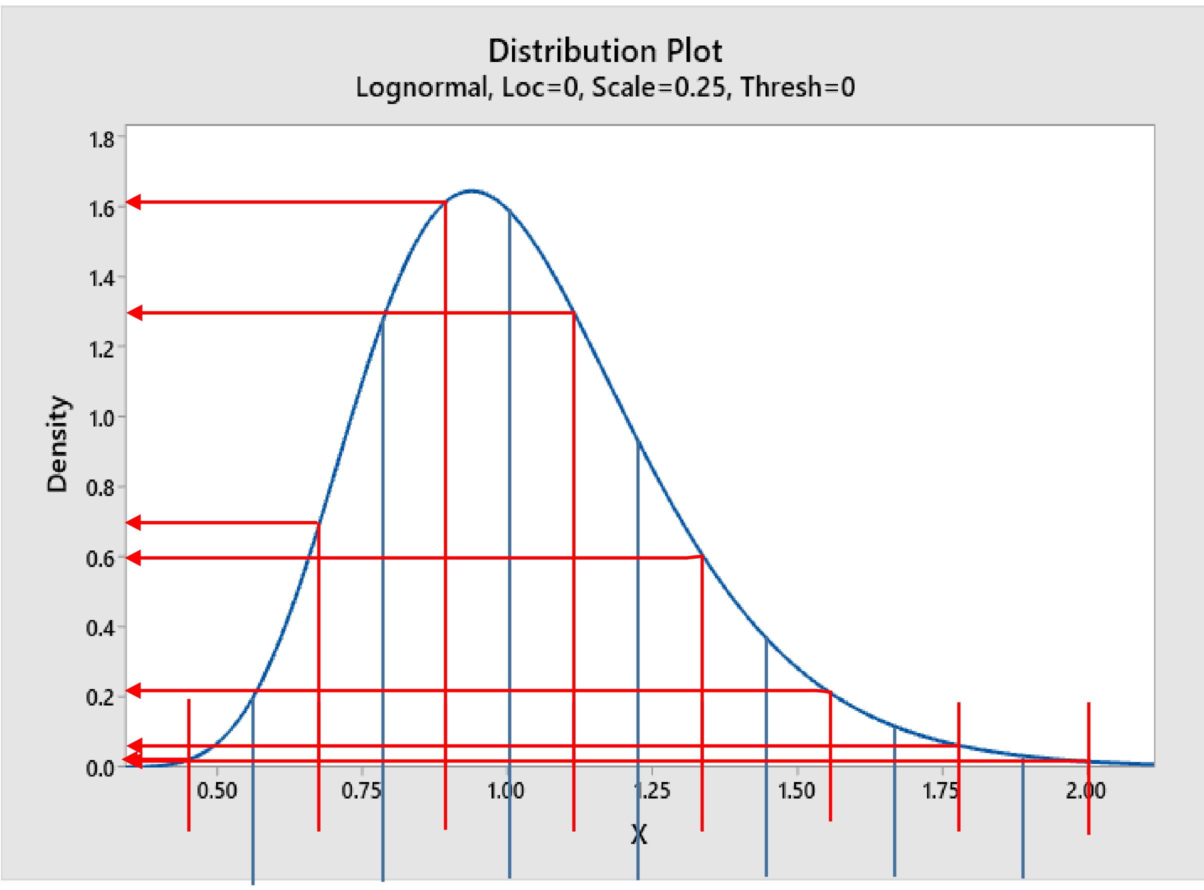

Figure 11. The height lines in the PDF data from Nicholson

The river flow data from Nicholson has a coefficient of skewness of 0.62. The predicted distribution is lognormal. Fig. 11 shows the PDF of a lognormal distribution. The high lines of the PDF are 0.0, 0.7, 1.6, 1.3, 0.6, 0.2, 0.05, and 0.0 for H1 to H8, respectively. The sum of high lines is 4.45. Table 11 shows the Chi-Square calculation of data from the Nicholson.

Table 11. The Chi-Square Calculation of Data From Nicholson

No	Class	OF	EF	(OF-EF)2 /EF
1	0 < X ≤304.5	16	0	0
2	304.5 < X ≤371	15	16	0.0625
3	371< X ≤ 437.5	25	36	3.36
4	437.5 < X ≤504	19	30	4.03
5	504 < X ≤ 570.5	15	14	0.07
6	570.5 < X ≤637	6	4	1
7	637 < X ≤ 703.5	2	1	1
8	703.5 < X ≤770	3	0	0
	Total	101	101	9.53

Table 11 shows that the calculated Chi-Square is 9.53. It is smaller than the critical Chi-Square of 15.51 from the degree of freedom of 7. Therefore, the lognormal distribution is accepted for the analysis of the river flow data from the Nicholson.

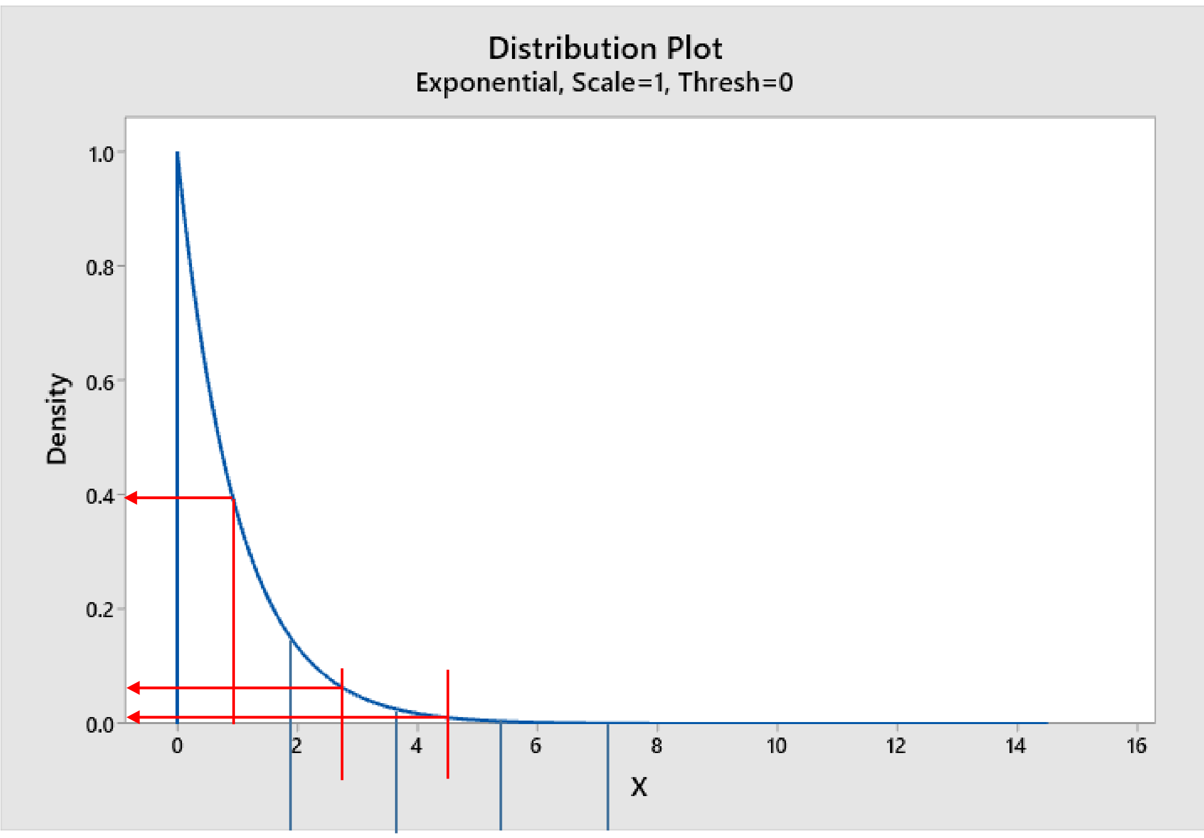

Figure 12. The height lines in the PDF data from Brewarrina

Table 12. The Chi-Square Calculation of Data From Brewarrina

No	Class	OF	EF	(OF-EF)2 /EF
1	0 < X ≤492	98	102	0.16
2	492 < X ≤982	18	23	1.09
3	982 < X ≤ 1471.98	8	5	1.8
4	1471.98 < X ≤1961.96	1	0	0
5	1961.96 < X ≤ 2451.95	3	0	0
6	2451.95 < X ≤2941.93	0	0	0
7	2941.93 < X ≤ 3431.91	0	0	0
8	3431.91 < X ≤3921.89	1	0	0
9	3921.89 < X ≤4411.875	1	0	0
	Total	130	130	3.04

The river flow data from Brewarrina has a coefficient of skewness of 3.71. The predicted distribution is exponential. Fig. 12 shows the PDF of an exponential distribution. The high lines of the PDF are 0.4, 0.09, 0.02, 0.0, 0.0, 0.2, 0.0, 0.0, and 0.0 for H1 to H9, respectively. The sum of high lines is 0.51. Table 12 shows the Chi-Square calculation of data from the Brewarrina.

Table 12 shows that the calculated Chi-Square is 3.04. It is smaller than the critical Chi-Square of 16.92 from the degree of freedom of 8. Therefore, an exponential distribution is accepted for the analysis of the river flow data from the Brewarrina.

3. Conclusions

Proper distribution of rainfall and river flow data is essential in water resource analysis. Improper data distribution will cause inaccuracies in the return period calculation. The Chi-Square test was created to identify the distribution of data. However, the earliest Chi-Square method has a weakness, namely the difficulty of mathematical calculations to obtain the area under the PDF curve. The area represents the data population. This paper has proposed a modification to the Chi-Square method. The paper has demonstrated the proposed modified technique for testing maximum daily rainfall and river flow data from several stations to represent regions worldwide. The results of this study indicate that the proposed modification simplifies the identification process of the rainfall and river flow data distribution.

Among the seven stations, five groups of data follow a lognormal distribution; one group of data follows a normal distribution, and one other group of data follows an exponential distribution.

Acknowledgements

The authors acknowledge the financial support for this research from the University of Mataram. The acknowledgment is also going to the River Basin Organization for Nusa Tenggara I in Indonesia and the Global Runoff Data Centre for providing the data.

REFERENCES

[1] Şen Z, "Flood Design Discharge and Case Studies," In: Flood Modeling, Prediction and Mitigation. Springer, Cham, 2018, pp. 70.

[2] Amatya D. M., Tian S., Marion D. A., Caldwell P., Laseter S., Youssef M. A., Grace J. M., Chescheir G. M., Panda S., Ouyang Y., Sun G., Vose J. M., "Estimates of precipitation IDF curves and design discharges for road-crossing drainage structures: Case study in four small-forested watersheds in the southeastern US". Journal of Hydrologic Engineering /Volume 26 Issue 4 - April 2021. Pp. 1-12, https://www.srs.fs.usda.gov/pubs/ja/2021/ja_2021_amatya_001.pdf. DOI: 10.1061/ (ASCE)HE.1943-5584.0002052.

[3] Flammini A., Dari J., Corradini C., Saltalippi C., and Morbidelli R.. Chapter 10 – "Areal reduction factor estimate for extreme rainfall events. Rainfall. Modeling, Measurement and Applications". Elsevier, 2022, Pages 285-306, ISBN 9780128225448,

[4] Shi L., Yi Q., Yixiu L., Xiaoyu S., Qiang L., and Ziwen L.. "Estimating the Design Flood under the Influence of Check Dams by Removing Nonstationarity from the Flood Peak Discharge Series". Hydrology Research Vol. 51 Issue 6, 2020, pp. 1261–1273. https://doi.org/10.2166/nh.2020.050

[5] Villaseñor J. A. R.. "Frequency Analyses of Natural Extreme Events: A Spreadsheets Approach". E-Book. Springer, 2021, pp. 7 – 38. https://link.springer.com/book/10.1007/978-3-030-86390-6

[6] Sharifi M., Tabatabai M. R. M., Najafabadi S. H. G..

"Determination of river design discharge (Tar River case study)". Journal Of Water And Climate Change. Vol. 12, no. 2. (Mar 2021), pp. 612 – 626. https://doi.org/10.2166/wcc.2020.278

[7] Mangukiya N. K., Mehta D. J., and Jariwala R.. "Flood frequency analysis and inundation mapping for lower Narmada basin, India". Water Practice and Technology, Vol.17, no. 2, 2022, pp. 612–622. https://doi.org/10.2166/wpt.2022.009

[8] Spuy D. V. D. and Plessis J. D.. "Flood frequency analysis – Part 1: Review of the statistical approach in South Africa". Water SA. Vol. 48 No. 2, April 2022. Pp. 110 – 119, https://doi.org/10.17159/wsa/2022.v48.i2.3848.1

[9] Kurniasari D., Warsono, Widiarti, Nabila S. U., Indryani N., Usman M., and Hadi S.. "The Comparison of the Efficiency of the Methods of Parameters Estimation for Generalized Beta of the Second Kind (Gb2) Distribution". ARPN Journal of Engineering and Applied Sciences. Vol. 17, No. 4, February 2022. Pp. 462 – 474, http://www.arpnjournals.org/jeas/research_papers/rp_2022/jeas_0222_8867.pdf

[10] Elhassanein A.. "On Statistical Properties of a New Bivariate Modified Lindley Distribution with an Application to Financial Data". Complexity, vol. 2022, Article ID 2328831, 19 pages, 2022. https://doi.org/10.1155/2022/2328831

[11] Buzaridah M. M., Ramadan D. A., El-Desouky B. S.. "Estimation of Some Lifetime Parameters of Flexible Reduced Logarithmic-Inverse Lomax Distribution under Progressive Type-II Censored Data". Journal of Mathematics, vol. 2022, 13 pages, Article ID 1690458, https://doi.org/10.1155/2022/1690458

[12] Harter H. L.. "Some Optimization Problems In Parameter Estimation, Optimizing Methods in Statistics", Academic Press, 1971, Pages 33-62, ISBN 9780126045505, https://doi.org/10.1016/B978-0-12-604550-5.50007-9

[13] Wu Y. and Chen P.. "Statistical Modeling in Biomedical Engineering". Encyclopedia of Biomedical Engineering. Elsevier, Pages 164-176, 2019. ISBN 978012805 1443. https://doi.org/10.1016/B978-0-12-801238-3.99970-7

[14] He X. S., Fan Q. W., and Yang X. S.. "Probability theory for analyzing nature-inspired algorithms". Nature-Inspired Computation and Swarm Intelligence. Academic Press, Pages 77-88, 2020. ISBN 9780128197141. https://doi.org/10.1016/B978-0-12-819714-1.00016-6

[15] Ramadan A., Ebeed M., Kamel S., and Nasrat L.. "Optimal power flow for distribution systems with uncertainty". Uncertainties in Modern Power Systems. Academic Press, Science Direct, Pages 145-162, 2021. ISBN 9780128204917,https://doi.org/10.1016/B978-0-12-820491-7.00005-0

[16] Fang L. and Zhang Y.. "Probability density function analysis based on logistic regression model". International Journal of Circuits. Systems and Signal Processing. Vol. 16. 2022. Pp: 60 – 69, https://npublications.com/journals/circuitssystemssignal/2022/a182005-009(2022).pdf DOI: 10.46300/9106.2022.16.9.

[17] Cohn T. A., Barth N. A., England J. F., Faber B. A., Mason R. R., and Stedinger J. R.. "Evaluation of recommended revisions to Bulletin 17B". USGS Report. USGS Publications Warehouse. 2019. 141 pages, DOI: 10.3133/ofr20171064.

[18] England J. F. Jr., Cohn T. A., Faber B. A., Stedinger J. R., Thomas W. O. Jr., Veilleux A. G., Kiang J. E., and Mason R. R. Jr. "Guidelines for determining flood flow frequency — Bulletin 17C", ver. 1.1, U.S. Geological Survey Techniques and Methods, book 4, chap. B5, 2019, 148 pages, https://doi.org/10.3133/tm4B5

[19] Pearson K.. "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling". Philosophical Magazine. Series, 5. 50 (302),1900. Pp.157–175, doi:10.1080/1478644000946 3897.

[20] Evita B. D.. Retention Pool Planning As an Alternative Flood Control In The Ranjok River (*Perencanaan Kolam Retensi Sebagai Alternatif Pengendali Banjir Di Sungai Ranjok*)". Thesis. Civil Engineering Bachelor Degree Program. Faculty of Engineering. University of Mataram press, 2022. 110 pages.

[21] Schober P. and Vetter T.. "Chi-square Tests in Medical Research". Statistical Minute: Statistical Minute. Anesthesia & Analgesia: November 2019 - Volume 129 - Issue 5 – page 1193, 2019. https://journals.lww.com/anesthesia-analgesia/fulltext/2019/11000/chi_square_tests_in_medical_research.3.aspx DOI: 10.1213/ANE.0000000000004410.

[22] Turhan N. S.. "Karl Pearson's Chi-Square Tests". Academic Journals. Article Number - 72074A964789. Vol.15 (9), pp. 575-580, September 2020. acfc8be54516124e51d45e0bf559812f8b66.pdf(semanticscholar.org)

[23] Sturges H. A.. "The choice of a class interval". Journal of the American Statistical Association. Vol. 21 (153), 1926, pp. 65–66,https://doi.org/10.1080/01621459.1926.10502161.

[24] Scott D. W.. "Sturges' rule". Wiley Interdisciplinary Reviews: Computational Statistics, Vol. 1(3), 2009, pp. 303 – 306, https://www.deepdyve.com/lp/wiley/sturges-rule-XlDTDSe4IQ, DOI: 10.1002/wics.35.

[25] Brereton R. G.. "Determining the significance of individual factors for orthogonal designs", Journal of Chemometrics, vol. 33, issue. 6, 2019, pp. 1–5, https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/epdf/10.1002/cem.3124

[26] Mohr D. L., Wilson W. J., and Freund R. J.. "Inferences for Two Populations", Statistical Methods (Fourth Edition), Academic Press, Pages 201-241, 2022. ISBN 9780128230435.https://www.coursehero.com/file/85444717/Inferences-for-Two-Populationspdf/

[27] Devore J. L.. "Probability and Statistics for Engineering and the Sciences". Duxbury Press. USA. 1995. 768 pages, https://www.amazon.com/Probability-Statistics-Engineering-Sciences-Devore/dp/0538733527?asin=0538733527&revisionId=&format=4&depth=1

[28] Rousseau R.. "Skewness for journal citation curves". Conference: STI-Conference at Leiden, the Netherlands. 2014. Researchgate. https://www.researchgate.net/publication/292404753_Skewness_for_journal_citation_curves

[29] Gatfaoui H., Nagot I., and Peretti P. D.. "Are Critical Slowing Down Indicators Useful to Detect Financial Crises?" Systemic Risk Tomography, Elsevier, 2017. Pages 73-93, ISBN 9781785480850, https://doi.org/10.1016/B978-1-78548-085-0.50003-0

[30] Romeo G.. "Elements of Numerical Mathematical Economics with Excel", Academic Press, Science Direct, 2020., Pages 763-795, ISBN 9780128176481, https://www.amazon.com/Elements-Numerical-Mathematical-Economics-Excel/dp/0128176482