Journal of Building Technology, 2023, Volume 5, Issue 2
DOI: 10.32629/jbt.v5i2.1278
ISSN Online: 2717-5103
ISSN Print: 2705-1390

Determination of the Probability Distribution of Concrete Mixing Components

Mariana Suárez Pereira, Fiamma Dubraska Peña Fabiani

Lisandro Alvarado Central Western University, Venezuela.

Abstract: The purpose of this research is to describe the statistical behavior and determine the probability distributions that best fit each component of conventional concrete mixtures, which are water, cement, fine aggregate and coarse aggregate, designed by the Porrero and ACI methods for compressive strengths between 250 and 280 kg/cm^2 and nominal maximum size of 1.0 inch, settlements up to 6.0 inches, Portland cement type I or type CPCA1 and natural sand. For this purpose, it was necessary to create a database by consulting the theses prepared in the Civil Engineering Department of the Lisandro Alvarado Central Western University, and 228 theses were reviewed, of which 66 complied with the defined scope. The descriptive statistical analysis reported low and intermediate dispersion, and it was concluded that the arithmetic mean obtained represents the data set by variable and the distributions obtained for water, cement, fine aggregate and coarse aggregate were Gen. Gamma (4P), Gen. Extreme Value, Weibull (3P) and Frechet respectively for the ACI method and for Porrero Hypersecant, Log-Pearson 3, Johnson SB and Chi-Square (2P) as appropriate.

Key words: concrete mixes; concrete components; ACI manual; Porrero manual

1. Introduction

Concrete is a material that has been used with great frequency worldwide for the construction projects. Concrete quality is a factor of great importance that depends on many variables, and requires the implementation of multiple controls and tests in order to guarantee it. The quality of concrete is determined by the aggregate materials and the proportions of these components, such as cement, water, fine aggregate and coarse aggregate, which directly influence the properties of the material and have effects not only on the appearance of the final product, but also on the workability and consistency at the plastic state, as well as on the durability, strength, elastic and thermal properties, volumetric changes and unit weight of the concrete after hardening. These concrete components constitute elements that can be modified in their quantities, representing a variable that can be considered random and describing a probabilistic behavior.

Although the dosages of the concrete components are obtained from a deterministic model such as the ACI method [1] and Porrero Structural Concrete Manual [2], according to the standard mixture database, behavior suitable for probability distribution can be captured for each component of the mixture. The importance of these probability distributions is that they serve as a basis for the development of models that simulate the concrete strength without the implementation of physical tests, and in turn open the way to new research in the field of statistics that can study other variables of this material. The present investigation was based on the data collection reflected in the degree theses elaborated in the Civil

Engineering Department of the Lisandro Alvarado Central Western University, Barquisimeto, Lara State, concerning the proportions of each concrete component used in the tests.

2. Development

2.1 Background

Regarding the probability distributions associated with the components of conventional concrete mix designs, some works are presented below, which are relevant to the research purpose to be developed:

García [3], analyzed the behavior of new statistical distributions never used in hydrology, for which he carried out a study with data of maximum annual rainfall from 53 meteorological stations in the province of Badajoz, comparing the fit goodness of the Kolmogorov-Smirnov and Anderson-Darling tests, and the flows reflected by the classical statistical distributions with other distributions more recently applied in other fields of science. He concluded that the new probability distributions, Dagum, Burr, Log-Logistic (3P), Pearson 5 (3P) and Frechet (3P) fit better statistically according to goodness-of-fittest than Gumbel, Log-Pearson (3P) and SQRT-ET max.

For his part, Cerón [4], proposed a solution that would save the country a considerable amount of money in the investment of maintenance and rehabilitation of infrastructure works. For this purpose, he conducted probability analysis on high-strength concrete to determine the distribution of controlled simple concrete, and developed a probability method from the study of HPC compressive strength test results. He concluded that such methodology provides results with more complete and realistic information about the strength level of the mix.

In another study, an analysis of the influence of some components in the production of high-strength concrete was carried out by using a database that made it possible to evaluate the different dosages for various w/c ratios and compressive strength [4]. The database used consisted of 487 concrete dosages from the collection of scientific journal articles from university, technical and scientific study centers, conference publications, doctoral and master's theses.

These studies essentially demonstrate the feasibility of conducting analyses related to the objectives proposed in this work, such as describing statistical behavior and determining the probability distribution most suitable for each component of traditional concrete mixtures.

2.2 Theoretical references

2.2.1 Concrete and its components

Venezuelan Standard COVENIN 221:2001 [5] defines concrete as "a mixture made up of binders, inert aggregates and water in adequate proportions to obtain pre-set strengths".Depending on the quantities and characteristics of the materials used, the properties and quality of the concrete will vary; for example, the water/cement ratio is directly related to strength and, together with the cement dosage, to workability. Thanks to its versatility, it is possible to obtain different plasticities, strengths and appearances by using different components or by varying their possible proportions, in order to meet the requirements of each project [2].

In this research work, the variables to be studied were the components of conventional concrete, which are water, cement, fine aggregate and coarse aggregate, according to the ACI and Porrero methods. In the case of cement, mixes made with Portland type I cement and cement with CPCA1 type additions were considered.

2.2.2 Mix design

According to Porrero [2], mix design is the process of determining the proportions of concrete constituents that are most suitable for obtaining the required quality, i.e. the desired behaviour in the plastic and hardened states. There are numerous methods for designing concrete mixtures, which vary according to the complexity of the variables they deal with and the relationships they establish; the most widely used in Venezuela are those of Porrero and ACI.

The method proposed by Porrero has been tested in laboratories and concrete plants and designed for the use of poorly controlled aggregates and for the calculation, preparation and control by professionals with little experience. It takes into account factors such as the site or environmental conditions, the type of work or part of the structure and its dimensions, the type of aggregate and the type of cement. A characteristic quality of this method is that it uses a ratio between fine and coarse aggregate, i.e., it handles the grain size of the combined aggregate [6].

The variables considered in this method include water/cement ratio, cement dosage, workability, and strength, which can improve concrete when combined with aggregates limited by particle size. Additionally, correction factors related to the maximum size and type of aggregate are taken into account for the calculation of the water/cement ratio and cement dosage. It is valid for concrete with slumps between 1 and 6, and compressive strengths between 180 and 430 kg/cm^2.

The method established by the American Concrete Institute (ACI) calculates the absolute volume of concrete components based on a logical sequence applicable to the available material properties. It determines the required strength and sets the water/cement ratio to ensure the durability and strength of the concrete. It is used for normal weight concrete and requires knowledge of strength, maximum aggregate size, maximum water cement ratio, cement content, admixtures, and air content [1].

3. Methodology

For the purpose of conducting research, the research subjects include all concrete mix designs for strengths between 250 - 280 kg/cm^2 up to 6" of slump, 1" crushed stone, natural river sand and Portland cement and CPCA1 type, according to the Porrero and ACI methods. The sample was made up by simple random probability sampling corresponding to infinite population and unknown variance. Thus, the sample size is as follows:

$$n = \left(\frac{Z_{\frac{\alpha}{2}} * pq}{e} \right)^2 \qquad (1)$$

Considering a confidence level of 0.95 and for an area under the curve of 0.025, it is obtained that $Z\ \alpha/2 = \pm\ 1.96$. Under a binomial distribution, the probabilities of success p and failure q, are $p = q = 0.5$. The estimation error of 0.05 was taken as the researcher's criterion, since it generally fluctuates between 0.04 and 0.06, according to the sampling theory. Substituting the values in the previous expression we obtain:

$$n = \left(\frac{1,96 * 0,5 * 0,5}{0,05} \right)^2 = 96 \qquad (2)$$

However, due to the limited availability of Special Degree Works (SDW) carried out at the Civil Engineering Department of the Lisandro Alvarado Central Western University, whose data complied with the characteristics defined above, the sample sizes for the Porrero method were 76 data points and for the ACI method, 30 data points for each component of each mix design. Although the data per component for the ACI method are few, they represent the population, since according to the sampling theory, a sample is representative when the number of observed elements of the population is greater than or equal to 30. In this research work, the variables under study are the quantities of water, cement, fine aggregate and coarse aggregate for each method of conventional concrete mix design.

3.1 Collection and processing of information

Data on the design of concrete mix proportions with strengths ranging from 250 to 280 kg/cm^2 in SDW were collected, and the component dosages of different concrete mix proportions in selected projects were compiled. Subsequently, the sample under study was formed and a descriptive statistical analysis was performed to estimate the measures of central tendency, dispersion and graphs for each component of concrete mix designs for strengths between 250 - 280 kg/cm^2. By recording the data collected in a Microsoft Excel spreadsheet, the tables, frequency polygons, and histograms are drew, and the mean, median, mode, variance, standard deviation, typical error, coefficient of variation, asymmetry, and kurtosis of descriptive statistical data are calculated.

The probability distribution that best fits the behavior of each component of concrete mix designs was determined by using software that adjusts the data to probability distributions EasyFit Version 5.6; introduce and analyze the data in software to obtain adjusted distributions and select the probability distribution that best fits each of the components by mix design method following the methodology applied by Garcia [3]:

- The data are entered into the program in order from smallest to largest.
- The adjustment of the different distributions is performed with the program.
- The program provided the most suitable distribution for input data through Kolmogorov-Smirnov, Anderson-Darling, and Chi-Square fitting tests. For each distribution, it offers the density function and the cumulative distribution function among other parameters.
- The result obtained is a table in which the fit order of the distributions is indicated, starting with one from best to worst fit, based on the goodness-of-fit tests indicated.
- A table is prepared in Microsoft Excel with the results of the goodness-of-fit tests and the ranks obtained are weighted to select the five distributions that best fit, taking into account the two tests simultaneously.
- Finally, for the first 5 distributions, their mean is calculated with the specified confidence interval, which is 0.95 in the present research work, and these values are compared with the mean of the 5 selected models, thus estimating their deviation with respect to the group mean, which generates a new order of fit. The distribution with the smallest deviation will be the one that best fits the input data.

4. Results

4.1 Descriptive statistical analysis

4.1.1 ACI method for conventional concrete mix design

The corresponding summary statistics are shown in Table 1. An average or arithmetic mean is a group characteristic and not an individual characteristic, therefore, in this table, the group characteristics of variable data for water, cement, fine aggregate and coarse aggregate in this method are: 196.9, 325.39, 916.13 and 902.51, representing the value around which the highest concentration of data for each variable revolves. The position values or medians of each variable are 195, 325, 952.62 and 874.75 respectively, indicating that 50% of each data set is less than these values and the other 50% is greater, dividing it into two equal parts. The most repeated value in each of the variables under study is 195, 325, 979.26 and 826.34 accordingly.

The standard deviation values for each variable are 7.96, 19.40, 100.81 and 75.64, respectively, which serve to measure whether there is low, intermediate or high dispersion. The coefficients of variation for water, cement and coarse aggregate are 0.04, 0.06 and 0.08 respectively, which indicates that these variables behave with a low dispersion, that is, for each of them their arithmetic mean represents the group, while for the variable fine aggregate an intermediate dispersion of 0.11 is observed with a tendency towards a low dispersion, for this reason we could consider that its mean also represents the group of data. Consequently, the value ranges for each of the variables are the mean plus or minus its standard deviation. These ranges are: for water 196.9 ± 7.96 lt, cement 325.39 ± 19.40 kg, fine aggregate 916.13 ± 100.81 kg and for coarse aggregate 902.51 ± 75.64 kg.

Table 1. Summary of descriptive statistics for the ACI method. Source: authors.

ESTADÍSTICO	VARIABLES			
	Agua (lt)	Cemento (Kg)	Agregado Fino (Kg)	Agregado Grueso (Kg)
Media	196,9	325,393	916,129333	902,511667
Mediana	195	325	952,62	874,75
Moda	195	325	979,26	826,34
Desviación estándar	7,96411261	19,3956592	100,810663	75,6389341
Varianza de la muestra	63,4270897	376,191594	10162,7898	5721,24836
Coeficiente de variación	0,0404475	0,05960687	0,11003977	0,08380937
Curtosis	7,92109398	2,18278064	-1,0756143	-1,21717641
Coeficiente de asimetría	2,36935337	1,27348938	-0,37061969	0,5440219
Rango	44,96	86	343,58	250,07
Mínimo	180	300	751,15	789,93
Máximo	224,96	386	1094,73	1040

The kurtosis of the water and cement variables are 7.92 and 2.18 respectively, because these values are greater than zero, their curves are leptokurtic, indicating that the data is often concentrated around the central value. For the fine and coarse aggregate variables, the kurtosis are -1.08 and -1.22, because they are less than zero, their curves are platykurtic, indicating that their behavior is opposite to the other two, that is, their data is dispersed.

For variables such as water, cement, and coarse aggregate, the asymmetry is 2.37, 1.27, and 0.54, with values greater than zero, indicating a positive asymmetry where values tend to cluster to the left of the arithmetic mean, i.e., the mode and median are lower than the arithmetic mean, as shown in Table 1. For the fine aggregate variable, the asymmetry coefficient is -0.37, and a value less than zero indicates a negative asymmetry where the values are clustered to the right of the mean, meaning that the mode and median are greater than the arithmetic mean of the variable. Figures 1 and 2 show a unimodal histogram because its bar graph is higher than the other bar graphs, where the absolute frequencies of one class are 188-196 and 315-330, respectively.

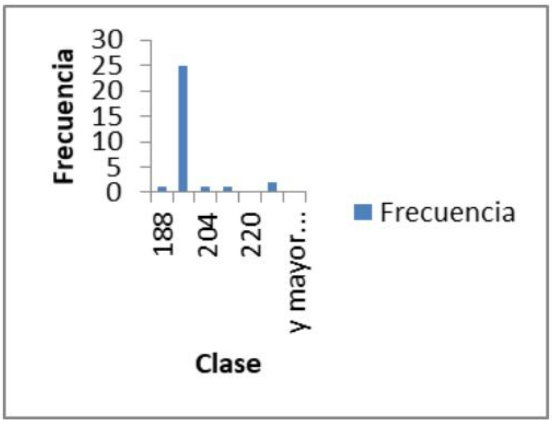

Figure 1. Absolute frequency histogram for the variable water by the ACI method. Source: authors.

In addition, it is skewed, showing a greater concentration of data to the left of the mean, i.e., the variable water and cement has an asymmetric behavior that agrees with the above mentioned for these variables. The class with the highest frequency of occurrence contains the mean, median and mode values for both variables.

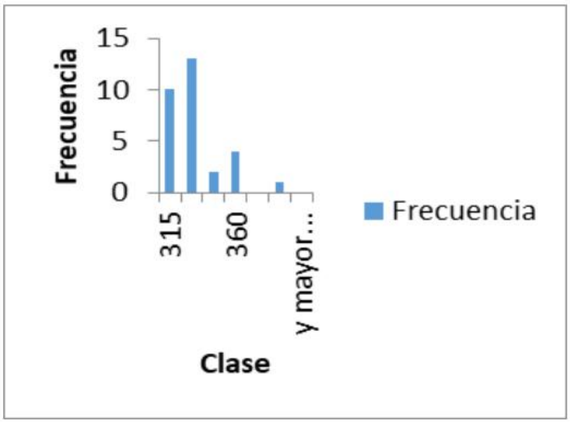

Figure 2. Absolute frequency histogram for the cement variable for the ACI method. Source: authors.

Figure 3 indicates a unimodal histogram, since it has one bar higher than the others, that is, one of the classes has a higher absolute frequency.

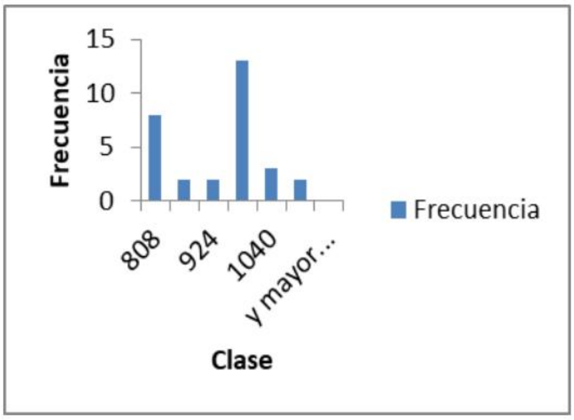

Figure 3. Absolute frequency histogram for the fine aggregate variable for the ACI method. Source: authors.

It is considered biased, meaning that the fine aggregate variable exhibits asymmetric behavior, with one of the two halves being larger than the other, and most of the data is concentrated on the right side, which is consistent with the dispersion previously pointed out by this variable. The class with the highest absolute frequency is 924-982, which includes the median and mode, but does not include the median. Figure 4 indicates a bimodal histogram, since it has two equal higher bars, that is, two classes with equal absolute frequency. In addition, it is skewed, which shows that the variable coarse aggregate has an asymmetric behaviour, with the data concentrated on the left side, which is consistent with the dispersion previously carried out for the same. The classes with the highest frequency of data occurrence are 788-830 and 872-914, with the latter containing the mean and median values, while the mode is contained in the first class.

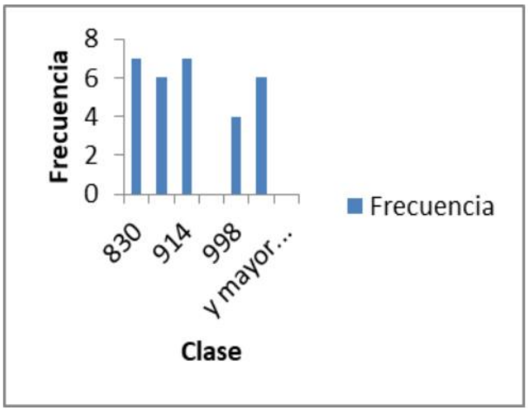

Figure 4. Absolute frequency histogram for the coarse aggregate variable for the ACI method. Source: authors.

4.1.2 Structural concrete manual or Porrero's method

Table 2 shows the summary statistics corresponding to the data recorded for the variables water, cement, fine aggregate, and coarse aggregate for the Porrero method, for a total of 76 data.

Table 2. Descriptive statistical summary of the Porrero method. Source: authors.

ESTADÍSTICO	VARIABLES			
	Agua (lt)	Cemento (Kg)	Agregado Fino (Kg)	Agregado Grueso (Kg)
Media	200,048132	389,5895	916,337276	845,062908
Mediana	202,345	384,11	894	841,525
Moda	210	340	1007,9	782,6
Desviación estándar	19,2776083	46,7371563	116,277188	93,067619
Varianza de la muestra	371,626183	2184,36178	13520,3844	8661,58171
Coeficiente de variación	0,09636485	0,11996513	0,12689344	0,110131
Curtosis	1,44265081	-0,32753058	-0,01933324	-0,10731227
Coeficiente de asimetría	0,21661204	0,22650782	0,7194439	-0,15451937
Rango	110,87	200	476,5	477,32
Mínimo	149,13	297	744,8	562,64
Máximo	260	497	1221,3	1039,96

The arithmetic mean grouping characteristics of variable data for water, cement, fine aggregate, and coarse aggregate in Table 2 are: 200.1, 389.6, 916.34 and 845.1, respectively, representing the highest degree of data concentration for each variable. The position or median of each variable is 202.3, 384.1, 894 and 841.5, indicating that 50% of each set of data is below these values and the other 50% is above these values, dividing it into two equal parts. The most repeated values in each study variable were 210, 340, 1007.9, and 782.6, respectively.

The standard deviation of each variable is 19.28, 46.74, 116.28, and 93.07, respectively, which is used to measure whether there is low, medium, or high dispersion. The variation coefficient of the water variable is 0.096, which is less than 0.10, indicating that the data dispersion of the variable is low, and its average value is representative of the group. For the cement variable, the values of fine aggregate and coarse aggregate are 0.12, 0.13, and 0.11, respectively, with values greater than 0.10 but less than 0.30, indicating an intermediate dispersion and a downward dispersion trend. Therefore, it can be considered that the average value also represents the data set. Therefore, the range of values for each variable is its standard deviation or the average of its standard deviations. For water 200.1 ± 19.28 lt, for cement 389.59 ± 46.74 kg, for fine aggregate 916.34 ± 116.28 kg, and for coarse aggregate 845.1 ± 93.07 kg.

The kurtosis of the water variable is 1.44, and because this value is greater than zero, its curve is leptokurtic, indicating that the data is often concentrated around the central value. For cement, fine aggregate, and coarse aggregate variables, the kurtosis are -0.33, -0.02, and -0.11, and because they are less than zero, their curves are platykurtic, indicating a low concentration of values, and meaning that their data is scattered. The platinum curve also shows a softening of its distribution form.

For variables such as water, cement, and fine aggregate, the asymmetry is 0.22, 0.23, and 0.72, with values greater than zero, indicating a positive asymmetry where values tend to cluster to the left of the arithmetic mean. For the coarse aggregate variable, the asymmetry coefficient is -0.15, and values less than zero indicate negative asymmetry, with their values clustered to the right of the mean.

Figures 5, 6 and 7 show a unimodal histogram, as they have a higher bar, indicating the highest absolute frequency class, which are 397-422, 205-219 and 804-864 respectively. Furthermore, it is skewed, indicating that these variables have an asymmetric behavior and in agreement with the aforementioned dispersion where the data are concentrated on the left side of the distribution.

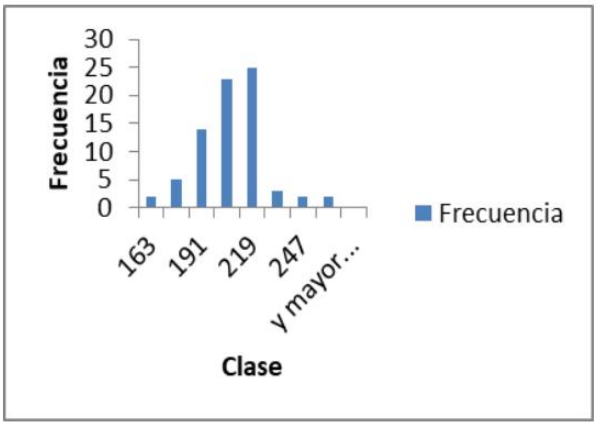

Figure 5. Absolute frequency histogram of water variable for the Porrero method. Source: authors.

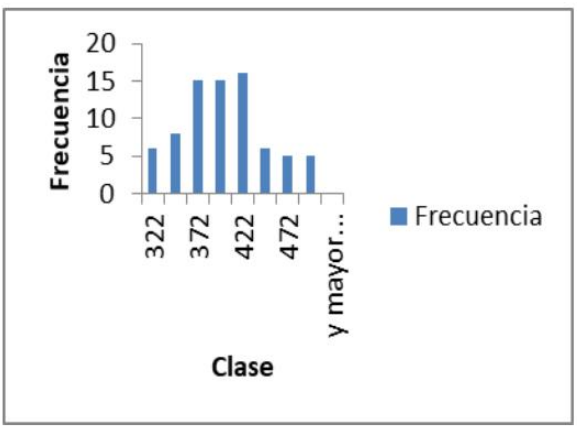

Figure 6. Absolute frequency histogram of the cement variable for the Porrero method. Source: authors.

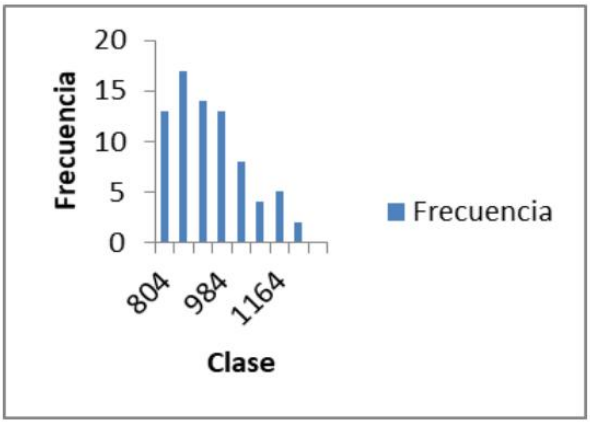

Figure 7. Absolute frequency histogram of fine aggregate variables for the Porrero method. Source: authors.

Figure 8 shows a bimodal histogram because it has two higher equality bars, namely two classes with the same absolute frequency. In addition, it is biased, indicating that the variable exhibits asymmetric behavior, with data concentrated on the right side, consistent with the aforementioned dispersion. The categories with the highest frequency of data occurrence are 802-862 and 862-922, with the first containing the mean and median.

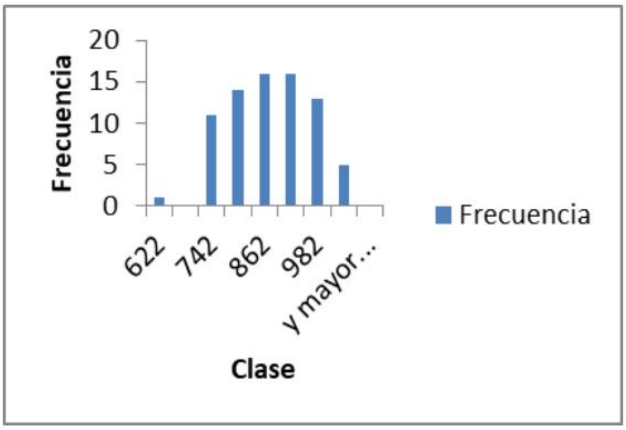

Figure 8. Absolute frequency histogram of coarse aggregate variables for the Porrero method. Source: authors.

4.2 Determination of probability distributions with Easyfit software version 5.6

The distribution that best fits the available data for each variable and design method was determined through the program (see Table 3). For the five most suitable, a new adjustment order was determined compared to the average level, with a confidence level of 0.95 (see Table 4). Next, only five best fit distributions will be given for the calculation and analysis results of variable water in the ACI method. The same procedure was used for other variables.

4.2.1 ACI method for conventional concrete mix design

Table 3 shows the adjustment order of the distribution, with the number one indicating that it is the best adjustment made for each test. In order to simultaneously consider the Kolmogorov-Smirnov and Anderson-Darling tests, their ranges were weighted to estimate the five distributions that best fit the data. As shown in the table, the Chi-Square test is not applicable to many distributions processed by the program, so its results were not used for analysis. In each goodness test, the five distributions that best fit the data set are indicated in red.

Table 3. Adjustment of the different distributions for the water variable of the ACI Method. Source: authors.

#	Distribución	Kolmogorov Smirnov		Anderson Darling		Chi-cuadrado		Valor Ponderado
		Estadística	Rango	Estadística	Rango	Estadística	Rango	
1	Beta	0,4072	14	8,8783	37	N/A		25,5
2	Burr	0,44178	35	6,8911	29	20,025	19	32
3	Burr (4P)	0,57948	53	14,602	54	N/A		53,5
4	Chi-Squared	0,45346	38	7,6365	33	11,016	10	35,5
5	Chi-Squared (2P)	0,42551	33	6,5273	11	N/A		22
6	Dagum	0,71365	57	17,572	57	N/A		57
7	Dagum (4P)	0,37526	1	6,2422	1	N/A		1
8	Erlang	0,4064	13	6,7848	17	7,2907	8	15
9	Erlang (3P)	0,41611	28	6,8068	19	20,007	18	23,5
10	Error	0,44653	36	7,0457	31	18,473	12	33,5
11	Error Function	1	59	N/A		N/A		59
12	Exponential	0,59915	54	13,136	52	N/A		53
13	Exponential (2P)	0,54714	51	10,661	44	N/A		47,5
14	Fatigue Life	0,41104	22	6,8449	25	7,1178	5	23,5
15	Fatigue Life (3P)	0,38282	6	6,5068	10	N/A		8
16	Frechet	0,46985	40	6,7802	16	N/A		28
17	Frechet (3P)	0,3884	9	6,4049	4	N/A		6,5
18	Gamma	0,41037	20	6,8192	23	7,2778	7	21,5
19	Gamma (3P)	0,38615	7	6,5454	12	N/A		9,5
20	Gen, Extreme Value	0,45574	39	10,436	43	N/A		41
21	Gen, Gamma	0,41276	24	6,875	28	19,207	17	26
22	Gen, Gamma (4P)	0,3814	4	6,4881	9	N/A		6,5
23	Gen, Pareto	0,47674	41	10,864	47	N/A		44
24	Gumbel Max	0,41476	25	6,3864	3	N/A		14
25	Gumbel Min	0,48577	43	12,337	51	N/A		47
26	Hypersecant	0,42781	34	6,7606	15	18,575	13	24,5
27	Inv, Gaussian	0,41576	27	6,8483	26	19,115	16	26,5
28	Inv, Gaussian (3P)	0,42253	32	6,8663	27	20,046	20	29,5
29	Johnson SB	0,49384	48	11,296	49	N/A		48,5
30	Kumaraswamy	0,49092	45	9,2799	39	N/A		42
31	Laplace	0,44653	37	7,0457	32	18,473	11	34,5
32	Levy	0,65015	55	15,45	55	N/A		55
33	Levy (2P)	0,47977	42	10,275	42	N/A		42
34	Log-Gamma	0,40916	18	6,8029	18	7,1413	6	18
35	Log-Logistic	0,41272	23	6,4289	5	N/A		14
36	Log-Logistic (3P)	0,37888	2	6,2459	2	N/A		2

37	Log-Pearson 3	0,49251	46	11,292	48	N/A		47
38	Logistic	0,42223	31	6,7504	14	18,66	14	22,5
39	Lognormal	0,41086	21	6,8428	24	7,1177	4	22,5
40	Lognormal (3P)	0,38053	**3**	6,47	8	N/A		**5,5**
41	Nakagami	0,40985	19	6,8146	22	7,4442	9	20,5
42	Normal	0,41563	26	6,9225	30	18,856	15	28
43	Pareto	0,5523	52	10,86	46	N/A		49
44	Pareto 2	0,67394	56	16,507	56	N/A		56
45	Pearson 5	0,40896	17	6,8104	21	7,0703	2	19
46	Pearson 5 (3P)	0,38244	**5**	6,4383	7	N/A		**6**
47	Pearson 6	0,40881	15	6,8098	20	7,0764	3	17,5
48	Pearson 6 (4P)	0,38703	8	6,4297	6	N/A		7
49	Pert	0,3979	10	8,6951	36	N/A		23
50	Power Function	0,40891	16	11,43	50	N/A		33
51	Rayleigh	0,50263	49	10,728	45	N/A		47
52	Rayleigh (2P)	0,41812	29	8,4003	35	N/A		32
53	Reciprocal	0,48657	44	9,0656	38	N/A		41
54	Rice	0,53971	50	9,4227	40	4,7238	1	45
55	Student's t	0,99998	58	306,91	58	N/A		58
56	Triangular	0,49255	47	10,242	41	N/A		44
57	Uniform	0,40219	12	14,12	53	N/A		32,5
58	Weibull	0,42062	30	7,6942	34	N/A		32
59	Weibull (3P)	0,40072	11	6,5923	13	N/A		12
60	Cauchy	No hay ajuste						
61	Johnson SU	No hay ajuste						

After weighting, the first five adjusted positions were selected from this column. For this variable, the results were Dagum (4p), Log-Logistic (3p), Log-Normal (3p), Pearson 5 (3p), and Gen. Gamma (4p). Under these five distributions, with a probability of 0.95, a new adjustment order was established considering the standard deviation (see Table 4).

Table 4. Distributions that best fit the data obtained for the water variable by the ACI design method for a confidence level of 95%. Source: authors.

#	Distribución	Orden de ajuste con la Media	Agua (lt)	Desviación con respecto a la media
1	Dagum (4P)	4	203,88	1,66
2	Log-Logistic (3P)	5	203,76	1,72
3	Lognormal (3P)	2	209,52	1,06
4	Pearson 5 (3P)	3	209,48	1,29
5	Gen, Gamma (4P)	1	210	1,04
		MEDIA	207,328	

The probability distribution with the smallest deviation will be the one closest to the estimated mean. For this variable, it is the Gen. Gamma (4P) distribution, with parameters k = 0.69112, α = 29.055, β = 0.20432 and γ = 169.86 (see Table 5). By comparing these distributions in Figure 9, it can be seen that the distribution with the highest contribution corresponds to the number 5, namely the Gen. Gamma (4p) distribution, while the distribution with the lowest contribution is the number 2, Log-Logistic (3p).

Table 5. Adjust the density function parameters of the distribution based on the data obtained from variable water using the ACI design method, with a confidence level of 95%. Source: authors.

#	Distribución	Parámetros
1	Beta	α_1=2,7267 α_2=4,3652 a=180,0 b=224,96
2	Burr	k=0,32465 α=154,89 β=193,66
3	Burr (4P)	k=0,51124 α=0,78193 β=1,6994 γ=180,0
4	Chi-Squared	ν=196
5	Chi-Squared (2P)	ν=25 γ=171,06
6	Dagum	k=173,05 α=14,577 β=125,63
7	Dagum (4P)	k=1,0455 α=11,304 β=27,931 γ=167,49
8	Erlang	m=611 β=0,32213
9	Erlang (3P)	m=14 β=1,9235 γ=170,5
10	Error	k=1,0 σ=7,9641 μ=196,9
11	Error Function	h=0,08879
12	Exponential	λ=0,00508
13	Exponential (2P)	λ=0,05917 γ=180
14	Fatigue Life	α=0,03822 β=196,76
15	Fatigue Life (3P)	α=0,2036 β=33,878 γ=162,32
16	Frechet	α=26,726 β=191,96
17	Frechet (3P)	α=1,3914E+8 β=8,7148E+8 γ=-8,7148E+8
18	Gamma	α=611,25 β=0,32213
19	Gamma (3P)	α=13,724 β=1,9235 γ=170,5
20	Gen, Extreme Value	k=0,4797 σ=2,0075 μ=193,95
21	Gen, Gamma	k=1,0101 α=652,15 β=0,32213
22	Gen, Gamma (4P)	k=0,69112 α=29,055 β=0,20432 γ=169,86
23	Gen, Pareto	k=0,36501 σ=2,9218 μ=192,3
24	Gumbel Max	σ=6,2096 μ=193,32
25	Gumbel Min	σ=6,2096 μ=200,48
26	Hypersecant	σ=7,9641 μ=196,9
27	Inv, Gaussian	λ=1,2035E+5 μ=196,9
28	Inv, Gaussian (3P)	λ=824,35 μ=34,562 γ=162,34
29	Johnson SB	γ=2,7234 δ=0,95617 λ=98,098 ξ=188,93
30	Kumaraswamy	α_1=1,05 α_2=1,15 a=180,0 b=224,96
31	Laplace	λ=0,17757 μ=196,9
32	Levy	σ=196,61
33	Levy (2P)	σ=13,058 γ=178,48
34	Log-Gamma	α=18483,0 β=2,8577E-4

35	Log-Logistic	α=36,285 β=195,85
36	Log-Logistic (3P)	α=11,258 β=27,504 γ=168,03
37	Log-Pearson 3	α=0,87018 β=0,04165 γ=5,2457
38	Logistic	σ=4,3908 μ=196,9
39	Lognormal	σ=0,0382 μ=5,2819
40	Lognormal (3P)	σ=0,21499 μ=3,4523 γ=164,55
41	Nakagami	m=140,28 Ω=38831,0
42	Normal	σ=7,9641 μ=196,9
43	Pareto	α=11,237 β=180
44	Pareto 2	α=165,62 β=26511,0
45	Pearson 5	α=703,22 β=1,3826E+5
46	Pearson 5 (3P)	α=34,102 β=1311,3 γ=157,23
47	Pearson 6	α_1=13971,0 α_2=737,68 β=10,382
48	Pearson 6 (4P)	α_1=57,287 α_2=38,152 β=21,181 γ=164,1
49	Pert	m=194,11 a=180 b=224,96
50	Power Function	α=0,72894 a=180,0 b=224,96
51	Rayleigh	σ=157,1
52	Rayleigh (2P)	σ=13,396 γ=180
53	Reciprocal	a=180,0 b=224,96
54	Rice	ν=193,28 σ=5,8943
55	Student's t	ν=2
56	Triangular	m=194,99 a=180,0 b=224,96
57	Uniform	a=183,11 b=210,69
58	Weibull	α=23,176 β=200,42
59	Weibull (3P)	α=2,338 β=20,731 γ=178,39
60	Cauchy	No hay ajuste
61	Johnson SU	No hay ajuste

By comparing these distributions in Figure 9, it can be seen that the distribution with the highest contribution corresponds to the number 5, namely the Gen. Gamma (4p) distribution, while the distribution with the lowest contribution is the number 2, Log-Logistic (3p).

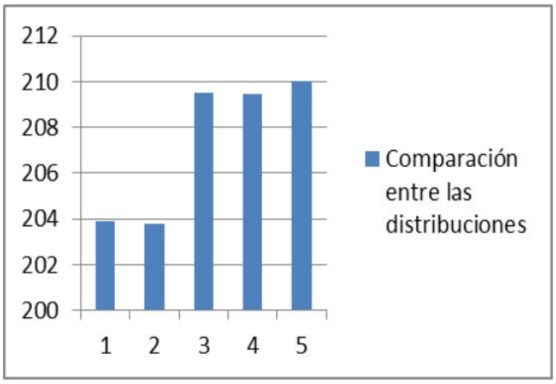

Figure 9. Differences of the water variable by the ACI design method between the different distributions. Source: authors.

Figure 10 presents the analysis of the fit tests for the hypothesis test of the distribution Gen. Gamma (4P) distribution.

Gen. Gamma (4P) [#22]					
Kolmogorov-Smirnov					
Tamaño de la muestra Estadística Valor P Rango	30 0.3814 1.9632E-4 4				
α	0.2	0.1	0.05	0.02	0.01
Valor crítico	0.19032	0.21756	0.2417	0.27023	0.28987
Rechazar?	Sí	Sí	Sí	Sí	Sí
Anderson-Darling					
Tamaño de la muestra Estadística Rango	30 6.4881 9				
α	0.2	0.1	0.05	0.02	0.01
Valor crítico	1.3749	1.9286	2.5018	3.2892	3.9074
Rechazar?	Sí	Sí	Sí	Sí	Sí

Figure 10. Analysis of the goodness-of-fit tests for the acceptance or rejection of the hypotheses of the distribution Gen. Gamma (4P) distribution. Source: authors.

In the case of Kolmogorov-Smirnov, the test statistic for cumulative observation and theoretical frequency is 0.3814, exceeding the critical value of 0.2417 at a significance level of 0.05. For example, the null hypothesis indicating that the data follows a specific distribution is rejected. When analyzing based on the P-value of 0.000196, it can be seen that according to the above situation, the P-value is lower than the five confidence levels used by the plan to reject the null hypothesis. Anderson Darling's test statistic is 6.4881, which is higher than the critical value at different significance levels, thus rejecting the assumption that the distribution of the dominant water variable data is a specific form. Following the entire program of cement, fine aggregate, and coarse aggregate mentioned above, the deviation of Gen distribution is relatively small. The extreme values of parameters $k = 0.07189$, $\sigma = 13.61$, and $\mu = 361$, Weibull (3P) with parameters $\alpha = 7.1563$, $\beta = 627.35$, and $\gamma = 329.93$, and Frechet with parameters $\alpha = 13.534$ and $\beta = 860.32$ (see Tables 6, 7, and 8, respectively).

Table 6. Distributions that best fit the data obtained for the cement variable by the ACI design method for a confidence level of 95%. Source: authors.

#	Distribución	Orden de ajuste con la Media	Cemento (Kg)	Desviación con respecto a la media
1	Cauchy	5	369,63	1,93
2	Log-Logistic (3P)	2	361,25	0,38
3	Burr	4	360,24	0,66
4	Gen, Extreme Value	1	361,56	0,3
5	Frechet (3P)	3	360,49	0,59
		MEDIA	362,634	

Table 7. Distributions that best fit the data obtained for the fine aggregate variable by the ACI design method for a confidence level of 95%. Source: authors

#	Distribución	Orden de ajuste con la Media	Agregado Fino (Kg)	Desviación con respecto a la media
1	Gumbel Min	5	1047,7	1,36
2	Gen. Extreme Value	2	1057,3	0,46
3	Weibull	3	1070,5	0,78
4	Error	4	1074,2	1,13
5	Weibull (3P)	1	1061,2	0,09
		MEDIA	1062,18	

Table 8. Distributions that best fit the data obtained for the coarse aggregate variable by the ACI design method for a confidence level of 95%. Source: authors.

#	Distribución	Orden de ajuste con la Media	Cemento (Kg)	Desviación con respecto a la media
1	Frechet	1	1071,4	0,54
2	Fatigue Life (3P)	3	1055,9	0,91
3	Inv, Gaussian (3P)	2	1056	0,9
4	Log-Logistic (3P)	5	1093	2,57
5	Gamma (3P)	4	1051,8	1,3
		MEDIA	1065,62	

When comparing the five best distributions in each of these variable cement, fine and coarse aggregate, it can be seen in Figure 11 that the distribution with the highest value is the one corresponding to number 1, i.e., the Cauchy distribution (4P) and the one that contributes the least is the Burr distribution. In Figure 12, the distribution with the highest value is the one corresponding to number 4, that is, the Error distribution, and the one with the lowest value is the Gumbel Min. In Figure 13, the distribution with the highest value is the one corresponding to number 4, that is, the Log-Logistic (3P) distribution, and the one with the lowest value is the Gamma (3P), respectively.

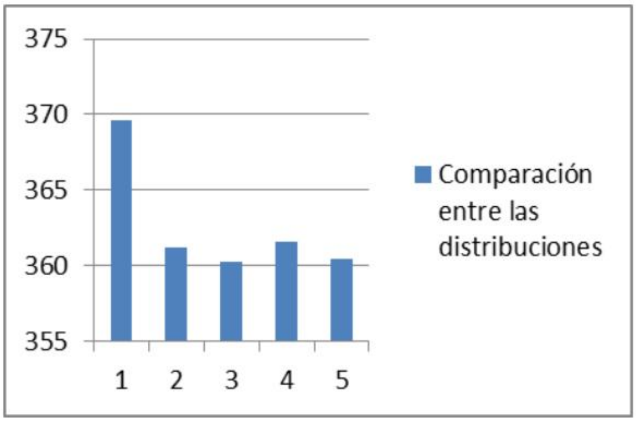

Figure 11. Differences of the cement variable by ACI design method between the different distributions. Source: authors.

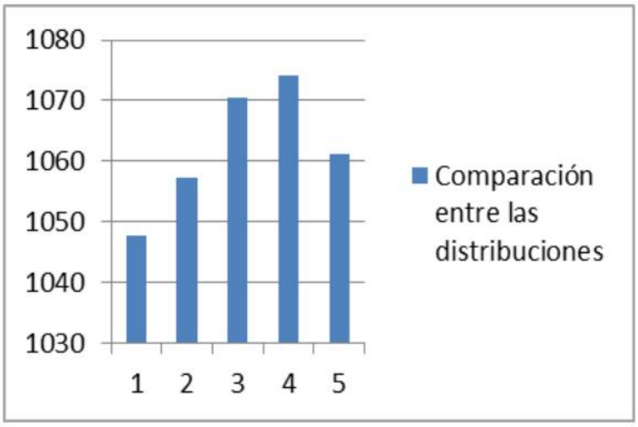

Figure 12. Differences of the fine aggregate variable by ACI design method between the different distributions. Source: authors.

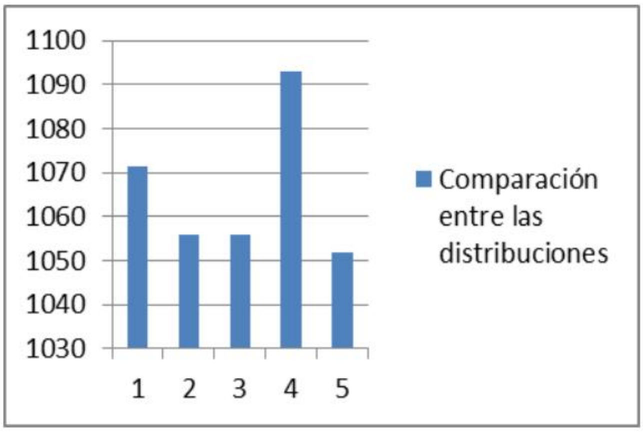

Figure 13. Differences of the coarse aggregate variable by ACI design method between the different distributions. Source: authors.

Figures 14, 15, and 16 show the density functions most suitable for water, cement, and fine aggregate variable data, as well as the corresponding unimodal and biased histograms, while Figure 17 shows the density functions most suitable for coarse aggregate variable data, with the histograms being bimodal and biased. This is consistent with the results of descriptive statistical analysis, indicating consistency.

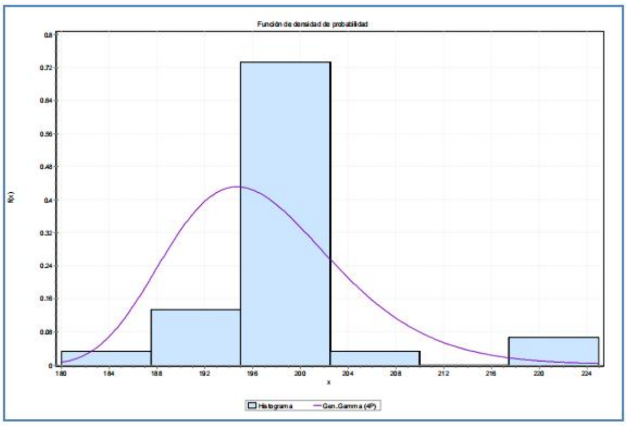

Figure 14. Density function for the Gen. Gamma (4P) distribution applicable to the data of the water variable by the ACI design method. Source: authors.

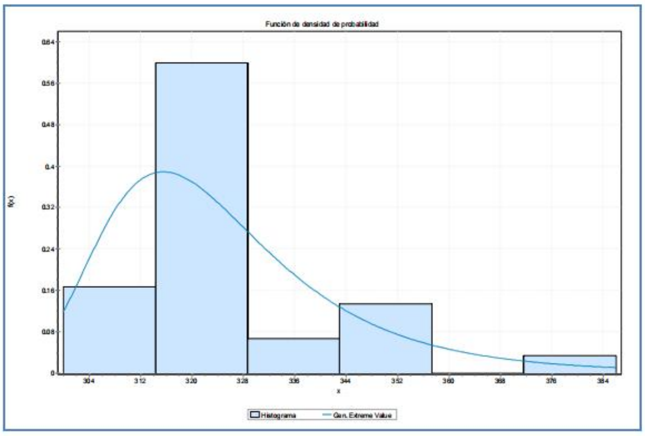

Figure 15. Density function for the Gen. Extreme value distribution applicable to the cement variable data by the ACI design method. Source: authors.

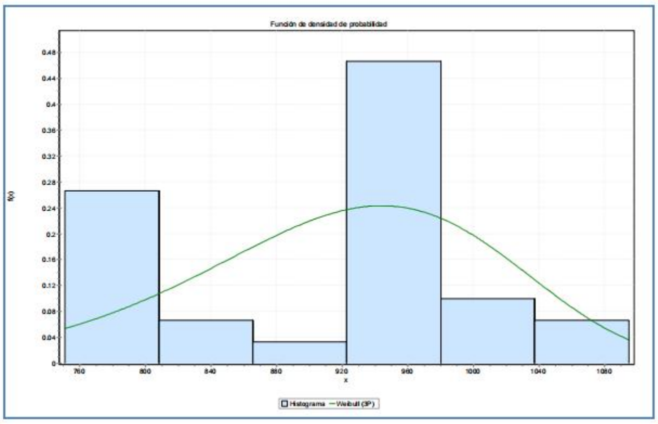

Figure 16. Density function of the Weibull distribution (3P) applicable to the fine aggregate variable data of the ACI design method. Source: authors.

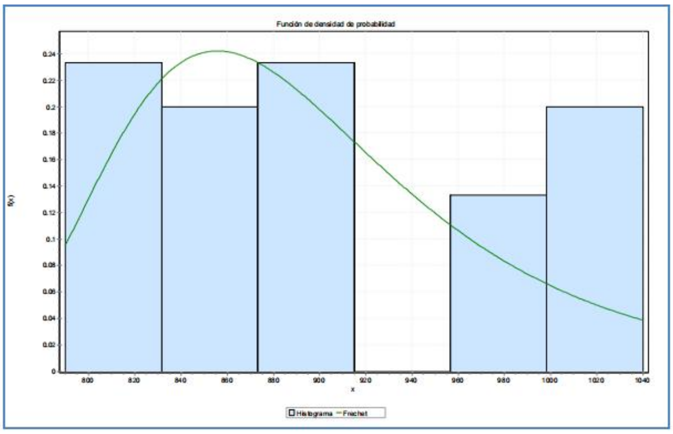

Figure 17. Density function for the Frechet distribution applicable to the data of the coarse aggregate variable by the ACI design method. Source: authors.

4.2.2 Structural concrete manual or Porrero's method

Table 9 shows the five distributions with the first positions in the fit after weighting with respect to the goodness-of-fit tests, which for the variable water are: Hypersecant, Error, Log-Logistic (3P), Burr and Johnson SU. Calculating the amount of water with these five distributions for a confidence level of 0.95 and comparing these values with their mean, a new order of adjustment was established that considers the deviation with respect to the mean, and in this case, the probability distribution that is closest to the estimated mean value is the Hypersecant or Secant Hyperbolic distribution whose parameters are: $\sigma = 19.278$ and $\mu = 200.05$.

Table 9. Distributions that best fit the data obtained for the water variable by Porrero's method for a confidence level of 95%. Source: authors.

#	Distribución	Orden de ajuste con la Media	Cemento (Kg)	Desviación con respecto a la media
1	Hypersecant	1	231,25	0,02
2	Error	2	231,84	0,23
3	Log-Logistic (3P)	3	230,75	0,24
4	Burr	4	230,66	0,28
5	Johnson SU	5	232,03	0,31
		MEDIA	231,306	

When comparing these distributions in Figure 18, the distribution with the highest value is the one corresponding to number 5, i.e., the Johnson SU distribution, and the one with the lowest value is the Burr distribution.

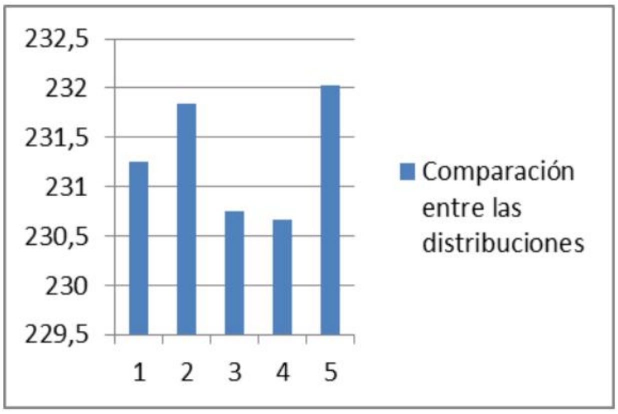

Figure 18. Differences of the variable water by Porrero's design method between the different distributions. Source: authors.

Following this procedure for cement, fine and coarse aggregates, the Log-Pearson 3 distribution with parameters $\alpha = 1027.3$, $\beta = -0.00375$, $\gamma = 9.8099$ was obtained with lower deviation.; Johnson SB, with parameters $\gamma = 1.0872$, $\delta = 1.09$, $\lambda = 684.38$, $\xi = 711.43$ and Chi-Squared (2P), with parameters $\nu = 4297$, $\gamma = -3452.2$ respectively (see Tables 10, 11 and 12). When comparing the five best distributions in each of these variable cement, fine and coarse aggregates, it can be seen in Figure 19 the distribution that contributes the highest value is the one corresponding to number 3, that is, the Log-Logistic distribution (3P) and the one that contributes the least is Gen. Gamma.

Table 10. Distributions that best fit the data obtained for the cement variable by the Porrero design method for a confidence level of 95%. Source: Authors.

#	Distribución	Orden de ajuste con la Media	Cemento (Kg)	Desviación con respecto a la Media
1	Gamma	3	469,52	0,23
2	Log-Pearson 3	1	470,37	0,05
3	Log-Logistic (3P)	5	474,05	0,73
4	Gen. Gamma	4	468,91	0,36
5	Burr	2	470,24	0,08
		MEDIA	470.618	

Table 11. Distributions that best fit the data obtained for the fine aggregate variable by the Porrero design method for a confidence level of 95%. Source: authors.

#	Distribución	Orden de ajuste con la Media	Cemento (Kg)	Desviación con respecto a la Media
1	Johnson SB	1	1139,3	0,01
2	Gen. Extreme Value	2	1139,9	0,04
3	Weibull (3P)	3	1140,1	0,06
4	Frechet (3P)	5	1135,6	0,34
5	Pearson 5 (3P)	4	1142,2	0,24
		MEDIA	1139.42	

Table 12. Distributions that best fit the data obtained for the coarse aggregate variable by the Porrero design method for a confidence level of 95%. Source: authors.

#	Distribución	Orden de ajuste con la Media	Cemento (Kg)	Desviación con respecto a la Media
1	Gen. Extreme Value	2	999,61	0,07
2	Error	4	997,88	0,1
3	Gen. Gamma	5	1000,4	0,15
4	Chi-Squared (2P)	1	998,39	0,05
5	Normal	3	998,15	0,07
		MEDIA	998.886	

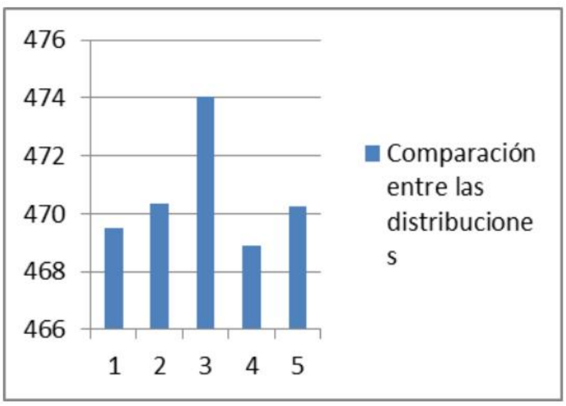

Figure 19. Differences of the cement variable by Porrero's method between the different distributions. Source: authors.

In Figure 20, the distribution with the highest value is the one corresponding to number 5, that is, the Pearson 5 (3P) distribution, and the one with the lowest value is the Frechet (3P) distribution. In Figure 21, the distribution with the highest value is the one corresponding to number 3, i.e., the Gen. Gamma and the one that contributes the least is the Error, respectively.

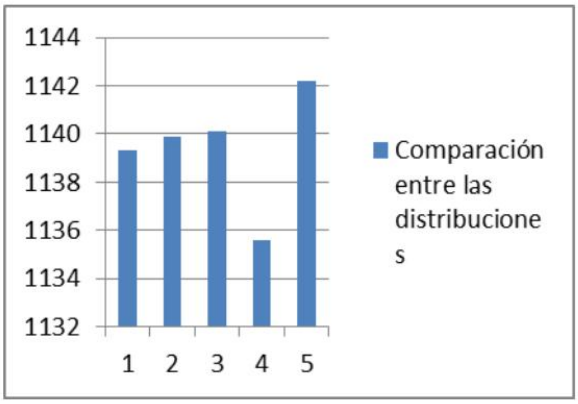

Figure 20. Differences of the variable fine aggregate by Porrero's method between the different distributions. Source: authors.

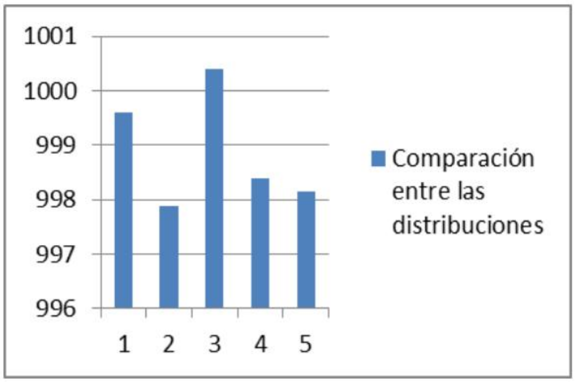

Figure 21. Differences of the coarse aggregate variable by Porrero's method between the different distributions. Source: authors.

Figures 22, 23, 24 and 25 show the density function that best fits the data of the variable water, cement and coarse and fine aggregate with the corresponding histograms, which are unimodal and skewed, in agreement with the results of the descriptive statistical analysis.

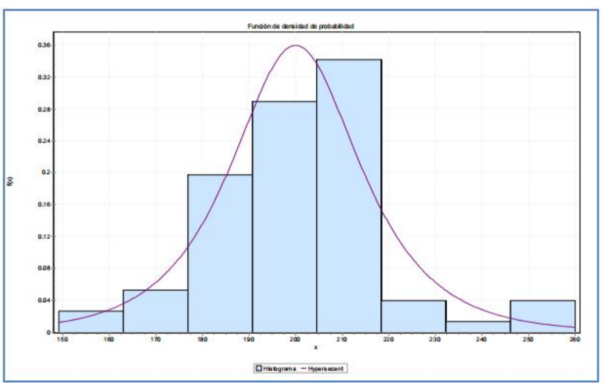

Figure 22. Density function for the Hypersecant distribution applicable to the data of the water variable by the Porrero design method. Source: authors.

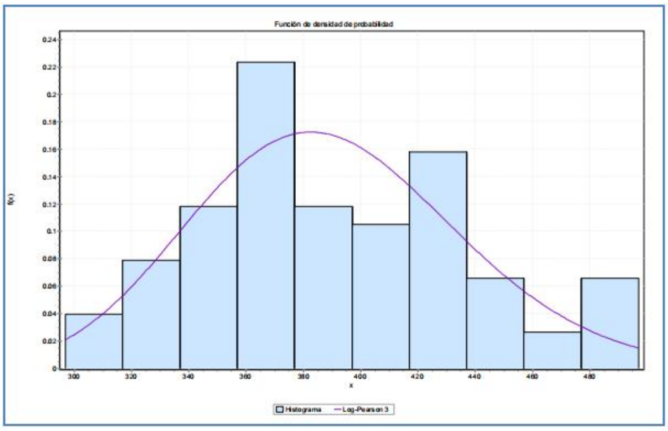

Figure 23. Density function for the Log-Pearson 3 distribution applicable to the data of the cement variable by Porrero's method. Source: authors.

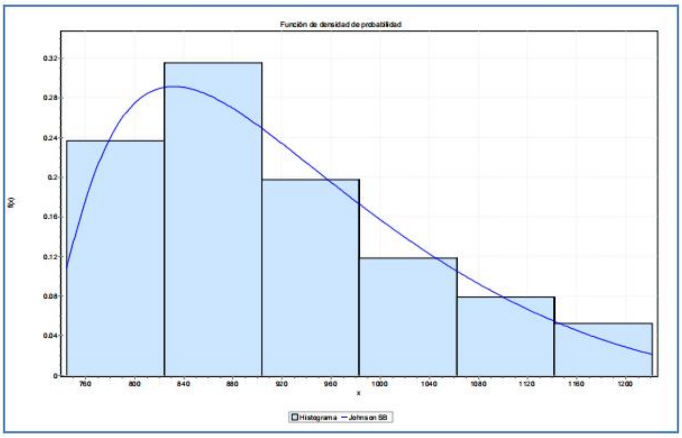

Figure 24. Density function for the Johnson SB distribution applicable to the data of the fine aggregate variable by Porrero's method. Source: authors.

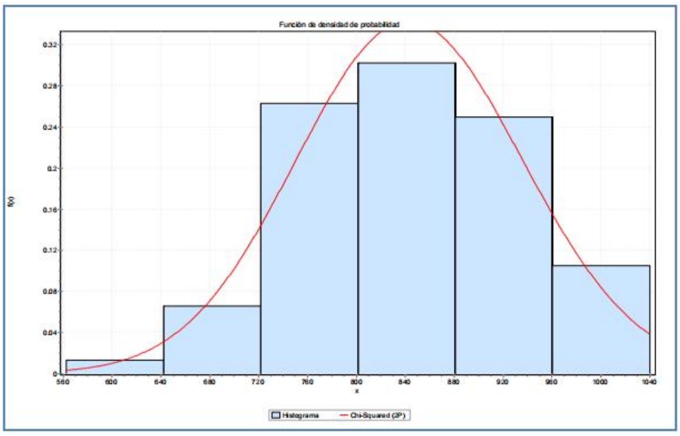

Figure 25. Density function for the Chi-Squared (2P) distribution applicable to the data of the coarse aggregate variable by Porrero's method. Source: authors.

Tables 13 and 14 present a summary of the mean and standard deviation statistics and the distributions determined for each of the concrete components by design method.

Table 13. Summary of statistics and distributions that best fit the variables under study for the ACI design method. Source: authors.

Variable	Media	Desviación Estándar	Distribución
Agua (Lt)	196,9	7,9641126	Gen. Gamma (4P)
Cemento (Kg)	325,393	19,3956592	Gen. Extreme Value
Agregado Fino (Kg)	916,12933	100,810663	Weibull (3P)
Agregado Grueso (Kg)	902,51167	75,6389341	Frechet

Table 14. Summary of statistics and distributions that best fit the data for the variables under study for Porrero's method. Source: authors.

Variable	Media	Desviación Estándar	Distribución
Agua (Lt)	200,048132	19,2776083	Hypersecant
Cemento (Kg)	389,5895	46,7371563	Log-Pearson 3
Agregado Fino (Kg)	916,337276	116,277188	Johnson SB
Agregado Grueso (Kg)	845,062908	93,067619	Chi-Squared (2P)

It should be noted that although the components are the same and the method of construction is different, the distributions obtained are not similar, nor are the statistical values.

4.3 Distributions of certain probabilities

4.3.1 ACI design method

The water variable value of this method ranges from 188.94 to 204.86 liter, with an average value of 196.9 liter and a deviation of 7.96 liter. The most suitable distribution for its data is the generalized Gamma of four parameters, which is a leptokurtic curve with positive skewness and its probability density function is as follows:

$$f(x) = \frac{(0.69112)*(x-169.86)^{0.69112*29.055-1}}{(0.20432)^{0.69112*29.055}*\Gamma(\alpha)} * e^{-\left(\frac{x-169.86}{0.20432}\right)^{0.69112}} \tag{3}$$

The values of the variable cement are in the range 305.99 - 344.79 kg, with an average value of 325.39 kg and a deviation of 19.40 kg. The distribution that best fits the data is the generalized extreme value, this is a leptokurtic curve with positive skewness and its probability density function is:

$$z = \frac{x - 316.5}{13.61} \tag{4}$$

$$f(x) = \frac{1}{13.61} * e^{-(1+0.07189*z)^{-\frac{1}{0.07189}}} * (1 + 0.07189 * z)^{-1-\frac{1}{0.07189}} \tag{5}$$

For the variable fine aggregate, the values are in the range 815.32 - 1016.94 kg, with an average value of 916.13 kg and a deviation of 100.81 kg. The distribution that best fits the data is the 3-parameter Weibull, which is a platykurtic curve with negative skewness and its probability density function is:

$$f(x) = \frac{13.534}{860.32} * \left(\frac{860.32}{x}\right)^{13.534+1} * e^{-\left(\frac{860.32}{x}\right)^{13.534}} \tag{6}$$

For the coarse aggregate variable, the values are in the range 826.87 - 978.25 kg, with an average value of 902.51 kg and a deviation of 75.64 kg. The distribution that best fits the data is the Frechet, which is a platykurtic curve with positive skewness and its probability density function is:

$$f(x) = \frac{13.534}{860.32} * \left(\frac{860.32}{x}\right)^{13.534+1} * e^{-\left(\frac{860.32}{x}\right)^{13.534}} \quad (7)$$

4.3.2 Porrero design method

The values of the variable water for this method are in the range 180.77 - 219.33 liter, with an average value of 200.05 liter and a deviation of 19.28 liter. The distribution that best fits your data is the Secant Hyperbolic, which is a leptokurtic curve with positive skewness and its probability density function is as follows:

$$f(x) = \frac{sec\ h\frac{\pi(x-200.05)}{2*19.278}}{2*19.278} \quad (8)$$

The values of the variable cement are in the range 342.85 - 436.33 kg, with an average value of 389.59 kg and a deviation of 46.74 kg. The distribution that best fits your data is the Log-Pearson 3, which is a platykurtic curve with positive skewness and its probability density function is:

$$f(x) = \frac{1}{x*|-0.00375|*\Gamma(1027.3)}\left(\frac{ln(x)-9.8099}{-0.00375}\right)^{1027.3-1} * e^{\left(-\frac{ln(x)-9.8099}{-0.00375}\right)} \quad (9)$$

For fine variables, the values range from 800.06 to 1032.62 kg, with an average value of 916.34 kg and a deviation of 11.28 kg. The distribution that best fits its data is Johnson SB, which is an asymmetric regular plane curve with a probability density function of:

$$z = \frac{x-711.43}{684.38} \quad (10)$$

$$f(x) = \frac{1.09}{684.38*\sqrt{2\pi}*z*(1-z)} * e^{\left(-\frac{1}{2}\left(1.0872*1.09*Ln\left(\frac{z}{1-z}\right)\right)^2\right)} \quad (11)$$

For the coarse aggregate variable, its value ranges from 751.99 to 938.13 kg, with an average of 845.06 kg and a deviation of 93.07 kg. The most suitable distribution for its data is the Chi-Square of two parameters, which is a negative asymmetric plastic curve with a probability density function of:

$$f(x) = \frac{(x-(-3452.2)^{\frac{4297}{2}-1} * e^{(-\frac{x-(-3452.2}{2})}}{2^{4297/2} * \Gamma(\frac{4297}{2})} \quad (12)$$

5. Conclusions

In the descriptive analysis of these two methods, the variables of water, cement, and fine aggregate exhibit similar behavior in histogram and symmetry, while coarse aggregate exhibits different behaviors. All research variables exhibit downward and intermediate dispersion, with minimal variability, reflecting the correctness of the obtained probability distribution. Most importantly, through its probability theory and sampling theory, it can be demonstrated through statistical science in both technical and scientific aspects.

Furthermore, as expected, they are of continuous type, as the variables being studied are continuous, indicating that the results obtained are consistent. The results of EasyFit software are consistent with those of descriptive statistical analysis, therefore, it is considered reliable for this study.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] ACI 211.1-91, Standard Practice for Selecting Proportions for Normal, Heavyweight, and Mass Concrete, American Concrete Institute, 2002.

[2] J. Porrero, et al. Manual del concreto estructural, SIDETUR, Caracas, Venezuela, 2004.

[3] A. García, Análisis de Distribuciones Estadísticas Alternativas a las Tradicionales para la Optimización de los Caudales de Cálculo Empleados en los Estudios Hidrológico (Tesis Doctoral), Universidad de Extremadura, Badajoz, 2013.

[4] Z. Cerón, Análisis Probabilístico del Concreto de Alta Resistencia (Trabajo de Grado), Universidad Católica de Colombia, 2013.

[5] COVENIN 221:2001, Materiales de Construcción. Terminología y definiciones FONDONORMA, Venezuela, 2001.

[6] N. Azuaje, et. al, Estimación de la constante de carbonatación "K" en concreto expuesto al ambiente en la ciudad de Nirgua, Estado Yaracuy, Trabajo Especial de Grado, Universidad Centroccidental Lisandro Alvarado, Venezuela, 2013.

[7] R. A. Dantas, Ingeniería de tasaciones una introducción a la metodología científica, premio Charles B. Akerson – UPAV 2000, Ed. Pini Ltda, 2002.

[8] M. Suarez, Curso Estadística, Universidad Centroccidental Lisandro Alvarado, Decanato de Ingeniería Civil, Venezuela, 2008.

[9] A. Rojas, Correlación entre el pulso ultrasónico y la resistencia a compresión en cilindros de concreto, Trabajo Especial de Grado, Universidad Centroccidental Lisandro Alvarado, Decanato de Ingeniería Civil Urbanismo, Venezuela. 2011.

[10] Mathwave Tecnologies, Ayuda del Software EasyFit versión 5.6., 2015.