CREAT. MATH. INFORM.
Volume **27** (2018), No. 2,
Pages 141 - 150

Online version at https://creative-mathematics.cunbm.utcluj.ro/
Print Edition: ISSN 1584 - 286X; Online Edition: ISSN 1843 - 441X
DOI: https://doi.org/10.37193/CMI.2018.02.06

Comparative statistical analysis on two types of acetylsalicylic acid tablets

THOMAS DIPPONG, ZOIŢA BERINDE, CRISTINA MIHALI, ANCA DUMUŢA and ZORICA VOSGAN

ABSTRACT. A statistical analysis was performed on two data sets, obtained by weighing 100 tablets each of acetylsalicylic acid (aspirin) and of calcium carbonate buffered acetylsalicylic acid. The frequency distribution histograms, the Gaussian distribution function and the tridimensional diagram of the Gaussian function and standard normal distribution were represented. The results showed that the weight of the buffered aspirin tablets varied in a larger range than the weights of the aspirin tablets. The novelty of the article consists in the transposition of the statistical processing of the results obtained by the measurement of the weight of acetylsalicylic acid tablets into computer programs specifically designed for this purpose.

1. INTRODUCTION

Over time many researches were conducted considering the therapeutic properties of aspirin on the human body. Aspirin represents the acetylsalicylic acid, well known as an important drug which can be taken in cases of acute and chronic inflammatory conditions due to its anti-inflammatory and analgesic-pyretic properties [22]. Also, aspirin taken in low-doses for a long term reduces the heart attack risk due to its anticoagulant effect [17]. Lately, aspirin is investigated considering its association to reduced cancer incidents and deaths [5]. There is evidence that aspirin may protect against the development of breast cancer, being performed observational studies indicating that breast cancer risk is reduced by 10 to 15% among users of aspirin [2, 23]. Not the same can be said about the results of some observational epidemiological studies about aspirin [1, 6, 8, 16, 19, 20, 21], in relation to pancreatic cancer risk [12]. However, the drug aspirin is known to affect the gastric mucosa by irritations and injuries. Considering a multicentre case-control study[11], the buffered aspirin is a softener variety of aspirin which is less likely to affect the gastric mucosa. Also, a new buffered chewable aspirin seems to provide a greater peak concentration of acetylsalicylate [7]. Therefore in this work a comparative statistical analysis was made between two types of aspirin tablets: simple aspirin and buffered aspirin. In the analysis, there were used the standard deviation, the Gaussian models, which have emerged as a powerful set of distributions for statistical modeling [3], also frequency and distribution Laplace. This paper aims to determine how values are distributed to the central value and on various integral values of the analyzed quantities; therefore we have chosen 100 values for each type of tablet.

2. EXPERIMENTAL

Two types of acetylsalicylic acid troches were taken into account, the simple tablet $(AA1)$ and the buffered tablet $(AA2)$. For each type of acetylsalicylic acid there were sampled 100 tablets which were weighed with an accuracy of 0.0001 mg on an analytical

Received: 07.11.2017. In revised form: 24.05.2018. Accepted: 31.05.2018
2010 *Mathematics Subject Classification.* 97M60, 97M10.
Key words and phrases. *aspirin, statistical analysis, standard deviation, histograms.*
Corresponding author: Thomas Dippong; th0mas2004@yahoo.com

balance. In order to downsize the inconveniences generated by their manual grouping and also values ranking a program was developed to accomplish these. Analytical balance used is Kern ABJ 80-4NM. Statistical programs were performed using the VISUAL BASIC 6.0 software.

3. Results and discussion

After the weighing procedure, the statistical processing of the obtained data was initiated by the determination of the mean and standard deviation. The average values (3.1) and standard deviation were calculated according to 3.1 and 3.2, respectively [4, 10, 15].

$$\mu = \frac{\sum X_i}{n} \tag{3.1}$$

where: μ is the average or arithmetic mean, n is the number of items and xi is the value of each individual item.

$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{n}} \tag{3.2}$$

where: σ is the standard deviation, xi is the value of each individual item, μ is the average of the data and n is the number of items. The statistical procedure began after the arrangement of the values in an increasing order and their separation into 10 intervals based on a computer software [10], as one can see in figures 1 and 2.

The highest values frequency can be observed in the acetylsalicylic acid tablets case, with 59 values in the interval $0.6376 - 0.6414$, while the $AA2$ troches recorded 51 values in the range of $0.8429 - 0.8492$. This fact indicates a lower variability of the weight of $AA1$ tablets, phenomenon backed also by the higher value of the standard deviation reached in the $AA2$ tablets case (figures 1 and 2).

A histogram presents a geometric shape representing the values distribution. On the ordinates are represented the number of values in each group and on the abscissa, the range of values. The mid range value of each class is the median of the values and sometimes is mistakenly thought to be the average of the group [10, 13, 15]. For determining the average value of the entire population of values one can use 3.3:

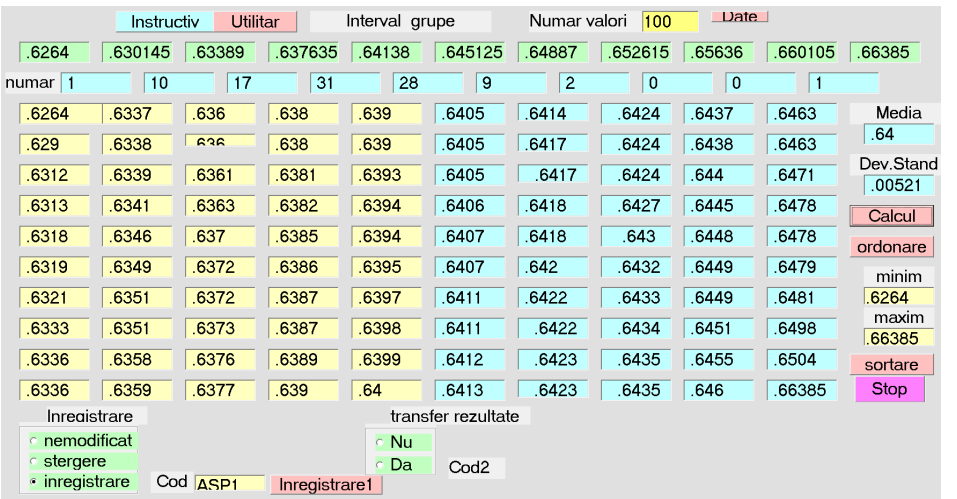

Figure 1. Ordered values obtained after weighing the AA1tablets

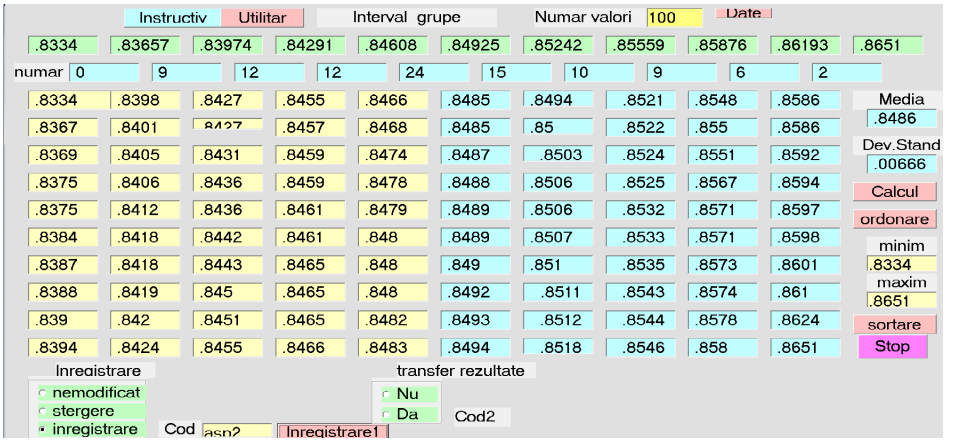

Figure 2. Ordered values obtained after weighing the AA2 buffered tablets

$$\mu = \frac{\sum((X)_{\frac{1}{2}} \cdot n_i)}{\sum n_i} \tag{3.3}$$

where

$X\frac{1}{2}$ represents the median in each group, n_i - the number of values in the group. Once the average is known one can calculate the standard deviation of the values that make up the local population using 3.4:

$$\mu = \sqrt{\frac{\sum\left((X_{\frac{1}{2}} - \mu)^2 n_i\right)}{\sum n_i}} \tag{3.4}$$

The histograms developed for the 2 sets of values are shown in figures 3 and 4. To compare the 2 different data the integrals of the 10 Gauss frequency function intervals are use, calculated by graphics integration. The frequency and Gaussian distribution curves were plotted using the average and the standard deviation. The integration of the frequency curve between two limits can determine the number of existing values within that range [10].

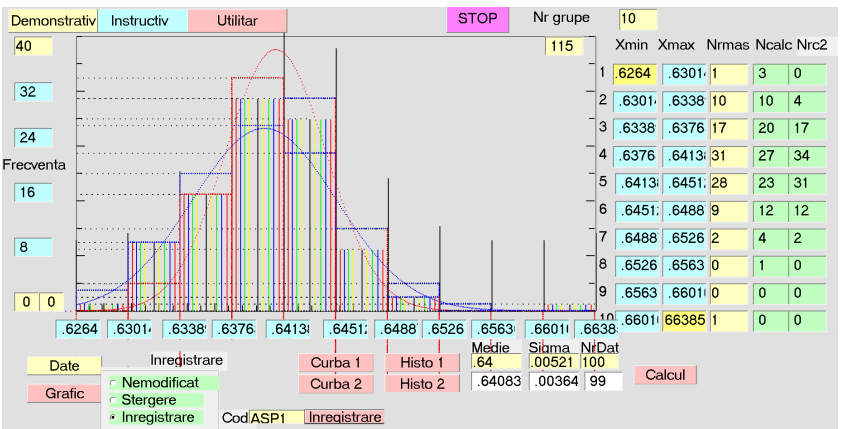

Figure 3. Histogram of the weighing values for AA1

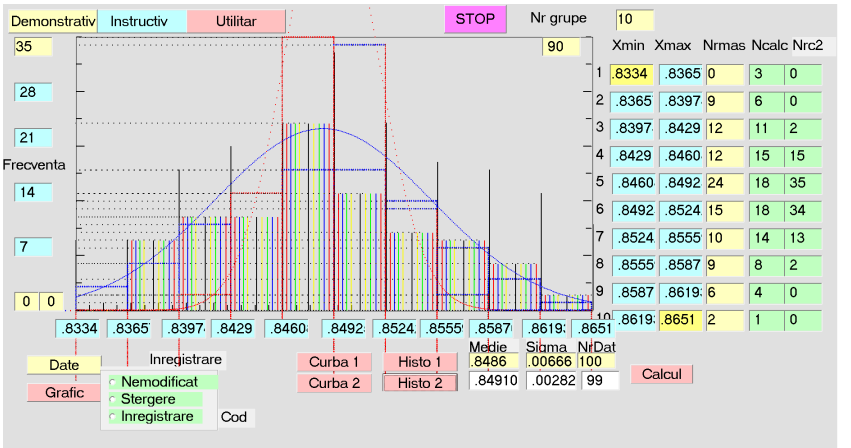

Figure 4. Histogram of the weighing values for AA2

From the above figures, there can be seen a higher variability in the case of the buffered acid acetylsalicylic tablets weight than in that of the acid acetylsalicylic tablets and this fact can be stated by a sharper distribution curve of $AA1$tablets and also by more scattered values recorded by $AA2$ tablets. By optimizing the values a sharper curve is acquired in both cases and thus a histogram with a much narrower distribution of values. The frequency function, $f(X)$ and the distribution function, $F(X)$ are developed according to equations 5 and 6.

$$f(X) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{\frac{\sum(X_i-\mu)^2}{2\sigma^2}} \tag{3.5}$$

$$F(X) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \int_{X_1}^{X_2} e^{\frac{\sum(X_i-\mu)^2}{2\sigma^2}} dX \tag{3.6}$$

Usually, considered to be a single variable function X_i, the Gaussian frequency function can be three-dimensional space represented, considering other values of σ, hence: X on OX, Y on OY and function $f(X, \sigma)$ on the Z axis [2, 3].

$$f(X) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{\frac{\sum(X_i-\mu)^2}{2\sigma^2}} \tag{3.7}$$

A single variable, Z, can be used so the calculations to be easier, by applying 3.8.

$$Z = \frac{X_i - \mu}{\sigma} \tag{3.8}$$

Thus the Gaussian distribution becomes: $f(X)dX = f(Z)dZ$ where $dX = \sigma \cdot dZ$
And the new frequency function can be calculated with 3.9 :

$$f(Z) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-Z^2}{2}}. \tag{3.9}$$

known as the normal distribution function Z or Laplace function, easier than the Gaussian distribution function, and not addicted to the average or standard deviation. The distribution function $(f(Z))$, defines the area under the bell and represents the integral of the frequency function, when the frequency function $(F(Z))$ can be calculated by equation 9 [10, 13, 14, 15].

$$F(Z) = \frac{1}{\sqrt{2\pi}} \int_{Z_1}^{Z_2} f(Z)dZ \tag{3.10}$$

While according to all statistical textbooks its integral values between 0 and a random value Z, is known as the Laplace function Eq. 11 [9, 18].

$$F(Z) = \frac{1}{\sqrt{2\pi}} \int_{0}^{Z} e\frac{-Z^2}{2}dZ \tag{3.11}$$

For the distribution function the graphical integration is applied using the trapezoids method. The chosen interval is divided into 400 parts and the surface of each part is calculated. By summing each interval the integral is obtained which is then plotted depending on X value (median of each interval). The distribution function takes values from 0 for$X = -$ infinity to 1 for $X = +$ infinity. The value of the distribution function can be calculated for any given period. The range of the distribution function values can be calculated only symmetrically to the average. It starts from the average to the extremities by summing the surfaces until the value of the calculated function exceeds the value of the given function. In this moment, the extremity function values are read. In the program are entered values for the extremes of the coordinate axes. For this purpose the computer displays some indicative values. The user will determine the coordinate axes value himself (slightly higher than the benchmarks). The frequency function is shown in figures 5 and 6.

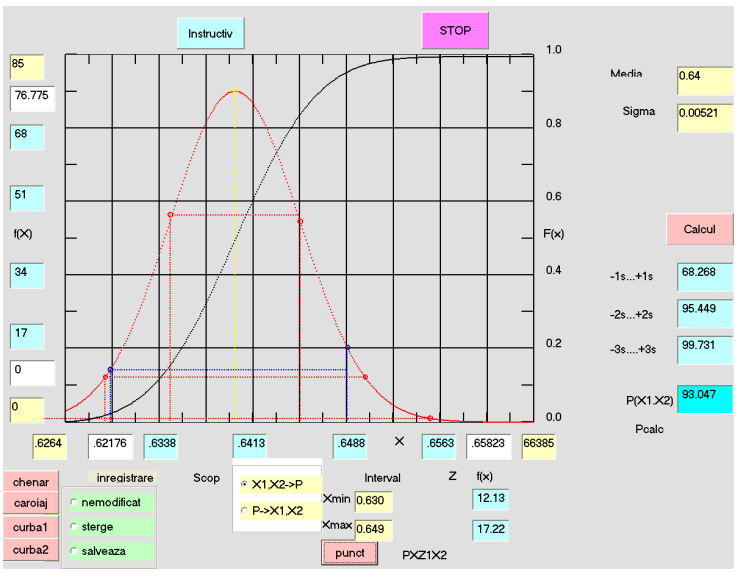

Figure 5. Frequency functions and Gaussian distribution curves for AA1

In the case of acid acetylsalicylic tablets values the software calculated a probability of 93.05% for the interval $0.630 - 0.649$ while for $AA2$ tablets, a 74.15% probability was established for the range of $0.831 \div 0.853$. One can note that in the acetylsalicylic acid $AA1$ case the frequency functions for two antipodal points are close, while in case of $AA2$ they are quite far away. Also the probability is greater in $AA1$ case than in $AA2$. The 3D graphs of the frequency function for $AA1$ and $AA2$ are represented in figures 7 and 8, based on random values for X [10].

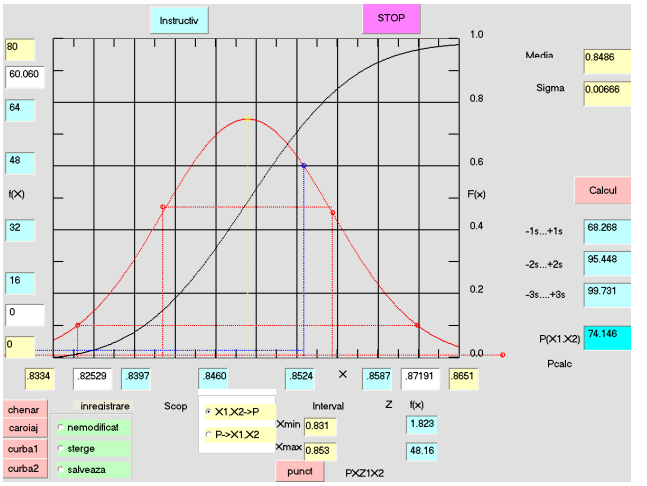

Figure 6. Frequency functions and Gaussian distribution curves for AA2

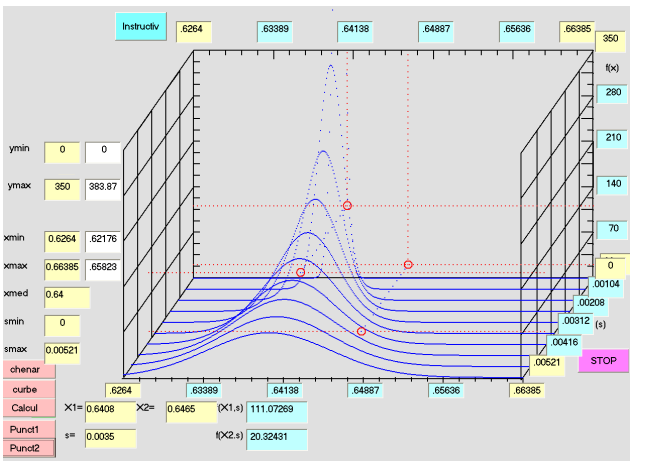

Figure 7. 3D Chart of the frequency function in relation to X value and standard deviation for AA1

In figure 7, is presented how $f(X_1) = 111.07$ and $f(X_2) = 20.32$ functions are obtained for $AA1$ tablets based on a standard deviation of 0.0035, $X1 = 0.6408$ and $X2 = 0.6465$. Also, figure 8 shows how $f(X_1) = 68.0$ and $f(X_2) = 48.2$ functions are obtained for $AA2$ tablets depending on a standard deviation of 0.0058, $X_1 = 0.8495$ and $X_2 = 0.8535$. Figures 9 and 10 illustrate the frequency function $f(Z)$, the integration between any limits based on the trapezoidal rule, the area under the bell in given limits and the variation of two distribution function values for 2 different values of the range. The integral function $f(Z)$ result is alike [10]. In figure 9, there can be seen that the probability of reproduction of weights on the interval $X_1 = 0.630(Z_1 = -1.9193)$ and $X_2 = 0.649(Z_2 = 1.7274)$ is 93.05%, equal to that presented in figure 5 which was determined by the Gaussian function $(P = 93.05)$. By calculating the reproduction probability from the interval for $AA2$ buffered aspirin, for values $X_1 = 0.831(Z_1 = -2.6426)$ and $X_2 = 0.853(Z_2 = 0.6606)$, the probability of the values reproduction is 74.146% (identical to that calculated with Gauss function $P = 74.146$, figure 6).

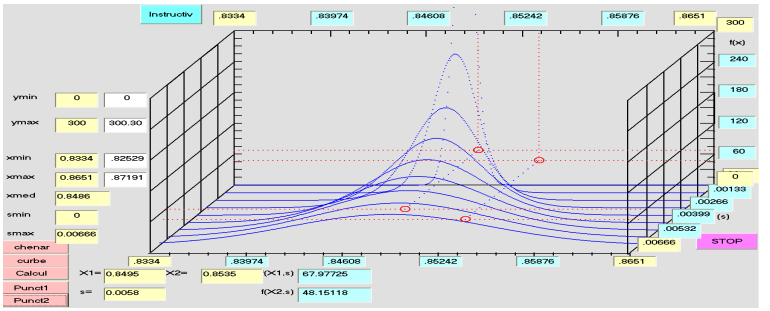

Figure 8. 3D Chart of the frequency function in relation to X value and standard deviation for AA2

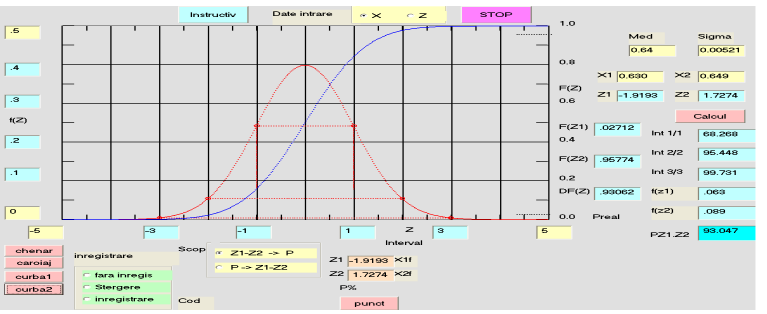

Figure 9. Frequency and distribution Laplace determining for AA1

The integral of two functions between two limit values of Z permits the probability determination for obtaining a specific value of the interval [10] and thus improving the frequency function $f(Z)$ importance. As shown in the graphic, the maximum value of the frequency function $f(Z)$, for $Z = 0$ is $\frac{1}{\sqrt{2\pi}} = 0.4$. For a range between $-$ infinity and $+$ infinity the integral value is equal to 1.

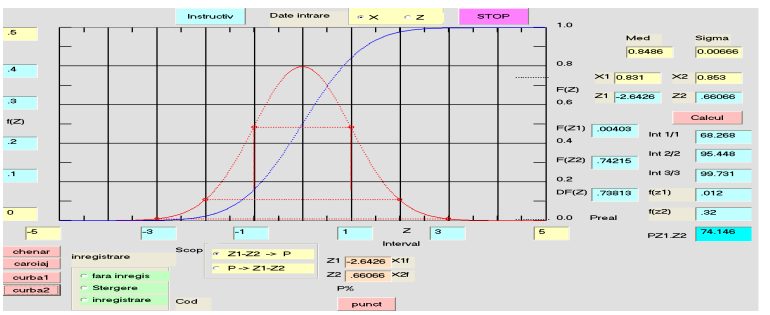

Figure 10. Frequency and distribution Laplace determining for AA2

Reverse calculation can be achieved, thus figures 11 and 12 present the same graphs like figures 9 and 10, except that in this case the probability value is introduced and $Z1$ and $Z2$ are calculated, in order to determine the difference between two values of Z for the value of a given area (between 0 and 100% of the total surface). Since that at the integration the exact requested surfaces cannot be taken, also the real surface for which the intervals were calculated is presented. Some reference points for values of σ equal to $\pm 1, \pm 2$ and ± 3 also appear in the graphic:

$$\sigma = \pm 1 \qquad P = 68.27\%$$
$$\sigma = \pm 2 \qquad P = 95.45\%$$
$$\sigma = \pm 3 \qquad P = 99.73\%$$

As shown in Figure 11, the calculation of the range from the reproduction probability of the values of 80.658% corresponds to $z = \pm 1$ for $X1 = 0.633$ and $X2 = 0646$ (identical to that obtained with the Gaussian function, figure 5). The integral value can also be found, which is nothing but the ratio of the number of the values between the two limits of the given range. This chart allows to find some reference data, namely the proportion of the values between $\sigma = \pm 1, \pm 2$ and ± 3.

The values of the distribution function marked on the graphic were calculated by both the trapezoidal method and also the difference of the integral value between two given boundaries. Also, specific information to choose can be introduced. Either are inserted the average, standard deviation and 2 values of X variable, from which Z values are calculated or Z values are entered directly.

The calcium carbonate buffered acetylsalicylic shows a higher variability of the weight due to the fact that it contains 2 ingredients and each of them contributes to the variability of data.

The benefits of the study are both in educational field as an application of the statistical analysis to a large set of real data and also in the analysis of the drugs quality to select or validate the number of a representative sample. The study can be extent to other data in the quality control of drugs.

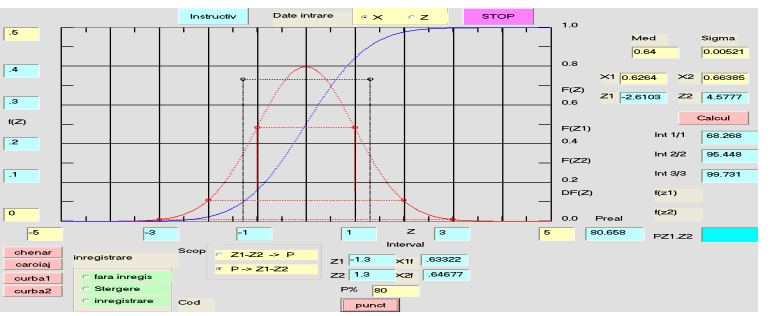

Figure 11. Laplace distribution and frequency function determination by reverse method for AA1

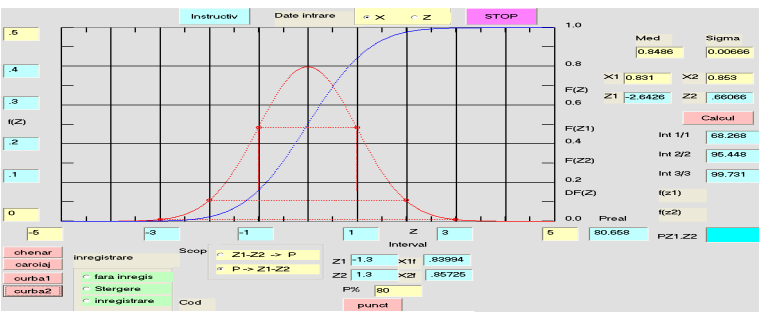

Figure 12. Laplace distribution and frequency function determination by reverse method for AA2

4. Conclusion

The study allows the assessment of the diversity of the weight of acid acetylsalicylic tablets tested upon 100 pieces of two different assortments for which the manufacturing technology is fully automated and even if the same molds are used to form the drugs, different results are obtained.

The standard deviation in the case of buffered acid acetylsalicylic $(AA2)$ is higher than for the acid acetylsalicylic tablet $(AA1)$ which indicates that for 100 weighed values, the weight of aspirin tablets has a lower range than that of buffered acid acetylsalicylic tablets. The histograms show that the weights of acid acetylsalicylic tablets are grouped in a smaller number of classes than those for the buffered acid acetylsalicylic tablets.

Based on the experimental data, the Laplace distribution and the frequency function were calculated, from which the probability can then be deduced. Knowing the probability, Laplace distribution and then the frequency function can be calculated.

The corresponding values of the histogram in the simple aspirin case are more tightly organized in a smaller number of classes while in the case of calcium aspirin the values are more dispersed, which shows a greater deviation between the tablets of calcium aspirin than simple aspirin.

The study can be used in further application allowing the assessment of other drugs with these statistical programs. In addition, this material can give a direction for future research concerning the statistical approach of the experimental rezults.

The present study on the variation of the aspirin tablets weight is addressed to the pharmaceutical companies, which could correct these differences by optmizing the techological process and thus obtaining aspirin tablets with closer mass values, a lower standard deviation and a better precision. In this way, the quality of the products could be improved. On the other hand, this approach could form the basis for other analytical studies on the reproducibility of analytical results such as an investigation regarding the degree of purity of aspirin tablets.

References

[1] Anderson, K. E., Johnson, T. W., Lazovich, D. and Folsom, A. R., *Association between nonsteroidal anti-inflammatory drug use and the incidence of pancreatic cancer*, J Natl Cancer Inst., **94** (2002), 1168–1171

[2] Bosetti, C., Rosato, V., Gallus, S., Cuzick, J. and La Vecchia, C., *Aspirin and cancer risk: a quantitative review to 2011*, Ann Oncol., **23** (2012), 1403–1415

[3] Eltoft, T., Kim, T. and Lee, T. W., *On the Multivariate Laplace Distribution*, Signal Processing Letters, **13** (2006), No. 5, 300–303

[4] Endo, Y., *Estimate of confidence intervals for geometric mean diameter and geometric standard deviation of lognormal size distribution*, Powder Technology, **193** (2009), No. 2, 154-–161

[5] Fraser, D. M., Sullivan, F. M., Thompson A. M. and McCowan, C., *Aspirin use and survival after the diagnosis of breast cancer: a population-based cohort study*, British Journal of Cancer, **111** (2014), 623–627

[6] Friis, S., S∅rensen H. T., McLaughlin, J. K., Johnsen, S. P. and Blot, W. J., *A population-based cohort study of the risk of colorectal and other cancers among users of low-dose aspirin*, Br. J. Cancer, **88** (2003), 684-–688

[7] Goke, B., Schmitz-Moorman, P., Boehme, K., Lange, K. and Arnold, R., *Magenvertraglichkeit von Acetylsalicylsaure bei Zusatz von Calciumcarbonat. Eine Studie in Zwei-Behandlungen*, Med. Klin., **84** (1989), 474–478

[8] Jacobs, E. J., Connell, C. J., Rodriguez, C., Patel, A. V., Calle, E. E. and Thun, M. J., *Aspirin use and pancreatic cancer mortality in a large United States cohort*, J. Natl Cancer Inst., **96** (2004), 524—528

[9] Julean, I., *Chimie analitică informatizată*, Ed. Mirton, Timişoara, 1996

[10] Julean, I. and Hoban, Ş., *Incertitudini la prelucrarea datelor experimentale şi în exprimarea rezultatelor*, Editura Politehnică Timişoara, 2009

[11] Kelly, J. P., Kaufman, D. W., Jurgelon, J. M., Sheehan, J., Koff, R. S. and Shapiro, S., *Risk of aspirin-associated major upper-gastrointestinal bleeding with enteric-coated or buffered product*, The Lancet, **348** (1996), No. 9039, 1413–1416

[12] Larsson, S. C., Giovannucci, E., Bergkvist, L. and Wolk, A., *Aspirin and Nonsteroidal Anti-inflammatory Drug Use and Risk of Pancreatic Cancer: A Meta-analysis*, Cancer Epidemiol Biomarkers Prev., **15** (2006), No. 12, 2561–2564
[13] Lepskiy, A., *On the Stability of Comparing Histograms with Help of Probabilistic Methods*, Procedia Computer Science, **31** (2014), 597–605
[14] López, J., Pagola, P. and Sinusía, E. P., *Applications to the Gamma function and Gauss hypergeometric function*, J. Appr. Theory, **161** (2009), 280–291
[15] Meier, P. and Zund, R., *Statistical Methods in anaytical chemistry*, John Wiley & Sons, 2000
[16] Menezes, R. J., Huber, K. R., Mahoney, M. C. and Moysich, K. B., *Regular use of aspirin and pancreatic cancer risk*, BMC Public Health, 2002, 2:18
[17] Motan, G. and Puia, A., *Studies of different types of aspirin by spectrophotometric methods*, Acta Chemica Iași, **22** 92014), No. 2, 155–164
[18] Olofsson, P., *Probabilities The Little Numbers that Rule Our Lives*, John Wiley & Sons, New Jersey, 2007
[19] Ratnasinghe, L. D., Graubard, B. I., Kahle, L., Tangrea, J. A., Taylor, P. R. and Hawk, E., *Aspirin use and mortality from cancer in a prospective cohort study*, Anticancer Res., **24** (2004), 3177–3184
[20] Schernhammer, E. S., Kang, J. H., Chan, A. T., Michaud, D. S., Skinner, H. G., Giovannucci, E., Colditz, G. A. and Fuchs, C. S.,*A prospective study of aspirin use and the risk of pancreatic cancer in women*, J. Natl. Cancer Inst., **96** (2004), 22–28
[21] Schreinemachers, D. M., Everson, R. B., *Aspirin use and lung, colon, and breast cancer incidence in a prospective study*, Epidemiology, **5** (1994), 138-–146
[22] Sennello, K. A. and Michael, S., *Leib Effects of Deracoxib or Buffered Aspirin on the Gastric Mucosa of Healthy Dogs*, J. Vet Intern Med., **20** (2006), 1291-–1296
[23] Takkouche, B., Regueira-Mendez, C. and Etminan, M., *Breast cancer and use of nonsteroidal anti-inflammatory drugs: a meta-analysis*, J. Natl Cancer Inst., **100** (2008), 1439–1447

Department of Chemistry and Biology
Faculty of Sciences
North University Center of Baia Mare
Technical University of Cluj-Napoca
Victoriei 76, Baia Mare, Romania
Email address: th0mas2004@yahoo.com
Email address: zoita_berinde@yahoo.com
Email address: mihali.cristina@gmail.com
Email address: codre_anca@yahoo.com
Email address: zori_v13@yahoo.com