JOURNAL OF ENGINEERING annals OF FACULTY OF ENGINEERING HUNEDOARA
TOME VI (year 2008), FASCICULE 3, (ISSN 1584 – 2673)

CAUCHY MODEL USING TWO-DIMENSIONAL QUADRANGULAR TRUNCATED DISTRIBUTION

[1.]MAKSAY Ştefan, [1.]BISTRIAN Diana Alina, [2.]MAKSAY Elisabeta

[1.]UNIVERSITY POLITEHNICA TIMISOARA, FACULTY OF ENGINEERING HUNEDOARA
[2.] „SIGISMUND TODUTA" HIGH SCHOOL DEVA

ABSTRACT:
Starting from the Three-sigma rule, according to which, in certain problems in probability theory and mathematical statistics, an event is considered to be practically impossible if it lies in the region of values of the normal distribution of a random variable at a distance from its mean value of more than three times the standard deviation, arose the idea of truncating the classic distributions. This paper introduces a truncated modelling of the two-dimensional Cauchy distribution. The authors suggest a method of truncation in which the limits are determined using the least squares method, under the conditions in which the corresponding probability density approximates ever better a set of numeric data. The results we obtained by simulating the two-dimensional Cauchy rule and that of the corresponding truncated one, are presented comparatively. Calculations have been done in MathCad computer algebra.
KEYWORDS:
Continuous distribution density ; Cauchy distribution ; Truncated data modelling

1. INTRODUCTION

Starting from the Three-sigma rule, according to which, in certain problems in probability theory and mathematical statistics, an event is considered to be practically impossible if it lies in the region of values of the normal distribution of a random variable at a distance from its mean value of more than three times the standard deviation, arose the idea of truncating the normal distribution (see [5]). We are going to give in Section 2 the rule of 2-dimensional Cauchy distribution and the way in which it can be truncated. In order to obtain an optimal truncation corresponding to a set of numerical data, we elaborated a friendly algorithm under MathCad. The modelling we obtained can be compared to the two-dimensional classic one, as it will be shown in Section 3. The application that we are going to analyze will prove that the rule of truncated Cauchy distribution, as it is built in this paper, can be used in data modelling more efficiently than the classic rule. Section 4 presents an analysis of the results we obtained, and includes the conclusions of the paper.

2. THE TRUNCATED 2-DIMENSIONAL CAUCHY LAW AGAINST THE CLASSIC LAW

In probability theory and statistics, the Cauchy distribution is a continuous probability distribution [2, 3, 4]. Its importance in physics is due to it being the solution to the differential equation describing forced resonance. The set of numeric data that are going to be processed is given bellow, where the first two lines represent the values of the independent variables x and y, and the last line represents the independent variable u.

date := augment(x, y, u)

date =

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.4829	1.0601	1.1437	1.3885	1.574	1.7049	1.8914	2.0392	2.1008	2.3057	2.5279	2.776	2.824	2.8866	3.9726
2	2.0933	2.1323	2.1534	2.2267	2.3801	2.7062	2.7818	3.0022	3.0829	3.1074	3.6696	3.836	4.0711	4.3291	4.5142
3	0.0197	0.0386	0.0437	0.0636	0.3004	0.1525	0.1829	0.2119	0.2126	0.1949	0.1042	0.0647	0.0458	0.0141	$279\cdot10^{-3}$

These variables are characterized by

$$mx := \mathrm{mean}(x) \qquad mx = 2.0452$$
$$my := \mathrm{mean}(y) \qquad my = 3.0724$$
$$mu := \mathrm{mean}(u) \qquad mu = 0.1102$$

where mx, my and mu represent the mean values of the respective variables, and

$$sx := \mathrm{stdev}(x) \qquad sx = 0.8518$$
$$sy := \mathrm{stdev}(y) \qquad sy = 0.8042$$
$$su := \mathrm{stdev}(u) \qquad su = 0.0887$$

where sx, sy and su are the standard deviations of variables x, y and u.

With these values, the classic Cauchy two-dimensional law, defined as $fclas: R \times R \rightarrow R$, has the expression

$$fclas(x,y) := \frac{sx \cdot sy}{\pi^2} \cdot \frac{1}{sx^2 + (x-mx)^2} \cdot \frac{1}{sy^2 + (y-my)^2} \tag{1}$$

We introduce the truncated two-dimensional Cauchy density function, given by

$$ftr(x,y,cx,cy) := \begin{cases} K \cdot \left[\frac{sx \cdot sy}{\pi^2} \cdot \frac{1}{sx^2+(x-mx)^2} \cdot \frac{1}{sy^2+(y-my)^2}\right] & \text{if } \left[(x-mx)^2 < (cx \cdot sx)^2\right] \wedge \left[(y-my)^2 < (cy \cdot sy)^2\right] \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where K is constant that will be determined under the circumstances that expression (2) represents a probability density function and therefore has to meet the followed conditions

$$ftr(x,y,cx,cy) \geq 0 \tag{3}$$

and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ftr(x,y,cx,cy) = 1 \tag{4}$$

3. DESCRIPTION OF THE METHOD OF TRUNCATED MODELING

Constants cx and cy are going to be determined by minimizing the error resulted from the use of the truncated function. In order to achieve this, the necessary condition is that the sum of the squares of the differences between the theoretical values of the function and the experimental values u, be minimum.

We shall find the minima with the help of the program we are giving hereinafter.

$$n := \mathrm{length}(x) \quad n = 15$$
$$dxi := 1 \quad dxf := 5 \quad dyi := 1 \quad dyf := 5$$
$$px := 4 \quad py := 4$$
$$fmin := 10^8$$

$$ftr1(x,y,cx,cy) := \begin{cases} \frac{sx \cdot sy}{\pi^2} \cdot \frac{1}{sx^2+(x-mx)^2} \cdot \frac{1}{sy^2+(y-my)^2} & \text{if } \left[(x-mx)^2 < (cx \cdot sx)^2\right] \wedge \left[(y-my)^2 < (cy \cdot sy)^2\right] \\ 0 & \text{otherwise} \end{cases}$$

Using this program, we obtain the values that minimize the function hereinbefore

$$indx = 3 \quad indy = 3 \quad cx\_tr = 3 \quad cy\_tr = 3 \quad Kmin = 1.5816$$

the calculated values of the 5x5 matrix being

fCT =

	1	2	3	4	5
1	0.6383	0.2115	0.1376	0.1113	0.0984
2	0.2098	0.0622	0.0463	0.0427	0.0417
3	0.1389	0.0463	0.0414	0.042	0.0431
4	0.1137	0.0427	0.042	0.0442	0.0461
5	0.1014	0.0417	0.0431	0.0461	0.0484

fCT − fmin=

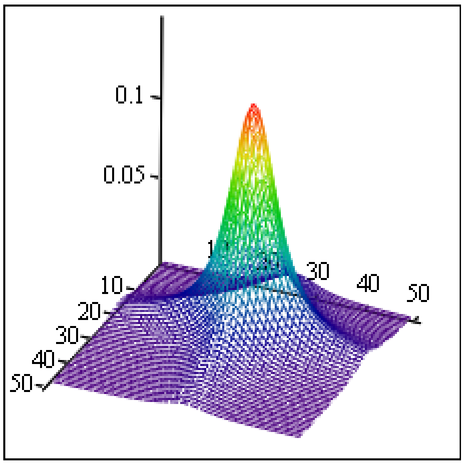
	1	2	3	4	5
1	0.5969	0.1701	0.0962	0.0699	0.057
2	0.1685	0.0208	$4.906\cdot10^{-3}$	$1.3163\cdot10^{-3}$	$2.8394\cdot10^{-4}$
3	0.0976	$4.9345\cdot10^{-3}$	0	$6.1043\cdot10^{-4}$	$1.7335\cdot10^{-3}$
4	0.0723	$1.3374\cdot10^{-3}$	$6.1043\cdot10^{-4}$	$2.8025\cdot10^{-3}$	$4.7295\cdot10^{-3}$
5	0.06	$3.0117\cdot10^{-4}$	$1.7335\cdot10^{-3}$	$4.7295\cdot10^{-3}$	$7.0625\cdot10^{-3}$

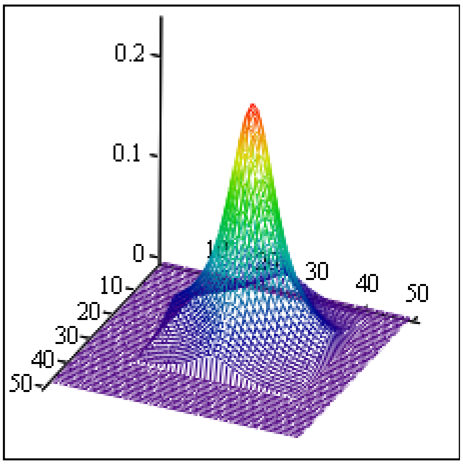
```
method :=  for j ∈ 1.. px + 1
             cx_j ← dxi + (j − 1) · (dxf − dxi)/px
             for k ∈ 1.. py + 1
               cy_k ← dyi + (k − 1) · (dyf − dyi)/py
               Q ← ∫_{mx−cx_j·sx}^{mx+cx_j·sx} ∫_{my−cy_k·sy}^{my+cy_k·sy} ftr1(x, y, cx_j, cy_k) dy dx
               K ← 1/Q
               fCT_{j,k} ← Σ_{i=1}^{n} (K · ftr1(x_i, y_i, cx_j, cy_k) − u_i)^2
               if fCT_{j,k} < fmin
                 fmin ← fCT_{j,k}
                 indx ← j
                 indy ← k
                 Kmin ← K
           for j ∈ 1.. px + 1
             for k ∈ 1.. py + 1
               if fCT_{j,k} < fmin
                 fmin ← fCT_{j,k}
                 indx ← j
                 indy ← k
                 Kmin ← K
           (indx  indy  cx_indx  cy_indy  fCT  fmin  Kmin)
```

These values of Kmin, cx_tr and cy_tr substituted in expression (2) lead to the truncated two-dimensional Cauchy density function in accordance to this modeling, its expression being

$$\mathrm{ftrunc}(x,y) := \begin{cases} 1.5816\cdot\left[\dfrac{0.6851}{\pi^2}\cdot\dfrac{1}{0.8518^2+(x-2.0452)^2}\cdot\dfrac{1}{0.8042^2+(y-3.0724)^2}\right] & \text{if } \left[(x-2.0452)^2<6.5304\right]\wedge\left[(y-3.0724)^2<5.8209\right] \\ 0 & \text{otherwise} \end{cases}$$

The graphic representation of the classic Cauchy density and the truncated probability density obtained for this model are depicted in Figure 1.





FCL FTR

FIGURE 1. Cauchy distribution density and Cauchy truncated two-dimensional law.

For the efficiency of our Cauchy truncated model, which we have introduced in this study, it is imperative that correlation coefficient to be as close as possible to one, greater than the classic correlation and, in the same time, the standard deviation of the truncated model has to be smaller than its classic counterpart. The numerical results are in accordance with our predictions, as followed

$$\text{coefcor\_fclas} = 0.6017 \qquad \text{coefcor\_ftrunc} = 0.8058$$
$$\text{abat\_fclas} = 0.0708 \qquad \text{abat\_ftrunc} = 0.0525$$

4. CLOSING REMARKS

The considerations presented in the paper may lead to the conclusion that the rule of truncated two-dimensional cAUCHY distribution obtained under the given conditions, preserving the properties of a probability density, leads to a higher correlation coefficient and to a smaller deviation than in the case of the classic distribution

Therefore, the optimal modeling of the existent numerical data set can be obtained by means of the truncated distribution. In practical cases the use of truncated modeling can be more expedient instead of using the classic law.

REFERENCES/BIBLIOGRAPHY

[1.] ABRAMOWITZ M, STEGUN IA, ZELEN M, SEVERO NC, „Handbook of mathematical functions-Probability functions". U.S. National Bureau of Standards, Applied Mathematics Series-55,1965.
[2.] McCULLAGH, P., "Conditional inference and Cauchy models", Biometrika, volume 79, 1992.
[3.] WEISSTEIN, E.W., "Cauchy Distribution", at MathWorld.
[4.] GUMBEL, E., GREENWOOD, J., DURAND, D., „The circular normal distribution: theory and tables", Journal of the American Statistical Association, 48, 131-152.
[5.] MAKSAY, Şt., „A probabilistic distribution law with practical applications", Mathematica – Revue D'Analyse numerique et de Theorie de L'Approximation, Tome 22 (45), Nr. 1, pp. 75-76, 1980.
[6.] MAKSAY, Şt., BISTRIAN, D., „Considerations upon the von Mises 2-dimensional distribution", Journal of Engineering, Annals of Faculty of Engineering Hunedoara, Tome V, Fascicole 3, pp. 175-180, 2007.