CONSIDERATIONS ON COMMON PRACTICAL APPLICATIONS OF NORMAL DISTRIBUTION

CĂTĂLIN-ILIE MITRAN [*]

ABSTRACT: *In the case of a large number of practical connsiderations, such as those related to the distribution coefficient of the population's intelligence, the ratio between motivation and performance, but also the distribution of the average of an exam in the case of a large number of competitors, it can be seen that these tend to follow similar results to those given by the normal distribution. For these reasons, the study of these phenomena can be described using the normal distribution.*

KEY WORDS: *normal distribution, average value, probability density function.*

JEL CLASSIFICATIONS: *C15, C46*

1. INTRODUCTION

The normal distribution is probably the most important and most widely used in probability theory and mathematical statistics. The distribution density function of the normal distribution has the form (Despa & Vişan, 2003):

$$f(x,m,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}, x \in \mathbb{R}, m \in \mathbb{R}, \sigma > 0 , \quad (1)$$

where *m* and σ having a meaning that we shall specify immediately.

2. THEORETICAL ASPECTS

Definition 1 The function $f(x,m,\sigma)$ is called a distribution function and needs to check the properties:

1) $f(x,m,\sigma) \geq 0, \forall x \in \mathbb{R}$. But this is obvious because $\sigma > 0$;
2) $\int_{-\infty}^{\infty} f(x,m,\sigma)dx = 1$.

We have that:

[*] *Lecturer, Ph.D., University of Petroșani, Romania*, cata.miha.ilie@gmail.com

$$\int_{-\infty}^{\infty} f(x,m,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = 1,$$

assuming that the variable change $\frac{x-m}{\sigma} = t$ has been made and also taking into account that

$$\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi},$$

which is the so-called Euler-Poisson integral.

Usually is denoted by $N(m,\sigma)$ the set of all variables with a normal distribution of parameters *m* and σ. It can be immediately shown that the average value of such a random variable is

$$M(X) = \int_{-\infty}^{\infty} x\, f(x,m,\sigma) dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} x e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx = \cdots = m \qquad (2)$$

and also that the dispersion of such a random variable is

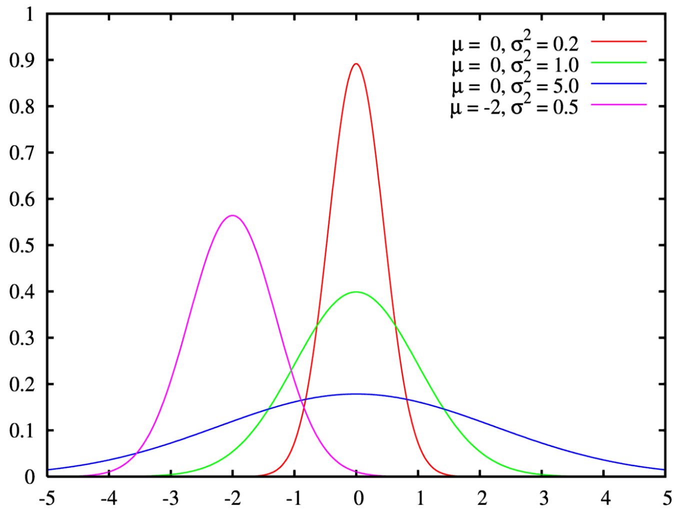
$$D(X) = \int_{-\infty}^{\infty} (x-m)^2 f(x,m,\sigma) dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} (x-m)^2 e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx = \cdots = \sigma^2 \quad (3)$$

That means that the parameters *m* and σ from the expression of the probability density function represents the mean value and the mean square deviation of normal random deviation.

Figure 2 shows representation of the normal distribution in cases:

1) $m = 0, \sigma^2 = 0.2$ (red curve);
2) $m = 0, \sigma^2 = 1.0$ (green curve);
3) $m = 0, \sigma^2 = 5{,}0$ (blue curve);
4) $m = -2, \sigma^2 = 0.5$ (pink curve).

We only mention that in the mentioned figure the mean value *m* is denoted by μ.



Source: htpps://ro.wikipedia.org/wiki/Distributia_Gauss#/media/Fisier:Normal_distribution_pdf.png

Figure 1. Graphical examples of some normal distributions

3. EXAMPLES OF OUTSTANDING APPLICATIONS OF THE NORMAL DISTRIBUTION

A first application of the normal distribution is related to the distribution of population intelligence. But perhaps an even more interesting result is the one related to the distribution of human imbecility. Interesting results and considerations about human imbecility we find in (Săvoiu, 2020), pp. 80-83 in which the author of the book speaks about the Carlo Cipolla's exceptional reflections on the laws of imbecility, stupidity and incompetence in his famous book *The basic Laws of Human Stupidity,* from 1987. Unfortunately, this book is translated in southern and eastern Europe quite late, precisely in 1998 and 2014.

As expected, the population intelligence follows the laws of normal distribution.

As far as we are concerned, related to the distribution of the imbecility, Cipolla's conclusion would be that the real percentage of imbecility always higher that an eventual forecast of it, the normal distribution turns into an abnormal one or long-tailed one to which the theory of the 6 sigma in which 99.73% of those investigated are belonging to are no longer verified due to some kurtosis and skewness coefficients.

In Figure 2 are presented the Gaussian or normal distributions of human intelligence (red curve) and the long-tailed distribution of human imbecility (htpps://researchgate.net/figure/Distributia-cu-coada-lunga-a-imbecilităţii-dupa-numar-si-impact-confruntate-cu-fig3-344103371):

1) Gaussian or normal distribution of human inteligence (red curve);
2) Long-tailed distribution of human imbecility (turquoise curve).

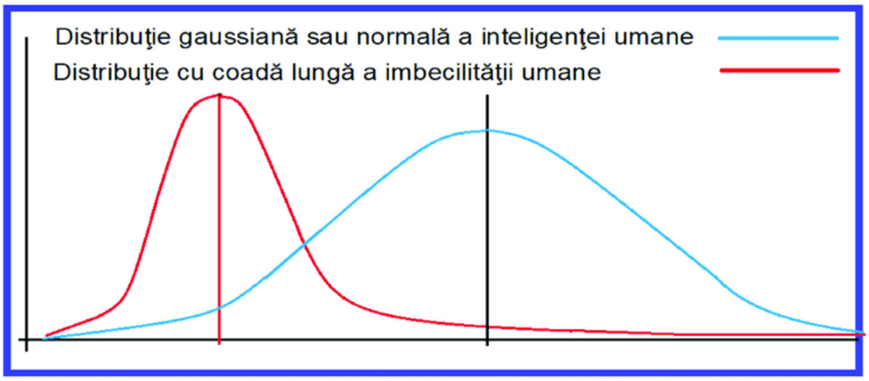


Figure 2. Normal distribution of human intelligence and long-tailed distribution of human imbecility

As a confirmation of Carlo Cipolla's theories, the opinions of the students and two researches carried out between 2017 and 2019 validate these ideas initially considered simple assumptions.

Table 1. Translation

Romanian text	*English translation*
Distribuție gaussiană sau normală a inteligenței umane	Gaussian or normal distribution of human inteligence
Distribuție cu coadă lungă a imbecilității umane	Long tail distribution of human imbecility

Another problem the ratio between motivation and performance. Results are also obtained here that verify the hypothesis of normal distribution (htpps://diplomade10.ro/distributia-normala-in-spps/). The method is called the forced distribution of results on the Gauss curve.

The method can be applied very well to large groups, but it is usually not valid for smaller groups.

The main result on the performance obtained from the study carried out of a large population is the next:

1) 40% of the population achieved average results;
2) 20% of the population achieved above-avererage results;
3) 20% of the population achieved bellow-average results;
4) 10% of the population achieved high results;
5) 10% of the population achieved low results.

Remark: It is obvious that a classification with 5 categories was used for the general case.

If we use a classification withonly three categories the result of the evaluation would be as follows:

1) Top performance, 10%;
2) Average performance, 80%;
3) Low performance, 10%.

The results obtained were used by the General Electric company to eliminate those with below expectations performance.

But the gradient for the forced distribution is different and looks like:

1) 10% -exceptional;
2) 15%-strong;
3) 50%-valuable;
4) 15%-borderline;
5) 10%-inefficient.

In the case of Sun Micro Systems a three-category scale came out as follows:

1) 20%- high performance;
2) 70%-full contributor;
3) 10%-below expectations.

Let us consider the distribution of the average of an exam in the case of a large numbers of competitors. Accessing the link *htpps://marastan.wordpress.com/top-content/uploads/2017/03/mngm-performantei-si-talente-globale.pdf* we can find a so-called *Empirical rule or rule 68- 95-99.7*. The main idea of this rule is the next:

1) Approximatively 68% of the values belongs to the interval $(m-\sigma, m+\sigma)$;

2) Approximatively 95% of the values belongs to the interval $(m - 2\sigma, m + 2\sigma)$;
3) Approximatively 99.73% of the values belongs to the interval $(m - 3\sigma, m + 3\sigma)$.

Example: As an example it is considerd the case where $m = 1150$ and $\sigma = 150$. The result in this case is:

1) Approximatively 68% of the values are between 1000 and 1300;
2) Approximatively 95% of the values are between 850 and 1450;
3) Approximatively 99.7% of the values are between 700 and 1600.

But if in the case of a large number of subjects it is veryclear that the results obtained follow the normal distribution, what happens in the case of analyzing the results on a small number of subjects?

In the case of a relatively small number of pupils subjects to examination we have the example that can be deduced from Figure 3 (ro.linkedin.com/pulse/normalitatea-datelor-solutii-pentru-analiza-constantin-bolocan):

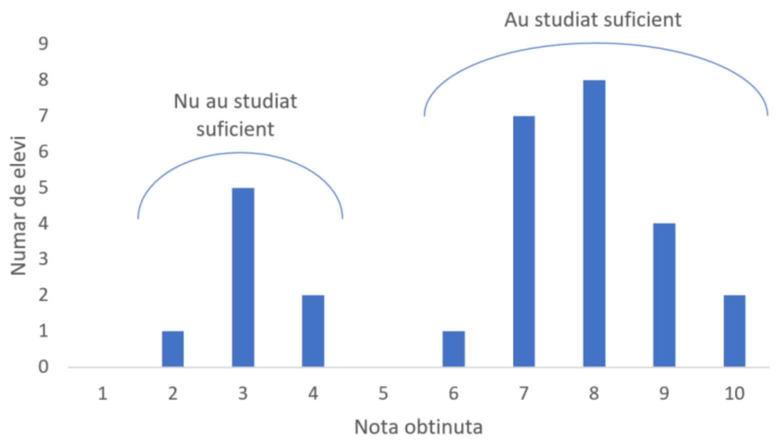


Figure 3. Distribution graphs obtained in the case of a small number of subjects

Table 2. Translation

Romanian text	***English translation***
Numar de elevi	Number of students
Nu au studiat suficient	They haven’t study enough
Au studiat suficient	They have study enough
Nota obtinuta	Grade obtained

The number of pupils to the grade obtained is as follows:

1) 1 pupil with a grase of 2;
2) 5 students with a grade of 3;
3) 2 pupils with a grade of 4;

4) 1 pupil with a grade of 6;
5) 7 pupils with a grade of 7;
6) 8 pupils with a grade of 8;
7) 4 pupils with a grade of 9;
8) 2 pupils with a grade of 10.

The pupils with grades 2,3 or 4 are considered with an insufficient grade. The pupils with the grades 6,7,8,9 or 10 are considered with a sufficient grade. Even in the example considered above, in the absence of information about grade 5, it is clear that the results do not follow a normal distribution. In the case of data considered non-normal, the author of the article recommends some available techniques to perform quality analyses (ro.linkedin.com/pulse/normalitatea-datelor-solutii-pentru-analiza-constantin-bolocan):

1) Parametric tests: 1-sample T, 2-sample T, One-Factor ANOVA, Two-Factor ANOVA, Pearson corelation;
2) Nonparametric tests: 1-sample Sign or Wilcoxon, Mann-Whitney, Kruskal-Walis or Mood, Friedman, Spearman corelation.

4. CONCLUSIONS

In the current paper, three remarkable examples of the practical applications of the known results related to the normal distribution have been presented. Although their list is significantly longer, the three were considered by the author to be among the most relevant checks. We also mention that, in the case of the graphs in figure 2 and figure 3, it was preferred to take them with terms and explanations in Romanian, of course with the mention of the provenance in the bibliography and the subsequent translation into English of Romanian terms used.

REFERENCES:

[1]. **Despa, R.; Vișan, C.** (2003) *Applied Mathematics in economics*, Sylvi Publisher, București
[2]. **Săvoiu, G.** (2020) *Statistical methods and cross-disciplinarity, 1st Edition*, Editura Universitară, Pitești
[3]. htpps://diplomade10.ro/distributia-normala-in-spps/
[4]. htpps://marastan.wordpress.com/top-content/uploads/2017/03/mngm-performantei-si-talente-globale.pdf
[5]. htpps://researchgate.net/figure/Distributia-cu-coada-lunga-a-imbecilităţii-dupa-numar-si-impact-confruntate-cu-fig3-344103371
[6]. htpps://ro.wikipedia.org/wiki/Distributia_Gauss#/media/Fisier:Normal_distribution_pdf.png
[7]. ro.linkedin.com/pulse/normalitatea-datelor-solutii-pentru-analiza-constantin-bolocan

This article was reviewed and accepted for presentation and publication within the 11th edition of the International Multidisciplinary Symposium „UNIVERSITARIA SIMPRO 2024”.