# Application and Properties of the Probability Density $Aexp(-(x-c)^2/(a(x-c)+2b^2))$

A.A. Kirillov, I.A. Kirillov
*D.V. Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, 1/2, Leninskie Gory, Moscow, 119992, Russia*


## Abstract

The form of functions of the probability density $Aexp((x-c)^2/(a(x-c)+2b^2))$ ( defined in a half-line, asymmetric, with region of maximum as that of the normal distribution and exponentially attenuating at large $x$ values) enables to describe a wider class of processes than just normal distribution. Similar to the normal distribution, this function can be used to describe the measurement error distribution, its asymmetry being taken into account. The function is useful for simulation process and is the basis for physical description of UHE cascades. Properties of the probability density, formulae for calculation its parameters are presented in brief. One can operate with this distribution as well as with normal distribution.


## 1. Introduction

There are many phenomena that have: 1) limitation of their characteristics, 2) large number of relatively small factors composing value of the characteristics, 3) velocity of development of the characteristics is proportional to their values; — as their principal intrinsic properties. Indeed, a phenomenon has start point in space or time and its geometry, time, or another characteristics (e. g. energy is positive) are limited; rate of development of characteristics is proportioning to the value of the characteristic defines its exponential development. To describe these phenomena, the main features of the function should be the following: 1) the domain of definition is half-line, 2) the maximum region is similar to the maximum region of the normal distribution, 3) exponential attenuation.

For example, this-kind phenomena are ultra high energy cascades. Approximation of the individual model cascades at $t \geq 0.5t_m$ by function $N_m exp(-(t-t_m)^2/(a(t-t_m)+2b^2))$ was successfully used [4]. The parameters of the function have interpretation: $t_m$, $N_m$ are for the cascade maximum position, $a$ is the asymmetry parameter, $2b^2$ is the square parameter. Paper [4] deals with methodic of construction of approximation by the form and estimates its accuracy. Paper [5] as well concernes using the form for a simulation method and discusses problems of distribution function of the parameters. For this approach each parameter is

pp. 535–538 

a random variable defined by its distribution function. An individual cascade is described by a set of concrete values (realization) of these random variables. That way, to describe the cascades, it is necessary and sufficient to produce distribution functions of the parameters. Investigation [2] of distribution functions of the parameters $t_m, N_m, a, b$ approved that their probability density can be approximated by form:

$$A(a,b)exp(-(x-c)^2/(a(x-c)+2b^2)) \ \ a \geq 0, \ \ b \geq 0, \ \ x \geq c - 2b^2/a. \qquad (1)$$

We hope that application of form (1) is not limited by the cascade theory , so in brief the paper presents main results of investigation of the general properties of form (1) as probability density.

## 2. Possibility application for approximation random values

One can see from (1) that it is unimodal, bell-shaped, left-asymmetric curve resembling the well-known and wide used probability density (the Maxwell distribution, $\chi^2$ - Pearson, $\gamma$ - distribution ...). It is obvious that form (1) is a general cause of the normal and exponential distributions but converging to them at $a = 0$ or $b = 0$, correspondingly. Approximation of any distribution by formulae simple and comfortable for analytical and calculation methods bought is an ordinary task often present in simulation and data processes. The normal distribution is often used not only when physical process leads to the stability distribution [3], but (in our opinion) the distribution is suitable: one look at a histogram is enough for estimating its parameters and then, if an accuracy of the approximation (standard error) is satisfied for the concrete conditional, the distribution is used. Probability density (1) inherits the suitability, and additional parameter $a$ essentially decreases the standard error in many causes.

The form of function (1) modifies the normal distribution in that direction as Edgeworth series [1], thus form (1) approximate probability density of finite set of random values better than normal ones in many causes. For example, approximation of composition for $n = 2$–$4$ exponentially distributed components by form (1) has an accuracy similar to approximation for $n = 20$–$30$ components by normal probability density, that is, standard errors of approximation by (1) are by a factor of 50 less then those obtained with the normal distribution and the factor is increasing with $n$.

## 3. Properties of the probability density

As generalization of normal distribution one can obtain for: 1) Normalizing coefficient $A(a,b) = (azexp(z)K_1(z))^{-1}$, where $z = (2b/a)^2$ – scaling invariat parameter, $K_1(z)$ is the modified third-kind Bessel function; 2) Mathematical expectation $< X > = c + a[1 - (z^{-1/2}/2)e^{-z}K_1(z)^{-1}F(z)]$, where $F(z) =$

$\int_{-z^{1/2}}^{\infty} exp(-x^2/(2z^{-1/2}x + 2))(z^{-1/2}x + 2)^{-2}dx$; 3) Variance (dispersion) $D = a^2[(< X > - c)/a + (1 - (< X > - c)/a)(z + (< X > - c)/a)]$; 4) Width of a distribuion $W(d)$ on level $d$ is defined as interval between values of the argument $x$, where probability density is less by the factor $d$ than at his maximum (mode) $W(d) = (a^2(lnd)^2 + 8b^2lnd)^{1/2}$. One can obtain useful and easy interpreted relations for parameters $a$ and $b$: $a = (W_r(d) - W_l(d))$, $b = (W_l(d)W_r(d)/(2lnd))^{1/2}$, where $W_l(d)$ and $W_r(d)$ are left and right (from the mode) parts of $W(d)$, thus $W(d) = W_l(d) + W_r(d)$; 5) At the area the closest to normal cause for feasible proximate of the central limiting theorem it can be shown that $< X > - c \geq (3/4)a/(1 - (2z)^{-1})$, $D \approx b^2 + a^2$.

To obtain numerical values of the normalizing factor, the mathematical expectation (mean value) and the variance, the scaling invariant parameter $z$ can be calculated and then the necessary values $A(a, b)$, $< X >$, $D$ can be calculated directly from the above formulae, or can be found using the table 1. For practical using it is convenient to prepare graphs-nomogramms based on the table 1.

**Table 1.** Values of functions for calculation of normalizing factor, a mathematical expectation and a variance.

| z | 0.001 | 0.01 | 0.1 | 1 | 10 | 100 | 1000 |
|---|---|---|---|---|---|---|---|
| $A(a, b) * a$ | .99900 | .98980 | .91826 | .61119 | .24345 | .07949 | .02522 |
| $(< X > - c)/a$ | .99950 | .99523 | .96230 | .84972 | .76708 | .75186 | .75019 |
| $D/a^2$ | 1.0000 | 1.0000 | 1.0023 | 1.1277 | 3.2750 | 25.752 | 250.75 |

## 4. Summary and example

Although the properties of probability density (1) make feasible their wide practical application, its usage in various fields (in metrology, for description of measurement error distribution when its asymmetry is taken into account, in the probability theory for construction estimates of composition of finite number of function) should be treated as preliminary.

Let us examine example of the study of the depth of maximum of cascades in the atmosphere, using the information from [6]. The function of distribution of the depth of maximum of 500 simulated cascades generated by protons of energy $10^{17}$(eV) in the atmosphere at depth $t(g/cm^2)$ is shown in [6] as the histogram presented in the left top fragment of fig. 1. The central part of fig. 1 shows the result of approximation of this initial information by distribution (1) at $A(a, b) = 7.15 * 10^{-3}$, $a = 36.56$, $b = 52.26$, $c = 670$. This figure shows interpretation of normalizing factor $A(a, b)$, parameters $a, b, c$, width $W = 152, 27$ and its parts $W_l = 57.85$ and $W_r = 94.41$ at the level $d^{-1} = e^{-1}$. There are also specified the

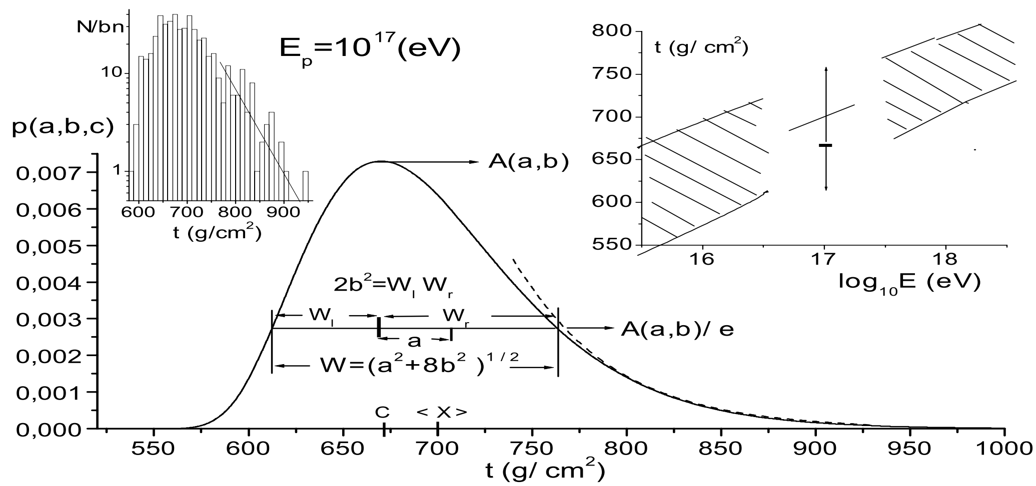


**Fig. 1.** Example of using of distribution with density $A(a,b)exp(-(x-c)^2/(a(x-c)+2b^2))$ and interpretation of its parameters.

distribution mode ($mode = c = 670(g/cm^2)$) and the mathematical expectation (mean ) ($< X >= 698.3(g/cm^2)$). The right top part of fig. 1 shows a fragment of the figure from [6], represented by the shaded strip corresponding to the central area, containing 68% of distributions, dependence on the energy of maximum depth of these cascades.

In the area $E = 10^{17}(eV)$ we give a version of representation of results, by the usual graph of the mean value and the values $W_l$ and $W_r$ at the level $d^{-1} = e^{-1}$ . The mode designated by horizontal stroke. This presentation is quite enough for reconsruction the distribution of form (1).

## 5. References

1. Cramer H. 1946, Mathematical methods of statistics. Stockholm.
2. Dedenko L.G., Kirillov A.A., Fedorova G.F. 2002, 18th European Cosmic Ray Symp. (Moscow), HE 56p.
3. Feller W. 1967, An introduction to probability theory and its applications.v.2. NewYork London Sidney
4. Kirillov A.A., Kirillov I.A. 2001, Proc. 27th Int. Cosmic Ray Conf. (Hamburg), v.2, 483-486
5. Kirillov A.A., Kirillov I.A. 2003, Astroparticle Physics, v.19, 101-114
6. Pryke C.L. 2001, Astroparticle Physics, v.14, 319