# Finding the Fine Structure of the Solutions of Complicate Logical Probabilistic Problems by the Frequent Distributions

Anatoly V. Belyakov
E-mail: belyakov.lih@gmail.com

The Author suggests that frequent distributions can be applied to the modelling the influences of stochastically perturbing factors onto physical processes and situations, in order to look for most probable numerical values of the parameters of the complicate systems. In this deal, very visual spectra of the particularly undetermined complex problems have been obtained. These spectra allows to predict the probabilistic behaviour of the system.

Normal distribution, also known as the Gauss distribution, is a distribution of the probabilities ruling physical quantities and any other parameters in general, if the parameters are affected by a large number of purely stochastic processes. The normal distribution plays a highly important rôle in many fields of knowledge and activity of the Mankind. This is because of all distributions, which may be met in the Nature, the most frequent is the normal distribution. In particular, the normal distribution sets up the law of the Brownian motion — the fluctuations of Brownian particles being affected by the probabilistically perturbing factors such as the heat motion of molecules. In these fluctuations, the consecutive changes of the particles' location are independent from the last events in them, and their any current location can be assumed to be the initially start-point.

As an example of another sort, a simplest situation of the theory of games can be provided. In this example, an initially rate $S_0$ increases proportionally to the progression coefficient $q_1$ with a probability of $p_1$, or decreases proportionally the progression coefficient $q_2$ with a probability of $p_2$. As is obvious, the pair of these numerical values are connected to each other here: these are the current and past values connected as $S_{i+1} = S_i\, q_i$.

However in the core of this problem, the examples are a manifestation of the same situation, because $S_0$ can be meant as any parameter under consideration in a process being affected by perturbing factors.

It is clear that, having duration of the process unbounded, the numerical value of the parameter $S_0$ will vary near an average value, then filling, step-by-step, the arc of the normal distribution.

The current value $S_i$ should return back to this average value each time after a number of the steps passed in the ways of different lengthes under stochastic alternating $q_1$ and $q_2$. Therefore, concerning the parameters of the perturbing effects in the perturbation series, the set of the current numerical values of the parameters is different in the cases of both sequent and parallel observations. Thus, it seems that there should not be "spectra" or "non-uniformities" in the Gauss arc. On the other hand, the Gauss distribution is a particular case of more complicate distributions, where the smooth form of the Gauss distribution is only an idealisation of those. Because some numerical values can meet each other in the series of the observations, the frequent distribution* of the *sum of all numerical values registered in many series* manifests the preferred numerical values of $S_i$ thus producing by this its own specific spectrum.

Note that the discrete nature of normal distributions was experimentally discovered in different physical processes in already the 1950's by S. E. Shnoll [1].

Figures 1–3 show examples of the frequent spectra which came from the normal distributions being affected by two, three, and four perturbing factors (the progression coefficients $q_i$). The ordinate axis shows the number of coincident numerical values. The axis of abscissas shows the current values of $S_i$ in doles of the initially value. These numerical values were given, for more simple and convenient comparing the histograms, in the same interval of abscissas from 0.0001 to 10000, while the initially parameters were assumed to be such that the axis of the distribution crosses the initially sum $S_0$. The diagrams were obtained by summing 500 series of 500 steps in each (so the common number of the values is $500 \times 500 = 250000$). The relative length of the current interval $g$ was assumed $10^{-6}$ of the current value $S_i$. The algorithmic language C++ was used in the calculation.

This is a fragment of a computer program

```
for ( int t = 1; t < 500; t ++ )   {
double  Si = 1;
for ( int u = 1; u < 500; u ++ )   {

if ( a >=  b && a >= c ) {qi = q1 ; goto nn ; }
if ( b >=  a && b >= c ) {qi = q2 ; goto nn ; }
if ( c >=  a && c >= b ) {qi = q3 ; goto nn ; }
nn:    Si = Si*qi ;
if ( Si < 10000 && Si > 0.0001 )
i++ ,  m[ i ] = Si ;               }
                                   }
```

*Frequent distributions provide a possibility for bonding the probability of the appearance of numerical values of a function in the area where it exists. That is, the frequent distributions show the reproducibility of numerical values of the function due to allowed varying its arguments. There is a ready-to-use function "frequency" in MS Excel; any other software can be applied as well.

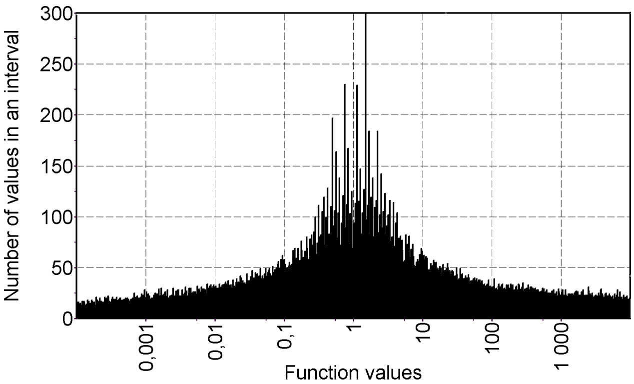


Fig. 1: Frequent distribution obtained with $q_1 = 1.5$, $q_2 = 0.5$, $p_1 = 0.555$, $p_2 = 0.444$; number of steps in the series is 500, number of the series is 500; number of the numerical values in the scale 190,000 (of those, nonzero intervals are 8,000).

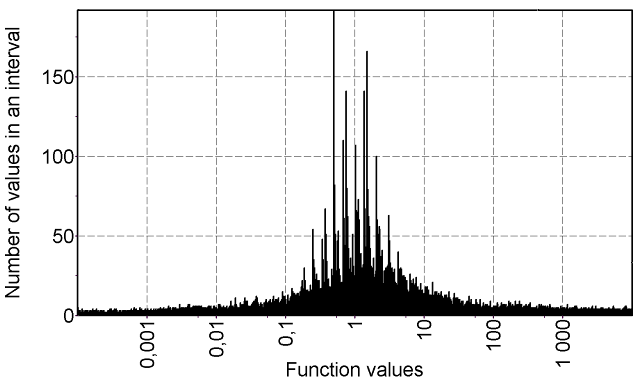


Fig. 2: Frequent distribution obtained with $q_1 = 1.5$, $q_2 = 0.5$, $q_3 = 1.37$, $p_1 = 0.333$, $p_2 = 0.333$, $p_3 = 0.333$; number of the steps in the series is 500, number of the series is 500; number of the numerical values in the scale is 180,000 (of those, nonzero intervals are 62,000).

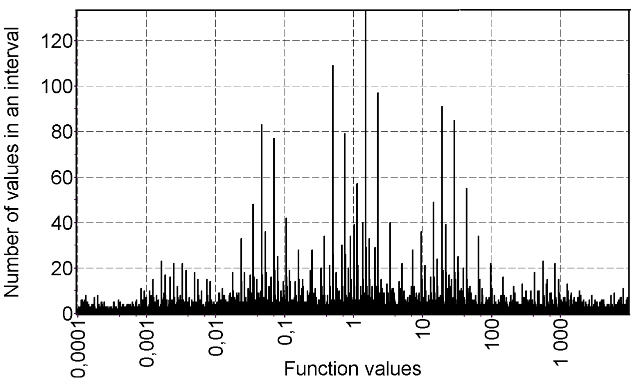


Fig. 3: Frequent distribution obtained with $q_1 = 1.5$, $q_2 = 0.5$, $q_3 = 19.3$, $q_4 = 0.047$, $p_1 = 0.294$, $p_2 = 0.235$, $p_3 = 0.235$, $p_4 = 0.235$; number of the steps in the series is 500, number of the series is 500; number of the numerical values in the scale is 67,000 (of those, nonzero intervals are 28,000).

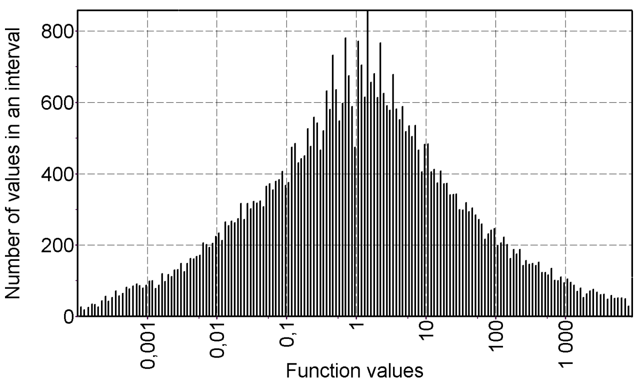


Fig. 4: Frequent distribution obtained with $g = 0.1$ from the current numerical value $S_i$; here $q_1 = 1.5$, $q_2 = 0.5$, $p_1 = 0.555$, $p_2 = 0.444$; number of the steps in the series is 100, number of the series is 500; number of the numerical values in the scale is 48,000 (of those, nonzero intervals are 173).

modelling the change of the parameter $S_i$ and the set of a massive data of $S_i$, in look for the frequent distributions obtained due to three perturbing factors $q_1$, $q_2$, $q_3$. Here *a*, *b*, *c* are prime numbers which stochastically change (the computer program contains a function which generates random numbers), in each single cycle of the observation, along the intervals whose length is proportional to their probabilities $p_1$, $p_2$, $p_3$.

The graphs manifest that fact that, in the common background of the numerical values of the current parameters, there is only minor number of those whose probability exceeds the average value in many times. Besides that, the exceeding numerical values depend on the numerical values of the progression coefficients, but are independent from the length of the series (the number of the steps). Increasing the number of the perturbing factors does not make the non-uniform distribution more smooth, as it should be expected. Contrary, the non-uniformity of the distribution increases: in this process the allowed current values $S_i$ occupy more square of the graph, while their number in the given section of the axis *x* decreases. Therefore a small probability of that the current values $S_i$ will valuable shift from their average positions appear due to the appearance of the long chains of the co-multipliers which have the progression coefficients larger (or lesser) than unit. If the progression coefficients differ valuable from each other, the histogram manifest distributions of high orders (see Figure 3).

Consider an ultimate case where all perturbing factors, i.e. the progression coefficients $q_i$, differ from each other by the numerical values, and there is not their coinciding numerical values in the series. This situation can easy be modelled, if setting up in the computer program that the progression coefficients have a connexion with the counters of the cycles *t* and *u*, or that they are varied by any other method. In this case, in a limit, the amplitude of the numerical values in the histogram will never exceed unit, nowhere, while the *frequent non-uniformity* will still remain in the distribution. Therefore, even if extending the length of the unit interval, the same dis-

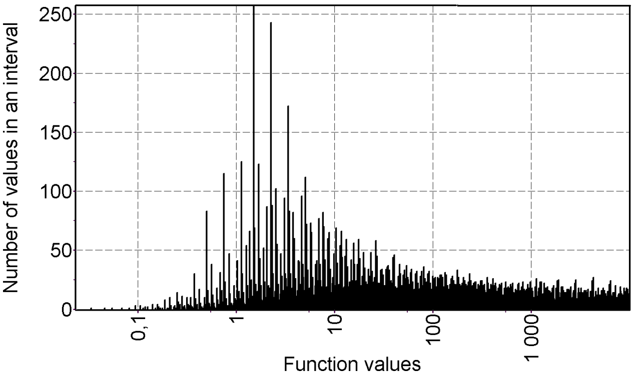


Fig. 5: Non-symmetric frequent distribution obtained according to the data of Fig. 2; number of the numerical values in the scale is 22,000 (of those, nonzero intervals are 2,400).

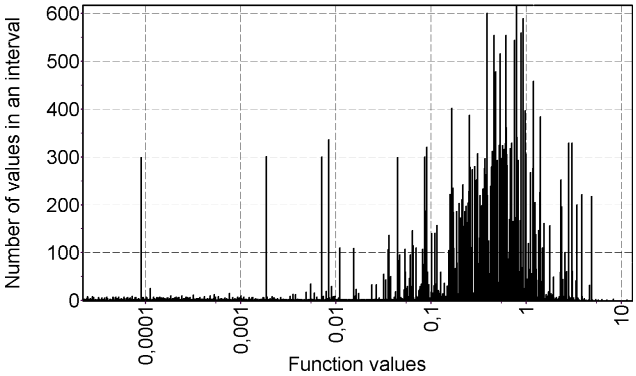


Fig. 6: Frequent distribution of the solutions of the quadratic equation $x^2 - 2Bx + C = 0$ with $q_1 = 1.33$, $q_2 = 0.71$, $q_3 = 1.33$, $q_4 = 0.71$; here $p_1 = p_2 = p_3 = p_4 = 0.25$; number of the steps in the series is 300, number of the series is 300; number of nonzero intervals is 16,000. All geometric coefficients of the progression are independent from each other.

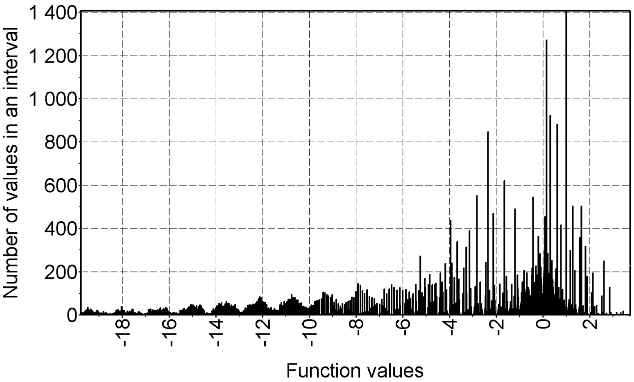


Fig. 7: Frequent distribution of the solutions of the quadratic equation $x2 - 2Bx + C = 0$ with $q_1 = 0.71$, $q_2 = -0.71$, $q_3 = 0.71$, $q_4 = -0.71$; here $p_1 = p_2 = p_3 = p_4 = 0.25$; number of the steps in the series is 250, number of the series is 250, number of nonzero intervals is 1,800. All arithmetic coefficients of the progression are independent from each other.

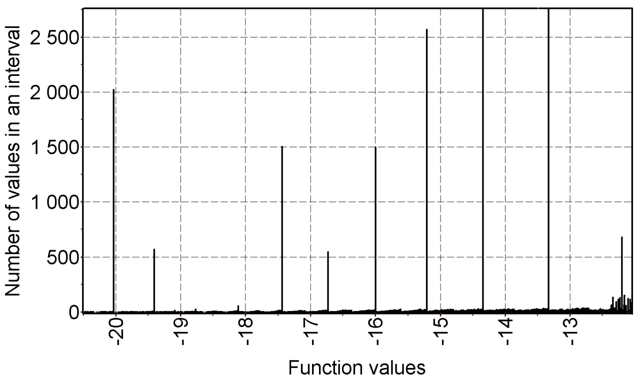


Fig. 8: Frequent distribution of the solutions of the quadratic equation $x2 - 2Bx + C = 0$ with $q_1 = 0.127$, $q_2 = 1.13$; $p_1 = 0.465$, $p_2 = 0.535$; number of the steps in the series is 500, number of the series is 10,000; number of the numerical values in the scale is 27,000, number of nonzero intervals is 1,350. All arithmetic coefficients of the progression are dependent on each other.

tribution takes the amplitudal discrete shape again. Finally, under truncating the number of the intervals (this, generally speaking, means analysis of the given process with a lower precision), the graph takes a shape of almost the smooth normal distribution (see Figure 4).

It is possible to suppose that the discreteness of normal distributions (and, as is obvious, any other distributions as well) is their core property originated from that the rational numbers are distributed with different density along the axis of numbers [2, 3].

Shapes of the histograms depend on specific parameters; they may be very spectacular. So, in the bit of the computer program that was given above, each perturbing factor realizes itself independent from the others. If however, for instance in the first condition, one replaces the logical "and" with the logical "or", the distribution changes its shape very much (see Figure 5).

So forth, Figures 6–9 show illustrative examples of the versions of the frequent distributions of one of the solutions of a quadratic equation $x^2 - 2Bx + C = 0$, where we see iterrationally correcting two parameters $B$ and $C$ whose initially numerical values are units.

In the example shown in Figure 6, the progression coefficients are geometric, and are independent from each other. The parameter $B$ is under a correction by the coefficients $q_1$ and $q_2$, while the parameter $C$ is under a symmetrical correction by the coefficients $q_3$ and $q_4$. Specific to the graph is that, somewhere left from the main distribution, in the background of many dense numerical values whose probabilities are very small, a small number of the numerical values having a very high probability appear (they experience a shift to the side of small numerical values of the function).

In the other examples shown in Figures 7 and 8, the progression coefficients are arithmetic. In the distribution shown

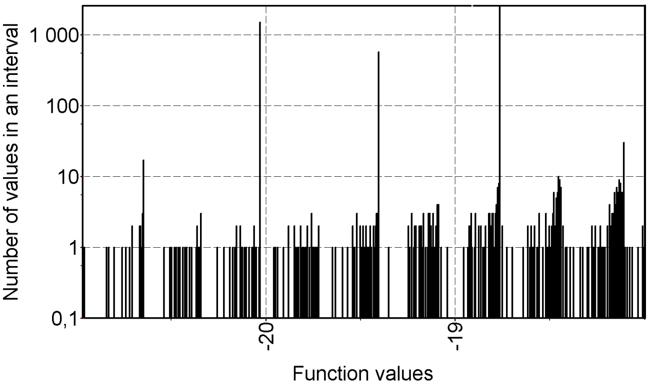


Fig. 9: A fragment of the frequent distrinution according to the data of Fig. 8; number of the numerical values in the scale is 5,300 (of those, nonzero intervals are 220).

in Figure 7, four progression coefficients are present; they are symmetric. The histogram is built by a set of the Gauss arcs of the first, second, and higher orders which fill the side of negative numerical values. The distance between the arcs, and their shape depend on the numerical value of the progression coefficients. In Figure 8, we give a part of the quadratic function distribution in the region of negative numerical values of its solutions taken under two coefficients $q_1$ and $q_2$, where the parameters $B$ and $C$ are additionally connected to each other, and their correction is produced commonly for them. The respective bit of the algorithm has the form:

```
for ( int t = 1; t < 10000; t ++ )      {
double  B  =  C = 1;
for ( int u = 1; u < 500; u ++ )        {

if ( a >=  b )   B = B + q1 , C = C - sqrt(q2) ;
if ( b >=  a )   C = C + q2 , B = B -  q1 ;
if ( B*B - C > 0 )
Si =  B - sqrt ( B*B - C ) ;
i++ ,  m[ i ] = Si ;                    }
                                        }
```

Here, as well as in the example shown in Figure 7 (but with more obvious visibility), that fact is manifested that the overwhelming number of the numerical values, i.e. the probable solutions of the function obtained under the variation of the parameters $B$ and $C$, have an infinitesimally small probability in the scale, while the probability of the solutions is concentrated in a very small number of the solutions where it thus is very high. In the fragment of the histogram taken in a semi-logarithmic scale (Figure 9), is is clearly seen that the peaks of the maxima "grow up" from the frequent concentrations of the numerical values of the functions in the axis $x$. Should this mean that, in the case of similar distributions of a macroscopic system having an arbitrary number of solutions (degrees of freedom), the macroscopic system under a specific set of the parameters acting in it can be in selected special discrete (quantum) states, i.e. the system can have discrete solutions?

It is absolutely obvious that, first, such maximally probable solutions are mostly interested in processes and phenomena we study. Finding these solutions by some other methods that the method given above would be very complicate. Of course, in formulating algorithms for similar problems (obtaining the massive of the required values and their distribution by the algorithm) it is expedient to introduce reasonable limitations on the intervals of the parameters, their relations, etc., in order to excluse some extra calculations non-useful in the problems.

The simple examples we considered here show that the logical mathematical models similar to those we considered can contain actually unbounded number (with a limit provided by the computer techniques only) of both stochastic influences (the parameters $q_i$) and the conditions of their appearance (the logical and other relations between the coefficients $q_i$ and also the parameters of the system). In the same way, very complicate complex influences of very different stochastic factors affecting any processes we study (not only physical processes) can be modelled if their formalization is possible. Moreover, it is probably we can set up the probabilistic system or process to be into a small number of stable states, which are necessary to our needs in the problem, by respective choice of the parameters affecting it.

Concerning the Brownian motion as a particular case of normal distributions, it can be also analysed if we know the spectrum of the factors perturbing it (the dole of each factor in their common sum, and the goal of each factor into the commonly perturbing influence). Concentration of the Brownian molecules and their momentum can be such factors in the problem.

Generalizing all that has been presented in this paper, I would like to say that frequent distributions provide a possibility for bonding the reaction of different parameters of a complicate system being affected by stochastic factors of the surrounding world, and also finding most probable states of the system thus predicting its behaviour. Having any problem, both those of physics, industry, economics, game, and others where numerous parameters are unknown, non-sufficiently determined, or are affected by stochastic changes, the method that presented in the paper leads to a spectrum of the most probable solutions of the problem.

Submitted on April 24, 2010 / Accepted on May 17, 2010

## References


1. Shnoll S. E. Cosmic physical factors in random processes. Svenska fysikarkivet, Stockholm, 2009, 388 pages.
2. Dombrowski K. I. *Bulletin of Soviet Atron. Geodesical Society*, 1956, no. 17(24), 46–50.
3. Dombrowski K. I. Rational numbers distribution and resonance. *Progress in Physics*, 2005, v.1, 65–67.