CDQM, An Int. J., Volume 17, Number 3, September 2014, pp. 26-34

COMMUNICATIONS
IN
DEPENDABILITY AND
QUALITY
MANAGEMENT
An International Journal

UDC 629.7.051.8:519.711
COBISS.SR-ID 213386508

The Use of Normal Probability Law In Problems of Affirmation of Safety Requirements on An Example of Automatic Landing Systems of Airplanes

***Lidiya N. Aleksandrovskaya*[1], *Anna E. Ardalionova*[2] *and* *Andrey V. Kirilin*[1]**

[1] Russian State Technological University Named After K. E. Tsiolkovsky - MATI, Faculty Flight Vehicles Tests, Bernikovskaja emb., 14-2, Moscow, Russia
E-mail: kirillinav@mati.ru
[2] Joint stock company Moscow Institute of Electromechanics and Automation, Aviatsionniy lane, 5, Moscow, Russia
E-Mail: ila-mati@yandex.ru

accepted August 20, 2014

Summary
The normal probability law of random quantities takes a central place both in classical mathematical statistics, and in practical applications. Completeness of the theoretical research relating the normal law, and also its rather simple mathematical properties make it the most attractive and convenient in application. Even in case of a diversion of explored experimental data from its normal law often it is possible to use as the first count stage; thus quite often it appears, that from the point of view of specific goals similar approach gives satisfactory results. In the article the examples of such use of the normal law gained on the basis of processing major data file of a statistical modelling of automatic landing of airplanes, conducted at Moscow institute of electromecanics and automatics with the purpose of affirming of demands in safety of landing to airworthiness standards are given.

Key words: *Normal, binomial, empirical probability law, approximation, a confidence interval, ordinal statistics.*

1. INTRODUCTION

At realization of a statistical modelling of automatic landing of airplanes empirical probability laws of such characteristics determining safety of landing as distance of a tangency (distance of tangency point an airplane of a flight strip (runway) from its beginning), a vertical velocity of touchdown, a lateral deviation from a runway center line, bank angles and a pitch were under construction. The example of a distribution law of distance of a tangency is given in Figure 1.

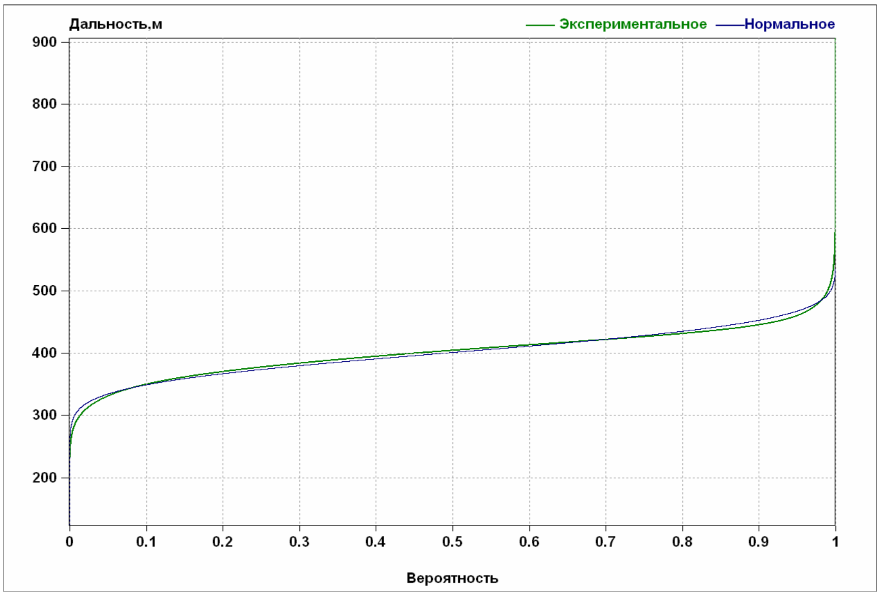

Figure 1. The distribution law of distance of a tangency of an airplane AN-148

As well-known, the empirical probability law of a random quantity x is determined by expressions [1]:

$$F(x)=\begin{cases}0, x \le x_{(1)} \\ \dfrac{i}{n}, x_{(i)} \le x \le x_{(i+1)}, 1 \le i \le n-1 \\ 1, x > x_{(n)}\end{cases}$$

Where:

$x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$ - variational series,

$x_{(i)}$ - the ordinal statistics,

n - a sample size.

In such form "tails" of allocation are cut off, т. е. Between $x_{(1)}$ the share of allocation, equal to one also $x_{(n)}$ is concluded, whence follows, that at any sample size the quantile of allocation matching as much as high probability can be discovered. Even intuitively it is clear, that precision of such estimation at a small sample size is small.

However in problems of the analysis of demands to safety us tails of allocations, т interest. To. Thus maximum permissible risks constitute $10^{-6} \div 10^{-8}$.

Therefore at build-up of an empirical distribution law other form at which tails are not cut off [2] is used, namely the empirical distribution law of a random quantity x is determined by assemblage $(x_{(i)}, p_i)$ where the ordinal statistics $x_{(i)}$ is a unbiased estimator of a quantile x_{p_i} at $p_i = \dfrac{i}{n+1}$. We shall note, that the interval of discretization $[x_{(i+1)} - x_{(i)}]$ is a random quantity, and the share of allocation $\Delta p_i = p_{i+1} - p_i$ concluded in each interval, is constant and equal $\Delta p_i = \frac{1}{n+1}$. Except for that: $p_1 = \frac{1}{n+1}$; $p_n = \frac{n}{n+1}$; $x_{(0)} = -\infty$; $x_{(n+1)} = +\infty$; $p_0 = 0$; $p_{n+1} = 1$, i.e. tails of allocation are not cut off.

For greater obviousness and comparison with a normal distribution law it is necessary to select gauge on an abscissa axis. As gauge the value of inverse function of the normal distribution, matching probability $p_i = \frac{i}{n+1}$ is selected**.**

By such selection the gaussian law of probabilities will represent a straight line transiting at $p_i = 0.5$ through an ensemble average with a declination, determined by a variance (Figure.2)

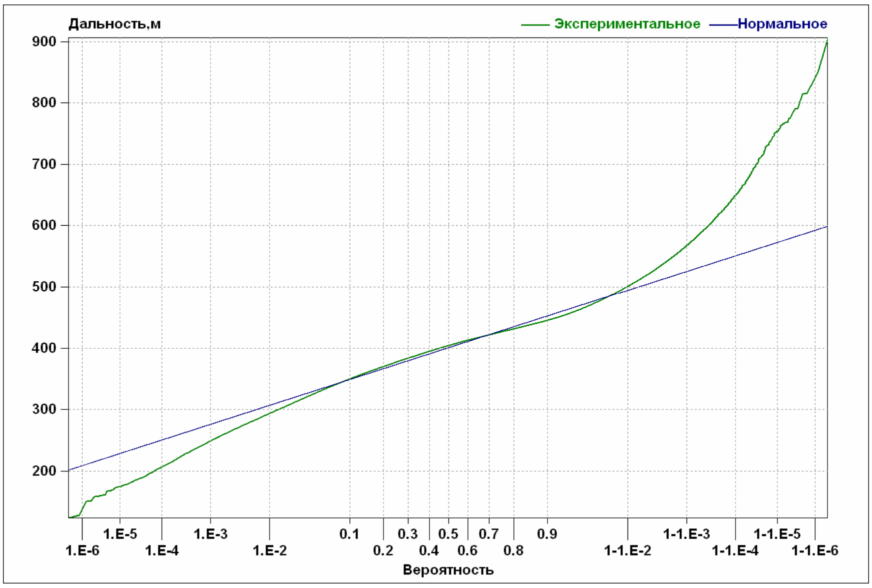

Figure 2. The distribution law of distance of a tangency of an airplane AN-148 in new coordinate system

In Figure 2 it is visible, that the mean part in the interval probabilities (0,05÷0,95) practically coincides with a normal distribution law, and tail parts considerably deviate normality.

2. PRECISION OF AN ESTIMATION OF A QUANTILE OF ALLOCATION

In problems of an estimation a compliance with requirements of safety we are interested in a quantile x_{p_i} precision estimation or at a preset value of $x_{(i)}$ precision of an estimation of matching probability p_i. The solution of this problem can be gained from the theory of ordinal statistics.

The distribution function of r-th ordinal statistics is determined as:

$F_r(x) = \Pr\{x_{(r)} \le x\}$ = *Probability* {at least r of $x_{(i)}$ is less or equated x} = $\sum_{i=r}^{n} C_n^i P^i(x)[1-P(x)]^{n-i}$,

i.e. featured by a binomial distribution, Where:

$C_n^i = \dfrac{n!}{i!(n-i)!}$; $P(x)$ - the true distribution law of a random quantity x.

The binomial totals are related to an incomplete beta function a ratio:

$F_r(x) = I_{p(x)}(r, n-r+1)$,

Where an incomplete beta function:

$$I_{p(x)} = \frac{1}{B(r, n-r+1)} \int_0^p t^{r-1}(1-t)^{n-r}\,dt;$$

Beta function:

$$B(r, n-r+1) = \int_0^1 t^{r-1}(1-t)^{n-r}\,dt, \quad r > 0;\, n-r > 0.$$

Tables of an incomplete beta function are present, for example in [3], however their volume is obviously insufficient for our problems. Therefore the program has been developed, allowing to evaluate necessary values of a beta function.

On the basis of a binomial distribution confidence bounds of an estimation p_i can be determined.

As a matter of convenience the subsequent analysis we shall inlet a label *d=n-r*. At *d=0* the maximal ordinal statistics $x_{(n)}$ is considered, at *d=1* - a statistics $x_{(n-1)}$, etc.

Then $F_r(x) = F_{n-d(x)} = \sum_{i=0}^{d} C_n^i P^{n-i}(x)[1-P(x)]^i = B_i[n, P(x), d]$, and values of the lower and upper confidence bounds are determined from equations of Clopper-Pearson [3]:

$$\begin{cases} B_i(n, p_в, d-1) = \gamma_1 \\ B_i(n, p_н, d) = 1-\gamma_2 \end{cases}, \gamma_1 + \gamma_2 - 1 = \gamma \text{ - confidential probability.}$$

In practice of statistical research normal approximation of a binomial distribution is widely used, thus it is considered, that this approximation yields good results in a mean part of a probability distribution and it is essential more the poor in tail parts.

Let's carry out a research:

- fluctuation of breadth of a confidence interval with magnification of the order *d*;
- fluctuation of symmetry of confidence bounds;
- comparison of precise confidence bounds with the boundaries gained at normal approximation

Let n=1000; $\gamma_1 = \gamma_2 = 0.95$. Using tables [3], we shall gain the results tabulated in Table 1.

Table 1. Tabulated results

d	$\hat{p}$	$p_н$	$p_в$	$\Delta p = p_в - p_н$	$\Delta p_н = p_н - \hat{p}$	$\Delta p_в = p_в - \hat{p}$
1	0.998	0.995265	0.999949	0.004684	-0.0027	0.002
100	0.899	0.883008	0.915215	0.0322	-0.0161	0.0161

Thus, with magnification of the order *d*, the confidence interval is dilated, however becomes more symmetrical concerning an estimation $\hat{p}$. Asymmetry of a confidence interval is subzero.

For transition from an estimation $\hat{p}_i$ to an estimation of a quantile $\hat{x}_{(i)}$, we shall use a procedure given in [2]. We read out from statistics $x_{(n-d)}$ number of steps, at which value $p_{(n-d-\Delta_H)} = p_н$ (or) $p_{(n-d-\Delta_в)} = p_в$. At normal approximation this number of steps is calculated by formula $\Delta = U_{\frac{1+\gamma}{2}} \cdot \sqrt{n \cdot \hat{p}_i \cdot (1-\hat{p}_i)}$ and the - quantile of a standard normal distribution is rounded off up to an integer $U_{\frac{1+\gamma}{2}}$. Matching values $x_{(n-d\pm\Delta)}$ also will be confidence bounds for a quantile $x_{(n-d)}$. Results of calculations for a viewed Example are tabulated in Table 2.

Table 2. Results of calculations for a viewed example

d	Δ_H binominal	$\Delta_H = \Delta_B = \Delta$ normal	Δ_B binominal
1	3	±1,64·1,41 = ±2,325 ≈ ±3	
100	16	1,64·9,52 = 15,6 ≈ 16	16

Thus, in view of a necessary rounding off the precise and approximate boundaries coincide.

Let's note, that at *d=1* to read out three steps aside magnifications of probability it is not possible, since the following ordinal statistics at d=0 loses only on one step.

In this and similar cases we shall use an assumption given in [2] about normal distribution law of quantiles $x_{(n-d)}$ provided that the matching estimation $p_{(n-d)}$ it is not equal to null or one, that is always executed at the received expedient of build-up of an empirical distribution law ($p_1 = \frac{1}{n+1} \neq 0$; $p_{(n)} = \frac{n}{n+1} \neq 1$). For a symmetrical normal distribution law we shall discover a difference $\Delta_H = x_{(n-d)} - x_{(n-d-\Delta)}$ and we shall add it to an estimation of a quantile $x_{(n-d)}$.

Example 1. For affirming demands to the maximal admissible distance of a tangency of an AN-148 airplane - 832 meters with probability 0,999999 had been conducted a statistical modelling by a volume n=1000000; thus the maximal value of distance has constituted $x_{(n)}$ =814.9м, and probabilities - $p_{(n)} = \frac{1000000}{1000001} = 0.999999$.

Precise lower confidence bound $P_н$ for probability $P_{(n)}$ is determined by formula $P_н^n = 1-\gamma$ and constitutes at $\gamma = 0.9$ $p_{н=0.999997}$. To this value there matches value of probability $P_{(n-2)} = \frac{999998}{1000001}$ and value of a quantile $x_{(n-2)}$ =802,39.

At normal approximation we shall gain $\Delta_H = 1.64\sqrt{1000000 * 0.999999(1-0.999999)} \approx 1.64$. Rounding off up to Δ_H =2, we gain the same values, as at precise calculation.

The upper confidence bound of a quantile $x_{(n)}$ will constitute (814,9-802,39) +814,9=827,41 m.

The value of a quantile $x_{(n+2)}$ at magnification of a sample size up to *n=1300000* has constituted 828.12 m, i.e. a lapse of the prognosis less 1 m.

Thus, demands to safety of landing on distance are confirmed.

The assumption about normality of allocation of the maximal value of sample makes experts doubt and requires additional research.

3. EXAMINATION OF NORMALITY OF ALLOCATION

Examination of normality of a probability distribution of the maximal ordinal statistics $x_{(n)}$ at $n = 10^6$ has been combined with a problem analytical the exposition of not observable tails of allocations on an example of the analysis of the maximal distance of the tangency, a bottleneck of automatic landing of an airplane AN-148 being most.

Considering, that at improvement of control laws it is necessary to carry out about ten adjustments, the method of a rapid analysis is necessary. The method should allow not to conduct a total storage of a statistical modelling to make the solution on overgrowth of a volume of model operation with the purpose of affirming of demands to safety, or on necessity of adjustments of a control law. Such mathematical method is offered in [4]: for the analytical exposition of the right tail of allocation of distance of a tangency it is offered to use the allocation of Pareto featuring a probability distribution of a random quantity, greater some fixed value C_0

$$F_{Pareto}(x) = 1 - \left(\frac{C_0}{x}\right)^{\alpha}, \text{ At } x \geq C_0. \qquad (1)$$

Allocation of Pareto more simple, does not require integration, and its analytic form allows to yield an estimation of precision of calculated values.

For example, α is determined by a method of moments $\hat{\alpha} = 1 + \sqrt{1 + \left(\frac{1}{\hat{V}}\right)^2}$,

Where:

$$\hat{V} = \frac{S}{\overline{x}}; \overline{x} = \frac{1}{n_{yc}} \sum_{i=1}^{n_{yc}} x_i; S^2 = \frac{1}{n_{yc} - 1} \sum_{i=1}^{n_{yc}} \left(x_i - \overline{x}\right)^2.$$

The degree of truncation of initial allocation is equal $F_{truncated} = \frac{n_{sample} - n_{truncated}}{n_{sample} + 1}$, where $n_{truncated}$ - number of measurings in an explored tail part.

The coordination of allocation of Pareto with the truncated initial allocation was yielded by formula:

$$F = F_{Pareto}\left(1 - F_{truncated}\right) + F_{truncated}.$$

Predicted on 10^6 value of distance $x_{пр}$ was determined on allocation of Pareto $F_{Pareto} = \frac{0{,}999999 - F_{truncated}}{1 - F_{truncated}}$, whence from (1) it is calculated x_{np}.

Parameter of truncation C_0 and therefore n_{yc} were selected from a requirement of the maximal coincidence experimental $x_э$ and calculated values x_p, determined of (1) at $F_{Pareto}(x) = \frac{n_{truncated}}{n_{truncated} + 1}$.

Example 2. Approximation of a tail part of allocation of distance of a tangency by allocation of Pareto.

$n_{\text{truncated}} = 20$; $C_0 = 800$ m; $n_{\text{sample}} = 300000$; $\overline{x}$ =828,765 m; S = 22,87 m; $\overline{\alpha}$ =37,25.

$$F_{truncated} = \frac{300000 - 20}{300001} = 0{,}99993.$$

$$F_{Pareto} = \frac{20}{21} = 0{,}9525381$$

$x_р$ = 867,96 m; $x_э$ = 869,9 m; $x_{пр.}$ = 896,9 m.

Example 3. A research of probability law of final value of distance of a tangency. Samples of a building up volume $3*350\cdot10^3$ $3*600\cdot10^3$, $3\cdot10^6$ for some intermediate version of a control system were explored.

On 9 values the calculated mean value and mean-square deviation, equal $\overline{x}_{пр} = 877{,}32$ m were found; S=17,1 m.

The mean calculated value was compared to the mean experimental value gained on three samples 10^6.

$\overline{x}_{экс} = 875{,}3м < \overline{x}_{пр}$

The difference between $\overline{x}_{пр}$ and $\overline{x}_{экс}$ was 2 m, that has shown a capability and expediency of exposition of a tangency distance allocation tail part by an allocation of Pareto.

Further examination of normality of predictable values according to GOST R ISO 5479-2002 on the basis of Shapiro-Wilke criterion was carried out. The criterion allowed to detect a diversion from normality even at small sample sizes.

According to this criterion statistics were calculated:

$$S^2 = \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 = 2956{,}78$$

$$b = a_n[x_{(n)} - x_{(1)}] + a_{n-1}[x_{(n-1)} - x_{(2)}] + a_{n-2}[x_{(n-2)} - x_{(3)}] + a_{n-3}[x_{(n-3)} - x_{(4)}] = 52.9,$$

Where coefficients a_n, a_{n-1}, a_{n-2}, a_{n-3} were selected from the table given in the indicated standard and equated: a_n=0,6058; a_{n-1}=0,3164; a_{n-2}=0,1743; a_{n-3}=0,0561. The statistics of criterion $W = \frac{b^2}{S^2} = \frac{2734.287}{2956.78} = 0.925$ was evaluated and compared to the critical values also given in the standard at various values n and a confidence level α.

At $W<W_\alpha$ probability of the fact that a sample is taken from the set meted under the normal law, constitutes α. Routinely the α confidence level is selected α=0,05÷0,1. In a viewed case value W exceeds Tabular value $W_{0,1}$=0.859, however it is a little bit less $W_{0,5}$=0.935. Thus, with enough high probability ≈0,5 sample of predictable values can be featured by normal probability law.

In conclusion we shall calculate the maximal value of distance of the tangency, compatible to the received statistical model.

According to Thompson's [5] criterion, measuring x_i is considered belonging the sample extracted from set with a normal distribution of probabilities if the inequality $\frac{|x_i - \overline{x}|}{S} \le W_\alpha\left(n-2\right)$ is executed, where $S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}$.

By selection α=0,1 $W_{0,1}(7)$=1.647 and the greatest expected value x_i constitutes 877,32+1,647*17,09=905,467 m, i.e. practically coincides with the maximal predicted value 905,514 m, that also confirms adequacy of the offered analytical exposition.

In Conclusion let's note, that the viewed intermediate version of control does not ensure a desired value of distance of a tangency, and was used only for the analysis of a capability of the analytical exposition of a tail of allocation of distance, thus parameters of an initial full distribution law have constituted: $\bar{x}$ = 433,2m; S=57.2m.

The conducted adjustments of a control law have allowed to reduce both the mathematical exposition, and a root-mean-square diversion up to values $\bar{x}$ = 400,7; S=40,3m, that has ensured affirming demands to distance of a tangency (an Example 1).

Besides allocation of Pareto the capability of the exposition of tails of allocations of probability was explored by the normal and truncated normal distribution.

4. USE OF THE NORMAL AND TRUNCATED NORMAL DISTRIBUTION LAW

It is known, that at truncation at the left the ensemble average and a variance of the initial normal and truncated allocations are interlinked by the ratio:

$$m_{yc} = m + \sigma \frac{Z}{F};$$

$$\sigma_{yc}^2 = \sigma^2 \left[1 - \frac{Z^2}{F^2} - \frac{m - C_0}{\sigma} \left(\frac{Z}{F} \right) \right], \qquad (2)$$

Where $Z = \varphi\left(\frac{C_0 - m}{\sigma}\right)$; $F = 1 - \Phi\left(\frac{C_0 - m}{\sigma}\right)$ - are the values of density and function of a normal distribution of probabilities [1]. As we have values of the empirical distribution law, matching calculations are simple enough.

Example 4. Approximation of a tail part of allocation of distance of a tangency truncated normal and normal allocations.

$n_{truncated}$ =20; C_0 = x $_{(1)}$ = 801,811 m; $\bar{x}$=828,765 m; S = 22,87 m;

$$\Phi = \frac{1}{21} = 0,0476;$$

$$1 - \Phi = 0,9524;$$

$$\frac{C_0 - m}{\sigma} = -1,669; \quad \varphi = 0,099;$$

$$\sigma = 25,33m; \; m = 826,13m.$$

The calculated value x_p = m + 1,669σ = 868,4 m, that practically coincides both with experimental value (867,98 m), and with a calculated value at approximation by allocation of Pareto (869,9 m).

Disregarding truncation of a normal distribution we shall gain a calculated value:

x_p = 828,765 + 1,669·22,87= 866,93 m, i.e. also we have good coincidence experimental value.

And, in conclusion, we shall calculate predicted value for $U_{0.999999}$ with allowance for and disregarding truncation.

$x_{пр\ ус}$ = 826,13 + 2,189·25,33= 881,58 m.
$x_{пр}$ = 828,765 + 2,189·22,87= 878,83 m.

More stringently at determination of computational and predictable values at a small sample size it is necessary to use not quantiles of a normal distribution, and quantile of a Student's distribution.

Then the predicted value disregarding truncation is determined, as:

$x_{пр}$ = 828,765 + 2,4·22,87= 883,6 m that is a little bit better, than at approximation by allocation of Pareto.

However, as approximation by allocation of Pareto gives "overestimate", and approximation by a normal distribution - "underestimation", use of allocation of Pareto is more preferable, since ensures some warranty of the gained results.

5. CONCLUSION

Thus, the assumption about normality of some distribution laws though is the confidant in the theoretical plot, however allows to solve simply enough a variety of practical problems, namely to create an effective method of a rapid analysis of demands to safety during improvement of automatic landing systems of airplanes by results of a statistical modelling of the limited volume and to yield a final estimation of conformity of these demands to airworthiness standards.

For additional demonstrating the achieved results in the end of an optimization it was possible to apply the stringent method regulated in Unified West-European airworthiness standards "fits - does not fit", however the necessary sample size for this purpose should three times exceed the volume recommended in the given research.

ACKNOWLEDGEMENT

The article is prepared at financial support of the Ministry of Education and Science of Russian Federation within the limits of the high-schools state research (theme №1636/14).

REFERENCES

[1] L. Bolshev, N. Smirnov: Tables of mathematical statistics, Nauka, Moscow, 1983.

[2] G. David: Ordinal statistics, Nauka, Moscow, 1979. 336 p.

[3] Statistical problems of an optimization of systems and tables for numerical calculations of parameters of reliability, Edited by R. S. Sudakov, Vyshaya Shkola, Moscow, 1975.

[4] L. N. Aleksandrovskaja, V. N. Masur, S. V. Hlgatjan, A. E. Ardalionova: A method of patch approximation of cumulative distribution functions in problems of estimation of correspondence of requests to airplanes' automatic landing systems safety to airworthiness standards // The Technological log-book. Works of FGUP "NPCAP". Systems and instruments of control. - FGUP "NPCAP of Academician N. A. Piljugin", Moscow, 2011. # 3., pp. 41-48.

[5] P. Muller, P. Neuman, R. Shtorm: Tables of mathematical statistics, Finance and statistics, Moscow, 1982.