Buychik, A., & Komissarov, P. V. (2023). Hypernormal distribution theory: Analysis of the set of extreme random variables models. *Actual Issues of Modern Science. European Scientific e-Journal, 24*, 56-81. Ostrava: Tuculart Edition & European Institute for Innovation Development.
Буйчик, А., Комиссаров, П. В. (2023). Теория гипернормального распределения: анализ множества моделей экстремальных случайных величин. *Actual Issues of Modern Science. European Scientific e-Journal*, *24*, 56-81. Ostrava: Tuculart Edition & European Institute for Innovation Development. (на англ.)

DOI: 10.47451/inn2023-03-03

The paper will be published in Crossref, ICI Copernicus, BASE, Zenodo, OpenAIRE, LORY, Academic Resource Index ResearchBib, J-Gate, ISI International Scientific Indexing, ADL, JournalsPedia, Mendeley, eLibrary, and WebArchive databases.

Alexander Buychik, Doctor of Science in Economics, PhD of Social and Political Sciences, Director for Sciences, Tuculart Holding. Ostrava, Czech Republic. ORCID: 0000-0002-2542-4198
Peter V. Komissarov, Graduate of Postgraduate Studies, Admiral Makarov State University of Maritime and Inland Shipping. St. Petersburg, Russia.

Hypernormal distribution theory: Analysis of the set of extreme random variables models

Abstract: The analysis of the set of extreme random variables models is still an extremely topical topic in many areas of mathematical research in the theory and practice of managing production processes due to its specificity and great interest in finding an expectation and stability indicators set studied in practical economics. Calculations of applied mathematics help to determine tentatively possible boundary parameters of various models, i.e., expectations, despite the fact that theoretical calculations do not have a direct association with practical data. Nevertheless, the consideration of extreme models of extreme random variables is still relevant in many areas of science and industry. The study subject was the hypernormal distribution theory. The study object was a set of extreme random variables models. The study purpose was a comprehensive analysis of many models of extreme random variables. To achieve the purpose and solve the tasks formulated on its basis, empirical, analytical and comparative methods of data analysis and the method of mathematical modelling, which contributed to the study of the materials presented in this article, were used. In the study course, materials from the works of such leading world experts in extreme value theory and programming K. Beck, M. Fowler, L. Tippett, E. Gumbel, K. Auer, R. Miller, and Scott W. Ambler and researchers as V.L. Khatskevich, B.V. Gnedenko, V.A. Akimov, V.A. Bykov, E.Yu. Shchetinin, K.M. Nazarenko, L.P. Kvashko, A.S. Losev, V.S. Mikhailov, V.A. Popov, E.R. Smolyakov. In the study course, the definition of an extreme value within the framework of the theory was refined, the typology of the distribution of maximum values was analysed, seven theories of the hypernormal distribution were identified and their proofs were presented, and practical examples of the application of each theory were given. The practical significance of the study of extreme random variables models in various areas of industrial human activity was confirmed. The materials of the study can be used in the widest range: from application in risk management of industrial production to predicting the probabilities of natural phenomena, which makes it possible to prevent significant economic and social losses of society, as well as make a tangible contribution to programming the probabilities of the development of the society of the future.

Keywords: extreme random variables, excess consumption, extreme values, entropy.

Александр Буйчик, доктор экономических наук, PhD социальных и политических наук, директор по науке, Tuculart Holding. Острава, Чехия. ORCID: 0000-0002-2542-4198

Пётр Вениаминович Комиссаров, выпускник аспирантуры, Государственный университет морского и речного флота им. адмирала С.О. Макарова. Санкт-Петербург, Россия.

Теория гипернормального распределения: анализ множества моделей экстремальных случайных величин

Аннотация: Анализ множества моделей экстремальных случайных величин до сих пор является крайне актуальной темой во многих областях математических исследований теории и практики управления производственными процессами в силу своей специфичности и большого интереса к поиску множества показателей ожидаемости и стабильности, исследуемые в практической экономике. Расчёты прикладной математики помогают определять ориентировочно возможные пограничные параметры всевозможных моделей, т.е., ожиданий, несмотря на то что теоретические расчёты не имеют прямой ассоциации с практическими данными. Тем не менее, рассмотрение экстремальных моделей экстремальных случайных величин до сих пор является актуальным во многих областях науки и промышленности. Предметом исследования являлась теорема гипернормального распределения. Объектом исследования являлось множество моделей экстремальных случайных величин. Целью данного исследования являлся комплексный анализ множества моделей экстремальных случайных величин. Для достижения поставленной цели и решения сформулированных на её основании задач использовались эмпирический, аналитический и сравнительный методы анализа данных и метод математического моделирования, которые способствовали исследованию материалов, представленных в данной статье. В ходе исследования были использованы материалы трудов таких отечественных исследователей как В.Л. Хацкевич, Б.В. Гнеденко, В.А. Акимов, В.А. Быков, Е.Ю. Щетинин, К.М. Назаренко, Л.П. Квашко, А.С. Лосев, В.С. Михайлов, В.А. Попов, Э.Р. Смольяков, а также материалы ведущих зарубежных специалистов в области теории экстремальных значений и программирования К. Бека, М. Фаулера, Л. Типпетта, Э. Гумбеля, К. Ауэра, Р. Миллер, и Скотта В. Эмблера. В ходе исследования было уточнено определение экстремальной величины в рамках теории, проанализирована типологию распределения максимальных величин, определены семь теорем гипернормального распределения и представлены их доказательства, а также даны практические примеры применения каждой из теорем. Тем самым, была подтверждена практическая значимость исследования вариативов экстремальных моделей экстремальных случайных величин в различных областях индустриальной деятельности человека. Материалы данного исследования могут быть использованы в самом широком спектре: от применения в области риск-менеджмента промышленного производства до предсказания вероятностей природных явлений, что позволяет предупредить значительные экономические и социальные потери общества, а также внести ощутимый вклад в программирование вероятностей развития общества будущего.

Ключевые слова: экстремальные случайные величины, гипернормальное распределение, теория экстремальных значений, энтропия.

Introduction

The analysis of the set of extreme random variables models is still an extremely topical topic in many areas of mathematical research in the theory and practice of managing production processes due to its specificity and great interest in finding an expectation and stability indicators set studied in practical economics. Calculations of applied mathematics help to determine tentatively possible boundary parameters of various models, i.e., expectations, despite the fact that theoretical calculations do not have a direct association with practical data. Nevertheless, the consideration of extreme models of extreme random variables is still relevant in many areas of science and industry.

The study subject was the hypernormal distribution theory.

The study object was a set of extreme random variables models.

The study purpose was a comprehensive analysis of many models of extreme random variables.

Based on the study purpose, the following tasks were formed:

- clarify the definition of an extreme value within the framework of the theory;
- analyse the typology of the distribution of maximum values;
- define hypernormal distribution theories and present their proofs;
- give a conclusion on the practical application of the evidence base of the hypernormal distribution theories.

To achieve the purpose and solve the tasks formulated on its basis, empirical, analytical and comparative methods of data analysis and the method of mathematical modelling, which contributed to the study of the materials presented in this article, were used.

In the study course, materials from the works of such leading world experts in extreme value theory and programming K. Beck, M. Fowler (*Beck & Fowler, 2001*), L. Tippett (*Tippett, 2013*), E. Gumbel (*Gumbel, 2012*), K. Auer, R. Miller (*Auer & Miller, 2001*), and Scott W. Ambler and researchers as V.L. Khatskevich (*Khatskevich, 2013*; *Khatskevich, 2020a*; *Khatskevich, 2020b*), B.V. Gnedenko (*Gnedenko, 1943*), V.A. Akimov, V.A. Bykov (*Akimov et al., 2009*), E. Yu. Shchetinin (*Akimov et al., 2009*; *Shchetinin & Nazarenko, 2008*), K.M. Nazarenko (*Shchetinin & Nazarenko, 2008*), L.P. Kvashko, A.S. Losev (*Kvashko & Losev, 2013*), V.S. Mikhailov (*Mikhailov, 2012*), V.A. Popov (*Popov, 2013*), E.R. Smolyakov (*Smolyakov, 2011*).

Materials and methods of research

Auxiliary information and basic definitions

The theory of extreme values is a branch of the science of statistics, which aims to study extreme deviations from the median of probability distributions, i.e., an assessment of phenomena based on an ordered selection of probability parameters for the most extreme events or processes. The concept of extreme value theory was introduced by Leonard Tippett (*Tippett, 2013*) in the first quarter of the 20th century and became the basis of many studies that have been going on for about 100 years. At that time, his research was based at the British Cotton Research Association, where he worked on strengthening the cotton thread. In his research, L. Tippett postulated that the strength of a thread is determined by the strength of its weakest fibers. He obtained three asymptotic limits that clearly described the distributions of extrema that considered independent variables (*Tippett, 2012*). It was the study that became the starting point in applying a qualitatively new approach to calculating extremeness in production and economic indicators. In the future, E.D. Gumbel codified this theory in his work *Statistics of Extremes* (*Gumbel, 2012*). There he gave the distribution concept, which now bears his name. In the second half of the 20th century, the results obtained were significantly expanded and began to consider insignificant correlations between variables. Strong correlations of the order of dispersion began to be actively studied already at the beginning of the 21st century with the use of artificial intelligence and a neural network.

With the potentially high probability of extreme manifestations, there naturally becomes an increased risk of redundant programming as a form of agile software development methodologies. The authors of this methodology are such prominent scientists today as Kent Beck, Ward Cunningham, Martin Fowler (*Beck & Fowler, 2001*) and others. Kent Beck pioneered the development of the methodology for the Chrysler Comprehensive Compensation System project (*Beck, 2003*). The goal was to apply theoretical methods and develop new and modern software for those times. As a result of the development, it was possible to raise and develop technology and programming at a new qualitative level. It should also note that it was in extreme programming that a departure from the long-term process of creating programmes was determined, which consisted in the fact that instead of one-time planning, analysis and design of a system for the calculated course of events, specialists now implement these operations in a phased complex during development.

The analysis of extreme values plays an important role in the study of many phenomena and in solving applied problems of the complex systems reliability and efficiency, structural mechanics, the theory of stability, dynamic strength, etc. Consideration of absolute extrema will begin with consideration of the maximum:

$$U = \max(X_1, X_2, \dots, X_n).$$

Values from a set of n random variables (random sequence). If all components of the sample $X_1, \dots X_n$ are independent and equally distributed random variables, then the distribution function of the largest $F_n(X)$ value is determined as follows:

$$F_n(X) = P\{U < x\} = P\{X_1 < X, X_2 < X, \dots, X_n < X\} = F^n(X)$$

where $F(X)$ is distribution function of the original random variable.

If V is the minimum value of a random variable from a set of *n* random variables:

$$V = \max(X_1, X_2, \dots, X_n).$$

And if the components of the sample $X_1, \dots, X_n$ are independent and equally distributed random variables, then the distribution function of the smallest $Q_n(X)$ value is determined similarly:

$$Q_n(X) = P\{V < x\} = 1 - P\{X_1 \geq X, X_2 \geq X, \dots, X_n \geq X\} = 1 - [1 - F^n(X)]^n.$$

Thus, extreme values distributions can be derived from the exact original distribution. In reality, the analytic properties of the original distribution are rarely known. This leads to the need to use the principle of maximum distribution and determine on this basis the extreme distribution of the extreme value (maximum or minimum). From a mathematical viewpoint, the maximum principle's application uncertainty leads to the solution of extremal (variational) problems with organic ones, determined by the form of setting the probabilistic characteristics and the range of random variable values.

According to the distribution, which has the greatest entropy under certain restrictions, is called extreme. Next, a brief description of 8 types of extreme distributions of extreme random variables with a certain degree of universality will be given. The common thing in the formation of such models is the definition of the Euler-Lagrange equations of variational problems, considering the specifics of specifying information about the initial random variable and allowing meaningful interpretation. The most obvious and, most importantly, the most practical application-oriented is the statistical interpretation of extremal distributions in terms of the theory of order statistics, the subject of which is the study of the properties and applications of ordered random variables and functions of them. To this end, we present some auxiliary information from the theory of order statistics. The source material for statistical analysis, obtained as a result of a simple random selection from the general population, determined by the random variable X, is a sample of a finite size n:

$$X_1, X_2, \dots, X_n.$$

A sequence of sample values ordered by magnitude $X_1^{(n)} \leq X_2^{(n)} \leq X_n^{(n)}$ is called a variation series.

If the initial distribution of the general population is characterised by the mathematical expectation m and the variance δ^2, the distribution of the rightmost member of the variation series is

$$P\left\{X_n^{(n)} < x\right\} = F(X).$$

It should let agree to call the distribution that delivers the entropy maximum an extremal distribution of type 1. An extremal distribution of type 2 is a limiting $n \to \infty$ distribution of type 1. If the initial distribution of the general population is characterised by only one mathematical expectation m, the distribution of the rightmost member of the variation series is

$$P\left\{X_n^{(n)} < x\right\} = G_n(X).$$

It should let agree to call the distribution that delivers the entropy maximum an extremal distribution of type 3. For the case $n \to \infty$ the distribution function $G_n(X)$ degenerates into a type 4 distribution function $G_\infty(X)$.

In a similar way, we introduce into consideration the extremal distributions of the minimum values. If the initial distribution of the general population is characterized by the mathematical expectation m and the variance δ^2, the distribution of the rightmost member of the variation series (the minimum of the random sequence):

$$P\left\{X_n^{(n)} < x\right\} = Q_n(X).$$

It should let agree to call the distribution that delivers the entropy maximum an extremal distribution of type 5.

As $n \to \infty$ it should say that an extremal distribution of type 5 degenerates into an extremal distribution of type 6. If it assumes that the original random variable is characterised only by the mathematical expectation m, the distribution of the leftmost member of the variational series $P\left\{X_n^{(n)} < x\right\} = R_n(X)$, which provides the maximum entropy, it should agree to call the extremal distribution of type 6. For the limiting case $n \to \infty$ it is useful to introduce a type 8 distribution. Thus, extreme distributions of types 1-4 are distributions of maximum values, and distributions of types 5-8 are distributions of minimum values of random sequences, extreme distributions of types 1, 2, 5, and 6 are distributions of extremes of random sequences of independent and identically distributed random variables of the general population, the distribution function of which is unknown and is characterised only by mathematical expectation and variance, extreme distributions of 3, 4, 6, and 8 types are distributions of extrema of random sequences of independent and identically distributed random variables of the general population, characterised by only one mean, extreme distributions of 3, 4, 7, and 8 types are asymptotic (limiting) distributions. In a compact form, the main notation and definitions are presented in the appendix (*Table 1*).

The principle of maximum values distribution

An extreme distribution of type 1 (hypernormal distribution) is a continuous distribution, the probability density of which is the solution of the differential equation:

$$n\sigma^2[F_n(X)]^{\frac{n-1}{n}}\ddot{F}_n(X) + (X-m)\dot{F}_n(X) = 0] \quad (1)$$

where m and σ^2 are the mathematical expectation and variance of the set of initial random variables. The nonlinear differential equation (1) satisfies the natural boundary conditions $F_n(-\infty) = 0, F_n(\infty) = 1$ (2) and is completely determined by the first two moments (m and σ^2) of the original random population and the sample size n.

The hypernormal distribution corresponds to the distribution function $F(X)$ of the original random variable, determined by solving the following differential equation with the same boundary conditions

$$n\sigma^2\ddot{F}_n(X)[F(X)]^{n-1}\sigma^2 n(n-1)\dot{F}^2(X)[F(X)]^{n-2} + (x-m)\dot{F}_n(X) = 0 \quad (3)$$

$$F(-\infty) = 0, F(\infty) = 1 \quad (4)$$

In the appendix, the figures 1-10 (*Figure 1*; *Figure 2*; *Figure 3*; *Figure 4*; *Figure 5*; *Figure 6*; *Figure 7*; *Figure 8*; *Figure 9*; *Figure 10*) show graphs of functions and numerical characteristics (expectation and variance) of the hypernormal distribution for integer parameters n from 1 to 10, obtained as a result of solving the nonlinear boundary value problem (1), (2). The calculation of the functions $F_n(X)$ is made for standard conditions (for the scale parameter $\sigma = 1$ the shift parameter $m = 0$).

For large values of the argument $(x > m + 3\sigma)$ the hypernormal distribution asymptotically approaches the normal distribution with density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma\sqrt{n}}e \quad (5)$$

With an extremal distribution of type 2 (hypernormal distribution), the random variable $X_n^{(n)}$ has a limiting (for $n > 10$) hypernormal distribution if its quantile function has the form

$$x_p = m + \sigma\sqrt{2n}\sqrt{-E_i(\ln P)} \quad (6)$$

where $E_i(\ln P)$ is integral exponential function, whose argument is the natural logarithm of the probability $p = F_\infty(x)$.

The second table 2 (*Table 2*) presents the values of the function of the limiting hypernormal distribution $F_\infty(x)$, whose argument is the value:

$$\tilde{x} = \frac{x-m}{\sigma\sqrt{n}}.$$

The mathematical expectation and variance of a random variable is determined by the formulas:

$$E\left[X_n^{(n)}\right] = m + \sigma\sqrt{2n}\int_0^1 \sqrt{-E_1(\ln P)}dp = m + 0.69676\sqrt{2n},$$

$$D = 2n\sigma^2\left\{\int_0^1\left[-E_i(\ln P)dp - \left[\int_0^1 E_i(\ln P)\right]dp\right]^2\right\} = 2n\sigma^2(2-0.6067^2) = 3.0292n\sigma^2.$$

An extremal distribution of type 3 is a continuous distribution whose probability density is a solution to the differential equation:

$$mn[G_n(x)]\ddot{G}_n(x) + G_n(x) = 0 \text{ (7)}$$

where m is mathematical expectation of the set of initial random variables.

Differential equation (7) satisfies the boundary conditions

$$G_n(0) = 0, G_n(\infty) = 1 \text{ (8)}$$

The density is determined by the parameters m and n (sample size).

An extremal distribution of type 3 corresponds to a quantile function that displays

$$p = \sigma_n(x)bx$$

$$x_p = mp_2F_1\left(l, n; n+1; \sqrt[n]{p}\right) \text{ (9)}$$

where $F_1(\lambda;\ \beta;\ \gamma; z)$ are Gaussian hypergeometric function.

Using the representation of the Gaussian hypergeometric function as a series, it is convenient to represent the calculation formula for function (9) in the form:

$$x = mp\sum_{r=0}^{\infty}\frac{n}{n+r}p^{\frac{r}{n}} \text{ (10)}$$

The mathematical expectation of a random variable $X_n^{(n)}$ is determined by the dependence:

$$E\left[X_n^{(n)}\right] = mn\sum_{r=0}^{\infty}\frac{1}{(n+r)(2n+r)} \text{ (11)}$$

The value of the sum S_n series (*Table 3*):

$$S_n = n^2\sum_{r=0}^{\infty}\frac{1}{(n+r)(2n+r)}.$$

An extreme distribution of type 4 is a continuous limiting $(n \to \infty)$ extreme distribution of type 3. The probability density of this distribution $G_\infty(x)$ is the solution of the differential equation:

$$mnG_\infty(x)\ddot{G}_\infty(x) + \dot{G}_\infty(x) = 0 \text{ (12)}$$

with boundary conditions $G_\infty(0) = 0; G_\infty(\infty) = 1.$

An extreme distribution of type 4 corresponds to a quantile function that displays:

$$P = G(x) \text{ в } x,$$

$$X = -\frac{m}{\ln\frac{n+1}{n}}E(\ln P) \text{ (13)}$$

The mathematical expectation of a random variable $X_n^{(n)}$ is determined by the dependence:

$$E\left[X_n^{(n)}\right] = \frac{m}{\ln\frac{n+1}{n}}\left[-\int_0^1 E_i(\ln P)\,dp\right] = \frac{m\ln 2}{\ln\frac{n+1}{n}} \text{ (14)}$$

Distribution of minimum values

An extremal distribution of type 5 is considered to be a continuous distribution whose probability density is a solution to the differential equation:

$$n\sigma^2[1 - Q_n(x)]^{\frac{n-1}{n}}\bar{Q}_n(x) + (x - m)Q(x) = 0 \text{ (15)}$$

where m and σ^2 are mathematical expectation and variance of the set of initial random variables.

Nonlinear differential equation (15) satisfies the natural boundary conditions:

$$Q_n(-\infty) = 0, Q_n(\infty) = 1 \text{ (16)}$$

and is completely determined by the first two moments (m and σ^2) of the initial random population and the sample size n.

The extremal distribution of type 5 corresponds to the distribution function $F(x)$ of the original random variable, determined as a result of solving differential equation (15) with boundary conditions (15).

In the appendix, the figures 11-20 (*Figure 11*; *Figure 12*; *Figure 13*; *Figure 14*; *Figure 15*; *Figure 16*; *Figure 17*; *Figure 18*; *Figure 19*; *Figure 20*) show graphs of functions and numerical characteristics (mathematical expectation and variance) of an extremal distribution of type 5 for integer parameters n from 1 to 10, obtained as a result of solving a nonlinear boundary value problem (15), (16). Calculation of the functions $Q_n(X)$ is made for standard conditions (for the scale parameter $\sigma = 1$ the shift parameter $m = 0$).

For the values of the argument $x < m - 3\sigma$, the extreme distribution of type 5 asymptotically approaches the normal distribution with density:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma\sqrt{n}} e^{-\frac{(x-m)}{2n\sigma^2}}.$$

As $n \to \infty$ the extremal distribution of type V approaches asymptotically the extremal distribution of type 6, whose quantile function has the form:

$$X_p = m - \sigma\sqrt{2n}\,\sqrt{-E(\ln(1-P))} \quad (17)$$

An extreme type 7 distribution defines a continuous distribution, or density, whose probabilities are the solution of the differential equation:

$$mn[1 - R_n(x)]^{\frac{n-1}{n}} \ddot{R}_n(x) + R_n(x) = 0 \quad (18)$$

where m is mathematical creation of a set of initial random variables. Differential equation (18) satisfies the boundary conditions and is completely determined by the first parameters m and n:

$$R(0) = 0, R_n\left(\frac{mn}{n-1}\right) = 1 \quad (19)$$

An extreme distribution of type 7 corresponds to a quantile function that displays:

$$p = R_n(x) \text{ в } x;$$

$$X_p = \frac{mn}{n-1}[1 - (1-p)]^{\frac{n-1}{n}} \quad (20)$$

The mathematical expectation of a random variable $X_n^{(n)}$ is determined by the dependence:

$$E\left[X_1^{(n)}\right] = \frac{mn}{2n-1} \quad (21)$$

As $n \to \infty$, the type 7 extremal distribution asymptotically approaches the type 8 extremal distribution.

An extremal distribution of type 8 is considered to be a uniform distribution with a distribution function:

$$R_\infty(x) = \begin{cases} xm^{-1}, x \le m \\ 1, x > m \end{cases} \quad (22)$$

And mathematical expectation:

$$E\left[X_1^{(n)}\right] = \frac{m}{2}. \quad (23)$$

Study results and discussion

This section contains a presentation of the most significant results of the study of the theory of extremal distributions of extremal random variables.

The differential equation (I) defining the hypernormal distribution function $F_n(x)$ is the Euler-Lagrange equation of the following variational problem:

$$H_r = -\int_{-\infty}^{\infty} fn(x)\ln(x)dx \rightarrow \max \quad (24)$$

$$\int_{-\infty}^{\infty} fn(x)dx = 1 \quad (25)$$

$$\int_{-\infty}^{\infty} f(x)dx = 1 \quad (26)$$

$$\int_{-\infty}^{\infty} xf(x)dx = 1 \quad (27)$$

$$\int_{-\infty}^{\infty} (x-m)^2 f(x)dx = \sigma^{\ 2} \quad (28)$$

$$F_n(x) = F^n(x). \quad (29)$$

The validity of this assertion follows from the proof of theories 1 and 2 below.

Theory 1. Let X is a random variable with density $f(x) > 0, x \in (-\infty, \infty), F_n(X)$ is the distribution function of the extreme member of the variational series constructed from a sample of a finite size n from the general population defined by the random variable X. Let, further, the first two central moments of the random variable X:

$$m = \int_{-\infty}^{\infty} sf(x)dx$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x-m)^2 f(x)dx.$$

Then the entropy maximum is reached on a distribution that satisfies the differential equation (1).

Proof. To do this, it is necessary to find the function $f(x)$ and $Fn(x)$, that ensure the maximum of functional (24) in the presence of holonomic constraint (29) and under isoperimetric conditions (25)-(28). According to the well-known theories of the calculus of variations, the Lagrange multipliers $\lambda(x), \upsilon_0, \upsilon_1, \upsilon_2$ is introduced and the Euler-Lagrange equations for the extended function is composed. So, if:

$$\frac{\partial L}{\partial f} = -\ln fn(x) - 1,$$

$$\frac{\partial L}{\partial f} = -\lambda(x)n[F(x)]^{n-1},$$

$$\frac{\partial L}{\partial f} = \lambda(x),$$

the Euler-Lagrange equation for the extended function has the form:

$$\frac{d\ \ln fn(x)}{dx} + \lambda(x) = 0,$$

$$\lambda(x)n[F(x)]^{n-1} + \upsilon_1 + 2\upsilon_2(x-m) = 0. \quad (30)$$

The last equation, taking into account (30), can be written as:

$$n\frac{d\ \ln f(x)}{dx}F^n(x) = F(x)[\upsilon_1 + 2\upsilon_2(x-m)].$$

After substituting $F_n(x) = [F_n(x)]^n$, the extremal equation in the considered variational problem has the form:

$$n[F_n(x)]^{\frac{n-1}{n}}\ddot{F}_n(x)[\upsilon_1 + 2\upsilon_2(x-m)]. \quad (31)$$

or

$$n\ddot{F}_n(x)[F_n(x)]^{n-1} + n(n-1)[F(x)]^{n-2}F^2(x) = \ddot{F}(x)[v_1 + 2v_2(x-m)]. \quad (32)$$

It should integrate equation (32) over the domain of the distribution function $F(\mathrm{x})$, applying integration by parts to the first term. Due to certain properties of the distribution function and boundary conditions, one can make sure that the integral of the left side of equation (32) will be equal to 0, and the Lagrange multipliers v_1 and v_2 will be related to the mathematical expectation by the following final relation:

$$v_1 + 2v_2 - 2v_2 m = 0.$$

It follows from here that $v_1 = 0$. Multiplying the left and right sides of equation (32) by the independent variable and integrating the resulting equation in a similar way, one can find the second final relation connecting the factor with the mathematical expectation k variance.

Indeed, since:

$$\int_{-\infty}^{\infty}\left[F^{n-1}(x)\ddot{F}(x) + (n-1)F^{n-2}(x)F^2(x)\right]dxF^{n-1}(x)F(x),$$

then integrating the left side of the new differential equation obtained as a result of multiplication by the independent variable gives the following result:

$$\int_{-\infty}^{\infty}\left[F^{n-1}(x)\ddot{F}(x) + (n-1)F^{n-2}(x)F^2(x)\right]dx = xF(x)F(x),$$

$$\int_{-\infty}^{x}\left[F^{n-1}(x)\ddot{F}(x) + (n-1)F^{n-2}(x)F^2(x)\right]dx = xF(x)F(x)\begin{cases} x = \infty \\ x = -\infty \end{cases} -$$

$$\int_{-\infty}^{\infty} F^{n-1}(x)F(x)dx = -\frac{1}{n}$$

(the first term after the disclosure of uncertainty gives 0).

Thus, $2v_2(m^2 - \sigma^2)2v_2m^2 = -1$. Hence it follows that $v_2 = \frac{1}{2\sigma^2}$.

Substitution of the Lagrange multipliers v_1 and v_2 into differential equation (31) makes it possible to verify the validity of differential equations (1) and (3). The theory has been proven.

Many problems of evaluating the efficiency of complex systems and probabilistic analysis of complex processes can be formulated in terms of the theory of order statistics and are related to the study of extreme values. As an illustration of the foregoing, two examples are given below that require such an approach.

Example 1. The average time to prepare a product for use is 100 minutes, the standard deviation is 10 minutes. Five departments simultaneously began preparing for the shipment of a batch of five products. Find the probability of preparing the entire batch of products by the time $T = 120\ min.$

Solution. Let $F(t)$ is the distribution function of the product preparation time for use (note that in the problem statement the product preparation time distribution law is not specified). Let t_1 is the time of product preparation by the first subdivision, and, и t_2 is by the second one, and so on.

Then the sequence $t_1, t_2, \ldots, t_5$ is a sample of independent identically distributed random variables. If to arrange this sequence by the value $t_1^{(5)} \le t_1^{(5)} \le \cdots \le t_5^{(5)}$, the distribution of the extreme member of the variational series $t_5^{(5)}$ $F_5(t) = F^5(t) = P\left\{t_5^{(5)} < T\right\}$ determines the probability that the random variable $t_5^{(5)}$ will be less than the number T.

According to the condition of the problem, it is required to find

$$F_5(t)P\left\{t_5^{(5)} < 120\right\}.$$

In this case, it is worth noting that such a schematization of a probabilistic experiment suggests using the hypernormal distribution function with parameters $n = 5$, to calculate the probability $P\left\{t_5^{(5)} < 120\right\}$.

Passing to normalised standard values, it is obtained:

$$x = \frac{T-m}{\sigma} = \frac{120-100}{10} = 2.0.$$

According to the graph for $F5(x)$ at $x = 2.0$, it is found:

$$F_5(2.0) = 0.75.$$

Example 2. “Model of collection on alarm”. There are eight individual means (units) of the same type (ships, aircraft, etc.) located at the initial moment at the base (airfield, etc.), which, according to the alarm signal, should arrive in a given area. The average time of moving one object to this region is $m = 3\ hrs$, the standard one is $0.5\ hr$. In how many hours it is necessary to give an alarm signal so that all eight units arrive in a given area with a probability P at least 80%.

Solution. According to the graph $F_8(x)$ для $P = 0.8$, the quantile is found $x_{0.8} = 3.1$.

Since $x = (t - m_t)\sigma^{-1}$, then $t = m_1 + \sigma_1 x_{0.8} = 3 + 0.5 * 3.1 = 4.55$.

It is necessary to dwell on one case of asymptotic behavior of the hypernormal distribution. Let $n \to \infty$ (practically for $x > 3$), a natural consequence of this condition is $F_n(x) \to 1$. Then the differential equation (1) can be represented for standard conditions ($m = 0, \sigma = 1$) in the following form:

$$n\ddot{F}_n(x) + x\dot{F}_n(x) = 0. \quad (33)$$

Separating the variables, it is found:

$$\frac{d \ln fn(x)}{dx} = -\frac{x}{n} \quad (34)$$

где $fn(x) = \frac{dFn(x)}{dx}$ – плотность гипернормального распределения.

Integration of equation (34) makes it possible to verify the validity of the following result, presented in the form of the following theory.

Theory 2. For large values of the argument, the hypernormal distribution asymptotically tends to the normal distribution with density (5).

The result obtained can be somewhat strengthened by considering, instead of the hypernormal distribution function, the normal distribution function that satisfies the differential equation (1), with a variance depending on the value of the argument:

$$fn(x) = \frac{1}{\sqrt{2\pi D(x)}} e^{-\frac{x^2}{2D(x)}}.$$

Using equation (1), it can be shown that the nature of the change in the variance $D(x)$ is determined by the following differential equation

$$(x^2 - D)\frac{dD}{dx} - 2xD + \frac{2xD^2}{n[\phi(x)]^{\frac{n-1}{n}}} = 0. \quad (35)$$

where $D(x)$ is Laplace function.

It follows from differential equation (35) that for

$$x \to \infty, D(x) \to \sigma^2 n.$$

We set ourselves the goal of determining the function of the limiting hypernormal distribution (extreme distribution of type 2) in the following form:

$$F_\infty(x) = P\{\sqrt{n}X < x\}. \quad (36)$$

For sufficiently large values of n. For standard conditions $(m = 0, \sigma = 1)$, the differentiated equation (1) can be represented (при $n \to 0$) in the following form:

$$nF_\infty(x)\ddot{F}_\infty(x) + x\dot{F}_\infty(x) = 0. \quad (37)$$

It is possible to check that the change of the independent variable $y = \frac{x^2}{n}$ allows to transform equation (37) to the following form:

$$2F_\infty(y)\frac{d^2F_\infty}{dy^2} + \frac{dF_\infty}{dy} = 0. \quad (38)$$

Separating the variables and integrating, it is found:

$$\frac{dF_\infty}{dy} = -\frac{1}{2}\ln F_\infty. \quad (39)$$

From this it follows:

$$y = -1\int_0^P \frac{dF_\infty}{\ln F_\infty} = -2E_i(\ln P) \quad (40)$$

where l_iP and $E_i(lnP)$ is integral logarithm and integral function respectively.

Making the reverse transition from y to x, it can be obtained the following result, presented in a compact form in the form of the following theory.

Theory 3. The quantile function of the hypernormal distribution asymptotically $(n \to 0)$ approaches the function

$$x_p = \sqrt{2n}\sqrt{-E_i(\ln P)}.$$

The proof of the theory follows from the above reasoning and confirms the validity of relation (6).

As an example, illustrating the applicability of an extreme type 2 distribution, consider a test planning problem.

Example 3. Tests for the failure-free operation of a product should be performed for no more than 100 days. It is assumed that by the end of the tests at least 90% of the ordered products should fail. A preliminary assessment of the reliability indicators showed that the operating time is average, but the failure T_c is 50 days, and its standard deviation σ_1 is equal to 5 days. How many items do you need to order and test?

Decision. If the law of the time-to-failure distribution is not known, and the ratio between the mathematical expectation and the standard expectation is such that there are grounds to consider the range of permissible values practically unlimited and the estimated number of ordered products $n > 10$, then using the function of the limiting hypernormal distribution, according to Table 2, we define the argument (*Table 2*):

$$F_\infty(\tilde{x}) = 0.9 \ (\tilde{x} \approx 1.9).$$

Then:

$$\frac{T_g - T_0}{\sigma_r\sqrt{n}} = 1.9$$

where $Tg = 100\ days$ (directive test time).

Therefore:

$$n = \left(\frac{T_g - T_0}{\sigma_r 1.9}\right)^2 = \left(\frac{100-50}{5*1.9}\right)^2 \approx 25\ items.$$

When solving practical problems, one can hardly hope for the completeness of information about the initial random variable, which makes it possible to estimate higher moments of its distribution. If the available information about the initial random variable allows us to give only an estimate of its mathematical expectation, then in the conditions of problem (24)-(29) the condition (28) is excluded, as will be shown below, the entropy maximum is achieved on a distribution that satisfies the differential equation (7). By inverting the distribution function $G_n(x)$, one can obtain the following result, presented as a theory.

Theory 4. Let X be a random variable with density $g(x) > 0, x \in (0, \infty), G(x)$ is the distribution function of the maximum value from the set of n random variables from the general population defined by the random variable X. Let, further, the mathematical expectation of the random variable X:

$$m = \int_0^\infty x g(x) dx.$$

Then the entropy maximum is reached on a distribution that satisfies the differential equation (7) with the quantile function (9).

The proof will be as follows. Differential equation (7) is the Euler-Lagrange equation:

$$n[G_n(x)]^{\frac{n-1}{n}} \ddot{G}_n(x) - \nu \dot{G}_n(x) = 0 \quad (41)$$

$$G_n(0) = 0, G_n(\infty)$$

where ν is undefined multiplier.

The next variational problem:

$$H_\varepsilon = -\int_0^\infty \dot{G}_n(x) \ln \dot{G}_n(x) dx \to max,$$

$$\int_0^\infty \dot{G}_n(x) dx = 1,$$

$$\int_0^\infty g(x) dx = 1,$$

$$\int_0^\infty x g(x) dx = m,$$

$$G_n(x) = \left[\int_0^\infty g(x) dx\right]^n. \quad (42)$$

Formal integration of equation (41) for $n = 1$ gives an exponential distribution law, and for $n \neq 1$ leads to a dependence of the form

$$x = \int_0^{G_n} \frac{dG_n}{\nu G_n^{\frac{1}{n}} + G_n(0)} \quad (43)$$

Using the substitution $z = G_n^{\frac{1}{n}}$ (43), преобразуем к табличному:

$$\int_0^u x^{m-1} (1 + \beta x)^{-\nu} dx = \frac{u^\mu}{\mu} \; {}_2F_1(\nu, \mu; 1 + \mu; -\beta u)$$

where ${}_2F_1(\alpha, \rho; \gamma; s)$ is hypergeometric Gaussian function. Therefore:

$$\dot{G}_n(0) x = Z^n \; {}_2F_1(1, n; n + 1; -\beta z) \quad (44)$$

or

$$\dot{G}_n(0) x = p_2 F_1\left(1, n; n + 1; \beta \sqrt[n]{p}\right),$$

where $\beta = \frac{\nu}{\dot{G}_n(0)}$ и $p = G_n(x) = P\left\{x_n^{(n)} < x\right\}$.

The normalisation condition for the distribution function $G_n(x)$ implies the following relation:

$$\beta = \frac{\nu}{\dot{G}_n(0)} = -1.$$

Indeed, the quantile functions determined from relation (44) have the form:

$$x = \frac{Z^n}{\dot{G}_n(0)} \ {}_2F_1(1, n; n+1; -\beta Z).$$

For the original random variable:

$$x_p = \frac{P}{\dot{G}_n} \ {}_2F_1\left(1, n; n+1; -\beta \sqrt[n]{p}\right).$$

For the maximum value from a set of n random variables. For $x_p \to \infty, z = \sqrt[n]{\dot{G}_n(x)} \to 1$ and $p = G_n(x) \to 1$ and vice versa.

The Gauss hypergeometric function can be represented as the following series:

$${}_2F_1(\alpha, \rho; \gamma; s) = \sum_{r=0}^{\infty} \frac{d_r \beta_r s^r}{\gamma_{(r)}} \frac{s^r}{r}$$

where $\lambda_{(r)} = \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)}$.

For $p = 1 (z = 1)$, series

$$\sum_{r=0}^{\infty} \frac{1_{(r)} n_{(r)}}{(1+n)_r} \frac{(-\beta)^r}{r!} \quad (45)$$

should diverge.

According to the d'Alembert test, series (45) will diverge if the relation follows:

$$\frac{1_{(r+1)} n_{(r+1)} (-\beta)^{(r+1)} (n+1)_{(r)} r!}{(1+n)_{(r+1)} 1_{(r)} n_{(r)} (-\beta)^r (r+1)!} \geq 1.$$

Or after transformation:

$$-\beta = \frac{n+r+2}{n+2} \geq 1.$$

Therefore, in order for the series to diverge (the normalisation condition is satisfied), it is necessary that for the sufficiently large $r (r \to \infty)$ complex:

$$\beta = \frac{\nu}{\dot{G}_n(0)} = -1.$$

To determine the density value at the initial point $\dot{G}_n(0)$, one integrates relation (44) taking into account the obtained value for ß:

$$\int_0^1 \dot{G}_n(0) x dz = \int_0^1 z^n \ {}_2F_1(1, n; n+1; z) dz.$$

The left side of the equation in accordance with the definition $z = [G_n(x)]^{\frac{1}{n}}$ is the product of the value of the distribution density at zero $\dot{G}_n(0)$ and the mathematical expectation of the original random variable. It can be shown that the right side of the equation is equal to one. Indeed, by representing the hypergeometric function as a series and changing the order of summation and integration, it is defined:

$$\int_0^1 z^n \sum_{r=0}^{\infty} \frac{1_{(r)} n_{(r)}}{(1+n)_r} \frac{(z)^r}{r!} dz = \sum_{r=0}^{\infty} \frac{1_{(r)} n_{(r)}}{(1+n)_r} \frac{1}{r!} dz = \frac{n}{(n+r)(n+r+1)}.$$

Series $\sum_{r=0}^{\infty} \frac{n}{(n+r)(n+r+1)}$ can be represented as the difference between two series:

$$\sum_{r=0}^{\infty} \frac{n}{(n+r)(n+r+1)} = n \sum_{r=0}^{\infty} \frac{1}{(n+r)} - n \sum_{r=0}^{\infty} \frac{1}{(n+r+1)},$$

which after transformation can be represented as follows:

$$n\sum_{r=0}^{\infty}\frac{1}{(n+r)} - n\sum_{r=0}^{\infty}\frac{1}{(n+r+1)} = n\sum_{m=n}^{\infty}\frac{1}{m} - n\sum_{m=n+1}^{\infty}\frac{1}{m} = n\frac{1}{n} + \sum_{m=n}^{\infty}\frac{1}{m} - n\sum_{m=n+1}^{\infty}\frac{1}{m} = 1.$$

Thus, $\dot{G}_n(0)m = 1$. Therefore, $\dot{G}_n(0) = m^{-1}$ and the Lagrange multiplier $\nu = \beta\dot{G}_n(0) = -m^{-1}$.

Substituting the value of the Lagrange multiplier into the differential equation (41) makes it possible to verify the validity of the differential equation (5) and the quantile function (9). The theory has been proven.

It is also necessary to clarify the method for determining the mathematical expectation of the largest value for the case under consideration. In a way analogous to that which was applied in the derivation of the formula for determining the value of the density at zero $\dot{G}_n(0)$, one can show that:

$$E\left[X_n^{(n)}\right] = m\sum_{r=0}^{\infty}\frac{1_{(r)}n_{(r)}}{(n+r)(2n+2)}.$$

To sum up, it suffices to integrate the quantile function (9):

$$E\left[X_n^{(n)}\right] = \int_0^1 x_p dp = m\int_0^1 p\,{}_2F_1(1, n; n+1; \sqrt[n]{p})dp.$$

Representing the hypergeometric function as a series and rearranging the operations of summation and integration, one can find:

$$E\left[X_n^{(n)}\right] = m\sum_{r=0}^{\infty}\frac{1_{(r)}n_{(r)}}{(n+r)r!}\int_0^1 p^{1+\frac{r}{n}}dp = m\sum_{r=0}^{\infty}\frac{n^2}{(n+r)(2n+2)}.$$

In table 3 shows the values of the sum S_n, which make it possible to estimate the mathematical expectation of the maximum value from samples of n to 10 (for $n > 10$, one can use, as will be shown below, the asymptotic properties of the obtained extreme distributions of extreme values) (*Table 3*).

If we set ourselves the goal of determining the function of the limiting extremal distribution of type 4, then, by performing constructions similar to the constructions used in the proof of Theory 3, we can obtain the result presented in the form of the following theory, given by virtue of obviousness without proof.

Theory 5. The quantile function of type 3 extremal distribution asymptotically $(n \to \infty)$ approaches the function:

$$x = \frac{m}{\ln\frac{n+1}{n}}E_i(\ln P).$$

As an example, illustrating the applicability of extremal distributions of types 3 and 4, consider the problem of "forecast by one point".

Example 4. A discrete random process $x(n)$ is observed, its value $x(1)$ is fixed at the first observation point. What is the expected value of the maximum value at the second point at the 20th point?

Solution. It is natural to take $m = x(1)$ as an estimate of the mathematical expectation of the average process under one observation. Then, using Table 3, one finds $E\left[\max X(2)\right] = x(1) * 1.1628$ and, using dependence (14), one determines:

$$E\left[\max X(20)\right] = x(1)\frac{\ln 2}{\ln\frac{21}{20}} = x(1) * 14.2067.$$

Further, it is necessary to consider some features of the construction of extremal distributions of minimal values. Differential distribution type 5 $Q_n(x)$ is the Euler-Lagrange equation of a variational problem similar to problem (24)-(29). The difference in the formulation of the variational problem lies in the replacement of relation (29) by the dependence:

$$Q_n(x) = 1 - [1 - F(x)]^n \quad (46)$$

defining the distribution function of the smallest value $Q_n(x)$ through the distribution function of the original random variable $F(x)$. This implies that

$$H_\varepsilon = -\int_0^\infty \dot{Q}_n(x) \ln \dot{Q}_n(x) dx \to max$$

under isoperimetric conditions (25)-(28) and holonomic constraint (46).

The solution of this variational problem allows us to formulate a result similar to that stated in Theory 1.

Theory 6. Let X be a random variable with density $f(x) > 0, x \in (-\infty, \infty), Q_n(x)$ is the distribution function of the smallest (leftmost) member of the variational series constructed from a sample of finite size n from the general set determined by the random variable X. Let, further, the two first central moments of the random variable X:

$$m = \int_{-\infty}^{\infty} x f(x) dx,$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x - m)^2 f(x) dx.$$

Then the entropy maximum is reached on a distribution that satisfies the differential equation (15).

Using dependence (46) and differential equation (15), one can verify that the extremal distribution $Q_n(x)$ corresponds to the distribution function $F(x)$ of the original random variable, determined by differential equation (3).

As an example of using a type 5 distribution, consider the following problem.

Example 5. Under the conditions of Example 1, find the probability of preparing the first product by the time $T = 100\ min$.

Solution. The solution to this problem is reduced to a sequence of reasoning and actions applied in solving Example 1. As a result, passing to the normalized standard values $x = \frac{T-m}{\sigma} = \frac{100-100}{10} = 0$ according to the schedule $Q_5(x)$ for $x = 0$, one finds $Q_5(x) = P\left(t_{(1)}^5 < 100\right) = 0.72$.

Analytic properties of the function $Q_n(x)$ are similar to those of the function $F_n(X)$. Therefore, for small values of the argument $x < m - 3\sigma$ the extreme distribution of type V asymptotically approaches the normal distribution about the parameters m and $n\sigma^2$, and for $n \to \infty$ it degenerates into an extreme distribution of type 6, the quantile function of which is described by dependence (17).

Generalising the results concerning extremal distributions of types 1 and 5, it seems appropriate to find the distribution function of the order statistics $X_m^{(n)} (m = 1, 2, \dots, n)$, which provides the entropy maximum under isoperimetric conditions (26)-(28). If the initial population distribution $F(x)$ has density $f(x)$, then the distribution of order statistics $X_m^{(n)}$ имеет плотность вида:

$$f_{nm}(x) = \frac{d}{dx} P\left\{X_n^{(n)} < x\right\} = \frac{n!}{(m-1)(n-m)} F(x)^{m-1}[1 - F(x)]^{n-m} f(x) \quad (47)$$

The extremal distribution of the order statistics (the favorite member of the variational series) is determined by the solution of the Euler-Lagrange equation of the following variational problem:

$$H_\varepsilon = -\int_0^\infty f_{mn}(x) \ln f_{nm}(x)\, dx \to max \quad (48)$$

$$1 = \int_{-\infty}^\infty f_{nm}(x)\, dx \quad (49)$$

$$m = \int_{-\infty}^\infty x f(x)\, dx \quad (50)$$

$$\sigma^2 = \int_{-\infty}^\infty (x-m)^2 f(x) dx \quad (51)$$

$$F_{nm}(x) = \frac{n!}{(m-1)!(n-m)!} \int_0^{F(x)} y^{m-1}(1-y)^{n-m}\, dy \quad (52)$$

Applying the course of reasoning and those transformations and constructions that were used in the proof of Theory 1, we can obtain a nonlinear differential equation with respect to the function $F_{nm}(x)$:

$$\sigma^2 \ddot{F}_{nm} + \frac{(x-m) \sum_{\nu=0}^\infty \frac{[F_{nm}(x) - F(x_0)]^\nu}{\nu!} D^\nu Fnm}{F_{nm}(x_0)} \quad (53)$$

Satisfying the boundary conditions $F_{nm}(-\infty) = 0, F_{nm}(\infty) = 1$.

Depending on (53) $D = \frac{(m-n)!(n-m)!}{n!} F^{1-m}(1-F)^{m-n} \frac{d}{dF}$ is conversion operator with properties $D^\nu = D(D^{\nu-1}), D^0 = F$.

The reference value x_0 for the function $F_{nm}(x)$ can be chosen at any point where the distribution density (47) does not vanish. Note that for $m = n$ differential equation (53) degenerates into equation (1), and for $m = n$ into equation (15).

Thus, the solution of the variational problem (48)-(53) and the representation of the nonlinear part of the differential equation using an operator series (S. Lie's series) allows us to represent the result obtained in a compact form in the form of the following theory.

Theory 7. Let X be a random variable with density $f(x) > 0, x \in (-\infty, \infty), F_{nm}(X)$ is the distribution function of the order statistics $X_m^{(n)}$ for a variational series constructed from a sample of finite size n from the general population determined by the random variable X. Further, let the first two central moments of the random variable X be known:

$$m = \int_{-\infty}^\infty x f(x) dx$$

$$\sigma^2 = \int_{-\infty}^\infty (x-m)^2 f(x) dx.$$

Then the entropy maximum is reached for a distribution that satisfies the differential equation (53).

In conclusion, it should note some features of extremal distributions of types 7 and 8. Differential equation (18) is the Euler-Lagrange equation of a variational problem similar to problem (42). The difference in the formulation of the problems lies in the replacement of the last holonomic constraint $G_n(x) = \left[\int_0^x g(x) dx\right]^n$ by the dependence $R_n(x) = 1 - \left[1 - \int_0^x r(x) dx\right]^n$, which determines the distribution function of the smallest value $R_n(x)$ through the distribution function of the original random variable$R(x) = \int_0^x r(x) dx$.

Integration of the differential equation (18) allows one to obtain the quantile function (20), and the passage to the limit $n \to \infty$ yields dependence (22). A characteristic feature of these distributions is that they are realised in the class of finite functions (truncated on the right). It should also be noted that the dependencies for the mathematical expectations of the smallest values (21) and (23) can find practical application in the express evaluation of samples from an unknown general population.

The asymptotic behavior of the largest observation in a sample of size n from a distribution with distribution function $F(x)$ was a problem in the "classical" theory of extreme values. The central result of this theory (the theorem on three types of limit distributions) was first obtained by Fisher and L. Tippett in 1928 (*Tippett, 2012*) and was later proved in full generality in 1943 by B.V. Gnedenko (*Gnedenko, 1943*). A systematic exposition of the theory of limiting distributions of extreme quantities and applications to technical problems can be found in Gumbel's monograph (*Gumbel, 2012*). However, here it is necessary to note some specific limitations of the asymptotic theory of extremal quantities. First of all, all extreme value distributions are derived either from the exact original distribution or from a distribution of some type. The original distribution from which the extreme values are selected must belong to one of the three types of distributions. In reality, the analytic properties of the original distribution are rarely known, and hence the conditions for using the asymptotic theory of extreme values do not always correspond to observations and practical applications. Note that in the "classical" theory of extreme values, when constructing parametric forms of distributions of extreme values, the key idea (stability postulates) used earlier by Fisher and L. Tippett is used, which consists in the following. Since the largest observation in a sample of size mn can be considered as the largest member in a sample of size n, consisting of the maximum members of samples of size m, and since in the case of the existence of a limit distribution $\Lambda(x)$ both of these distributions will tend to $\Lambda(x)$ at $m \to \infty$, then $\Lambda(x)$ must satisfy the relation $\Lambda^n(a_n x + b) = \Lambda(x)$, i.e., the largest observation in the sample of size n from the distribution with the distribution function $\Lambda(x)$ must, after appropriate normalisation, itself have a limit distribution function $\Lambda(x)$. The solution of this functional equation with respect to $\Lambda(x)$ allows us to obtain the following three parametric forms (three types of distributions of extreme values built on the stability postulate):

Type 1: $\Lambda_1(x) = exp(-e^{-x}), -\infty < x\infty.$

Type 2: $\Lambda_2(x) = \begin{cases} 0, x \le 0 \\ exp, (-x^{-a}), a > 0, x > 0 \end{cases}.$

Type 3: $\Lambda_3(x) = \begin{cases} exp(-(-x)^a), a > 0, x \le 0 \\ 1, x > 0 \end{cases}.$

In this case, it is logical to compare the distributions of extreme values based on different construction principles. As an example, consider the graphs of distribution functions $F_\infty(x)$ and $\Lambda_1(x)$ $\Lambda_1(x) = exp\left(-e^{-\frac{x-a}{b}}\right)$. Naturally, the functions must be compared with the same scale and position parameters. Note that if a sample $x_1, x_2, \dots, x_k$ is given from the population with the distribution function $\Lambda_1(x) = \exp\,(-\exp\,(-\frac{x-a}{b}))$, as estimates a^* and b^* of unknown parameters a and b, one can take solutions of equations (6):

$$\frac{1}{k}\sum_{i=1}^{k} \frac{x_i - a^*}{b^*} = c$$

$$\frac{1}{k}\sum_{i=1}^{k}\left\{-\frac{x_i-a^*}{b^*}\right\}=1,$$

where C is Euler constant.

Equating the empirical values of the moments of random variables $\frac{1}{k}\sum_{i=1}^{k}x_i$ and $\frac{1}{k}\sum_{a}^{k}\exp\left(-\frac{x_i}{b}\right)$ with the theoretical ones for the distribution $F_\infty(x)$, after some algebraic transformations, one can find equations relating the parameters a^* and b^* of the distribution function $\Lambda_1(x)$ and m, and $\sigma\sqrt{2n}$ of the distribution function $F_\infty(x)$:

$$m+\sigma\sqrt{2n}I_1=a^*+cb^* \quad (54)$$

$$e^{-\frac{m}{b^*}}I_2\left(\frac{\sigma\sqrt{2n}}{b}\right)=e^{-\frac{a^*}{b^*}}$$

where

$$I_2=\int_0^1\sqrt{-E_i(\ln p)}dp=0.6967,$$

$$I_2\left(\frac{\sigma\sqrt{2n}}{b}\right)=\int_0^1 e^{-\frac{\sigma\sqrt{2n}\sqrt{-Ei(\ln p)}}{b}}dp.$$

With appropriately chosen normalizing and centering parameters $a^*=0$ and $b^*=1$, the limiting hypernormal distribution corresponds from equations (54) to the values of the parameters $\sigma\sqrt{2n}$ and m, determined from the following relations obtained from the systems of equations (54):

$$\sigma\sqrt{2n}+\ln I_2\left(\sigma\sqrt{2n}\right)=0,$$

$$m=0-\sigma\sqrt{2n}I_1.$$

The figure (*Figure 21*) shows the graph of the function $\Lambda(x)=e^{-e^{-x}}$ and the graph of the function of the limiting hypernormal distribution corresponding to it in terms of parameters $F_\infty(x)=-1.48+3\sqrt{-E_i(\ln P)}$.

A comparison of the graphs indicates the closeness of the statistical laws of extreme values obtained under different assumptions. On the other hand, it is not possible to introduce partial ordering into the set of distribution functions of the form $\Lambda_1(x)$ and $F_\infty(x)$ (the graphs illustrate this circumstance): the function $\Lambda_1(x)$ does not dominate the function in the sense of the first order. Therefore, when choosing a mathematical model of the mechanism for the formation of extreme random variables, when constructing estimates that guarantee their effectiveness, etc., one should be guided by the stability postulate, if it is known that the distribution of the initial random variable belongs to the exponential type distribution (4), and by the principle of maximum uncertainty, if only estimates of the mathematical expectation and the variance of the original random variable (1) are known.

As an example, illustrating the possibility of using the considered laws of distributions of extreme values, it is presented the distribution function of the levels of water rise at the mouth of the Neva, constructed according to statistical data fixed since 1703 (7):

$$\Lambda_1(x)=exp\left\{-e^{-\frac{x-221,4}{30,38}}\right\},$$

$$x=183+7.97\sqrt{2n}\sqrt{-E_i(\ln P)}$$

where x is level of Neva's mouth water rise in centimeters.

Comparison of the probabilities of occurrence of events $P\{x \leq 200\}^n$, built on these dependencies, is similar to the nature of dependencies (*Figure 21*).

The discussion of the applications of the theory of extrema of sequences of random variables is given fragmentarily and does not yet cover the whole variety of possible problems related to extreme values, because "... nature speaks to us in the language of mathematics" (G. Galileo) and it is quite appropriate that this area is worthy of study on its own. yourself. Therefore, it seems appropriate to proceed to the consideration of the basic problems of constructing models and some generalisations, confining ourselves mainly to the formulation of the main results.

Conclusion

Thus, in the study course, the definition of an extreme value within the framework of the theory was refined, the typology of the distribution of maximum values was analysed, seven theories of the hypernormal distribution were identified and their proofs were presented, and practical examples of the application of each theory were given. The practical significance of the study of extreme random variables models in various areas of industrial human activity was confirmed.

The materials of the study can be used in the widest range: from application in risk management of industrial production to predicting the probabilities of natural phenomena, which makes it possible to prevent significant economic and social losses of society, as well as make a tangible contribution to programming the probabilities of the development of the society of the future.

References:

Akimov, V. A., Bykov, V. A., & Shchetinin, E. Yu. (2009). *An introduction to extreme value statistics and its applications*. Russian Emergency Situations Ministry. Moscow. (In Russian)

Ambler, S. (2002). *Agile Modeling: Effective Practices for EXtreme Programming and the Unified Process*. J. Wiley.

Auer, K., & Miller, R. (2001). *Extreme Programming Applied: Playing to Win* (1st ed.). Addison-Wesley Professional.

Beck, K. (2003). *Test Driven Development: By Example* (2nd ed.). Addison-Wesley Professional.

Beck, K., & Fowler, M. (2001). *Planning Extreme Programming*. Addison-Wesley Professional.

Buychik, A. (2021). Updating the parameters of the development of effective economic thought to motivate society to finance innovative activities. *European Scientific e-Journal*, *4*(10), 7-16.

Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, *44*(3), 423-453. (In French)

Gumbel, E. J. (2012). *Statistics of Extremes*. New York: Columbia University Press.

Komissarov, P. V. (2021). Determination of the centric rate of the economic stability domain for manufacturing enterprises. *European Scientific e-Journal*, *4*(10), 27-36.

Khatskevich, V. L. (2013). On some extreme properties of mean values and mathematical expectations of random variables. *Bulletin of the Voronezh State Technical University*, *9*(3-1), 39-44. (In Russian)

Khatskevich, V. L. (2020a) Average characteristics of fuzzy random variables and their extreme properties. *Proceedings of the International Conference of the Voronezh Spring Mathematical School "Pontryagin Readings – XXXI"*, 229-230. (In Russian)

Khatskevich, V. L. (2020b). On extremal properties of mean fuzzy-random variables. *S.G. Krein Voronezh Winter School of Mathematics*, 307-312. (In Russian)

Kvashko, L. P., & Losev, A. S. (2013). Introduction to the theory of extreme values. *In the World of Scientific Research*, 57-62. (In Russian)

Mikhailov, V. S. (2012). Theory of extreme values as a risk management tool. *Economics of Nature Management*, *6*, 130-133. (In Russian)

Popov, V. A. (2013). *Probability Theory. Part 2. Random Variables: Textbook*. Kazan: Kazan University. (In Russian)

Smolyakov, E. R. (2011). Fundamentals of extremal theory of dimensions and fundamental new results. *Proceedings of the Institute for System Analysis of the Russian Academy of Sciences*, *61*(4), 110-120. (In Russian)

Shchetinin, E. Yu., & Nazarenko, K. M. (2008). *Mathematical models, methods for estimating distribution functions of extreme values*. Dubna: Publishing House of the Branch of the Joint Institute for Nuclear Research. (In Russian)

Tippett, L. (2012). *Statistics* (2nd ed.). London: Oxford University Press.

Tippett, L. (2013). *Random Sampling Numbers* (3rd ed.). London: Cambridge University Press.

Appendix

Table 1. Extreme distributions and statistical characteristics of the original random variable

Determination	Extreme distributions and statistical characteristics of the original random variable			
	Mathematical expectation m and variance σ^2		Expected value m	
$P\{X_n^{(n)} < x\}$	$F_n(X)$ type 1	$F_\infty(X)$ type 2 (limiting $(n \to \infty)$)	$G_n(X)$ type 3	$G_\infty(X)$ type 4 (limiting $(n \to \infty)$)
$P\{X_1^{(n)} \geq x\}$	$Q_n(X)$ type 5	$Q_\infty(X)$ type 6 (limiting $(n \to \infty)$)	$R_n(X)$ type 7	$R_\infty(X)$ type 8 (limiting $(n \to \infty)$)

Table 2. Values of the function of the limiting hypernormal distribution $F_\infty(x)$, whose argument is the value $\tilde{x}$

$\tilde{x}$	$F_\infty(\tilde{x})$	$\tilde{x}$	$F_\infty(\tilde{x})$
0,1	0,0230	1,6	0,830
0,2	0,0707	1,7	0,858
0,3	0,126	1,8	0,873
0,4	0,192	1,9	0,904
0,5	0,259	2,0	0,921
0,6	0,326	2,1	0,936
0,7	0,393	2,2	0,949
0,8	0,457	2,3	0,959
0,9	0,518	2,4	0,967
1,0	0,573	2,5	0,975
1,1	0,628	2,6	0,981
1,2	0,676	2,7	0,985
1,3	0,722	2,8	0,988
1,4	0,762	2,9	0,992
1,5	0,798	3,0	0,996

Table 3. The value of the sum S_n of a series from 1 to 5

n	1	2	3	4	5
S_n	1	1,628	2,2643	2,8878	3,4967
n	6	7	8	9	10
S_n	4,0904	4,6694	5,2338	5,7844	6,3214

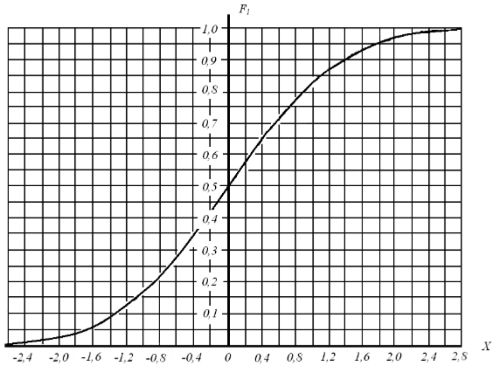

Figure 1. Distribution function $F_1(x) = P\left\{x_1^{(1)} < x\right\}$; $(m = 0, \sigma^2 = 1)$

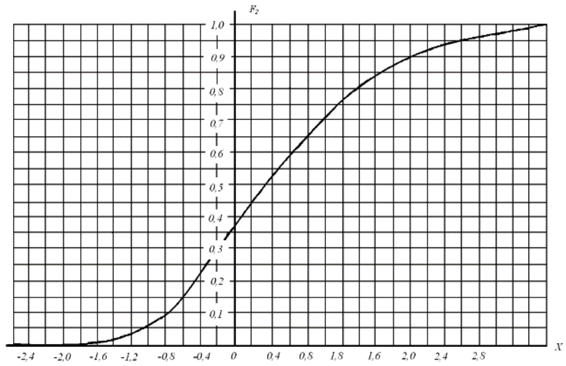

Figure 2. Distribution function $F_2(x) = P\left\{x_2^{(2)} < x\right\}$; $(m = 0.4634;\ \sigma^2 = 1.1077)$

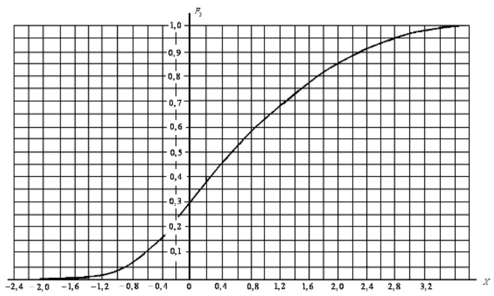

Figure 3. Distribution function $F_3(x) = P\left\{x_3^{(3)} < x\right\}$; $m = 0.5355;\ \sigma^2 = 1.3594$

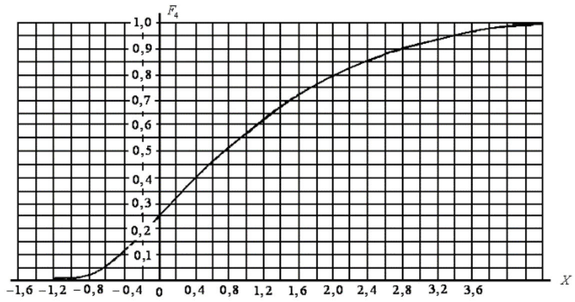

Figure 4. Distribution function $F_4(x) = P\left\{x_4^{(4)} < x\right\}$; $m = 0.9764;\ \sigma^2 = 1.6622$

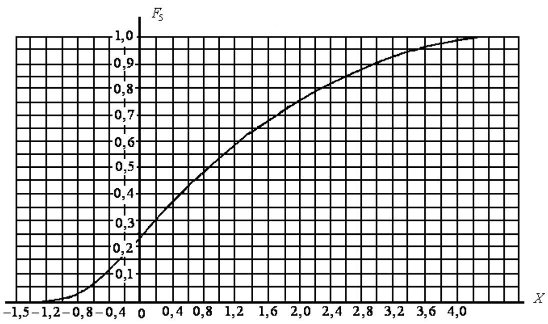

Figure 5. Distribution function $F_5(x) = P\left\{x_5^{(5)} < x\right\}$; $m = 1.1355;\ \sigma^2 = 2.0190$

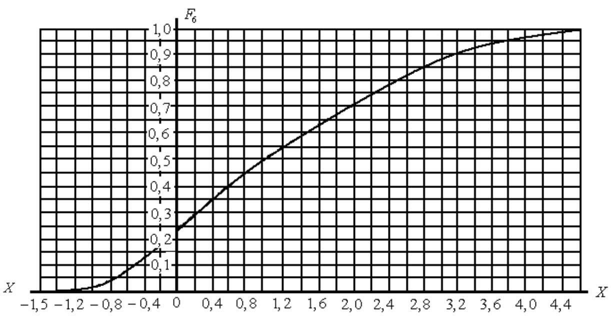

Figure 6. Distribution function $F_6(x) = P\left\{x_6^{(6)} < x\right\}$; $m = 1.2458;\ \sigma^2 = 2.3316$

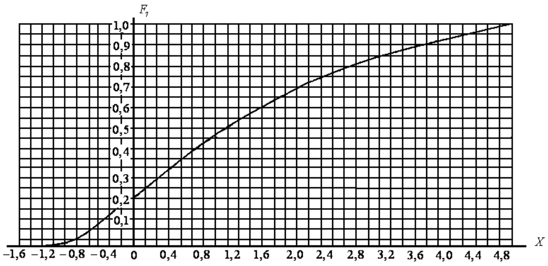

Figure 7. Distribution function $F_7(x) = P\left\{x_7^{(7)} < x\right\}$; $m = 1.4656$; $\sigma^2 = 2.9520$

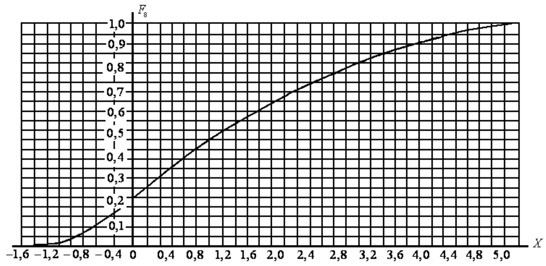

Figure 8. Distribution function $F_8(x) = P\left\{x_8^{(8)} < x\right\}$; $m = 1.5504$; $\sigma^2 = 3.3127$

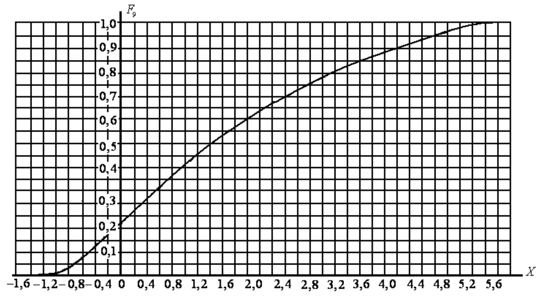

Figure 9. Distribution function $F_9(x) = P\left\{x_9^{(9)} < x\right\}$; $m = 1.5748$; $\sigma^2 = 3.5122$

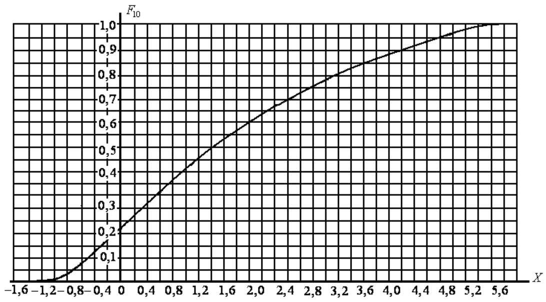

Figure 10. Distribution function $F_{10}(x) = P\left\{x_{10}^{(10)} < x\right\}$; $m = 1.5792$; $\sigma^2 = 3.6438$

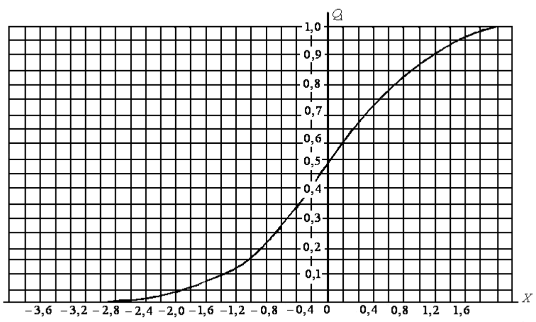

Figure 11. Distribution function $Q_1(x) = P\left\{x_1^{(1)} < x\right\}$; $m = 0$; $\sigma^2 = 1$

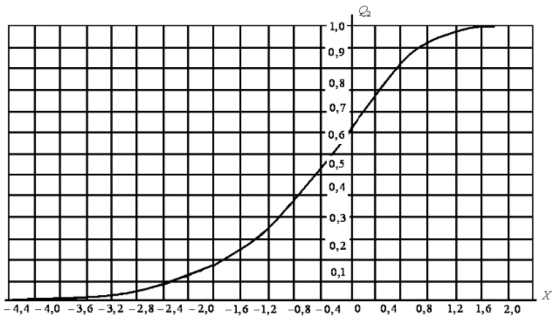

Figure 12. Distribution function $Q_2(x) = P\left\{x_1^{(2)} < x\right\}$; $m = -0.5321$; $\sigma^2 = 1.0492$

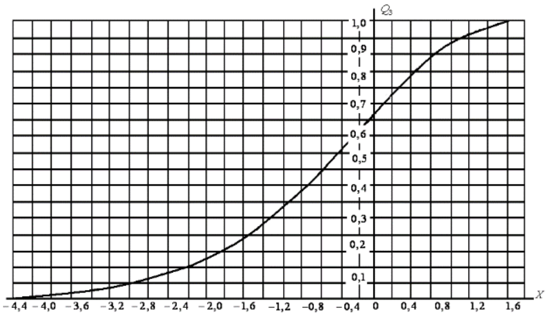

Figure 13. Distribution function $Q_3(x) = P\left\{x_1^{(3)} < x\right\}$; $m = -0.7330$; $\sigma^2 = 1.1250$

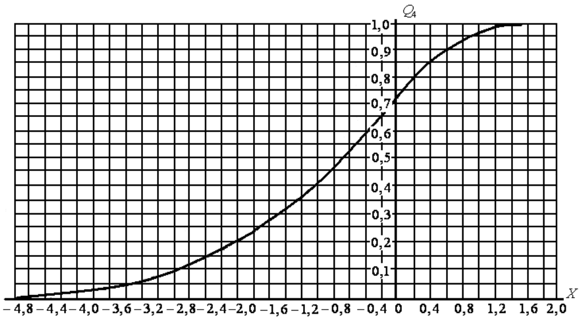

Figure 14. Distribution function $Q_4(x) = P\left\{x_1^{(4)} < x\right\}$; $m = -0.9445$; $\sigma^2 = 1.1546$

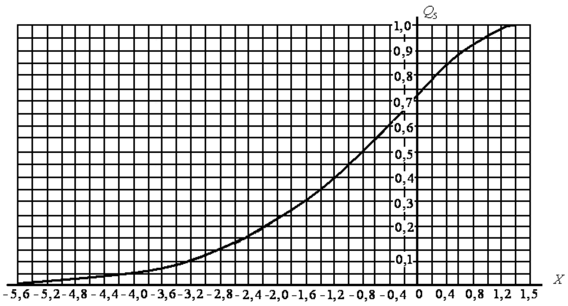

Figure 15. Distribution function $Q_5(x) = P\left\{x_1^{(5)} < x\right\}$; $m = -1.0692$; $\sigma^2 = 2.1477$

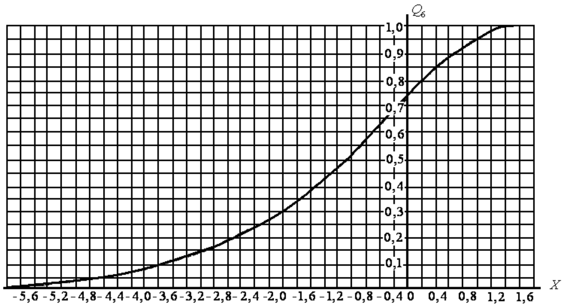

Figure 16. Distribution function $Q_6(x) = P\left\{x_1^{(6)} < x\right\}$; $m = -1.1950$; $\sigma^2 = 2.5258$

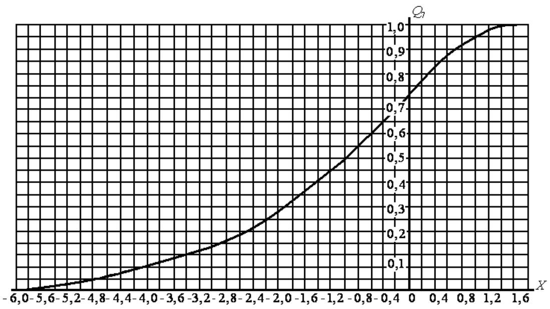

Figure 17. Distribution function $Q_7(x) = P\left\{x_1^{(7)} < x\right\}$; $m = -1.2977$; $\sigma^2 = 2.9194$

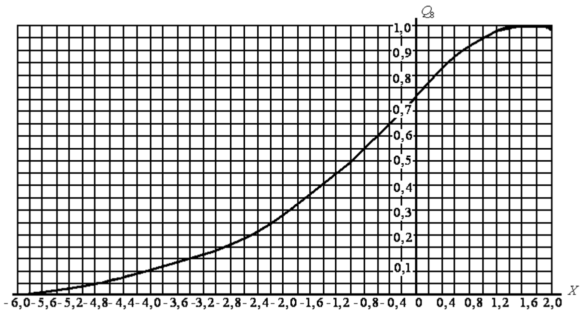

Figure 18. Distribution function $Q_8(x) = P\left\{x_1^{(8)} < x\right\}$; $m = -1.4451$; $\sigma^2 = 3.2427$

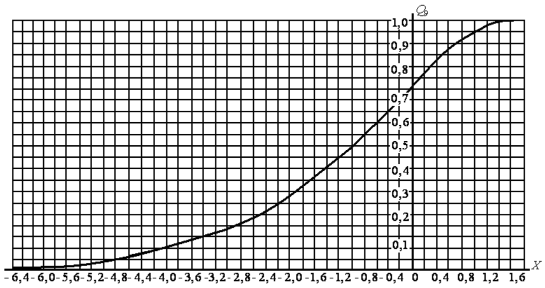

Figure 19. Distribution function $Q_9(x) = P\left\{x_1^{(9)} < x\right\}; m = -1.5573;\ \sigma^2 = 3.6019$

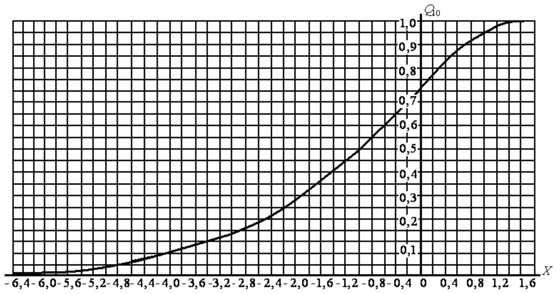

Figure 20. Distribution function $Q_{10}(x) = P\left\{x_1^{(10)} < x\right\}; m = -1.5703;\ \sigma^2 = 4.0949$

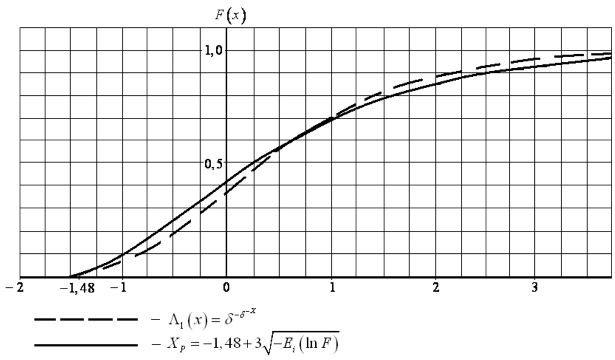

Figure 21. Distribution function $F_\infty(x) * \Lambda_1(x); m = -1.5573;\ \sigma^2 = 3.6019$