УДК 625.7/8

Osiaiev Iu.M. Candidate of Technical Sciences, **Bryk Douglas (Canada)**

THE USE OF THE DIRICHLET DISTRIBUTION IN PROBABILITY STATISTIC METHODS OF RESEARCH

Анотація. У статті розглянуто основні методи обробки статистичних даних за допомогою бета-розподілу,узагальненням якого є розподіл Діріхле.

Ключові слова: оцінка новітніх технологій при будівництві та ремонтах автомобільних доріг,.бета – розподіл,розподіл Діріхле.

Аннотация. В статье рассмотрены основные методы обработки статистических данных с помощью бета-распредения, обобщением которого является распределение Дирихле

Ключевые слова:оценка новых технологий при строительстве и ремонтах автомобильных дорог, бета – распределение, распределение Дирихле.

Annotation. The article considers the basic methods of statistic data procession with the help of the beta-distribution, the generalization of which is the Dirichlet distribution.

Keywords: evaluation of new technologies in automotive road construction and repair, beta - distribution, Dirichlet distribution.

Stating the problem. One of the current directions in the field of conducting applied research is the substantiation of the possibility to apply relevant mathematical knowledge to solve technical and economic tasks with the purpose of the increase in the objectiveness of the results of research.

The essence of the problem. High information uncertainty in the period of the new technologies introduction in automotive road construction and repair and the forecast of the determination of the relevant criteria of efficiency.

The objective of the article. The possibility to apply relevant mathematical knowledge, in particular the beta-distribution, to solve technical and economic tasks with the purpose of the increase in the objectiveness of results. It becomes especially important in the evaluation of the new technologies introduction in automotive road construction and repair.

The exposition of the main material. One of the current directions in the field of conducting applied research is the substantiation of the possible application of relevant mathematical knowledge to solve technical and economical tasks with the purpose of increasing efficiency. The application of relevant mathematical knowledge is becoming especially important in the evaluation of modern technologies in automotive road construction and repair as high information uncertainty in the period of the modern technologies introduction admits a possibility of different forecasts of technical and economic indicators.

In the theory of probability and mathematical statistics the Dirichlet distribution, often denoted as Dir (α), is a family of continuous multivariate probability distributions parametrized by a vector α of positive reals. The Dirichlet distribution is the multivariate generalization of the beta-distribution. That is, its probability density function returns the belief that the probabilities of K rival events are x_i given that each event has been observed α_i-1 times.

The probability density function for the Dirichlet distribution of order *K* is:

$$f(x_1, \dots x_{k-1}; \alpha_1, \dots \alpha_k) = \frac{1}{B(\alpha)} \prod_{i=1}^{k} x_i^{\alpha_{i-1}} \quad , \qquad (1)$$

where, $x_i \geq 0, \ \ \alpha_i \geq 0, \sum_{i=1}^{k} x_i = 1$

Let X=(X_1,…X_k) ~ Dir(α) i $\alpha_0 = \sum_{i=1}^{k} \alpha_i$, then

$$E[X_i|\alpha] = \frac{\alpha_i}{\alpha_0} \qquad (2)$$

$$Var[X_i|\alpha] = \frac{\alpha_i(\alpha_0-\alpha_i)}{\alpha_0^2(\alpha_0+1)} \quad (3)$$

$$Cov[X_iX_j|\alpha] = \frac{-\alpha_i\alpha_j}{\alpha_0^2(\alpha_0+1)}$$

The mode of the distribution is the vector x ($x_1, ..., x_K$) with

$$x_i = \frac{\alpha_i-1}{\alpha_0-K} \qquad \alpha_i > 1 \quad (4)$$

The Dirichlet distribution is the conjugate prior distribution of the multinomial distribution, that is: if

$$\beta|X = (\beta_1, \dots, \beta_k)|X \sim \mathrm{Mult}(X) \quad (5)$$

where β_i – the number of entries *i* in the sample from *n* points of the categorical distribution on {1, ..., *K*} determined through *X*, then

$$X|\beta \sim \mathrm{Dir}\,(\alpha+\beta) \quad (6)$$

This relationship is used in Bayesian statistics to estimate the underlying parameters, X, of the discrete probabilitydistribution given a collection of *n* samples. Obviously, if the prior distribution denoted as Dir (α), then Dir (α+β) is the posterior distribution after a number of observations with histogram β

The normal law is most often used in studying the distribution of technical and economical parameters. But the given approach is not always grounded and, as a result, may lead to approximate evaluations. Besides, many researchers admit the existence of a big mass of initial data, which is not always possible with a limited number of objects being studied. In a number of works, the beta-distribution is said to be possibly accepted as the typical distribution of certain parameters in time. [1,3].

Thus, in the forecast of specific technical and economical parameters with the substantiation of the algorithm of determining the numerical values of its statistic characteristics. It is known that a random quantity has the beta-distribution with parameters (α, β) $(\alpha > 0, \beta < 0)$, if

$$f(t) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1}(1-t)^{\beta-1} \quad (7)$$

where $t \in [0,1]$; 0, $t \notin [0,1]$

With the use of the normalized quantity of the parameter being studied t (in the interval of changes [0; 1]) the probability density of the forecast parameter looks:

$$f(t) = ct^{\alpha-1}(1-t)^{\beta-1} \quad (8)$$

where α, β – statistic parameters of the distribution,

c – constant

the quantity of the constant can be determined with the help

$$c = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}, \quad (9)$$

де Г(n) = (n-1), (10)

denote $$\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \mathrm{B}(\alpha,\beta). \quad (11)$$

Then

$$f(t) = \frac{1}{B(\alpha,\beta)} t^{\alpha-1}(1-t)^{\beta-1}, \quad (12)$$

and $$\int_0^1 f(t)\,dt = 1$$

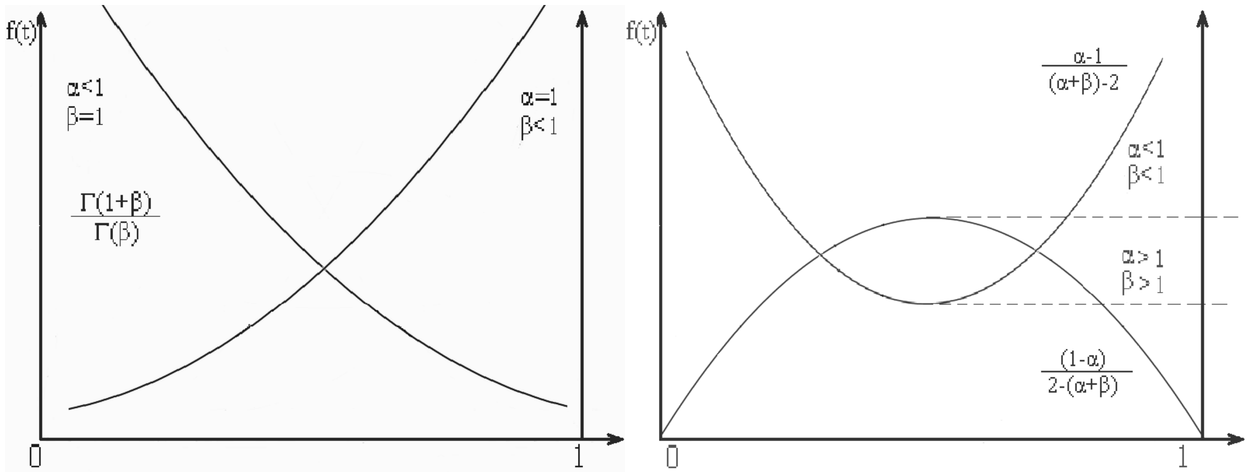
The expectation value and variance of the random quantity in this case equal to

$$\mathrm{M}(t) = \frac{\alpha}{\alpha+\beta} \quad (13)$$

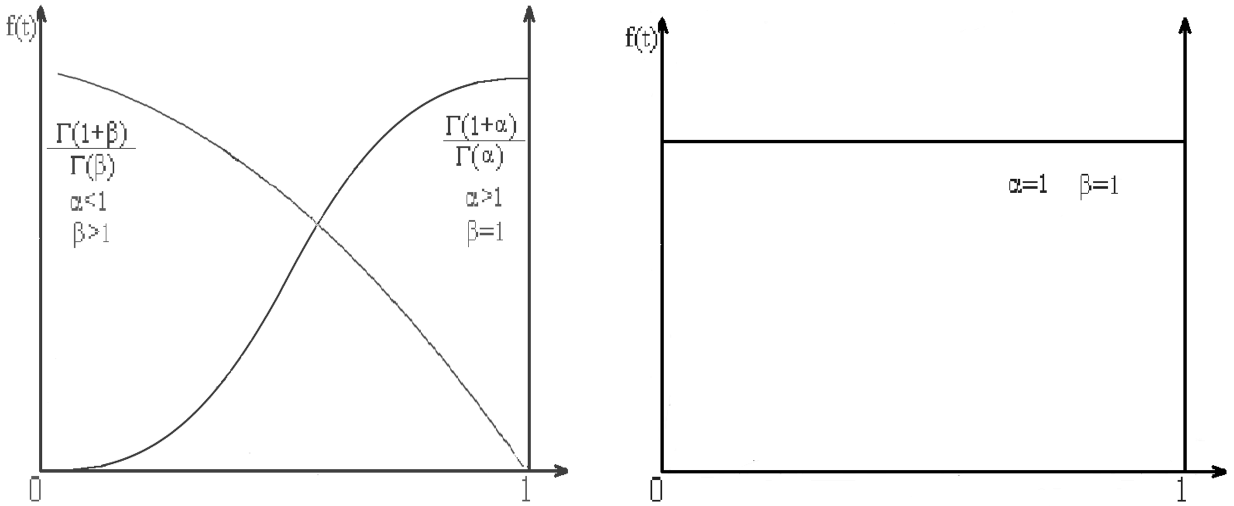
$$D(t) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \quad (14)$$

In picture. 1 there are examples of the curved functions of the beta-distribution density with different values of its statistic parameters:

a) b)



c) d)



Picture. 1 – Density of the beta-distribution

Let us look at the change of the parameter being studied in the interval [a,b] that is $a \le x \le b$. In this case:

$$t = \frac{x-a}{b-a}, \quad x = a + (b-a)\cdot t$$

$$f(x) = \frac{1}{(b-a)^{\alpha+\beta-1}\cdot B(\alpha;\beta)}(x+a)^{\alpha-1}(b-x)^{\beta-1} = c(x-a)^{\alpha-1}(b-x)^{\beta-1}, \qquad (15)$$

That is determining the constant of the beta-distribution is conducted as

$$c = \frac{1}{(b-a)^{\alpha+\beta-1}B(\alpha;\beta)}, \qquad (16)$$

$$\mathrm{M}(x) = a + (b-a)$$

$$\mathrm{M}(t) = a + (b-a)\frac{\alpha}{\alpha+\beta} = \frac{a\beta+b\alpha}{\alpha+\beta} \qquad (17)$$

The variance of the random quantity is:

$$D(x) = (b-\alpha)^2 \qquad D(t) = \frac{2\beta(b-a)^2}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{18}$$

Given that $M(x) = \bar{x}_b \ i \ D(x) = \sigma_b{}^2$ (where $\bar{x}_b$ and $\sigma_b{}^2$ - respectively the parameter evaluation, its variance which can be found on the data of the selective totality). Evaluations of statistic parameters α i β can be received by solving the system of equation:

$$\begin{cases} \dfrac{a\beta+b\varphi}{\varphi+\beta} = \bar{x}_b \\ \dfrac{\alpha\beta(b-a)^2}{(\alpha+\beta)^2(\alpha+\beta+1)} = \sigma_b{}^2 \end{cases} \tag{19}$$

Besides, accepting, on the empirical data, the type of the curved line of the beta-distribution density and giving different values of the parameters α i β (picture.1), it is possible to determine the quantity of the distribution constant with the formula (16) and check the relevance of the empirical distribution to the theoretical assumption with the help of the known criteria of mathematical statistics.

Thus, in studying the distribution of technical and economic parameters of the modern technologies being studied it is possible to study the beta-distribution with the probability density determined by the dependence (15), where a - the minimal evaluation of the parameter, and b - the maximum evaluation of the parameter.

CONCLUSIONS

1. The beta-distribution (the Dirichlet distribution) can be accepted as the typical distribution in studying technical and economic parameters of modern technologies in time in automotive road construction and repair, which allows to receive the probability density function of the parameter being studied with the

absence of a meaningful mass of initial data and to objectively determine its evaluations.

2. Determining specific statistic characteristics of the distribution being studied can be conducted on the basis of the estimation stated in the article, and also with the help of the analysis of the empirical data distribution form and the further analysis with the help of the statistic criteria of the empirical and theoretical distributions relevance.

3. The considered research method expands the possibilities of the application of the probability methods of the evaluation of technical and economic processes and indicators, which leads to the increase of the objectiveness and substantiation of the received results.

LITERATURE

1. Hovanov N., Kornikov V., Seregin I. Multicriteria estimation under uncertainty // Proceedings of the International Conference Signals, Data, Systems. Heiderabad (India). December 12-14, 1994. Vol. 1. Heiderabad: AMSE Press, 1994. P. 83–91
2. Годенко Д.И. Статистические методы в экономических системах. – М. “Статистика “,1970 – С.320
3. Мізіна О.В. Прогнозування техніко-економічних параметрів на основі бета – розподілу. Научные труды ДонНТУ.2004. – Выпуск 87- С.125 – 134
4. Грушко И.М., Сиденко В.М. Основы научных исследований. 3-е изд. – Харьков: Вища школа. Изд-во при Харьк.ун-те,1983.- 224с.
5. Ковалев Д.,Плетникова И. Количественная оценка уровня єкономической безопасности предприятия // Экономика Украины. 2001. - №4 – С.35-40