DETERMINATION OF A GENERAL NORMAL DISTRIBUTION

Subotić Borivoj[1], Stanković Ivica[2]
Faculty for Education of the Executives, Novi Sad, SERBIA,
e-mail: borivojesubotic@gmail.com
Faculty for Education of the Executives, Novi Sad, SERBIA,
e-mail: ivica.stankovic@fpsp.edu.rs

Abstract: *In this paper, one way of the generalization of the normal distribution curve, using interpolation system and Chebyshev polynomials because it represents an indispensable arsenal in the economic analysis and econometrics that each serious economist is using, in order to get a quick and elegant result of its research. Therefore, it is a non-standard apparatus known only to rare economists. The name of the Russian mathematician Chebyshev in mathematical statistics is related mainly to so-called Chebyshev inequality, which also called Chebyshev law of large numbers, while his polynomials remain unknown to many theorists of statistics and econometrics. Therefore, the aim of this article is to fill this gap and to put closer Chebyshev polynomial approximation to a large number of economists.*

Keywords: *normal curve, Chebyshev polynomials, generalized normal distribution, recurrent formulas, characteristic of orthogonal polynomials*

1. INTRODUCTION

In 1855, Chebyshev proposes a general interpolation formula (Chebyshev order, which satisfies the conditions of the method of the smallest squares and is expressed using orthogonal polynomials. It has the following form:

$$F(x) = \frac{\sum[\psi_0(x_i)\theta^2(x_i)F(x_i)}{\sum[\psi_0^2(x_i)\theta^2(x_i)]}\Psi_0(x_i) + \cdots + \frac{\sum[\psi_m(x_i)\theta^2(x_i)F(x_i)}{\sum[\psi_m^2(x_i)\theta^2(x_i)]}\psi_m(x_i)$$

Where:
$F(x)$ – required function,
$F(x_i)$ - known values of the required function,
θ^2 - weight (for example, the distribution of frequencies), and
$\psi_m(x_i)$ - Chebyshev polynomials

Chebyshev showed that his polynomials have three very important characteristics:

2. TH FEATURES OF CHEBYSHEV POLYNOMIALS

1. Zero polynomial is always one:
$\psi_0(x) = 1$

2. Polynomials have the property of orthogonality:
$\sum[\psi_m(x)\psi_k(x)] = 0$ za $\neq k$; $\sum[\psi_k^2(x) = B_k^2$

3. Successive polynomials are associated by recurrent formula:
$\psi_m(x) = (\gamma_m x + \beta_m)\psi_{m-1}(x) - \alpha_m \psi_{m-2}(x)$

Subsequent research has shown that the Chebyshev interpolation order substantially simplifies if we measure the distance x - a from the mean value (in the form $x_{1\cdot 0} = x_1 - \overline{x_1}$) and k in medium square deviation units, i.e . shaped into $k = -\frac{1}{2\sigma^2} = -\frac{\sum x_{1\cdot 0}^2}{2\sum x_{0\cdot 0}^2}$

If we introduce a new variable t using the formula:
$t^2 = \frac{x_{1\cdot 0}^2 \sum x_{0\cdot 0}^2}{2\sum x_{1\cdot 0}^2}$ where: $\sum x_{0\cdot 0}$ -is the number of cases or observed objects (analyzed) (5)
we have:
$\theta^2(x) = \varphi_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ - a formula for the function of the normal distribution density (6)

$\psi_1(t) = -t = \frac{x_{1\cdot 0}}{\sqrt{\sum x_{0\cdot 0}^2}} \sum \psi_1(t) F(x) = 0$ (7)

$\psi_2(t) = t^2 - 1 \sum \psi_2(t) F(x) = 0$ (8)

$\psi_3(t) = 3t - t^3 \sum \psi_3(t) F(x) = \frac{\sum x_{1\cdot 0}^3}{\sum x_{0\cdot 0}^2} = \mu_3$ (9)

$\psi_4(t) = t^4 - 6t^2 + 3 \sum \psi_4(t) F(x) = \frac{\sum x_{1\cdot 0}^4}{\sum x_{0\cdot 0}^2} - 3 = \mu_4 - 3$ (10)

3. THE EXAMPLE OF CHEBYSHEV INTERPOLATION ORDER

Let us select one numerical example to show how to decompose a function of frequency allocation into the order Chebyshev interpolation order. Let it be in the field of forest management.
Determination of the normal general curves is shown in the table. The calculations in the table are based on the use of functions $\varphi_l(x)$, which are, in fact, the products of Chebyshev polynomial $\psi_l(x)$ and the function $\varphi_0(x)$. Values of the function $\varphi_0(x)$ were taken from the regular table for a normal probability distribution curve .
It is useful to calculate the coefficients of generalized curves using central moments, because in this case, zero, third and forth term remain in the expression. So we get a curve:
$y = \frac{N}{\sigma}\left[\varphi_0(x) + \frac{\mu_3}{\sigma^3} \cdot \frac{\varphi_3(x)}{6} + \left(\frac{\mu_4}{\sigma^4} - 3\right) \cdot \frac{\varphi_4(x)}{24} + \cdots\right]$ (11)

Table 1: Determination of general normal curve (frequency of pine trees by diameter)

x	$x-m$	$t=\frac{x-m}{\sigma}$	$\varphi_0(x)$	$\varphi_3(x)$	$\varphi_4(x)$	$\frac{\mu_3}{\sigma^3}\cdot\frac{\varphi_3(x)}{6}$	$\frac{\mu_3}{\sigma^3}\cdot\frac{\varphi_3(x)}{6}$	(4)+(7)+(8)	$\frac{N}{\sigma}\times 9\overline{y_t}$	Observed frequencies y
6	- 6,25	- 2,338	0,02594	+ 0,14955	+ 0,00213	- 0,01237	- 0,00000	0,013	44,82	9
7	- 5,25	- 1,964	0,05799	+ 0,09762	- 0,30529	- 0,00807	+ 0,00029	0,050	172,39	132
8	- 4,25	- 1,590	0,11270	- 0,08456	- 0,65113	+ 0,00699	+ 0,00063	0,120	413,74	370
9	- 3,25	- 1,216	0,19047	- 0,35235	- 0,70193	+ 0,02914	+ 0,00067	0,220	758,52	868
10	- 2,25	- 0,842	0,27987	- 0,53987	- 0,21020	+ 0,04465	+ 0,00020	0,325	1120,54	1226
11	- 1,25	- 0,468	0,35756	- 0,46535	+ 0,61993	+ 0,03848	- 0,00060	0,395	1361,87	1332
12	- 0,25	- 0,094	0,39718	- 0,11167	+ 1,17046	+ 0,00921	- 0,00112	0,405	1396,36	1325
13	+ 0,75	+ 0,281	0,38350	+ 0,31477	+ 0,97118	- 0,02603	- 0,00093	0,356	1227,42	1266
14	+ 1,75	+ 0,655	0,32192	+ 0,54207	+ 0,19635	- 0,04483	- 0,00019	0,277	955,04	908
15	+ 2,75	+ 1,029	0,23496	+ 0,46931	- 0,52438	- 0,03881	+ 0,00050	0,197	679,22	630
16	+ 3,75	+ 1,403	0,14910	+ 0,21580	- 0,73589	- 0,01785	+ 0,00071	0,132	455,11	470
17	+ 4,75	+ 1,777	0,08227	- 0,02305	- 0,49154	+ 0,00191	+ 0,00047	0,085	293,06	321
18	+ 5,75	+ 2,151	0,03947	- 0,13809	- 0,13235	+ 0,01142	+ 0,00013	0,051	175,84	185
19	+ 6,75	+ 2,525	0,01646	- 0,14030	+ 0,08879	+ 0,01160	- 0,00009	0,028	96,54	95
20	+ 7,75	+ 2,899	0,00597	- 0, 09353	+ 0,13853	+ 0,00773	- 0,00013	0,013	44,82	57
21	+ 8,75	+ 3,273	0,00188	- 0,04752	+ 0,10068	+ 0,00393	- 0,00010	0,006	20,69	22
Σ	-	-	2,65724	-	-	-	-	2,673	9215,98	9216

Determination of central moments is carrying out in a conventional manner, based on the following formulas:

$$n = 9216 \quad (12)$$

$$\frac{\sum yx}{n} = 0{,}025 \quad (13)$$

$$\frac{\sum yx^2}{n} = 7{,}210 \quad (14)$$

$$\frac{\sum yx^3}{n} = 14{,}845 \quad (15)$$

$$\frac{\sum yx^4}{n} = 164{,}234 \quad (16)$$

$$\mu_2 = \frac{\sum yx^2}{n} - \frac{[yx]^2}{n^2} = 7{,}147 \quad (17)$$

$$\sigma = \sqrt{\mu_2} = 2{,}673 \quad (18)$$

$$\mu_3 = \frac{\sum yx^3}{n} - \frac{3(\sum yx^2)\cdot(\sum yx)}{n^2} + \frac{2[\sum yx]^3}{n^3} = 9{,}469 \qquad (19)$$

$$\mu_4 = \frac{\sum yx^4}{n} - \frac{4(\sum yx^3)\cdot(\sum yx)}{n^2} + \frac{6[\sum yx]^2[\sum yx]^2}{n^3} - \frac{3(\sum yx)^4}{n^4} = 152{,}081 \qquad (20)$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\sigma_3} = 0{,}496 \qquad (21)$$

$$E = \left(\frac{\mu_4}{\sigma_4} - 3\right) = -0{,}023 \qquad (22)$$

$$A_3 = \frac{\sqrt{\beta_1}}{6} = 0{,}08267 \qquad (23)$$

$$A_4 = \frac{E}{24} = -0{,}00096 \qquad (24)$$

Putting these values in the interpolation polynomial (11), we finally obtain the following equation:

$$\overline{y_5} = \frac{9216}{2{,}673}[\varphi_0(x) + 0{,}08267\cdot\varphi_3(x) - 0{,}00096\cdot\varphi_4(x)] = 3447{,}81\cdot\varphi_0(x) + 285{,}03\cdot\varphi_3(x) - 3{,}308\cdot\varphi_4(x) \qquad (25)$$

The calculated values of the distribution of the pinewood by the diameter using the obtained formula are given also in table 1 (penultimate column).
We define quadratic approximation using the following formula:

$$\sum d_4^2 = \sum(y - \overline{y_4})^2 = 41363{,}13 \qquad (26)$$

wherein :

$$n - 5 = 11 \qquad (27)$$

$$\overline{d_4^2} = 3760{,}2 \qquad (28)$$

$$\overline{d_4} = 61{,}3 \qquad (29)$$

The resulting approximation is quite satisfactory.

4. RESUME

This paper showed the basic theory and in a practical way demonstrated why interpolation order and Chebyshev polynomials have fundamental importance for the theory of the distribution function probability and the distribution of computation.

BIBLIOGRAPHY:

[1] Немчинов, В.: Экономико–математические методы и модели. Москва: Мысль, (1965).
[2] Vranić,V.: Vjerojatnost i statistika. Zagreb:Tehnička knjiga, (1970).
[3] Hogg, C.: Introductionto mathematical statistics. NewYork: Macmillan, (1959).
[4] Munroe, M.E.: Theory of Probability. NewYork: McGraw-Hill, (1951).