A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S
FOLIA OECONOMICA 3(314)2015
http://dx.doi.org/10.18778/0208-6018.314.02

Tadeusz Gerstenkorn[*]

REMARKS ON THE FORMULA FOR THE MOMENTS OF THE PÓLYA PROBABILITY DISTRIBUTION

Abstract. The probability distribution of a random variable can be characterized by some numbers called parameters of the distribution. Moments belong to the most frequently used parameters. We focus on the Pólya distribution because we can easily obtain from it, as special cases, some distributions important in statistics such as: binomial, negative binomial and Poisson (the last one in the limit procedure).

In 1972 G. Mühlbach gave very interesting formulae for the moments of the Pólya distribution. The author did not investigate the evaluation of the numerical efficacy of the formula for the moments. We will show that it is possible to demonstrate this formula in a simpler form, which is important from practical point of view.

Keywords: moments of the probability distribution, Pólya probability distribution.

1. INTRODUCTION

In principle a random variable is described exactly enough by its probability distribution. However, practical reasons impose the need to find some numerical characteristics of a distribution since they are short descriptions and give quick comparison of the distributions with themselves.

In theoretical statistics as well as in statistics used in economy, a need frequently arises to get the basic properties of an investigated population. Then we do not mention many particulars and we characterize the needed properties many a time with the help of one or a few numbers. First of all, we use here the arithmetic mean, mean deviation and, in the case of a distribution, evidently the moments.

2. THE G. PÓLYA DISTRIBUTION

In the paper we study a formula for the moments of the G. Pólya probability distribution given in 1972 by G. Mühlbach. For this purpose we should bear in mind that this distribution is expressed by the formula

[*] Professor Emeritus, University of Łódź, Poland, e-mail: tadger@math.uni.lodz.pl

$$P(X=k)=\binom{n}{k}\frac{p(p+a)...[p+(k-1)a]q(q+a)...[q+(n-k-1)a])}{1(1+a)(1+2a)...[1+(n-1)a]},$$

where $0<p<1$, $n=1,2,\ldots,$ $k=0,1,2,\ldots,n,$ $a-$ any number but for $a<0$ we assume $-an\le\min(p,q),$ $q=1-p.$

To facilitate the notation of that expansible formula we usually make use of the so called factorial polynomials of the degree r in respect of x (also known as generalized power of the degree r of the number x) in the following way

$$x^{[0,a]}=1,\quad x^{[r,a]}=x^{[r-1,a]}.\left[x-\left(r-1\right)a\right],$$

where r=1,2,…, a – denotes any number.

It follows from the given recurrence definition that

$$x^{[r,a]}=x\left(x-a\right)\left(x-2a\right)\ldots\left[x-\left(r-1\right)a\right].$$

Basing on these formulae we can denote the Pólya distribution as follows

$$P\left(X=k\right)=\binom{n}{k}\frac{p^{[k,-a]}}{1^{[n,-a]}}.q^{[n-k,-a]}$$

G. Mühlbach denoted that distribution with slightly different symbols

$$q_{n,x}\left(x,a\right)=\binom{n}{k}\frac{\varphi_k\left(x,a\right)\varphi_{n-k}\left(1-x,a\right)}{\varphi_n\left(1,a\right)},$$

where we can remember that

$$\varphi_n\left(x,a\right)=x^{[k,-a]}.$$

3. G. MÜHLSBACH FORMULA FOR THE MOMENTS

To find a formula for the moments of the Pólya probability distribution G. Mühlbach used an operator $Q_n[f;x,a]$ which is presented as follows

$$Q_n\left[f;x,a\right]=\sum_{l=0}^{n}\binom{n}{l}\Delta^l f\left(x_{n,l}\right)q_{n,l}\left(x,a\right),$$

where $\Delta^l f\left(x_{n,k}\right)$ denotes a finite difference of degree l of an function $f(x_{n,k})$ with $x_{n,k} \in [0,1]$ defined as follows

$$\Delta^0 f\left(x_{n,k}\right) = f\left(x_{n,k}\right),$$

$$\Delta^{l+1} f(x_{n,k}) = \Delta^l f\left(x_{n,k+1}\right) - \Delta^l f\left(x_{n,k}\right),\ l = 0,1,2,\ldots,$$

while

$$q_{n,l}\left(x,a\right) = P\left(x = l\right),\ \text{as given previously.}$$

On the basis of that operator the author has obtained the following formula for the moments

$$m_r = Q_n\left[g_r;x,a\right] = \sum_{l=0}^{n}\binom{n}{l} g_r\left(t_{n,l}\right).q_{n,l}\left(x,a\right),$$

where

$$g_r\left(t_{n,l}\right) = \left(nt_{n,l}\right)^r, \quad t_{n,l} = \frac{l}{n},$$

what we can also denote in the form

$$m_r = \sum_{l=0}^{n}\frac{\binom{n}{l} g_r\left(t_{n,l}\right)(\varphi_l\left(x,a\right))}{\varphi_l\left(1,a\right)},$$

or in better known symbols

$$m_r = \frac{\sum_{l=0}^{n}\binom{n}{l}\Delta^l g_r\left(t_{n,l}\right)x^{[l,-a]}}{1^{[l,-a]}},$$

where at this notation used $x = p$.

4. A MODIFICATION OF THE MÜHLBACH FORMULA

The author did not get an insight into the mathematical efficacy of the given formula. That formula can be presented in a simpler, more handy and easier form by use of the Stirling numbers of the second kind which are defined (see: J.Łukasiewicz and M. Warmus (1956), p. 50) as the coefficients at the factorial polynomials in the identity

$$x^r = S_0^r + S_1^r x^{[1]} + S_2^r x^{[2]} + \ldots S_r^r x^{[r]} = \sum_{l=0}^{r} S_l^r x^{[l]} ,$$

where $x^{[r]} = x(x-1)(x-2)...(x-(r-1))$ and getting

$$S_0^0 = 1, \quad S_0^r = 0 \text{ for } r = 1,2,...,$$

$$S_r^r = 1 \text{ for } r = 1,2,...,$$

$$S_k^r = 0 \text{ for } r < k,$$

and using the finite difference of zero, i.e. the difference of the function $y = x^k$ at the point x= 0 with the step 1, that is

$$\Delta 0^k = 1^k - 0^k, \qquad \Delta^2 0^k = \Delta\left(1^k - 0^k\right) = 2^k - 2.1^k + 0^k$$

and so on.

We know the following formula which is used for preparing the tables of the finite difference of zero

$$\Delta^l 0^{k+1} = l\left(\Delta^l 0^k + \Delta^{l-1} 0^k\right), \qquad l \le k.$$

Taking into account

$$\binom{n}{l} = \frac{n^{[l]}}{l!}$$

and the fact that the difference of the degree l in the formula for the moments is calculated at the point zero we obtain the notation of the formula in the form

$$m_r = \sum_{l=0}^{n} \frac{n^{[l]}}{l!} \Delta^l 0^r \, \frac{x^{[l,-a]}}{1^{[l,-a]}} .$$

The relation between the finite differences of zero and Stirling numbers of the second kind is as follows:

$$\frac{\Delta^l 0^r}{l!} = S_l^r$$

Therefore, the formula for the moments can be written in the final form

$$m_r = \sum_{l=0}^{n} n^{[l]} S_l^r \frac{x^{[l,-a]}}{1^{[l,-a]}} .$$

The Stirling numbers are tabulated, for example at A. Kaufmann (1968), p. 52, what allows to calculate the moment of the needed degree efficiently enough . For example

$$m_1 = \sum_{l=0}^{n} n^{[l]} S_l^1 \frac{x^{[l,-a]}}{1^{[l,-a]}} = nx$$

$$m_2 = nx + n(n-1) \cdot \frac{x(x+a)}{1+a}$$

$$m_3 = nx + \frac{3n(n-1)(x(x+a))}{1+a} + n(n-1)(n-2)\frac{x(x+a)(x+2a)}{(1+a)(1+2a)}$$

$$m_4 = nx + 7n(n-1)\frac{x(x+a)}{1+a} + 6n(n-1)(n-2)\frac{x(x+a)(x+2a)}{(1+a)(1+2a)} +$$

$$+ n(n-1)(n-2)(n-3)\frac{x(x+a)(x+2a)(x+3a)}{(1+a)(1+2a)(1+3a)}$$

5. A RECURRENCE RELATION FOR THE MOMENTS

In the academic book of the probability theory by T. Gerstenkorn and T. Śródka (1972) a recurrence relation for the moments (formula (6.5.11), p. 227)is given, together with a proof in the Pólya random urn scheme that is when N is a number of balls in the urn, b is a number of white balls in the urn, c is a number of black balls in the urn, $b + c = N$, s is a number of the putting or taking out of the balls of a given colour according to the colour of the ball drawn by then and turned to the urn. The formula is of the form

$$m_{r+1} = \frac{1}{N+rs} \sum_{i=0}^{r} \left(bn\binom{r}{i} - (b-sn)\binom{r}{i+1} - s\binom{r}{i+2} \right) m_{r-i} ,$$

where r= 0,1,2,…, n is the number of realized experiences (draws). In the case s<0 it should be assumed: $-ks \le b$ and $-(n-k)s \le c,\ k = 0,1,2,\ldots n.$

In the Pólya scheme we ask about the probability of obtaining k white balls in n draws. If we regard the known in that scheme relation: $\frac{b}{N} = p,\ \frac{c}{N} = q,\ \frac{s}{N} = a$, we obtain a comfortable form of the formula for the moments of the Pólya distribution

$$m_{r+1} = \frac{1}{1+ra}\sum_{i=0}^{r}\left(np\binom{r}{i} - (p-an)\binom{r}{i+1} - a\binom{r}{i+2}\right)m_{r-i}\,.$$

From that formula for the moments of the Pólya distribution one can easily obtain as special cases the formulae for the moments of the distributions: binomial (Bernoulli), hypergeometrical, negative binomial and, in the limit case, also for the Poisson distribution.

REFERENCES

Gerstenkorn T., Śródka T. (1972), *Kombinatoryka i rachunek prawdopodobieństwa*, PWN, Warszawa.

Kaufmann A. (1968), *Introduction à la Combinatorique en Vue des Applications*, Paris, Dunod.

Łukasiewicz J., Warmus M. (1956), *Metody numeryczne i graficzne*, Część I, Warszawa, PWN.

Mühlbach G. (1972), *Rekursionsformeln für die zentralen Momente der Pólya- und der Beta-Verteilung*, Metrika 19, Fasc. 2–3, 171–177.

Tadeusz Gerstenkorn

UWAGI O WZORZE NA MOMENTY ROZKŁADU PRAWDOPODOBIEŃSTWA G. PÓLYI

Streszczenie. Rozkład prawdopodobieństwa zmiennej losowej może być scharakteryzowany przez podanie pewnych liczb zwanych parametrami rozkładu. Do najczęściej używanych parametrów należą momenty. Uwaga nasza jest skoncentrowana na rozkładzie prawdopodobieństwa G. Pólyi, bowiem można z niego łatwo uzyskać jako przypadki szczególne, lub odpowiednio graniczne, ważne w statystyce rozkłady jak dwumianowy, ujemny dwumianowy lub Poissona.

W 1972 r. G. Mühlbach podał interesujące wzory na momenty rozkładu Pólyi. Autor ten nie wnikał w ocenę efektywności rachunkowej podanego wzoru na momenty zwykłe. Pokażemy, co ma znaczenie praktyczne, że wzór ten można przedstawić w prostszej, wygodnej formie.

Słowa kluczowe: rozkład prawdopodobieństwa Pólyi, momenty rozkładu.