UWAGI O WZORZE NA MOMENTY ROZKŁADU PRAWDOPODOBIEŃSTWA PÓLYI

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 13(19)

Tadeusz Gerstenkorn

Emerytowany profesor Uniwersytetu Łódzkiego

ISSN 1644-6739
e-ISSN 2449-9765

DOI: 10.15611/sps.2015.13.09

Streszczenie: Rozkład prawdopodobieństwa zmiennej losowej może być scharakteryzowany przez podanie pewnych liczb zwanych parametrami rozkładu. Do najczęściej używanych parametrów należą momenty. W pracy skoncentrujemy się na rozkładzie Pólyi, bowiem można z niego łatwo uzyskać jako przypadki szczególne lub odpowiednio graniczne ważne w statystyce rozkłady, takie 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 Polyi, moment zwykły.

1. Wstęp

Zmienna losowa jest zasadniczo wystarczająco dokładnie opisana przez jej rozkład prawdopodobieństwa. Względy praktyczne dyktują jednak potrzebę znalezienia charakterystyk liczbowych rozkładu, ponieważ są to opisy krótkie i umożliwiają szybkie porównanie rozkładów ze sobą.

W statystyce teoretycznej, a także w statystyce użytkowanej w ekonomii zachodzi często potrzeba uchwycenia zasadniczych własności badanej zbiorowości. Abstrahujemy wtedy od wielu szczegółów, a własności, o których podkreślenie nam chodzi, charakteryzujemy niejednokrotnie też za pomocą jednej lub kilku liczb. Należą do nich w pierwszym rzędzie średnia arytmetyczna i odchylenie przeciętne, a przypadku rozkładu momenty.

2. Rozkład G. Pólyi

W niniejszej pracy analizujemy wzór na momenty rozkładu G. Pólyi podany w 1972 r. przez G. Mühlbacha. W tym celu przypomnimy, że rozkład ten wyraża się wzorem

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 13(19)

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

gdzie n = 1,2,…, k = 01,2,…,n, a – dowolna liczba, przy czym dla $a<0$ zakładamy

$$-an \le \min(p,q)\,, q=1-p.$$

W celu ułatwienia zapisu tego dość rozciągliwego wzoru, posługujemy się zwykle tzw. wielomianami czynnikowymi stopnia r względem x (nazywanymi także uogólnioną r-tą potęgą liczby x) w sposób następujący

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

gdzie r =1,2,…, zaś a oznacza dowolną liczbę.

Z podanego tu określenia rekurencyjnego wynika, że

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

W oparciu o powyższe wzory rozkład Pólyi można zapisać następująco

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

G. Mühlbach zapisywał ten rozkład w nieco innej symbolice

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

co sprowadza się do zapisu

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

3. Wzór na momenty G. Mühlbacha

Dla znalezienia wzoru na momenty zwykłe w rozkładzie Pólyi, G. Mühlbach posłużył się operatorem $Q_n[f;x,a]$, który przedstawia się następująco

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

gdzie $\Delta^l f(x_{n,k})$ oznacza różnicę rzędu l określoną następująco

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

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

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 13(19)

natomiast

$$q_{n,l}(x,a) = P(x = l), \text{ jak uprzednio podano.}$$

W oparciu o podany operator autor uzyskał następujący wzór na momenty

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

gdzie

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

co można także zapisać w postaci

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

lub w bardziej znanej symbolice

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

gdzie przy tej stosowanej notacji $x = p$.

4. Modyfikacja wzoru Mühlbacha

Autor nie wnikał w ocenę efektywności rachunkowej podanego wzoru. Wzór ten można przedstawić w prostszej i wygodniejszej do obliczeń formie przy pomocy liczb Stirlinga S_l^r drugiego rodzaju, które określamy jako współczynniki przy wielomianach czynnikowych w tożsamości

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

przyjmując

$$S_0^0 = 1, S_0^r = 0 \text{ dla } r = 1,2,\ldots,$$

$$S_r^r = 1 \text{ dla } r = 1,2,\ldots,$$

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

oraz korzystając z różnic skończonych zera, to jest różnicy funkcji $y = x^k$ w punkcie x=0 z krokiem 1, tzn.

$$\Delta 0^k = 1^k - 0^k, \Delta^2 0^k = \Delta(1^k - 0^k) = 2^k - 2.\,1^k + 0^k \text{ itd.}$$

Zachodzi następujący wzór rekurencyjny, który jest wykorzystywany przy układaniu tablic różnic skończonych zera

$$\Delta^l 0^{k+1} = l(\Delta^l 0^k + \Delta^{l-1} 0^k), l \leq k.$$

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 13(19)

Uwzględniając, że

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

oraz fakt, że różnica rzędu l we wzorze na momenty jest liczona w punkcie zero, otrzymujemy zapis tego wzoru w postaci

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

Zachodzi następujący związek między różnicami skończonymi zera a liczbami Stirlinga drugiego rodzaju

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

więc wzór na momenty można zapisać ostatecznie w postaci

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

Liczby Stirlinga są stablicowane, np. przez Kaufmanna [1968], więc pozwala to dość sprawnie wyliczyć moment potrzebnego rzędu. Na przykład

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

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

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

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

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

5. Wzór rekurencyjny na momenty

W podręczniku z rachunku prawdopodobieństwa Gerstenkorn i Śródka [1972] podany jest wzór rekurencyjny na momenty (wzór 6.5.11, s. 227) wraz z dowodem w schemacie Pólyi losowego pobierania kul z urny, tzn. gdy N – liczba kul w urnie, b – liczba białych kul w urnie, c – liczba czarnych kul w urnie, b+c=N, s – liczba dodawanych lub wyjmowanych kul z urny danego koloru w zależności od koloru uprzednio wylosowanej i zwróconej kuli do urny. Wzór ten jest postaci

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

gdzie r=0,1,2,…, n – liczba przeprowadzonych doświadczeń (losowań). W przypadku $s < 0$ należy przyjąć założenie $-ks \leq b\ i - (n-k)s \leq c, k = 0,1,2, \dots n.$

W schemacie Pólyi pytamy o prawdopodobieństwo otrzymania k kul białych na n losowań.

Jeśli uwzględnimy znaną w tym schemacie zależność: $\frac{b}{N} = p, \frac{c}{N} = q, \frac{s}{N} = a$, to otrzymamy wygodną formę wzoru dla rozkładu Pólyi

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

Z podanych tu wzorów na moment rozkładu Pólyi uzyskuje się łatwo jako przypadki szczególne wzory na momenty rozkładów dwumianowego (Bernoulliego), hipergeometrycznego, ujemnego dwumianowego, a w przypadku granicznym także dla rozkładu Poissona.

Literatura

Gerstenkorn T., Środka T., *Kombinatoryka i rachunek prawdopodobieństwa*, PWN, Warszawa 1972.

Kaufmann A., *Introduction a la Combinatorique en Vue des Applications,*Paris, Dunod 1968.

Mühlbach G., *Rekursionsformeln fur die zentralen Momente der Pólya- und der Beta--Verteilung*, Metrika 19, 1972, vol. 2–3, s. 171–177.

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

Summary: The probability distribution of a random variable can be characterized by some numbers called parameters of the distribution. The most commonly used parameters are the moments. Our attention is concentrated on the Pólya distribution because it is easily possible to obtain from it some special cases very important in the statistics distributions such as binomial, negative binomial and Poisson (in the limit procedure). In 1972 G. Mühlbach introduced very interesting formulae for the moments of the Pólya distribution. The author did not investigate an appreciation of the numerical efficacy of the formula for the simple moments. We will show that it is possible to demonstrate this formula in a simpler form. It has a practical significance and importance.

Keywords: the Pólya distribution, moments.