

Tadeusz Gerstenkorn

Wyższa Szkoła Kupiecka w Łodzi
 Wyższa Szkoła Kultury Społecznej i Medialnej w Toruniu

UWAGI O ROZKŁADACH INFLACYJNYCH

1. Zarys problematyki rozkładów inflacyjnych

Dawno już zauważono, że klasyczne typowe rozkłady prawdopodobieństwa nie zawsze są idealnymi modelami przy rozpatrywaniu różnorodnych zagadnień ekonomicznych lub biologicznych. Szukano więc pewnych modyfikacji znanych rozkładów. Już dość wcześnie interesujące propozycje zaprezentowali A.C. Cohen i A.C. Cohen Jr. [1; 2], biorąc pod uwagę rozkłady: dwumianowy, Poissona, geometryczny i ujemny dwumianowy. Dla przykładu przytoczę tu jego propozycję odnoszącą się do rozkładu geometrycznego:

$$P_x = \begin{cases} (1-\alpha q)p, & x=0 & 0 \leq q < 1 \\ (1-\alpha q)pq, & x=1 & 0 \leq \alpha < 1 \\ pq^x, & x=2, 3, \dots & p=1-q. \end{cases}$$

W 1963 r. S.N. Singh zaproponował rozpatrywanie w badaniach statystycznych rozkładu Poissona z pewnym zniekształceniem (nadmiarem wartości) w początkowej wartości i dlatego nazwanym inflacyjnym [13]. M.P. Singh wskazał na to, że istnieją analogiczne sytuacje opisywane w zasadzie dobrze przez rozkład dwumianowy z wyłączeniem zera (lub innego punktu), w którym następuje zniekształcenie teoretycznie spodziewanej wartości prawdopodobieństwa [14]. Dokładnie mówiąc, w punkcie tym zauważa się większą częstość obserwowanego zjawiska, niżby należało oczekiwać na podstawie przyjętego probabilistycznego dwumianowego modelu (inflacja częstości). Omawiany autor podał przykład takiej sytuacji. Badano populację rodzin składających się z 4 osób. Zbiór rozpatrywanych rodzin składał się z dwóch grup: 1 – rodzin narażonych, 2 – nie narażonych na niebezpieczeństwo zapadnięcia na pewną chorobę w danej miejscowości. Notowano liczby przypadków zachorowań w rodzinach obu grup stanowiących statystycznie mieszaninę. Liczby te wynosiły: 0-4. Okazało się, że rozpatrywana tu sytuacja była dobrze opisana przez rozkład dwumianowy z wyjątkiem liczby zachorowań przy-

bierającej wartość 0 (tzn. braku zachorowania), natomiast przypadek był dobrze modelowany przez rozkład dwumianowy inflacyjny. W przypadku niezachorowania żadnego członka rodziny otrzymano większą częstość, niżby to wynikało ze zwykłego rozkładu dwumianowego.

Inflacyjny rozkład dwumianowy określamy następująco:

$$P(X = k) = \begin{cases} \beta + \alpha \binom{n}{l} p^l q^{n-l} & \text{dla } k = l \\ \alpha \binom{n}{k} p^k q^{n-k} & \text{dla } k = 0, 1, 2, \dots, l-1, l+1, \dots, n, \end{cases}$$

gdzie: $0 < \alpha \leq 1$, $\alpha + \beta = 1$, $0 < p < 1$, $p + q = 1$.

Rozkład ten był nazwany uogólnionym, gdyż zniekształcenie mogło nastąpić w dowolnym punkcie rozkładu. Jeśli następowało w początkowym, to nie używano tego miana.

W analogiczny sposób definiuje się inne rozkłady inflacyjne. Łatwo zauważyć, iż rozkład inflacyjny jest szczególnym przypadkiem mieszaniny rozkładów $P_1(X = k)$ i $P_2(X = k)$, z których pierwszy jest tzw. rozkładem zdegenerowanym, tj. postaci:

$$P(X = k) = \begin{cases} 1, & \text{gdy } k = l \\ 0, & \text{gdy } k \neq l, \end{cases}$$

a drugi dowolnym rozkładem dyskretnym. Mieszaniną rozkładów P_1 oraz P_2 określamy bowiem taki rozkład $P(X = k)$, który przy współczynniku α udziału rozkładu P_2 w mieszaninie (często nieznanym i podlegającym estymacji) wyraża się wzorem

$$P(X = k) = (1 - \alpha)P_1(X = k) + \alpha P_2(X = k).$$

Widać stąd, że badania dotyczące rozkładów inflacyjnych są szczególnym przypadkiem badań mieszanin rozkładów.

Znaczenie praktyczne tych badań jest więc uzasadnione znaczeniem mieszanin w problematyce statystycznej.

Zasadniczymi problemami rozpatrywanymi w badaniu rozkładów inflacyjnych są:

a) estymacja współczynnika α (frakcji rozkładu niezdegenerowanego w mieszaninie),

b) estymacja parametru rozkładu (np. p lub q w przypadku rozkładu dwumianowego),

c) obliczenie momentów lub innych charakterystyk, a także ich estymatorów.

W 1973 r. L. Grzegórska w pracy [12]:

1) omówiła sumy rozkładów inflacyjnych, oraz w odniesieniu do rozkładu typu szeregu potęgowego, tj.

$$P(X = k) = p(x; \theta) = \frac{a(x)\theta^x}{f(\theta)}, \quad x \in T,$$

gdzie $f(\theta) = \sum a(x)\theta^x$ i sumowanie rozciąga się na pewien podzbiór T liczb całkowitych, $a(x) \geq 0$ i $\theta \in \Omega = \{\theta : 0 < \theta < R\}$ jest przestrzenią parametru, R jest promieniem zbieżności szeregu potęgowego $f(\theta)$,

2) omówiła momenty centralne rzędu $(r+1)$ tegoż rozkładu inflacyjnego;

3) ustaliła związek między momentami rozkładu inflacyjnego i nieinflacyjnego.

W pracy [9] podane są związki rekurencyjne na momenty zwykłe i centralne rozkładów inflacyjnych dwumianowego i Poissona. Wzory na momenty uzyskano korzystając bądź to z funkcji charakterystycznej odpowiedniego rozkładu, bądź to przez różniczkowanie wzoru definicyjnego na momenty względem parametru rozkładu (p lub λ). Również inne własności inflacyjnego rozkładu dwumianowego podane są w [16]. Wyniki autora niniejszego artykułu zawarte są w pracach [3-11].

2. Momenty inflacyjnego rozkładu Pólyi

Między momentami zwykłymi m_r określonego rozkładu a momentami $\bar{m}_r$ rozkładu inflacyjnego zachodzi łatwy do wykazania związek

$$\bar{m}_r = \beta l^r + \alpha m_r.$$

Wykorzystując ten wzór, można otrzymać moment dowolnego rzędu rozkładu inflacyjnego, znając moment rozkładu zwykłego. Pokażemy to na przykładzie rozkładu Pólyi, z tego rozkładu bowiem można otrzymać jako przypadki szczególne inne znane, podstawowe rozkłady.

Wzór rekurencyjny na momenty zwykłe m_{r+1} , w przypadku gdy zmienna losowa X podlega rozkładowi Pólyi, jest następujący:

$$P(X = k) = \binom{n}{k} \frac{b(b+s) \dots [b+(k-1)s] \cdot c(c+s) \dots [c+(n-k-1)s]}{N(N+s) \dots [N+(n-1)s]},$$

gdzie b , c , s są to parametry schematu urnowego typu Pólyi (b – liczba białych kul, c – czarnych, s – dodawanych i odejmowanych danego typu po każdym losowaniu), n – liczba losowań, N – liczba kul.

m_{r+1} jest określone następująco [17]:

$$m_{r+1} = \frac{1}{N+rs} \sum_{i=0}^r [bn \binom{r}{i} - (bn - sn) \binom{r}{i+1} - s \binom{r}{i+2}] m_{r-i}, \quad r = 0, 1, 2, \dots$$

Jeśli wprowadzimy oznaczenia:

$$\frac{b}{N} = p, \quad \frac{c}{N} = q, \quad \frac{S}{n} = a,$$

to otrzymamy inną postać wzoru zarówno na prawdopodobieństwo, jak i na momenty:

$$m_{r+1} = \frac{1}{1+rs} \sum_{i=0}^r [pn \binom{r}{i} - (p-an) \binom{r}{i+1} - a \binom{r}{i+2}] \cdot m_{r-i}.$$

Wzór ten umożliwia otrzymanie momentu zwykłego dowolnego rzędu w kolejnym postępowaniu. Mamy np.:

$$m_1 = np, \quad m_2 = \frac{np}{1+a} (np + an + q),$$

$$m_3 = \frac{np}{(1+a)(1+2a)} [p^2(n-1)(n-2) + 3p(n-1)(an-1) + (2an+1)(an+1)].$$

Wypisanie kolejnych wzorów $\bar{m}_r$ rozkładu inflacyjnego nie sprawia już żadnych trudności.

Między momentami centralnymi $\bar{\mu}_r$ a μ_r , odpowiednio rozkładu inflacyjnego i zwykłego zachodzi związek:

$$\bar{\mu}_r = \beta(l - \bar{m}_1)^r + \alpha \sum_{i=0}^r \binom{r}{i} \beta^{r-i} (m_1 - l)^{r-i} \mu_i,$$

także łatwy do uzasadnienia, choć nieco bardziej rachunkowo skomplikowany.

Wzór rekurencyjny na momenty centralne w rozkładzie Pólyi jest następujący [3]:

$$\mu_r = np \left[\left(E + q \frac{1+na}{1+a} \right)^{r-1} \mu_0^* - \mu_{r-1} \right]$$

dla $r = 2, 3, \dots$, gdzie E jest operatorem przesunięcia, tj. $E\mu_r = \mu_{r+1}$, $\mu_0^* = 1$, $\mu_1^* = 0$, a $E^i \mu_0 = \mu_i$ i ogólnie: μ_r^* oznacza moment centralny rozkładu Pólyi z pomniejszoną o jeden liczbą losowań, tj.

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

$k = 0, 1, 2, \dots, n-1$.

Obliczając odpowiednie momenty gwiazdkowe zastępujemy n przez $n - 1$, p przez $p + a$ oraz 1 przez $1 + a$.

Obliczając kolejno według podanych wzorów, otrzymujemy:

$$\bar{\mu}_2 = \alpha \left[\beta(l - np)^2 + npq \frac{1 + na}{1 + a} \right],$$

$$\begin{aligned} \bar{\mu}_3 = \alpha & \left\{ \beta(\alpha^2 - \beta^2)(l - np)^3 + 3\beta(np - l)npq \frac{1 + na}{1 + a} + \right. \\ & + np \left[\frac{(n - 1)(p + a)}{1 + a} \left(1 + \frac{(n - 2)(p + 2a)}{1 + 2a} - \frac{(n - 1)(p + a)}{1 + a} \right) + \right. \\ & \left. \left. + q \frac{1 + na}{1 + a} \left(q \frac{1 + na}{1 + a} - np \right) \right] \right\}. \end{aligned}$$

Wiele informacji dotyczących zwłaszcza danych bibliograficznych odnoszących się do rozkładów inflacyjnych lub innych zmodyfikowanych podano w [18].

Literatura

- [1] Cohen A.C., *Estimation in the Poisson Distribution when Values of $c+1$ Are Sometimes Erroneously Reported as c* , „Annals of the Institute of Statistical Mathematics” (Tokyo) 1959, 11, 189-193.
- [2] Cohen A.C., *Estimating the Parameters of a Modified Poisson Distribution*, „Journal of American Statistical Association” 1960, 55, 139-143.
- [3] Gerstenkorn T., *The Recurrence Relations for the Moments of the Discrete Probability Distributions*, „Dissertationes Mathematicae” z. 83, 1971, PWN, Warszawa.
- [4] Gerstenkorn T., *Jednowymiarowe rozkłady dyskretne ze zniekształceniem*, [w:] *Metody statystyczne w sterowaniu jakością*, red. S. Firkowicz, Ossolineum, Wrocław 1977.
- [5] Gerstenkorn T., *Moment Recurrence Relations for the Inflated Negative Binomial, Poisson and Geometric Distributions*, „Demonstratio Mathematica” 1979, 12 (2), 389-410.
- [6] Gerstenkorn T., *On the Recurrence Relation for the Moments About an Arbitrary Point of a Class of Discrete Inflated Distributions*, „Annales Societatis Mathematicae Polonae, Seria I: Commentationes Mathematicae” 1979, 21(2), 291-298.
- [7] Gerstenkorn T., *On Some Statistical Problem Joining the Pólya Inflated Distribution to Bayes' Rule*, „Fasciculi Mathematici” 1990, 19(236), 77-85.
- [8] Gerstenkorn T., *The Gini's Mean Difference of an Inflated Discrete Distribution*, [w:] *16th International Conference on Multivariate Statistical Analysis MSA'97*, November 27-29, 1997, p. 147-151.
- [9] Gerstenkorn T., Gerstenkorn J., *Die Gini-Mitteldifferenz in der statistischen Praxis. Anwendung auf einige inflationistische Verteilungen*, Acta Universitatis Lodziensis, Folia Oeconomica 2003, 164, p. 15-28.
- [10] Gerstenkorn T., Gerstenkorn J., *Gini's Mean Difference in the Theory and Applications to Inflated Distributions*, „Statistica” 2003, 63(3), 469-488.

- [11] Gerstenkorn T., *A Compound of an Inflated Pascal Distribution with the Poisson One*, Acta Universitatis Lodzianis, Folia Oeconomica 2005, 194, 21-32.
- [12] Grzegórska L., *Distribution of Sums of The So-called Inflated Distributions*, Annales Universitatis Mariae Curie-Skłodowska, Lublin 1973, 27(4), 31-42.
- [13] Singh S.N., *A Note on Inflated Poisson Distribution*, „Journal of the Indian Statistical Association” 1963, 1(3), 140-144.
- [14] Singh M.P., *Inflated Binomial Distribution*, „Journal of Scientific Research of the Banares Hindu University” (1965/66), 16(1), 87-90.
- [15] Sobich L., *On the Recurrence Relation for the Moments of Inflated Binomial and Inflated Poisson Distribution*, Annales Societatis Mathematicae Polonae, Seria I: Commentationes Mathematicae 1974, 17(2), 459-468.
- [16] Sobich L., Szynal D., *Some Properties of Inflated Binomial Distribution*, „Canadian Mathematical Bulletin” 1974, 17, 609-611.
- [17] Śródka T., *Wzór rekurencyjny na momenty zwykłe w rozkładzie Pólyi*, „Roczniki Polskiego Towarzystwa Matematycznego, Seria I, Prace Matematyczne” 1964, 8(2), 217-220.
- [18] Wimmer G., Altmann G., *Thesaurus of Discrete Probability Distributions*, Stamm, Essen 1999.

REMARKS ON THE INFLATED DISTRIBUTIONS

Summary

This is a review paper on inflated distributions. Relations between moments of ordinary and inflated Pólya distributions are particularly discussed.