

Разпределение на разликата, модулът ѝ и частното на геометрично разпределени случайни величини*

Павлина Йорданова

Abstract: In this paper we obtain distributions, mean values, variances and probability generating functions of module of difference, product and fraction of independent geometrically distributed random variables. In the case of identically distributed random variables we show that the difference has two-tailed and it's modulo has zero-modified geometric distribution. We use method of moments to obtain consistent estimators for the unknown parameters of these distributions. By stochastic machinery we find explicit expressions for series indexed by ordered pairs of mutually prime numbers.

Key words: Geometric distribution, independence, functions of random variables.

ВЪВЕДЕНИЕ

Понятието "геометрично разпределение" изглежда е въведено за първи път през 1943 г. от N. Arley [3]. От тогава до сега то е изследвано от много автори и е доказало своето широко приложение. Тук получаваме разпределението на някои функции от такива величини.

Дефиниция 1.1. Случайната величина ξ е геометрично разпределена с вероятност за успех $p \in (0, 1)$ ако има следния ред на разпределение:

$$P(\xi = k) = (1 - p)^k p, \quad k = 0, 1, \dots$$

Едни от първите автори, които правят характеристика на геометричното разпределение са: 1965 Т. S. Ferguson [7], 1974 г. Dallas [6], 1977 г., К. М. Chong [5], 1980г. В. С. Arnold [4].

През 1984г. М. Ahsanullah and В. Holland [2] изследват рекордните стойности в редица от геометрично разпределени случайни величини. През 1991г. А. Adatia [1] пресмята дисперсията и ковариациите на техните порядкови статистики.

Мястото на геометричното разпределение сред другите дискретни разпределения е добре описано например от Johnson N. L. et al, 1992 [8]. Заедно с експоненциалното разпределение то притежава свойството "липса на последствие".

В тази статия намираме разпределението на разликата, модулът на разликата и частното на независими и геометрично разпределени случайни величини. Поради ограничение в обема тук не са дадени доказателствата на твърденията. Показано е, че при еднакво разпределени случайни величини разликата има двустранно геометрично разпределение, а модулът на разликата има модифицирано в нулата геометрично разпределение. Виж [9].

Дефиниция 1.2. Случайната величина ξ има двустранно геометрично разпределение с вероятност за успех $p \in (0, 1)$ (накратко $\xi \sim TTGeom(p)$) ако има следния ред на разпределение:

$$P(X_1 - X_2 = i) = q^{|i|} \frac{p}{1 + q}, \quad i = 0, \pm 1, \pm 2, \dots$$

Дефиниция 1.3. Случайната величина ξ има модифицирано в нулата геометрично разпределение с вероятност за успех при първия опит $p_0 \in (0, 1)$ и с вероятност за успех $p \in (0, 1)$ при всеки следващ опит (накратко $\xi_0 \sim Geom_0(p_0, p)$) ако има следния ред на разпределение:

$$P(\xi_0 = 0) = p_0, \\ P(\xi_0 = k) = (1 - p_0)(1 - p)^{k-1} p, \quad k = 1, 2, \dots$$

* Статията е финансирана от НИП № РД-05-285/11.03.2009

Навсякъде по-долу $p_i \in (0, 1)$ и $q_i = 1 - p_i$, $i = 1, 2$.

ОСНОВНИ РЕЗУЛТАТИ

1. Разпределение на разликата на геометрично разпределени случайни величини
Теорема 2.1. Ако X_1 и X_2 са независими геометрично разпределени случайни величини с параметри съответно $p_1 \in (0, 1)$ и $p_2 \in (0, 1)$, т.е. $X_i \sim \text{Geom}(p_i)$, $i = 1, 2$, то

$$P(X_1 - X_2 = i) = q_2^{|i|} \frac{p_1 p_2}{1 - q_1 q_2}, \quad i = 0, \pm 1, \pm 2, \dots$$

$$\text{ii) } E(X_1 - X_2) = \frac{p_2 - p_1}{p_1 p_2};$$

$$\text{iii) } D(X_1 - X_2) = \frac{q_1}{p_1^2} + \frac{q_2}{p_2^2};$$

$$\text{iv) } \text{mod}(X_1 - X_2) = 0;$$

$$\text{v) } E(z^{(X_1 - X_2)}) = \frac{p_1 p_2}{(1 - q_1 z)(1 - q_2 z^{-1})}.$$

Забележка: Това разпределение е удачно да се нарече модифицирано в нулата двустранно геометрично.

Следствие 2.2. Ако X_1 и X_2 са независими еднакво геометрично разпределени случайни величини с параметър $p \in (0, 1)$ и ако $q = 1 - p$, то

i) $\xi \sim \text{TTGeom}(p)$, т.е.

$$P(X_1 - X_2 = i) = q^{|i|} \frac{p}{1 + q}, \quad i = 0, \pm 1, \pm 2, \dots$$

$$\text{ii) } E(X_1 - X_2) = 0;$$

$$\text{iii) } D(X_1 - X_2) = \frac{2q}{p^2};$$

$$\text{iv) } \text{mod}(X_1 - X_2) = 0;$$

$$\text{v) } E(z^{(X_1 - X_2)}) = \frac{p^2 z}{(1 - zq)(z - q)}.$$

2. Разпределение на модула на разликата на геометрично разпределени случайни величини

Теорема 2.3. Ако X_1 и X_2 са независими геометрично разпределени случайни величини с параметри съответно $p_1 \in (0, 1)$ и $p_2 \in (0, 1)$, т.е. $X_i \sim \text{Geom}(p_i)$, $i = 1, 2$, то

$$P(|X_1 - X_2| = i) = \frac{p_1 p_2}{1 - q_1 q_2} (q_1^i + q_2^i), \quad i = 0, 1, 2, \dots$$

$$\text{ii) } E|X_1 - X_2| = \frac{q_1 p_2^2 + q_2 p_1^2}{p_1 p_2 (1 - q_1 q_2)};$$

$$\text{iii) } D|X_1 - X_2| = \frac{q_1}{p_1^2} + \frac{q_2}{p_2^2} + \frac{(p_2 - p_1)^2}{p_1^2 p_2^2} - \frac{(q_1 p_2^2 + q_2 p_1^2)^2}{p_1^2 p_2^2 (1 - q_1 q_2)^2};$$

$$\text{iv) } \text{mod}|X_1 - X_2| = \begin{cases} 0, & q_1 + q_2 \leq 1; \\ 1, & q_1 + q_2 > 1; \end{cases}$$

$$\text{v) } E(z^{|X_1 - X_2|}) = \frac{p_1 p_2}{1 - q_1 q_2} \left(\frac{1}{1 - zq_1} + \frac{1}{1 - zq_2} \right).$$

Следствие 2.4. Ако X_1 и X_2 са независими еднакво геометрично разпределени случайни величини с параметър $p \in (0, 1)$ и $q = 1 - p$, то

$$i) |X_1 - X_2| \sim \text{Geom}_0\left(\frac{2q}{1+q}, p\right), \text{ т.е.}$$

$$P(|X_1 - X_2| = i) = \begin{cases} \frac{2pq^i}{1+q}, & i = 1, 2, \dots \\ \frac{p}{1+q}, & i = 0 \end{cases};$$

$$ii) E|X_1 - X_2| = \frac{2q}{p(1+q)};$$

$$iii) D|X_1 - X_2| = \frac{2q(1+q^2)}{p^2(1+q)^2};$$

$$iv) \text{mod}|X_1 - X_2| = \begin{cases} 0, & q \leq 0.5 \\ 1, & q > 0.5 \end{cases}$$

$$v) E(z^{|X_1 - X_2|}) = \frac{2p}{(1+q)(1-zq)}.$$

3. Разпределение на частното на геометрично разпределени случайни величини.

Теорема 2.5. Ако X_1 и X_2 са независими геометрично разпределени случайни величини с параметри съответно $p_1 \in (0, 1)$ и $p_2 \in (0, 1)$, т.е. $X_i \sim \text{Geom}(p_i)$, $i = 1, 2$, то

$$P\left(\frac{X_1}{X_2} = \frac{i}{j} \mid X_2 \neq 0\right) = \begin{cases} \frac{p_1 p_2 q_1^i q_2^j}{q_2(1 - q_1^i q_2^j)}, & (i, j) \in \Delta, \\ p_1, & i = 0 \end{cases},$$

където Δ е множеството на наредените двойки от взаимно прости положителни числа.

$$ii) E\left(\frac{X_1}{X_2} \mid X_2 \neq 0\right) = \frac{q_1 p_2}{q_2 p_1} (-\ln p_2);$$

$$iii) D\left(\frac{X_1}{X_2} \mid X_2 \neq 0\right) = \frac{q_1 p_2}{p_1 q_2} Li_2(q_2) - \left(\frac{q_1 p_2}{p_1 q_2}\right)^2 (\ln p_2)^2, \text{ където } Li_2(z) = \sum_{j=1}^{\infty} \frac{z^j}{j^2}, \quad |z| < 1 \text{ е}$$

известната като dilogarithm или Spence функция;

$$iv) \text{mod}\left(\frac{X_1}{X_2} \mid X_2 \neq 0\right) = 0;$$

$$v) E(z^{X_1 \cdot X_2} \mid X_2 \neq 0) = p_1 p_2 \sum_{j=1}^{\infty} \frac{q_2^{j-1}}{1 - q_1^j \sqrt{z}}.$$

Нататък с Δ ще означаваме множеството на наредените двойки от взаимно прости положителни числа.

От свойствата на вероятностните мерки получаваме следното следствие.

Следствие 2.6.
$$\sum_{i,j \in \Delta} \frac{q_1^i q_2^j}{1 - q_1^i q_2^j} = \frac{q_1 q_2}{(1 - q_1)(1 - q_2)}.$$

В частност за $q_1 = q_2 = q \in (0, 1)$,

$$\sum_{i,j \in \Delta} \frac{q^{i+j}}{1 - q^{i+j}} = \frac{q^2}{(1 - q)^2}.$$

Следващото следствие се получава от дефиницията за математическо очакване.

Следствие 2.7.
$$\sum_{i,j \in \Delta} \frac{i}{j} \cdot \frac{q_1^i q_2^j}{1 - q_1^i q_2^j} = \frac{-q_1 \ln(1 - q_2)}{(1 - q_1)^2}.$$

В частност за $q_1 = q_2 = q \in (0, 1)$,

$$\sum_{i,j \in \Delta} \frac{i}{j} \cdot \frac{q^{i+j}}{1-q^{i+j}} = \frac{-q \ln(1-q)}{(1-q)^2}.$$

От изразяването на трансформацията на Лаплас-Стилтес по два различни начина получаваме следващото следствие.

Следствие 2.8.

$$\sum_{i,j \in \Delta} \frac{i}{j} \cdot \frac{q_1^i q_2^{j-1}}{1-q_1^i q_2^j} = \sum_{j=1}^{\infty} \frac{q_2^{j-1}}{1-q_1 \sqrt[j]{z}}.$$

Следствие 2.9. Ако X_1 и X_2 са независими еднакво геометрично разпределени случайни величини с параметър $p \in (0, 1)$ и $q = 1 - p$, то

i)

$$P\left(\frac{X_1}{X_2} = \frac{i}{j} / X_2 \neq 0\right) = \begin{cases} \frac{p^2 q^{i+j-1}}{1-q^{i+j}}, & (i, j) \in \Delta. \\ p, & i = 0 \end{cases}$$

ii) $E\left(\frac{X_1}{X_2} / X_2 \neq 0\right) = -\ln p;$

iii) $D\left(\frac{X_1}{X_2} / X_2 \neq 0\right) = \sum_{j=1}^{\infty} \frac{q^j}{j^2} - (\ln p)^2.$

4. Оценки на неизвестните параметри.

Ще използваме методът на моментите.

4.1. Нека разполагаме с проста извадка $Y_1, Y_2, \dots, Y_n$ от наблюдения върху *разликите* на две независими геометрично разпределени случайни величини ξ_1 и ξ_2 с параметри съответно $p_1 \in (0, 1)$ и $p_2 \in (0, 1)$.

От Теорема 2.1 получаваме, че $\hat{p}_1 = \frac{c}{1 + \hat{m}_1 c}$ и $\hat{p}_2 = c$, където

$$c = \frac{-1 + \sqrt{1 + (2 + \hat{m}_1)(\hat{m}_1 + \hat{S}_n^2)}}{\hat{m}_1 + \hat{S}_n^2},$$

$\hat{m}_1$ е средното аритметично на наблюденията $Y_1, Y_2, \dots, Y_n$, а $\hat{S}_n^2$ е тяхната емпирична дисперсия, т.е.

$$\hat{S}_n^2 = \frac{\sum_{i=1}^n (Y_i - \hat{m}_1)^2}{n-1}.$$

В случая на равни параметри, т.е. при $p_1 = p_2 = p$ от Следствие 2.2. получаваме, че $E(\xi_1 - \xi_2) = 0$. В горните формули заместваем $\hat{m}_1$ с 0 и стигаме до

$$\hat{p} = \frac{-1 + \sqrt{1 + 2\hat{S}_n^2}}{\hat{S}_n^2}.$$

4.2. Нека $Y_1, Y_2, \dots, Y_n$ е проста извадка от наблюдения върху *модулите на разликите* на две независими еднакво геометрично разпределени случайни величини ξ_1 и ξ_2 с параметър $p \in (0, 1)$.

От Следствие 2.4 получаваме следната оценка за неизвестния параметър p

$$\hat{p} = 1 - \frac{\hat{m}_2 - \sqrt{\hat{m}_2^2 - \hat{m}_1^4}}{\hat{m}_1^2}.$$

Тъй като са получени по метода на моментите, описаните по-горе оценки са слабо състоятелни.

ЗАКЛЮЧЕНИЕ

От резултатите в теоремите се вижда, че в случая на съвпадащи параметри само разликата на независими геометрично разпределени случайни величини е симетрична. Наличието на повече параметри в останалите случаи позволява да се постигне до асиметрия на разпределението и по-точна оценка на наблюдаваните величини.

Не е трудно да се намерят двойките от съвместни разпределения на $\min(X_1, X_2)$, $X_1 - X_2$, $|X_1 - X_2|$ и $\frac{X_1}{X_2}$ и да се забележи, че тези случайни величини са зависими две по две.

До сега в литературата изглежда няма метод за изследване на суми, чиито индекси са наредени двойки от взаимно прости числа. Тук такива са сумите в Следствия 2.6, 2.7 и 2.8. Стохастиката и по-точно, разпределението от Теорема 2.5 ни дават един възможен подход за намирането им. При работа с други разпределения на случайни величини, с разпределението на частното им и числовите му характеристики, бихме могли да получим други подобни суми.

ЛИТЕРАТУРА

- [1] Adatia, A. Computation of variances and covariances of order statistics from the geometric distribution. Journal of Statistical Computation and Simulation, 1991, 39, 91-94.
- [2] Ahsanullah, M. and Holland B. Record values and geometric distribution. Statistische hefte, 25, 319-327.
- [3] Arley, N. On the Theory of Stochastic Processes and Their Application to the Theory of Cosmic Radiation. Copenhagen: Cads Forlag, 1943.
- [4] Arnold, B. C. Two Characterizations of the Geometric Distribution. Journal of Applied Probability, 1980, 17, 570-573.
- [5] Chong, K. M. On Characterization of the Exponential and Geometric Distributions by Expectations. Journal of the American Statistical Association, 1977, 72, 160-161.
- [6] Dallas, A. C. A Characterization of the Geometric Distribution. Journal of Applied Probability, 1974, 11, 609-611.
- [7] Ferguson, T. S. A Characterization of the Geometric Distribution. American Mathematical Monthly, 1965, 72, 256-260.
- [8] Johnson, N. L., Adrienne, W., K., Kotz, S. Univariate Discrete Distributions. John Wiley Sons, 1992.
- [9] O'Neal, J. Predictive quantizing differential pulse code modulation for the transmission of television signals, Bell. Syst. Tech.J., 45, 689 - 722, 1966.

За контакти:

Гл. ас. д-р Павлина Йорданова, Катедра "Икономика и моделиране", Шуменски университет "Епископ Константин Преславски", тел.: 054-830 495, вѣтр.138, е-mail: pavlina_kj@abv.bg

Докладът е рецензиран.