УДК 519.2 Вестник СПбГУ. Математика. Механика. Астрономия. 2022. Т. 9 (67). Вып. 2
MSC 60E05, 62E10, 62G32

О некоторых вероятностных распределениях, связанных с классической схемой Бернулли

С. М. Ананьевский, В. Б. Невзоров

Санкт-Петербургский государственный университет,
Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7–9

Для цитирования: *Ананьевский С. М., Невзоров В. Б.* О некоторых вероятностных распределениях, связанных с классической схемой Бернулли // Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия. 2022. Т. 9 (67). Вып. 2. С. 201–208. https://doi.org/10.21638/spbu01.2022.202

Классическая схема независимых испытаний Бернулли уже более трех столетий, начиная с работ Якова Бернулли, представляет одну из самых популярных тем в теории вероятностей. Она идеально подходит для постановки и решения различных практических задач. Получено множество результатов, связанных с модификациями этой схемы, но появляются новые ситуации, новые проблемы, которые требуют дальнейших продвижений в изучении разных случайных величин, связанных в той или иной степени с независимыми бернуллиевскими испытаниями. Решения некоторых таких задач представлены в данной статье.

Ключевые слова: схема Бернулли, биномиальное распределение, геометрическое распределение, производящие функции.

1. Рассмотрим последовательность независимых случайных величин (с. в.) $X_1, X_2, \ldots$, принимающих значение 1 с некоторой вероятностью $0 < p < 1$ и значение 0 с вероятностью $q = 1 - p$.

Часто событие $\{X_n = 1\}$ трактуется как «успех в n-м испытании», а событие $\{X_n = 0\}$ формулируется как «неудача».

Пусть μ_n обозначает число успехов в первых n попытках-испытаниях. Распределение этого числа имеет $B(n, p)$-биномиальное распределение с вероятностями

$$P_r = P\{\mu_n = r\} = C_n^r p^r q^{n-r}, \qquad r = 0, 1, \ldots, n-1, n.$$

Математические ожидания даются равенствами

$$E\mu_n = np, \qquad n = 1, 2, \ldots .$$

Со схемой Бернулли тесно связаны геометрические $\mathrm{Geom}(q)$-распределения с параметром $0 < q < 1$. Если проводить испытания до появления среди них первого «успешного», то получаем, что число неудач имеет геометрическое распределение, поскольку

$$P\{X_1 = 0, X_2 = 0, \ldots, X_n = 0, X_{n+1} = 1\} = pq^n, \qquad n = 0, 1, 2, \ldots,$$

и математическое ожидание числа неудачных попыток равно q/p.

Если интересует момент ν, когда в последовательности $X_1, X_2, \ldots$ обнаружены оба исхода, то получаем

$$P\{\nu = n\} = pq^{n-1} + qp^{n-1}, \qquad n = 2, 3, \ldots.$$

Рассмотрим случайную величину $W(k)$ — число неудач до момента появления k-го по счету успеха. В этом случае имеем дело с суммой, состоящей из k независимых геометрически $\mathrm{Geom}(q)$-распределенных случайных слагаемых. Эта сумма имеет отрицательное биномиальное распределение с вероятностями

$$P\{W(k) = n\} = C_{n+k-1}^{n} p^k q^n, \qquad n = 0, 1, 2, \ldots,$$

и математическим ожиданием

$$EW(k) = kq/p, \qquad k = 1, 2, \ldots.$$

Производящая функция с. в. $W(k)$ имеет вид

$$Es^{W(k)} = pqs^k/(1-qs)^k, \qquad k = 1, 2, \ldots.$$

2. Рассмотрим следующую ситуацию. Проводятся испытания не до k-го успеха, а до момента появления первой серии в последовательности $X_1, X_2, \ldots$, содержащей не менее k подряд появляющихся успехов.

В первом томе двухтомника В. Феллера [1, гл. 13, формула (7.6)] приведен вид производящей функции $F(s)$ для

$$f(k, n) = P\{L_n\}$$

— вероятностей того, что первая серия, состоящая из k успешных испытаний, связана с событием

$$L_n = \{X_{n-k} = 0, X_{n-k+1} = 1, \ldots, X_{n-1} = 1, X_n = 1\}.$$

Показано, что

$$F(s) = p^k s^k (1-ps)/(1 - s + qp^k s^{k+1}), \qquad k = 1, 2, \ldots. \tag{1}$$

Если немного изменить условие и рассмотреть число испытаний до момента появления такой серии, то получим, что соответствующая производящая функция имеет вид

$$F_1(s) = p^k (1-ps)/(1 - s + qp^k s^{k+1}), \qquad k = 1, 2, \ldots. \tag{2}$$

Далее будем иметь дело с событиями

$$A_0 = \{X_1 = 0\},$$

$$A_r = \{X_1 = 1, X_2 = 1, \ldots, X_{r-1} = 1, X_r = 1, X_{r+1} = 0\}, \qquad r = 1, 2 \ldots, k-1,$$

$$A_k = \{X_1 = 1, X_2 = 1, \ldots, X_{k-1} = 1, X_k = 1\}.$$

Отметим, что

$$P\{A_r\} = qp^r, \quad r = 0, 1, \ldots, k-1, \qquad \text{и} \qquad P\{A_k\} = p^k.$$

Что можно сказать о числе неудач $W(k)$, если имеет место событие A_r? Понятно, что условное распределение этой величины при условии, что произошло событие A_r, $r = 0, 1, 2, \ldots, k-1$, совпадает с распределением случайной величины $W_1(k) = 1 + W(k)$. Если же произошло событие A_k, то никаких неудачных испытаний не наблюдалось.

Пусть $B_1(k) = EW(k)$. Получаем, что

$$B_1(k) = (1-p^k)(1+EW(k)) = (1-p^k) + (1-p^k)B_1(k).$$

Следовательно,

$$B_1(k) = (1-p^k)/p^k, \qquad k = 1, 2, \ldots.$$

Пусть теперь

$$P(k,s) = Es^{W(k)}$$

обозначает производящую функцию случайной величины $W(k)$. Учитывая приведенные выше возможные варианты, получаем, что справедливы равенства

$$P(k,s) = p^k + (1-p^k)Es^{(1+W(k))} = p^k + (1-p^k)sP(k,s)$$

и

$$P(k,s) = p^k/(1-(1-p^k)s),$$

т. е. величина $W(k)$ имеет геометрическое $\mathrm{Geom}(1-p^k)$-распределение с вероятностями

$$P\{W(k) = n\} = p^k(1-p^k)^n, \qquad n = 0, 1, 2, \ldots.$$

В частности,

$$P\{W(k) = 0\} = p^k$$

и

$$P\{W(k) = 1\} = p^k(1-p^k).$$

Перейдем к случайной величине $R(k)$ — числу успехов до появления первой интересующей нас серии, содержащей, как минимум, k подряд идущих успешных испытаний. Справедливо равенство $R(k) = 0$, если имеет место событие A_k. Если имеем дело с событием A_r, $r = 0, 1, \ldots, k-1$, то условное распределение $R(k)$ совпадает с распределением суммы $R(k) + r$. Для математического ожидания $B_2(k) = ER(k)$ получаем соотношение

$$B_2(k) = \sum_{r=0}^{k-1} (ER(k) + r)\, p^r q = (1-p^k)B_2(k) + (1 - kp^{k-1} + (k-1)p^k)/q^2.$$

Отсюда следует, что

$$B_2(k) = (1 - kp^{k-1} + (k-1)p^k)/p^{k-1}q.$$

Для производящей функции $R(k,s)$ случайной величины $R(k)$ справедливо равенство

$$R(k,s) = qR(k,s) + \sum_{r=1}^{k-1} s^r R(k,s) p^r q + p^k,$$

из которого получаем, что

$$R(k,s) = p^{k-1}(1-ps)/(1-s+qp^{k-1}s^k). \tag{3}$$

При $k = 1$ имеем естественное соотношение

$$R(1,s) = 1,$$

что соответствует вырожденному в нуле распределению. Если $k = 2$, то

$$R(2,s) = p(1-ps)/(1-s+qps^2) = p/(1-qs),$$

т. е. получаем геометрическое распределение с вероятностями

$$P\{R(2) = n\} = pq^n, \qquad n = 0,1,2\dots.$$

Рассмотрим теперь случайную величину $M(k)$ — число серий успехов до момента появления первой серии из k или большего числа последовательных успешных испытаний.

Вернемся к событиям $A_0, A_1, \dots, A_k$. Если имеет место событие A_0, то условное распределение случайной величины $M(k)$ совпадает с ее безусловным распределением. В случае появления события A_k получаем, что $M(k) = 0$. Во всех остальных ситуациях условное распределение случайной величины $M(k)$ при условии, что имеет место событие A_r, $\quad r = 1,2,\dots,k-1$, совпадает с безусловным распределением величины $M(k)+1$. Для математических ожиданий $a(k) = EM(k)$ получаем, что

$$\begin{aligned} a(k) &= a(k)P(A_0) + (1+a(k))(P(A_1)+P(A_2)+\ldots+P(A_{k-1})) = \\ &= qa(k) + (1+a(k))(qp+qp^2+qp^{k-1}) = qa(k) + p(1-p^{k-1}) + a(k)p(1-p^{k-1}) = \\ &= a(k) - p^k a(k) + p(1-p^{k-1}), \end{aligned}$$

откуда следует, что

$$a(k) = (1-p^{k-1})/p^{k-1}. \tag{4}$$

Если $k = 1$, то естественно получаем, что $a(1) = 0$. Если $k = 2$, то $a(2) = q/p$. Если $k = 3$, то $a(3) = (1-p^2)/p^2$.

Отметим, что распределение числа серий успехов $M(k)$ не зависит от того, является ли первое испытание успешным или неудачным.

Если рассмотреть теперь число $N(k)$ серий неудач до появления группы из k последовательных успехов, то здесь нужно выделить две возможные ситуации. Если первое испытание было успешным, то $N(k) = M(k)$. Иначе получаем, что $N(k) = M(k)+1$. Для математических ожиданий $b(k) = EN(k)$ приходим к равенству

$$b(k) = pB_2(k) + q(1+B_2(k)) = q + B_2(k) = (q^2p^{k-1} + 1 - kp^{k-1} + (k-1)p^k)/p^{k-1}q. \tag{5}$$

Если рассмотреть также суммарное число $S(k)$ серий из успехов и серий из неудач до появления k подряд идущих успешных испытаний, то распределение этой величины совпадает со смесью с весами p и q распределений случайных величин $2M(k)$ и $2M(k)+1$.

Найдем вид производящих функций $M(k,s)$, $N(k,s)$ и $S(k,s)$ случайных величин $M(k)$, $N(k)$ и $S(k)$. Отметим, что справедливы соотношения

$$N(k,s) = pM(k,s) + qsM(k,s) = (p+qs)M(k,s) \tag{6}$$

и

$$S(k,s) = pM(k,s^2) + qsM(k,s^2) = (p+qs)M(k,s^2). \qquad (7)$$

Возвращаясь к событиям $A_0, A_1, \ldots, A_k$, можем написать следующее равенство для функции $M(k,s)$:

$$M(k,s) = qM(k,s) + (pq + p^2q + \ldots + p^{k-1}q)sM(k,s) + p^k = qM(k,s) + (p - p^k)sM(k,s) + p^k,$$

из которого следует, что

$$M(k,s) = p^{k-1}/(1 - s + sp^{k-1}). \qquad (8)$$

Из (8) получаем, в частности, что

$$P\{M(k) = n\} = p^{k-1}(1 - p^{k-1})^n, \qquad n = 0, 1, 2, \ldots. \qquad (9)$$

Получаем также, учитывая равенства (6) и (7), что

$$N(k,s) = (p+qs)p^{k-1}/(1 - s + sp^{k-1}) \qquad (10)$$

и

$$S(k,s) = (p+qs)p^{k-1}/(1 - s^2 + p^{k-1}s^2). \qquad (11)$$

В [1, гл. 13] представлен следующий результат (задача 24).

Пусть $q_{n,l} = P\{Q(n) = l\}$ — вероятность иметь в n испытаниях Бернулли ровно l серий успехов длины r. Тогда производящая функция $Q_n(s) = Es^{Q(n)}$ является коэффициентом при s^n в выражении

$$(1 - p^r s^r)/[1 - s + qp^r s^{r+1} - (1 - ps)p^r s^r q].$$

Приведем рассуждения, позволяющие найти соответствующие распределения для числа $V(k,r)$ серий, состоящих ровно из $r < k$ успехов, в представленной нами схеме. В этой ситуации событию A_r соответствует случайная величина $V(k,r) + 1$. Если имеет место событие A_k, то таких серий нет. Во всех остальных ситуациях имеем дело с условным распределением, совпадающим с распределением случайной величины $V(k,r)$. Для производящей функции

$$V(k,r,s) = Es^{V(k,r)}$$

справедливо равенство

$$V(k,r,s) = p^k + p^r qsV(k,s) + (1 - p^r q - p^k)V(k,r,s),$$

из которого следует, что

$$V(k,r,s) = p^{k-r}/(p^{k-r} + q - sq). \qquad (12)$$

Для $r = 1$ получаем, в частности, что

$$V(k,1,s) = p^{k-1}/(p^{k-1} + q - sq). \qquad (13)$$

Из равенства (13) следует, в частности, что

$$P\{V(k,1) = n\} = q^n p^{k-1}/(q + p^{k-1})^{n+1}, \quad n = 0, 1, 2, \ldots. \qquad (14)$$

Отметим также, что выражения для производящих функций

$$V(k,k-1,s) = p/(1-sq), \quad k = 2,3,\ldots,$$

не зависят от k. В этих случаях получаем, что

$$P\{V(k,k-1) = n\} = pq^n, \quad n = 0,1,2,\ldots,$$

т. е. имеем дело с геометрическим распределением.

3. Продолжим рассматривать последовательность независимых случайных величин $X_1, X_2, \ldots$, принимающих значение 1 с некоторой вероятностью $0 < p < 1$ и значение 0 с вероятностью $q = 1 - p$, до появления первой серии из k успехов. Выше описали распределения числа успехов и числа неудач, числа серий успехов и числа серий неудач в последовательности до момента появления k подряд идущих успешных испытаний. Рассмотрели также распределение числа серий успехов определенной длины $r = 1, 2, \ldots, k-1$, предшествующих появлению первой серии из k успехов.

Вернемся к рассмотрению серий неудач, наблюдаемых до момента появления k подряд фиксируемых успехов. Отметим, что такого рода серии уже могут состоять из любого числа последовательных неудач. В равенстве (10) было дано выражение для производящей функции случайной величины $N(k)$ — числа серий неудач. Теперь мы можем рассмотреть серии из таких неудач фиксированной длины $r = 1, 2, \ldots$. Пусть $W(r,k)$ обозначает их число.

Рассмотрим два возможных случая.

В ситуациях, когда $X_1 = 0$ и $X_1 = 1$, число серий неудач длины r будем обозначать как $W_1(r)$ и $W_2(r)$ соответственно. Потребуются также соответствующие производящие функции $W(r,k,s)$, $W_1(r,s)$ и $W_2(r,s)$ случайных величин $W(r,k)$, $W_1(r)$ и $W_2(r)$. Отметим, что

$$W(r,k) = qW_1(r) + pW_2(r). \tag{15}$$

Пусть $X_1 = 0$. Тогда

$$P\{W(r,k) = m | X_1 = 0\} = P\{W_1(r) = m\}. \tag{16}$$

Если $X_1 = 1$, то

$$P\{W(r,k) = m | X_1 = 1\} = P\{W_2(r) = m\}. \tag{17}$$

Получаем, что

$$P\{W(r,k) = m\} = qP\{W_1(r) = m\} + pP\{W_1(r) = m\}. \tag{18}$$

Рассмотрим следующий набор исходных событий:

$$A_j = \{X_1 = 0, X_2 = 0, \ldots, X_{j-1} = 0, X_j = 1\}, \quad j = 1, 2, \ldots, r-1, r, r+1 \ldots.$$

Если имеет место событие A_{r+1}, которому соответствует условная вероятность

$$P\{A_{r+1} | X_1 = 0\} = q^{r-1}p,$$

то распределение случайной величины $W_1(r)$ совпадает с распределением случайной величины $1 + W_2(r)$. В остальных случаях распределение $W_1(r)$ совпадает с распределением $W_2(r)$.

Пусть теперь $X_1 = 1$. Рассмотрим события

$$B_j = \{X_1 = 1, X_2 = 1, \dots, X_{j-1} = 1, X_j = 0\}, \qquad j = 2, 3, \dots, k,$$

$$B_{k+1} = \{X_1 = 1, X_2 = 1, \dots, X_k = 1\}.$$

В ситуациях, когда имеет место любое из событий $B_2, B_3, \dots, B_k$, условное распределение $W_2(r)$ совпадает с распределением случайной величины $W_1(r)$. В случае B_{k+1} получаем, что $W_2(r) = 0$.

Для первой из этих двух ситуаций получаем соотношение для производящих функций вида

$$W_1(r,s) = pq^{r-1}sW_2(r,s) + (1 - pq^{r-1})W_2(r,s). \tag{19}$$

Во втором случае приходим к равенству

$$W_2(r,s) = (1 - p^{k-1})W_1(r,s) + p^{k-1}. \tag{20}$$

Из двух последних соотношений, связывающих две производящие функции, следуют равенства

$$W_1(r,s) = \frac{p^{k-1}(1 + pq^{r-1}s - pq^{r-1})}{p^{k-1} + (pq^{r-1} - p^k q^{r-1})(1-s)}, \tag{21}$$

$$W_2(r,s) = \frac{p^{k-1}}{p^{k-1} + pq^{r-1}(1-s) - p^k q^{r-1}(1-s)} \tag{22}$$

и

$$W(r,k,s) = qW_1(r,s) + pW_2(r,s) = \frac{qp^{k-1} + p^k q^r s - p^k q^r + p^k}{p^{k-1} + (pq^{r-1} - p^k q^{r-1})(1-s)}. \tag{23}$$

В частном случае при $k = 2, r = 1$ получаем производящую функцию в виде

$$W(1,2,s) = \frac{1 - pq + pqs}{1 + q - qs}.$$

С другими задачами, связанными со схемой независимых испытаний Бернулли, можно ознакомиться в работах [2, 3].

Авторы благодарят рецензентов за замеченные мелкие ошибки в рукописи, исправление которых улучшило качество работы.

Литература

1. Феллер В. *Введение в теорию вероятностей и ее приложения*, пер. с англ. Т. 1. Москва, Мир (1984).

2. Balakrishnan N., Nevzorov V. B. *A primer on statistical distributions*. Hoboken, New Jersey, John Wiley & Sons (2003).

3. Иванов О. А. Схема Бернулли и обобщения чисел Фибоначчи. *Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия* **1 (59)**, вып. 4, 529–532 (2014).

Статья поступила в редакцию 13 сентября 2021 г.;
доработана 22 октября 2021 г.;
рекомендована к печати 2 декабря 2021 г.

К о н т а к т н а я   и н ф о р м а ц и я:

Ананьевский Сергей Михайлович — канд. физ.-мат. наук, доц.; ananjevskii@mail.ru
Невзоров Валерий Борисович — д-р физ.-мат. наук, проф.; vanev@mail.ru

On some probability distributions, related to the classical Bernoulli scheme

S. M. Ananjevskii, V. B. Nevzorov

St Petersburg State University, 7–9, Universitetskaya nab., St Petersburg, 199034, Russian Federation

For citation: Ananjevskii S. M., Nevzorov V. B. On some probability distributions, related to the classical Bernoulli scheme. *Vestnik of Saint Petersburg University. Mathematics. Mechanics. Astronomy*, 2022, vol. 9 (67), issue 2, pp. 201–208. https://doi.org/10.21638/spbu01.2022.202 (In Russian)

For more than three centuries, the classical scheme of independent Bernoulli trials, starting with the work of Jacob Bernoulli, has been one of the most popular topics in probability theory. It is ideal for setting and solving various practical problems. Many results have been obtained related to modifications of this scheme, but new situations and new problems appear that require further progress in the study of various random variables associated to one degree or another with independent Bernoulli trials. Solutions to some of these problems are presented in this paper.

Keywords: Bernoulli scheme, binomial distribution, geometric distribution, generating functions.

References

1. Feller W. *An introduction o probability theory and its applications.* Vol. 1. New York, John Wiley & Sons (1968). [Rus. ed.: Feller W. *Vvedenie v teoriju verojatnostej i ee prilozhenija.* Moscow, Mir Publ. (1984)].

2. Balakrishnan N., Nevzorov V. B. *A primer on statistical distributions.* Hoboken, New Jersey, John Wiley & Sons (2003).

3. Ivanov O. A. Bernoulli scheme and generalizations of Fibonacci numbers. *Vestnik of Saint Petersburg University. Mathematics. Mechanics. Astronomy* **1 (59)**, iss. 4, 529–532 (2014). (In Russian)

Received: September 13, 2021
Revised: October 22, 2021
Accepted: December 2, 2021

A u t h o r s’   i n f o r m a t i o n:

Sergey M. Ananjevskii — ananjevskii@mail.ru
Valery B. Nevzorov — vanev@mail.ru