Записки научных
семинаров ПОМИ
Том 515, 2022 г.

С. М. Ананьевский, В. Б. Невзоров

О СЕРИЯХ УСПЕХОВ И НЕУДАЧ В СХЕМАХ БЕРНУЛЛИ

Введение

Одним из классических направлений в теории вероятностей являются исследования, связанные с так называемыми бернуллиевскими последовательностями случайных величин (с.в.). Речь идет о последовательностях независимых с.в. $X_1, X_2, \dots,$ принимающих значение 1 с некоторой вероятностью $p, \quad 0 < p < 1$, и значение 0 с вероятностью $q = 1 - p$.

Обычно событие $\{X_n = 1\}$ трактуется как "успех в n-ом испытании", а его дополнение – событие $\{X_n = 0\}$ – как "неудача в этом испытании".

С этой схемой тесно связаны, например, геометрическое распределение числа испытаний (числа неудач) до первого успеха и биномиальные распределения числа неудач или числа успехов, появляющихся в результате проведения фиксированного числа n независимых испытаний. Несмотря на множество работ, в которых решаются различные проблемы для схем Бернулли и для некоторых обобщений этих схем, появляются все новые и новые задачи в этой области, представляющие несомненный интерес с теоретической и практической точек зрения. Ряд таких задач рассмотрим ниже.

Речь пойдет о взаимоотношениях между сериями успехов и сериями неудач в последовательностях бернуллиевских величин. Будут приведены обобщения некоторых результатов, полученных ранее в работе [1].

С сериями успехов конкретной длины $k, \quad k = 1, 2, \dots,$ связаны события вида

$$\{X_1 = 1, X_2 = 1, \dots, X_k = 1, X_{k+1} = 0\}$$

и

$$\{X_r = 0, X_{r+1} = 1, \dots, X_{r+k} = 1, X_{r+k+1} = 0\}, r = 1, 2, \dots.$$

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

Будем проводить независимые испытания до момента образования первой серии успехов, длина которой не меньше k. В [2] приведен вид производящей функции

$$F(s) = \frac{p^k s^k (1 - ps)}{(1 - s + qp^k s^{k+1})} \tag{1}$$

для числа испытаний, которые нужно провести, чтобы получить группу из k последовательных успехов. Там же ([2], глава 13, задача 24) даны соотношения, позволяющие находить производящие функции $Q_n(s)$ для числа серий успехов определенной длины, наблюдаемых в наборе $X_1, X_2, \ldots, X_n$.

В [1] был получен вид производящей функции $V(k, r, s)$ для числа $V(k, r)$ серий , состоящих ровно из r, $r = 1, 2, \ldots, k-1$, успехов, появляющихся до момента образования первой серии, насчитывающей не менее k успехов. Показано, что

$$V(k, r, s) = \frac{p^{k-r}}{(p^{k-r} + q - sq)}. \tag{2}$$

Здесь можно выделить частный случай ($r = 1$) для числа серий, состоящих ровно из одного успеха:

$$V(k, 1, s) = \frac{p^{k-1}}{(p^{k-1} + q - sq)}. \tag{3}$$

В этой ситуации

$$P\{V(k, 1) = n\} = \frac{q^n p^{k-1}}{(q + p^{k-1})^{n+1}}, \quad n = 0, 1, 2, \ldots. \tag{4}$$

В [1] был также получен вид производящей функции

$$W(k, r, s) = \frac{(qp^{k-1} + p^k q^r s - p^k q^r + p^k)}{(p^{k-1} + pq^{r-1} - p^k q^{r-1}(1 - s))} \tag{5}$$

для числа серий неудач определенной длины r, наблюдаемых до появления первой серии из k успехов.

Продолжим исследования такого рода и приведем некоторые обобщения полученных ранее результатов для числа различных вариантов серий в рассматриваемой схеме.

Пусть $N = N(l, m, k)$ обозначает суммарное число серий успехов, длина которых принимает значения $l, l+1, \ldots, m-1$ или m, где $1 \leqslant l \leqslant m \leqslant k-1$.

Рассмотрим следующую полную группу событий:

$$A_0 = \{X_1 = 0\},$$
$$A_r = \{X_1 = 1, X_2 = 1, \ldots, X_r = 1, X_{r+1} = 0\}, \quad r = 1, 2, \ldots, k-1,$$
$$A_k = \{X_1 = 1, X_2 = 1, \ldots, X_k = 1\}$$

и образуем события

$$B_1 = \bigcup_{r=0}^{l-1} A_r, \qquad B_2 = \bigcup_{r=l}^{m} A_r$$
$$B_3 = \bigcup_{r=m+1}^{k-1} A_r, \qquad B_4 = A_k,$$

вероятности которых, соответственно, равны

$$p_1 = P\{B_1\} = q + qp + \cdots + qp^{l-1} = 1 - p^l,$$
$$p_2 = P\{B_2\} = qp^l + qp^{l+1} + \cdots + qp^m = p^l - p^{m+1},$$
$$p_3 = P\{B_3\} = qp^{m+1} + \cdots + qp^{k-1} = p^{m+1} - p^k,$$
$$p_4 = P\{B_4\} = p^k.$$

Если имеет место любое событие, входящее в B_1 или в B_3, то условное распределение случайной величины N в этом случае совпадает с ее безусловным распределением. В случае события B_4 получаем, что $N = 0$. Наконец, если взять любое из событий, образующих B_2, то в этих ситуациях условное распределение N совпадает с безусловным распределением случайной величины $N + 1$. Собирая эти варианты, приходим к следующему соотношению для производящей функции $P(s) = Es^N$:

$$\begin{aligned} P(s) &= (p_1 + p_3)P(s) + p_2 sP(s) + p_4 \\ &= (1 - p^l - p^k + p^{m+1})P(s) + (p^l - p^{m+1})sP(s) + p^k, \end{aligned}$$

из которого следует, что

$$\begin{aligned} P(s) &= \frac{p^k}{(p^l + p^k - p^{m+1} - p^l s + p^{m+1} s)} \\ &= \frac{p^{k-l}}{(1 + p^{k-l} - p^{m+1-l} + (p^{m+1-l} - 1)s)}. \end{aligned} \tag{6}$$

Из равенства (6) получаем, что

$$P\{N(l,m,k)=n\}=\frac{p^{k-l}(1-p^{m+1-l})^n}{(1+p^{k-l}-p^{m+1-l})^{n+1}}, \qquad n=0,1,2,\dots. \tag{7}$$

Приведем несколько частных случаев формул (6) и (7).

1) Пусть $l=1,\ m=k-1$. Тогда $N(1,k-1)$ представляет собой суммарное число всех серий успехов, фиксируемых до момента образования первой группы из k подряд идущих успешных испытаний. Получаем, что

$$Es^{N(1,k-1)}=\frac{p^{k-1}}{(1-(1-p^{k-1})s)}, \tag{8}$$

$$P\{N(1,k-1)=n\}=p^{k-1}(1-p^{k-1})^n, \quad n=0,1,2,\dots,k=2,3,\dots, \tag{9}$$

и

$$EN(1,k-1)=\frac{(1-p^{k-1})}{p^{k-1}}. \tag{10}$$

2) Пусть $1\leqslant l=m<k$. В этом случае $N(m,m)$ – число серий успехов фиксированной длины m, получаемых до первой серии из k успехов. Из (6) и (7) следуют результаты, полученные в [1]:

$$Es^{N(m,m)}=\frac{p^{k-m}}{(1+p^{k-m}-p-(1-p)s)}, \tag{11}$$

что соответствует равенству (2), и

$$P\{N(m,m)=n\}=\frac{p^{k-m}(1-p)^n}{(1-p+p^{k-m})^{n+1}}, \qquad n=0,1,2,\dots \tag{12}$$

В частности,

$$P\{N(1,1)=n\}=\frac{p^{k-1}(1-p)^n}{(1-p+p^{k-1})^{n+1}}, \qquad n=0,1,2,\dots \tag{13}$$

и

$$P\{N(k-1,k-1)=n\}=p(1-p)^n, \qquad n=0,1,2,\dots \tag{14}$$

Изучим также некоторые ситуации с числом серий, составленных из последовательных неудач, до появления первой серии из k успехов. Обозначим $W=W(l,m,k)$ – число серий неудач до первой группы из k успехов, длина которых от l до m, и пусть

$$W(s)=W(l,m,k,s)=Es^W$$

будет обозначать производящую функцию этой случайной величины.

Рассмотрим два варианта для W в зависимости от того, какое значение принимает случайная величина X_1.

Пусть условное распределение W при условии, что $X_1 = 1$, совпадает с распределением некоторой случайной величины W_1, а при условии, что $X_1 = 0$, пусть оно совпадает с распределением другой случайной величины W_2. Обозначим $W_1(s)$ и $W_2(s)$ – производящие функции этих двух величин. Интересующая нас производящая функция W(s) связана с этими функциями равенством

$$W(s) = pW_1(s) + (1-p)W_2(s). \tag{15}$$

Начнем с варианта, когда $X_1 = 0$. Рассмотрим набор из следующих событий:

$$\begin{aligned} D_1 &= \{X_2 = 1\}, \\ D_2 &= \{X_2 = 0, X_3 = 1\}, \\ &\dots, \\ D_{l-1} &= \{X_2 = 0, X_3 = 0, \dots, X_{l-1} = 0, X_l = 1\}, \\ D_l &= \{X_2 = 0, X_3 = 0, \dots, X_l = 0, X_{l+1} = 1\}, \\ &\dots, \\ D_m &= \{X_2 = 0, \dots, X_m = 0, X_{m+1=1}\}, \\ D_{m+1} &= \{X_2 = 0, \dots, X_m = 0, X_{m+1} = 0, X_{m+2} = 1\}, \\ &\dots \end{aligned} \tag{16}$$

Получаем, что в ситуациях, когда справедливы события $D_l, D_{l+1}, \dots, D_m$, суммарная вероятность которых равна $q^{l-1} - q^m$, распределение случайной величины W_2 совпадает с распределением случайной величины $1 + W_1$. В остальных ситуациях, имеющих суммарную вероятность $1 - q^{l-1} + q^m$, это распределение и распределение случайной величины W_1 одинаковы.

Пусть теперь $X_1 = 1$. Рассмотрим следующую полную группу событий:

$$\begin{aligned} &\{X_2 = 0\},\ \{X_2 = 1, X_3 = 0\},\ \dots, \\ &\{X_2 = 1, X_3 = 1,\ \dots,\ X_{k-1} = 1, X_k = 0\}, \\ &\{X_2 = 1, X_3 = 1, \dots, X_k = 1\}. \end{aligned}$$

Если имеет место последнее из этих событий, вероятность которого равна p^{k-1}, то $W_1 = 0$. Во всех остальных ситуациях, суммарная вероятность осуществления которых $1 - p^{k-1}$, распределение случайной величины W_1 совпадает с распределением случайной величины W_2. Собирая приведенные варианты, получаем для производящих функций $W_1(s)$ и $W_2(s)$ соотношения

$$W_2(s) = (1 - q^{l-1} + q^m)W_1(s) + (q^{l-1} - q^m)sW_1(s) \tag{17}$$

и

$$W_1(s) = p^{k-1} + (1 - p^{k-1})W_2(s). \tag{18}$$

Решая эту систему из двух уравнений и вспоминая, что

$$W(s) = pW_1(s) + (1 - p)W_2(s),$$

получаем в итоге, что

$$W(s) = W(l, m, k, s) = \frac{p^{k-1}(1 - (q^l - q^{m+1})(1 - s))}{(p^{k-1} + (q^{l-1} - q^m)(1 - p^{k-1})(1 - s))}. \tag{19}$$

Рассмотрим некоторые частные случаи формулы (19).

1) Имеем дело с числом всех серий неудач ($l = 1, m = \infty$) до первой серии из k успехов. Тогда

$$W(1, \infty, k, s) = \frac{p^{k-1}(p + qs)}{(1 - s + sp^{k-1})}. \tag{20}$$

В этом случае

$$EW = \frac{(1 - p^k)}{p^k}. \tag{21}$$

2) Для числа серий неудач, длина которых равна $m \quad (l = m)$, получаем равенство

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

3) Число серий неудач, длина которых не больше $m \quad (l = 1)$:

$$W(1, m, k, s) = \frac{p^{k-1}(1 - q(1 - q^m)(1 - s))}{(p^{k-1} + (1 - q^m)(1 - p^{k-1})(1 - s))}. \tag{23}$$

4) Число серий неудач, длина которых не меньше m:

$$W(m, \infty, k, s) = \frac{p^{k-1}(1 - q^m(1 - s))}{(p^{k-1} + (1 - p^{k-1})q^{m-1}(1 - s))}. \tag{24}$$

Из (10) и (21) получаем, что математическое ожидание суммарного числа $S(k)$ серий из успехов и серий неудач до первой группы, содержащей не менее k последовательных успешных испытаний, имеет вид

$$S(k) = \frac{(1+p-2p^k)}{p^k}. \tag{25}$$

Имея перед глазами равенство (24), можно ответить на вопрос, какова вероятность того, что серия из k подряд идущих успехов будет получена раньше серии, содержащей не менее, чем m неудач? Для этого в (24) достаточно подставить $s = 0$ в выражение для производящей функции $W(m, \infty, k, s)$ и получить, что эта вероятность равна

$$\frac{p^{k-1}(1-q^m)}{p^{k-1}+(1-p^{k-1})q^{m-1}}.$$

Полученные результаты позволяют ответить и на вопросы, связанные с ситуацией, когда наблюдения заканчиваются после появления первой серии, состоящей не из k успехов, а из k последовательных неудач. Достаточно в приведенных выше соотношениях поменять местами вероятности p и q. В частности, получим, что вероятность того, что серия из k неудач будет предшествовать серии из m успехов, равна

$$\frac{q^{k-1}(1-p^m)}{q^{k-1}+(1-q^{k-1})p^{m-1}}.$$

Отсюда следует, например, что вероятность получить серию из k неудач раньше серии из k успехов равна

$$\frac{q^{k-1}(1-p^k)}{q^{k-1}+(1-q^{k-1})p^{k-1}},$$

а вероятность того, что серия из k успехов предшествует комбинации из k неудач , имеет вид

$$\frac{p^{k-1}(1-q^k)}{p^{k-1}+(1-p^{k-1})q^{k-1}}.$$

Естественно, что сумма этих вероятностей равна 1.

Равенство (24) позволяет ответить и на следующий вопрос: какова длина максимальной серии неудач, получаемой до появления первой группы из k успехов?

Обозначим эту случайную величину как $M(k)$. Понятно, что

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

Найдем вероятности

$$P\{M(k) = n\}, \qquad n = 1, 2, \ldots.$$

Получаем, используя равенство (24) для производящей функции $W(m, \infty, k, s)$, что

$$P\{W(m+1, \infty, k) = 0\} = W(m+1, \infty, k, 0) = \frac{p^{k-1}(1-q^{m+1})}{(p^{k-1} + (1-p^{k-1})q^m)}$$

и

$$P\{W(m, \infty, k) = 0\} = W(m, \infty, k, 0) = \frac{p^{k-1}(1-q^m)}{(p^{k-1} + (1-p^{k-1})q^{m-1})}.$$

Отсюда следует, что для вероятности того, что максимальная длина из имеющихся серий неудач равна m, справедливо равенство

$$\begin{aligned} P\{M(k) = m\} &= P\{W(m+1, \infty, k) = 0\} - P\{W(m, \infty, k) = 0\} \\ &= \frac{p^{k-1}(1-q^{m+1})}{(p^{k-1} + (1-p^{k-1})q^m)} - \frac{p^{k-1}(1-q^m)}{(p^{k-1} + (1-p^{k-1})q^{m-1})}. \end{aligned}$$

Меняя местами вероятности p и q, можно получить и распределение длины максимальной серии из успехов до появления первой группы из k неудач.

Список литературы

1. С. М. Ананьевский, В. Б. Невзоров, *О некоторых вероятностных распределениях, связанных с классической схемой Бернулли.* — Вестник Санкт-Петербургского университета, Математика, Механика, Астрономия, т. 9(67), вып. 2 (2022), 201–208.
2. В. Феллер, *Введение в теорию вероятностей и ее приложения*, том 1, Москва, Мир, 1984.

Ananjevskii S. M., Nevzorov V. B. On the series of successes and failures in Bernoulli schemes.

The classic scheme of independent trials by Bernoulli has been one of the most popular topics in probability theory for more than three centuries, starting with the work of Jacob Bernoulli. It is ideal for setting and solving various practical problems. Many results have been obtained related to modifications of this scheme, but new situations, new problems appear that require further progress in the study of various random variables related to one degree or another with independent Bernoulli tests. In this paper, we continue to study some problems related to the series of successes and failures in sequences of Bernoulli random variables. This work is a direct continuation of the authors' article "On some probability distributions related to the classical Bernoulli scheme", published in 2022.

С.-Петербургский
государственный университет,
Университетская наб., 7-9
199034, Санкт-Петербург, Россия
E-mail: ananjevskii@mail.ru
E-mail: valnev@mail.ru

Поступило 25 октября 2022 г.