Записки научных
семинаров ПОМИ
Том 525, 2023 г.

С. М. Ананьевский, В. Б. Невзоров

О ЧИСЛЕ УСПЕХОВ И ИХ СЕРИЙ В БЕРНУЛЛИЕВСКИХ ПОСЛЕДОВАТЕЛЬНОСТЯХ СЛУЧАЙНЫХ ВЕЛИЧИН

Введение

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

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

Эта классическая схема названа в честь Якова Бернулли (Jacob Bernoulli), который рассматривал такие последовательности. Полученные им результаты были опубликованы его племянником Николаем Бернулли (Nicolaus Bernoulli) в 1713 году в Базеле в книге Ars Conjectandi (“искусство предположений”).

С тех пор, хоть и появилось множество результатов для различных построений, основанных на бернуллиевских с.в., возникает необходимость в рассмотрении новых схем для них, в решении новых задач. Ниже рассмотрим взаимоотношения таких двухточечных распределений с рядом других вероятностных законов. Будет приведен небольшой обзор полученных ранее результатов в этой области и добавлены несколько новых. Будут продолжены исследования, начатые в работах [2–4], некоторых возможных вариантов появления успешных и неудачных испытаний и их серий в последовательностях бернуллиевских величин.

Начнем с простейших классических ситуаций.

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

1) Если проводим n независимых испытаний, то суммарное число успехов $N_1(n)$ в этих испытаниях среди первых n, $n = 1, 2, \ldots$, наблюдений имеет $B(n,p)$ – биномиальное распределение с вероятностями

$$\mathbf{P}\{N_1(n) = m\} = C_n^m p^m q^{n-m}, \quad m = 0, 1, 2, \ldots, n,$$

и математическим ожиданием

$$\mathbf{E}\, N_1(n) = np, \quad n = 1, 2, \ldots.$$

2) Рассмотрим число успешных испытаний $N_2(p)$ в последовательности $X_1, X_2, \ldots$ до появления первого неудачного исхода. В этой ситуации, как известно, имеем дело с $\mathrm{Geom}(p)$ – геометрическими распределениями.

Здесь:

$$\mathbf{P}\{N_2(p) = n\} = (1-p)p^n, \quad n = 0, 1, 2, \ldots,$$

и

$$\mathbf{E}\, N_2(p) = p/(1-p).$$

Математическое ожидание общего числа испытаний до появления первой неудачи, включая ее, имеет вид

$$\mathbf{E}\, N_2(p) + 1 = 1/(1-p).$$

3) Число успехов $N_3(r,p)$ до появления r-ой неудачи имеет так называемое отрицательное биномиальное распределение с целым параметром r. В этом случае

$$\mathbf{P}\{N_3(r,p) = n\} = C_n^{n+r-1}(1-p)^r p^n, \quad n = 0, 1, 2, \ldots,$$

и

$$\mathbf{E}\, N_3(r,p) = rp/(1-p), \quad n = 0, 1, 2, \ldots.$$

Естественно, что при $r = 1$ имеем дело с геометрическими распределениями.

Перейдем к нескольким неклассическим схемам.

В работах [2-4] рассматривались последовательности испытаний, которые проводятся до момента появления либо k-го успеха, либо r-ой неудачи.

4) Пусть $N_4(k,r)$, $\quad k = 1, 2, \ldots$, $r = 1, 2, \ldots$, обозначает число успехов, зафиксированных в момент первого появления в последовательности $X_1, X_2, \ldots$ либо k-го успеха, либо r-ой по счету неудачи, а

$$M_4(k,r) = \mathbf{E}\, N_4(k,r), \quad k = 0, 1, 2, \ldots, \quad r = 0, 1, 2, \ldots.$$

В этой схеме

$$M_4(k,0)=0, \quad k=0,1,2,\dots, \quad M_4(0,r)=0, \quad r=0,1,2,\dots,$$

и

$$M_4(1,1)=p.$$

Информация о других таких математических ожиданиях была представлена в виде двойной производящей функции

$$P_4(s,t)=\sum_{k=0}^{\infty}\sum_{r=0}^{\infty}M_4(k,r)s^k t^r=\sum_{k=1}^{\infty}\sum_{r=1}^{\infty}M_4(k,r)s^k t^r,$$

которая имеет вид

$$P_4(s,t)=\frac{pst}{(1-s)(1-t)(1-ps-qt)}. \tag{1}$$

5) Были рассмотрены и числа успехов $N_5(k,r)$ с математическими ожиданиями $M_5(k,r)=\mathbf{E}\,N_5(k,r)$, получаемых в серии испытаний, останавливаемых в первый момент, когда зафиксировано не менее k успехов и не менее r неудач. В этом варианте

$$M_5(k,0)=k, \quad k=0,1,2,\dots, \tag{2}$$

$$M_5(0,r)=rp/(1-p), \tag{3}$$

$$M_5(1,1)=1+p^2/q, \tag{4}$$

а для двойной производящей функции

$$P_5(s,t)=\sum_{k=0}^{\infty}\sum_{r=0}^{\infty}M_5(k,r)s^k t^r$$

было получено равенство

$$P_5(s,t)=\frac{s(1-sp)/(1-s)^2+pt(1-qt)/(1-t)^2+pst/(1-s)(1-t)}{1-st-qt}. \tag{5}$$

Интерес представляют и ситуации, связанные с образованием в последовательности испытаний различных серий из успехов или неудач.

6) Пусть $N_6(k)$ обозначает число успехов до момента, предшествующего появлению серии, содержащей, как минимум, k подряд идущих

успешных испытаний. В работе [2] показано, что соответствуюшая производящая функция $P_6(k,s)$ для математических ожиданий $\mathbf{E}\, N_6(k)$ имеет следующий вид:

$$P_6(k,s) = p^{k-1}(1-ps)/(1-s+qp^{k-1}s^k). \tag{6}$$

Отметим, что

$$\mathbf{E}\, N_6(k) = (1-p^{k-1}+(k-1)p^k)/p^{k-1}, \quad k=1,2,\dots. \tag{7}$$

В той же работе приведен вид производящей функции

$$P_7(l,s) = q^l/(1-(1-q^l)s) \tag{8}$$

для числа успехов $N_7(l)$, фиксируемых до появления серии из l неудач. В этом случае

$$\mathbf{E}\, N_7(l) = (1-q^l)/q^l, \quad l=1,2,\dots. \tag{9}$$

Когда речь идет о числе испытаний, успехов или неудач на момент появления первой из этих двух серий, то здесь важным является следующее представленное в [1], глава VIII, равенство для вероятностей события $A(k,l)$, состоящего в том, что серия из k успехов появляется раньше серии из l неудач:

$$\mathbf{P}\{A(k,l)\} = p^{k-1}(1-q^l)/(p^{k-1}+q^{l-1}-p^{k-1}q^{l-1}). \tag{10}$$

Рассмотрим события, зависящие от моментов появления обеих серий – из k успехов и из l неудач.

Пусть $N(k,l)$ обозначает число успехов, появившихся к моменту, когда образовалась первая из этих двух серий, а $M(k,l) = \mathbf{E}\, N(k,l)$. Обозначим через $M(1,k,l)$ и $M(2,k,l)$ математические ожидания чисел этих успехов в ситуациях, соответственно, когда первое испытание было успешным или неудачным.

Пусть $\{X_1 = 1\}$. При этом условии рассмотрим следующую полную группу событий:

$$A_1 = \{X_2 = 0\}, A_2 = \{X_2 = 1, X_3 = 0\}, A_3 = \{X_2 = 1, X_3 = 1, X_4 = 0\},$$
$$\dots, A_{k-2} = \{X_2 = 1, X_3 = 1, \dots, X_{k-2} = 1, X_{k-1} = 0\},$$
$$A_{k-1} = \{X_2 = 1, X_3 = 1, \dots, X_{k-2} = 1, X_{k-1} = 1, X_k = 0\},$$
$$A_k = \{X_2 = 1, X_3 = 1, \dots, X_{k-2} = 1, X_{k-1} = 1, X_k = 1\} \tag{11}$$

с вероятностями

$$\mathbf{P}\{A_r\} = qp^{r-1}, \quad r = 1,2,\dots,k-1,$$

и

$$\mathbf{P}\{A_k\} = p^{k-1}.$$

В ситуации, когда имеем дело с событием A_r, $r = 1, 2, \dots, k-1$, получаем, что условное математическое ожидание числа успехов до момента появления первой из двух интересующих групп равно $r + M(2, k, l)$. В случае A_k уже сформирована одна серия, состоящая в данной ситуации из k успехов. Рассматривая все эти варианты с их вероятностями, получаем, что

$$\begin{aligned} M(1,k,l) &= (1 + M(2,k,l))q + (2 + M(2,k,l))pq + \dots \\ &\quad + (k-1 + M(2,k,l))p^{k-2}q + kp^{k-1} \\ &= qM(2,k,l)(1 + p + \dots + p^{k-2}) + q(1 + 2p + \dots + (k-1)p^{k-2}) + kp^{k-1} \\ &= M(2,k,l)(1-p^{k-1}) + (1-p^k)/(1-p). \qquad (12) \end{aligned}$$

Пусть теперь $\{X_1 = 0\}$. Здесь, если имеет место событие

$$\{X_2 = 0, X_3 = 0, \dots, X_l = 0\},$$

вероятность которого равна q^{l-1}, то сформирована уже серия без единого успешного испытания. В остальных ситуациях типа

$$\{X_2 = 1\}; \quad \{X_2 = 0,\ X_3 = 0, \dots, X_{r-1} = 0,\ X_r = 1\}, \quad r = 3, 4, \dots, l,$$

интересующее нас математическое ожидание совпадает с рассмотренным выше вариантом, когда отсчет успехов начинался с события $\{X_1 = 1\}$, т.е. оно равно $M(1, k, l)$.

Получаем равенство

$$M(2,k,l) = (1-q^l)M(1,k,l), \qquad (13)$$

и соответствующее выражение для $M(k, l)$ сводится к соотношению

$$M(k,l) = pM(1,k,l) + qM(2,k,l) = (1-q^l)M(1,k,l). \qquad (14)$$

Получаем, что

$$M(1,k,l) = \frac{(1-p^k)}{q(q^{l-1} + p(1-q^{l-1}))}, \qquad (15)$$

и, следовательно,

$$M(k,l) = \frac{(1-p^k)(1-q^l)}{q(q^{l-1} + p^{k-1}(1-q^{l-1}))}. \qquad (16)$$

Отметим, что $M(k,1)$ – среднее число успехов до того момента, когда произойдет первая неудача или k-ый успех, имеет вид

$$M(k,1) = \frac{p(1-p^k)}{q}, \quad k = 1,2,\ldots, \tag{17}$$

а математическое ожидание успешных испытаний, заканчивающихся в момент появления либо первого успеха, либо l-ой неудачи, дается равенством

$$M(1,l) = 1 - q^l, \quad l = 1,2,\ldots. \tag{18}$$

Пусть теперь $N^*(k,l)$ обозначает число успехов, появившихся к моменту, когда уже сформировались обе интересующие нас серии, а

$$M^*(k,l) = \mathbf{E}\, N^*(k,l).$$

Также рассматриваем два варианта. Пусть $M^*(1,k,l)$ и $M^*(2,k,l)$ – математические ожидания чисел этих успехов в ситуациях, когда первое испытание было успешным или неудачным. Начнем с ситуации, когда $\{X_1 = 1\}$, и также рассмотрим полную группу из событий

$$A_1, \quad A_2, \quad \ldots, \quad A_k.$$

Если имеет место какое-то из событий

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

то

$$M^*(1,k,l) = r + M^*(2,k,l). \tag{19}$$

В случае события $A_k = \{X_2 = 1, \ldots, X_k = 1\}$ серия из k успехов уже получена, но какое-то число успешных испытаний может еще добавиться до появления серии из l неудач. Математическое ожидание числа таких дополнительных успехов равно $(1-q^l)/q^l$. Объединяя все эти варианты, получаем, что

$$\begin{aligned} M^*(1,k,l) &= q(1 + M^*(2,k,l)) + pq(2 + M^*(2,k,l)) + \ldots \\ &\quad + p^{k-2}q(k-1 + M^*(2,k,l)) + p^{k-1}\Big(k + \frac{1-q^l}{q^l}\Big) \\ &= \frac{1-p^{k-1}}{q} + \frac{p^{k-1}}{q^l} + (1-p^{k-1})M^*(2,k,l). \end{aligned} \tag{20}$$

Пусть теперь имеет место событие $\{X_1 = 0\}$. Здесь также потребуется рассмотреть полную группу из событий

$$B_1 = \{X_2 = 1\},\ B_2 = \{X_2 = 0,\ X_3 = 1\},\ \ldots,$$
$$B_{l-1} = \{X_2 = 0,\ X_3 = 0,\ \ldots,\ X_{l-1} = 0,\ X_l = 1\},$$
$$B_l = \{X_2 = 0,\ X_3 = 0, \ldots, X_{l-1} = 0,\ X_l = 0\},$$

с вероятностями

$$\mathbf{P}\{B_r\} = pq^r, \quad r = 1, 2, \ldots, l-1, \quad \text{и} \quad \mathbf{P}\{B_l\} = q^{l-1}.$$

Если происходит событие B_l, то группа из неудач уже сформирована и остается вспомнить, что математическое ожидание числа успешных испытаний на момент образования требуемой еще серии из k успехов равно $q(1-p^k)/p^{k-1}$. В случае событий B_1, B_2, …, B_{l-1} с их суммарной вероятностью $1 - q^{l-1}$ величина $M^*(2, k, l)$ совпадает с $M^*(1, k, l)$. Приходим к равенствам

$$M^*(2, k, l) = (1 - q^{l-1})M^*(1, k, l) + \frac{q^{l-2}(1-p^k)}{p^{k-1}}. \tag{21}$$

Из соотношений (20) и (21) следует, что

$$M^*(1, k, l) = \frac{(1-p^{k-1})p^{k-1}q^{l-1} + p^{2k-2} + q^{2l-2}(1-p^{k-1})(1-p^k)}{q^l p^{k-1}(q^{l-1} + p^{k-1} - p^{k-1}q^{l-1})}. \tag{22}$$

В итоге получаем, что

$$M^*(k, l) = pM^*(1, k, l) + qM^*(2, k, l) = (1 - q^l)M^*(1, k, l) + \frac{q^{l-1}(1-p^k)}{p^{k-1}},$$

или

$$M^*(k, l) = \frac{(1-q^l)(1 - p^{k-1}q^{l-1} + p^{2k-2} + q^{2l-2}(1-p^{k-1})(1-p^k))}{q^l p^{k-1}(q^{l-1} + p^{k-1} - p^{k-1}q^{l-1})} + \frac{q^{l-1}(1-p^k)}{p^{k-1}}. \tag{23}$$

Здесь были представлены результаты, в которых приведены выражения для математических ожиданий чисел успешных испытаний в нескольких новых бернуллиевских схемах. Соответствующие результаты для математических ожиданий чисел неудачных испытаний в большинстве подобных схем получаются обычно перестановкой в этих формулах вероятностей p и q.

Список литературы

1. В. Феллер, *Введение в теорию вероятностей и ее приложения*, том 1, Москва, Мир, 1984
2. С. М. Ананьевский, В. Б. Невзоров, *О некоторых вероятностных распределениях, связанных с классической схемой Бернулли.* — Вестник Санкт-Петербургского университета, Математика, механика, астрономия **9** (67), вып. 2 (2022), 201–208.
3. С. М. Ананьевский, В. Б. Невзоров, *О сериях успехов и неудач в схемах Бернулли.* — Зап. научн. семин. ПОМИ **515** (2022), 30–38.
4. С. М. Ананьевский, В. Б. Невзоров, *О некоторых вероятностных распределениях, связанных с классической схемой Бернулли.* II. — Вестник Санкт-Петербургского университета, Математика, механика, астрономия **10** (68), вып. 1 (2022), 14–20.

Ananjevskii S. M., Nevzorov V. B. On the number of successes and their series in Bernoulli sequences of random variables.

One of the classical directions in probability theory is research related to the so-called Bernoulli sequences of random variables. We are talking about independent random variables $X_1, X_2, \ldots,$ taking the value 1 with some probability p, $0 < p < 1$, and the value 0 with probability $q = 1-p$. Often the event $\{X_n = 1\}$ is interpreted as "success in the nth trial", and its complement—the event $\{X_n = 0\}$—as "failure in this trial".

This classic scheme is named after Jacob Bernoulli, who considered such sequences. His results were published by his nephew Nicolaus Bernoulli in 1713 in Basel in the book Ars Conjectandi ("The Art of Conjecture"). Since then, although many results have appeared for various constructions based on Bernoulli random variables, there is a need to consider new schemes for them and to solve new problems. The article examines the relationship of such two-point distributions with a number of other probabilistic laws. A small review of previous results in this area and some new ones are added. The research started in previous works of the authors is continued.

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

Поступило 29 сентября 2023 г.