Записки научных
семинаров ПОМИ
Том 535, 2024 г.

С. М. Ананьевский, В. Б. Невзоров

ОБ ОДНОМ ОБОБЩЕНИИ СХЕМЫ БЕРНУЛЛИ

§1. Введение

Одни из первых исследований в теории вероятностей были связаны со схемами Бернулли – последовательностями независимых одинаково распределенных случайных величин (с.в.) $X_1, X_2, \ldots$, принимающих значения 1 с некоторой вероятностью $0 < p < 1$ и 0 с вероятностью $q = 1 - p$. Часто для удобства события $\{X_k = 0\}$ и $\{X_k = 1\}$, $k = 1, 2, \ldots$, трактуются, соответственно, как "неудача" и "успех" в k-ом испытании. Суммы $S_n = X_1 + X_2 + \ldots + X_n$ определяют число успехов в n испытаниях и имеют биномиальные $B(n,p)$-распределения с вероятностями

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

и математические ожидания

$$\mathbf{E}\, S_n = np.$$

Еще одним классическим семейством вероятностных распределений, тесно связанным с бернуллиевскими величинами, является семейство геометрических $\mathrm{Geom}(p)$ случайных величин $N_1(p)$, представляющих собой числа успехов до первой неудачи, с вероятностями

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

и математическими ожиданиями

$$\mathbf{E}\, N_1(p) = p/(1-p).$$

Если рассмотреть число успехов $N_2(r,p)$, предшествующих моменту проведенного r-ого неудачного испытания, то оно имеет так называемое отрицательное биномиальное распределение с целым параметром r. В этом случае

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

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

и

$$\mathbf{E}\, N_2(r,p) = rp/(1-p), \quad r = 1, 2, \ldots,$$

В [2–4] исследовались и несколько неклассических схем. В частности, рассматривались математические ожидания $\mathbf{E}\, N_3(k,r)$, $k = 1, 2, \ldots$, $r = 1, 2, \ldots$, чисел успехов, зафиксированных на момент первого появления в последовательности X-ов либо k-го успеха, либо r-ой по счету неудачи. Для двойной производящей функции последовательности $\mathbf{E}\, N_3(k,r)$

$$P_3(s,t) = \sum_{k=0}^{\infty}\sum_{r=0}^{\infty} \mathbf{E}\, N_3(k,r) s^k t^r$$

получено соотношение

$$P_3(s,t) = \frac{pst}{(1-s)(1-t)(1-ps-qt)}. \tag{1}$$

В [5] были рассмотрены и числа успехов $N_4(k,r)$ до появления не менее k успешных испытаний и не менее r неудачных. Показано, что двойная производящая функции для последовательности математических ожиданий $\mathbf{E}\, N_4(k,r)$ имеет следующий вид:

$$\begin{aligned} P_4(s,t) &= \sum_{k=0}^{\infty}\sum_{r=0}^{\infty} \mathbf{E}\, N_4(k,r) s^k t^r \\ &= \frac{(s(1-sp)/(1-s)^2 + pt(1-qt)/(1-t)^2 + pst/(1-s)(1-t))}{(1-sp-qt)}. \end{aligned} \tag{2}$$

§2. Основные результаты

Перейдем к некоторому обобщению схемы Бернулли. Рассмотрим следующую последовательность независимых одинаково распределенных с.в. $X_1, X_2, \ldots$, принимающих значения -1, 0, 1 с вероятностями

$$\mathbf{P}\{X_n = -1\} = p_1, \quad \mathbf{P}\{X_n = 0\} = p_2, \quad \mathbf{P}\{X_n = 1\} = p_3,$$

где

$$0 < p_1 < 1, \quad 0 < p_2 < 1, \quad 0 < p_3 < 1 \ \text{ и } \ p_1 + p_2 + p_3 = 1.$$

Можно, например, трактовать события

$$A_n = \{X_n = -1\}, \quad B_n = \{X_n = 0\}, \quad C_n = \{X_n = 1\}, \ n = 1, 2, \ldots,$$

как проигрыш, ничья или выигрыш шахматиста в своей n-ой партии. Эти события могут описывать и случайное блуждание по целочисленной решетке на прямой, когда A_n соответствует сдвигу на единицу влево в момент времени n, C_n обозначает шаг вправо, а B_n означает, что сохраняется положение в достигнутой к этому моменту точке.

Если интересоваться только числом значений -1 в наборе из n с.в. $X_1, X_2, \ldots, X_n$, то к такого рода событиям можно применять формулы, используемые для схем Бернулли с вероятностями успеха p_1. Аналогично, к появлениям значений $+1$ можно подходить, как к появлению успехов в схеме Бернулли с вероятностью успеха p_3. Если интересует появление только нулевых значений X-ов, то их число в n проводимых испытаниях имеет биномиальное $B(n, p_2)$-распределение, а математическое ожидание числа таких появлений равно np_2.

Но в схеме с тремя возможными вариантами значений случайных величин появляется и ряд новых, по сравнению с бернуллиевской схемой, задач. Рассмотрим некоторые из них, ограничившись ситуациями, связанными с появлениями в данной схеме нулевых значений случайных величин. Аналогичные результаты для значений -1 или $+1$ получатся просто заменой в получаемых формулах вероятности p_2 на p_1 или p_3.

a) Уже упомянуто, что математическое ожидание числа появлений нулевых значений среди величин $X_1, X_2, \ldots, X_n$ равно np_2.

Рассмотрим ситуацию, когда подсчет нулевых значений заканчивается в момент появления в последовательности X-ов n-го по счету значения -1. Пусть $E_5(n)$ обозначает соответствующее среднее значение появляющихся в этой ситуации нулевых значений. Для $n = 1$, рассматривая варианты $\{X_1 = -1\}$, при котором число наблюдаемых нулей равно нулю, $\{X_1 = 0\}$ и $\{X_1 = 1\}$, приходим к равенству

$$E_5(1) = p_2(1 + E_5(1)) + p_3 E_5(1), \tag{3}$$

из которого следует, что

$$E_5(1) = \frac{p_2}{1 - p_2 - p_3} = \frac{p_2}{p_1}. \tag{4}$$

Отсюда естественно следует, что

$$E_5(n) = \frac{np_2}{p_1}, \quad n = 1, 2, \ldots. \tag{5}$$

b) Для нахождения величины $E_6(n)$ – математического ожидания числа появлений нулей к моменту получения n единичных значений в последовательности X-ов достаточно в (5) заменить вероятность p_1 на p_3 и получить, что

$$E_6(n) = \frac{np_2}{p_3}, \quad n = 1, 2, \dots . \tag{6}$$

c) Пусть $E_7(k, l)$ обозначает математическое ожидание числа нулей, насчитываемых в момент, когда происходит первое из следующих двух событий:

$\{$появление k-го по счету значения $-1\}$

или

$\{$появление l-го по счету значения $+1\}$.

Отметим, что

$$E_7(k, 0) = 0, \quad k = 1, 2, \dots, \quad \text{и} \quad E_7(0, l) = 0, \quad l = 1, 2, \dots,$$

а

$$E_7(1, 1) = p_2(1 + E_7(1, 1)),$$

то есть

$$E_7(1, 1) = \frac{p_2}{1 - p_2}. \tag{7}$$

Для $k \geq 1$, $l \geq 1$ справедливы равенства

$$E_7(k, l) = p_1 E_7(k - 1, l) + p_2(1 + E_7(k, l)) + p_3 E_7(k, l - 1). \tag{8}$$

Из равенства (8), в частности, получаем:

$$E_7(2, 1) = \frac{p_1 p_2}{(1 - p_2)^2} + \frac{p_2}{1 - p_2}, \quad E_7(1, 2) = \frac{p_3 p_2}{(1 - p_2)^2} + \frac{p_2}{1 - p_2}$$

и

$$E_7(2, 2) = \frac{2p_1 p_2 p_3}{(1 - p_2)^3} + \frac{p_1 p_2}{(1 - p_2)^2} + \frac{p_2}{1 - p_2} + \frac{p_2 p_3}{(1 - p_2)^2}.$$

Если $p_1 = p_2 = p_3 = \frac{1}{3}$, то

$$E_7(2, 1) = E_7(1, 2) = \frac{3}{4} \quad \text{и} \quad E_7(2, 2) = \frac{5}{4}.$$

Далее рассмотрим двойную производящую функцию

$$\begin{aligned} \Pi_1(s, t) &= \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} E_7(k, l) s^k t^l = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} E_7(k, l) s^k t^l \\ &= p_1 \sum_{k=1}^{\infty} \sum_{l=0}^{\infty} E_7(k - 1, l) s^k t^l + p_2 \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} s^k t^l + p_2 \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} E_7(k, l) s^k t^l \end{aligned}$$

$$+ p_3 \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} E_7(k,l-1)s^k t^l = p_1 \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} E_7(k,l)s^k t^l + \frac{p_2 st}{(1-s)(1-t)}$$

$$+ p_2 \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} E_7(k,l)s^k t^l + p_3 t \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} E_7(k,l)s^k t^l$$

$$= \frac{p_2 st}{(1-s)(1-t)} + \Pi_1(s,t)(p_1 s + p_2 + p_3 t) \quad (9)$$

Из (9) получаем, что

$$\begin{aligned}\Pi_1(s,t) &= \frac{p_2 st}{(1-s)(1-t)(1-p_2-p_1 s - p_3 t)} \\ &= \frac{p_2 st}{(1-s)(1-t)(p_1(1-s)+p_3(1-t))}.\end{aligned} \quad (10)$$

d) Рассмотрим теперь ситуацию, когда подсчет числа нулей прекращается в первый момент, когда зафиксировано не менее k значений -1 и не менее l единиц.

Пусть $E_8(k,l)$ обозначает математическое ожидание такого числа нулевых значений.

Заметим, что

$$E_8(k,0) = E_5(k) = \frac{kp_2}{p_1}, \quad k = 1,2,\ldots,$$

и

$$E_8(0,l) = E_6(l) = \frac{lp_2}{p_3}, \quad l = 1,2,\ldots.$$

Если рассмотреть математическое ожидание $E_8(1,1)$, то получим, что

$$\begin{aligned}E_8(1,1) &= p_1 E_8(0,1) + p_2(1+E_8(1,1)) + p_3 E_8(1,0) \\ &= p_2 + \frac{p_1 p_2}{p_3} + \frac{p_3 p_2}{p_1} + p_2 E_8(1,1),\end{aligned} \quad (11)$$

откуда следует, что

$$E_8(1,1) = \frac{(p_2 + p_1 p_2/p_3 + p_3 p_2/p_1)}{1-p_2} = \frac{p_2(1+p_1/p_3+p_3/p_1)}{(1-p_2)} \quad (12)$$

Рассмотрим эти математические ожидания для значений $k \geq 1$ и $l \geq 1$. Для них справедливы соотношения

$$E_8(k,l) = p_1 E_8(k-1,l) + p_2(1+E_8(k,l)) + p_3 E_8(k,l-1) \quad (13)$$

при $k = 1, 2, \ldots, \quad l = 1, 2, \ldots$. Из (12) и (13) можно получить

$$E_8(2,1) = \frac{p_1 p_2 (1 + p_1/p_3 + p_3/p_1)}{(1-p_2)^2} + \frac{p_2 + 2p_3}{1 - p_2},$$

$$E_8(1,2) = \frac{p_3 p_2 (1 + p_1/p_3 + p_3/p_1)}{(1-p_2)^2} + \frac{p_2 + 2p_1}{1 - p_2}$$

и

$$E_8(2,2) = \frac{p_1 E_8(1,2) + p_2 + p_3 E_8(2,1)}{1 - p_2}.$$

При $p_1 = p_2 = p_3 = \frac{1}{3}$

$$E_8(2,1) = E_8(1,2) = \frac{9}{4} \quad \text{и} \quad E_8(2,2) = \frac{11}{4}.$$

Перейдем к двойной производящей функции

$$\Pi_2(s,t) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} E_8(k,l) s^k t^l.$$

Используя равенства (11)–(13), получаем, что

$$\begin{aligned}
\Pi_2(s,t) &= \sum_{k=0}^{\infty} E_8(k,0)s^k + \sum_{l=0}^{\infty} E_8(0,l)t^l + \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} E_8(k,l)s^k t^l \\
&= \sum_{k=0}^{\infty} E_8(k,0)s^k + \sum_{l=0}^{\infty} E_8(0,l)t^l + p_1 \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} E_8(k-1,l)s^k t^l \\
&+ p_2 \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} s^k t^l + p_2 \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} E_8(k,l)s^k t^l + p_3 \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} E_8(k,l-1)s^k t^l \\
&= \sum_{k=0}^{\infty} E_8(k,0)s^k + \sum_{l=0}^{\infty} E_8(0,l)t^l + p_1 s \sum_{k=0}^{\infty}\sum_{l=1}^{\infty} E_8(k,l)s^k t^l \\
&+ p_3 t \sum_{k=1}^{\infty}\sum_{l=0}^{\infty} E_8(k,l)s^k t^l + p_2 \frac{st}{(1-s)(1-t)} + p_2 \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} E_8(k,l)s^k t^l \\
&- p_2 \sum_{k=0}^{\infty} E_8(k,0)s^k - p_2 \sum_{l=0}^{\infty} E_8(0,l)t^l = \sum_{k=0}^{\infty} E_8(k,0)s^k + \sum_{l=0}^{\infty} E_8(0,l)t^l \\
&+ p_2 \frac{st}{(1-s)(1-t)} + p_2 \Pi_2(s,t) - p_2 \sum_{k=0}^{\infty} E_8(k,0)s^k - p_2 \sum_{l=0}^{\infty} E_8(0,l)t^l
\end{aligned}$$

$$
\begin{aligned}
&+ p_1 s(\Pi_2(s,t) - \sum_{k=0}^{\infty} E_8(k,0)s^k) + p_3 t(\Pi_2(s,t) - \sum_{l=0}^{\infty} E_8(0,l)t^l) \\
&= (1 - p_2 - p_1 s) \sum_{k=0}^{\infty} E_8(k,0)s^k) + (1 - p_2 - p_3 t) \sum_{l=0}^{\infty} E_8(0,l)t^l \\
&\qquad + p_2 \frac{st}{(1-s)(1-t)} + \Pi_2(s,t)(p_2 + p_1 s + p_3 t). \qquad (14)
\end{aligned}
$$

Заметим, что

$$
\sum_{k=0}^{\infty} E_8(k,0)s^k = \frac{p_2}{p_1} \sum_{k=0}^{\infty} ks^k = \frac{p_2 s}{p_1(1-s)^2}, \qquad (15)
$$

$$
\sum_{l=0}^{\infty} E_8(0,l)t^l = \frac{p_2}{p_3} \sum_{l=0}^{\infty} lt^l = \frac{p_2 t}{p_3(1-t)^2}. \qquad (16)
$$

Из равенств (14)–(16) следует, что

$$
\begin{aligned}
\Pi_2(s,t) = &\frac{p_2 st}{(1-s)(1-t)(1-p_2-p_1 s-p_3 t)} \\
&+ \frac{p_2 s(1-p_2-p_1 s)}{p_1(1-s)^2(1-p_2-p_1 s-p_3 t)} \\
&+ \frac{p_2 t(1-p_2-p_3 t)}{p_3(1-t)^2(1-p_2-p_1 s-p_3 t)}. \qquad (17)
\end{aligned}
$$

Список литературы

1. J. Bernoulli, *Ars conjectandi, opus posthumum. Accedit Tractatus de seriebus infinitis, et epistola gallicé scripta de ludo pilae reticularis*, Basel, Thurneysen Brothers, 1713.
2. С. М. Ананьевский, В. Б. Невзоров, *О некоторых вероятностных распределениях, связанных с классической схемой Бернулли.* — Вестник СПбГУ, сер. 1. Математика, механика, астрономия **9(67)**, вып. 2 (2022), 201–208.
3. С. М. Ананьевский, В. Б. Невзоров, *О некоторых вероятностных распределениях, связанных с классической схемой Бернулли.II.* — Вестник СПбГУ, сер. 1. Математика, механика, астрономия **10(68)**, вып. 1 (2023), 14–20.
4. С. М. Ананьевский, В. Б. Невзоров, *О сериях успехов и неудач в схемах Бернулли.* — Зап. научн. семин. ПОМИ **515** (2022), 30–38.
5. С. М. Ананьевский, В. Б. Невзоров, *О числе успехов и их серий в бернуллиевских последовательностях случайных величин.* — Зап. научн. семин. ПОМИ **525** (2023), 22–29.

Ananjevskii S. M., Nevzorov V. B. On a generalization of the Bernoulli scheme.

The paper considers a generalization of Bernoulli's scheme. We consider a sequence of independent identically distributed random variables (r.v.) $X_1, X_2, \ldots,$ taking values $-1,\ 0,\ 1$ with probabilities

$$\mathbf{P}\{X_n = -1\} = p_1, \quad \mathbf{P}\{X_n = 0\} = p_2, \quad \mathbf{P}\{X_n = 1\} = p_3,$$

where

$$0 < p_1 < 1, \quad 0 < p_2 < 1, \quad 0 < p_3 < 1 \text{ and } p_1 + p_2 + p_3 = 1.$$

If we are only interested in the number of -1 values in a set of n r.v. $X_1, X_2, \ldots, X_n$, then the formulas used for Bernoulli schemes with success probabilities p_1 can be applied to such events. Similarly, occurrences of values $+1$ can be treated as occurrences of successes in a Bernoulli scheme with probability of success p_3. If you are interested in the appearance of only zero values of X, then their number in n trials has a binomial $B(n, p_2)$ distribution, and the mathematical expectation of the number of such appearances is equal to np_2.

But in a scheme with three possible variants of values of random variables, a number of new problems appear, in comparison with the Bernoulli scheme. The paper examines some of them, limiting ourselves to situations associated with the appearance of zero values of random variables in a given scheme. Similar results for values -1 or $+1$ can be obtained simply by replacing the probability of p_2 with p_1 or p_3 in the resulting formulas. The article examines the relationship of such three-point distributions with a number of other probability laws. A short review of previously obtained results in this area is given and several new ones are added. The research begun in the previous works of the authors was continued.

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

Поступило 26 сентября 2024 г.