Afrika Statistika

Afrika Statistika
Vol. 18 (2), 2023, pages 3467 - 3474.
DOI: http://dx.doi.org/10.16929/as/2023.3467.313

ISSN 2316-090X

About successes and failures in Bernoulli schemes

Mohammad Ahsanullah [(1,*)] , **Sergey M. Ananjevskii** [(2)], **Valery B. Nevzorov** [(2)]

[(1)] Rider University. Lawrenceville, New Jersey, USA
[(2)] St.Petersburg State University, St.Petersburg, Russian Federation

Received on August 14, 2023; Accepted on September 01, 2023; Published on June 15, 2024

Abstract. One of the most popular areas in the probability theory is the research related to Bernoulli random variables. In this paper we will consider the random variable $N_{1,\ell}(n)$ that will take on a value equal to the number of successes before the ℓ-th failure occurs if the ℓ-th failure occurred before time n and a value equal to the number of successes that occurred by time n if the ℓ-th failure did not occur by n-th trial. Some variations of $N_{1,\ell}(n)$ will also be considered.

Key words: Bernoulli random variables; geometric distribution and binomial distribution
AMS 2010 Mathematics Subject Classification Objects : 60A10; 60E05; 60E99

[(*)] Correspondaning author: Mohammad Ahsanullah (ahsan@rider.edu)
Sergey M. Ananjevskii: ananjevskii@mail.ru
Valery B. Nevzorov : valnev@mail.ru

Résumé (Abstract in French) La recherche liée aux variables aléatoires de Bernoulli est l'un des thèmes les plus populaires en théorie des probabilités. Dans cet article, nous considérerons la variable aléatoire $N_{1,\ell}(n)$ qui est le nombre de succès avant que le ℓ-ième échec ne se produise si le ℓ-ième échec s'est produit avant le temps n et qui est le nombre de succès survenus au temps n si le ℓ-ième échec ne s'est pas produit au n-ième essai. Certaines variantes de $N_{1,\ell}(n)$ seront également prises en compte.

Presentation of authors.

Mohammad Ahsanullah, Ph.D., Professor Emeritus of Statistics at: Rider University, USA.

Valery B. Nevzorov, Ph.D., Professor of Statistics, St. Petersburg State Unuversity, Russian Federation.

Sergey M. Ananjevskii, Ph.D., Professor of Statistics, St. Petersburg State Unuversity, Russian Federation.

1. Introduction

Consider a sequence of independent random variables $X_1, X_2, \ldots$ which take values 1 with some probability p, $0 < p < 1$, and values 0 with probability $q = 1 - p$. Usually in this kind of events $\{X_n = 1\}$ and $\{X_n = 0\}$ are interpreted as a success and a failure in the n-th trial. This Bernoulli scheme, in addition to two-point ones, is associated with a number of classical probability distributions.

The sum $S_n = X_1 + X_2 + \ldots + X_n$, $n \geq 1$, denotes the number of successes in a series of independent Bernoulli trials and have binomial distributions with probabilities:

$$\mathbb{P}(S_n = m) = C_m^n p^m q^{n-m}, \quad m \geq 1 \;\; and \;\; n \geq 1$$

and the mathematical expectations $\mathbb{E}(S_n) = np$. The number of successes $N(1,p)$ before the first failure has a geometric distribution:

$$\mathbb{P}(N(1,p) = n) = p^n q, \quad n \geq 1$$

and $\mathbb{E}(N(1,p)) = p/q$. More generally, let $r \geq 1$ and let $N(r,p)$ be the number of successful trials $N(r,p)$ before the appearance of the r-th failure has a negative binomial distribution with probabilities

$$\mathbb{P}(N(r,p) = m) = C_{r-1}^{m+r-1} p^m q^r, \quad m \geq 1$$

and mathematical expectations $\mathbb{E}(N(r,p)) = rp/q$.

More than three centuries have passed since the publication by Nicholas Bernoulli (see Basel in 1913) in the book *Arc Conjesttandi (The Art of Conjecture)* of Bernoulli, Jacob. Ars (1713) some results which later called Bernoulli sequences of random variables. But until now, new results appear for various variations of Bernoulli schemes (see, for example, in reviews in Ananjevskii and Nevzorov(2022), Ananjevskii and Nevzorov (2023a), Ananjevskii and Nevzorov(2023b)). In those reviews, they studied, in particular, the number of trials, the number of successes and the number of failures until the moments of the appearance of various series of successful or unsuccessful states. When setting up most of these problems, the possible number of tests carried out before the implementation of certain events was usually not limited. For example, the number of trials before the first series of a certain number of failures appears can be arbitrarily large. In this paper we will consider schemes in which the number of tests cannot exceed a certain fixed value n, and schemes in which at least two tests are carried out.

2. Main Results

Let $N_{1,1}(n)$ be denote the number of successes, counting which ends either at time n, or when the first failure occurs, if this occurs before time n, and $E_{1,1}(n)$ is the corresponding mathematical expectation of this number of successes. The random variable $N_{1,1}(n)$ can take the values in $\{0, 1, 2, \ldots, n-1\}$ with probabilities

$$\mathbb{P}(N_{1,1}(n) = m) = (1-p)p^m$$

and

$$\mathbb{P}(N_{1,1}(n) = n) = p^n.$$

In this case we get that the expected value $E_{1,1}(n)$ as

$$\begin{aligned} E_{1,1}(n) &= np^n + (1-p)\sum_{m=0}^{n-1} mp^m && (1)\\ &= np^n + \sum_{m=0}^{n-1} mp^m - \sum_{m=0}^{n-1} mp^{m+1} \\ &= p + p^2 + p^3 + \ldots + p^n \\ &= p(1-p^n)/(1-p) && (2) \end{aligned}$$

Let us consider a more general situation with numbers of successes $N_{1,\ell}(n)$ and their mathematical expectations $E_{1,\ell}(n)$. The random variable $N_{1,\ell}(n)$ will take on a value equal to the number of successes before the ℓ-th failure occurs if the ℓ-th failure occurred before time n and a value equal to the number of successes that occurred at time n if the ℓ-th failure occurred by this moment not yet recorded. The mathematical expectations $E_{1,\ell}(n)$ are related by the equalities

$$E_{1,\ell}(n) = p(1 + E_{1,\ell}(n-1)) + qE_{1,\ell-1}(n-1)), \quad \ell \geq 1,$$

Let us denote

$$\pi_1(s,t) = \sum_{l=0}^{\infty}\sum_{n=0}^{\infty} E_{1,\ell}(n)s^{\ell}\, t^n,$$

the double generating function for quantities $E_{1,\ell}(n)$. Using Equality , we obtain that

$$\begin{aligned}
\pi_1(s,t) &= p\sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty}(1 + E_{1,\ell}(n-1))s^{\ell}\, t^n + q\sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty} E_{1,\ell-1}(n-1)s^{\ell}\, t^n \\
&= pst/(1-s)(1-t) + pt\sum_{\ell=1}^{\infty}\sum_{n=0}^{\infty} E_{1,\ell}(n)s^{\ell}\, t^n + qst\sum_{\ell=0}^{\infty}\sum_{n=0}^{\infty} E_{1,\ell}(n)s^{\ell}\, t^n \\
&= pst/(1-s)(1-t) + (pt + qst)\pi_1(s,t).
\end{aligned}$$

Whence it follows that

$$\pi_1(s,t) = pst/(1-s)(1-t)(1-pt-qst). \tag{3}$$

Consider the number of successes $S_1(k,n)$, which is equal to k if the k-th successful test is completed before time n, and otherwise $S_1(k,n)$ coincides with the number of successes recorded at time n. Let also

$$M_1(k,n) = ES_1(k,n),$$

and

$$P_1(s,t) = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} M_1(k,n)s^k t^n \tag{4}$$

denotes the double generating function for the mathematical expectations of random variables $S_1(k,n)$. If $k \geq n$ then $M_1(k,n) = np$. The following recurrence relations are valid

$$M_1(k,n) = p(1 + M_1(k-1,n-1)) + qM_1(k,n-1) \tag{5}$$

Using them, for the generating function $P_1(s,t)$, we obtain the following equalities

$$\begin{aligned} P_1(s,t) &= pst/(1-s)(1-t) + pst\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} M_1(k,n)s^k t^n \\ &+ qt\sum_{k=1}^{\infty}\sum_{n=0}^{\infty} M_1(k,n)s^k t^n \\ &= pst/(1-s)(1-t) + (pst+qt)P_1(s,t), \end{aligned}$$

from which we obtain the form of the generating function

$$P_1(s,t) = pst/(1-s)(1-t)(1-pst-qt). \tag{6}$$

We considered options when the number of tests was limited to a certain number n. Let us move on to situations where at least n tests are already being carried out.

Now let the tests be completed either at time n, if the ℓ-th failure has already been recorded, or at the time of the l-th unsuccessful attempt, if it is later than time n. Let us denote $N_{2,\ell}(n)$ the random number of successful trials at this moment and its mathematical expectation as

$$E_{2,\ell}(n) = \mathbb{E}(N_{2,\ell}(n)).$$

Note that

$$E_{2,0}(n) = np, \ \ n \geq 0$$

and

$$E_{2,\ell}(n) = \ell\ p/q \ \ \textit{where}\ \ell \geq n.$$

The following recurrence relations are valid

$$E_{2,\ell}(n) = p(E_{2,\ell}(n-1)+1) + qE_{2,\ell-1}(n-1)), \quad \ell \geq 1. \tag{7}$$

We will consider the generating function

$$\pi_2(s,t) = \sum_{\ell=0}^{\infty}\sum_{n=0}^{\infty} E_{2,\ell}s^{\ell}\ t^n.$$

By taking into account Equality (7), we obtain that

$$\begin{aligned}
\pi_2(s,t) &= \sum_{\ell=0}^{\infty} E_{2,0}(n)t^n + \sum_{\ell=0}^{\infty} E_{2,\ell}(0)s^\ell + p\sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty} E_{2,\ell}(n-1)s^\ell\, t^n \\
&+ q\sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty} E_{2,\ell-1}(n-1)s^\ell\, t^n \\
&= p\sum_{n=0}^{\infty} nt^n + (p/q)\sum_{\ell=0}^{\infty} \ell\, s^\ell + pst/(1-s)(1-t) \\
&+ pt\sum_{\ell=1}^{\infty}\sum_{n=0}^{\infty} E_{2,\ell}(n)s^\ell\, t^n + qst\pi_2(s,t) \\
&= pt/(1-t)^2 + ps/q(1-s)^2 + +pst/(1-s)(1-t) + pt(\pi_2(s,t) \\
&- \sum_{n=0}^{\infty} E_{2,0}(n)t^n) + qst\pi_2(s,t) \\
&= pt/(1-t)^2 + ps/q(1-s)^2 + pst/(1-s)(1-t) + (pt+qst)\pi_2(s,t) - pt\sum_{n=0}^{\infty} np\; t^n,
\end{aligned}$$

whence, it follows that

$$\pi_2(s,t)(1-pt-qst) = pt/(1-t)^2 + ps/q(1-s)^2 + pst/(1-s)(1-t) - p^2t^2/(1-t)^2.$$

Finally, we get that

$$\pi_2(s,t) = (pt(1-pt)/(1-t)^2 + ps/q(1-s)^2 + pst/(1-s)(1-t))/(1-pt-qst).$$

3) Let us move on to considering the numbers of successes in situations where their counting is completed at the moment when at least n tests have been carried out and at least k successes have been received by this moment. Let $S_2(k,n)$ denote the corresponding number of successes. Note that it is equal to k if the kth successful test is completed after time n.

Let

$$M_2(k,n) = \mathbb{E}(S_2(k,n)),$$

and

$$P_2(s,t) = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} M_2(k,n)s^k t^n$$

denotes the double generating function for the mathematical expectations of success numbers $S_2(k,n)$. Notice that

$$\begin{aligned}
&M_2(k,n)=k,\ \textit{if}\ k\geq n,\\
&M_2(0,n)=np,\quad n\geq 0.
\end{aligned}$$

and

$$M_2(k,0)=k,\quad k\geq 1$$

The following equalities are valid, taking into account the conditions whether the first attempt was successful or unsuccessful:

$$M_2(k,n)=p(1+M_2(k-1,n-1))+qM_2(k,n-1)),\quad k=\geq 1,\quad n\geq 1.$$

By consider the generating function

$$P_2(s,t)=\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}M_2(k,n)s^kt^n,$$

we get that

$$\begin{aligned}
P_2(s,t) &= \sum_{n=0}^{\infty}M_2(0,n)t^n+\sum_{k=1}^{\infty}M_2(k,0)s^k\\
&+ pst\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}M_2(k,n)s^kt^n+\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}ps^kt^n+qt\sum_{k=1}^{\infty}\sum_{n=0}^{\infty}M_2(k,n)s^kt^n\\
&= \sum_{n=1}^{\infty}npt^n+\sum_{k=1}^{\infty}ks^k+pstP_2(s,t)+pst/(1-s)(1-t)+qtP_2(s,t)\\
&= pt\sum_{n=1}^{\infty}nt^{n-1}+s\sum_{k=1}^{\infty}ks^{k-1}+pst/(1-s)(1-t)+pstP_2(s,t)+qtP_2(s,t)\\
&- pt/(1-t)^2+s/(1-s)^2+pst/(1-s)(1-t)+pstP_2(s,t)+qtP_2(s,t) \qquad (8)
\end{aligned}$$

From (8), it follows that

$$P_2(s,t)=(s/(1-s)^2+pt(1-qt)/(1-t)^2+pst/(1-s)(1-t))/(1-qt-pst).$$

3. Conclusion

We hope the result of this paper will be able to provide new ideas about more researches in Bernoulli problems in probability.

Declaration. The authors declare that there is no conflicting interest in this work.

Acknowledgment. The authors thank the Editor-in-chief and the referees for valuable comments that improved the presentation of the paper.

References

Bernoulli, Jacob. Ars conjectandi (1713); opus posthumum. Accedit Tractatus de seriebus infinitis, et epistola gallica scripta de ludo pilae reticularis.- Basel: Thurneysen Brothers.

Ananjevskii,S.M and Nevzorov, V.B (2022). On some probabilistic distributions related to the classical Bernoulli schema,-Vestn. ST. Petersburg Univ.: Math 55, 141-146.

Ananjevskii,S.M and Nevzorov, V.B (2023a). On some probabilistic distributions related to the classical Bernoulli schema II,-Vestn. ST. Petersburg Univ.:56, 9-14.

Ananjevskii,S.M and Nevzorov, V.B.(2023b). On the number of successes and their series in Bernoulli sequences of random variables, Zapiski Nauchn. seminar.POMI, 2023, 525.