J. Appl. Prob. **46,** 1201–1208 (2009)
Printed in England

ON CONSECUTIVE RECORDS IN CERTAIN BERNOULLI SEQUENCES

LARS HOLST,* *Royal Institute of Technology*

Abstract

In an infinite sequence of independent Bernoulli trials with success probabilities $p_k = a/(a+b+k-1)$ for $k = 1, 2, 3, \ldots$, let N_r be the number of $r \geq 2$ consecutive successes. Expressions for the first two moments of N_r are derived. Asymptotics of the probability of no occurrence of r consecutive successes for large r are obtained. Using an embedding in a marked Poisson process, it is indicated how the distribution of N_r can be calculated for small r.

Keywords: Bernoulli trial; embedding in Poisson process; random permutation; records; runs of successes; sums of indicators

2000 Mathematics Subject Classification: Primary 60C05
Secondary 60K99

1. Introduction

In this paper we study the number of runs of successes or consecutive records in an infinite sequence of independent Bernoulli trials where the success or record probability in the kth trial is $p_k = a/(a+b+k-1)$ for $k = 1, 2, 3, \ldots$, $a > 0$, and $b \geq 0$. Note that, for $a = 1$ and $b = 0$, the ps are the probabilities of records in a sequence of independent, identically distributed, continuous random variables. For this case, results on the distributions of the number of consecutive records are given in Chern *et al*. (2000), Chern and Hwang (2005), and Hahlin (1995).

For $a > 0$ and $b = 0$, it follows from general results of Arratia *et al.* (1992), see also Arratia *et al.* (2003), on the limiting distributions of the cycle lengths in certain random permutations that the number of double records, that is, two consecutive successes, is Poisson with mean a. For the general case $a > 0$ and $b \geq 0$, Móri (2001) proved that the distribution of the number of double records is a beta mixture of Poisson distributions. Some recursion formulae concerning multiple records are also given there. An explicit expression for the binomial moment of the number of double records in the first n trials is derived in Holst (2008a). Letting $n \to \infty$ gives the limit distribution of Móri. This is also proved by other methods in Holst (2007), Holst (2008b), and Huffer *et al*. (2009).

Apart from the works of Chern *et al*. (2000), Chern and Hwang (2005), and Hahlin (1995) for the case in which $a = 1$ and $b = 0$, the author is not aware of other detailed studies on multiple records. In Section 2 we calculate for general $a > 0$ and $b \geq 0$ the first two moments of the number of r consecutive records and obtain an asymptotic expansion for the probability of no occurrence of r consecutive records. In Section 3 we illustrate how we can obtain distributions for the number of multiple records for small r, triple and quadruple records in particular. For

Received 6 April 2009.
* Postal address: Department of Mathematics, Royal Institute of Technology, SE–100 44 Stockholm, Sweden.
Email address: lholst@math.kth.se

https://doi.org/10.1239/jap/1261670698 Published online by Cambridge University Press

$a = 1$ and $b = 0$, we compare our numerical results with those of Chern *et al*. (2000). An embedding of the Bernoulli sequence in a marked Poisson process is used in our derivations.

2. Notation, moments, and occurrence of runs

In the rest of the paper $I_1, I_2, I_3, \ldots$ are independent Bernoulli random variables with

$$\mathrm{P}(I_k = 1) = 1 - \mathrm{P}(I_k = 0) = \frac{a}{a+b+k-1},$$

where $a > 0$ and $b \geq 0$. An infinite random sequence $\mathcal{S}$ of 1s and 0s is generated by the Is. The number of $r \geq 2$ consecutive 1s in $\mathcal{S}$ can be written as

$$N_r = \sum_{k=1}^{\infty} I_k I_{k+1} \cdots I_{k+r-1}.$$

By the Borel–Cantelli lemma, it follows that $N_r < \infty$ with probability 1. In the following the notation $s^{\overline{n}} = s(s+1)\cdots(s+n-1)$ is used for rising factorials.

Theorem 2.1. *The first two moments of the number of r consecutive records satisfy*

$$\mathrm{E}(N_r) = \frac{a^r}{(r-1)(a+b)^{\overline{r-1}}},$$

$$\mathrm{E}(N_r^2) = \mathrm{E}(N_r) + 2\sum_{j=r+1}^{2r-1} \mathrm{E}(N_j) + \frac{2a}{r-1}\,\mathrm{E}(N_{2r-1}).$$

Proof. Using the fact that the Is are independent, we obtain a telescoping sum, i.e.

$$\begin{aligned}
\mathrm{E}(N_r) &= \sum_{k=1}^{\infty} \mathrm{E}(I_k \cdots I_{k+r-1}) \\
&= \sum_{k=1}^{\infty} \frac{a^r}{(a+b+k-1)^{\overline{r}}} \\
&= \frac{a^r}{r-1} \sum_{k=1}^{\infty} \frac{1}{(a+b+k)^{\overline{r-2}}} \left(\frac{1}{a+b+k-1} - \frac{1}{a+b+k+r-2} \right) \\
&= \frac{a^r}{(r-1)(a+b)^{\overline{r-1}}}.
\end{aligned}$$

The independence of the Is, the identity $I_j^2 \equiv I_j$, and the formula for $\mathrm{E}(N_r)$ give

$$\begin{aligned}
\mathrm{E}(N_r^2) &= \mathrm{E}\bigg(\sum_{k=1}^{\infty} I_k I_{k+1} \cdots I_{k+r-1} \sum_{\ell=1}^{\infty} I_\ell I_{\ell+1} \cdots I_{\ell+r-1} \bigg) \\
&= \mathrm{E}(N_r) + 2\,\mathrm{E}\bigg(\sum_{k=1}^{\infty} I_k \cdots I_{k+r-1} \sum_{\ell=k+1}^{k+r-1} I_\ell \cdots I_{\ell+r-1} \bigg) \\
&\quad + 2\,\mathrm{E}\bigg(\sum_{k=1}^{\infty} I_k \cdots I_{k+r-1} \sum_{\ell=k+r}^{\infty} I_\ell \cdots I_{\ell+r-1} \bigg)
\end{aligned}$$

https://doi.org/10.1239/jap/1261670698 Published online by Cambridge University Press

$$
\begin{aligned}
&= \mathrm{E}(N_r) + 2\sum_{j=1}^{r-1} \mathrm{E}\bigg(\sum_{k=1}^{\infty} I_k \cdots I_{k+r-1+j}\bigg) \\
&\quad + 2\sum_{k=1}^{\infty} \mathrm{E}\bigg(I_k \cdots I_{k+r-1} \sum_{\ell=k+r}^{\infty} I_\ell \cdots I_{\ell+r-1}\bigg) \\
&= \mathrm{E}(N_r) + 2\sum_{j=1}^{r-1} \mathrm{E}(N_{r+j}) + 2\sum_{k=1}^{\infty} \mathrm{E}(I_k \cdots I_{k+r-1})\, \mathrm{E}\bigg(\sum_{\ell=k+r}^{\infty} I_\ell \cdots I_{\ell+r-1}\bigg) \\
&= \mathrm{E}(N_r) + 2\sum_{j=1}^{r-1} \mathrm{E}(N_{r+j}) + 2\sum_{k=1}^{\infty} \frac{a^r}{(a+b+k-1)^{\overline{r}}}\, \frac{a^r}{(r-1)(a+b+k+r)^{\overline{r-1}}} \\
&= \mathrm{E}(N_r) + 2\sum_{j=1}^{r-1} \mathrm{E}(N_{r+j}) + \frac{2a}{r-1}\sum_{k=1}^{\infty} \frac{a^{2r-1}}{(a+b+k-1)^{\overline{2r-1}}} \\
&= \mathrm{E}(N_r) + 2\sum_{j=r+1}^{2r-1} \mathrm{E}(N_j) + \frac{2a}{r-1}\,\mathrm{E}(N_{2r-1}),
\end{aligned}
$$

proving the assertion.

In particular, for $a = 1$ and $b = 0$, we have

$$
\mathrm{E}(N_r) = \frac{1}{(r-1)!\,(r-1)},
$$

$$
\mathrm{var}(N_r) = \frac{1}{(r-1)!\,(r-1)} + \sum_{j=r}^{2r-2} \frac{2}{j!\,j} + \frac{1}{(2r-2)!\,(r-1)^2} - \frac{1}{(r-1)!^2(r-1)^2},
$$

in agreement with Chern *et al.* (2000). For triple records, we obtain $\mathrm{E}(N_3) = \frac{1}{4}$ and $\mathrm{var}(N_3) = \frac{95}{288} > \frac{1}{4}$. Thus, the distribution of N_3 is not Poisson. Recall that the number of double records, N_2, is Poisson with mean 1.

Corollary 2.1. *For the mean of N_r, as $r \to \infty$,*

$$
\mathrm{E}(N_r) = \frac{\Gamma(a+b)a^r}{(r-1)!\,r^{a+b}}\bigg(1 + O\Big(\frac{1}{r}\Big)\bigg).
$$

Proof. Using Stirling's formula for the gamma function, we obtain, as $r \to \infty$,

$$
\begin{aligned}
\mathrm{E}(N_r) &= \frac{a^r}{(r-1)(a+b)^{\overline{r-1}}} \\
&= \frac{a^r}{(r-1)!}\bigg(1 + \frac{a+b}{r-1}\bigg)\frac{(r-1)!}{(a+b)^{\overline{r}}} \\
&= \frac{a^r}{(r-1)!}\bigg(1 + \frac{a+b}{r-1}\bigg)\frac{\Gamma(a+b)\Gamma(r)}{\Gamma(a+b+r)} \\
&= \frac{\Gamma(a+b)a^r}{(r-1)!\,r^{a+b}}\bigg(1 + O\Big(\frac{1}{r}\Big)\bigg).
\end{aligned}
$$

Corollary 2.2. *The probability of* $r \geq 2$ *consecutive records satisfies*

$$\mathrm{E}(N_r) > \mathrm{P}(N_r \geq 1) > \mathrm{E}(N_r)\left(1+\frac{b}{a+b+r}+\frac{(a+b)(b+1)}{(a+b+r)(a+b+r+1)}\right)\left(1+\frac{a+b}{r-1}\right)^{-1}.$$

Proof. As $\mathrm{P}(N_r \geq 2) \geq \mathrm{P}(I_1 = \cdots = I_{r+1} = 1) > 0$, the left-hand side inequality in the assertion follows. The right-hand side inequality follows from

$$\begin{aligned}\mathrm{P}(N_r \geq 1) &> \mathrm{P}(I_1 = \cdots = I_r = 1) + \mathrm{P}(I_1 = 0,\ I_2 = \cdots = I_{r+1} = 1)\\ &\quad + \mathrm{P}(I_2 = 0,\ I_3 = \cdots = I_{r+2} = 1)\\ &= \frac{a^r}{(a+b)^{\overline{r}}}\left(1 + \frac{b}{a+b+r} + \frac{(a+b)(b+1)}{(a+b+r)(a+b+r+1)}\right)\end{aligned}$$

and the above formula for $\mathrm{E}(N_r)$.

Combining the above results we obtain the following asymptotic expansion.

Corollary 2.3. *The probability of no occurrence of* r *consecutive records satisfies*

$$\mathrm{P}(N_r = 0) = 1 - \frac{\Gamma(a+b)a^r}{(r-1)!\, r^{a+b}}\left(1 + O\left(\frac{1}{r}\right)\right)$$

as $r \to \infty$.

For $a = 1$ and $b = 0$, we obtain

$$\mathrm{E}(N_r) = \frac{1}{(r-1)!\,(r-1)} > \mathrm{P}(N_r \geq 1) > \frac{1}{r!} + \frac{1}{(r+2)!},$$

which implies that

$$\mathrm{P}(N_r = 0) = 1 - \frac{1}{(r-1)!\,(r-1)} + O\left(\frac{1}{r!\,r}\right),$$

in agreement with Corollary 2 of Chern and Hwang (2005), proved in a different way. Using recursions of generating functions, $\mathrm{P}(N_r = 0)$ is numerically computed in Table 1 of Chern *et al.* (2000). With these values in the above inequality for $\mathrm{P}(N_r \geq 1)$ we obtain, for $r = 3, 4, 5, 6$,

$$\begin{aligned}0.250 > 0.192 > 0.175, &\qquad 0.0556 > 0.0451 > 0.0431,\\ 0.0104 > 0.008\,75 > 0.008\,53, &\qquad 0.001\,67 > 0.001\,44 > 0.001\,41.\end{aligned}$$

The simple approximation $\mathrm{P}(N_r \geq 1) \approx 1/r! + 1/(r+2)!$ is accurate; for example, $\mathrm{P}(N_7 \geq 1) = 0.000\,203$ compared to $1/7! + 1/9! = 0.000\,201$. This is not surprising: if a long run of records occurs then, with high probability, it should start in the first or third trial.

3. Embedding and multiple records

Apart from special cases, no simple formula for the distribution of N_r seems possible to give. However, using the embedding below, we will indicate how the distribution can be calculated for any $a > 0$ and $b \geq 0$.

In Holst (2008b) the following embedding of the sequence $\mathcal{S} = I_1 I_2 I_3 \cdots$ of 1s and 0s is given. First, generate an outcome p of a random variable P with a Beta(a, b) distribution, that is, P has density

$$f(p) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1-p)^{b-1}, \qquad 0 < p < 1;$$

Beta$(a, 0)$ is interpreted as $P \equiv 1$. Conditional on $P = p$, generate independent Poisson processes $\Pi_{1p}, \Pi_{2p}, \ldots$ on the positive real line, where $\Pi_{\ell p}$ has intensity

$$\lambda_\ell(t) = (1 - p\mathrm{e}^{-t/a})^{\ell-1} p\mathrm{e}^{-t/a}, \qquad t > 0.$$

Independent of $\Pi_{1p}, \Pi_{2p}, \ldots$, let L_0 be a geometric random variable with

$$\mathrm{P}(L_0 = \ell) = (1-p)^{\ell-1} p, \qquad \ell = 1, 2, \ldots.$$

The superposition Π_p of $\Pi_{1p}, \Pi_{2p}, \ldots$ is a Poisson process of intensity 1. Suppose that the kth point of Π_p originates from the process $\Pi_{L_k p}$. Define a sequence $\mathcal{S}_p$ of 1s and 0s by letting the first 1 occur at position L_0, the second at position $L_0 + L_1$, the third at $L_0 + L_1 + L_2$, and so on. It was proved in Holst (2008b) that the sequences $\mathcal{S}$ and $\mathcal{S}_P$ have the same distribution. We can identify $\mathcal{S}$ and $\mathcal{S}_P$ with each other.

Note that a point in Π_p originating from Π_{1p} corresponds to a double record in $\mathcal{S}$. As the number of points in Π_{1p} is Poisson with mean $\int_0^\infty p\mathrm{e}^{-t/a}\,\mathrm{d}t = ap$, it follows that, conditional on the Beta(a, b) random variable $P = p$, the number of double records, N_2, is Poisson with mean ap, as was proved in Móri (2001). Also, note that, conditional on $P = p$, the superposition $\Pi_{\geq 2p}$ of $\Pi_{2p}, \Pi_{3p}, \ldots$ is a Poisson process of intensity $1 - p\mathrm{e}^{-t/a}$ and independent of Π_{1p}.

We have, for $j = 0, 1, 2, \ldots$,

$$\begin{aligned}
\mathrm{P}(N_2 = j) &= \mathrm{E}(\mathrm{P}(N_2 = j \mid P)) \\
&= \mathrm{E}\biggl(\frac{(aP)^j \mathrm{e}^{-aP}}{j!}\biggr) \\
&= \int_0^1 \frac{(ap)^j \mathrm{e}^{-ap}}{j!} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1-p)^{b-1}\,\mathrm{d}p \\
&= \frac{a^j}{j!} \sum_{\ell=0}^{\infty} \frac{(-a)^\ell}{\ell!} \frac{a^{\overline{j+\ell}}}{(a+b)^{\overline{j+\ell}}}
\end{aligned}$$

and

$$\mathrm{E}(N_2) = \mathrm{E}(\mathrm{E}(N_2 \mid P)) = \mathrm{E}(aP) = \frac{a^2}{a+b}.$$

For the number of r consecutive records, $\mathrm{P}(N_r = k) = \sum_{j=0}^{\infty} \mathrm{P}(N_2 = j,\ N_r = k)$, where the probability $\mathrm{P}(N_2 = j,\ N_r = k)$ can, at least in principle, be found using the embedding above. As an illustration, we will study triple records and briefly quadruple records.

3.1. Triple records

First consider the probability of no triple record:

$$\mathrm{P}(N_3 = 0) = \mathrm{P}(N_2 \leq 1) + \sum_{j=2}^{\infty} \mathrm{P}(N_2 = j,\ N_3 = 0).$$

https://doi.org/10.1239/jap/1261670698 Published online by Cambridge University Press

Proposition 3.1. *For two or three double records and no triple record,*

$$\mathrm{P}(N_2 = 2,\ N_3 = 0) = \mathrm{P}(N_2 = 2) - \sum_{\ell=0}^{\infty} \mathrm{P}(N_2 = 2 + \ell) \frac{(\ell + 1)!}{(a + 1)^{\overline{\ell+1}}}$$

and

$$\begin{aligned}
&\mathrm{P}(N_2 = 3,\ N_3 = 0) \\
&= \mathrm{P}(N_2 = 3) - \sum_{\ell=0}^{\infty} \mathrm{P}(N_2 = 3 + \ell) \bigg(\frac{(\ell + 1)!}{(a + 1)^{\overline{\ell+1}}} + \frac{(\ell + 2)!}{(a + 2)^{\overline{\ell+1}}} - \frac{(\ell + 2)!\,(\ell + 1)}{(a + 1)^{\overline{\ell+2}}} \bigg).
\end{aligned}$$

Proof. Conditional on $P = p$, we can argue as follows. Given two fixed numbers $0 < u < v$, the number of points of $\Pi_{\geq 2p}$ in the interval (u, v) is Poisson with mean

$$\int_u^v (1 - p\mathrm{e}^{-t/a})\,\mathrm{d}t = v - u - ap(\mathrm{e}^{-u/a} - \mathrm{e}^{-v/a}).$$

The event '$N_2 = j$, $N_3 = 0$' means that the Poisson process Π_{1p} has j points and that between all adjacent points there is at least one point from $\Pi_{\geq 2p}$. Given that the j points of Π_{1p} occur at $t_1 < t_2 < \cdots < t_j$, the probability that $N_3 = 0$ is

$$\prod_{k=1}^{j-1} (1 - \exp(-t_{k+1} + t_k - ap\mathrm{e}^{-t_{k+1}/a} + ap\mathrm{e}^{-t_k/a})).$$

As the points of the Poisson process Π_{1p} occur with intensity $p\mathrm{e}^{-t/a}$, the probability of having no points of Π_{1p} outside a fixed discrete finite set is e^{-ap}. Hence,

$$\begin{aligned}
&\mathrm{P}(N_2 = j,\ N_3 = 0 \mid P = p) \\
&= \int \cdots \int \mathbf{1}(0 < t_1 < \cdots < t_j) \mathrm{e}^{-ap} p\mathrm{e}^{-t_1/a} \cdots p\mathrm{e}^{-t_j/a} \\
&\qquad \times \prod_{k=1}^{j-1} (1 - \exp(-t_{k+1} + t_k - ap\mathrm{e}^{-t_{k+1}/a} + ap\mathrm{e}^{-t_k/a}))\,\mathrm{d}t_1 \cdots \mathrm{d}t_j \\
&= \mathrm{e}^{-ap} \int_0^{ap} \int_0^{x_j} \cdots \int_0^{x_2} \prod_{k=1}^{j-1} \bigg(1 - \bigg(\frac{x_k}{x_{k+1}}\bigg)^a \mathrm{e}^{x_{k+1} - x_k}\bigg)\,\mathrm{d}x_1 \cdots \mathrm{d}x_j.
\end{aligned}$$

For $j = 2$, we find, with $u = x_1/x_2$, the series expansion and the beta function:

$$\begin{aligned}
&\mathrm{P}(N_2 = 2,\ N_3 = 0 \mid P = p) \\
&= \mathrm{e}^{-ap} \int_0^{ap} x_2 \int_0^1 (1 - u^a \mathrm{e}^{x_2(1-u)})\,\mathrm{d}u\,\mathrm{d}x_2 \\
&= \mathrm{e}^{-ap} \int_0^{ap} x_2 \bigg(1 - \sum_{\ell=0}^{\infty} \frac{x_2^\ell}{\ell!} \int_0^1 u^a (1 - u)^\ell\,\mathrm{d}u\bigg)\,\mathrm{d}x_2 \\
&= \mathrm{e}^{-ap} \int_0^{ap} x_2 \bigg(1 - \sum_{\ell=0}^{\infty} \frac{x_2^\ell}{(a + 1)^{\overline{\ell+1}}}\bigg)\,\mathrm{d}x_2 \\
&= \mathrm{P}(N_2 = 2 \mid P = p) - \sum_{\ell=0}^{\infty} \frac{\mathrm{P}(N_2 = 2 + \ell \mid P = p)(\ell + 1)!}{(a + 1)^{\overline{\ell+1}}},
\end{aligned}$$

which gives the unconditional probability in the assertion. With similar calculations we obtain the formula for $\mathrm{P}(N_2 = 3,\ N_3 = 0)$.

Analogously, it is possible to calculate $\mathrm{P}(N_2 = j,\ N_3 = 0)$ for $j = 4, 5, \ldots$.

At least two double records are needed to obtain a triple record. Therefore,

$$\mathrm{P}(N_3 = 1) = \sum_{j=2}^{\infty} \mathrm{P}(N_2 = j,\ N_3 = 1),$$

which can be calculated in a similar way as $\mathrm{P}(N_3 = 0)$. Note that

$$\mathrm{P}(N_2 = 2,\ N_3 = 1) = \mathrm{P}(N_2 = 2) - \mathrm{P}(N_2 = 2,\ N_3 = 0).$$

The integral

$$\mathrm{P}(N_2 = j,\ N_3 = 1 \mid P = p)$$
$$= \mathrm{e}^{-ap} \int_0^{ap} \int_0^{x_j} \cdots \int_0^{x_2} \sum_{\ell=1}^{j-1} \prod_{\substack{k=1 \\ k \neq \ell}}^{j-1} \biggl(1 - \biggl(\frac{x_k}{x_{k+1}}\biggr)^a \mathrm{e}^{x_{k+1}-x_k}\biggr) \biggl(\frac{x_\ell}{x_{\ell+1}}\biggr)^a \mathrm{e}^{x_{\ell+1}-x_\ell}\, \mathrm{d}x_1 \cdots \mathrm{d}x_j$$

has to be evaluated. Similar expressions can be given for $\mathrm{P}(N_3 = 2), \mathrm{P}(N_3 = 3), \ldots$.

3.2. Triple records for $a = 1$ and $b = 0$

For record probabilities in a sequence of independent, identically distributed, continuous random variables, we have $a = 1$ and $b = 0$. Using the above results, we find, after some computations, that

$$\mathrm{P}(N_2 = j) = \frac{\mathrm{e}^{-1}}{j!}, \qquad \mathrm{P}(N_2 \leq 1) = 0.735\,758\,88,$$
$$\mathrm{P}(N_2 = 2,\ N_3 = 0) = 0.066\,990\,05, \qquad \mathrm{P}(N_2 = 3,\ N_3 = 0) = 0.005\,047\,15,$$
$$\mathrm{P}(N_2 = 4,\ N_3 = 0) = 0.000\,207\,57, \qquad \mathrm{P}(N_2 = 5,\ N_3 = 0) = 0.000\,005\,32.$$

Summing these probabilities gives

$$\mathrm{P}(N_2 \leq 5,\ N_3 = 0) = 0.808\,008\,98.$$

By subtle asymptotic analyses of a recurrence and a differential equation, a complicated formula for the probability generating function of N_3 is given in Chern *et al.* (2000); see also Chern and Hwang (2005). From this, they obtained

$$\mathrm{P}(N_3 = 0) = 0.808\,009\,125\,346\,368\ldots.$$

For general $a > 0$ and $b \geq 0$, the methods used in Chern *et al.* are not applicable.

3.3. Quadruple records

To obtain a quadruple record, that is, four consecutive records, at least three double records are needed. Hence,

$$\mathrm{P}(N_4 = 0) = \mathrm{P}(N_2 \leq 2) + \sum_{j=3}^{\infty} \mathrm{P}(N_2 = j,\ N_4 = 0).$$

Proposition 3.2. *For three double records and no quadruple record,*

$$\mathrm{P}(N_2 = 3,\ N_4 = 0) = \mathrm{P}(N_2 = 3) - \sum_{\ell=0}^{\infty} \mathrm{P}(N_2 = 3 + \ell) \frac{(\ell+2)!\,(\ell+1)}{(a+1)^{\overline{\ell+2}}}.$$

Proof. As in the proof of Proposition 3.1, we see that

$$\begin{aligned}
&\mathrm{P}(N_2 = 3,\ N_4 = 1 \mid P = p) \\
&= \mathrm{e}^{-ap} \iiint \mathbf{1}(0 < x_1 < x_2 < x_3 < ap) \left(\frac{x_1}{x_3}\right)^a \mathrm{e}^{x_3 - x_1}\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3 \\
&= \mathrm{e}^{-ap} \int_0^{ap} \int_0^{x_3} \left(\frac{x_1}{x_3}\right)^a \mathrm{e}^{x_3 - x_1} (x_3 - x_1)\,\mathrm{d}x_1\,\mathrm{d}x_3 \\
&= \sum_{\ell=0}^{\infty} \frac{(ap)^{\ell+3}\mathrm{e}^{-ap}}{\ell!} \frac{1}{\ell+3} \int_0^1 u^a (1-u)^{\ell+1}\,\mathrm{d}u \\
&= \sum_{\ell=0}^{\infty} \frac{(ap)^{\ell+3}}{(\ell+3)!} \mathrm{e}^{-ap} \frac{(\ell+2)!\,(\ell+1)}{(a+1)^{\overline{\ell+2}}},
\end{aligned}$$

from which the assertion follows.

Similar formulae can be derived for $\mathrm{P}(N_2 = j,\ N_4 = 0)$ for $j = 4, 5, \dots$.

For $a = 1$ and $b = 0$, we find that $\mathrm{P}(N_2 \le 2) = 0.919\,70$, $\mathrm{P}(N_2 = 3,\ N_4 = 0) = 0.030\,97$, and, after computations similar to those for triple records, $\mathrm{P}(N_2 = 4,\ N_4 = 0) = 0.003\,91$. Thus, $\mathrm{P}(N_2 \le 4,\ N_4 = 0) = 0.954\,58$. Chern *et al.* (2000) obtained the value $\mathrm{P}(N_4 = 0) = 0.954\,90$. Recall that

$$1 - \mathrm{E}(N_4) = 1 - \frac{1}{3!\,3} = 0.944\,44 < \mathrm{P}(N_4 = 0) = 0.954\,90 < 1 - \frac{1}{4!} - \frac{1}{6!} = 0.956\,94;$$

see the end of Section 2.

References

Arratia, R., Barbour, A. D. and Tavaré, S. (1992). Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Prob.* **2,** 519–535.

Arratia, R., Barbour, A. D. and Tavaré, S. (2003). *Logarithmic Combinatorial Structures: A Probabilistic Approach.* European Mathematical Society, Zürich.

Chern, H.-H. and Hwang, H.-K. (2005). Limit distribution of the number of consecutive records. *Random Structures Algorithms* **26,** 404–417.

Chern, H.-H., Hwang, H.-K. and Yeh, Y.-N. (2000). Distribution of the number of consecutive records. *Random Structures Algorithms* **17,** 169–196.

Hahlin, L.-O. (1995). Double records. Res. Rep. 1995:12, Department of Mathematics, Uppsala University.

Holst, L. (2007). Counts of failure strings in certain Bernoulli sequences. *J. Appl. Prob.* **44,** 824–830.

Holst, L. (2008a). The number of two consecutive successes in a Hoppe–Pólya urn. *J. Appl. Prob.* **45,** 901–906.

Holst, L. (2008b). A note on embedding certain Bernoulli sequences in marked Poisson processes. *J. Appl. Prob.* **45,** 1181–1185.

Huffer, F., Sethuraman, J. and Sethuraman, S. (2009). A study of counts of Bernoulli strings via conditional Poisson processes. *Proc. Amer. Math. Soc.* **137,** 2125–2134.

Móri, T. F. (2001). On the distribution of sums of overlapping products. *Acta Sci. Math.* **67,** 833–841.