

RUNS IN COIN TOSSING: A GENERAL APPROACH FOR DERIVING DISTRIBUTIONS FOR FUNCTIONALS

LARS HOLST,* *Royal Institute of Technology*

TAKIS KONSTANTOPOULOS,** *Uppsala University*

Abstract

We take a fresh look at the classical problem of runs in a sequence of independent and identically distributed coin tosses and derive a general identity/recursion which can be used to compute (joint) distributions of functionals of run types. This generalizes and unifies already existing approaches. We give several examples, derive asymptotics, and pose some further questions.

Keywords: Regeneration; coin tossing; runs; longest run; Poisson approximation; Laplace transform; Rouché's theorem

2010 Mathematics Subject Classification: Primary 60G40

Secondary 60G50; 60F99; 60E10

1. Introduction

The tendency of ‘randomly occurring events’ to clump together is a well-understood chance phenomenon which has occupied people since the birth of probability theory. In tossing independent and identically distributed coins, we will, from time to time, see ‘long’ stretches of heads. The phenomenon has been studied and quantified extensively. For a bare-hands approach; see Erdős and Rényi [3], its sequel paper by Erdős and Révész [4], and the review paper by Révész [15].

We will consider a sequence $(\xi_n, n \in \mathbb{N})$ of Bernoulli random variables with $\mathbb{P}(\xi_n = 1) = p$, $\mathbb{P}(\xi_n = 0) = q = 1 - p$, and let

$$S(n) := \xi_1 + \cdots + \xi_n, \quad n \geq 1, \quad S(0) := 0.$$

Throughout this paper, a ‘run’ refers to an interval $I \subset \mathbb{N} := \{1, 2, \dots\}$ such that $\xi_n = 1$ for all $n \in I$ and there is no interval $J \supset I$ such that $\xi_n = 1$ for all $n \in J$. There has been an interest in computing the distribution of runs of various types such as the number of runs of a given length in n coin tosses. Feller [5, Section XIII.7] considered the probability that a run of a given length ℓ first appears at the n th coin toss and, using renewal theory, computed the distribution of the number of runs of a given length [5, Problem 26, Section XIII.12] as well as asymptotics [5, Problem 25, Section XIII.12]. (Warning: his definition of a run is slightly different.) He attributed this result to von Mises [16]. (In his classic work [17, p. 138], von Mises refers to a 1916 paper of the philosopher Karl Marbe who reported that in 200 000 birth registrations in a town in Bavaria there was only one ‘run’ of 17 consecutive births of children of the same sex.

Received 13 February 2014; revision received 2 July 2014.

* Postal address: Department of Mathematics, Royal Institute of Technology, SE-10044 Stockholm, Sweden.

Email address: lholst@kth.se

** Postal address: Department of Mathematics, Uppsala University, SE-75106 Uppsala, Sweden.

Email address: takiskonst@gmail.com

Note that $\log_2(200\,000) \approx 17.61$ and see Section 5 below.) Philippou and Makri [14] derived the joint distribution of the longest run and the number of runs of a given length. More detailed computations are considered in [11]. The literature is extensive and there are two books on the topic; see [2] and [6].

In this paper we take a more broad view. We study real- or vector-valued functionals of runs of various types and derive, using elementary methods, an equation which can be specified at will to provide a formula for the quantity of interest. To be more specific, let $R_\ell(n)$ be the number of runs of length ℓ in the first n coin tosses. Consider the vector

$$\mathbf{R}(n) := (R_1(n), R_2(n), \dots)$$

as an element of the set

$$\mathbb{Z}_+^* := \{x \in \mathbb{Z}_+^{\mathbb{N}} : x_k = 0 \text{ eventually}\},$$

which is identified with the set $\bigcup_{\ell=1}^{\infty} \mathbb{Z}_+^\ell$ of nonempty words from the alphabet of nonnegative integers, but, for the purpose of our analysis, it is preferable to append, to each finite word, an infinite sequence of zeros. This set is countable, and so the random variable $\mathbf{R}(n)$ has a discrete distribution. If $h: \mathbb{Z}_+^* \rightarrow \mathbb{R}^d$ is any function then we refer to the random variable $h(\mathbf{R}(n))$ as a d -dimensional functional of a run-vector. For example, for $d = 1$, if $h_1(x) = \sup\{\ell : x_\ell > 0\}$ (with $\sup \emptyset = 0$), then $h_1(\mathbf{R}(n))$ is the length of the longest run of heads in n coin tosses. If $h_2(x) := \sum_{\ell=1}^{\infty} \mathbf{1}_{\{x_\ell > 0\}}$, then $h_2(\mathbf{R}(n))$ is the total number of runs of any length in n coin tosses. Letting $d = 2$, we may consider $h(x) := (h_1(x), h_2(x))$ as a two-dimensional functional; a formula for the distribution of $h(\mathbf{R}(n))$ would then be a formula for the joint distribution of the number of runs of a given length together with the size of the longest run. It is useful to keep in mind that $\mathbb{Z}^* := \{x \in \mathbb{Z}^{\mathbb{N}} : x_k = 0, \text{ eventually}\}$ is a vector space and that \mathbb{Z}_+^* is a cone in this vector space. If $x, y \in \mathbb{Z}^*$ then $x + y$ is defined component-wise. The symbol 0 denotes the origin $(0, 0, \dots)$ of this vector space. For $j = 1, 2, \dots$, we let $e_j = (e_j(1), e_j(2), \dots) \in \mathbb{Z}^*$ be defined by

$$e_j(n) := \mathbf{1}_{\{n=j\}}, \quad n \in \mathbb{N}.$$

It is convenient, and logically compatible with the last display, to set

$$e_0 := (0, 0, \dots),$$

thus having two symbols for the origin of the vector space \mathbb{Z}^* .

The paper is organized as follows. Theorem 1 in Section 2 is a general formula for functionals of \mathbf{R}^* , defined as $\mathbf{R}(\cdot)$ stopped at an independent geometric time. We call this formula a ‘portmanteau identity’ because it contains many special cases of interest. To explain this, we provide, in the same section, formulae for specific functionals. In Section 3 we compute binomial moments and distribution of $G_\ell(n) := \sum_{k \geq \ell} R_k(n)$, the number of runs of length at least ℓ in n coin tosses. In particular, we point out its relationship with hypergeometric functions. In Section 4 we translate the portmanteau identity into a ‘portmanteau recursion’ which provides, for example, a method for recursive evaluation of the generating function of the random vector $\mathbf{R}(n)$. In Section 5 we take a closer look at the most common functional of $\mathbf{R}(n)$, namely the length $L(n)$ of the longest run in n coin tosses. We discuss the behavior of its distribution function and its relation to a Poisson approximation theorem, given in Proposition 2 stating that $R_{\ell_1}(n), \dots, R_{\ell_\nu}(n)$ become asymptotically independent Poisson random variables as $n \rightarrow \infty$ when, simultaneously, $\ell_1, \dots, \ell_\nu \rightarrow \infty$. A second approximation for the distribution function $\mathbb{P}(L(n) < \ell)$ of $L(n)$, which works well at small values of ℓ , is obtained in Section 5.2, using complex analysis. We numerically compare the two approximations in Section 5.3 and

finally pose some further questions in the last section. Although, in this paper, our method has been applied to finding very detailed information about the distribution function of the specific functional $L(n)$, many other functionals, mentioned above and in Section 2, can be treated analogously if detailed information about their distribution function is desired.

2. A portmanteau identity

Let N^* be a geometric random variable

$$\mathbb{P}(N^* = n) = w^{n-1}(1 - w), \quad n \in \mathbb{N},$$

independent of the sequence ξ_1, ξ_2, \dots . We let

$$R^* := R(N^* - 1).$$

Thus, R^* is a random element of \mathbb{Z}_+^* which is distributed like $R(n)$ with probability $w^n(1 - w)$ for $n = 0, 1, \dots$. Note that $R(0) = (0, 0, \dots)$, which is consistent with our definitions. To save some space, we use the abbreviations

$$\alpha := wp, \quad \beta := wq, \quad \gamma := 1 - w, \tag{1}$$

throughout the paper, noting that if α, β, γ are three nonnegative real numbers adding up to 1 with γ strictly positive, then $w, p, q = 1 - p$ are uniquely determined.

Theorem 1. *For any $h: \mathbb{Z}_+^* \rightarrow \mathbb{R}$ such that $\mathbb{E}h(R^*)$ is defined, we have the Stein–Chen type of identity*

$$\mathbb{E}h(R^*) = \gamma \sum_{j \geq 0} \alpha^j h(e_j) + \beta \sum_{j \geq 0} \alpha^j \mathbb{E}h(R^* + e_j). \tag{2}$$

Proof. The equation becomes apparent if we think probabilistically, using an ‘explosive coin’. Consider a usual coin (think of a British pound coin—a British pound coin is sufficiently thick so that the chance of landing on its edge is nonnegligible, especially at the hands of a skilled coin tosser; if a US (thinner) nickel is used then the chance of landing on its edge is estimated to be $\frac{1}{6000}$ [12]) but equip it with an explosive mechanism which is activated if the coin touches the ground on its edge. An explosion occurs with probability $\gamma = 1 - w$. When explosion occurs the coin is destroyed immediately. As long as explosion does not occur then the coin lands heads or tails, as usual. Clearly, $\alpha = wp$ is the probability that we observe heads and $\beta = wq$ is the probability that we observe tails. We let $\mathbf{E}, \mathbf{H}, \mathbf{T}$ denote ‘explosion’, ‘heads’, ‘tails’, respectively, for the explosive coin. The possible outcomes in tossing such a coin comprise the set

$$\Omega^* := \bigcup_{k \geq 0} \{\mathbf{H}, \mathbf{T}\}^k \times \{\mathbf{E}\}.$$

Indeed, the repeated tossing of an explosive coin results in an explosion (which may happen immediately), in which case the coin is destroyed. We can then naturally define R^* on Ω^* . Let $\mathbf{H}^j\mathbf{E} \subset \Omega^*$ be an abbreviation for the event of seeing heads j times followed by explosion. Similarly, for $\mathbf{H}^j\mathbf{T}$. Clearly, $\Omega^* = \bigcup_{j \geq 0} (\mathbf{H}^j\mathbf{E}) \cup \bigcup_{j \geq 0} (\mathbf{H}^j\mathbf{T})$ and all events involved in the union are mutually disjoint. Hence,

$$\mathbb{E}h(R^*) = \sum_{j \geq 0} \mathbb{E}[\mathbf{H}^j\mathbf{E}; h(R^*)] + \sum_{j \geq 0} \mathbb{E}[\mathbf{H}^j\mathbf{T}; h(R^*)],$$

where, as usual, $\mathbb{E}[A; Y] := \mathbb{E}[\mathbf{1}_A Y]$, if A is an event and Y a random variable. For $j \geq 0$, on the event $H^j E$, we have $R^* = e_j$. Hence, $\mathbb{E}[H^j E; h(R^*)] = \alpha^j \gamma h(e_j)$. On the event $H^j T$, we have $R^* = e_j + \theta^{j+1} R^*$, where $\theta^{j+1} R^* = (R_{j+1}^*, R_{j+2}^*, \dots)$, which is independent and identical in law to R^* . Hence, $\mathbb{E}[H^j T; h(R^*)] = \alpha^j \beta \mathbb{E}h(e_j + R^*)$.

The easiest way to see that the identity we just proved actually characterizes the law of R^* is by direct computation. If $x \in \mathbb{Z}_+^*$, we let

$$z^x := z_1^{x_1} z_2^{x_2} \dots,$$

for any sequence z_1, z_2, \dots of real or complex numbers such that $z_k \neq 0$ for all k . (This product is a finite product, by definition of \mathbb{Z}_+^* .)

Theorem 2. *There is a unique (in law) random element R^* of \mathbb{Z}_+^* such that (2) holds for all nonnegative h . For this R^* , we have*

$$\mathbb{E}_{z^{R^*}} = \gamma \frac{1 + \sum_{j \geq 1} \alpha^j z_j}{1 - \beta - \beta \sum_{j \geq 1} \alpha^j z_j}.$$

Moreover, for any $\ell \in \mathbb{N}$, the law of (R_1^*, \dots, R_ℓ^*) is specified by

$$\mathbb{E}_{z_1^{R_1^*} \dots z_\ell^{R_\ell^*}} = \gamma \frac{1 + \sum_{j=1}^{\ell} \alpha^j z_j + \sum_{j>\ell} \alpha^j}{1 - \beta - \beta \sum_{j=1}^{\ell} \alpha^j z_j - \beta \sum_{j>\ell} \alpha^j}.$$

Proof. Let $h(x) := z^x$ in (2). Then $h(e_j) = z_j$, and $h(R^* + e_j) = z_j h(R^*)$. Substituting into (2) gives the result. Taking $z_j = 1$ for all $j \geq \ell$ gives the second formula.

We can now derive distributions of various functionals of R^* quite easily. For example, to deal with the one-dimensional marginals of R^* , set $z_\ell = \theta$ and let $z_k = 1$ for $k \neq \ell$:

$$\mathbb{E} \theta^{R_\ell^*} = \gamma \frac{1 + \sum_{j \neq \ell} \alpha^j + \alpha^\ell \theta}{1 - \beta - \beta \sum_{j \neq \ell} \alpha^j - \beta \alpha^\ell \theta}. \tag{3}$$

This is a geometric-type distribution (with mass at 0), and we give it a name for our convenience.

Definition 1. For $0 \leq \alpha, \beta \leq 1$ let $\text{geo}(\alpha, \beta)$ denote the probability measure Q on $\mathbb{Z}_+ \cup \{+\infty\} = \{0, 1, \dots, +\infty\}$ with

$$Q\{0\} = \alpha, \quad Q\{n\} = (1 - \alpha)(1 - \beta)\beta^{n-1} \quad \text{for } n \geq 1.$$

For example, N^* has $\text{geo}(0, w)$ distribution and $N^* - 1$ has $\text{geo}(1 - w, w)$ distribution. Abusing notation and letting $\text{geo}(\alpha, \beta)$ denote a random variable with the same law, we easily see that

$$\begin{aligned} \mathbb{E} \text{geo}(1 - \beta, \beta) &= \frac{\beta}{1 - \beta}, & \mathbb{E} \binom{\text{geo}(\alpha, \beta)}{r} &= \frac{1 - \alpha}{\beta} \left(\frac{\beta}{1 - \beta} \right)^r \quad \text{for } r \geq 1, \\ \mathbb{E} \theta^{\text{geo}(\alpha, \beta)} &= \frac{\alpha + (1 - \alpha - \beta)\theta}{1 - \beta\theta} = \left[1 - \frac{1 - \alpha - \beta}{1 - \beta} (1 - \theta) \right] \left[1 + \frac{\beta}{1 - \beta} (1 - \theta) \right]^{-1}. \end{aligned}$$

Therefore, comparing with (3), we have the following corollary.

Corollary 1. *It holds that R_ℓ^* has $\text{geo}(\alpha_\ell, \beta_\ell)$ distribution with*

$$\alpha_\ell = \gamma \frac{1 + \sigma_\ell}{1 - \beta - \beta\sigma_\ell}, \quad \beta_\ell = \frac{\beta\alpha^\ell}{1 - \beta - \beta\sigma_\ell},$$

where $\sigma_\ell := \sum_{j \geq 1, j \neq \ell} \alpha^j$.

As a reality check, observe that $\mathbb{E}R_\ell^* = \beta_\ell / (1 - \beta_\ell) = (1 - \alpha)^2 \alpha^\ell / \gamma$ and so $\sum_{\ell=1}^\infty \ell \mathbb{E}R_\ell^* = \alpha / \gamma$. On the other hand, $\sum_{\ell=1}^\infty \ell R_\ell^* = S(N^* - 1)$. Since $S(n)$ is binomial and N^* is independent geometric, we have, by elementary computation, $S(N^* - 1) \sim \text{geo}(\gamma / (1 - \beta), \alpha / (1 - \beta))$ and so $\mathbb{E}S(N^* - 1) = \alpha / \gamma$, agreeing with the above.

As another example, consider the following functional $\bar{\lambda}: \mathbb{Z}_+^* \rightarrow \mathbb{R}$:

$$\bar{\lambda}(x) = \sup\{i > 0: x_i > 0\}.$$

Corollary 2. *Let $L^* := \bar{\lambda}(R^*)$ be the longest run in $N^* - 1$ coin tosses. Then*

$$\mathbb{P}(L^* < \ell) = \frac{\gamma(1 - \alpha^\ell)}{\gamma + \beta\alpha^\ell}, \quad \ell \in \mathbb{N}.$$

Proof. With 0 denoting the zero element of \mathbb{Z}_+^* , we have $\bar{\lambda}(0) = 0$ since $\sup \emptyset = 0$. Also

$$\bar{\lambda}(x + e_j) = \bar{\lambda}(x) \vee j \quad \text{for } j \geq 0, x \in \mathbb{Z}_+^*.$$

Fix $\ell \in \mathbb{N}$ and use (2) with $h(x) := \mathbf{1}_{\{\bar{\lambda}(x) < \ell\}}$. Then $\mathbb{P}(L^* < \ell) = \mathbb{E}h(R^*)$. Since $h(x + e_j) = \mathbf{1}_{\{\bar{\lambda}(x) \vee j < \ell\}} = h(x) \mathbf{1}_{\{j < \ell\}}$, we have $\mathbb{E}h(R^* + e_j) = \mathbb{P}(L^* < \ell) \mathbf{1}_{\{j < \ell\}}$. Substituting into (2), we have

$$\mathbb{P}(L^* < \ell) = \gamma \sum_{j \geq 0} \alpha^j \mathbf{1}_{\{j < \ell\}} + \beta \sum_{j \geq 0} \alpha^j \mathbf{1}_{\{j < \ell\}} \mathbb{P}(L^* < \ell),$$

which completes the proof.

See also [9, Section 5.12, Problems 46 and 47] for an alternative way of obtaining the distribution of L^* .

Alternatively, we can look at the functional

$$\underline{\lambda}(x) := \inf\{i > 0: x_i > 0\},$$

which takes value $+\infty$ at the origin of \mathbb{Z}_+^* , but this poses no difficulty.

Corollary 3. *Let $\underline{\lambda}(R^*)$ be the run of least length in $N^* - 1$ coin tosses. Then*

$$\mathbb{P}^*(\underline{\lambda}(R^*) \geq \ell) = \frac{\gamma(1 - \alpha + \alpha^\ell)}{\gamma - \beta(1 - \alpha + \alpha^\ell)}, \quad \ell \in \mathbb{N}.$$

The random variable $\underline{\lambda}(R^*)$ is defective with $\mathbb{P}(\underline{\lambda}(R^*) = \infty) = \gamma(1 - \alpha) / (\gamma - \beta(1 - \alpha))$.

Proof. Fix $\ell \in \mathbb{N}$ and let $h(x) = \mathbf{1}_{\{\underline{\lambda}(x) \geq \ell\}}$ in (2). We compute that $h(0) = 1$ and, for $j \in \mathbb{N}$, $h(e_j) = j, h(x + e_j) = h(x) \mathbf{1}_{\{j \geq \ell\}}$. The rest is elementary algebra.

Corollary 4. *If $h: \mathbb{Z}^* \rightarrow \mathbb{R}$ is a linear function then*

$$\mathbb{E}h(R^*) = \frac{(1 - \alpha)^2}{\gamma} \sum_{j \geq 0} \alpha^j h(e_j). \tag{4}$$

As another example of the versatility of the portmanteau formula, we specify the joint distribution of finitely many components of R^* together with L^* .

Corollary 5. *It holds that*

$$\mathbb{E}\left[z_1^{R_1^*} \cdots z_{\ell-1}^{R_{\ell-1}^*}; L^* < \ell\right] = \gamma \frac{1 + \sum_{j=1}^{\ell-1} \alpha^j z_j}{1 - \beta - \beta \sum_{j=1}^{\ell-1} \alpha^j z_j}. \tag{5}$$

Proof. Let $h(x) = z_1^{x_1} \cdots z_{\ell-1}^{x_{\ell-1}} \mathbf{1}_{\{\bar{\lambda}(x) < \ell\}}$ in (2). Then $h(0) = 0$, $h(e_j) = z_j \mathbf{1}_{\{j < \ell\}}$, and $h(x + e_j) = h(x)h(e_j)$, $j \in \mathbb{N}$. Again, substitution into (2) and simple algebra yields the result.

For verification, note that taking $\ell \rightarrow \infty$ in (5) gives (4) for $\mathbb{E}_z^{R^*}$, while letting $z_1 = \cdots = z_{\ell-1} = 1$ gives the result of Corollary 2.

The joint moments and binomial moments of the components of R^* can be computed explicitly.

Corollary 6. *Consider positive integers v, ℓ_1, \dots, ℓ_v , and nonnegative integers r_1, \dots, r_v , such that $r_0 := r_1 + \cdots + r_v \geq 1$. Let $\ell := (\ell_1, \dots, \ell_v)$ and $\mathbf{r} := (r_1, \dots, r_v)$ and set $\ell \cdot \mathbf{r} = \ell_1 r_1 + \cdots + \ell_v r_v$. Then*

$$\mathbb{E}_{z_1}^{R_{\ell_1}^*} \cdots \mathbb{E}_{z_v}^{R_{\ell_v}^*} = \frac{1 + (1 - \alpha) \sum_{j=1}^v \alpha^{\ell_j} (z_j - 1)}{1 - ((1 - \alpha)\beta/\gamma) \sum_{j=1}^v \alpha^{\ell_j} (z_j - 1)} \tag{6}$$

and

$$\mathbb{E}\binom{R_{\ell_1}^*}{r_1} \cdots \binom{R_{\ell_v}^*}{r_v} = \binom{r_0!}{r_1! \cdots r_v!} \left(\frac{\alpha^{\ell \cdot \mathbf{r}} \beta^{r_0-1} (1 - \alpha)^{r_0+1}}{\gamma^{r_0}} \right). \tag{7}$$

Proof. By Theorem 1,

$$\begin{aligned} & \mathbb{E}_{z_1}^{R_{\ell_1}^*} \cdots \mathbb{E}_{z_v}^{R_{\ell_v}^*} \\ &= \gamma \left(\sum_{j=1}^v \alpha^{\ell_j} z_j + \sum_{j \notin \{\ell_1, \dots, \ell_v\}} \alpha^j \right) + \beta \left(\sum_{j=1}^v \alpha^{\ell_j} z_j + \sum_{j \notin \{\ell_1, \dots, \ell_v\}} \alpha^j \right) \mathbb{E}_{z_1}^{R_1^*} \cdots \mathbb{E}_{z_v}^{R_v^*} \\ &= \gamma \left(\sum_{j=1}^v \alpha^{\ell_j} (z_j - 1) + \sum_{j \geq 0} \alpha^j \right) + \beta \left(\sum_{j=1}^v \alpha^{\ell_j} (z_j - 1) + \sum_{j \geq 0} \alpha^j \right) \mathbb{E}_{z_1}^{R_1^*} \cdots \mathbb{E}_{z_v}^{R_v^*}, \end{aligned}$$

from which (6) follows. Expanding the denominator in (6), we obtain

$$\begin{aligned} \mathbb{E}_{z_1}^{R_{\ell_1}^*} \cdots \mathbb{E}_{z_v}^{R_{\ell_v}^*} &= \left(1 + (1 - \alpha) \sum_{j=1}^v \alpha^{\ell_j} (z_j - 1) \right) \sum_{k=0}^{\infty} \left(\frac{(1 - \alpha)\beta}{\gamma} \right)^k \left(\sum_{j=1}^v \alpha^{\ell_j} (z_j - 1) \right)^k \\ &= 1 + \sum_{k=1}^{\infty} \left(\frac{(1 - \alpha)\beta}{\gamma} \right)^k \left(\sum_{j=1}^v \alpha^{\ell_j} (z_j - 1) \right)^k \\ &\quad + \frac{\gamma}{\beta} \sum_{k=1}^{\infty} \left(\frac{(1 - \alpha)\beta}{\gamma} \right)^k \left(\sum_{j=1}^v \alpha^{\ell_j} (z_j - 1) \right)^k \end{aligned}$$

$$= 1 + \frac{1 - \alpha}{\beta} \sum_{k=1}^{\infty} \left(\frac{(1 - \alpha)\beta}{\gamma} \right)^k \times \sum_{\substack{i_1, \dots, i_v \\ i_1 + \dots + i_v = k}} \frac{k!}{i_1! \dots i_v!} \alpha^{\ell_1 i_1 + \dots + \ell_v i_v} (z_1 - 1)^{i_1} \dots (z_v - 1)^{i_v}.$$

Now,

$$\begin{aligned} \mathbb{E}z_1^{R_{\ell_1}^*} \dots z_v^{R_{\ell_v}^*} &= \mathbb{E}(1 + (z_1 - 1))^{R_{\ell_1}^*} \dots (1 + (z_v - 1))^{R_{\ell_v}^*} \\ &= \sum_{i_1, \dots, i_v} \mathbb{E} \binom{R_{\ell_1}^*}{r_1} \dots \binom{R_{\ell_v}^*}{r_v} (z_1 - 1)^{i_1} \dots (z_v - 1)^{i_v} \end{aligned}$$

and so (7) is obtained by inspection.

It is of interest to study the distribution of the number of runs exceeding a given length (see [2] and [11]),

$$G_{\ell}(n) := \sum_{k \geq \ell} R_k(n).$$

Consider the \mathbb{Z}_+^* -valued random variable

$$G(n) := (G_1(n), G_2(n), \dots).$$

We work up to a geometric random variable. Thus, let

$$G^* := G(N^* - 1).$$

We can compute $\mathbb{E}z^{G^*}$ easily from the first equation of Theorem 2 by replacing z_j with $z_1 \dots z_j$.

Corollary 7. *It holds that*

$$\mathbb{E}z^{G^*} = \gamma \frac{1 + \sum_{j \geq 1} \alpha^j z_1 \dots z_j}{1 - \beta - \beta \sum_{j \geq 1} \alpha^j z_1 \dots z_j}.$$

Marginalizing, we have the following corollary.

Corollary 8. *It holds that G_{ℓ}^* has $\text{geo}(\tilde{\alpha}_{\ell}, \tilde{\beta}_{\ell})$ distribution with $\tilde{\alpha}_{\ell} = \gamma(1 - \alpha^{\ell})/(\gamma + \beta\alpha^{\ell})$, $\tilde{\beta}_{\ell} = \beta\alpha^{\ell}/(\gamma + \beta\alpha^{\ell})$.*

3. Number of runs of given (or exceeding a given) length in n coin tosses

Our interest next is in obtaining information about the distributions of $R_{\ell}(n)$ and $G_{\ell}(n)$. Since R_{ℓ}^* and G_{ℓ}^* are both of $\text{geo}(\alpha, \beta)$ type with explicitly known parameters, and since (if X is a random variable, we let $\mathcal{L}\{X\}$ be its law)

$$\mathcal{L}\{G_{\ell}^*\} = (1 - w) \sum_{n \geq 0} w^n \mathcal{L}\{G(n)\}$$

(likewise for R_{ℓ}^*), the problem is, in principle, solved. Moreover, such formulae exist in the literature; see, e.g. [2] and [13]. Our intent in this section is to provide an independent derivation of the formulae but also point out their relations with hypergeometric functions.

It turns out that (i) equations for $G_\ell(n)$ are simpler than those for $R_\ell(n)$ and (ii) binomial moments for both variables are simpler to derive than moments. Therefore, we start by computing the r th binomial moment of $G_\ell(n)$. By Corollary 8, G_ℓ^* is a $\text{geo}(\tilde{\alpha}_\ell, \tilde{\beta}_\ell)$ random variable and from the equations following Definition 1, we have

$$\begin{aligned} \mathbb{E}\binom{G_\ell^*}{r} &= \frac{1 - \tilde{\alpha}_\ell}{\tilde{\beta}_\ell} \left(\frac{\tilde{\beta}_\ell}{1 - \tilde{\beta}_\ell} \right)^r = \frac{1 - \alpha}{\beta} \left(\frac{\beta\alpha^\ell}{\gamma} \right)^r \\ &= (1 - wp)(wq)^{r-1}(wp)^{\ell r}(1 - w)^{-r} \\ &= (1 - w)p^{\ell r}q^{r-1} \times \underbrace{(1 - wp)w^{\ell r+r-1}(1 - w)^{-r-1}}. \end{aligned}$$

Now use the Taylor expansion

$$(1 - w)^{-r-1} = \sum_{k=0}^{\infty} \binom{r+k}{r} w^k \tag{8}$$

to express the under-braced term above as

$$\begin{aligned} (1 - wp) \sum_{k=0}^{\infty} \binom{r+k}{r} w^{\ell r+r-1+k} &= \sum_{k=0}^{\infty} \binom{r+k}{r} w^{\ell r+r-1+k} - p \sum_{k=0}^{\infty} \binom{r+k}{r} w^{\ell r+r+k} \\ &= \sum_n \binom{n+1-\ell r}{r} w^n - p \sum_n \binom{n-\ell r}{r} w^n \\ &= \sum_n \left[\binom{n+1-\ell r}{r} - p \binom{n-\ell r}{r} \right] w^n. \end{aligned}$$

So, by inspection

$$\mathbb{E}\binom{G_\ell(n)}{r} = p^{\ell r}q^{r-1} \left[\binom{n+1-\ell r}{r} - p \binom{n-\ell r}{r} \right]. \tag{9}$$

In particular, we have

$$\mathbb{E}G_\ell(n) = p^\ell [(n - \ell + 1) - p(n - \ell)] \quad \text{for } n \geq \ell$$

and, since $R_\ell(n) = G_\ell(n) - G_\ell(n + 1)$,

$$\mathbb{E}R_\ell(n) = p^\ell [(n - \ell + 1) - 2(n - \ell)p + (n - \ell - 1)p^2] \quad \text{for } n > \ell,$$

while $\mathbb{E}R_n(n) = p^n$. Note that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}R_\ell(n) = p^\ell q^2$$

as expected by the ergodic theorem.

We now use the following standard formula relating probabilities to binomial moments (the binomial coefficient $\binom{a}{b}$ is taken to be zero if $b > a$ or if $a < 0$):

$$\mathbb{P}(G_\ell(n) = x) = \sum_{r \geq x} (-1)^{r-x} \binom{r}{x} \mathbb{E}\binom{G_\ell(n)}{r}. \tag{10}$$

Substituting the equation for the binomial moment and changing variable from $r \geq x$ to $m = r - x \geq 0$, we obtain

$$\begin{aligned} \mathbb{P}(G_\ell(n) = x) &= \sum_{m \geq 0} (-1)^m \binom{x+m}{x} p^{\ell(x+m)} q^{x+m-1} \\ &\quad \times \left[\binom{n+1-\ell(x+m)}{x+m} - p \binom{n-\ell(x+m)}{x+m} \right] \\ &= p^{\ell x} q^{x-1} \left[\sum_{m \geq 0} (-p^\ell q)^m \binom{x+m}{x} \binom{n+1-\ell(x+m)}{x+m} \right. \\ &\quad \left. - p \sum_{m \geq 0} (-p^\ell q)^m \binom{x+m}{x} \binom{n-\ell(x+m)}{x+m} \right]. \end{aligned}$$

It is interesting to note the relation of the distribution of $G_\ell(n)$ to hypergeometric functions. Recall the notion of the hypergeometric function [8, Section 5.5] (the notation is from this book and is not standard):

$$F \left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z \right) = \sum_{k \geq 0} \frac{a_1^{\bar{k}} \dots a_m^{\bar{k}} z^k}{b_1^{\bar{k}} \dots b_n^{\bar{k}} k!},$$

where $m, n \in \mathbb{Z}_+$, $a_1, \dots, a_m \in \mathbb{C}$, $b_1, \dots, b_n \in \mathbb{C} \setminus \{0, -1, -2, \dots\}$, $z \in \mathbb{C}$, and $x^{\bar{k}} := x(x+1) \dots (x+k-1)$. A little algebra gives

$$\begin{aligned} H_\ell(x, y; z) &:= \sum_{m \geq 0} z^m \binom{x+m}{x} \binom{x+y-\ell m}{x+m} \\ &= \binom{x+y}{x} F \left(\begin{matrix} \mathbf{V}_{\ell+1}(y) \\ \mathbf{V}_\ell(x+y) \end{matrix} \middle| -\frac{(\ell+1)^{\ell+1}}{\ell^\ell} z \right), \end{aligned} \tag{11}$$

where $\mathbf{V}_{\ell+1}(y)$ and $\mathbf{V}_\ell(x+y)$ denote arrays of sizes $\ell+1$ and ℓ , respectively, defined via

$$\mathbf{V}_k(u) := -\frac{1}{k}(u, u-1, \dots, u-k+1).$$

Returning to (10), we recognize that the two terms in the bracket are expressible in terms of the function H_ℓ ,

$$\mathbb{P}(G_\ell(n) = x) = p^{\ell x} q^{x-1} [H_\ell(x, n+1 - (\ell+1)x - p^\ell q) - H_\ell(x, n - (\ell+1)x - p^\ell q)].$$

The point is that the probabilities $\mathbb{P}(G_\ell(n) = x)$ are expressible in terms of the function H_ℓ which is itself expressible in terms of a hypergeometric function as in (11). Hypergeometric functions are efficiently computable via computer algebra systems (we use MAPLE®).

Ultimately, the hypergeometric functions appearing above are nothing but polynomials. So the problem is, by nature, of combinatorial character. Instead of digging in the literature for recursions for these functions, we prefer to transform the portmanteau identity into a recursion which can be specialized and iterated.

4. Portmanteau recursions in the time domain

Recall the identity (2). We pass from ‘frequency domain’ (variable ‘ w ’) to ‘time domain’ (variable ‘ n ’), we do obtain a veritable recursion in the space \mathbb{Z}_+^* . Recalling that α, β, γ are

given by (1) and that $\mathcal{L}\{h(R^*)\} = \sum_{n \geq 0} (1-w)w^n \mathcal{L}\{h(\mathbf{R}(n))\}$, we take each of the terms in (2) and bring out its dependence on w explicitly. The left-hand side of (2) is

$$\mathbb{E}h(R^*) = (1-w) \sum_{n \geq 0} w^n \mathbb{E}h(\mathbf{R}(n)). \tag{12}$$

The first term on the right-hand side of (2) is

$$\gamma \sum_{n \geq 0} \alpha^j h(e_n) = (1-w) \sum_{n \geq 0} w^n p^n h(e_n). \tag{13}$$

As for the second term of (2), we have

$$\begin{aligned} \beta \sum_{j \geq 0} \alpha^j \mathbb{E}h(R^* + e_j) &= wq \sum_{j \geq 0} w^j p^j (1-w) \sum_{n \geq 0} w^n \mathbb{E}h(\mathbf{R}(n) + e_j) \\ &= (1-w)q \sum_{j \geq 0} \sum_{n \geq 0} w^{1+j+n} p^j \mathbb{E}h(\mathbf{R}(n) + e_j). \end{aligned}$$

Change variables by $(j, n) \mapsto (j, m = 1 + j + n)$ to further write

$$\begin{aligned} \beta \sum_{j \geq 0} \alpha^j \mathbb{E}h(R^* + e_j) &= (1-w)q \sum_{m \geq 0} \sum_{0 \leq j \leq m-1} w^m p^j \mathbb{E}h(\mathbf{R}(m-j-1) + e_j) \\ &= (1-w) \sum_{n \geq 0} w^n q \sum_{0 \leq j \leq n-1} p^j \mathbb{E}h(\mathbf{R}(n-j-1) + e_j). \end{aligned} \tag{14}$$

Using (2) and (12)–(14), we obtain the following theorem.

Theorem 3. *Let $h : \mathbb{Z}_+^* \rightarrow \mathbb{R}$ be any function. Then, for all $n \in \mathbb{N}$,*

$$\mathbb{E}h(\mathbf{R}(n)) = q \sum_{j=0}^{n-1} p^j \mathbb{E}h(\mathbf{R}(n-j-1) + e_j) + p^n h(e_n). \tag{15}$$

Remark 1. (i) We say ‘any function’ because $\mathbf{R}(n)$ takes finitely many values for all n .

(ii) This is a linear recursion but, as expected, it does not have bounded memory.

(iii) It can easily be programmed. It is initialized with $\mathbb{E}h(\mathbf{R}(0)) = h(0)$.

(iv) Of course, this recursion is nothing else but ‘explicit counting’.

(v) One could provide an independent proof of Theorem 3 and obtain the result of Theorem 1. This is a matter of taste.

(vi) We instructed MAPLE to run the recursion a few times and here are the results:

$$\begin{aligned} \mathbb{E}h(\mathbf{R}(1)) &= qh(0) + ph(e_1), \\ \mathbb{E}h(\mathbf{R}(2)) &= q^2h(0) + 2qph(e_1) + p^2h(e_2), \\ \mathbb{E}h(\mathbf{R}(3)) &= q^3h(0) + 3q^2ph(e_1) + qp^2h(2e_1) + 2qp^2h(e_2) + p^3h(e_3), \\ \mathbb{E}h(\mathbf{R}(4)) &= q^4h(0) + 4q^3ph(e_1) + 3q^2p^2h(2e_1) + 3q^2p^2h(e_2) + 2qp^3h(e_1 + e_2), \\ &\quad + 2qp^3h(e_3) + p^4h(e_4), \end{aligned}$$

which could be interpreted combinatorially (it may be worth carrying out the combinatorial approach further).

Since $G(n) = \sigma(\mathbf{R}(n))$, where $\sigma : \mathbb{Z}_+^* \rightarrow \mathbb{Z}_+^*$ is given by

$$\sigma(x)_k := \sum_{j \geq k} x_j,$$

if $f : \mathbb{Z}_+^* \rightarrow \mathbb{R}$ is any function, then letting $h = f \circ \sigma$ in the recursion of Theorem 3 and noting that $\sigma(e_n) = e_1 + \dots + e_n$, we have the following corollary.

Corollary 9. *Let $f : \mathbb{Z}_+^* \rightarrow \mathbb{R}$ be any function. Then, for all $n \in \mathbb{N}$,*

$$\mathbb{E}f(G(n)) = q \sum_{j=0}^{n-1} p^j \mathbb{E}f(G(n-j-1) + e_j) + p^n f(e_1 + \dots + e_n).$$

These two recursions can be transformed into recursions for probability generating functions. Recalling that $z^x = \prod_{j \geq 1} z_j^{x_j}$ for $x \in \mathbb{Z}^*$, we consider

$$\Phi_n(z) := \mathbb{E}z^{\mathbf{R}(n)}, \quad \Psi_n(z) := \mathbb{E}z^{G(n)}, \quad z \in \mathbb{C}^{\mathbb{N}},$$

and immediately obtain the following corollary.

Corollary 10. *The probability generating functions Φ_n and Ψ_n of the random elements $\mathbf{R}(n)$ and $G(n)$, respectively, of \mathbb{Z}_+^* satisfy $\Phi_0(z) = \Psi_0(z) = 1$, and, for $n \in \mathbb{N}$,*

$$\begin{aligned} \Phi_n(z) &= q\Phi_{n-1}(z) + q \sum_{1 \leq j \leq n-1} p^j z_j \Phi_{n-j-1}(z) + p^n z_n, \\ \Psi_n(z) &= q\Psi_{n-1}(z) + q \sum_{1 \leq j \leq n-1} p^j z_1 \cdots z_j \Psi_{n-j-1}(z) + p^n z_1 \cdots z_n. \end{aligned}$$

Let us now look at $G_\ell(n)$. Consider the probability generating function

$$\Psi_{n,\ell}(\theta) := \mathbb{E}\theta^{G_\ell(n)}, \quad \theta \in \mathbb{C}.$$

Clearly, $\Psi_{n,\ell}(\theta) = 1$ for $n < \ell$ and $\Psi_{n,\ell}(\theta) = \Psi_n(1 + (\theta - 1)e_\ell)$, where $1 \in \mathbb{Z}^{\mathbb{N}}$ is the infinite repetition of ones. We thus have the following corollary.

Corollary 11. *For $n < \ell$, we have $\Psi_{n,\ell}(\theta) = 1$, and*

$$\Psi_{n,\ell} = \begin{cases} q \sum_{j=0}^{n-\ell-1} p^j \Psi_{n-j-1,\ell} + p^{n-\ell} + (\theta - 1)p^\ell, & \ell \leq n \leq 2\ell, \\ q \sum_{j=0}^{\ell-1} p^j \Psi_{n-j-1,\ell} + q\theta \sum_{j=\ell}^{n-\ell-1} p^j \Psi_{n-j-1,\ell} \\ \quad + q\theta(p^{n-\ell} - p^\ell) + \theta p^m, & n \geq 2\ell + 1. \end{cases}$$

5. Longest run, Poisson, and other approximations

Recall that $L(n) = \bar{\lambda}(\mathbf{R}(n))$ is the length of the longest run in n coin tosses. Although there is an explicit equation (see Corollary 2) for

$$(1 - w) \sum_{n=0}^{\infty} w^n \mathbb{P}(L(n) < \ell) = \mathbb{P}(L^* < \ell) = \frac{(1 - w)(1 - (wp)^\ell)}{1 - w + (wp)(wp)^\ell}, \tag{16}$$

inverting this does not result in explicit expressions. To see what we obtain, let us, instead, note that

$$\mathbb{P}(L(n) < \ell) = \mathbb{P}(G_\ell(n) = 0) = \sum_{r \geq 0} (-1)^r \mathbb{E} \binom{G_\ell(n)}{r}$$

and use the binomial moment formula (9) to obtain

$$F_\ell(n) := \mathbb{P}(L(n) < \ell) = 1 + \sum_{r \geq 1} (-1)^r \left[\binom{n - \ell r}{r} p^{\ell r} q^r + \binom{n - \ell r}{r - 1} p^{\ell r} q^{r-1} \right]. \tag{17}$$

It is easy to see that the function $n \mapsto F_\ell(n)$ satisfies a recursion.

Proposition 1. *Let $\ell \in \mathbb{N}$. Define $F_\ell(0) = 1$ and, for $n \geq 1$, $F_\ell(n) = \mathbb{P}(L(n) < \ell)$. Then*

$$F_\ell(n) = qF_\ell(n - 1) + qpF_\ell(n - 2) + \dots + qp^{\ell-1}F_\ell(n - \ell).$$

Proof. This can be proved directly by induction. But, since Theorem 3 is available, set $h(x) := \mathbf{1}_{\{\bar{\lambda}(x) < \ell\}}$, observe that $h(x + e_j) = h(x) \mathbf{1}_{\{j < \ell\}}$ and substitute into (15).

5.1. The Poisson regime for large lengths

According to Feller [5, Section XIII.12, Problem 25, p. 341], asymptotics for $L(n)$ date back to von Mises [16]. Very sharp asymptotics for $L(n)$ are also known; see Erdős and Rényi [3], its sequel paper by Erdős and Révész [4], and the review paper by Révész [15]. But it is a matter of elementary analysis to see that the distribution function $\ell \mapsto F_\ell(n)$ exhibits a cutoff at ℓ of the order of magnitude of $\log n$. To see this in a few lines, consider (9) for the binomial moment of $G_\ell(n)$. We have the following lemma.

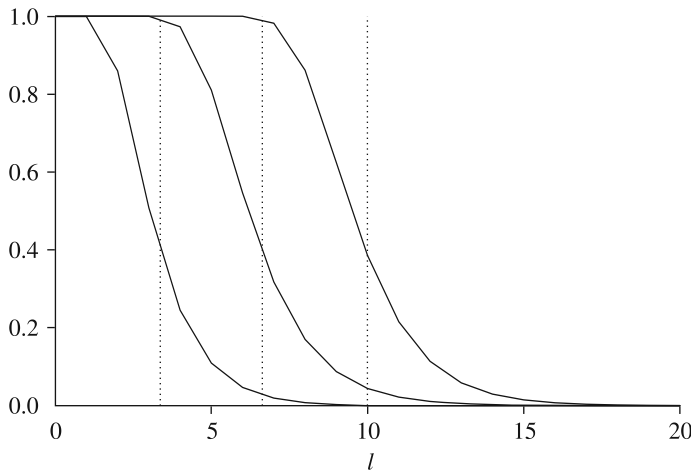


FIGURE 1: Plot (piecewise linear interpolation) of $\ell \mapsto \mathbb{P}(L(n) \geq \ell)$ for $n = 10, 100, 1000$, and $p = \frac{1}{2}$. The vertical lines are at $\ell = \log_2(n)$.

TABLE 1: Comparing exact and approximate values for $\mathbb{P}(L(n) \geq \ell)$ when $p = \frac{1}{2}$ and $n = 10^4$.

| ℓ | $\mathbb{P}(L(n) \geq \ell)$ | $1 - \exp(-(n - \ell)p^\ell q - p^\ell)$ |
|--------|--|--|
| 10 | 0.992 583 894 386 551 | 0.992 394 672 192 560 |
| 12 | 0.705 167 040 532 444 | 0.704 616 988 848 744 |
| 14 | 0.262 835 671 849 087 | 0.262 736 242 068 365 |
| 20 | 0.004 748 524 931 253 | 0.004 748 478 671 106 |
| 50 | 4.419 575 816 418 15×10^{-12} | 4.420 000 000 000 01×10^{-12} |

Lemma 1. Fix $0 < p < 1$ and let $\ell = \ell(n) \rightarrow \infty$ so that $np^{\ell(n)}q \rightarrow \theta$ as $n \rightarrow \infty$ for some $\theta > 0$. Then

$$\mathbb{E}\binom{G_{\ell(n)}(n)}{r} \rightarrow \frac{\theta^r}{r!}, \quad \mathbb{P}(G_{\ell(n)}(n) = 0) \rightarrow e^{-\theta}.$$

The proof is elementary. Since $\mathbb{P}(G_\ell(n) = 0) = \mathbb{P}(L(n) < \ell)$ for all n and ℓ , the last asymptotic result can be translated immediately into the following threshold behavior.

Corollary 12. Let $0 < \alpha < \infty, 0 < \beta \leq +\infty$. Then

$$\mathbb{P}(L(n) < \alpha \log_{1/p} n + \log_{1/p} \beta) \rightarrow \begin{cases} e^{-q/\beta} & \text{if } \alpha = 1, \\ 1 & \text{if } \alpha > 1, \\ 0 & \text{if } \alpha < 1. \end{cases}$$

In Figure 1, we take $p = \frac{1}{2}$ and plot $\ell \mapsto \mathbb{P}(L(n) \geq \ell)$ for three values of n .

Corollary 12 suggests the practical approximation

$$\mathbb{P}(L(n) < \ell) \doteq \exp(-\mathbb{E}G_\ell(n)) = \exp(-(n - \ell)p^\ell q - p^\ell), \tag{18}$$

valid for large n and ℓ , roughly when ℓ is of order $\log_{1/p} n$ or higher. In Table 1 we compare the exact result with the approximation for $n = 10^4, p = \frac{1}{2}$, and ℓ ranging from slightly below $\log_2 10^4 \approx 13.288$ to much higher values. We programmed (17) in MAPLE in order to obtain the exact values of $\mathbb{P}(L(n) \geq \ell)$.

In Figure 2, we plot $\ell \mapsto \mathbb{P}(L(n) \geq \ell)$ for $n = 1000$ and three different values of p . We also plot the analytical approximation given by the right-hand side of (18). Note that, visually at least, there is no way to tell the difference between real values and the approximating curves.

The result of Lemma 1 easily implies that the law of $G_{\ell(n)}(n)$ converges weakly as $n \rightarrow \infty$ to a Poisson law with mean θ .

Corollary 13. Under the assumptions of Lemma 1, we have

$$\mathcal{L}\{G_{\ell(n)}(n)\} \rightarrow \text{Poisson}(\theta).$$

Proof. It is enough to establish convergence of binomial moments to those of a Poisson law. Recall that if N is $\text{Poisson}(\theta)$ then $\mathbb{E}\binom{N}{r} = \theta^r/r!$. Lemma 1 tells us that the r th binomial moment of $G_{\ell(n)}(n)$ converges to $\theta^r/r!$ and this establishes the result.

More interestingly, using the result of Corollary 6, we arrive at the following proposition.

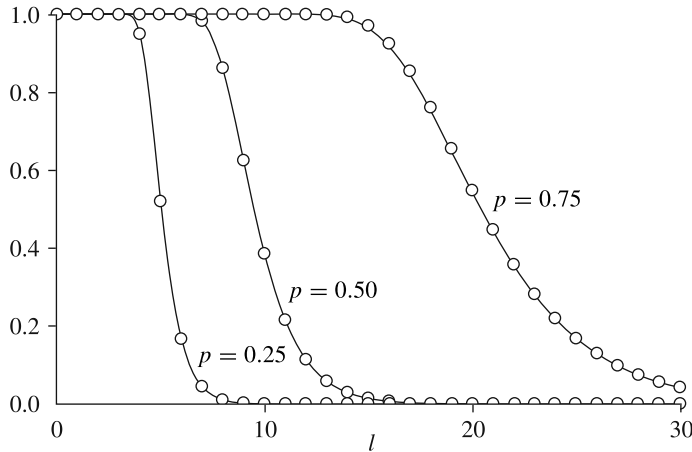


FIGURE 2: Plot of $l \mapsto \mathbb{P}(L(1000) \geq l)$ and $p = 0.25, 0.5, 0.75$. The dots correspond to the actual values. The solid lines correspond to the analytical approximation (18).

Proposition 2. Consider $v \in \mathbb{N}$, positive real numbers $\theta_1, \dots, \theta_v$, and sequences $\ell_j(n)$, $j = 1, \dots, v$ of positive integers, such that

$$\lim_{n \rightarrow \infty} np^{\ell_j(n)}q = \theta_j \quad \text{for } j = 1, \dots, v.$$

Then

$$\mathcal{L}\{R_{\ell_1(n)}(n), \dots, R_{\ell_v(n)}(n)\} \longrightarrow \text{Poisson}(\theta_1q) \times \dots \times \text{Poisson}(\theta_vq) \longrightarrow \quad \text{as } n \rightarrow \infty.$$

Proof. It suffices to show that the joint binomial moments converge to products of Poisson binomial moments. Fix $\ell_1, \dots, \ell_v, r_1, \dots, r_v$, set $r_0 = r_1 + \dots + r_v$, and, using the abbreviations (1) for α, β and γ , write the expression (7) for the joint binomial moments as

$$\mathbb{E} \binom{R_{\ell_1}^*}{r_1} \dots \binom{R_{\ell_v}^*}{r_v} = (1-w) \frac{r_0!}{r_1! \dots r_v!} p^{\ell \cdot r} q^{r_0-1} w^{\ell \cdot r + r_0 - 1} \frac{(1-wp)^{r_0+1}}{(1-w)^{r_0+1}}.$$

Expand $(1-wp)^{r_0+1}$ using the binomial formula, and $(1-w)^{r_0+1}$ using (8), to write

$$\frac{(1-wp)^{r_0+1}}{(1-w)^{r_0+1}} = \sum_{k=0}^{\infty} \sum_{s=0}^{r_0+1} (-p)^s w^{k+s} \binom{r_0+1}{s} \binom{r_0+k}{r_0}$$

and obtain

$$\begin{aligned} & \mathbb{E} \binom{R_{\ell_1}^*}{r_1} \dots \binom{R_{\ell_v}^*}{r_v} \\ &= (1-w) \frac{r_0!}{r_1! \dots r_v!} p^{\ell \cdot r} q^{r_0-1} \sum_{k=0}^{\infty} \sum_{s=0}^{r_0+1} (-p)^s w^{\ell \cdot r + r_0 - 1 + k + s} \binom{r_0+1}{s} \binom{r_0+k}{r_0} \\ &= (1-w) \sum_{n=0}^{\infty} w^n \sum_{s=0}^{r_0+1} \frac{r_0!}{r_1! \dots r_v!} p^{\ell \cdot r} q^{r_0-1} (-p)^s \binom{r_0+1}{s} \binom{n+1-\ell \cdot r-s}{r_0}. \end{aligned}$$

Therefore,

$$\begin{aligned} &\mathbb{E}\binom{R_{\ell_1(n)}}{r_1} \dots \binom{R_{\ell_v(n)}}{r_v} \\ &= \frac{r_0!}{r_1! \dots r_v!} p^{\ell \cdot r} q^{r_0-1} \sum_{s=0}^{r_0+1} (-p)^s \binom{r_0+1}{s} \binom{n+1-\ell \cdot r-s}{r_0} \\ &= \frac{q^{-1}}{r_1! \dots r_v!} \sum_{s=0}^{r_0+1} (-p)^s \binom{r_0+1}{s} (n+1-\ell \cdot r-s)_{r_0} \prod_{j=1}^v (p^{\ell_j} q)^{r_j}, \end{aligned}$$

where $(N)_{r_0} = N(N-1) \dots (N-r_0+1)$. Using the assumptions, we have

$$\lim_{n \rightarrow \infty} (n+1-\ell \cdot r-s)_{r_0} \prod_{j=1}^v (p^{\ell_j} q)^{r_j} = \theta_1^{r_1} \dots \theta_v^{r_v},$$

and so

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}\binom{R_{\ell_1(n)}(n)}{r_1} \dots \binom{R_{\ell_v(n)}(n)}{r_v} &= q^{-1} \sum_{s=0}^{r_0+1} (-p)^s \binom{r_0+1}{s} \frac{\theta_1^{r_1}}{r_1!} \dots \frac{\theta_v^{r_v}}{r_v!} \\ &= q^{-1} (1-p)^{r_0+1} \frac{\theta_1^{r_1}}{r_1!} \dots \frac{\theta_v^{r_v}}{r_v!} \\ &= q^{r_1+\dots+r_v} \frac{\theta_1^{r_1}}{r_1!} \dots \frac{\theta_v^{r_v}}{r_v!}, \end{aligned}$$

thus establishing the assertion.

5.2. A better approximation for small length values

We now pass on to a different approximation for $F_\ell(n) = \mathbb{P}(L(n) < \ell)$. We again consider (16),

$$\sum_{n=0}^{\infty} w^n F_\ell(n) = \frac{1 - (wp)^\ell}{1 - w + (wq)(wp)^\ell}, \tag{19}$$

and look at the denominator

$$f(w) := 1 - w + p^\ell q w^{\ell+1},$$

considered as a polynomial in $w \in \mathbb{C}$, of degree $\ell + 1$. The smallest (in magnitude) zeros of $f(w)$ govern the behavior of $n \mapsto F_\ell(n)$ for large n (and all ℓ).

Proposition 3. *The equation*

$$f(w) = 0, \quad w \in \mathbb{C},$$

has two real roots $w_0 = w_0(\ell)$ and $1/p$, such that

$$\begin{aligned} 1 < w_0 < \frac{\ell+1}{\ell} < \frac{1}{p((\ell+1)q)^{1/\ell}} < \frac{1}{p} \quad \text{for } p < \frac{\ell}{\ell+1}, \\ 1 < \frac{1}{p} < \frac{\ell+1}{\ell} < \frac{1}{p((\ell+1)q)^{1/\ell}} < w_0 \quad \text{for } p > \frac{\ell}{\ell+1}, \\ 1 < w_0 = \frac{1}{p} \quad \text{for } p = \frac{\ell}{\ell+1}, \end{aligned}$$

and all other roots are outside the circle with radius $\max(w_0, 1/p)$ in the complex plane. Moreover, $\lim_{\ell \rightarrow \infty} w_0(\ell) = 1$.

Proof. We check the behavior of $f(w)$ for real w . First, we have $f(1/p) = 0$. Now, $f'(w) = -1 + (\ell + 1)p^\ell q w^\ell$, and so the only real root of $f'(w) = 0$ is $w^* = 1/p((\ell + 1)q)^{1/\ell}$. Since $f''(w) = \ell(\ell + 1)p^\ell q w^{\ell-1}$, the function f is strictly convex on $[0, \infty)$ and so $f(w^*)$ is a global minimum of f on $[0, \infty)$. Note that $f(w^*) = 1 - \ell w^*/(\ell + 1)$. We claim that $f(w^*) \leq 0$, or, equivalently, that $w^* \geq (\ell + 1)/\ell$. Upon substituting with the value of w^* , this last inequality is equivalent to $p^\ell(1 - p) \leq (\ell/(\ell + 1))^\ell$. But this is true, since $\max_{0 \leq p \leq 1} p^\ell(1 - p) = (\ell^\ell/(\ell + 1))^{\ell+1} \leq (\ell/(\ell + 1))^\ell$. Hence, $f(w^*) \leq 0$ with equality if and only if $p = \ell/(\ell + 1)$. On the other hand, $f(0) = 1$ and $\lim_{w \rightarrow \infty} f(w) = \infty$. Therefore, $f(w) = 0$ has two positive real roots straddling w^* . One of them is $1/p$. Denote the other root by w_0 . Since $f((\ell + 1)/\ell) = -(1/\ell) + ((\ell + 1)/\ell)^{\ell+1} p^\ell q < 0$, and $f(w^*) < 0$, provided that $p \neq \ell/(\ell + 1)$, it follows that, in this case, w_0 and $1/p$ are outside the interval $[(\ell + 1)/\ell, w^*]$. Depending on whether p is smaller or larger than $\ell/(\ell + 1)$, we have $w_0 < 1/p$ or $w_0 > 1/p$, respectively. If $p = \ell/(\ell + 1)$ then $w^* = (\ell + 1)/\ell$ and then $1/p = w_0 = w^*$. Since $f(1) = p^\ell q$, it follows that $w_0 > 1$, in all cases. Finally, for all sufficiently large ℓ , we have $p < \ell/(\ell + 1)$ and so $1 < w_0 < (\ell + 1)/\ell$, showing that the limit of w_0 , as $\ell \rightarrow \infty$, is 1. To show that the only roots $f(w) = 0$ with $|w| \leq \max(w_0, 1/p)$ are w_0 and $1/p$, we need an auxiliary lemma which is probably well-known but whose proof we supply for completeness.

Lemma 2. Consider the polynomial $P(z) := c_0 + c_1z + \dots + c_nz^n$, $z \in \mathbb{C}$ with real coefficients such that $c_0 > c_1 > \dots > c_n > 0$. Then all the zeros of $P(z)$ lie outside the closed unit ball centered at the origin.

Proof. Fix $\lambda > 1$ such that $c_0 > c_1/\lambda > c_2/\lambda^2 > \dots > c_n/\lambda^n$ and note that

$$c_0 + (z - 1)P\left(\frac{z}{\lambda}\right) = \left(c_0 - \frac{c_1}{\lambda}\right)z + \left(\frac{c_1}{\lambda} - \frac{c_2}{\lambda^2}\right)z^2 + \dots + \left(\frac{c_{n-1}}{\lambda^{n-1}} - \frac{c_n}{\lambda^n}\right)z^n + \frac{c_n}{\lambda^n}z^{n+1}.$$

Therefore, on $|z| = 1$,

$$\left|c_0 + (z - 1)P\left(\frac{z}{\lambda}\right)\right| \leq \left(c_0 - \frac{c_1}{\lambda}\right) + \left(\frac{c_1}{\lambda} - \frac{c_2}{\lambda^2}\right) + \dots + \left(\frac{c_{n-1}}{\lambda^{n-1}} - \frac{c_n}{\lambda^n}\right) + \frac{c_n}{\lambda^n} = c_0 = |-c_0|.$$

Rouché’s theorem [1, p. 153] implies that $(z - 1)P(z/\lambda)$ and $-c_0$ have the same number of zeros inside the open unit ball centered at the origin. That is, all zeros of $P(z/\lambda)$ lie outside the open unit ball. Since $\lambda > 1$, it follows that all zeros of $P(z)$ lie outside the closed unit ball.

Completion to the proof of Proposition 3. Firstly, assume that $w_0p \neq 1$. By polynomial division, write $f(w)$ as

$$f(w) = -q(1 - pw)(w - w_0)S(w),$$

where

$$S(w) := \sum_{j=0}^{\ell-1} \frac{1 - (w_0p)^{\ell-j}}{1 - w_0p} p^j w^j.$$

Since the positive numbers $(1 - (w_0p)^{\ell-j})/(1 - w_0p)$ decrease with j , it follows from Lemma 2 that if $S(w) = 0$ then $|pw| > 1$, i.e. $|w| > 1/p$. On the other hand, since

$$S(w) = \sum_{j=0}^{\ell-1} \frac{(1 - (w_0p)^{\ell-j})(w_0p)^j}{(1 - w_0p)(w/w_0)^j}$$

and since the positive numbers $(1 - (w_0p)^{\ell-j})(w_0p)^j/(1 - w_0p)$ decrease with j it follows, again by Lemma 2, that $S(w) = 0$ implies $|w/w_0| > 1$, i.e. $|w| > w_0$. If $w_0p = 1$ then $S(w) = \sum_{j=0}^{\ell-1} (\ell - j)p^j w_j$ and so $S(w) = 0$ implies $|w| > 1/p = w_0$.

We translate this result into an approximation for the distribution function of $L(n)$.

Proposition 4. *Let $w_0 = w_0(\ell)$ be the root of the equation $f(w) = 0$ defined in Proposition 3. If $p \neq \ell/(\ell + 1)$ (where $a(n) \sim b(n)$ stands for $\lim a(n)/b(n) = 1$), then*

$$\mathbb{P}(L(n) < \ell) \sim \frac{1 - (w_0p)^\ell}{1 - (\ell + 1)q(w_0p)^\ell} w_0^{-n-1} \quad \text{as } n \rightarrow \infty.$$

If $p = \ell/(\ell + 1)$ then

$$\mathbb{P}(L(n) < \ell) \sim 2 \left(\frac{\ell}{(\ell + 1)} \right)^{n+1} \quad \text{as } n \rightarrow \infty.$$

Proof. Firstly, suppose that $p \neq \ell/(\ell + 1)$ and, using partial fraction expansion, write (19) as

$$\frac{g(w)}{f(w)} = \frac{1 - (wp)^\ell}{1 - w + (wq)(wp)^\ell} = \frac{c_0}{1 - w/w_0} + \frac{h(w)}{j(w)}. \tag{20}$$

To do this, we use the fact that w_0 is a zero of the denominator $f(w) = 1 - w + (wq)(wp)^\ell$ but not a zero of the numerator $g(w) = 1 - (wp)^\ell$. Also, $w_0 \neq 1/p$, and both $g(w)$ and $f(w)$ have a zero at $1/p$. Hence, $h(w)$ and $j(w)$ are polynomials with degrees $\ell - 2$ and $\ell - 1$, respectively, and $j(w_0) \neq 0$. Hence,

$$c_0 = \frac{1}{w_0} \lim_{w \rightarrow w_0} \frac{(w_0 - w)g(w)}{f(w)} = - \left(\frac{g(w_0)}{f'(w_0)} \right) \left(\frac{1}{w_0} \right) = \left(\frac{1 - (w_0p)^\ell}{1 - (\ell + 1)q(w_0p)^\ell} \right) \left(\frac{1}{w_0} \right).$$

From Proposition 3, we see that the zeros of $j(w)$ are all outside the circle with radius w_0 . Hence, from (19) and (20), we obtain

$$\mathbb{P}(L(n) < \ell) \sim c_0 w_0^{-n} \quad \text{as } n \rightarrow \infty, \tag{21}$$

which proves the first assertion. If $p = \ell/(\ell + 1)$ then $w = 1/p = w_0$ is a simple zero for $g(w)$ and a double zero for $f(w)$. Hence, from (20), we have

$$\begin{aligned} c_0 &= - \left(\frac{g'(w_0)}{(1/2)f''(w_0)} \right) \left(\frac{1}{w_0} \right) \\ &= - \left(\frac{p^\ell q w_0^{\ell-1}}{(1/2)p^\ell q \ell (\ell + 1) w_0^{\ell-1}} \right) \left(\frac{1}{w_0} \right) \\ &= \frac{2}{q(\ell + 1)w_0} \\ &= \frac{2\ell}{\ell + 1}. \end{aligned}$$

Since the zeros of $j(w)$ are outside the circle with radius $w_0 = (\ell + 1)/\ell$, (21) holds. This proves the second assertion.

Since these approximations are valid for all ℓ , they nicely complement the Poisson approximation discussed earlier. For $n, \ell \rightarrow \infty$, such that $np^\ell \asymp 1$, we have $w_0(\ell) = 1 + p^\ell q + O(\ell/n^2)$. From the approximation above, we find $\mathbb{P}(L_\ell(n) < \ell) \approx e^{-np^\ell q}$ which is asymptotically equivalent to the Poisson approximation.

TABLE 2: Values of $\mathbb{P}(L(n) \geq \ell)$ for different values of p and ℓ .

| n | Exact | Poisson approx. (error) | Second approx. (error) |
|---|-----------------|--|---|
| For $p = \frac{1}{2}$ and $\ell = 2$. | | | |
| 5 | 0.59375 | 0.46474 (22%) | 0.59426 (0.086%) |
| 7 | 0.73438 | 0.58314 (20.6%) | 0.73445 (0.01%) |
| 10 | 0.85938 | 0.71350 (17%) | 0.8594 (0.002%) |
| 20 | 0.98311 | 0.91792 (6.63%) | 0.98312 (0.0010%) |
| For $p = \frac{1}{3}$ and $\ell = 2$. | | | |
| 5 | 0.32510 | 0.28347 (12.8%) | 0.32557 (0.14%) |
| 7 | 0.44033 | 0.38213 (13.2%) | 0.44080 (0.11%) |
| 10 | 0.57730 | 0.50525 (12.5%) | 0.57779 (0.08%) |
| 20 | 0.83415 | 0.76411 (8.4%) | 0.83453 (0.05%) |
| For $p = \frac{4}{5}$ and $\ell = 2$. | | | |
| 5 | 0.94208 | 0.64084 (32.0%) | 0.94386 (0.189%) |
| 7 | 0.98509 | 0.72196 (26.71%) | 0.98526 (0.0173%) |
| 10 | 0.998 023 2 | 0.810 620 1 (18.78%) | 0.998 017 9 (0.00052%) |
| 20 | 0.999 997 5 | 0.947 345 3 (5.265%) | 0.999 997 5 (0.000 003%) |
| For $p = \frac{1}{2}$ and $\ell = 7$. | | | |
| 100 | 0.31752 | 0.31002 (2.36%) | 0.19644 (38.13%) |
| 500 | 0.86364 | 0.85537 (0.96%) | 0.8372 (3.06%) |
| 1500 | 0.99757 | 0.99709 (0.048%) | 0.99700 (0/057%) |
| 3000 | 0.999 994 198 6 | 0.999 991 699 7 (0.00025%) | 0.999 993 192 8 (0.00010%) |
| For $p = \frac{2}{3}$ and $\ell = 10$. | | | |
| 100 | 0.43531 | 0.41583 (4.475%) | 0.46433 (6.667%) |
| 500 | 0.95209 | 0.94214 (1.045%) | 0.95480 (0.285%) |
| 1500 | 0.999 900 | 0.999 821 (0.00790%) | 0.999 905 (0.00050%) |
| 3000 | 0.999 999 990 4 | 0.999 999 969 4 ($2.1 \times 10^{-6}\%$) | 0.999 999 990 8 ($0.04 \times 10^{-6}\%$) |

5.3. Numerical comparisons of the two approximations

We numerically compute $\mathbb{P}(L(n) \geq \ell)$, first using the exact formula (17), then using the Poisson approximation (18), and finally using the approximation suggested from Proposition 4. We see, as expected, that for small values of ℓ compared to n , the second approximation outperforms the first one.

First, we let $\ell = 2$. Then $f(w) = 1 - w + p^2qw^3 = (pw - 1)(pqw^2 + qw - 1)$, and so

$$w_0 = \frac{\sqrt{1 + 4p/q} - 1}{2p}.$$

The approximation suggested from Proposition 4 is

$$\mathbb{P}(L(n) < 2) \sim \begin{cases} \frac{1 - (w_0p)^2}{1 - 3q(w_0p)^2} w_0^{-n+1} & \text{if } p \neq \frac{2}{3}, \\ 2\left(\frac{2}{3}\right)^{-(n+1)} & \text{if } p = \frac{2}{3}. \end{cases}$$

Some numerical values are shown in Table 2.

Next, we increase the value of ℓ and pick two different values for p . We solve, in each case, the equation $f(w) = 0$ numerically.

6. Discussion and open problems

Gordon *et al.* [7] developed an extreme value theory for long runs. As mentioned therein, it is intriguing that the longest run possesses no limit distribution, and this is based on an earlier paper by Guibas and Odlyzko [10].

We have not touched upon the issue of more general processes generating heads and tails, for example, Markovian processes. The portmanteau identity can be generalized to include the Markovian dependence and this can be the subject for future work provided that a suitable motivation be found.

Another set of natural questions arising is to what extent do we have a weak approximation of $\mathbf{R}(n)$ on a function space (convergence to a Brownian bridge?), as well as the quality of such an approximation.

Acknowledgements

We would like to thank Joe Higgins for the short proof of Lemma 2. T. Konstantopoulos was supported by the Swedish Research Council (grant no. 2013-4688).

References

- [1] AHLFORS, L. V. (1978). *Complex Analysis*. McGraw-Hill, New York.
- [2] BALAKRISHNAN, N. AND KOUTRAS, M. V. (2002). *Runs and Scans with Applications*. John Wiley, New York.
- [3] ERDŐS, P. AND RÉNYI, A. (1970). On a new law of large numbers. *J. Analyse Math.* **23**, 103–111.
- [4] ERDŐS, P. AND RÉVÉSZ, P. (1977). On the length of the longest headrun. In *Topics in Information Theory* (Colloq. Math. Soc. János Bolyai **16**), North-Holland, Amsterdam, pp. 219–228.
- [5] FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd edn. John Wiley, New York.
- [6] FU, J. C. AND LOU, W. Y. W. (2003). *Distribution Theory of Runs and Patterns and Its Applications*. World Scientific, River Edge, NJ.
- [7] GORDON, L., SCHILLING, M. F. AND WATERMAN, M. S. (1986). An extreme value theory for long head runs. *Prob. Theory Relat. Fields* **72**, 279–287.
- [8] GRAHAM, R. L., KNUTH, D. E. AND PATASHNIK, O. (1994). *Concrete Mathematics*, 2nd edn. Addison-Wesley, Reading, MA.
- [9] GRIMMETT, G. R. AND STIRZAKER, D. R. (2001). *Probability and Random Processes*, 3rd edn. Oxford University Press.
- [10] GUIBAS, L. J. AND ODLYZKO, A. M. (1980). Long repetitive patterns in random sequences. *Z. Wahrscheinlichkeitsth.* **53**, 241–262.
- [11] MAKRI, F. S. AND PSILLAKIS, Z. M. (2011). On success runs of length exceeded a threshold. *Methodol. Comput. Appl. Prob.* **13**, 269–305.
- [12] MURRAY, D. B. AND TEARE, S. W. (1993). Probability of a tossed coin landing on edge. *Phys. Rev. E* **48**, 2547–2552.
- [13] MUSELLI, M. (1996). Simple expressions for success run distributions in Bernoulli trials. *Statist. Prob. Lett.* **31**, 121–128.
- [14] PHILIPPOU, A. N. AND MAKRI, F. S. (1986). Success runs and longest runs. *Statist. Prob. Lett.* **4**, 101–105, 211–215.
- [15] RÉVÉSZ, P. (1980). Strong theorems on coin tossing. In *Proceedings of the International Congress of Mathematicians* Acad. Sci. Fennica, Helsinki, pp. 749–754.
- [16] VON MISES, R. (1921). Zur theorie der iterationen. *Z. Angew. Math. Mech.* **1**, 298–307.
- [17] VON MISES, R. (1981). *Probability, Statistics and Truth*. Dover, New York.