# Runs in coin tossing: a general approach for deriving distributions for functionals

Lars Holst        Takis Konstantopoulos*

February 10, 2014

**Abstract**

We take a fresh look at the classical problem of runs in a sequence of i.i.d. 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 and phrases.* regeneration, coin tossing, runs, longest run, Poisson approximation, Laplace transforms, Rouché's theorem

*AMS* 2010 *subject classifications.* 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 i.i.d. 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 [13].

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

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

*Corresponding author; Supported by Swedish Research Council grant 2013-4688

$S(0) := 0$. Throughout the 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 $n \in J$. People have been interested in computing the distribution of runs of various types. For example, we may ask for the distribution of the number of runs of a given length in $n$ coin tosses. Feller [5, Section XIII.7] considers the probability that a run of a given length $\ell$ first appears at the $n$-th coin toss and, using renewal theory, computes 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 attributes this result to von Mises [14].[1] Philippou and Makri [12] derive the joint distribution of the longest run and the number of runs of a given length. More detailed computations are considered in [10]. The literature is extensive and there is even a 452-page book on the topic [2].

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 result into a formula for the quantity of interest. To be more specific, let $R_\ell(n)$ be the number of runs of length $\ell$ in the first $n$ coin tosses. Consider the vector

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

as an element of the set

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

The set can be 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. The set is countable, and so the random variable $R(n)$ has a discrete distribution. If $h : \mathbb{Z}_+^* \to \mathbb{R}^d$ is any function then we refer to the random variable $h(R(n))$ as a $d$-dimensional functional of a run-vector. For example, for $d = 1$, if $h_1(x) = \sup\{\ell : x_\ell > 0\}$ (with $\sup \varnothing = 0$), then $h_1(R(n))$ 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 $h_2(R(n))$ 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

[1] In his classic work [15, p, 138], von Mises, refers to a 1916 paper of the philosopher Karl Marbe who reports that in 200,000 birth registrations in a town in Bavaria, there is only one 'run' of 17 consecutive births of children of the same sex. Note that $\log_2(200000) \approx 17.61$ and see Section 5 below.

a 2-dimensional functional; a formula for the distribution of $h(R(n))$ 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 $\mathbb{Z}^* := \{x \in \mathbb{Z}^N : x_k = 0, \text{ eventually}\}$, 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 $(0, 0, \ldots)$ of this vector space. For $j = 1, 2, \ldots$, we let $e_j = (e_j(1), e_j(2), \ldots) \in \mathbb{Z}^*$ 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

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

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 $R^*$, defined as $R(\cdot)$ stopped at an independent geometric time. We call this formula a "portmanteau identity" because it contains lots of special cases of interest. To explain this, we give, in the same section, formulas 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 $\ell$ in $n$ coin tosses. In particular, we point out its relationship with hypergeometric functions. Section 4 translates the portmanteau identity into a "portmanteau recursion" which provides, for example, a method for recursive evaluation of the generating function of the random vector $R(n)$. In Section 5 we take a closer look at the most common functional of $R(n)$, 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. This roughly states that, in a certain approximating regime, the random variables $R_{\ell_1}(n), \ldots, R_{\ell_\nu}(n)$ become asymptotically independent Poisson random variables, as $n \to \infty$, when, simultaneously, $\ell_1, \ldots, \ell_\nu \to \infty$. A second approximation for the distribution function $\mathbb{P}(L(n) < \ell)$ of $L(n)$, which works well at small values of $\ell$, 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. We point out that, although our method is applied to finding very detailed information about the distribution function of the specific functional $L(n)$, 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 $\xi_1, \xi_2, \ldots$. 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 $n = 0, 1, \ldots$ Note that $R(0) = (0, 0, \ldots)$, 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 $\alpha, \beta, \gamma$ are three nonnegative real numbers adding up to 1 with $\gamma$ strictly positive, then $w, p, q = 1 - p$ are uniquely determined.

**Theorem 1.** *For any $h : \mathbb{Z}_+^* \to \mathbb{R}$ 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[2]) but equip it with an explosion mechanism which is activated if the coin touches the ground on its edge. An explosion occurs with probability $\gamma = 1 - w$. When an explosion occurs the coin is destroyed immediately and we do not observe heads or tails. If an 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

[2] A British pound is sufficiently thick so that the chance of landing on its edge is non-negligible, 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 1/6000 [11].

the probability that we observe tails. We let $\mathsf{E}, \mathsf{H}, \mathsf{T}$ denote "explosion", "heads", "tails", respectively, for the explosive coin. The possible outcomes in tossing such a coin comprise the set

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

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

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

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 $\{\mathsf{H}^j\mathsf{E}\}$, we have $R^* = e_j$. Hence $E[\mathsf{H}^j\mathsf{E}; h(R^*)] = \alpha^j \gamma h(e_j)$. On the event $\{\mathsf{H}^j\mathsf{T}\}$ we have $R^* = e_j + \theta^{j+1}R^*$, where $\theta^{j+1}R^* = (R^*_{j+1}, R^*_{j+2}, \ldots)$, which is independent and identical in law to $R^*$. Hence $\mathbb{E}[\mathsf{H}^j\mathsf{T}; h(R^*)] = \alpha^j \beta \mathbb{E}h(e_j + R^*)$. $\square$

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

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

for any sequence $z_1, z_2, \ldots$ 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 $(R_1^*, \ldots, R_\ell^*)$ is specified by*

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

*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. $\square$

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 $\mathrm{geo}(\alpha, \beta)$ denote the probability measure $Q$ on $\mathbb{Z}_+ \cup \{+\infty\} = \{0, 1, \ldots, +\infty\}$ with*

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

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

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

Therefore, comparing with (3), we have

**Corollary 1.** *$R_\ell^*$ has $\mathrm{geo}(\alpha_\ell, \beta_\ell)$ 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*

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

As a reality check, observe that

$$\mathbb{E}R_\ell^* = \frac{\beta_\ell}{1-\beta_\ell} = \frac{(1-\alpha)^2\alpha^\ell}{\gamma},$$

and so

$$\sum_{\ell=1}^\infty \ell \mathbb{E}R_\ell^* = \frac{\alpha}{\gamma}.$$

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 computations,

$$S(N^*-1) \sim \text{geo}\left(\frac{\gamma}{1-\beta}, \frac{\alpha}{1-\beta}\right)$$

and so

$$\mathbb{E}S(N^*-1) = \frac{\alpha}{\gamma},$$

as above.

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

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

**Corollary 2.** *Let $L^* := \overline{\lambda}(R^*)$ 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 $\overline{\lambda}(0) = 0$, since $\sup \varnothing = 0$. Also

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

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

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

which immediately gives the announced formula. $\square$

See also Grimmett and Stirzaker [6, Section 5.12, Problems 46,47] for another 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*

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

*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 $h(x) = \mathbf{1}\{\underline{\lambda}(x) \ge \ell\}$ in (2). We work out that $h(0)=1$ and, for $j \in \mathbb{N}$, $h(e_j) = j$, $h(x+e_j) = h(x)\mathbf{1}\{j \ge \ell\}$. The rest is elementary algebra. □

Since $\mathbb{Z}^*$ is a vector space, there are linear functions from $\mathbb{Z}^*$ into $\mathbb{R}$.

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

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

Setting $h(x) = x_\ell$ gives again the earlier formula for $\mathbb{E}R^*_\ell$.

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.**

$$\mathbb{E}\big[z_1^{R_1^*} \cdots z_{\ell-1}^{R^*_{\ell-1}};\, L^* < \ell\big] = \gamma \frac{1+\sum_{j=1}^{\ell-1} \alpha^j z_j}{1-\beta-\beta\sum_{j=1}^{\ell-1}\alpha^j z_j}.$$

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

For verification, note that taking $\ell \to \infty$ in the last display gives the previous formula for $\mathbb{E}z^{R^*}$, while letting $z_1 = \cdots = z_{\ell-1} = 1$ gives the previous formula for $\mathbb{P}(L^* < \ell)$.

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

**Corollary 6.** *Consider positive integers $\nu$, $\ell_1, \dots, \ell_\nu$, and nonnegative integers $r_1, \dots, r_\nu$, such that $r_0 := r_1 + \cdots + r_\nu \ge 1$. Let $\boldsymbol{\ell} := (\ell_1, \dots, \ell_\nu)$ and $\boldsymbol{r} := (r_1, \dots, r_\nu)$ and set $\boldsymbol{\ell} \cdot \boldsymbol{r} = \ell_1 r_1 + \cdots + \ell_\nu r_\nu$. Then*

$$\mathbb{E}z_1^{R^*_{\ell_1}} \cdots z_\nu^{R^*_{\ell_\nu}} = \frac{1+(1-\alpha)\sum_{j=1}^\nu \alpha^{\ell_j}(z_j-1)}{1-\frac{(1-\alpha)\beta}{\gamma}\sum_{j=1}^\nu \alpha^{\ell_j}(z_j-1)} \tag{4}$$

*and*

$$\mathbb{E}\binom{R^*_{\ell_1}}{r_1}\cdots\binom{R^*_{\ell_\nu}}{r_\nu} = \frac{r_0!}{r_1!\cdots r_\nu!}\,\frac{\alpha^{\boldsymbol{\ell}\cdot\boldsymbol{r}}\beta^{r_0-1}(1-\alpha)^{r_0+1}}{\gamma^{r_0}} \tag{5}$$

*Proof.* By Theorem 1,

$$\begin{aligned}
\mathbb{E}z_1^{R^*_{\ell_1}}\cdots z_\nu^{R^*_{\ell_\nu}} &= \gamma\left(\sum_{j=1}^{\nu}\alpha^{\ell_j}z_j + \sum_{j\notin\{\ell_1,\dots,\ell_\nu\}}\alpha^j\right) + \beta\left(\sum_{j=1}^{\nu}\alpha^{\ell_j}z_j + \sum_{j\notin\{\ell_1,\dots,\ell_\nu\}}\alpha^j\right)\mathbb{E}z_1^{R^*_1}\cdots z_\nu^{R^*_\nu}\\
&= \gamma\left(\sum_{j=1}^{\nu}\alpha^{\ell_j}(z_j-1) + \sum_{j\geq 0}\alpha^j\right) + \beta\left(\sum_{j=1}^{\nu}\alpha^{\ell_j}(z_j-1) + \sum_{j\geq 0}\alpha^j\right)\mathbb{E}z_1^{R^*_1}\cdots z_\nu^{R^*_\nu},
\end{aligned}$$

from which the formula (4) follows. Expanding the denominator in (4), we obtain

$$\begin{aligned}
\mathbb{E}z_1^{R^*_{\ell_1}}\cdots z_\nu^{R^*_{\ell_\nu}} &= \left(1+(1-\alpha)\sum_{j=1}^{\nu}\alpha^{\ell_j}(z_j-1)\right)\sum_{k=0}^{\infty}\left(\frac{(1-\alpha)\beta}{\gamma}\right)^k\left(\sum_{j=1}^{\nu}\alpha^{\ell_j}(z_j-1)\right)^k\\
&= 1+\sum_{k=1}^{\infty}\left(\frac{(1-\alpha)\beta}{\gamma}\right)^k\left(\sum_{j=1}^{\nu}\alpha^{\ell_j}(z_j-1)\right)^k + \frac{\gamma}{\beta}\sum_{k=1}^{\infty}\left(\frac{(1-\alpha)\beta}{\gamma}\right)^k\left(\sum_{j=1}^{\nu}\alpha^{\ell_j}(z_j-1)\right)^k\\
&= 1+\frac{1-\alpha}{\beta}\sum_{k=1}^{\infty}\left(\frac{(1-\alpha)\beta}{\gamma}\right)^k\sum_{\substack{i_1,\dots,i_\nu\\ i_1+\cdots+i_\nu=k}}\frac{k!}{i_1!\cdots i_\nu!}\,\alpha^{\ell_1\nu_1+\cdots+\ell_\nu i_\nu}(z_1-1)^{i_1}\cdots(z_\nu-1)^{i_\nu}.
\end{aligned}$$

Now,

$$\begin{aligned}
\mathbb{E}z_1^{R^*_{\ell_1}}\cdots z_\nu^{R^*_{\ell_\nu}} &= \mathbb{E}(1+(z_1-1))^{R^*_{\ell_1}}\cdots(1+(z_\nu-1))^{R^*_{\ell_\nu}}\\
&= \sum_{i_1,\dots,i_\nu}\mathbb{E}\binom{R^*_{\ell_1}}{r_1}\cdots\binom{R^*_{\ell_\nu}}{r_\nu}(z_1-1)^{i_1}\cdots(z_\nu-1)^{i_\nu},
\end{aligned}$$

and so formula (5) is obtained by inspection. □

Sometimes [2, 10] people are interested in the distribution of the number of runs exceeding a given length:

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

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 formula of Theorem 2 by replacing $z_j$ by $z_1\cdots z_j$:

**Corollary 7.**
$$\mathbb{E}z^{G^*} = \gamma \, \frac{1+\sum_{j\ge 1}\alpha^j z_1\cdots z_j}{1-\beta-\beta\sum_{j\ge 1}\alpha^j z_1\cdots z_j}.$$

Marginalizing, we see that

**Corollary 8.** *$G_\ell^*$ has* $\mathrm{geo}(\widetilde{\alpha}_\ell, \widetilde{\beta}_\ell)$ *distribution with*
$$\widetilde{\alpha}_\ell = \frac{\gamma(1-\alpha^\ell)}{\gamma+\beta\alpha^\ell}, \quad \widetilde{\beta}_\ell = \frac{\beta\alpha^\ell}{\gamma+\beta\alpha^\ell}$$

This follows from direct comparison of the $\mathbb{E}\theta^{G_\ell^*}$ with the formula for the probability generating function of a $\mathrm{geo}(\alpha,\beta)$ random variable.

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

Our interest is obtaining in information about the distributions of $R_\ell(n)$ and $G_\ell(n)$. Since $R_\ell^*$ and $G_\ell^*$ have both distribution of $\mathrm{geo}(\alpha,\beta)$ type, with explicitly known parameters, and since[3]
$$\mathcal{L}\{G_\ell^*\} = (1-w)\sum_{n\ge 0} w^n\,\mathcal{L}\{G(n)\}$$
(likewise for $R_\ell^*$) the problem is, in principle, solved. Moreover, such formulas exist in the numerous references. See, e.g., [2]. Our intent in this section is to give an independent derivation of the formulas but also point out their relations with hypergeometric functions.

It turns out that (i) formulas 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. We therefore start by computing the $r$-th binomial moment of $G_\ell(n)$. By Corollary 2, $G_\ell^*$ is a $\mathrm{geo}(\widetilde{\alpha}_\ell, \widetilde{\beta}_\ell)$ random variable, and, from the formulas following Definition 1, we have
$$\begin{aligned}\mathbb{E}\binom{G_\ell^*}{r} &= \frac{1-\widetilde{\alpha}_\ell}{\widetilde{\beta}_\ell}\left(\frac{\widetilde{\beta}_\ell}{1-\widetilde{\beta}_\ell}\right)^r = \frac{1-\alpha}{\beta}\left(\frac{\beta\alpha^\ell}{\gamma}\right)^r \\ &= (1-wp)(wq)^{r-1}(wp)^{\ell r}(1-w)^{-r} \\ &= (1-w)p^{\ell r}q^{r-1} \times \underbrace{(1-wp)w^{\ell r+r-1}(1-w)^{-r-1}}\end{aligned}$$

[3]If $X$ is a random variable, we let $\mathcal{L}\{X\}$ be its law.

Now use

$$(1-w)^{-r-1} = \sum_{k=0}^{\infty} \binom{r+k}{r} w^k \tag{6}$$

to get that the under-braced term above equals

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

So, by inspection,

$$\mathbb{E}\binom{G_\ell(n)}{r} = p^{\ell r} q^{r-1}\left[\binom{n+1-\ell r}{r} - p\binom{n-\ell r}{r}\right]. \tag{7}$$

In particular, we have

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

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

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

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

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

as expected by the ergodic theorem.

We now use the standard formula relating probabilities to binomial moments: [4]

$$\mathbb{P}(G_\ell(n) = x) = \sum_{r\geq x} (-1)^{r-x}\binom{r}{x}\mathbb{E}\binom{G_\ell(n)}{r}. \tag{8}$$

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

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

[4] The binomial coefficient $\binom{a}{b}$ is taken to be zero if $b > a$ or if $a < 0$.

It is perhaps interesting to notice the relation of the distribution of $G_\ell(n)$ to hypergeometric functions. Recall the notion of the hypergeometric function [9, Section 5.5.] (the notation is from this book and is not standard):

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

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 $x^{\overline{k}} := x(x+1) \cdots (x+k-1)$. A little algebra gives

$$H_\ell(x, y; z) := \sum_{m \ge 0} z^m \binom{x+m}{x} \binom{x+y-\ell m}{x+m} = \binom{x+y}{x} F\left(\begin{matrix} \boldsymbol{V}_{\ell+1}(y) \\ \boldsymbol{V}_\ell(x+y) \end{matrix} \middle| -\frac{(\ell+1)^{\ell+1}}{\ell^\ell} z\right), \tag{9}$$

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

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

Looking back at the expression for $\mathbb{P}(G_\ell(n) = x)$ we recognize that the two terms in the bracket are expressible in terms of the function $H_\ell$:

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

The point is that the probabilities $\mathbb{P}(G_\ell(n) = x)$ are expressible in terms of the function $H_\ell$ which is itself expressible in terms of a hypergeometric function as in (9). This does not solve the problem other than it expresses the distribution of $G_\ell(n)$ via hypergeometric functions which are efficiently computable in standard computer packages (we use Maple$^{\mathrm{TM}}$.)

Ultimately, the hypergeometric functions appearing above are nothing but polynomials. So the problem is, by nature, of combinatorial character. Instead of digging 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 $\alpha, \beta, \gamma$

are given by (1) and that

$$\mathcal{L}\{h(R^*)\} = \sum_{n\geq 0}(1-w)w^n\,\mathcal{L}\{h(R(n)\},$$

we take each of the terms in (2) and bring out its dependence on $w$ explicitly. The left-hand side of (2) is

$$\mathbb{E}h(R^*) = (1-w)\sum_{n\geq 0} w^n\,\mathbb{E}h(R(n)). \tag{10}$$

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

$$\gamma\sum_{n\geq 0}\alpha^j h(e_n) = (1-w)\sum_{n\geq 0} w^n p^n h(e_n). \tag{11}$$

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

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

Change variables by

$$(j,\,n)\mapsto(j,\,m=1+j+n)$$

to further write

$$\begin{aligned}\beta\sum_{j\geq 0}\alpha^j\mathbb{E}h(R^*+e_j) &= (1-w)q\sum_{m\geq 0}\sum_{0\leq j\leq m-1} w^m p^j\mathbb{E}h(R(m-j-1)+e_j)\\ &= (1-w)\sum_{n\geq 0} w^n q\sum_{0\leq j\leq n-1} p^j\mathbb{E}h(R(n-j-1)+e_j). \end{aligned}\tag{12}$$

Using (2) and (10), (11), (12), we obtain

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

$$\mathbb{E}h(R(n)) = q\sum_{j=0}^{n-1} p^j\mathbb{E}h(R(n-j-1)+e_j) + p^n h(e_n). \tag{13}$$

**Remark 1.** *(i) We say "any function" because $R(n)$ 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 $\mathbb{E}h(R(0)) = h(0)$.*

*(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 asked Maple to run the recursion a few times and here is what it found:*

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

*which could be interpreted combinatorially.*

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

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

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

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

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

These two recursions can be made 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

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

and immediately obtain

**Corollary 10.** *The probability generating functions $\Phi_n$ and $\Psi_n$ of the random elements $R(n)$ 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 $\ell$-th marginal of the law of $G(n)$. 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

$$\Psi_{n,\ell}(\theta) = \Psi_n(\mathbf{1} + (\theta - 1)e_\ell)$$

where $\mathbf{1} \in \mathbb{Z}^{\mathbb{N}}$ is the infinite repetition of 1's. We thus have

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

$$\Psi_{n,\ell} = \begin{cases} q \displaystyle\sum_{j=0}^{n-\ell-1} p^j \Psi_{n-j-1,\ell} + p^{n-\ell} + (\theta-1)p^\ell, & \ell \le n \le 2\ell \\ q \displaystyle\sum_{j=0}^{\ell-1} p^j \Psi_{n-j-1,\ell} + q\theta \sum_{j=\ell}^{n-\ell-1} p^j \Psi_{n-j-1,\ell} + q\theta(p^{n-\ell} - p^\ell) + \theta p^m, & n \ge 2\ell+1 \end{cases}$$

## 5 Longest run, Poisson and other approximations

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

$$(1-w)\sum_{n=0}^\infty w^n \mathbb{P}(L(n) < \ell) = \mathbb{P}(L^* < \ell) = \frac{(1-w)(1-(wp)^\ell)}{1-w+(wq)(wp)^\ell}, \tag{14}$$

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

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

and use the binomial moment formula (7) to obtain

$$F_\ell(n) := \mathbb{P}(L(n) < \ell) = 1 + \sum_{r \ge 1} (-1)^r \left[\binom{n-\ell r}{r} p^{\ell r} q^r + \binom{n-\ell r}{r-1} p^{\ell r} q^{r-1}\right]. \tag{15}$$

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

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

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

*Proof.* This can be proved directly by induction. But, since Theorem 3 is available, set $h(x) := \mathbf{1}\{\overline{\lambda}(x) < \ell\}$, observe that $h(x + e_j) = h(x)\mathbf{1}\{j < \ell\}$ and substitute into (13). $\square$

### 5.1 The Poisson regime for large lengths

According to Feller, [5, Section XIII.12, Problem 25, page 341] asymptotics for $L(n)$ go back to von Mises [14]. 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 [13]. But it is a matter of elementary analysis to see that the distribution function $\ell \mapsto F_\ell(n)$ exhibits a cutoff at $\ell$ of the order of magnitude of $\log n$. To see this in a few lines, consider the formula (7) for the binomial moment of $G_\ell(n)$. Then

**Lemma 1.** *Keep $0 < p < 1$ fixed and let $\ell = \ell(n) \to \infty$ so that*

$$np^{\ell(n)}q \to \theta,$$

*as $n \to \infty$, for some $\theta > 0$. Then*

$$\mathbb{E}\binom{G_{\ell(n)}(n)}{r} \to \frac{\theta^r}{r!},$$

*and*

$$\mathbb{P}(G_{\ell(n)}(n) = 0) \to e^{-\theta}.$$

*Proof.* Expanding the binomial coefficients in (7),

$$\begin{aligned}\mathbb{E}\binom{G_{\ell(n)}(n)}{r} &= \frac{1}{r!}\prod_{j=0}^{r-1}\left[(n-\ell(n)r-j)p^{\ell(n)}q\right] + \frac{1}{(r-1)!}\prod_{j=0}^{r-2}\left[(n-\ell(n)r-j)p^{\ell(n)}q\right] \times p^{\ell(n)} \\ &\to \frac{1}{r!}\theta^r + \frac{1}{(r-1)!}\theta^{r-1} \times 0 = \frac{\theta^r}{r!}, \qquad \text{as } n \to \infty.\end{aligned}$$

For the second assertion, use (8):

$$\mathbb{P}(G_{\ell(n)}(n) = 0) = \sum_{r\geq 0}(-1)^r\mathbb{E}\binom{G_{\ell(n)}(n)}{r} \to \sum_{r\geq 0}(-1)^r\frac{\theta^r}{r!} = e^{-\theta}, \qquad \text{as } n \to \infty.$$

□

Since $\mathbb{P}(G_\ell(n) = 0) = \mathbb{P}(L(n) < \ell)$ for all $n$ and $\ell$, the last asymptotic result can be translated immediately into the following:

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

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

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

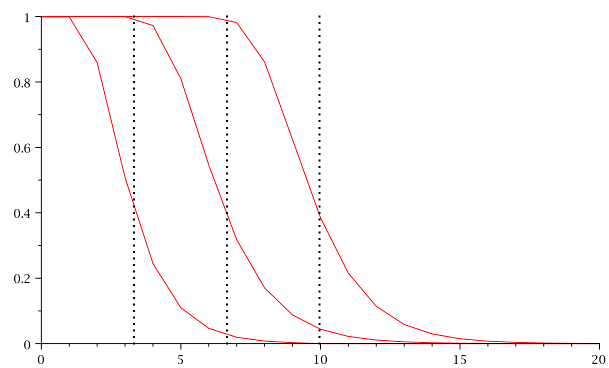


Figure 1: *Plot (piecewise linear interpolation) of* $\ell \mapsto \mathbb{P}(L(n) \geq \ell)$ *for* $n = 10, 100, 1000$ *and* $p = 1/2$. *The vertical lines are at* $\ell = \log_2(n)$.

Corollary 12 suggests the following practical approximation:

$$\mathbb{P}(L(n) < \ell) \doteq \exp(-\mathbb{E}G_\ell(n)) = \exp(-(n-\ell)p^\ell q - p^\ell), \tag{16}$$

which is valid for large $n$ and $\ell$, roughly when $\ell$ is of order $\log_{1/p} n$ or higher. In table 5.1 we compare the exact result with the approximation for $n = 10^4$, $p = 1/2$, and $\ell$ ranging from slightly below $\log_2 10^4 \approx 13.288$ to much higher values. We programmed (15) in Maple to obtain the exact values of $\mathbb{P}(L(n) \geq \ell)$. We note that for values of $\ell$ smaller than 10, a straightforward coding of (15) is too time-consuming and point out to Section 5.2 for better approximations in this case.

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 (16). Notice 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 $n \to \infty$ to a Poisson law with mean $\theta$.

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

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

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

| $\ell$ | $\mathbb{P}(L(n) \geq \ell)$ | $1 - \exp(-(n-\ell)p^\ell q - p^\ell)$ |
|---|---|---|
| 10 | 0.992583894386551 | 0.992394672192560 |
| 11 | 0.913367688920047 | 0.912770175666911 |
| 12 | 0.705167040532444 | 0.704616988848744 |
| 13 | 0.456748458590744 | 0.456475326366319 |
| 14 | 0.262835671849087 | 0.262736242068365 |
| 15 | 0.141377186760083 | 0.141346252684305 |
| 20 | 0.004748524931253 | 0.004748478671106 |
| 50 | $4.41957581641815 \times 10^{-12}$ | $4.42000000000001 \times 10^{-12}$ |

*Proof.* It is enough to establish convergence of binomial moments to those of a Poisson law. Recall that if $N$ is 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. $\square$

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

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

$$\lim_{n\to\infty} np^{\ell_j(n)}q = \theta_j, \quad j = 1, \dots, \nu.$$

*Then*

$$\mathcal{L}\{R_{\ell_1(n)}(n), \dots, R_{\ell_\nu(n)}(n)\} \to \text{Poisson}(\theta_1 q) \times \cdots \times \text{Poisson}(\theta_\nu q),$$

*as $n \to \infty$.*

*Proof.* It suffices to show that the joint binomial moments converge to the right thing, namely, that

$$\lim_{n\to\infty} \mathbb{E}\binom{R_{\ell_1(n)}(n)}{r_1} \cdots \binom{R_{\ell_\nu(n)}(n)}{r_\nu} = \frac{(\theta_1 q)^{r_1}}{r_1!} \cdots \frac{(\theta_\nu q)^{r_\nu}}{r_\nu!},$$

for all nonnegative integers $r_1, \dots, r_\nu$. Fix $\ell_1, \dots, \ell_\nu$, $r_1, \dots, r_\nu$, set $r_0 = r_1 + \cdots + r_\nu$, and, using the abbreviations (1) for $\alpha, \beta$ and $\gamma$, write the expression (5) for the joint binomial

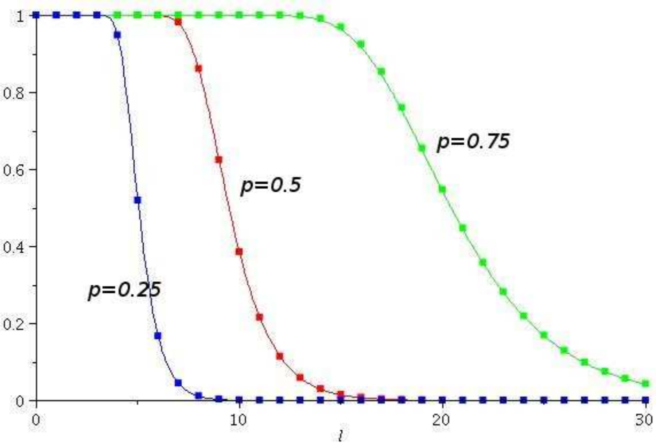


Figure 2: *Plot of* $\ell \mapsto \mathbb{P}(L(1000) \geq \ell)$ *and* $p = 0.25, 0.5, 0.75$. *The dots correspond to the actual values. The solid lines correspond to the analytical approximation* (16)

moments as

$$\mathbb{E}\binom{R^*_{\ell_1}}{r_1}\cdots\binom{R^*_{\ell_\nu}}{r_\nu} = (1-w)\,\frac{r_0!}{r_1!\cdots r_\nu!}\,p^{\boldsymbol{\ell}\cdot\boldsymbol{r}}q^{r_0-1}w^{\boldsymbol{\ell}\cdot\boldsymbol{r}+r_0-1}\frac{(1-wp)^{r_0+1}}{(1-w)^{r_0+1}}$$

Expand $(1-wp)^{r_0+1}$ using the binomial formula, and $(1-w)^{r_0+1}$ using (6) 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

$$\begin{aligned}\mathbb{E}\binom{R^*_{\ell_1}}{r_1}\cdots\binom{R^*_{\ell_\nu}}{r_\nu} &= (1-w)\,\frac{r_0!}{r_1!\cdots r_\nu!}\,p^{\boldsymbol{\ell}\cdot\boldsymbol{r}}q^{r_0-1}\sum_{k=0}^{\infty}\sum_{s=0}^{r_0+1}(-p)^s w^{\boldsymbol{\ell}\cdot\boldsymbol{r}+r_0-1+k+s}\binom{r_0+1}{s}\binom{r_0+k}{r_0}\\ &= (1-w)\sum_{n=0}^{\infty}w^n\sum_{s=0}^{r_0+1}\frac{r_0!}{r_1!\cdots r_\nu!}\,p^{\boldsymbol{\ell}\cdot\boldsymbol{r}}q^{r_0-1}(-p)^s\binom{r_0+1}{s}\binom{n+1-\boldsymbol{\ell}\cdot\boldsymbol{r}-s}{r_0}.\end{aligned}$$

Therefore,

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

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

$$\lim_{n\to\infty}(n+1-\boldsymbol{\ell}\cdot\boldsymbol{r}-s)_{r_0}\prod_{j=1}^{\nu}(p^{\ell_j}q)^{r_j} = \theta_1^{r_1}\cdots\theta_\nu^{r_\nu},$$

and so

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\binom{R_{\ell_1(n)}(n)}{r_1}\cdots\binom{R_{\ell_\nu(n)}(n)}{r_\nu} &= q^{-1}\sum_{s=0}^{r_0+1}(-p)^s\binom{r_0+1}{s}\frac{\theta_1^{r_1}}{r_1!}\cdots\frac{\theta_\nu^{r_\nu}}{r_\nu!}\\ &= q^{-1}(1-p)^{r_0+1}\frac{\theta_1^{r_1}}{r_1!}\cdots\frac{\theta_\nu^{r_\nu}}{r_\nu!}\\ &= q^{r_1+\cdots+r_\nu}\frac{\theta_1^{r_1}}{r_1!}\cdots\frac{\theta_\nu^{r_\nu}}{r_\nu!},\end{aligned}$$

establishing the assertion. □

## 5.2 A better approximation for small length values

We now pass on to a different approximation for $F_\ell(n) = \mathbb{P}(L(n) < \ell)$. Consider again (14),

$$\sum_{n=0}^{\infty} w^n F_\ell(n) = \frac{1-(wp)^\ell}{1-w+(wq)(wp)^\ell}, \tag{17}$$

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 $\ell+1$. It is the smallest zeros of $f(w)$ which govern the behavior of $n \mapsto F_\ell(n)$, for $n$ large (and all $\ell$.)

**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*

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

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

*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 $[0,\infty)$. Notice that $f(w^*) = 1 - \ell w^*/(\ell+1)$. We claim that $f(w^*) \le 0$, or, equivalently, that $w^* \ge (\ell+1)/\ell$. Upon substituting with the value of $w^*$, this last inequality is equivalent to $p^\ell(1-p) \le (\ell/(\ell+1))^\ell$. But this is true, since $\max_{0\le p\le 1} p^\ell(1-p) = \ell^\ell/(\ell+1))^{\ell+1} \le (\ell/(\ell+1)^\ell)$. Hence $f(w^*) \le 0$ with equality if and only if $p = \ell/(\ell+1)$. On the other hand, $f(0) = 1$ and $\lim_{w\to\infty} f(w) = \infty$. 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 $f((\ell+1)/\ell) = -\frac{1}{\ell} + ((\ell+1)/\ell)^{\ell+1}p^\ell q < 0$, and $f(w^*) < 0$, provided that $p \neq \ell/(\ell+1)$, it actually 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 $\ell$, we have $p < \ell/(\ell+1)$ and so $1 < w_0 < (\ell+1)/\ell$, showing that the limit of $w_0$, as $\ell \to \infty$, is 1. To show that the only roots $f(w) = 0$ with $|w| \le \max(w_0, 1/p)$ 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 $P(z) := c_0 + c_1 z + \cdots + c_n z^n$, $z \in \mathbb{C}$, with real coefficients such that $c_0 > c_1 > \cdots > c_n > 0$. Then all the zeros of $P(z)$ lie outside the closed unit ball centered at the origin.*

*Proof.* Fix $\lambda > 1$ such that $c_0 > c_1/\lambda > c_2/\lambda^2 > \cdots > c_n/\lambda^n$ and notice that

$$c_0 + (z-1)\,P(z/\lambda) = (c_0 - \frac{c_1}{\lambda})z + (\frac{c_1}{\lambda} - \frac{c_2}{\lambda^2})z^2 + \cdots + (\frac{c_{n-1}}{\lambda^{n-1}} - \frac{c_n}{\lambda^n})z^n + \frac{c_n}{\lambda^n}z^{n+1}.$$

Therefore, on $|z| = 1$,

$$\big|c_0 + (z-1)\,P(z/\lambda)\big| \le (c_0 - \frac{c_1}{\lambda}) + (\frac{c_1}{\lambda} - \frac{c_2}{\lambda^2}) + \cdots + (\frac{c_{n-1}}{\lambda^{n-1}} - \frac{c_n}{\lambda^n}) + \frac{c_n}{\lambda^n} = c_0 = |-c_0|.$$

Rouché's theorem [1, page 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. □

*End of proof of Proposition 3.* By polynomial division, write $f(w)$ as

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

where

$$P(z) = c_0 + c_1 z + \cdots + c_{\ell-1} z^{\ell-1},$$

and

$$c_j = \frac{1-(w_0 p)^{\ell-j}}{1-w_0 p}\, p^j, \quad j = 0, 1, \ldots, \ell-1.$$

The polynomial $P(z)$ satisfies the assumptions of Lemma 2 and thus all its zeros lie outside the closed unit circle centered at the origin. $\square$

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 $p \neq \ell/(\ell+1)$ then* [5]

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

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

$$\mathbb{P}(L(n) < \ell) \sim 2(\ell/(\ell+1))^{n+1}, \text{ as } n \to \infty.$$

*Proof.* Suppose first that $p \neq \ell/(\ell+1)$ and, using partial fraction expansion, write the expression (17) as

$$\frac{g(w)}{f(w)} = \frac{1-(wp)^{\ell}}{1-w+(wq)(wp)^{\ell}} = \frac{c_0}{1-w/w_0} + \frac{h(w)}{j(w)}. \tag{18}$$

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

$$c_0 = \frac{1}{w_0} \lim_{w \to w_0} \frac{(w_0 - w)g(w)}{f(w)} = \frac{g(w_0)}{-f'(w_0)} \frac{1}{w_0} = \frac{1-(w_0 p)^{\ell}}{1-(\ell+1)q(w_0 p)^{\ell}} \frac{1}{w_0}.$$

Now, the zeros of $j(w)$ are, from Proposition 3, all outside the circle with radius $w_0$. Using this result, and inverting the expression (18), we obtain

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

which proves the first assertion.

[5] $a(n) \sim b(n)$ stands for $\lim a(n)/b(n) = 1$.

The case $p = \ell(\ell+1)$ corresponds, again from the result of Proposition 3, to the case where the numerator $g(w)$ has a simple zero at $1/p$ and the denominator $f(w)$ a double zero at $w_0 = 1/p$. Taking this into account and the partial fraction expansion (18), we find

$$c_0 = -\frac{g'(w_0)}{\frac{1}{2}f''(w_0)}\,\frac{1}{w_0}.$$

But $g'(w) = -p^\ell q w^{\ell-1}$, $f''(w) = p^\ell q \ell(\ell+1) w^{\ell-1}$, so

$$c_0 = \frac{2}{q(\ell+1)}\,\frac{1}{w_0} = \frac{2}{(1-p)(\ell+1)} = \frac{2}{w_0}.$$

Therefore,

$$\mathbb{P}(L(n) < \ell) \sim \frac{2}{w_0}\,w_0^{-n} = 2w_0^{-n-1} = 2(\ell/(\ell+1))^{-n-1},$$

proving the second assertion. □

Since these approximations are valid for all $\ell$, they nicely complement the Poisson approximation discussed earlier.

For $n, \ell \to \infty$, 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.

### 5.3 Numerical comparisons of the two approximations

We numerically compute $\mathbb{P}(L(n) \geq \ell)$, first using the exact formula (15), then using the Poisson approximation (16), and finally using the approximation suggested by Proposition 4. We see, as expected, that for small values of $\ell$ compared to $n$, the second approximation outperforms the first one.

First, we let $\ell = 2$. Then

$$f(w) = 1 - w + p^2 q w^3 = (pw-1)(pqw^2 + qw - 1),$$

and so

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

The approximation suggested by Proposition 4 is

$$\mathbb{P}(L(n)<2) \sim \frac{1-(w_0p)^2}{1-3q(w_0p)^2}\, w_0^{-n+1},$$

assuming that $p \neq 2/3$, whereas, for $p = 2/3$, we have

$$\mathbb{P}(L(n)<2) \sim 2(2/3)^{-(n+1)}.$$

Let $p = 1/2$, $1/3$ and $4/5$.

*$\mathbb{P}(L(n) \geq \ell)$ for $p = 1/2$ and $\ell = 2$*

| $n$ | exact | Poisson approx. (error) | second approx. (error) |
|---|---|---|---|
| 5 | 0.59375 | 0.59375 (21.7%) | 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%) |

*$\mathbb{P}(L(n) \geq \ell)$ for $p = 1/3$ and $\ell = 2$*

| $n$ | exact | Poisson approx. (error) | second approx. (error) |
|---|---|---|---|
| 5 | 0.32510 | 0.28347 (12.8%) | 0.64762 (99%) |
| 7 | 0.44033 | 0.38213 (13.2%) | 0.44080 (0.11%) |
| 10 | 0.57730 | 0.50525 (12.5%) | 0.57780 (0.09%) |
| 20 | 0.83415 | 0.76411 (8.4%) | 0.83454 (0.05%) |

*$\mathbb{P}(L(n) \geq \ell)$ for $p = 4/5$ and $\ell = 2$*

| $n$ | exact | Poisson approx. (error) | second approx. (error) |
|---|---|---|---|
| 5 | 0.94208 | 0.64084 (32.0%) | 0.94382 (0.18%) |
| 7 | 0.98509 | 0.72196 (26.71%) | 0.98526 (0.0173%) |
| 10 | 0.9980232 | 0.8106201 (18.78%) | 0.9980173 (0.00059%) |
| 20 | 0.9999975 | 0.9473453 (5.265%) | 0.9999975 ($10^{-7}$%) |

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

$\mathbb{P}(L(n) \geq \ell)$ *for* $p = 1/2$ *and* $\ell = 7$

| $n$ | exact | Poisson approx. (error) | second approx. (error) |
|---|---|---|---|
| 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 | 0.9999941986 | 0.9999916997 (0.00025%) | 0.9999931928 (0.00010%) |

$\mathbb{P}(L(n) \geq \ell)$ *for* $p = 2/3$ *and* $\ell = 10$

| $n$ | exact | Poisson approx. (error) | second approx. (error) |
|---|---|---|---|
| 100 | 0.43531 | 0.41583 (4.475%) | 0.46433 (6.667%) |
| 500 | 0.95209 | 0.94214 (1.045%) | 0.95480 (0.285%) |
| 1500 | 0.999900 | 0.999821 (0.00790%) | 0.999905 (0.00050%) |
| 3000 | 0.9999999904 | 0.9999999694 ($2.1\times10^{-6}$%) | 0.9999999908 ($0.04\ \times10^{-6}$%) |

# 6 Discussion and open problems

We mention the paper of Gordon, Schilling and Waterman [7] developing 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 older paper by Guibas and Odlyzko [8].

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, especially in the light of suitable applications.

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


## Acknowledgments

We would like to thank Joe Higgins for the short proof of Lemma 2.

## References


[1] L. V. Ahlfors (1979). *Complex Analysis.* Mc-Graw Hill, New York.

[2] N. Balakrishnan and M. Koutras (2002). *Runs and Scans with Applications.* Wiley, New York.

[3] P. Erdős and A. Rényi (1970). On a new law of large numbers. *J. Analyse Math.* **22**, 103-111.

[4] P. Erdős and P. Révész (1976). On the length of the longest headrun. *Topics in Information Theory (Second Colloq., Keszthely, 1975)*, pp. 219-228. Colloq. Math. Soc. János Bolyai, Vol. 16, North-Holland, Amsterdam, 1977.

[5] W. Feller (1968). *An Introduction to Probability Theory and its Applications, Vol. I.* Third edition, John Wiley, New York.

[6] G. Grimmett and D. Stirzaker (2001). *Probability and Random Processes,* 3d Ed. Oxford University Press, Oxford.

[7] L. Gordon, M.F. Schilling and M.S. Waterman (1986). An extreme value theory for long head runs. *Probab. Theory Relat. Fields* **72**, 279-287.

[8] L.J. Guibas and A.M. Odlyzko (1980). Long repetitive patterns in random sequences. *Z. Wahrsch. verw. Gebiete* bf 53, 241-262.

[9] R.L. Graham, D.E. Knuth and O. Patashnik (1994). *Concrete Mathematics.* Addison-Wesley, Reading, Massachusetts.

[10] F.S. Makri and Z.M. Psillakis (2011). On success runs of length exceeded a threshold. *Methodol. Comput. Appl. Probab.* **13**, 269-305.

[11] D.B. Murray and S.W. Scott (1993). Probability of a tossed coin landing on edge. *Physical Review E (Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics)*, **48**, 2547-2552.

[12] A. Philippou and F.S. Makri (1986). Success runs and longest runs. *Stat. Prob. Letters*, **4**, 101-105.

[13] P. Révész (1980). Strong theorems on coin tossing. *Proc. International Congress of Mathematicians (Helsinki, 1978)*, pp. 749-754, *Acad. Sci. Fennica, Helsinki.*

[14] R. von Mises (1921). Zur Theorie der Iterationen. *Zeitschr. für angew. Mathem. und Mechan.* **1**, 298-307.

[15] R. von Mises (1981). *Probability, Statistics and Truth.* Dover, New York. (Originally published in German in 1928; this is a translation from the third 1957 German edition.)

Lars Holst
Department of Mathematics
Royal Institute of Technology
SE-10044 Stockholm
Sweden
Email: lholst@kth.se

Takis Konstantopoulos
Department of Mathematics
Uppsala University
SE-75106 Uppsala
Sweden
E-mail: takis@math.uu.se