HOW MANY COIN TOSSES WOULD YOU NEED UNTIL YOU GET n HEADS OR m TAILS

SVANTE JANSON, LUCY MARTINEZ, AND DORON ZEILBERGER

Abstract. We harness both human ingenuity and the power of symbolic computation to study the number of coin tosses until reaching n Heads or m Tails. We also talk about the closely related problem of reaching n Heads and m Tails. This paper is accompanied by a Maple package that enables fast computation of expectations, variances, and higher moments of these quantities.

1. Preface

If you toss a coin whose probability of heads is p, until you reach n Heads, you should expect to make n/p coin tosses, and the variance and higher moments are easily derived from the explicit probability generating function, (as usual $q := 1 - p$)

$$\sum_{k=0}^{\infty} \binom{n+k-1}{n-1} (px)^n (qx)^k \;=\; \left(\frac{px}{1-qx}\right)^n ,$$

which is essentially the *negative-binomial distribution* [4] (note that usually one only counts the number of Tails until you reach n Heads, but we are interested in the total number of coin-tosses, so we add the n Heads). From this probability generating function we can extract not only the expectation, n/p, but also the variance $\frac{n(1-p)}{p^2}$, and by repeated differentiation with respect to x, and plugging in $x = 1$, we can easily derive explicit expressions of as many as desired *factorial moments*, that in turn, yield the *moments*, and from them the *central moments*. Then we can compute the *scaled central moments*, take the limit as $n \to \infty$ and prove that for a **fixed** p it tends to the good old Normal Distribution. Of course, in this simple case we can also derive a *local limit law*.

But what if you are not a *Headist*? What if you like Tails just as much, and stop as soon as you get n Heads OR m Tails? Another interesting stopping rule is to make **both** Heads and Tails happy and keep tossing until you get n Heads AND m Tails. Now things are not as nice and simple. Nevertheless, using Wilf-Zeilberger algorithmic proof theory [3], we can derive the next-best thing, linear recurrences that enable very fast computation of these quantities. These will be presented in Section 2.

In the special case of a *fair* coin, and the same desired number of Heads and Tails (let's call it n), we get, **explicit** expressions not only for the expectation and variance, but for as many as-desired moments (we went up to 200^{th}, but could go much further). Then we (or rather Maple) can compute the scaled limit, and **surprise!** they coincide **exactly** with the central-scaled moments of $-|N(0,1)|$, the continuous probability distribution whose *probability density function* (pdf) is

$$\frac{e^{-x^2/2}}{\sqrt{\pi/2}},$$

supported in $-\infty < x < 0$. This will be accomplished in Section 3.

While we (or rather our computer) can prove this convergence for the first 200 moments, and with a larger computer, the first 2000, we can **not** prove it for *all* moments. In the last three sections of this article we will prove it completely, using purely human-generated, paper-and-pencil mathematics.

1.1. The Maple package.

`CoinToss.txt` is freely available from
`https://sites.math.rutgers.edu/~zeilberg/tokhniot/CoinToss.txt`,

2. Recurrences for the Duration with a Loaded Coin and Different Target Goals

You are tossing a coin whose probability of Heads is p (and hence the probability of Tails is $q := 1 - p$). We consider two random variables

- $X_1(n, m; p)$: The number of tosses until reaching (for the first time) either n Heads OR m tails, and
- $X_2(n, m; p)$: The number of tosses until reaching (for the first time) n Heads AND m tails.

The **probability generating function** of X_1, in x, let's call it $F_1(n, m; p)(x)$ is

$$F_1(n,m;p)(x) = (qx)^m \sum_{h=0}^{n-1} \binom{h+m-1}{m-1} (px)^h + (px)^n \sum_{t=0}^{m-1} \binom{t+n-1}{n-1} (qx)^t.$$

On the other hand, that of $X_2(n, m; p)$, let's call it $F_2(n, m; p)(x)$ is:

$$F_2(n,m;p)(x) = (qx)^m \sum_{h=n}^{\infty} \binom{h+m-1}{m-1} (px)^h + (px)^n \sum_{t=m}^{\infty} \binom{t+n-1}{n-1} (qx)^t.$$

Unlike the probability generating function for the Negative Binomial distribution [4], $F_1(n, m; p)(x)$ and $F_2(n, m; p)(x)$ do not have **closed-form**, but thanks to **Wilf-Zeilberger Algorithmic Proof theory** [3,5], they have the next-best thing, linear recurrences with polynomial coefficients (that happen to be third-order), in each of n and m, that enable an efficient compilation of a table of these. We observe that $F_1 + F_2$ *does* have a closed-form.

$$F_1(n,m;p)(x) + F_2(n,m;p)(x) = \left(\frac{qx}{1-px}\right)^m + \left(\frac{px}{1-qx}\right)^n.$$

The actual recurrences are too complicated to reproduce here but can be looked up from the output file, `https://sites.math.rutgers.edu/~zeilberg/tokhniot/oCoinToss2.txt`.

Let $L_1(n, m; p)$ be the expectation of $X_1(n, m; p)$ and let $L_2(n, m; p)$ be the expectation of $X_2(n, m; p)$, then both satisfy the **same** system of third-order linear pure recurrences. We have the following pure recurrences, in n and m, respectively (below $L(n, m)$ stands for either $L_1(n, m; p)$ and $L_2(n, m; p)$):

$$\begin{aligned}
L(n,m) =& \frac{(pn+pm-2p+2n-2)}{n-1} \cdot L(n-1,m) \\
& - \frac{(2pn+2pm-4p+n-1)}{n-1} \cdot L(n-2,m) + \frac{p\,(m-2+n)}{n-1} \cdot L(n-3,m)\,, \\
L(n,m) =& - \frac{(pn+pm-2p-n-3m+4)}{m-1} \cdot L(n,m-1) \\
& + \frac{(2pn+2pm-4p-2n-3m+5)}{m-1} \cdot L(n,m-2) - \frac{(-1+p)\,(m-2+n)}{m-1} \cdot L(n,m-3)\,.
\end{aligned}$$

Of course $L_1(n, m; p)$ and $L_2(n, m; p)$ differ in the *initial conditions*. Here they are:

$$\begin{aligned}
&[[L_1(1,1), L_1(1,2), L_1(1,3)], [L_1(2,1), L_1(2,2), L_1(2,3)], [L_1(3,1), L_1(3,2), L_1(3,3)]] = \\
&[[1, -p+2, p^2-3p+3], [p+1, -2p^2+2p+2, 3p^3-7p^2+3p+3], \\
&[p^2+p+1, -3p^3+2p^2+2p+2, 6p^4-12p^3+3p^2+3p+3]],
\end{aligned}$$

and

$$\begin{aligned}
&[[L_2(1,1), L_2(1,2), L_2(1,3)], [L_2(2,1), L_2(2,2), L_2(2,3)], [L_2(3,1), L_2(3,2), L_2(3,3)]] = \\
&\left[\left[-\frac{p^2-p+1}{p\,(p-1)}, \frac{p^3-3p^2+p-1}{p\,(p-1)}, -\frac{p^4-4p^3+6p^2-p+1}{p\,(p-1)}\right],\right. \\
&\left[-\frac{p^3-2p+2}{p\,(p-1)}, \frac{2p^4-4p^3+2p-2}{p\,(p-1)}, -\frac{3p^5-10p^4+10p^3-2p+2}{p\,(p-1)}\right], \\
&\left.\left[-\frac{p^4-3p+3}{p\,(p-1)}, \frac{3p^5-5p^4+3p-3}{p\,(p-1)}, -\frac{3\left(2p^6-6p^5+5p^4-p+1\right)}{p\,(p-1)}\right]\right].
\end{aligned}$$

These recurrences are implemented in procedures `fAveF(n,m,p)` and `FaveF(n,m,p)` respectively. For example to find the expected number of coin-tosses it takes if you toss a loaded coin whose probability of Heads is $\frac{1}{3}$ until it reaches, for the first time $100\,i$ Heads OR $200\,i$ Tails, for $1 \leq i \leq 7$, type:

```
restart:  read ‘CoinToss.txt‘:t0:=time():
evalf([seq(faveF(100*i,200*i,1/3),i=1..7)]);time()-t0;
```

getting

$$[285.3561686, 579.2804255, 874.6196952, 1170.690974, 1467.229920, 1764.101012, 2061.223764],$$

and it took 0.563 seconds.

For comparison, if you do it directly, **not** using the recurrence, but rather the definition as a sum, typing

```
restart:  read ‘CoinToss.txt‘:  t0:=time():
evalf([seq(fave(100*i,200*i,1/3),i=1..7)]);time()-t0;
```

you would get the same output, but it took more than 12 seconds.

What about $L_2(r,s,p)$, i.e. $L_2(100i, 200i, \frac{1}{3})$? Type:

```
restart:  read ‘CoinToss.txt‘:
t0:=time():evalf([seq(FaveF(100*i,200*i,1/3),i=1..7)]);time()-t0;
```

getting

$$[314.6438314, 620.7195745, 925.3803048, 1229.309026, 1532.770080, 1835.898988, 2138.776236],$$

and this took 0.561 seconds, and the direct way took more than 12 seconds.

We observe that for positive integers a and b we have the **explicit** expressions

$$L_1(an, bn, \frac{a}{a+b}) = (a+b)n\left(1 - \frac{((a+b)n)!}{(an)!(bn)!}\cdot\left(\frac{a^a b^b}{(a+b)^{a+b}}\right)^n\right),$$

$$L_2(an, bn, \frac{a}{a+b}) = (a+b)n\left(1 + \frac{((a+b)n)!}{(an)!(bn)!}\cdot\left(\frac{a^a b^b}{(a+b)^{a+b}}\right)^n\right).$$

These are asymptotically

$$(a+b)n\left(1 \pm \sqrt{\frac{a+b}{2ab\pi}}\cdot\frac{1}{\sqrt{n}}\right).$$

So $L_1(an, bn, \frac{a}{a+b})/((a+b)(n))$ and $L_2(an, bn, \frac{a}{a+b})/((a+b)(n))$ converge slowly (as $n^{-1/2}$) to 1 as n goes to infinity. On the other hand if $p > \frac{a}{a+b}$ then $L_1(an,bn,p)/n$ and $L_2(an,bn,p)/n$ converge exponentially fast to bp, and if $p < \frac{a}{a+b}$ then they converge exponentially fast to ap. This makes sense, since when a coin is loaded in favor of achieving your goal you should expect to achieve your goal only a bit later than if the other side of the coin didn't matter.

3. Explicit expressions for the moments of the number of tosses until getting n Heads or n Tails with a fair coin

The probability generating function for the number of tosses until a fair coin reaches n Heads or n Tails is

$$(\frac{1}{2}x)^n \sum_{h=0}^{n-1} \binom{h+n-1}{n-1} (\frac{1}{2}x)^h + (\frac{1}{2}x)^n \sum_{t=0}^{n-1} \binom{t+n-1}{n-1} (\frac{1}{2}x)^n = (\frac{1}{2})^{n-1} \sum_{h=0}^{n-1} \binom{h+n-1}{n-1} (\frac{1}{2})^h x^{h+n}.$$

Recall that the r-th **factorial moment** of a random variable is

$$\mathbb{E}\left[X(X-1)\dots(X-r+1)\right] = r!\,\mathbb{E}\left[\binom{X}{r}\right].$$

Let $A(n,r)$ be the the r-th factorial moment of our random variable (number of tosses of a fair coin until you get for the first time n Heads or n Tails). We have:

$$A(n,r) = (\frac{1}{2})^{n-1} \sum_{h=0}^{n-1} \binom{h+n-1}{n-1} r! \binom{h+n}{r} (\frac{1}{2})^h.$$

For each *specific* r, this can be evaluated as a closed-form expression in n, and Maple can do it easily for small r, but as r gets larger, it becomes harder and harder. There is no closed-form expression in r. Luckily, thanks to the Zeilberger algorithm [3, 5], one can get the following linear recurrence equation for $A(n,r)$ in r: (Let us abbreviate $C \coloneqq (n+1)\binom{2n}{n+1}/4^n$):

$$A(n,r) = 2nA(n,r-1) + (r-1)(r-2)A(n,r-2) - 4n\binom{2n-1}{r-2}(r-2)!\,C,$$

subject to the initial conditions,

$$A(n,1) = 2n - 2C, \quad A(n,2) = 4n^2 - 8nC.$$

This enables a very fast computation of $A(n,r)$ for many r. Once we have them, Maple can easily compute the (usual) moments

$$\mathbb{E}\,[X^r] = \sum_{i=0}^{r} S(r,i)A(n,i),$$

where $S(r,i)$ are the Stirling numbers of the second kind.

Now, Maple can easily compute the *central moments* where $\mu \coloneqq \mathbb{E}\,[X]$ (that happens to be $2n-2C$)

$$\mathbb{E}\,[(X-\mu)^r] = \sum_{i=0}^{r}\binom{r}{i}(-\mu)^{r-i}\,\mathbb{E}\,[X^i].$$

In particular the variance $\sigma^2 \coloneqq \mathbb{E}\,[(X-\mu)^2]$. Finally it can take the limits of the scaled central moments

$$\lim_{n\to\infty}\frac{\mathbb{E}\,[(X-\mu)^r]}{\sigma^r},$$

and **surprise!** they are exactly the same as the central scaled moments of $-|N(0,1)|$, that are easily computed by Maple. We verified it up to 200 moments, but could have easily gone further. See the output file: `https://sites.math.rutgers.edu/~zeilberg/tokhniot/oCoinToss4.txt`.

But in order to prove it for *all* moments we need some human ingenuity and paper-and-pencil good-old-traditional math.

4. The Human Touch

Let $(\xi_i)_0^\infty$ be an infinite sequence of independent fair coin tosses, with $\xi_i = 1$ representing "Heads" and $\xi_i = -1$ representing "Tails". Let $S_n \coloneqq \sum_{i=1}^{n}\xi_i$.

Let H_n and T_n be the number of Heads and Tails, respectively, in the first n tosses. Thus

$$H_n = \frac{n+S_n}{2}, \qquad T_n = \frac{n-S_n}{2}. \tag{1}$$

Let

$$X(n) \coloneqq \min\{N : H_N = n \text{ or } T_N = n\} = \min\{N : \max(H_N, T_N) = n\}.$$

Note that

$$H_n + T_n = n, \tag{2}$$
$$H_n - T_n = S_n. \tag{3}$$

In particular,

$$X(n) = H_{X(n)} + T_{X(n)} \leqslant 2n. \tag{4}$$

Furthermore, at time $X(n)$, one of $H_{X(n)}$ and $T_{X(n)}$ equals n while the other is smaller. By (3), the smaller one is $n - |S_{X(n)}|$, and thus (4) yields

$$X(n) = 2n - |S_{X(n)}|. \tag{5}$$

Hence, the random variable $2n - X(n)$ that we are interested in is $|S_{X(n)}|$. In particular, the centered variable

$$X(n) - \mathbb{E}\,[X(n)] = -\big(S_{X(n)} - \mathbb{E}\,[S_{X(n)}]\big). \tag{6}$$

The idea to analyse $S_{X(n)}$ is that $X(n) \approx 2n$, and thus $S_{X(n)} \approx S_{2n}$, which has a nice normal limit by the central limit theorem. One elegant way to do this rigorously is by using Donsker's theorem on convergence of the entire process $(S_i)_{i=1}^{2n}$, after suitable scaling, to a Brownian motion. But I guess that you might be more comfortable with the following, I hope more pedestrian, approach.

We split the process of coin tosses into two phases, in the first we toss $n_1 := 2n - 2n^{2/3}$ times, and in the second we proceed to the end.

By the central limit theorem, since $\mathbb{E}\,\xi_i = 0$ and $\operatorname{Var}\xi_i = 1$,

$$\frac{S_{n_1}}{\sqrt{n}} = \sqrt{\frac{n_1}{n}}\frac{S_{n_1}}{\sqrt{n_1}} = \bigl(\sqrt{2} + o(1)\bigr)\frac{S_{n_1}}{\sqrt{n_1}} \xrightarrow{\mathrm{d}} N(0,2). \tag{7}$$

In particular, w.h.p. (with high probability, meaning with probability tending to 1 as $n \to \infty$), $|S_{n_1}| < 2n^{2/3}$, and thus $H_{n_1}, T_{n_1} < n$ by (1), so we have not reached the stopping time $X(n)$ before n_1. We may thus condition on this event, so we will assume it in the sequel.

Let $n_2 = 2n^{2/3}$, so $n_1 + n_2 = 2n$. Then, by the assumption just made and (4),

$$n_1 \leqslant X(n) \leqslant 2n = n_1 + n_2. \tag{8}$$

Let

$$S'_k := S_{n_1+k} - S_{n_1} = \sum_{i=1}^{k} \xi_{n_1+i}.$$

By Kolmogorov's inequality [2, Theorem 3.1.6], for every $x > 0$,

$$\mathbb{P}\bigl(\max_{1\leqslant k\leqslant n_2} |S'_k| > x\bigr) \leqslant \frac{\sum_{i=1}^{n_2} \operatorname{Var}(\xi_{i+n_1})}{x^2} = \frac{n_2}{x^2} = \frac{2n^{2/3}}{x^2}. \tag{9}$$

In particular,

$$\mathbb{P}\bigl(\max_{1\leqslant k\leqslant n_2} |S'_k| > n^{0.4}\bigr) \to 0,$$

and thus w.h.p., recalling (8),

$$|S_{X(n)} - S_{n_1}| = |S'_{X(n)-n_1}| \leqslant \max_{1\leqslant k\leqslant n_2} |S'_k| \leqslant n^{0.4}.$$

Hence,

$$\frac{S_{X(n)} - S_{n_1}}{\sqrt{n}} \xrightarrow{\mathrm{p}} 0,$$

which together with (7) yields

$$\frac{S_{X(n)}}{\sqrt{n}} \xrightarrow{\mathrm{d}} N(0,2). \tag{10}$$

If we let Z denote a standard normal variable, then (10) can also be written as

$$\frac{S_{X(n)}}{\sqrt{n}} \xrightarrow{\mathrm{d}} \sqrt{2}Z.$$

We have proved the following, recalling (5) and (6),

Theorem 4.1. We have, with notation defined above,

$$\frac{2n - X(n)}{\sqrt{n}} \xrightarrow{\mathrm{d}} \sqrt{2}|Z|, \tag{11}$$

and, for the centered variables,

$$\frac{X(n) - \mathbb{E}\,X(n)}{\sqrt{n}} \xrightarrow{\mathrm{d}} -\sqrt{2}\bigl(|Z| - \mathbb{E}\,|Z|\bigr). \tag{12}$$

5. n Heads and n Tails

We can argue similarly with

$$Y(n) := \min\{N \ : H_N \geqslant n \text{ and } T_N \geqslant n\} = \min\{N : \min(H_N, T_N) = n\}.$$

Now, $Y(n) \geqslant 2n$. Furthermore, one of $H_{Y(n)}$ and $T_{Y(n)}$ is n, and the other is $n + |S_{Y(n)}|$. Consequently, by (3),

$$Y(n) = 2n + |S_{Y(n)}|. \tag{13}$$

Let $n_2 = 2n^{2/3}$ as above. It follows from the central limit theorem, similarly to (7), that

$$\frac{S_{2n+n_2}}{\sqrt{n}} \xrightarrow{\text{d}} N(0,2).$$

and it follows easily that w.h.p. $|S_{2n+n_2}| < n_2$ and thus, by (1), $H_{2n+n_2}, T_{2n+n_2} > n$. Hence, w.h.p.,

$$2n \leqslant Y(n) < 2n + n_2.$$

We use Kolmogorov's inequality (9) again, but now for $1 \leqslant k \leqslant 2n_2$, and obtain as above

$$\frac{S_{Y(n)} - S_{n_1}}{\sqrt{n}} \xrightarrow{\text{p}} 0,$$

and thus

$$\frac{S_{Y(n)}}{\sqrt{n}} \xrightarrow{\text{d}} N(0,2),$$

and equivalently

$$\frac{S_{Y(n)}}{\sqrt{n}} \xrightarrow{\text{d}} \sqrt{2}Z.$$

Hence we obtain by (13), in analogy to Theorem 4.1:

Theorem 5.1. We have, with notation defined above,

$$\frac{Y(n) - 2n}{\sqrt{n}} \xrightarrow{\text{d}} \sqrt{2}|Z|, \tag{14}$$

and, for the centered variables,

$$\frac{Y(n) - \mathbb{E}\,Y(n)}{\sqrt{n}} \xrightarrow{\text{d}} \sqrt{2}\big(|Z| - \mathbb{E}\,|Z|\big). \tag{15}$$

Remark. Thus $X(n)$ and $Y(n)$ have, apart from a sign, the same (centered) asymptotic distribution. Moreover, it is really "the same" Z in Theorem 4.1 and 5.1: it follows from the profs above that

$$\frac{X(n) + Y(n) - 4n}{\sqrt{n}} = \frac{Y(n) - 2n}{\sqrt{n}} - \frac{2n - X(n)}{\sqrt{n}} = \frac{|S_{Y(n)}| - |S_{X(n)}|}{\sqrt{n}} \xrightarrow{\text{p}} 0.$$

In fact, an extension of the arguments above shows that $X(n) + Y(n) - 2n$ is of the order $n^{1/4}$. More precisely,

$$\frac{X(n) + Y(n) - 4n}{n^{1/4}} \xrightarrow{\text{d}} W,$$

for a random variable W with a distribution that is a mixture of normal distributions. (I omit the proof.)

6. Moment convergence

The results in Theorem 4.1 and 5.1 are convergence in distribution, and as always, this does not by itself imply convergence in distribution. In this case, it is easy to give supplementary arguments showing uniform integrability of powers of the random variables in (11)–(12) and (14)–(15), and consequently these hold with convergence of all moments. Here is one such argument:

Theorem 6.1. We have convergence of all moments (both ordinary and absolute) in (11), (12), (14), and (15).

Proof. Let

$$\nu_H(n) := \inf\{k : H_k \geqslant n\}, \qquad \nu_T(n) := \inf\{k : T_k \geqslant n\}.$$

Then

$$X(n) = \nu_H(n) \wedge \nu_T(n), \tag{16}$$

$$Y(n) = \nu_H(n) \vee \nu_T(n). \tag{17}$$

Note that $\nu_H(n)$ and $\nu_T(n)$ have the same distribution. They are dependent, so the representation (16)–(17) is of limited use, but it is nevertheless very helpful here.

Let $r \geqslant 2$; it is then well-known from renewal theory, see e.g. [1, Theorem 3.7.4(ii)], that the sequence of random variables

$$\left|\frac{\nu_H(n) - 2n}{\sqrt{n}}\right|^r, \qquad n \geqslant 1,$$

is uniformly integrable. (See e.g. [2, Section 5.4] for the definition.) The same is true for $\nu_T(n)$, since it has the same distribution as $\nu_H(n)$, and it follows from (16)–(17) that

$$\left|\frac{X(n) - 2n}{\sqrt{n}}\right|^r \quad \text{and} \quad \left|\frac{Y(n) - 2n}{\sqrt{n}}\right|^r, \qquad n \geqslant 1, \tag{18}$$

also are uniformly integrable.

This implies that all moments converge in (11) and (14), see e.g. [2, Theorem 5.5.9]. In particular, this shows that $(\mathbb{E}[X(n)] - 2n)/\sqrt{n}$ converges to $\mathbb{E}[\sqrt{2}|Z|]$, so $\mathbb{E}[X(n)] = 2n + O(\sqrt{n})$. Similarly, $\mathbb{E}[Y(n)] = 2n + O(\sqrt{n})$. This and (18) implies that

$$\left|\frac{X(n) - \mathbb{E}[X(n)]}{\sqrt{n}}\right|^r \quad \text{and} \quad \left|\frac{Y(n) - \mathbb{E}[Y(n)]}{\sqrt{n}}\right|^r, \qquad n \geqslant 1,$$

also are uniformly integrable. Hence we have moment convergence also in (11) and (14). □

Acknowledgements

S. Janson was supported by the Knut and Alice Wallenberg Foundation and the Swedish Research Council. L. Martinez was supported by the NSF Graduate Research Fellowship Program under Grant No. 2233066.

References

[1] Allan Gut, *Stopped Random Walks. Limit theorems and applications*, 2nd ed., Springer, New York, 2009.

[2] Allan Gut, *Probability: A Graduate Course*, 2nd ed., Springer, New York, 2013.

[3] Marko Pektovsek, Herbert S. Wilf, and Doron Zeilberger, *"A=B"*, AK Peters, 1996. Available at `https://sites.math.rutgers.edu/~zeilberg/AeqB.pdf`.

[4] Wikipedia, *Negative binomial distribution*. Available at `https://en.wikipedia.org/wiki/Negative_binomial_distribution`.

[5] Doron Zeilberger, *The method of creative telescoping*, J. Symbolic Comput. **11** (1991), 195-204. Available at `https://sites.math.rutgers.edu/~zeilberg/mamarimY/ct1991.pdf`.

(S. Janson) Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
Email address: svante.janson@math.uu.se

(L. Martinez & D. Zeilberger) Department of Mathematics, Rutgers University, Piscataway, NJ 08854
Email address: lucy.martinez@rutgers.edu

Email address: doronzeil@gmail.com