SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

The ballot theorems

av

Mathieu Thuillier

2017 - No 4

The ballot theorems

Mathieu Thuillier

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Yishao Zhou

2017

The ballot theorems

January 16, 2017

Abstract

The ballot problem was first enunciated by the French mathematician Bertrand in 1887. It solves the following question: in case of an election between two opposing candidates, what is the probability that the winner has been ahead throughout the whole count? Bertrand solved immediately the problem and concluded that the probability was equal to $\frac{\alpha+\beta}{\alpha-\beta}$.

This single result led to a flurry of research, aiming at generalizing that result. In this paper, we will focus on the original ballot theorem and its first generalization, that is the probability for a candidate to have throughout the count k times as many votes as the losing candidate. Beyond these results and the proofs attached to them, we will explore the link between ballot problems and the Catalan numbers, as well as some of the direct consequences of the results presented, especially in the theory of random walks. We will also investigate one direct application of the theorems in the field of electronics.

In a final section, we will focus on the perspectives for researchers, as they attempt to generalize the ballot theorems, eliminating the restrictions, among which the fact that the original ballot problem deals only with integers.

Acknowledgements

I wish to express my deep gratitude to Yishao Zhou for her clever advice, her precious guidance and her unending patience.

Contents

1 Introduction

It all started with a game. In 1887, a French mathematician publishes the result of a question, the so-called ballot problem, that was a mere exercise in basic combinatorics.

The ballot problem has indeed rapidly given birth to a series of different theorems, by changing some of the parameters of the original problem. In this essay, we will focus on the two most basic theorems: Bertrand's ballot theorem and the generalized ballot theorem, adding only a few connecting results. These two theorems are usually classified as "discrete time ballot theorems", in the sense that they rely on calculations that are valid only with integer parameters.

Our main goal with this essay is to show that a result that seemed harmless, a mere game for mathematicians has since had quite important repercussions in the field of combinatorics, especially within the theory of random walks.

The theorems also have applications in various fields, like biology and physics to name a few, and we will focus on one of them, in the field of electronics. It is indeed possible to compute by relatively easy means the probability of losing a given amount of packets of informations while going through a buffer before being injected into a network. The method presented in this thesis is, to the knowledge of its author, the fastest available.

We will also look into the generalizations that have been made around this theorem and try to show where the research is nowadays headed thanks to the ballot-style theorems.

Indeed, as we will sketch in the final section of this essay, there has been a flurry of publications during the last few decades around the so-called "continuous time ballot theorems" which generalize the results that will be presented in this essay. Nevertheless, as researchers focus on that subject, a truly general ballot theorem still eludes them.

But before we delve into the applications of the theorems, we will highlight the connection between this problem and a series of ubiquitous integers: the Catalan numbers. These numbers are used to count a myriad of phenomena and have a wide array of applications within mathematics. Proving this link between Catalan numbers and the ballot problem is another important goal for this essay.

Throughout this essay, we will try to keep a historical perspective on the subject of the ballot theorems and their applications, to keep track of the rapid developments in this specific field. We will see that, over the course of a century, this problem has evolved from a simple object of amusement for a few probabilists to a vivid topic of research.

2 At the origins of the ballot theorems

2.1 The classical ballot theorem

In 1887, The Comptes rendus de l'Académie des sciences published a short article entitled "Theory of probability. Solution of a problem by Mr. J. Betrand" between a vehement accusation of plagiarism thrown at a German physicist and the abstract of a study on the ways to reduce the number of vessels collisions at sea. These few lines will be the start of a new development in combinatorics, remembered nowadays as the first ballot theorem, which we reproduce here in its original statement.

Theorem 1 (Bertrand's ballot theorem) *Suppose that two candidates A and B are in an election. The number of the voters is μ. A obtains m votes and is elected, B obtains $\mu - m$. Find the probability that, during the counting of the votes, the number of votes for A always exceeds those of his competitor. The required probability is* $(2m - \mu)/\mu$.

The author did not produce in this article a detailed proof and contents himself with noting that "If $P_{m,\mu}$ indicates the number of combinations which, in the counting of the votes, are favorable to the required event, one has $P_{m+1,\mu+1} = P_{m,\mu} + P_{m+1,\mu}$." Indeed, the number of favorable combinations with $n+1$ votes for A and $\mu+1$ votes for B is equal to the number of favorable combinations with m votes for A and μ votes for B added to the number of favorable combinations with $n+1$ votes for A and μ votes for B.

A few weeks later another French mathematician, Désiré André, brought a formal proof to the result presented by Joseph Betrand.

André starts by noting that the total of possible events is equal to the number of words one can form with the letters A and B representing each a vote for candidate A or a vote for candidate B. Let us assume that the number of votes for A is equal to α and the number of votes for B is equal to β. André focuses on the number of unfavourable events, that is the number of permutations of the letters A and B that do not give A the lead throughout the whole count. These permutations begin either by A or by B. There are as many words describing a unfavourable event starting with B as there are permutations with α A and $\beta - 1$ B.

André proves then that the number of unfavourable permutations starting with A is also equal to that same number, by creating a bijection between unfavourable events represented by a word starting with A and unfavourable events starting represented by a word starting with a B.

The correspondence is based on the following rule:

1. One can associate any unfavourable permutation starting with an A with α A and β B to a permutation with α A and $\beta - 1$ B as follows. One starts by removing the first B that violates the law of the problem - that is the first B for which the number of A equals the number of B. One can then exchange the place of the two groups of letters thus obtained to obtain the desired permutation of α A and $\beta - 1$ B.

For example, the permutation AABBABAA is split in two groups AAB and ABAA by removing the second B in the word - since at that point there are 2 A and 2 B. The groups then switch places to give the word ABAAAAB.

2. One can associate any permutation containing α A and $\beta - 1$ B to an unfavourable permutation starting with an A with α A and β B. One has to consider the word from right to left and when, for the first time, the number of A surpasses the number of B by one unit, one splits the word in two at that point, switches places for the two groups thus obtained and adds a letter B between them. One obtains then the desired permutation.

For example, ABAAAAB is split in two after the second A starting on the right - since at that point the number of A exceeds the number of B by one unit. The two groups ABAA and AAB switch places and one adds a B between the two to get the unfavourable permutation AABBABAA.

André concludes his proof by computing the desired probability, knowing that there are twice as many unfavourable events with α A and β B as there are permutations with α A and $\beta - 1$ B. This gives the number of favorable events Q:

$$
\begin{aligned}
Q &= \frac{(\alpha+\beta)!}{\alpha!\beta!} - 2\frac{(\alpha+\beta-1)!}{\alpha!(\beta-1)!} \\
&= \frac{(\alpha+\beta)!}{\alpha!\beta!} - 2\frac{\frac{(\alpha+\beta)!}{\alpha+\beta}}{\frac{\alpha!\beta!}{\beta}} \\
&= \frac{(\alpha+\beta)!}{\alpha!\beta!} - 2\frac{(\alpha+\beta)!}{\alpha!\beta!} \cdot \frac{\beta}{\alpha+\beta} \\
&= \frac{(\alpha+\beta)!}{\alpha!\beta!}(1 - \frac{2\beta}{\alpha+\beta}) \\
&= \binom{\alpha+\beta}{\alpha}(\frac{\alpha+\beta-2\beta}{\alpha+\beta}) \\
&= \binom{\alpha+\beta}{\alpha}(\frac{\alpha-\beta}{\alpha+\beta})
\end{aligned}
$$

The first term if the last expression being the number of words made of $\alpha+\beta$ letters and containing exactly α A and β B, he deduces the desired probability P:

$$P = \frac{\alpha-\beta}{\alpha+\beta}$$

2.2 The generalized ballot theorem

The same year, another French mathematician, Joseph-Émile Barbier proposes a generalized version of the problem. Instead of stating the problem with A leading throughout the count, that is $\alpha > \beta$, Barbier considers the problem under the formulation $\alpha > k\beta$, for some integer k.

Theorem 2 (Generalized ballot theorem) *Suppose that two candidates A and B are in an election. The number of the voters is μ. A obtains m votes and is elected, B obtains $\mu - m$ votes. The probability that, during the counting of the votes, the number of votes for A always exceeds k times the number of votes for his/her opponent, is equal to*

$$P = \frac{\alpha - k\beta}{\alpha+\beta}.$$

This implies that the number of ways in which we can count the ballots so that the number of ballots for A exceeds, at any moment during the count, k times the number of ballots counted for the candidate B is equal to:

$$\frac{\alpha - k\beta}{\alpha+\beta}\binom{\alpha+\beta}{\alpha}$$

This results from the fact that there are $\binom{\alpha+\beta}{\alpha}$ ways to count $\alpha+\beta$ ballots.

Barbier did not provide any proof to this generalized ballot theorem. But we can easily proceed by induction to prove that this result is exact, as Takacs did in an article published in 1997.

Let us call $N_k(\alpha,\beta)$ the number of ways the ballots can be ordered so that the number of ballots for A exceeds at any moment during the count k times the number of ballots counted for the candidate B for a given constant k.

We find that $N_k(\alpha, 0) = 1$ for all $\alpha > 0$, since there is only one way to arrange the count if A got all the votes. Moreover, we have $N_k(k\beta, \beta) = 0$ for all $\beta > 0$. Both of these statements satisfy the condition

$$N_k(\alpha,\beta) = \frac{\alpha - k\beta}{\alpha+\beta}\binom{\alpha+\beta}{\alpha}.$$

In the case $\beta > 0$ and $\alpha > k\beta$, we know that $N_k(\alpha, \beta) = N_k(\alpha, \beta - 1) + Nk(\alpha - 1, \beta)$. By induction, we deduce that the number of ways we can count the ballots is:

$$\frac{\alpha - k(\beta - 1)}{\alpha + \beta - 1} \binom{\alpha + \beta - 1}{\alpha} + \frac{\alpha - 1 - k\beta}{\alpha + \beta - 1} \binom{\alpha + \beta - 1}{\alpha - 1}.$$

with the help of a recursive definition of the binomial coefficient, we find that this expression simplifies to

$$\frac{\alpha - k\beta}{\alpha + \beta} \binom{\alpha + \beta}{\alpha}.$$

2.3 The weak ballot theorem

A little transformation of the initial problem gives us an important result, often called "the weak ballot theorem". Instead of counting the ways the votes can be counted so that candidate A is always *more than* k times ahead of B, one can determine the number of ways the votes can be counted so that A is ahead by *at least* k times the number of votes for B.

In fact, adding a single vote for A at the beginning of the count gives us a count where A is always more than k times ahead. The reverse is also true. Based on this bijection, we realize that the number of ways A can lead throughout the count with at least k times more votes than for candidate B is equal to the number of ways A is ahead by more than k times as many votes as for B in an election with a final tally of $\alpha + 1$ votes for A and β votes for B. The latter can be counted with the help of the classical ballot theorem:

$$\frac{\alpha + 1 - k\beta}{\alpha + 1\beta} \binom{(\alpha + 1) + \beta}{\alpha + 1} = \frac{\alpha + 1 - k\beta}{\alpha + 1} \binom{\alpha + \beta}{\alpha}.$$

Since there are $\binom{\alpha+\beta}{\alpha}$ ways to count the ballots with α votes for A and β votes for B, the probability P of A being ahead by at least k times the number of votes for B throughout the count is:

$$P = \frac{\alpha + 1 - k\beta}{\alpha + 1}.$$

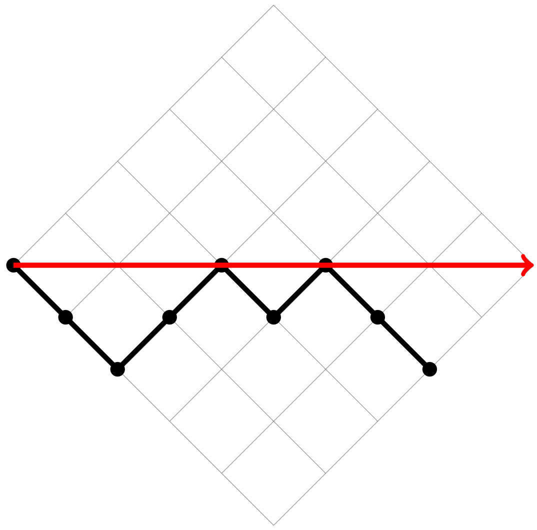

Figure 1: Representing the count AABBABAA in terms of a path on the integral lattice. The path touches the red line when $\alpha_k = \beta_k$ after $2k$ votes have been counted.

3 The ballot theorems and the reflection method

3.1 The count as a path on the integral lattice

It is possible to restate the ballot problem in terms of paths on the integral lattice in the coordinate plane. One can indeed represent the count as a path, i.e. a sequence of lattice points (points with integer coordinates) from the origin (0,0) to $(\alpha+\beta, \beta-\alpha)$, where $\alpha+\beta$ is the number of ballots cast in the election.

The sequence of lattice points is determined in that fashion: Every ballot cast for A implies a step northeastwards (1, 1) and every ballot cast for B implies a step southeastwards (1, -1) in the coordinate plane. This sequence of points is thereafter called a path.

In the case of the original ballot problem, that is, A has to be ahead of B throughout the whole count, a count satisfying the conditions enunciated in the problem will be represented as a path that stays all the way under the x-axis. Indeed, when a path crosses that line - or touches it -, the number of ballots counted for the candidate A at that point is inferior - or equal - to the number of ballots counted for B.

A path satisfying this condition, i.e. a path staying below the x-axis, will from now on be called a good path, while a path that crosses at some point (except at the origin) that line will be called a bad path. The original ballot problem can now be restated as a way of counting the good paths.

3.2 The reflection method

We saw in the first section that André was the first to provide a proof to the solution of the ballot problem. His proof relies on a count of permutations. Rather strangely, his name remains associated to another proof of the ballot theorem, usually referred to as "André's reflection method". The reflection method relies on the representation of counts as paths on the integral lattice and geometric arguments to establish a bijection that allows one to count the bad paths. There are no indications in André's works that he had used such a method to solve the problem. As Addario-Berry and Reed state (1) "The proof is often incorrectly attributed to André (1887), who established the same bijection in a different way - its true origins remain unknown."

Marc Renault (12) traces the origin of the method to 1923, when Aebly published an article illustrating different permutations as paths starting from a corner on a rectangular grid and designs a symmetry of bad paths along the diagonal.

To prove the result via this method, we start by noticing that every good path starts with an step to the point with coordinates (1,-1), that is with a step southeastwards on the integral lattice. We then construct a one-to-one correspondence between the set of all paths starting from (1,-1) that touch somewhere the x-axis and the set of paths from (1,-1) to $(\alpha+\beta, \beta-\alpha)$, the terminal point. The existence of this one-to-one correspondence can be easily verified by reflecting across the x-axis any bad path starting from (1,-1) up to the point where it touches the x-axis.

It is therefore possible to determine the number of good paths by subtracting the number of these bad paths starting from (1,-1) to the number of all paths from (1,1) to our terminal point. There are $\binom{\alpha+\beta-1}{\alpha-1}$ such paths and $\binom{\alpha+\beta-1}{\beta-1}$

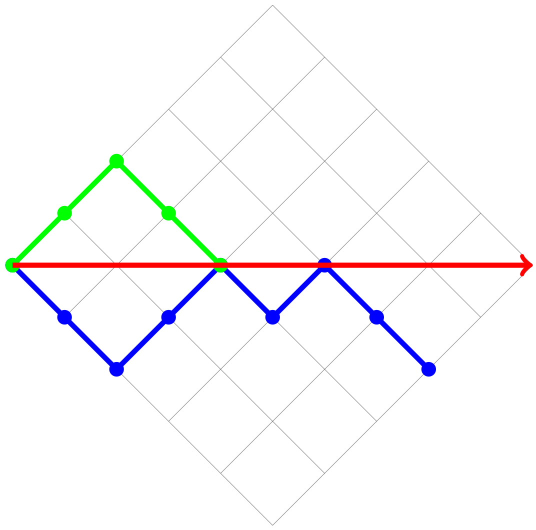

Figure 2: Reflecting a bad path. The green line corresponds to the part of the bad count where B had more votes than A. Reflecting that green part into the x-axis, one obtains the blue path.

bad paths starting from (1,-1). The number of good paths is therefore:

$$\begin{aligned}
\binom{\alpha+\beta-1}{\alpha-1} - \binom{\alpha+\beta-1}{\beta-1} &= \frac{(\alpha+\beta-1)!}{(\alpha-1)!+\beta!} - \frac{(\alpha+\beta-1)!}{(\beta-1)!+\alpha!} \\
&= \frac{(\alpha+\beta-1)!}{\frac{\alpha!\beta!}{\alpha}} - \frac{(\alpha+\beta-1)!}{\frac{\alpha!\beta!}{\beta}} \\
&= \frac{\frac{(\alpha+\beta)!}{\alpha+\beta}}{\frac{\alpha!\beta!}{\alpha}} - \frac{(\alpha+\beta-1)!}{\frac{\alpha!\beta!}{\beta}} \\
&= \frac{(\alpha+\beta)!}{\alpha!+\beta!}\frac{\alpha}{\alpha+\beta} - \frac{(\alpha+\beta)!}{\alpha!+\beta!}\frac{\beta}{\alpha+\beta} \\
&= \frac{\alpha-\beta}{\alpha+\beta}\frac{(\alpha+\beta)!}{\alpha!+\beta!} \\
&= \frac{\alpha-\beta}{\alpha+\beta}\binom{\alpha+\beta}{\alpha}
\end{aligned} .$$

3.3 Using lattice paths to solve the generalized ballot problem

Extending the reflection method to solve the generalized ballot problem is not without difficulties. In 2003, Goulden and Serrano noted indeed that, while there are many ways to solve the generalized ballot problem, "there appears to be no solution which is in the spirit of the reflection principle".

Nevertheless, it is possible to use the lattice path representation to solve the problem. In the generalized ballot problem, one looks for the number of ways to count the ballots so that candidate A has at any moment more than k times the number of votes for B, for some integer k. One may indeed represent a count as a lattice path with a series of upsteps (1,1) for each vote for A and downsteps (1,-k). A path is considered a good path if no step end on or below the x-axis.

For each integer i, $0 \leq i \leq k$, we denote B_i the set of bad paths whose first bad step ends i steps below the x-axis. The set B_k is the set of paths that start with a downstep. It is obvious that these $k+1$ sets are disjoint and their union is the set of all bad paths.

We can now create a one-to-one correspondence between the set A of paths consisting of α upsteps and $\beta - 1$ downsteps and each of the sets B_i. One can indeed consider any path $P = CDE$ belonging to B_i as the composition of three partial paths C, D and E, where C is the part of the path before the first step the crosses or lands on the x-axis, D is the step that crosses of lands on the x-axis and E is the terminal part of the path that follows that step.

One can then rearrange this path, suppressing the part D of the path, to get a path EC, that is uniquely determined and that possesses α upsteps and $\beta - 1$ downsteps. Hence this new path $P' = EC$ is an element of A.

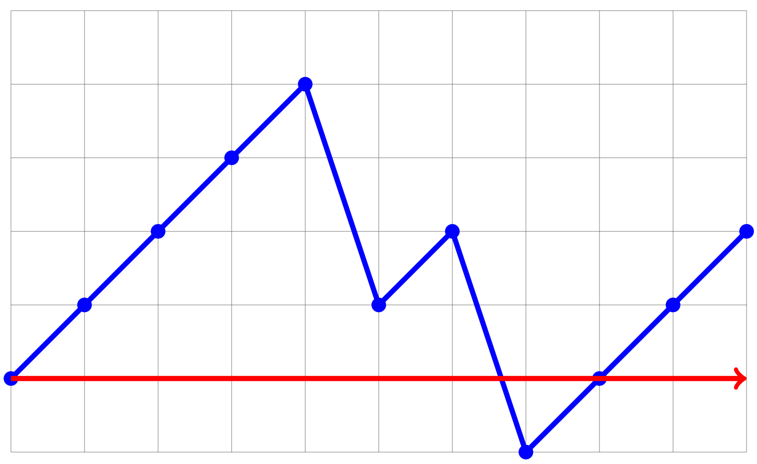

Figure 3: A path $P = CDE$ belonging to B_1. The partial path C includes the six first steps, the partial part D includes the step that goes below the x-axis and the partial path E includes the three last steps.

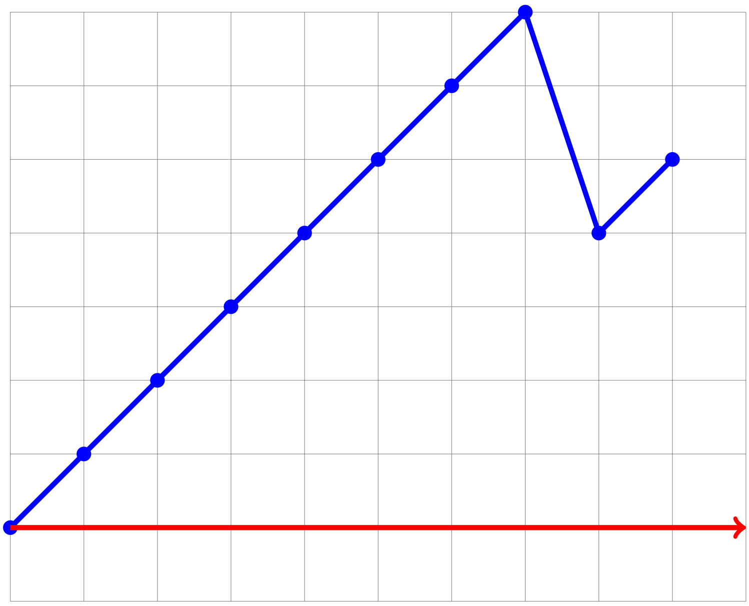

Figure 4: The representation of the path $P' = EC$.

On the other hand, a path Q belonging to A can be transformed into a path Q' belonging to B_i, by scanning the path from the right to the left until one finds a vertex lying $k-i$ units below the terminal vertex. Clearly, such a vertex must exist, since the path starts with a vertex that is more than k units below the terminal vertex. One decomposes then the path in two parts E and C joined at that vertex. One interchanges C:s and E:s places and inserts a downstep D between them to find a path Q, uniquely determined, with $Q = CDE$. One does translate then this path to start at the origin.

The path thus obtained touches the x-axis only at the origin and ends i units below the x-axis. We have thus found a path Q' that is uniquely determined and belongs to B_i.

The number of all possible paths with α upsteps and β downsteps is equal to $\binom{\alpha+\beta}{\alpha}$. The number of paths belonging to A with α upsteps and $\beta-1$ downsteps is equal to $\binom{\alpha+\beta-1}{\alpha}$ and there are $k+1$ sets B_k. Hence, we find the number of good paths to be:

$$\binom{\alpha+\beta}{\alpha} - (k+1)\binom{\alpha+\beta-1}{\alpha} = \frac{\alpha-k\beta}{\alpha+\beta}\binom{\alpha+\beta}{\alpha}.$$

4 The Catalan numbers

In this section, we will explore the link between the ballot problem and the Catalan numbers.

4.1 Definition

Richard P. Stanley, who has dedicated a major part of his research to the Catalan numbers, presents them as "probably the most ubiquitous sequence of numbers in mathematics". In a monograph on this subject, he describes 214 different kinds of objects that can be counted with the help of Catalan numbers.

In 1751, Euler was the first to introduce a closed formula for what we call nowadays the Catalan numbers based on the number of possible triangulations of convex $(n+2)$-gons. We will develop this definition, as it is the original one. A triangulation is a set of diagonals which do not cross their interiors and partition the $(n+2)$-gon into a series of different triangles.

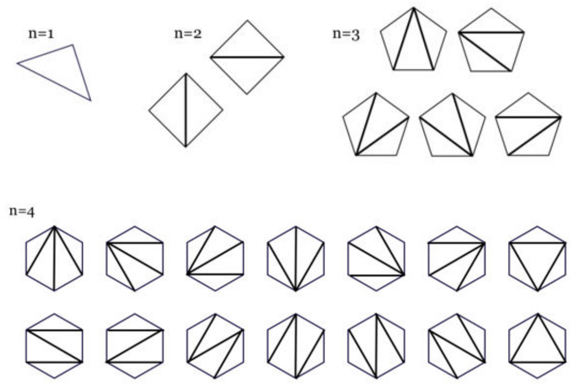

Figure 5: Triangulated polygons.

4.2 The fundamental recurrence

To count the ways in which a $(n+2)$-gon can be divided into triangles so that diagonals do not cross their interiors, one can define a recurrence relation.

We consider a convex $(n+3)$-gon and fix an edge. If we observe a particular triangulation of this figure and remove the designated edge, we obtain two triangulated polygons that share a common vertex. To compute the number of possible triangulations of the figure obtained when ones takes away a given edge, one can multiply the number of triangulations for each polygon obtained. The first polygon (the bigger one of the two) contains $a_i + 2$ edges and the second (the smaller one) $a_j + 2$, with $a_i + a_j = n$. If the first polygon can be triangulated in C_{a_i} ways, the other can be triangulated in $C_{a_i} = C_{a_{n-i}}$ ways. We get thus $C_{a_i} C_{a_{n-i}}$ ways.

If we choose another triangulation, we may obtain two other polygons of different sizes. The total number of triangulations for the $(n + 3)$-gon will therefore be equal to

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$$

.

In some cases, one of the two polygons can be a single edge, when the removed edge is part, in the original triangulation, of a triangle built by a diagonal and two adjacent edges. In these cases, we consider that the number of possible triangulations for the single edge (that is a "2-gon") is $C_0 = 1$. This initial condition completes the fundamental recurrence to calculate Catalan numbers.

4.3 A generating function

With the help of this recurrence, we can define a generating function for the Catalan numbers.

Theorem 3 *Let*

$$C(x) = \sum_{n \geq 0} C_n x^n = 1+x+2x^2+x+5x^3+14x^4+42x^5+132x^6+429x^7+1430x^8...$$

Then

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

Proof. We use the recurrence relation obtained earlier and multiply it by x^n and sum on $n \geq 0$. On the left-hand side we have

$$\sum_{n \geq 0} C_{n+1} x^n = \frac{1}{x} \sum_{n \geq 0} C_{n+1} x^{n+1} = \frac{1}{x} \sum_{n \geq 1} C_n x^n = \frac{1}{x} \sum_{n \geq 0} C_n x^n - C_0 = \frac{C(x) - 1}{x}.$$

On the right-hand side, we get

$$\sum_{n \geq 0} (\sum_{k=0}^{n} C_k C_{n-k}) x^n = C(x)^2.$$

We get thus

$$xC(x)^2 - C(x) + 1 = 0$$

whose roots are:

$$C(x) = \frac{1 \pm \sqrt{1-4x}}{2x}. \tag{1}$$

We must now determine if the positive or the negative sign applies. We know that $C_0 = 1$. We could therefore plug in $x = 0$ in the equation, but the numerator becomes 0 then. To go around this obstacle, we evaluate the limit of the expression for each sign when $x \to 0^+$. Applying L'Hôpital's rule, we find that only the negative sign gives us the required limit, that is, 1. We can conclude that

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x},$$

which is the desired generating function.

4.4 An explicit formula

With the help of this generating function, it is possible to derive an explicit formula for the calculation of the Catalan numbers.

Theorem 4

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{n!(n+1)!}.$$

Proof. We have

$$\sqrt{1-4x} = (1-4x)^{1/2} = \sum_{n \geq 0} \binom{\frac{1}{2}}{n} (-1)^n (4x)^n$$

With the help of the binomial expansion for fractional exponents, we find:

$$\sqrt{1-4x} = 1 - \frac{1/2}{1!}4x + \frac{(1/2)(-1/2)}{2!}(4x)^2 - \frac{(1/2)(-1/2)(-3/2)}{3!}(4x)^3$$

$$+\frac{(1/2)(-1/2)(-3/2)(-5/2)}{4!}(4x)^4 - \frac{(1/2)(-1/2)(-3/2)(-5/2)(-7/2)}{5!}(4x)^5 + ...$$

This simplifies to

$$\sqrt{1-4x} = 1 - \frac{1}{1!}2x + \frac{1}{2!}4x^2 - \frac{3 \cdot 1}{3!}8x^3 + \frac{5 \cdot 3 \cdot 1}{4!}16x^4 - \frac{7 \cdot 5 \cdot 3 \cdot 1}{5!}32x^5 + ...$$

We notice that the numerators in the coefficients look like factorials but with the even numbers missing. Using the fact that $2^2 \cdot 2! = 4 \cdot 2, 2^3 \cdot 3! = 6 \cdot 4 \cdot 2$

and $2^4 \cdot 4! = 8 \cdot 6 \cdot 4 \cdot 2$, we rewrite the numerators to get a factorial. Inserting in (1), we get:

$$\begin{aligned} C(x) &= 1 + \frac{1}{2}(\frac{2!}{1!1!})x + + \frac{1}{3}(\frac{4!}{2!2!})x^2 + + \frac{1}{4}(\frac{6!}{3!3!})x^3 + + \frac{1}{5}(\frac{8!}{4!4!})x^4 + ... \\ &= \sum_{n=0}^{\infty} \frac{1}{n+1}\binom{2n}{n}x^n \end{aligned}$$

We conclude that the n^{th} Catalan number is

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

An alternative expression is:

$$C_n = \frac{1}{n}\binom{2n}{n-1}$$

4.5 A generalization of the Catalan numbers

Provided this definition, we can extend the concept of Catalan numbers. The generalized Catalan numbers, or k-Catalan numbers, are defined in that fashion:

$$C_n^k = \frac{1}{kn+1}\binom{k(n+1)}{n}$$

They can be defined as the number of ways to count the k-ary trees, that is, ordered trees in which each node has either out degree 0 or k.

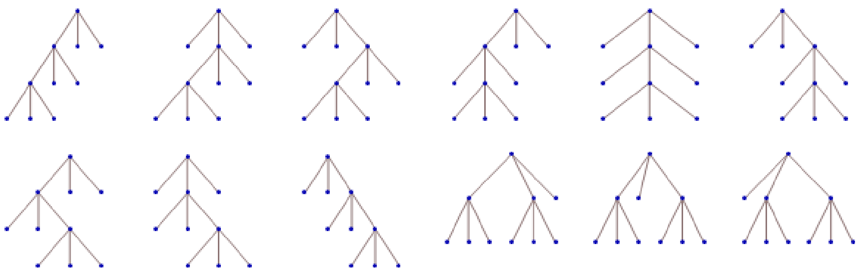

Figure 6: Ternary trees.

These numbers are also linked to the ballot problem, this time in its generalized version. Indeed, if we consider the case where the candidate A gets

$\alpha = kn + 1$ votes and the candidate B gets $\beta = n$ votes, the number of ways to count the ballot that satisfy the conditions of the problem are:

$$\frac{\alpha - k\beta}{\alpha + \beta}\binom{\alpha+\beta}{\alpha} = \frac{kn+1-kn}{kn+1+n}\binom{kn+1+n}{kn+1} = \frac{1}{kn+1}\binom{k(n+1)}{n} = C_n^k.$$

4.6 The link between the ballot problem and Catalan numbers

Earlier, we saw that the number of ways to count the ballots so that A is ahead throughout the count is:

$$\binom{\alpha+\beta-1}{\alpha-1} - \binom{\alpha+\beta-1}{\beta-1} = \frac{\alpha-\beta}{\alpha+\beta}\binom{\alpha+\beta}{\alpha}$$

Let us examine the closest possible situation with B getting k votes and A getting $k+1$ votes. In that case, the number of ways to count the votes satisfying the conditions enuciated in the problem are:

$$\begin{aligned}
\binom{k+1+k-1}{k+1-1} - \binom{k+1+k-1}{k-1} &= \binom{2k}{k} - \binom{2k}{k-1} \\
&= \frac{(2k)!}{k!k!} - \frac{(2k)!}{(k-1)!(k+1)!} \\
&= \frac{(2k)!}{\frac{k}{k+1}(k+1)!(k-1)!} - \frac{(2k)!}{(k-1)!(k+1)!} \\
&= (\frac{k+1}{k})(\frac{(2k)!}{(k-1)!(k+1)!}) - \frac{(2k)!}{(k-1)!(k+1)!} \\
&= (\frac{k+1}{k} - 1)\binom{2k}{k-1} \\
&= \frac{1}{k}\binom{2k}{k-1}
\end{aligned}$$

We notice that this result is equivalent to the alternative expression of the k^{th} Catalan number. We can therefore conclude that one can define the k^{th} Catalan number as the number of possible ways to count the ballots in a closely fought election, where the losing candidate got k votes and the winning candidate $k+1$ votes, so that the winning candidate is ahead throughout the count.

As we stated earlier, there are many ways to interpret the Catalan numbers. As Richard P. Stanley states, the Catalan number C_n count the following:

(i) Triangulations of a convex polygon with $n+2$ vertices.
(ii) Binary trees with n vertices.
(iii) Plane trees with $n+1$ vertices.
(iv) Ballot sequences of length $2n+1$.
(v)Parenthesizations (or bracketings) of a string of $n+1$ x's subject to a nonassociative binary operation.
(vi) Dyck paths of length $2n$.

The details and proof of this statement are outside the scope of this essay.

5 Applications

5.1 Direct applications

Though its scope might seem restricted at first sight, the ballot theorem can be used in many different situations - and many of them have nothing to do with elections. We will dedicate this section to solving a few problems with the help of the ballot problem, starting with some direct applications of the theorem.

Problem 1. Suppose that an urn contains r red balls and b black balls, with $r > b$. The balls are sampled without replacement from the urn. What is the probability that the number of red balls is strictly greater than the number of black balls removed after each draw?

This situation is a simple transposition of the original ballot problem. Indeed, if we represent the drawing of a red ball as step southeastwards on the lattice path and the drawing of a black ball as a step northeastwards, then the problem at hand becomes finding the probability of the path to remain throughout the whole game below the x-axis. The problem at hand is therefore equivalent to the ballot problem, with the drawing of a red ball the equivalent of a vote for A.

We can find

$$\frac{r-b}{r+b}\binom{r+b}{r}$$

such paths. The number of possible drawings being equal to $\binom{r+b}{r}$, the desired probability P is therefore

$$P = \frac{r-b}{r+b}.$$

Problem 2. A wholesale house does business only in 100-dollar units, one transaction per customer per month. All bills are paid on the last day of the month. On this day, n creditors and n debtors appear one at a time. The house has no cash to start with. What is the probability P that a creditor will have to wait?

Once again, it is quite obvious that the problem at hand is equivalent to the ballot problem. Indeed, if we consider the succession of creditors and debtors the last day of the month as the count, the coming of a debtor as a vote of A and the coming of a creditor as a vote for B, we realize that we are dealing with the ballot problem.

We can indeed represent the 100-dollar transactions as a succession of moves northeastwards (when a creditor presents himself/herself) and southeastwards (when a debtor presents himself/herself). A creditor has to wait if the wholesale house has no cash left. If, every time the wholehouse has no cash left, and a debtor shows up, then no creditor has to wait. Hence, the path can touch the x-axis, but cannot go above it. Therefore, the problem at hand is the ballot problem with ties allowed, with $2n$ "votes", $\alpha = n$ and $\beta = n$.

We can count therefore the probability $P*$ of the path to be good (with ties allowed), with the help of the weak ballot theorem:

$$P* = \frac{n+1-n}{n+1} = \frac{1}{n+1}.$$

The probability P that a creditor will have to wait is the complement of $P*$

$$P = 1 - \frac{1}{n+1} = \frac{n+1}{n+1} - \frac{1}{n+1} = \frac{n}{n+1}.$$

Problem 3. A pack of cards consist of 26 reds and 26 blacks. A person is given the chance to win a prize by correctly predicting when the next card to be dealt will be red. This person is allowed to make a single prediction and has to bet on a red card. An obvious stategy is to wait until the number of black cards already dealt is greater than the number of red cards. Then one has more than 50% chance to win. The problem is that there is a risk that the number of blacks cards never surpasses the number of red cards picked. How big is that risk?

We need to count the number of ways one can draw the 52 cards so that the number of black cards drawn is, at any moment, at least equal to the number of red cards picked. By the weak version of the ballot theorem, we find that there are $\frac{26+1-26}{26+1}\binom{26+26}{26} = \frac{1}{27}\binom{52}{26}$ ways to achieve this.

This gives us the following probability P of this happening:

$$P = \frac{1}{27}.$$

Problem 4. A club opens in the evening and closes in the morning. People arrive alone and leave in pairs. What is the possible number of scenarios from dusk to dawn as seen from the club's entry?

In this case, we get a variant of the ballot problem, where votes are weighted in two different manners. It can be translated into lattice paths starting at the origin that never cross the x-axis, but where the steps are either (1,-1) or (1,2), with $2n$ people entering and n pairs leaving the club. The answer is given by the generalized ballot theorem, with $2n$ votes for A and n votes for B.

We get the following number of scenarios:

$$\frac{2n+1-2n}{2n+1}\binom{2n+n}{2n} = \frac{1}{2n+1}\binom{3n}{2n}.$$

5.2 An application in the field of electronics

Beyond these basic applications, the ballot theorem and its variant have had important consequences in various fields, from physics, through the study of the distribution of ionized particles in a gas tube, to biology, in particular within angiogenesis. We will study more in depth an application in the field of electronics.

The transmission of data in a variety of networking applications relies on the existence of buffers that can stock packets of data before their transmission into the network. The buffer can only contain a given number of packets. If the buffer is full, all the subsequent packets arriving will be lost until a packet leaves the buffer (or queue) and enters the network. The loss of packets may result in the loss of the entire message.

The probability distribution of the number of packet losses is therefore an essential characteristic of a network. This can be studied through recursive models, but it implies a high number of calculations. The ballot theorem offers a simpler and faster method to study this probability distribution.

In an article published in 2000, Omer Gurewitz, Moshe Sidi and Israel Cidon presented the following method.

The model.
The model studies systems with packets of variable length whose transmission time into the network is exponentially distributed with parameter μ. The buffer accomodates up to m packets and, in the event the buffer is full, and a new packet arrives, the packet is lost. Packets are grouped in blocks of a given size of n consecutive packets. The packets arrive according to a Poisson process with rate λ and the average load ρ is defined as $\rho = \lambda/\mu$.

The goal is to compute the probabilities $P_i(j,n), i = 0, 1, ..., m, n \geq 1, 0 \leq j \leq n$ of j losses in a block of n packets, given that there are i packets in the buffer before the first packet arrives. This constitutes a first step towards the calculation of $P(j,n), n \geq 1, 0 \leq j \leq n$ of j losses in a block of n packets.

The latest can indeed be computed in the following manner:

$$P(j,n) = \sum_{i=0}^{m} \Pi(i) P_i(j,n)$$

where $\Pi(i)$ is the probability of having i packets in the buffer at the start of the process. This probability can be computed as follows:

$$\Pi(i) = \rho^i / \sum_{l=0}^{m} \rho^l.$$

Moreover, if we denote $P_i^k(j,n)$ the probability that the kth packet of the block is the jth of that block to be lost, we note that:

$$P_i(j,n) = \sum_{k=1}^{n} P_i^k(j,n).$$

We can represent the state of the buffer in terms of lattice paths, as a function of time, where one step northwards represent the arrival of a packet and a step southwards represent the departure of a packet. The following figure represents an example of such a lattice path.

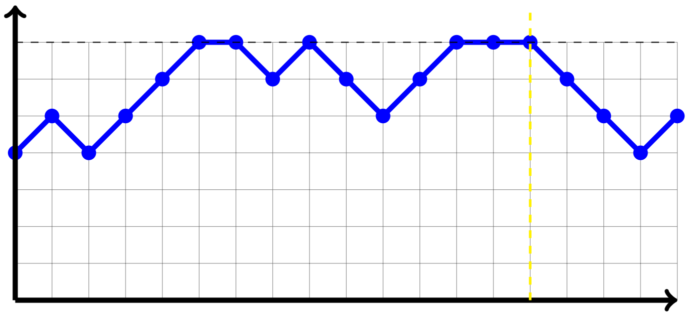

Figure 7: The state of a buffer containing $i = 4$ packets at the start, with a capacity of $m = 7$ packets. In this example, three packets are lost. When already full, the buffer cannot accept another packet. This loss of a packet is represented by the line ”staying” flat.

This path - just like any other path - can be divided into two differents events:

- Event A: from the start up to the last loss.
- Event B: from the last loss to the end.
(In the graphic above, the yellow dashed line marks the end of event A and the start of event B)

In the following, we will denote as p the probability of a departure when the buffer is not empty and q its complement. The probability that a path has u arrivals and v departures is $p^u q^v$ if the buffer does not empty and it becomes $p^{u-h} q^v$ if the buffer empties h times.

To compute the probabilities of events A and B, the authors use two rather obvious results:

1. The kth packet is the last to be lost if and only if upon the arrival of this packet, the number of arrivals exceeds the numbers of departures by $m - i + j$ for the first time, where m is the maximal capacity of the buffer, i is the number of packets in the buffer at the origin and j is the number of packets eventually lost.

2. No packet is lost after the kth packet if and only if, afterwards, the number of departures always exceeds or is equal to the number of arrivals.

Two cases have to be distinguished:
- The buffer can contain a complete block.
- The block is larger than the buffer.

Case 1: The buffer can contain a complete block.

This case is quite straightforward. We start by computing the probabilities $P_i(j, n)$ before we deduce the probability $P(j, n)$. To start, we consider the probability of event A, that is the probability that the kth packet in the block is the jth packet to be lost. The paths all end with an arrival and are composed of k arrivals and l departures.

The first result presented (The kth packet is the last to be lost if and only if upon the arrival of this packet, the number of arrivals exceeds the numbers of departures by $m - i + j$ for the first time) yields:

$$k - l = m - i + j$$

This implies that exactly j packets will be lost along such paths. These paths have $2k - (m - i + j)$ upwards and downwards steps. Since they all start with an arrival, the probability of each one is $p^{k-1} q^{k-(m-i+j)-1}$. Moreover, we can deduce that there are exactly $\binom{2k-(m-i+j)-1}{k-1}$ paths.

We can count the relevant paths among this number, with the help of the classical ballot theorem, considering the path backwards. Indeed, in this direction, the number of arrivals has to exceed at all time the number of departures (otherwise, the condition exposed in the result 1 is not fulfilled).

The proportion of the number of paths that fulfill the conditions is equal to $\frac{m-i+j-1}{2k-(m-i+j)-1}$. We deduce the desired probability for event A (that is that the number of arrivals exceed the number of departures by $m-i+j-1$ for the first time by the kth arrival) is equal to

$$\frac{m-i+j-1}{2k-(m-i+j)-1}\binom{2k-(m-i+j)-1}{k-1}p^{k-1}q^{k-(m-i+j)}. \tag{2}$$

To determine the probability of event B, one considers the fact that the number of departures exceeds the number of arrivals until the last packet arrives. Moreover, we note that, once the number of departures equals $n-k$, the path becomes irrelevant, since the risk of it reaching m is 0. The number of paths, with $n-k$ departures and l arrivals, given that the last step is a departure is equal to $\binom{n-k-1+l}{l}$. The probability of each of them is equal to $qp^lq^{n-k-1} = p^lq^{n-k}$. The ballot theorem gives us the probability to obtain a "good" path $\frac{n-k-l}{n-k}$.

This gives us the probability of event B, that is that no more packets will be lost after the kth packet is lost

$$\sum_{l=0}^{n-k-1}\frac{n-k-l}{n-k}\binom{n-k-1+l}{l}p^lq^{n-k} \tag{3}$$

Combining (2) and (3) and using the fact that $P_i^k(j,n) = 0$ for $1 \le k < m-i+j$, we obtain

$$\begin{aligned}P_i(j,n) &= \\ &\sum_{k=m-i+j}^{n}\frac{m-i+j-1}{2k-(m-i+j)-1}\binom{2k-(m-i+j)-1}{k-1}p^{k-1}q^{k-(m-i+j)-1} \\ &\cdot\sum_{l=0}^{n-k-1}\frac{n-k-l}{n-k}\binom{n-k-1+l}{l}p^lq^{n-k}\end{aligned} \tag{4}$$

Computing $P_i(\ge j,n)$, that is the probability of losing more than j packets, one needs to sum (2) on all $k > m-i+j$, since event B becomes irrelevant in that case. The result can be therefore obtained in $O(n-j+1)$ computations.

Finally, using the same type of reasoning, it is possible to compute $P(\ge j,n)$

with the following result:

$$
\begin{aligned}
P(j,n) = & \\
& \Pi(m)\sum_{k=j}^{n}\frac{j-1}{2k-j-1}\binom{2k-j-1}{k-1}p^{k-1}q^{k-j}(n-k+1) \\
& -\Pi(m)\sum_{k=j+1}^{n}\frac{j}{2k-j-2}\binom{2k-j-2}{k-1}p^{k-1}q^{k-j-1}(n-k+1) \quad (5)
\end{aligned}
$$

This probability can be computed in only $O(m+n-j+1)$ steps.

Case 2: The block is larger than the buffer.

This case is more complex than the previous, because of the fact that the buffer can empty itself one or several times during the transmission, which affects the path, and after that, some packets can still be lost. A typical path can be decomposed in a series of events of three different types:
- Event $V_i(k_1)$: from the first packet arrival to the first loss (arrival of packet k_1)
- Events $S(k_l, k_{l+1})$: this event comprises the arrivals and departures between the the loss of the k_lth packet and the k_{l+1}th packet.
- Event $U(k_j)$: from the loss of the jth packet to the last packet's arrival.

The probability of the loss of one packet $P_i(1,n)$ is clearly equal to:

$$P_i(1,n) = \sum_{k_1=1}^{n} v_i(k_1)u(k_1)$$

while the loss of two packets has a probability equal to:

$$P_i(2,n) = \sum_{k_1=1}^{n-1}\sum_{k_2=k_1+1}^{n} v_i(k_1)s(k_1,k_2)u(k_2)$$

Continuing in that fashion, we find that :

$$
\begin{aligned}
P_i(j,n) &= \sum_{k_1=1}^{n-j}\sum_{k_2=k_1+1}^{n-j+1}\cdots\sum_{k_j=k_{j-1}+1}^{n} v_i(k_1)s(k_1,k_2)...s(k_{j-1},k_j)u(k_j) \\
&= \sum_{k_1=1}^{n}\sum_{k_2=k_1+1}^{n}\cdots\sum_{k_j=k_{j-1}+1}^{n} v_i(k_1)s(k_1,k_2)...s(k_{j-1},k_j)u(k_j).
\end{aligned}
$$

This last equality is obtained using the fact that $s(k_l,k_{l+1})=0$ when $k_l \geq k_{l+1}$ for $1 \leq l \leq j-1$. This can also be rewritten in a matrix form as:

$$P_i(j,n) = V_i \cdot S^{j-1} \cdot U^T,$$

where V_i and U are n-length row vectors with elements $v_i(k_j)$ and $u(k_j)$ respectively and S is an $n \times n$ matrix with elements $s(k_j, k_{j+1})$. We have now to determine the elements of this matrix. In the following, we call $\zeta_\nu(i,\mu)$ the probability of a path with ν steps, starting with a buffer containing i packets and finishing with μ packets.

The probability $v_i(k_1)$ is the probability of a path containing $\nu = 2k_1 - m + i - 3$ steps (since there are $k_1 - 2$ arrivals up to the time where the buffer is full and $m - i - 1$ more arrivals than departures (otherwise the buffer would not be full), that is $p \cdot \zeta_{2k_1 - m + i - 3}(i+1, m)$. It is obviously equal to 0 when $k_1 \leq m - i$.

The probability $s(k_l, k_{l+1})$ is the probability of a path from the point where the buffer is full to a point where the buffer is also full. This path is composed of $k_{l+1} - k_l - 1$ arrivals and $k_{l+1} - k_l - 1$ departures. One has to add a last move, which is necessarily an arrival (the loss of the k_{l+1}th packet). That is why the sought probability is equal to $p \cdot \zeta_{2(k_{l+1} - k_l - 1)}(m, m)$.

The problem of determining the probability of $u(k_j)$ is restricted to the only cases where k_j is not the last packet. The path ends necessarily with an arrival, so we only consider the possible paths until the last move and multiply the probability of each of them with q. Until the last move, we count $n - k_j - 1$ departures. If we call h the number of departures, we get

$$u(k_j) = \sum_{h=max(0, n-k_j-m)}^{n-k_j-1} \zeta_n - k_j - 1 + h(m, m - n + k_j + h + 1)$$

unless $k_j = n$ (then $u(k_j) = 1$).

It remains now to compute the probabilities $\zeta_\nu(i, \mu)$. To that end, we will use another version of the ballot theorem.

Theorem 5 *The number of possible arrangements of votes so that A is always ahead of B by less than c votes and more than $c - d$ votes ($0 < c < d$) is equal to*

$$\sum_{\Upsilon} [\binom{a+b}{b - \Upsilon d} - \binom{a+b}{b + c - \Upsilon d}],$$

where $-\infty < \Upsilon < \infty$ takes values so that the binomial coefficients are proper (in the first sum $\Upsilon d < b$ and $a > -\Upsilon d$ and in the second $\Upsilon d < b + c$ and $a - c > -\Upsilon d$).

Using this version of the ballot theorem, with $a = (\nu - i + \mu)/2$ being the number of arrivals, $b = (\nu + i - \mu)/2$ being the number of departures, $c = m + 1 - i$ the threshold by which the number of arrivals has to exceed the number of departures to lose a packet and $c - d = -i$ the threshold to empty the buffer, we get the number of possible paths:

$$\sum_{\Upsilon}[\binom{\nu}{\frac{\nu+i-\mu}{2}-\Upsilon(m+1)}-\binom{\nu}{\frac{\nu-i-\mu}{2}-\Upsilon(m+1)}]$$

The probability of a given path is equal to

$$p^{\frac{\nu-i+\mu}{2}}q^{\frac{\nu+i-\mu}{2}}$$

since there are $\frac{\nu-i+\mu}{2}$ arrivals and $\frac{\nu+i-\mu}{2}$ departures.

We obtain the probability of paths that do not cause the buffer to be empty to be equal to:

$$\xi_\nu(i,\mu)=\sum_{\Upsilon}[\binom{\nu}{\frac{\nu+i-\mu}{2}-\Upsilon(m+1)}-\binom{\nu}{\frac{\nu-i-\mu}{2}-\Upsilon(m+1)}]p^{\frac{\nu-i+\mu}{2}}q^{\frac{\nu+i-\mu}{2}}$$

This expression is only valid if $i \geq 1$ and $\mu \geq 1$.

Considering the case when $i=0$ and $\mu=0$, one notes that the probability of a path with ν steps starting with an empty buffer ending with μ packets in the buffer is equal with the probability of a path with $\nu-1$ steps starting with one packet in the buffer and ending with μ packets in the buffer (since the first step in a path with an empty buffer is necessarily an arrival).

One notes also that the probability of a path starting with $i \geq 1$ packets and ending with an empty buffer is equal to the probability of a path starting with i packets and ending with 1 packet in the buffer multiplied by q (the probability of a departure on the last step).

With these two results in mind, it is possible to see that the probability $\xi_\nu(0,0)$ of starting with an empty path and ending with an empty path is equal to

$$\xi_\nu(0,0)=q\xi_{\nu-2}(1,1)$$

This completes the determination of the probability of a path starting with i packets in the buffer and ending with μ packets. The only remaining point now is to consider the paths where the buffer is emptied one or several times.

Any such path is a composition of the following parts:
- Event W: from the start to the first time the path is emptied,
- Events Y: from a point where the buffer is empty to the next occasion where the buffer is empty,
- Event Z: from the last point where the buffer is empty to the end of the path.

Just as in the previous case (with j losses), one can sum the probability of a path that empties itself one or several times as the product of three matrices $W_i \cdot Y^{r-1} \cdot Z^T$.

The elements of the row matrix W_i is the probability for a path starting with i packets in the buffer and emptied after $i+2l$ events. This probability is $\xi_{i+2l}(i,0)$.

The lth element of the vector matrix Z is the probability of a path starting with an empty buffer and finishing with μ packets after $\nu - i + 2l$ steps. This probability is just $\xi_{\nu-i+2l}(0,\mu)$.

Finally, the lth element of the matrix Y is the probability of a path starting and ending with an empty buffer after $2(l_n - l_{n-1})$ events. This probability is by definition $\xi_{2(l_n-l_{n-1})}(0,0)$. We get

$$Y = \begin{pmatrix} \xi_2(0,0) & \xi_4(0,0) & \cdots & \xi_{2(R-1)}(0,0) \\ 0 & \xi_2(0,0) & \cdots & \xi_{2(R-2)}(0,0) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \xi_2(0,0) \\ 0 & 0 & 0\cdots & 0 \end{pmatrix}$$

where R is the maximal number of times the buffer can empty itself.

To compute this matrix, one needs to perform $O(j \cdot n^2)$. We obtain eventually $\zeta_\nu(i,\mu)$ by summing over all paths that cause the buffer to empty itself over r $(0 \leq r \leq R)$. We get that:

$$\zeta_\nu(i,\mu) = \xi_\nu(i,\mu) + \sum_{r=1}^{R} W_i \cdot Y^{r-1} \cdot Z^T$$

This computation requires $O(R^2 \cdot n^2)$ operations. We can now compute all elements of the matrices V_i, U and S. We obtain then $P_i(j,n)$ for any j, which gives in turn $P_i(\geq j,n)$, the probability of losing at least j packets, using the following formula:

$$P_i(\geq j,n) = V_i \cdot S^{j-1} \cdot E^T$$

where E is a unit row vector that replaces U in our calculations, since event U becomes irrelevant once one is looking for the probability of at least j losses. This probability can be computed with $O(R^2 \cdot n^3)$ operations.

6 Some consequences of the ballot theorem

In this section, we will examine some of the consequences of the ballot theorem in terms of random walks models.

6.1 First time passage

The classical ballot theorem can be used to analyze the first passage time distribution. Speaking in terms of lattice walks and considering a particular random walk, we say that a first passage through m occurs at step k if the walk touches for the first time the line $y = m$. In terms of the ballot problem, we see that this happens when the difference in votes counted for A and B reaches m for the first time.

If S_c represents the difference of votes between the two candidates A and B, we have $S_c = \alpha_c - \beta_c$, where α_c and β_c are respectively the number of votes for A and B after c ballots have been counted. Then, a first passage through m occurs at step k if

$$S_1 < m, S_2 < m, S_3 < m, ..., S_{k-1} < m, S_k = m.$$

This series of inequalities is defined for $m > 0$. If $m < 0$, the inequalities are simply reversed.

Using the representation aforementioned (with a vote for A represented as a step southeastwards and a vote for B represented as a step northeastwards), we can define a first passage through m as the first time the path reaches the line $y = -m$.

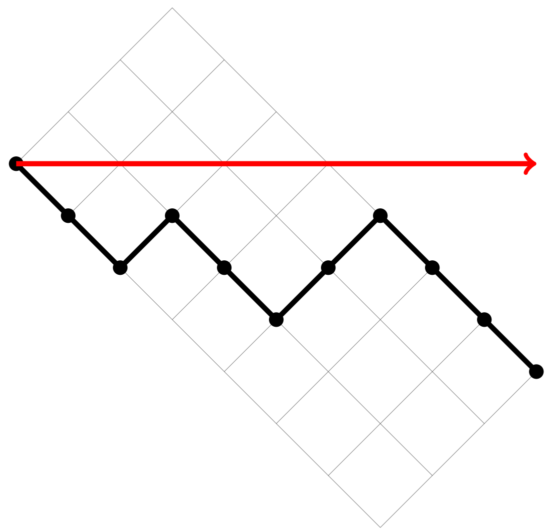

Figure 8: The representation of a ballot sequence AABAABBAAA. We notice that the maximal difference between the votes cast for A and the votes cast for B is achieved when the last ballot has been counted.

In the litterature on the subject, another representation is often used to represent such ballot sequences. Instead of a move southeastwards for each vote for A and a move northeastwards for each vote for B, each vote for B is represented as a move northwards (0,1), while a vote for A is represented as a move eastwards. This gives us the following representation:

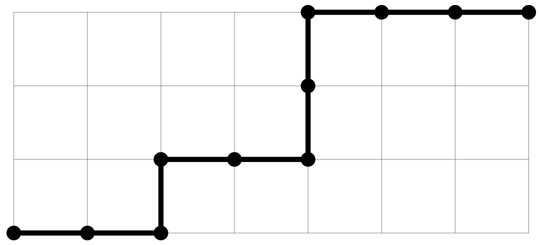

Figure 9: Alternative representation for the same ballot sequence.

Using this representation, the walk ends at the node (α, β).

The problem of first passage can be reduced to ballot problems. Indeed, in figure 8, we have an example of count that achieves first passage through $m = -4$ when the last ballot has beed tallied. If we consider, within a longer count, the series of ballots that achieves a first passage through $m = -4$ at the 10th step, then we can count $N_{10,4}$ the number of ways to achieve such a result:

$$N_{10,4} = \frac{4}{10}\binom{10}{7}$$

This result can be explained in the following way: the number of ways to achieve a first passage through $m = -4$ after 10 ballots have been counted is equal to the number of ways one can count 10 ballots, with 7 for candidate A and 3 for candidate B.

To prove this result, we use the latter representation (figure 9) of that very same count. We begin by reversing the steps, obtaining a reflexion of the random walk showed in figure 8.

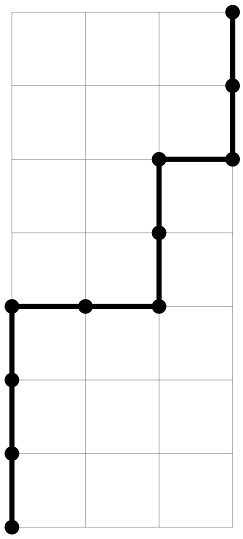

Figure 10: Reversing the steps of the count.

This representation takes us back to a traditional ballot problem and, generalizing this example, we realize that the number of walks with first passage through $m \geq 1$ at step $2n+m$ is equal to the number of ways of counting ballots with A keeping an advantage throughout the count with $m = \beta - \alpha$, $\alpha = n$ and $\beta = n + m$.

We get therefore the number of walks $N_{2n+m,m}$ with first passage $m \geq 1$ at step $2n + m$:

$$N_{2n+m,m} = \frac{m}{2n+m}\binom{2n+m}{n}$$

Hence, in the so-called symmetric random walk model where a step in both move is equally plausible, the probability $P_{2n+m,m}$ of first passage through $m \geq 1$ at step $2n + m$ is therefore:

$$P_{2n+m,m} = \frac{m}{2n+m}\binom{2n+m}{n}2^{-(2n+m)},$$

since there are $2^{(2n+m)}$ possible walks by the rule of product.

6.2 The duration of games

A few months after he had published his short article on what would become remembered as the "ballot problem", Bertrand wrote:

> "Though I proposed this curious question as an exercise in reason and calculation, in fact it is of great importance. It is linked to the important question of the duration of games, previously considered by Huygens, Moivre, Laplace, lagrange and Ampère. The problem is this: a gambler plays a game in which in each round he wagers $\frac{1}{n}$'th of his initial fortune. What is the probability he is eventually ruined and that he spends his last coin in the $2n+m$'th round?" (reference)

Considering the rounds in reverse order, Bertrand notes that the probability equals

$$P_{2n+m,m} = \frac{m}{2n+m}\binom{2n+m}{n}2^{-(2n+m)}.$$

He goes on to notice that the probability of ruin occuring before the $2n+m$'th round is approximately $1 - \frac{\sqrt{2/n\pi}}{\sqrt{n+2m}}$. This implies that, for this probability to be large, m has to be large compared to n^2.

6.3 Number of visits

The question of the number of visits, that is, the number of times a random walk reaches a certain height in our model, is closely linked to the results we have just shown. In terms of ballot problems, we are here interested in knowing how many times the difference $S_k = \beta_k - \alpha_k$ of votes separating A from B equals a given value j.

Using the same correspondence as in section 5.2, we find the following distribution of the number of visits to c.

Theorem 6 *If $c \geq 0$ then we have*

$$P_1 = 2^j \frac{\alpha-(\beta-j)}{\alpha+(\beta-j)}\binom{\alpha+(\beta-j)}{\alpha}/\binom{\alpha+\beta}{\alpha} for\ \alpha \geq \beta\ and\ 0 \leq j \leq \beta.$$

Proof: To prove this theorem, we take advantage of the notion of representative path developed by Feller. A path is said to be representative if it starts at the origin and ends at the point $(n,0)$ with all sides below the x-axis and touches the axis precisely j times. One can decompose this path into j sections with endpoints on the x-axis. By the rule of product, we deduct that we can construct 2^j by mirroring the sections on this line and preserving the number of encounters with the x-axis.

Using this representative path, we remove then the j moves that end on the x-axis. One obtains therefore a path with $\alpha + \beta - j$ moves that never touches the x-axis and remains under it throughout the whole count. The original ballot theorem gives us that there are $\frac{\alpha-(\beta-j)}{\alpha+(\beta-j)}\binom{\alpha+(\beta-j)}{\alpha}/\binom{\alpha+\beta}{\alpha}$ such paths. Multiplying with 2^j, we get Theorem 6.

7 Perspectives

The ballot theorems presented in this essay are only the oldest - and most elementar - ones. They have the particularity that they apply only in the field of discrete mathematics, since they are only verified for integers. Indeed, in the generalized ballot theorem, the result only applies if k is an integer.

The question that comes directly to mind concerns the existence of such results extended to the real numbers. What it the probability that candidate A has been ahead by, with a number of ballots that was, at each point during the count, at least 1,5 time the number of ballots for candidate B?

Mathematicians have dedicated a lot of research to the question of extending the results presented in the first ballot theorems to the field of real numbers. These researches have led to a number of so-called "continuous time ballot theorems", as opposed to the ones presented in this essay, qualified as "discrete time ballot theorems".

In 1965, Takàcs presented the first one of these theorems, showing that:

Theorem 7 *If $X_t, 0 \leq t \leq r$ is a separable stochastic process with cyclically interchangeable increments whose sample functions (i.e. a function $x_\omega : T \to \mathbb{R}$ given by fixing some $\omega \in \Omega$ and letting $x_\omega(t) = X_t(\omega)$) are almost surely non-decreasing step functions, then*

$$P\{X_t - X_0 \leq t \; for \; 0 \leq t \leq r \mid X_r - X_0 = s\} = \begin{cases} \frac{t-s}{t} & if \; 0 \leq s \leq t \\ 0 & otherwise. \end{cases} \tag{6}$$

This theorem is directly derived from the discrete time case. But, according to Addario-Berry and Reed, the first and only real ballot-style result proved for random walks that may take non-integer values is the following:

Theorem 8 *If X is a real random variable with maximum value 1 and $\{X_1, X_2, ..., X_n\}$ are independent and identically distributed copies of X with corresponding partial sums $\{0 = S_0, S_1, ..., S_n\}$, then*

$$P\{S_i > 0 \;\forall\; 1 \leq i \leq n \mid S_n\} = \frac{S_n}{n}.$$

As Addario-Berry and Reed notice, research focuses nowadays on moving "towards making ballot theorems part of the general theory of random walks - part of the body of results that hold for *all* random walks (with identically distributed steps), regardless of the precise distribution of their steps". As of today, the theorems hold for a large class of random walks, but not all of them. The authors went on to conclude: "A *truly* general ballot theorem, however, remains beyond our grasp".

8 References

L. ADDARIO-BERRY, B.A. REED: *Ballot theorems, old and new*, Horizons of Combinatorics vol. 17 (2007). Springer, Berlin.

D. ANDRÉ: *Solution directe du problème résolu par M. J. Bertrand*, Comptes Rendus de l'Académie des Sciences vol. 105 (1887), p. 436-437.

É. BARBIER: *Généralisation du problème résolu par M. J. Bertrand*, Comptes Rendus de l'Académie des Sciences vol. 105 (1887), p. 407.

J. BERTRAND: *Solution d'un problème*, Comptes Rendus de l'Académie des Sciences vol. 105 (1887), pp. 369.

J. BERTRAND: *Observations*, Comptes Rendus de l'Académie des Sciences vol. 105 (1887), pp. 437-439.

S. R. DUNBAR: *Topics in probability and stochastic processes: The ballot theorem and the reflexion principle*, 2013.

W. FELLER: *An Introduction to Probability Theory and its Applications*, John Wiley and Sons Inc., 1968.

I.P. GOULDEN, L. G. SERRANO: *Maintaining the spirit of the reflection principle when the boundary has arbitrary integer slope*, Journal of Combinatorial Theory, Series A vol. 104 (2003), pp. 317-326.

O. GUREWITZ, M. SIDI, I. CIDON *The ballot theorem strikes again: packet loss process distribution*, IEEE Transactions on Information Theory, vol. 46, issue 7 (2000), pp. 2588-2595.

P. HILTON, J. PEDERSEN: *Catalan numbers, their generalization and their uses*, The Mathematical Intelligencer vol. 13, issue 2 (1991), pp. 64-75. Springer-Verlag, New York.

M.L. LACKNER, M. WALLNER: *An invitation to AC and LP counting*, Lecture notes of a mini-course held at the ALEA in Europe Young Researchers' Workshop, University of Bath, Bath, UK, December 2015.

T. LENGYEL: *Direct consequences of the basic ballot theorem*, Statistic and Probability Letters vol. 81, issue 10 (2011), pp. 1476–1481. Elsevier, Amsterdam.

J. K. PERCUS: *Combinatorial methods* (1971), Springer Verlag, New York.

M. RENAULT: *Four proofs of the ballot theorem*, Mathematics Magazine, vol. 80, issue 5 (2007), pp.345-352.

M. RENAULT: *Lost (and found) in translation: André's actual method and its application to the generalized ballot problem*, American Mathematical Monthly vol. 115, issue 4 (2008), pp.358-363.

R. P. STANLEY: *Catalan numbers*, (2015), Cambridge University Press, Cambridge.