

Bertrand's ballot theorem

From Wikipedia, the free encyclopedia

In combinatorics, **Bertrand's ballot problem** is the question: "In an election where candidate A receives p votes and candidate B receives q votes with $p > q$, what is the probability that A will be strictly ahead of B throughout the count?" The answer is

$$\frac{p - q}{p + q}.$$

The result was first published by W. A. Whitworth in 1878, but is named after Joseph Louis François Bertrand who rediscovered it in 1887.^{[1][2]}

In Bertrand's original paper, he sketches a proof based on a general formula for the number of favourable sequences using a recursion relation. He remarks that it seems probable that such a simple result could be proved by a more direct method. Such a proof was given by Désiré André,^[3] based on the observation that the unfavourable sequences can be divided into two equally probable cases, one of which (the case where B receives the first vote) is easily computed; he proves the equality by an explicit bijection. A variation of his method is popularly known as **André's reflection method**, although André did not use any reflections.^[4]

Contents

- 1 Example
- 2 Equivalent problems
- 3 Proof by reflection
- 4 Proof by induction
- 5 Bertrand's and André's proofs
- 6 Variant: ties allowed
- 7 Notes
- 8 References
- 9 External links

Example

Suppose there are 5 voters, of whom 3 vote for candidate A and 2 vote for candidate B (so $p = 3$ and $q = 2$). There are ten possibilities for the order of the votes cast:

- AAABB
- AABAB
- ABAAB
- BAAAB
- AABBA
- ABABA
- BAABA
- ABBAA
- BABAA
- BBAAA

For the order AABAB, the tally of the votes as the election progresses is:

Candidate	A	A	B	A	B
A	1	2	2	3	3
B	0	0	1	1	2

For each column the tally for *A* is always larger than the tally for *B* so the *A* is always strictly ahead of *B*. For the order *AABBA* the tally of the votes as the election progresses is:

Candidate	A	A	B	B	A
A	1	2	2	2	3
B	0	0	1	2	2

For this order, *B* is tied with *A* after the fourth vote, so *A* is not always strictly ahead of *B*. Of the 10 possible orders, *A* is always ahead of *B* only for *AAABB* and *AABAB*. So the probability that *A* will always be strictly ahead is

$$\frac{2}{10} = \frac{1}{5},$$

and this is indeed equal to $\frac{3-2}{3+2}$ as the theorem predicts.

Equivalent problems

Rather than computing the probability that a random vote counting order has the desired property, one can instead compute the number of favourable counting orders, then divide by the total number of ways in which the votes could have been counted. (This is the method used by Bertrand.) The total number of ways is the binomial coefficient $\binom{p+q}{p}$; Bertrand's proof shows that the number of favourable orders in which to count the votes is $\binom{p+q-1}{p-1} - \binom{p+q-1}{p}$ (though he does not give this number explicitly). And indeed after division this gives $\frac{p}{p+q} - \frac{q}{p+q} = \frac{p-q}{p+q}$.

Another equivalent problem is to calculate the number of random walks on the integers that consist of *n* steps of unit length, beginning at the origin and ending at the point *m*, that never become negative. Assuming *n* and *m* have the same parity and *n* ≥ *m* ≥ 0, this number is

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

When *m* = 0 and *n* is even, this gives the Catalan number $\frac{1}{\frac{n}{2} + 1} \binom{n}{\frac{n}{2}}$.

Proof by reflection

For *A* to be strictly ahead of *B* throughout the counting of the votes, there can be no ties. Separate the counting sequences according to the first vote. Any sequence that begins with a vote for *B* must reach a tie at some point, because *A* eventually wins. For any sequence that begins with *A* and reaches a tie, reflect the votes up to the point of the first tie (so any *A* becomes a *B*, and vice versa) to obtain a sequence that begins with *B*. Hence every sequence that begins with *A* and reaches a tie is in one to one correspondence with a sequence that begins with *B*, and the probability that a sequence begins with *B* is $q/(p+q)$, so the probability that *A* always leads the vote is

$$\begin{aligned}
&= 1 - \text{the probability of sequences that tie at some point} \\
&= 1 - \text{the probability of sequences that tie at some point and begin with A or B} \\
&= 1 - 2 \frac{q}{p+q} = \frac{p-q}{p+q}
\end{aligned}$$

Proof by induction

Another method of proof is by mathematical induction:

- Clearly the theorem is true if $p > 0$ and $q = 0$ when the probability is 1, given that the first candidate receives all the votes; it is also true when $p = q > 0$ since the probability is 0, given that the first candidate will not be *strictly* ahead after all the votes have been counted.
- Assume it is true both when $p = a - 1$ and $q = b$, and when $p = a$ and $q = b - 1$, with $a > b > 0$. Then considering the case with $p = a$ and $q = b$, the last vote counted is either for the first candidate with probability $a/(a+b)$, or for the second with probability $b/(a+b)$. So the probability of the first being ahead throughout the count to the penultimate vote counted (and also after the final vote) is:

$$\frac{a}{(a+b)} \frac{(a-1)-b}{(a+b-1)} + \frac{b}{(a+b)} \frac{a-(b-1)}{(a+b-1)} = \frac{a-b}{a+b}.$$

- And so it is true for all p and q with $p > q > 0$.

Bertrand's and André's proofs

Bertrand expressed the solution as

$$\frac{2m - \mu}{\mu}$$

where $\mu = p + q$ is the total number of voters and $m = p$ is the number of voters for the first candidate. He states that the result follows from the formula

$$P_{m+1, \mu+1} = P_{m, \mu} + P_{m+1, \mu},$$

where $P_{m, \mu}$ is the number of favourable sequences, but "it seems probable that such a simple result could be shown in a more direct way". Indeed, a more direct proof was soon produced by Désiré André. His approach is often mistakenly labelled "the reflection principle" by modern authors but in fact uses a permutation. He shows that the "unfavourable" sequences (those that reach an intermediate tie) consist of an equal number of sequences that begin with A as those that begin with B. Every sequence that begins with B is unfavourable, and there are $\binom{p+q-1}{q-1}$ such sequences with a B followed by an arbitrary sequence of $(q-1)$ B's and p A's. Each unfavourable sequence that begins with A can be transformed to an arbitrary sequence of $(q-1)$ B's and p A's by finding the first B that violates the rule (by causing the vote counts to tie) and deleting it, and interchanging the order of the remaining parts. To reverse the process, take any sequence of $(q-1)$ B's and p A's and search from the end to find where the number of A's first exceeds the number of B's, and then interchange the order of the parts and place a B in between. For example, the unfavourable sequence **A**AB**B**ABAA corresponds uniquely to the arbitrary sequence ABAA**A**AB. From this, it follows that the number of favourable sequences of p A's and q B's is

$$\binom{p+q}{q} - 2 \binom{p+q-1}{q-1} = \binom{p+q}{q} \frac{p-q}{p+q}$$

and thus the required probability is

$$\frac{p - q}{p + q}$$

as expected.

Variant: ties allowed

The original problem is to find the probability that the first candidate is always strictly ahead in the vote count. Consider now the problem to find the probability that the second candidate is never ahead (i.e. ties are allowed); the solution is

$$\frac{p + 1 - q}{p + 1}.$$

The variant problem can be solved by the reflection method in a similar way to the original problem. First note that the number of possible vote sequences is $\binom{p+q}{q}$. Call a sequence "bad" if the second candidate is ever ahead, and if the number of bad sequences can be enumerated then the number of "good" sequences can be found by subtraction and the probability can be computed.

Represent a voting sequence as a lattice path on the Cartesian plane as follows:

- Start the path at (0, 0)
- Each time a vote for the first candidate is received move right 1 unit.
- Each time a vote for the second candidate is received move up 1 unit.

Each such path corresponds to a unique sequence of votes and will end at (p, q). A sequence is 'good' exactly when the corresponding path never goes above the diagonal line $y = x$; equivalently, a sequence is 'bad' exactly when the corresponding path touches the line $y = x + 1$.

For each 'bad' path P , define a new path P' by reflecting the part of P up to the first point it touches the line across it. P' is a path from (-1, 1) to (p, q). The same operation applied again restores the original P . This produces a one-to-one correspondence between the 'bad' paths and the paths from (-1, 1) to (p, q). The number of these paths is $\binom{p+q}{q-1}$ and so that is the number of 'bad' sequences. This leaves the number of 'good' sequences as

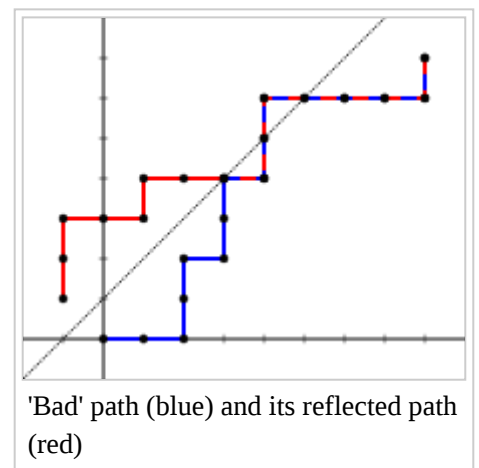
$$\binom{p+q}{q} - \binom{p+q}{q-1} = \binom{p+q}{q} \frac{p+1-q}{p+1}.$$

Since there are $\binom{p+q}{q}$ altogether, the probability of a sequence being good is $\frac{p+1-q}{p+1}$.

In fact, the solutions to the original problem and the variant problem are easily related. For candidate A to be strictly ahead throughout the vote count, they must receive the first vote and for the remaining votes (ignoring the first) they must be either strictly ahead or tied throughout the count. Hence the solution to the original problem is

$$\frac{p}{p+q} \frac{p-1+1-q}{p-1+1} = \frac{p-q}{p+q}$$

as required.



Conversely, the tie case can be derived from the non-tie case. Note that the *number* of non-tie sequences with $p+1$ votes for A is equal to the number of tie sequences with p votes for A. The number of non-tie votes with $p+1$ votes for A votes is $\frac{p+1-q}{p+1+q} \binom{p+1+q}{q}$, which by algebraic manipulation is $\frac{p+1-q}{p+1} \binom{p+q}{q}$, so the *fraction* of sequences with p votes for A votes is $\frac{p+1-q}{p+1}$.

Notes

1. Feller, William (1968), *An Introduction to Probability Theory and its Applications, Volume I* (3rd ed.), Wiley, p. 69.
2. J. Bertrand, Solution d'un problème, Comptes Rendus de l'Académie des Sciences, Paris 105 (1887), 369.
3. D. André, Solution directe du problème résolu par M. Bertrand, Comptes Rendus de l'Académie des Sciences, Paris 105 (1887) 436–437.
4. Renault, Marc, Lost (and found) in translation: André's actual method and its application to the generalized ballot problem. Amer. Math. Monthly 115 (2008), no. 4, 358--363. (<http://webpace.ship.edu/msrenault/ballotproblem/montly358-363-renault.pdf>)

References

- Ballot theorems, old and new (<http://www.dms.umontreal.ca/~addario/papers/btsurvey.pdf>), L. Addario-Berry, B.A. Reed, 2007, in Horizons of combinatorics (<http://books.google.com/books?id=kIKW18ENfUMC>), Editors Ervin Györi, G. Katona, Gyula O. H. Katona, László Lovász, Springer, 2008, ISBN 978-3-540-77199-9

External links

- The Ballot Problem (<http://webpace.ship.edu/msrenault/ballotproblem/>) (includes scans of the original French articles and English translations)
- Bernard Bru, Les leçons de calcul des probabilités de Joseph Bertrand (<http://www.jehps.net/Decembre2006/Bru.pdf>), history of the problem (in French)
- Weisstein, Eric W. "Ballot Problem". *MathWorld*.

Retrieved from "https://en.wikipedia.org/w/index.php?title=Bertrand%27s_ballot_theorem&oldid=703412476"

Categories: Named probability problems | Enumerative combinatorics | Theorems in combinatorics | Probability theorems

-
- This page was last modified on 5 February 2016, at 10:39.
 - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.