How to Count

2

2.1 Permutations and combinations

The key to elementary probability calculations is an ability to count outcomes of interest among a given set of possibilities called the sample space. Exhaustive enumeration of cases often is not feasible. There are, fortunately, systematic methods of counting that do not require actual enumeration. In this chapter we introduce these methods, giving some applications to fairly challenging probability problems, but defer to the next chapter the formal theory of probability.

Basic formula

A basic counting principle is this: if A can occur in m different ways and B in n different ways, then, absent any restrictions on either A or B, both can occur in m times n different ways. For example, if in a given political party there are three potential candidates for president, and four for vice-president, then, barring any prohibited marriages, there are 12 possible tickets for the two executive positions.

Permutations

A permutation is an ordered arrangement of distinct items into distinct positions or slots. If there are n items available to fill slot 1, there will be $n-1$ items remaining to fill slot 2, resulting in $n\,(n-1)$ ways to fill two slots. Continuing in this manner, there are $n\cdot(n-1)\cdots(n-r+1)$ ways to fill $r\leq n$ slots. This permutation number is given the symbol ${}_nP_r$. In particular, there are ${}_nP_n = n\cdot(n-1)\cdots 2\cdot 1$ ways to arrange all n items, and this number is given the special symbol $n!$, read "n factorial." Factorial notation is well defined for positive integers. For $n=0$, because there is precisely one way to arrange an empty set of items into zero slots (the arrangement is vacuous, but it is an arrangement), the definition $0!=1$ is adopted as consistent and useful. Note that we can write ${}_nP_r$ with factorial notation as $n!/(n-r)!$

© Springer Science+Business Media, LLC 2015

M.O. Finkelstein, B. Levin, *Statistics for Lawyers*, Statistics for Social and Behavioral Sciences, DOI 10.1007/978-1-4419-5985-0_2

Combinations

Suppose now that in selecting r items out of n our only interest is in the final collection of items as opposed to the number of arrangements. Then among the ${}_nP_r$ possible arrangements of these r items, each collection or combination is repeated $r!$ times (once for each rearrangement of the same r items), so that the number of combinations of r items out of n is ${}_nP_n/r!$. This combination number, sometimes given the symbol ${}_nC_r$, is most often denoted by the binomial coefficient, $\binom{n}{r}$, read "n choose r." Thus

$$\binom{n}{r} = {}_nC_r = \frac{{}_nP_r}{r!} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)\cdots(n-r+1)}{r(r-1)\cdots 2\cdot 1}.$$

Note that in the final expression both numerator and denominator contain r factors. Also note that $\binom{n}{r} = \binom{n}{n-r}$. As an example, there are $\binom{5}{2} = 5!/(2!\cdot 3!) = (5\cdot 4)/(2\cdot 1) = 10$ committees of two that can be formed from a pool of five people. There are also $\binom{5}{3} = 10$ committees of three that can be formed (e.g., those left in the pool after selecting committees of two).

The binomial coefficient also counts the number of different patterns of n items that are of two types, say r of type 1 and $n-r$ of type 0, when items of the same type are *indistinguishable*. Among the $n!$ potential arrangements, there are $r!$ repetitions due to indistinguishable rearrangements of type 1 items, and there are $(n-r)!$ repetitions due to rearrangements of type 0 items. Thus the $r!\cdot(n-r)!$ repetitions leave a total of only $n!/[r!\cdot(n-r)!]$ distinguishable patterns. For example, if a coin is tossed n times, there are $\binom{n}{r}$ distinct arrangements of outcomes with r heads and $n-r$ tails. More generally, if there are k types among n items—say, r_1 of type 1... and r_k of type k—with objects of the same type indistinguishable, then the number of distinct patterns is $n!/(r_1!\ldots r_k!)$; this is known as a *multinomial coefficient*. For example, there are 2,520 distinct gene sequences of length eight that can be formed from the genome $\{a, a, b, b, c, c, d, d\}$. Do you see why?

Stirling's formula

A beautiful approximation known as Stirling's formula may be used to calculate $n!$ when n is large:

$$n! \approx \sqrt{2\pi n}\cdot n^n \cdot e^{-n} = \sqrt{2\pi}\cdot n^{n+1/2}\cdot e^{-n}$$

where e is Euler's constant ≈ 2.718. A slightly more accurate version of Stirling's formula is

$$n! \approx \sqrt{2\pi n}\cdot\left(n+\frac{1}{2}\right)^{n+\frac{1}{2}}\cdot e^{-\left(n+\frac{1}{2}\right)}.$$

For example, when $n=10$, $n! = 10\cdot 9\cdot 8\cdots 3\cdot 2\cdot 1 = 3{,}628{,}800$. Stirling's formula gives 3,598,696 (99.2% accuracy), while the second approximation gives 3,643,221

(99.6% accuracy). Stirling's formula may be applied to the factorials in a binomial coefficient to yield, for large n and r,

$$\binom{n}{r} \approx \frac{1}{\sqrt{2\pi}} \cdot \frac{n^{n+\frac{1}{2}}}{r^{r+\frac{1}{2}} \cdot (n-r)^{n-r+\frac{1}{2}}}$$

For example, when $n = 2r$, the binomial coefficient $\binom{n}{r}$ is approximately $2^n/(\pi r)^{1/2}$. This result is used in Sections 2.1.2 and 2.2.

Occupancy problems

Some probability problems correspond to placing balls at random into cells. The number of balls in each cell is known as the occupancy number for that cell, or the cell frequency. To compute the number of ways of obtaining a set of occupancy numbers (without regard to the order of the cells or the order of the balls), multiply (i) the number of ways of distributing the balls to arrive at a specific sequence of cell frequencies by (ii) the number of different sequences of cell frequencies with the given set of occupancy numbers. To calculate the probability of obtaining a particular set of occupancy numbers, assuming equally likely outcomes, divide the total number of distributions that produce that set of occupancy numbers by the total number of ways of putting the balls into the cells. If there are n balls distributed into k cells, the total number of possible distributions is k^n.

The birthday problem is an example of an occupancy problem. Given a room with N people, what is the probability that two or more people have the same birthday? How large must N be for this chance to be at least 1/2? In this problem there are $k = 365$ cells and N balls. No coincidences implies a set of occupancy numbers all zero or one. Assuming that all birth dates are equally likely, the probability of the sequence of cell frequencies (1, 1,..., 1, 0, 0,..., 0) is $[N!/(1!...0!)]/k^N = N!/k^N$, and there are $\binom{k}{N} = k!/[N!(k-N)!]$ such sequences. Thus the probability of obtaining a set of occupancy numbers all zero or one is $(N!/k^N)k!/[N!(k-N)!] = k(k-1)...(k-N+1)/k^N$, or

$$\left(\frac{365}{365}\right)\left(\frac{364}{365}\right)\cdots\left(\frac{365-N+1}{365}\right).$$

For $N = 23$, this works out to just under 0.50 so that the chance of at least one coincidence is just over 0.50.

2.1.1 DNA profiling

DNA profiling has become standard in criminal proceedings in which identification of a defendant is an issue. We describe briefly the genetic background and introduce the statistical aspects raised by the method. Further issues are discussed at Sections 3.1.2 and 3.2.2.

Human beings have 23 pairs of chromosomes in every cell except egg and sperm cells, which have 23 single chromosomes. A chromosome is a very thin thread of DNA (deoxyribonucleic acid). The thread consists of two long strings of four chemical bases twisted to form a double helix. The four bases are abbreviated A, T,G, and C (which stand for adenine, thymine, guanine, and cytosine). In double-stranded DNA, the bases line up in pairs, an A opposite a T and a G opposite a C, so that if a sequence on one strand is known the other is determined. Before cell division, the two strands separate into single strands. Each strand then picks up free-floating bases from the cell in accordance with the A-T and G-C pairing, thus creating two identical double-stranded DNA helixes, one for each cell. The process is completed when the replicated chromosome pairs separate into daughter cells. This process assures uniformity of DNA throughout the cells of the body.

A *gene* is a stretch of DNA, ranging from a few thousand to tens of thousands of base pairs, that produces a specific product, usually a protein. The position that a gene occupies on the DNA thread is its *locus*. Genes are interspersed along the length of the DNA and actually compose only a small fraction of the total molecule. Most of the rest of the DNA has no known function.

Alternative forms of genes at the same locus, like those producing both normal blood and sickle-cell anemic blood, are called alleles. A person has two genes at each locus, one from the maternal chromosome and the other from the paternal chromosome: the two genes together are referred to as the person's *genotype* at the locus. If the same allele is present in both chromosomes of a pair, the genotype is said to be *homozygous*. If the two are different, the genotype is said to be *heterozygous*. A heterozygous genotype with allele A from the maternal chromosome and allele B from the paternal chromosome cannot be distinguished from one in which allele A is from the paternal chromosome and allele B is from the maternal chromosome. However, genes on the Y chromosome can only have come from the father, which permits some spectacular lineage tracing–such as the evidence that Thomas Jefferson was indeed the father of Eston Hemings Jefferson, the younger son of his slave Sally Hemings. *DNA Test Finds Evidence of Jefferson Child by Slave*, N.Y. Times, November 1, 1998, at A1, col 5. Matrilineal descent can be traced using genes from mitochondrial DNA. Mitochondria are microscopic organelles responsible for energy storage and release found in the cell, but outside the nucleus, so they are not associated with the chromosomes. The transmission of mitochondria is from mother to child because the sperm has very little material other than chromosomes. All children of one woman will have identical mitochondrial DNA and this will be passed down through the female line to successive generations.

VNTR analysis

One group of DNA loci that were used extensively in forensic analysis are those containing Variable Numbers of Tandem Repeats (VNTRs). Technically, these are not genes because they have no known effect on the person. This is an important attribute for forensic work because it makes it less likely that the VNTRs would be

influenced by natural selection or selection of mates, which could lead to different frequencies in different populations.

A typical VNTR region consists of 500–10,000 base pairs, comprising many tandemly repeated units, each some 15–35 base pairs in length. The number of repeats, and hence the length of the VNTR region, varies from one region to another, and different regions can be distinguished by their lengths. The variation in length of VNTR regions is a form of *length polymorphism*. The term *restriction fragment length polymorphism* (RFLP) refers to length polymorphism found in fragments of DNA snipped out by a biochemical method (using restriction enzymes) which isolates the same regions of DNA on the two chromosomes. The term allele is usually applied to alternative forms of a gene; here we extend the term to include nongenic regions of DNA, such as VNTRs. Each VNTR allele is distinguished by a characteristic band on an autoradiograph in which the molecular weight of the band reflects the number of repeats in the VNTR and determines its location. If the genotype is homozygous, only a single band will appear; if heterozygous, two bands will appear. To allow for measurement error that is roughly proportional to the fragment size, preset "match windows" (e.g. $\pm 2.5\%$) around each autorad band are used and two bands are declared to match only if their match windows overlap.

Typically, there is a large number of alleles at VNTR regions (usually 15–25 can be distinguished), no one of which is common. The number of genotypes (pairs of alleles) is far larger; and when the possibilities for different loci are combined, the number of allelic combinations quickly becomes astronomical. An example of an autorad in an actual case is shown in Fig. 2.1.1.

PCR-based methods

The *polymerase chain reaction* (PCR) is a laboratory process for copying a chosen short segment of DNA millions of times. The process is similar to the mechanism by which DNA duplicates itself normally, except that by the use of enzymes only a segment of the DNA is reproduced. At present, the method is used to reproduce relatively short segments of DNA, up to 1,000 nucleotides in length, which is much shorter than most VNTRs.

There are significant advantages to this process over the VNTR process. First, it is possible to work with much smaller amounts of DNA, which is significant because forensic traces may involve minute amounts of DNA. Second, amplification of samples of degraded DNA is possible, which permits analysis of old and decayed samples. Third, it is usually possible to make an exact identification of each allele copied so that the measurement uncertainties associated with the identification of VNTRs by weight are largely eliminated.

There are also some disadvantages. The amplification process is so efficient that even a few molecules of contaminating DNA can be amplified with the intended DNA. Second, most markers used in PCR-based typing have fewer alleles than VNTRs and the distribution of allele frequencies is not as flat. Hence, more loci are required to produce the same amount of information about the probability that two persons share a profile. Furthermore, some of these loci are functional (they are

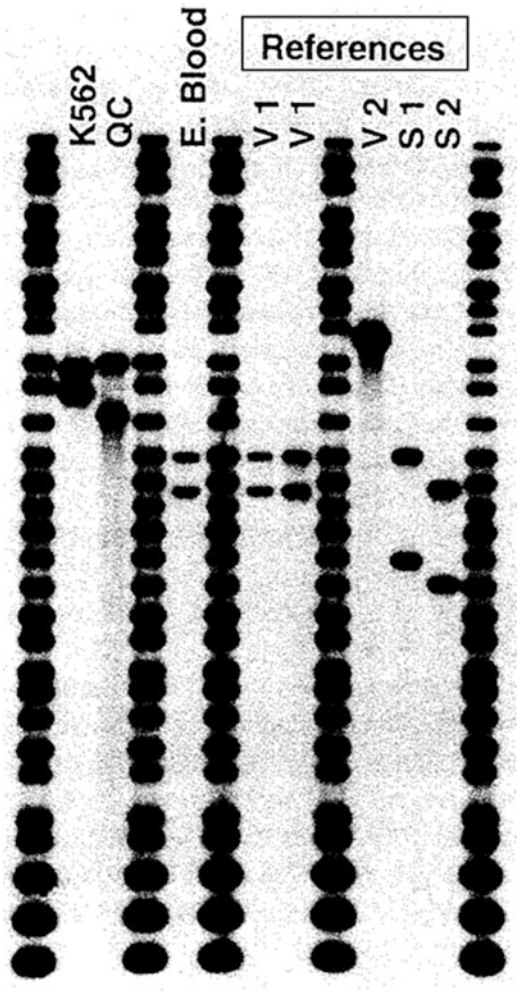

Fig. 2.1.1. An autoradiograph from an actual case illustrating RFLP at the D1S7 locus. In the case, the suspects (S-1 and S-2) were charged with having beaten to death the two victims (V-1 and V-2). Blood stains were found on the clothing of one of the suspects. The lane marked E Blood is the DNA from those stains. The lanes marked V-1 and V-2 are DNA from the first and second victims, respectively; the lanes marked S-1 and S-2 are DNA from the first and second suspects, respectively. The other lanes are for molecular sizing and quality control purposes. Note that E blood matches Victim 1's blood, a result confirmed in the case by matching on 10 loci.

genes, not just markers). Consequently, they are more likely to be subject to mating and natural selection and may not conform to the population-genetics assumptions used in evaluating the significance of a match. (For a discussion of such assumptions, see Section 3.1.2.)

On balance, PCR methods are definitely superior to VNTR methods and have largely replaced them in forensic analyses.

Questions

1. If there are 20 possible alleles at a locus, how many distinguishable homozygous and heterozygous genotypes are possible at that locus?
2. If four similar loci are considered, how many possible distinguishable genotypes are there?

Source

National Research Council, *The Evaluation of Forensic DNA Evidence* (1996).

Notes

Works on DNA profiling abound. Besides the NRC Report cited above, see, e.g., Kaye & Sensabaugh, *Reference Guide on DNA Evidence*, in Federal Judicial Center, *Reference Manual on Scientific Evidence*, (3rd ed. 2011).

2.1.2 Weighted voting

A nine-person board of supervisors is composed of one representative from each of nine towns. Eight of the towns have approximately the same population, but one larger town has three times the population of each of the others. To comply with the one-man, one-vote doctrine of the federal constitution, the larger town's supervisor has been given 3 votes, so that there is a total of 11 votes.

In *Iannucci v. Board of Supervisors*, 20 N.Y.2d 244, 282 N.Y.S.2d 502 (1967), the New York Court of Appeals held that, "The principle of one man one vote is violated, however, when the power of a representative to affect the passage of legislation by his vote, rather than by influencing his colleagues, does not roughly correspond to the proportion of the population in his constituency. . . Ideally, in any weighted voting plan, it should be mathematically possible for every member of the legislative body to cast the decisive vote on legislation in the same ratio which the population of his constituency bears to the total population." *Id.*, 20 N.Y.2d at 252, 282 N.Y.S.2d at 508.

Test whether the weighted-voting scheme complies with the *Iannucci* standard by computing the number of ways the supervisors can vote so as to permit the larger town, on the one hand, or a smaller town, on the other, to cast a decisive vote. A decisive vote may be defined as a vote which, when added to the tally, could change the result. Measures are carried by a majority; a tie defeats a measure. Note that the larger town can cast a decisive vote if the other eight towns are evenly split (4-4) or if they are divided 5-3, either for or against a measure. A smaller town can cast a decisive vote if the larger town is joined by two smaller towns, again either for or against a measure.

Questions

1. Counting affirmative and negative votes as separate ways, in how many ways can the eight small-town supervisors vote on a measure so that the larger-town supervisor has a deciding vote?
2. Using the same counting convention, in how many ways can the larger-town supervisor and seven small-town supervisors vote on a measure so that the eighth small-town supervisor has a deciding vote?

3. Does the ratio of the two results indicate that by the Iannucci standard the larger town has been correctly compensated for its larger size?
4. Using Banzhaf's theory and Stirling's formula, as applied at the voter level, show that a voter in the larger town has $1/\sqrt{3}$ the voting power of a voter in one of the smaller towns.

Notes

The measure of voting power referred to in the problem was first suggested in Banzhaf, *Weighted Voting Doesn't Work: A Mathematical Analysis*, 19 Rutgers L. Rev. (1965). The *Iannucci* case was the first to adopt Banzhaf's theory.

At the voter level, the Banzhaf theory is in effect a square root rule: the probability of breaking a tie in a district with N votes is approximately $[2/(\pi N)]^{1/2}$ by Stirling's formula (see Section 2.1 at p. 48). At this level, the probability of breaking a tie seems too remote to influence a voter. The U.S. Supreme Court has so held. See *Whitcomb v. Chavis*, 403 U.S. 124 (1971); *Board of Estimate v. Morris*, 489 U.S. 688 (1989).

However, voters are probably influenced by the size of the anticipated plurality and may feel that individual votes are more important if the plurality will be smaller rather than larger. Under quite general conditions—particularly in close elections—the plurality will be approximately normally distributed, with an ascertainable probability P of exceeding a given value d in a district of a given size (in which everyone votes). If a second district is k times larger than the first, it follows from the assumption of normality that the plurality event in the larger district with the same probability P is not dk, but $d\sqrt{k}$. Thus, the square root relation derived from the tiebreaking situation has a more general justification in terms of pluralities than Banzhaf had given to it.

2.1.3 Was the bidding rigged?

Every six months the U.S. Maritime Administration issued requests for sealed bids for sale of its obsolete ships. Seven firms in the ship dismantling business ostensibly compete in bidding for these ships. In the last nine requests for bids, five firms submitted lowest bids once each, and two submitted lowest bids twice each. The firms deny collusion, arguing that the work is standard and the bidders have the same cost structure.

Questions

1. Assuming that each firm has the same probability of success on a bid and that success on one bid is independent of success on another, use a simple probability model to argue that the distribution of successes suggests collusive allocation.
2. Is the observed distribution of successes the most probable?
3. Are the assumptions of the model reasonable?

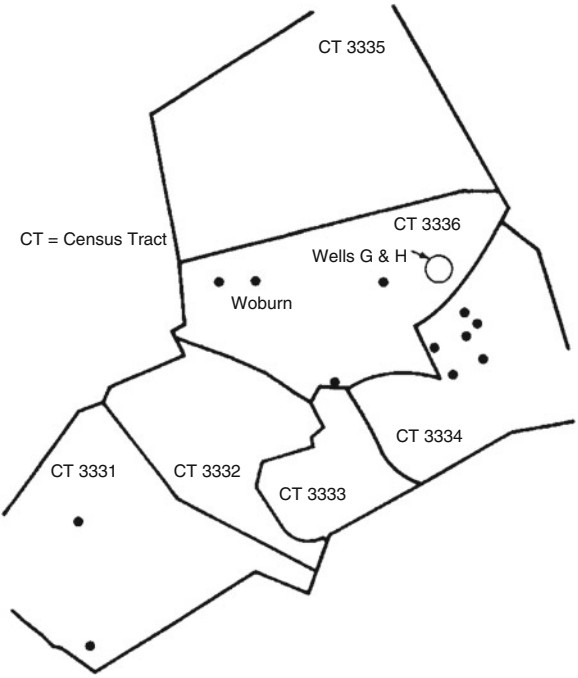

Fig. 2.1.4. Residences of childhood leukemia patients at time of diagnosis, Woburn, Massachusetts, 1969–1979

2.1.4 A cluster of leukemia

Between 1969 and 1979 there were 12 cases of childhood leukemia in Woburn, Massachusetts, when only 5.3 were expected on the basis of national rates. There are six approximately equal census tracts in Woburn. Six cases were clustered in census tract 3334. Lawsuits were brought on the theory that the leukemia had been caused by contaminated well water, although census tract 3334 did not receive the largest amount of this water. See Fig. 2.1.4. For additional facts, see Section 11.2.2.

Question

In principle, how would you calculate the probability that a cluster of six or more leukemia cases would occur as a matter of chance in one district, assuming that the probability of leukemia is the same in all tracts and that 12 cases occurred?

Source

See Section 11.2.2.

2.1.5 Measuring market concentration

The U.S. Department of Justice guidelines to its enforcement policy for horizontal mergers relate enforcement to a measure of market concentration known as the Herfindahl Index (HHI). The HHI is an index of concentration calculated by

squaring the percentage market share of each firm in the market and then adding those squares. To interpret the HHI, note that if all firms have equal market shares, the index divided into 100^2, or 10,000, is the number of such firms. Thus, when firms have unequal market shares, dividing the HHI into 10,000 gives the number of equal firms in a market of equivalent concentration. The index divided by 10,000 has another interpretation: it gives the probability that two customers selected at random with replacement would be customers of the same firm. This probability is also called the "repeat rate."

Use of the HHI depends on the degree of concentration in the market and the effect of the merger. Here are two examples from the guidelines.

- Where the post-merger market is "unconcentrated", which is defined as a market with an HHI that is below 1,000. In such an "unconcentrated" market, the Department would be unlikely to challenge any merger. An index of 1,000 indicates the level of concentration that exists, for instance, in a market shared equally by 10 firms.
- Where the post-merger market is "moderately concentrated," with an HHI between 1,000 and 1,800. A challenge would still be unlikely, provided the merger increases the HHI by less than 100 points. If the merger increased the index by more than 100 points, it would "potentially raise significant competitive concerns," depending on the presence or absence of other relevant factors specified in the guidelines.

Questions

1. If there are N equal customers in a market, and the i^{th} firm has a_i customers, write an expression for the number of different ways the customers may be distributed among firms without changing market shares.
2. Write the above expression in terms of market shares, with the i^{th} firm having $1/n_i$ share of market, so that $1/n_i = a_i/N$.
3. Use the approximation[1] $n! \approx n^n/e^n$ to eliminate the number of customers, leaving only market shares, and take the N^{th} root to make it equal the number of firms in the market when each firm has the same percentage share of market. Is this "entropy" index a plausible measure of concentration?
4. Compare the "entropy" measure with the HHI for the situation in which four large firms share equally 80% of the market, with the remaining 20% shared first by 10, and then by 20, firms. Which measure shows the more concentrated market in terms of an equivalent number of equal firms as the number of small firms increases?

Source

Merger Guidelines of Department of Justice, Trade Reg. Rep. (CCH) ¶13,104 (1997); Finkelstein & Friedberg, *The Application of an Entropy Theory of*

[1] This is a crude form of Stirling's formula (see Section 2.1 at p. 48).

Concentration to the Clayton Act, 76 Yale L.J. 677, 696-97 (1967), reprinted in Finkelstein, *Quantitative Methods in Law*, ch. 5 (1978).

2.2 Fluctuation theory

A sequence of equal deposits and withdrawals to and from an account can be likened to a series of coin tosses, with a deposit corresponding to heads and a withdrawal corresponding to tails. The amount on deposit at any time is represented by the excess of heads over tails. This process can be represented as a polygonal line that starts at the origin, with vertices at abscissas 0, 1, 2... representing the aggregate number of tosses and ordinates at each vertex equal to the net lead of heads over tails (or tails over heads). The behavior of these paths is the subject of fluctuation theory, which deals with problems such as leads in coin tossings and random walks. The results of this theory are frequently counterintuitive because people tend to believe that a random process that is balanced in probability will remain closely balanced in result throughout a sequence. The more correct view, reflected in various theorems, is that leads in one direction, even over long stretches, are surprisingly likely to occur. This behavior is illustrated by a result of 10,000 tosses of a fair coin, shown in Fig. 2.2a. The result is that, in a coin-tossing game, large leads (on the order of the square root of the number of tosses) in favor of one

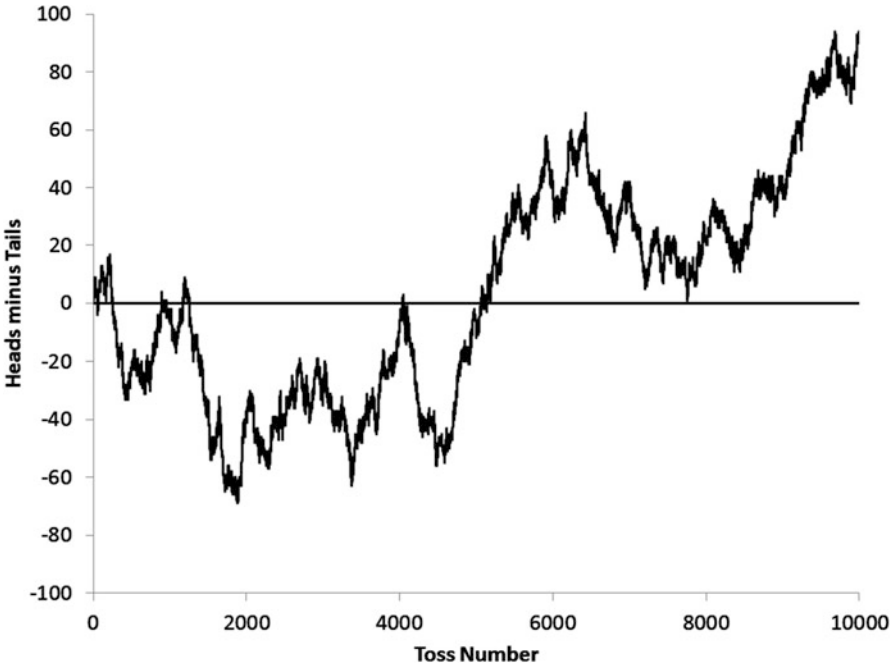

Fig. 2.2a. Record of 10,000 tosses of an ideal coin

party or the other, and the persistence of a lead in favor of one party or the other, are more probable than one might imagine. W. Feller gives the following example: If a fair coin is tossed once a second for 365 days a year, "in one out of 20 cases the more fortunate player will be in the lead for more than 364 days and 10 hours. Few people believe that a perfect coin will produce preposterous sequences in which no change in lead occurs for millions of trials in succession, and yet this is what a good coin will do rather regularly." Feller, *An Introduction to Probability Theory and Its Applications* 81 (3d ed. 1968).

Two basic points about paths representing fluctuations are noted here. First, the number of possible paths representing n coin tosses in which there are a heads and b tails is $\binom{n}{a} = \binom{n}{b}$. Second, the number of paths that start at ordinate A, touch or cross the line corresponding to ordinate B, and end at ordinate C (B lying below A and C) is equal to the number of paths that start at A and end at a point C' that is as far below B as B is below C. This is the "reflection" principle of D. André. See Fig. 2.2b.

Assume that n is the total number of deposits and withdrawals, not necessarily equal in number, but each deposit or withdrawal equal in amount; w is the size of each single deposit or withdrawal; and aw is the amount by which the closing

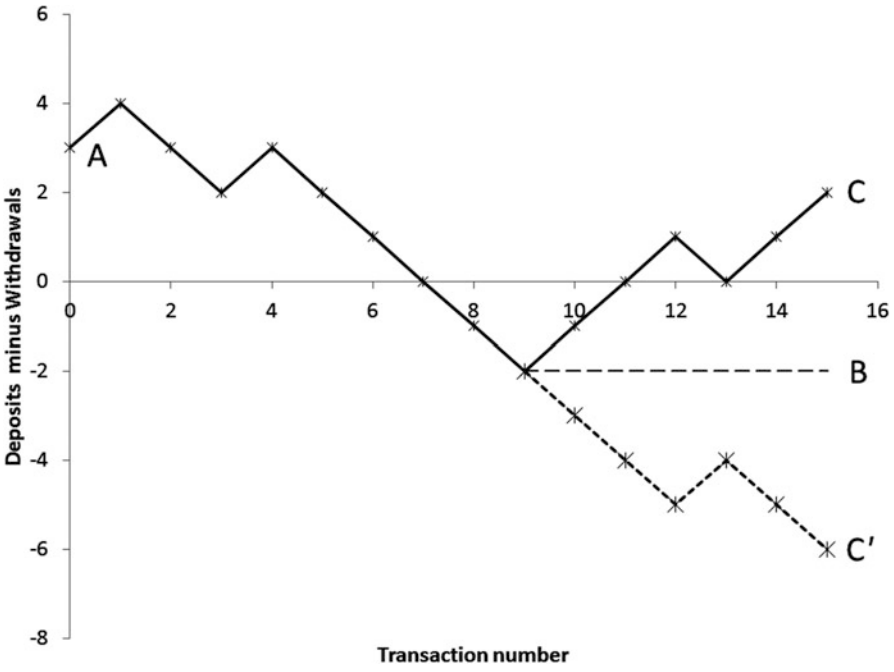

Fig. 2.2b. The André Reflection Principle

balance is less than the opening balance (a is a negative integer if the closing balance exceeds the opening balance). The probability, P, that at some point the opening balance was decreased by kw or more is: $P = exp[-2k(k-a)]$, approximately, for $k \geq a$. (The term exp[x] is alternate notation for the exponential function e^x). In the special case in which the opening and closing amounts are equal and $w = 1$, the expected maximum dissipation is $0.627\sqrt{n}$ and the median maximum dissipation is $0.589\sqrt{n}$.

If deposits and withdrawals are not equal in amount, no closed form approximation is known for the probability of any given dissipation of the account. In such cases, computer simulation must be used instead. See Finkelstein and Robbins, *infra*, for an example.

Further Reading

Feller, *An Introduction to Probability Theory and Its Applications*, ch. 14 (3d ed. 1968).

2.2.1 Tracing funds for constructive trusts

A person who defrauds another of money may be declared a constructive trustee of the funds for the defrauded party—if that party can trace the funds. There are various rules for tracing. If the funds are deposited in a bank account, withdrawals are deemed made first from non-trust funds (if they cannot be traced into new assets), but to the extent the account is reduced at any point a subsequent restoration by fresh deposits will not restore the trust. The sequence of deposits and withdrawals may therefore be critical. Suppose that there are $10 in trust funds in an account, and that during a day there are 50 one-dollar deposits and 50 one-dollar withdrawals. As in the case of most banks, the order of deposits and withdrawals within the day cannot be determined.

Questions

1. What is the probability that at some point during the day the trust funds were reduced to 0?
2. What is the expected maximum reduction in the trust?
3. If this situation persists for 10 days (i.e., 50 repeated deposits and withdrawals leaving the account even at the end of the day), is it reasonable to presume that the fund has been reduced to 0 at some point during the period?

Further Reading

Finkelstein & Robbins, *A Probabilistic Approach to Tracing Presumptions in the Law of Restitution*, 24 Jurimetrics 65 (1983).

http://www.springer.com/978-1-4419-5984-3

Statistics for Lawyers

Finkelstein, M.O.; Levin, B.
2015, XXXII, 657 p. 56 illus., Hardcover
ISBN: 978-1-4419-5984-3