The Little End of the Stick

(a) If a stick is broken in two at random, what is the average length of the smaller piece?

(b) (For calculus students.) What is the average ratio of the smaller length to the larger?

Solution for The Little End of the Stick

(a) Breaking "at random" means that all points of the stick are equally likely as a breaking point (uniform distribution). The breaking point is just as likely to be in the left half as the right half. If it is in the left half, the smaller piece is on the left; and its average size is half of that half, or one-fourth the length of the stick. The same sort of argument applies when the break is in the right half of the stick, and so the answer is one-fourth of the length.

(b) We might suppose that the point fell in the right-hand half. Then $(1 - x)/x$ is the fraction if the stick is of unit length. Since x is evenly distributed from $\frac{1}{2}$ to 1, the average value, instead of the intuitive $\frac{1}{3}$, is

$$2\int_{\frac{1}{2}}^{1} \frac{1-x}{x}\,dx = 2\int_{\frac{1}{2}}^{1}\left(\frac{1}{x} - 1\right)dx$$
$$= 2\log_e 2 - 1 \approx 0.386.$$

Average Number of Matches

The following are two versions of the matching problem:

(a) From a shuffled deck, cards are laid out on a table one at a time, face up from left to right, and then another deck is laid out so that each of its cards is beneath a card of the first deck. What is the average number of matches of the card above and the card below in repetitions of this experiment?

(b) A typist types letters and envelopes to n different persons. The letters are randomly put into the envelopes. On the average, how many letters are put into their own envelopes?

Solution for Average Number of Matches

Let us discuss this problem for a deck of cards. Given 52 cards in a deck, each card has 1 chance in 52 of matching its paired card. With 52 opportunities for a match, the expected number of matches is $52(\frac{1}{52}) = 1$; thus, on the average you get 1 match. Had the deck consisted of n distinct cards, the expected number of matches would still be 1 because $n(1/n) = 1$. The result leans on the theorem that the mean of a sum is the sum of the means.

More formally, each pair of cards can be thought of as having associated with it a random variable X_i that takes the value 1 when there is a match and the value 0 when there is not. Then

$$E(X_i) = 1\left(\frac{1}{n}\right) + 0\left(1 - \frac{1}{n}\right) = \frac{1}{n}.$$

The Clumsy Chemist

In a laboratory, each of a handful of thin 9-inch glass rods had one tip marked with a blue dot and the other with a red. When the laboratory assistant tripped and dropped them onto the concrete floor, many broke into three pieces. For these, what was the average length of the fragment with the blue dot?

Solution for The Clumsy Chemist

Assuming these rods broke at random, the principle of symmetry says that each fragment—blue-dotted, middle, and red-dotted segment—would have the same distribution and the same mean. Since the means have to add to 9 inches, the blue-dotted segments average about 3 inches.

The First Ace

Shuffle an ordinary deck of 52 playing cards containing four aces. Then turn up cards from the top until the first ace appears. On the average, how many cards are required to produce the first ace?

•

Solution for The First Ace

Assume that the principle of symmetry holds for discrete as well as continuous events. The four aces divide the pack into 5 segments of size from 0 to 48 cards. If two aces are side by side, we say the segment between them is of length 0. If the first card is an ace, the segment before it is of length zero, and similarly for the segment following an ace that is a last card. The principle of symmetry says the 5 segments should average $\frac{48}{5} = 9.6$ cards. The next card is the ace itself, so it is the 10.6th card on the average.

The Little End of the Stick

(a) If a stick is broken in two at random, what is the average length of the smaller piece?

(b) (For calculus students.) What is the average ratio of the smaller length to the larger?

Solution for The Little End of the Stick

(a) Breaking "at random" means that all points of the stick are equally likely as a breaking point (uniform distribution). The breaking point is just as likely to be in the left half as the right half. If it is in the left half, the smaller piece is on the left; and its average size is half of that half, or one-fourth the length of the stick. The same sort of argument applies when the break is in the right half of the stick, and so the answer is one-fourth of the length.

(b) We might suppose that the point fell in the right-hand half. Then $(1 - x)/x$ is the fraction if the stick is of unit length. Since x is evenly distributed from $\frac{1}{2}$ to 1, the average value, instead of the intuitive $\frac{1}{3}$, is

$$2\int_{\frac{1}{2}}^{1} \frac{1 - x}{x}\,dx = 2\int_{\frac{1}{2}}^{1} \left(\frac{1}{x} - 1\right) dx$$
$$= 2\log_e 2 - 1 \approx 0.386.$$

Winning an Unfair Game

A game consists of a sequence of plays; on each play either you or your opponent scores a point, you with probability p (less than $\frac{1}{2}$), he with probability $1 - p$. The number of plays is to be even—2 or 4 or 6 and so on. To win the game you must get *more than* half the points. You know p, say 0.45, and you get a prize if you win. You get to choose in advance the number of plays. How many do you choose?

Solution for Winning an Unfair Game

Don't balk just because the game is unfair; after all you are the only one eligible for a prize. Let us call you player A and your opponent player B. Let the total number of plays be $N = 2n$. On a given play, your chance of winning a point is p, your opponent's $q = 1 - p$.

At first blush, most people notice that the game is unfair and therefore that, as N increases, the expected value of the difference (A's points $-$ B's points) grows more and more negative. They conclude that A should play as little as he can and still win—that is, two plays.

Had an odd number of plays been allowed, this reasoning based on expected values would have led to the correct answer, and A should choose only one play. With an even number of plays, two opposing effects are at work: (1) the bias in favor of B, and (2) the redistribution of the probability in the middle term of the binomial distribution (the probability of a tie) as the number of plays increases.

Consider, for a moment, a fair game ($p = \frac{1}{2}$). Then the larger N, the larger A's chance to win because as $2n$ increases, the probability of a tie tends to zero, and the limiting value of A's chance to win is $\frac{1}{2}$. For $N = 2, 4, 6$, his probabilities are $\frac{1}{4}, \frac{5}{16}, \frac{22}{64}$. Continuity suggests that for p slightly less than $\frac{1}{2}$, A should choose a large but finite number of plays. But if p is small, $N = 2$ should be optimum for A. It turns out that for $p < \frac{1}{3}$, $N = 2$ is optimum.

Your probability of winning in a game of $2n$ trials is the sum of the probabilities of getting $n + 1, n + 2, \ldots, 2n$ points, a sum given by

$$P_{2n} = \sum_{x=n+1}^{2n} \binom{2n}{x} p^x q^{2n-x}.$$

In a game of $2n + 2$ plays, the probability of winning at least $n + 2$ points and the game is

$$P_{2n+2} = \sum_{x=n+2}^{2n+2} \binom{2n+2}{x} p^x q^{2n+2-x}.$$

A game composed of $2n + 2$ plays can be regarded as having been created by adding two plays to a game of $2n$ plays. Unless player A has won either n or $n + 1$ times in the $2n$ game, his status as a winner or loser cannot differ in the $2n + 2$ game from that in the $2n$ game.

Except for these two possibilities, P_{2n+2} would be identical with P_{2n}. These exceptions are: (1) having $n + 1$ successes in the first $2n$ plays, A loses the next two, thus reducing his probability of winning in the $2n + 2$ game by

$$q^2\binom{2n}{n+1} p^{n+1}q^{n-1};$$

or (2) having won n plays in the $2n$ game, he wins the next two, increasing his probability by

$$p^2\binom{2n}{n} p^n q^n.$$

If $N = 2n$ is the optimum value, then both $P_{N-2} \leq P_N$ and $P_N \geq P_{N+2}$ must hold. The results of the previous paragraph imply that these inequalities are equivalent to the following two inequalities:

(1)
$$q^2\binom{2n-2}{n}p^nq^{n-2} \le p^2\binom{2n-2}{n-1}p^{n-1}q^{n-1},$$
$$q^2\binom{2n}{n+1}p^{n+1}q^{n-1} \ge p^2\binom{2n}{n}p^nq^n.$$

After some simplifications, which you may wish to verify (we exclude the trivial case $p = 0$), we reduce inequalities (1) to

(2) $$(n-1)q \le np; \qquad nq \ge (n+1)p.$$

These inequalities yield, after a little algebra, the condition

(3) $$\frac{1}{1-2p} - 1 \le 2n \le \frac{1}{1-2p} + 1.$$

Thus unless $1/(1-2p)$ is an odd integer, N is uniquely determined as the nearest even integer to $1/(1-2p)$. When $1/(1-2p)$ is an odd integer, both adjacent even integers give the same optimum probability. And we can incidentally prove that when $1/(1-2p) = 2n + 1$, $P_{2n} = P_{2n+2}$.

Consequently for $p = 0.45$, we have $1/(1-0.9) = 10$ as the optimum number of plays to choose.

This material is abbreviated from "Optimal length of play for a binomial game," *Mathematics Teacher*, Vol. 54, 1961, pp. 411–412.

P. G. Fox originally alluded to a result which gives rise to this game in "A primer for chumps," which appeared in the *Saturday Evening Post*, November 21, 1959, and discussed the idea further in private correspondence arising from that article in a note entitled "A curiosity in the binomial expansion—and a lesson in logic." I am indebted to Clayton Rawson and John Scarne for alerting me to Fox's paper and to Fox for helpful correspondence.

Doubling Your Accuracy

An unbiased instrument for measuring distances makes random errors whose distribution has standard deviation σ. You are allowed two measurements all told to estimate the lengths of two cylindrical rods, one clearly longer than the other. Can you do better than to take one measurement on each rod? (An unbiased instrument is one that on the average gives the true measure)

Solution for Doubling Your Accuracy

Yes. Let A be the true length of the longer one, B that for the shorter. You could lay them side by side and measure their difference in length, $A - B$, and then lay them end to end and measure the sum of their lengths, $A + B$.

Let D be your measurement of $A - B$, S of $A + B$. Then an estimate of A is $\frac{1}{2}(D + S)$, and of B is $\frac{1}{2}(S - D)$. Now $D = A - B + d$, where d is a random error, and $S = A + B + s$, where s is a random error. Consequently,

$$\tfrac{1}{2}(D + S) = \tfrac{1}{2}(A - B + A + B + d + s) = A + \tfrac{1}{2}(d + s).$$

On the average, the error $\frac{1}{2}(d + s)$ is zero because both d and s have mean zero. The variance of the estimate of A is the variance of $\frac{1}{2}(d + s)$, which is $\frac{1}{4}(\sigma_d^2 + \sigma_s^2) = \frac{1}{4}(\sigma^2 + \sigma^2) = \frac{1}{2}\sigma^2$. This value is identical with the variance for the average of a sample of two independent measurements. Thus both our measurements have contributed their full value to measuring A. In the same manner you can show that the variance of the estimate of B is also $\frac{1}{2}\sigma^2$. Consequently, taking two measurements, one on the difference and one on the sum, gives estimates whose precision is equivalent to that where 4 measurements are used, two on each rod separately.

To achieve such good results, we must be able to align the ends of the rods perfectly. If we cannot, instead of two alignments for each measurement, we have three. If each alignment contributes an independent error with standard deviation $\sigma/\sqrt{2}$, then one measurement of the sum or difference has standard deviation $\sigma\sqrt{3/2}$. Then the variance of our estimate of A would be

$$\tfrac{1}{4}[\tfrac{3}{2}\sigma^2 + \tfrac{3}{2}\sigma^2] = \tfrac{3}{4}\sigma^2 = \sigma^2/\tfrac{4}{3}.$$

Under these assumptions our precision is only as good as $1\frac{1}{3}$ independent measurements instead of 2, but still better than a single direct measurement.

We may rationalize the assignment of standard deviation $\sigma/\sqrt{2}$ to each alignment by thinking of s (or d) as composed of the sum of two independent unbiased measurement errors, each having variance $\sigma^2/2$. Then the sum of the component errors would produce the variance assumed earlier of σ^2. When we also assign the third alignment the variance $\sigma^2/2$, our model is completed.

You can read about variances of means and sums of independent variables in PWSA, pp. 318–322.

Random Quadratic Equations

What is the probability that the quadratic equation

$$x^2 + 2bx + c = 0$$

has real roots?

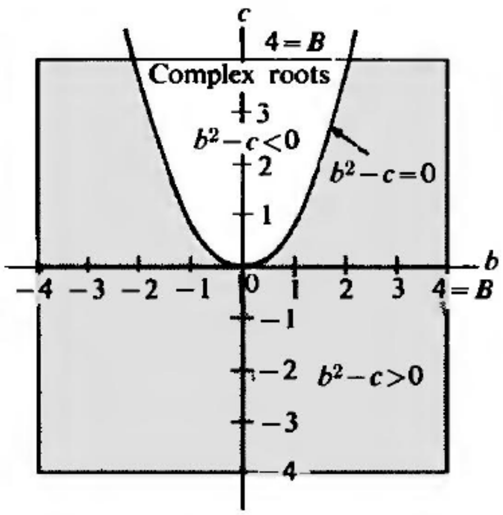
Solution for Random Quadratic Equations

To make this question meaningful, we shall suppose that the point (b, c) is randomly chosen from a uniform distribution over a large square centered at the origin, with side $2B$ (see the figure). We solve the problem for a given value of B; then we let B grow large so that b and c can take any values.

For the quadratic to have real roots, we must have

$$b^2 - c \geq 0.$$

In the figure, we plot the parabola $b^2 = c$ and show the regions in the square, for $B = 4$, where the original equation has real roots.



Regions yielding complex and real roots. Shaded region gives real roots; unshaded, complex.

It is an easy exercise in calculus to show that the area of the unshaded region is $\frac{4}{3}B^{\frac{3}{2}}$ (for $B \geq 1$), and, of course, the whole square has area $4B^2$. Consequently, the probability of getting complex roots is $1/(3\sqrt{B})$. When $B = 4$, the result is $\frac{1}{6}$. As B grows large, $1/\sqrt{B}$ tends to zero, and so the probability that the roots are real tends to 1!

I should warn you that the problem we have just solved is not identical with that for $ax^2 + 2bx + c = 0$. You might think you could divide through by a. You can, but if the old coefficients a, b, c were independently uniformly distributed over a cube, then b/a and c/a are neither uniformly nor independently distributed.

Buffon's Needle with Horizontal and Vertical Rulings

Suppose we toss a needle of length $2l$ (less than 1) on a grid with both horizontal and vertical rulings spaced one unit apart. What is the mean number of lines the needle crosses? (I have dropped $2a$ for the spacing because we might as well think of the length of the needle as measured in units of spacing)

Solution for Buffon's Needle with Horizontal and Vertical Rulings

The mean number of vertical rulings crossed is the same as the probability of crossing a vertical ruling. From the previous problem (with $a = \frac{1}{2}$), it is $4l/\pi$. The mean number of horizontal rulings crossed must also be $4l/\pi$ because it is the same problem if you turn your head through 90°. The mean of a sum is the sum of the means, and so the mean total number of crossings is $8l/\pi$.

If the needle is of length 1, the mean number of crosses is $4/\pi \approx 1.27$.

Up to now we have worked with needles shorter than the spacing, what about longer needles?

Long Needles

In the previous problem let the needle be of arbitrary length, then what is the mean number of crosses?

Solution for Long Needles

Let the needle be divided into n pieces of equal lengths so that all are less than 1. If we toss each of these little needles at random, each will have a mean number of crosses obtained from the previous problem. The mean of the sum is the sum of the means, and so their expected number of crosses is 4(original length)/π. The fact that the needle was not tossed as a rigid structure does not matter to the mean.

For purposes of estimating π the experiment of tossing a long needle on a grid of squares represents a substantial improvement over the original Buffon problem. Why not get some graph paper and try it? I used a toothpick and graph paper ruled in half-inch squares. The toothpick was 5.2 half-inches long. I decided on 10 tosses, got 8, 6, 7, 6, 5, 6, 7, 5, 5, 7 crosses, totaling 62. My estimate for π is $4(5.2)/(62/10) \approx 3.35$, instead of 3.14. A friend of mine also made 10 tries, producing 67 crosses, yielding the estimate 3.10.

Molina's Urns

Two urns contain the same total numbers of balls, some blacks and some whites in each. From each urn are drawn n (≥ 3) balls with replacement. Find the number of drawings and the composition of the two urns so that the probability that all white balls are drawn from the first urn is equal to the probability that the drawing from the second is either all whites or all blacks.

Discussion for Molina's Urns

E. C. Molina invented this problem to display Fermat's famous conjecture in number theory as a probability problem.

Let z be the number of white balls in the first urn, x the number of whites and y the number of blacks in the second. Then we want to find integers n, x, y, and z so that

$$\left(\frac{z}{x+y}\right)^n = \left(\frac{x}{x+y}\right)^n + \left(\frac{y}{x+y}\right)^n,$$

or

$$z^n = x^n + y^n.$$

Although, for many values of n, it is known that this equation cannot be satisfied, it is not known whether it is impossible for all values of $n \geq 3$. But it is known to be impossible for $n < 2000$.

The Sock Drawer

A drawer contains red socks and black socks. When two socks are drawn at random, the probability that both are red is $\frac{1}{2}$ (a) How small can the number of socks in the drawer be? (b) How small if the number of black socks is even?

Just to set the pattern, let us do a numerical example first. Suppose there were 5 red and 2 black socks; then the probability of the first sock's being red would be $5/(5 + 2)$. If the first were red, the probability of the second's being red would be $4/(4 + 2)$, because one red sock has already been removed. The product of these two numbers is the probability that both socks are red:

$$\frac{5}{5+2} \times \frac{4}{4+2} = \frac{5(4)}{7(6)} = \frac{10}{21}.$$

This result is close to $\frac{1}{2}$, but we need exactly $\frac{1}{2}$. Now let us go at the problem algebraically.

Let there be r red and b black socks. The probability of the first sock's being red is $r/(r+b)$; and if the first sock is red, the probability of the second's being red now that a red has been removed is $(r-1)/(r+b-1)$. Then we require the probability that both are red to be $\frac{1}{2}$, or

$$\frac{r}{r+b} \times \frac{r-1}{r+b-1} = \frac{1}{2}.$$

One could just start with $b = 1$ and try successive values of r, then go to $b = 2$ and try again, and so on. That would get the answers quickly. Or we could play along with a little more mathematics. Notice that

$$\frac{r}{r+b} > \frac{r-1}{r+b-1}, \qquad \text{for } b > 0.$$

Therefore we can create the inequalities

$$\left(\frac{r}{r+b}\right)^2 > \frac{1}{2} > \left(\frac{r-1}{r+b-1}\right)^2.$$

Taking square roots, we have, for $r > 1$,

$$\frac{r}{r+b} > \frac{1}{\sqrt{2}} > \frac{r-1}{r+b-1}.$$

From the first inequality we get

$$r > \frac{1}{\sqrt{2}}(r+b)$$

or

$$r > \frac{1}{\sqrt{2}-1}b = (\sqrt{2}+1)b.$$

From the second we get

$$(\sqrt{2}+1)b > r-1$$

or all told

$$(\sqrt{2}+1)b + 1 > r > (\sqrt{2}+1)b.$$

For $b = 1$, r must be greater than 2.414 and less than 3.414, and so the candidate is $r = 3$. For $r = 3$, $b = 1$, we get

$$P(2 \text{ red socks}) = \tfrac{3}{4} \cdot \tfrac{2}{3} = \tfrac{1}{2}.$$

And so the smallest number of socks is 4.

Beyond this we investigate even values of b.

b	r is between	eligible r	P(2 red socks)
2	5.8, 4.8	5	$\frac{5(4)}{7(6)} \neq \frac{1}{2}$
4	10.7, 9.7	10	$\frac{10(9)}{14(13)} \neq \frac{1}{2}$
6	15.5, 14.5	15	$\frac{15(14)}{21(20)} = \frac{1}{2}$

And so 21 socks is the smallest number when b is even. If we were to go on and ask for further values of r and b so that the probability of two red socks is $\frac{1}{2}$, we would be wise to appreciate that this is a problem in the theory of numbers. It happens to lead to a famous result in Diophantine Analysis obtained from Pell's equation.* Try $r = 85$, $b = 35$.

*See for example, W. J. LeVeque, *Elementary theory of numbers*, Addison-Wesley, Reading, Mass., 1962, p. 111.

Successive Wins

To encourage Elmer's promising tennis career, his father offers him a prize if he wins (at least) two tennis sets in a row in a three-set series to be played with his father and the club champion alternately: father-champion-father or champion-father-champion, according to Elmer's choice. The champion is a better player than Elmer's father. Which series should Elmer choose?

Solution for Successive Wins

Since the champion plays better than the father, it seems reasonable that fewer sets should be played with the champion. On the other hand, the middle set is the key one, because Elmer cannot have two wins in a row without winning the middle one. Let C stand for champion, F for father, and W and L for a win and a loss by Elmer. Let f be the probability of Elmer's winning any set from his father, c the corresponding probability of winning from the champion. The table shows the only possible prize-winning sequences together with their probabilities, given independence between sets, for the two choices.

Set with:	Father first F	C	F	Probability	Champion first C	F	C	Probability
	W	W	W	fcf	W	W	W	cfc
	W	W	L	$fc(1-f)$	W	W	L	$cf(1-c)$
	L	W	W	$(1-f)cf$	L	W	W	$(1-c)fc$
Totals				$fc(2-f)$				$fc(2-c)$

Since Elmer is more likely to best his father than to best the champion, f is larger than c, and $2-f$ is smaller than $2-c$, and so Elmer should choose CFC. For example, for $f = 0.8$, $c = 0.4$, the chance of winning the prize with FCF is 0.384, that for CFC is 0.512. Thus the importance of winning the middle game outweighs the disadvantage of playing the champion twice.

Many of us have a tendency to suppose that the higher the expected number of successes, the higher the probability of winning a prize, and often this supposition is useful. But occasionally a problem has special conditions that destroy this reasoning by analogy. In our problem the expected number of wins under CFC is $2c + f$, which is less than the expected number of wins for FCF, $2f + c$. In our example with $f = 0.8$ and $c = 0.4$, these means are 1.6 and 2.0 in that order. This opposition of answers gives the problem its flavor. The idea of independent events is explained in PWSA, pp. 81–84.

The Flippant Juror

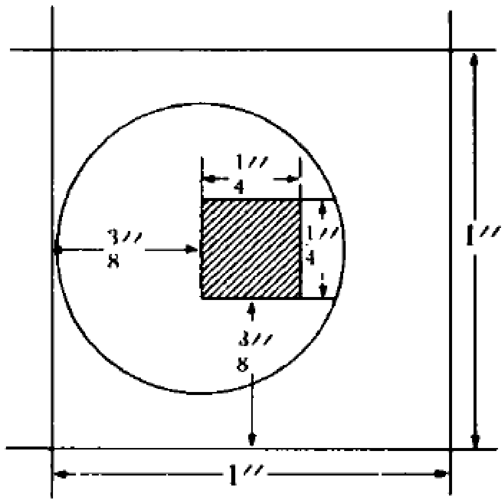
A three-man jury has two members each of whom independently has probability p of making the correct decision and a third member who flips a coin for each decision (majority rules). A one-man jury has probability p of making the correct decision. Which jury has the better probability of making the correct decision?

Solution for The Flippant Juror

The two juries have the same chance of a correct decision. In the three-man jury, the two serious jurors agree on the correct decision in the fraction $p \times p = p^2$ of the cases, and for these cases the vote of the joker with the coin does not matter. In the other correct decisions by the three-man jury, the serious jurors vote oppositely, and the joker votes with the "correct" juror. The chance that the serious jurors split is $p(1-p) + (1-p)p$ or $2p(1-p)$. Halve this because the coin favors the correct side half the time. Finally, the total probability of a correct decision by the three-man jury is $p^2 + p(1-p) = p^2 + p - p^2 = p$, which is identical with the probability given for the one-man jury.

Coin in Square

In a common carnival game a player tosses a penny from a distance of about 5 feet onto the surface of a table ruled in 1-inch squares. If the penny ($\frac{3}{4}$ inch in diameter) falls entirely inside a square, the player receives 5 cents but does not get his penny back; otherwise he loses his penny. If the penny lands on the table, what is his chance to win?



Shaded area shows where center of coin must fall for player to win.

Solution for Coin in Square

When we toss the coin onto the table, some positions for the center of the coin are more likely than others, but over a very small square we can regard the probability distribution as uniform. This means that the probability that the center falls into any region of a square is proportional to the area of the region, indeed, is the area of the region divided by the area of the square. Since the coin is $\frac{3}{8}$ inch in radius, its center must not land within $\frac{3}{8}$ inch of any edge if the player is to win. This restriction generates a square of side $\frac{1}{4}$ inch within which the center of the coin must lie for the coin to be in the square. Since the probabilities are proportional to areas, the probability of winning is $(\frac{1}{4})^2 = \frac{1}{16}$. Of course, since there is a chance that the coin falls off the table altogether, the total probability of winning is smaller still. Also the squares can be made smaller by merely thickening the lines. If the lines are $\frac{1}{16}$ inch wide, the winning central area reduces the probability to $(\frac{3}{16})^2 = \frac{9}{256}$ or less than $\frac{1}{28}$.

Craps

The game of craps, played with two dice, is one of America's fastest and most popular gambling games. Calculating the odds associated with it is an instructive exercise.

The rules are these. Only totals for the two dice count The player throws the dice and wins at once if the total for the first throw is 7 or 11, loses at once if it is 2, 3, or 12. Any other throw is called his "point." If the first throw is a point, the player throws the dice repeatedly until he either wins by throwing his point again or loses by throwing 7. What is the player's chance to win?

Solution for Craps

The game is surprisingly close to even, as we shall see, but slightly to the player's disadvantage.

Let us first get the probabilities for the totals on the two dice. Regard the dice as distinguishable, say red and green. Then there are $6 \times 6 = 36$ possible equally likely throws whose totals are shown in the table (next page).

By counting the cells in the table we get the probability distribution of the totals:

Total	2	3	4	5	6	7	8	9	10	11	12
P(total)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Here P means "probability of."

Cell entries give totals for game of craps
Throw of green die

Throw of red die		1	2	3	4	5	6
	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

Thus the probability of a win on the first throw is

$$P(7) + P(11) = \tfrac{6}{36} + \tfrac{2}{36} = \tfrac{8}{36}.$$

The probability of a loss on the first throw is

$$P(2) + P(3) + P(12) = \tfrac{1}{36} + \tfrac{2}{36} + \tfrac{1}{36} = \tfrac{4}{36}.$$

For later throws we need the probability of making the point. Since no throws except either the point or 7 matter, we can compute for each of these the conditional probability of making the point given that it has been thrown initially. Sometimes such an approach is called the method of reduced sample spaces because, although the actual tosses produce the totals 2 through 12, we ignore all but the point and 7.

For example, for four as the point, there are 3 ways to make the point and 6 ways to make a seven, and so the probability of making the point is $3/(3 + 6) = 3/9$.

Similarly, we get the conditional probabilities for the other points and summarize:

$$4: \quad \frac{3}{3+6} = \frac{3}{9} \qquad 8: \quad \frac{5}{5+6} = \frac{5}{11}$$

$$5: \quad \frac{4}{4+6} = \frac{4}{10} \qquad 9: \quad \frac{4}{4+6} = \frac{4}{10}$$

$$6: \quad \frac{5}{5+6} = \frac{5}{11} \qquad 10: \quad \frac{3}{3+6} = \frac{3}{9}$$

Each probability of winning must be weighted by the probability of throwing the point on the initial throw to give the unconditional probability of winning for that point. Then we sum to get for the probability of winning by throwing a point

$$\tfrac{3}{36}(\tfrac{3}{9}) + \tfrac{4}{36}(\tfrac{4}{10}) + \tfrac{5}{36}(\tfrac{5}{11}) + \tfrac{5}{36}(\tfrac{5}{11}) + \tfrac{4}{36}(\tfrac{4}{10}) + \tfrac{3}{36}(\tfrac{3}{9}) \approx 0.27071.$$

To this we add the probability of winning on the first throw, $\frac{6}{36} \approx 0.22222$, to get 0.49293 as the player's probability of winning. His expected loss per unit stake is $0.50707 - 0.49293 = 0.01414$, or 1.41%. I believe that this is the most nearly even of house gambling games that have no strategy. And 1.4% doesn't sound like much, but as I write, the stock of General Motors is selling at 71, and their dividend for the year (before extras) is quoted as \$2, or about 2.8%. So per two plays at craps your loss is at a rate equal to the yearly dividend payout by America's largest corporation.

Some readers may not be satisfied with the conditional probability approach used for points and may wish to see the series summed.

Let the probability of throwing the point be P and let the probability of a toss that does not count be $R(=1 - P - \frac{1}{6})$. The $\frac{1}{6}$ is the probability of throwing 7. The player can win by throwing a number of tosses that do not count and then throwing his point. The probability that he makes his point in the $(r + 1)$st throw (after the initial throw) is R^rP, $r = 0, 1, 2, \ldots$. To get the total probability, we sum over the values of r:

$$P + RP + R^2P + \cdots = P(1 + R + R^2 + \cdots).$$

Summing this infinite geometric series gives

$$\text{Probability of making point} = P/(1 - R).$$

$$\text{Probability of making point} = P/(1 - R).$$

For example, if the point is 4, $P = \frac{3}{36}$, $R = 1 - \frac{3}{36} - \frac{6}{36} = \frac{27}{36}$, $1 - R = \frac{9}{36}$, $P(\text{making the point 4}) = (3/36)/(9/36) = 3/9$, as we got by the simpler approach of reduced sample spaces.

The first time I met this problem, I summed the series and was quite pleased with myself until a few days later the reduced sample space approach occurred to me and left me deflated.

An Experiment in Personal Taste for Money

(a) An urn contains 10 black balls and 10 white balls, identical except for color. You choose "black" or "white." One ball is drawn at random, and if its color matches your choice, you get $10, otherwise nothing. Write down the maximum amount you are willing to pay to play the game. The game will be played just once.

(b) A friend of yours has available many black and many white balls, and he puts black and white balls into the urn to suit himself. You choose "black" or "white." A ball is drawn randomly from this urn. Write down the maximum amount you are willing to pay to play this game. The game will be played just once.

Discussion for An Experiment in Personal Taste for Money

No one can say what amount is appropriate for you to pay for either game. Even though your expected value in the first game is $5, you may not be willing to pay anything near $5 to play it. The loss of $3 or $4 may mean too much to you. Let us suppose you decided to offer 75¢.

What we *can* say is that you should be willing to pay at least as much to play the second game as the first. You can always choose your own color at random by the toss of a coin and thus assure that you have a fifty-fifty chance of being right and therefore an expectation of $5. Furthermore, if you have any information about your friend's preferences, you can take advantage of that to improve your chances.

Most people feel that they would rather play the first game because the conditions of the second seem more vague. I am indebted to Howard Raiffa for this problem, and he tells me that the idea was suggested to him by Daniel Ellsberg.

Silent Cooperation

Two strangers are separately asked to choose one of the positive whole numbers and advised that if they both choose the same number, they both get a prize. If you were one of these people, what number would you choose?

Discussion for Silent Cooperation

I have not met anyone yet who would choose more than a one-digit number; and of these only 1, 3, and 7 have been chosen. Most of my informants choose 1, which seems on the face of it to be the natural choice. But 3 and 7 are popular choices.

Quo Vadis?

Two strangers who have a private recognition signal agree to meet on a certain Thursday at 12 noon in New York City, a town familiar to neither, to discuss an important business deal, but later they discover that they have not chosen a meeting place, and neither can reach the other because both have embarked on trips. If they try nevertheless to meet, where should they go?

Discussion for Quo Vadis?

My daughter when asked this question said enthusiastically "Why, they should meet in the most famous place in New York!" "Fine," I said, "where?" "How would I know that?" she said, "I'm only 9 years old."

Places that come to mind in 1964 are top of the Empire State Building, airports, information desks at railroad stations, Statue of Liberty, Times Square. The Statue of Liberty will be eliminated the moment the strangers find out how hard it is to get there. Airports suffer from distance from town and numerosity. That there are two important railroad stations seems to me to remove them from the competition. That leaves the Empire State Building or Times Square. I would opt for the Empire State Building, because Times Square is getting vaguely large these days. I think their problem would have been easier if they had been meeting in San Francisco or Paris, don't you?

The Prisoner's Dilemma

Three prisoners, *A*, *B*, and *C*, with apparently equally good records have applied for parole. The parole board has decided to release two of the three, and the prisoners know this but not which two. A warder friend of prisoner *A* knows who are to be released. Prisoner *A* realizes that it would be unethical to ask the warder if he, *A*, is to be released, but thinks of asking for the name of *one* prisoner *other than himself* who is to be released He thinks that before he asks, his chances of release are $\frac{2}{3}$. He thinks that if the warder says "*B* will be released," his own chances have now gone down to $\frac{1}{2}$, because either *A* and *B* or *B* and *C* are to be released. And so *A* decides not to reduce his chances by asking. However, *A* is mistaken in his calculations Explain.

Solution for The Prisoner's Dilemma

Of all the problems people write me about, this one brings in the most letters.

The trouble with *A*'s argument is that he has not listed the possible events properly. In technical jargon he does not have the correct sample space. He thinks his experiment has three possible outcomes: the released pairs *AB*, *AC*, *BC* with equal probabilities of $\frac{1}{3}$. From his point of view, that is the correct sample space for the experiment conducted by the parole board given that they are to release two of the three. But *A*'s own experiment adds an event—the response of the warder. The outcomes of his proposed experiment and reasonable probabilities for them are:

1. *A* and *B* released and warder says *B*, probability $\frac{1}{3}$.
2. *A* and *C* released and warder says *C*, probability $\frac{1}{3}$.
3. *B* and *C* released and warder says *B*, probability $\frac{1}{6}$.
4. *B* and *C* released and warder says *C*, probability $\frac{1}{6}$.

If, in response to *A*'s question, the warder says "*B* will be released," then the probability for *A*'s release is the probability from outcome 1 divided by the sum of the probabilities from outcomes 1 and 3. Thus the final probability of *A*'s release is $\frac{1}{3}/(\frac{1}{3} + \frac{1}{6})$, or $\frac{2}{3}$, and mathematics comes round to common sense after all.

Collecting Coupons

Coupons in cereal boxes are numbered 1 to 5, and a set of one of each is required for a prize. With one coupon per box, how many boxes on the average are required to make a complete set?

Solution for Collecting Coupons

We get one of the numbers in the first box. Now the chance of getting a new number from the next box is $\frac{4}{5}$. Using the result of Problem 4, the second new number requires $1/(4/5) = \frac{5}{4}$ boxes. The third new number requires an additional $1/(3/5) = \frac{5}{3}$; the fourth $\frac{5}{2}$, the fifth $\frac{5}{1}$.

Thus the average number of boxes required is

$$5(\tfrac{1}{5} + \tfrac{1}{4} + \tfrac{1}{3} + \tfrac{1}{2} + 1) \approx 11.42.$$

Euler's Approximation for Harmonic Sums

Though it is easy to add up the reciprocals here, had there been a large number of coupons in a set, it might be convenient to know Euler's approximation for the partial sum of the harmonic series:

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \approx \log_e n + \frac{1}{2n} + 0.57721\ldots.$$

(The 0.57721 . . . is known as Euler's constant.) For n coupons in a set, the average number of boxes is approximately

$$n \log_e n + 0.577\, n + \tfrac{1}{2}.$$

Since $\log_e 5 \approx 1.6094$, Euler's approximation for $n = 5$ yields 11.43, very close to 11.42. Often we omit the term $1/2n$ in Euler's approximation.

An Even Split at Coin Tossing

When 100 coins are tossed, what is the probability that exactly 50 are heads?

Solution for An Even Split at Coin Tossing

Let us order the 100 coins from left to right, and then toss each one. The probability of any particular sequence of 100 tosses, a sequence of 100 heads and tails, is $(\frac{1}{2})^{100}$ because the coins are fair and the tosses independent. For example, the probability that the first 50 are heads and the second 50 are tails is $(\frac{1}{2})^{100}$. How many ways are there to arrange 50 heads and 50 tails in a row? In the Solution to the Perfect Bridge Hand (Problem 8) we found we could use binomial coefficients to make the count. We get $\binom{100}{50} = \frac{100!}{50!50!}$.

Consequently, the probability of an even split is

$$P(\text{even split}) = \frac{100!}{50!50!}\left(\frac{1}{2}\right)^{100}.$$

Evaluating this with logarithms, I get 0.07959 or about 0.08.

Stirling's Approximation

Sometimes, to work theoretically with large factorials, we use Stirling's approximation

$$n! \approx \sqrt{2\pi}\, n^{n+\frac{1}{2}}e^{-n},$$

where e is the base of the natural logarithms. The percentage error in the approximation is about $100/12n$. Let us use Stirling's approximation on the probability of an even split

$$P(\text{even split}) \approx \frac{\sqrt{2\pi}\,100^{100+\frac{1}{2}}e^{-100}}{(\sqrt{2\pi}\,50^{50+\frac{1}{2}}e^{-50})^2 2^{100}} = \frac{100^{\,100+\frac{1}{2}}}{\sqrt{2\pi}\,50^{100}50(2^{100})}$$

$$= \frac{\sqrt{100}}{\sqrt{2\pi}\,50} = \frac{1}{\sqrt{50\pi}} = \frac{1}{5\sqrt{2\pi}}.$$

Since $1/\sqrt{2\pi}$ is about 0.4, the approximation gives about 0.08 as we got before. More precisely the approximation gives to four decimals 0.0798 instead of 0.0796.

Stirling's approximation is discussed in advanced calculus books. For one nice treatment see R. Courant, *Differential and integral calculus*, Vol. I, Translated by E. J. McShane, Interscience Publishers, Inc., New York, 1937, pp. 361–364.