

BERKELEY MATH CIRCLE 2000-2001

Graph Theory, Oct. 8, 2000

PAUL ZEITZ

UNIVERSITY OF SAN FRANCISCO

- (1) Why is the product of three consecutive integers always divisible by 6?
- (2) Why is $1^3 + 2^3 + 3^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4}$?
- (3) Why is the n th Fibonacci number equal to

$$\frac{1}{\sqrt{5}} \left\{ \left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right\}?$$

- (4) Fermat's Little Theorem states that if p is a prime, then $a^p \equiv a \pmod{p}$; in other words, $a^p - a$ is divisible by p . Why is Fermat's Little Theorem true?
- (5) De Moivre's Theorem states that

$$(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta.$$

Why is De Moivre's Theorem true?

- (6) Euler's formula states that

$$e^{i\theta} = \cos \theta + i \sin \theta.$$

Why is it true?

- (7) Why is the Fundamental Theorem of Calculus true?
- (8) *A Problem from BAMM 2000.* Consider the following experiment:
 - (a) First a random number p between 0 and 1 is chosen by spinning an arrow around a dial which is marked from 0 to 1. (This way, the random number is "uniformly distributed"—the chance that p lies in the interval, say, from 0.45 to 0.46 is exactly 1/100; and the chance that p lies in the interval from 0.324 to 0.335 is exactly 11/1000, etc.)
 - (b) Then an unfair coin is built so that it lands "heads up" with probability p .
 - (c) This coin is then flipped 2000 times, and the number of heads seen is recorded.

What is the probability that exactly 1000 heads were recorded?

Four Solutions to The Coin Problem (#8 on previous page). The amazing answer is that the probability is $1/2001$. Indeed, it doesn't matter how many heads we wish to see—for any k between 0 and 2000, the probability that r heads occur is $1/2001$.

We will present three “how” solutions and one “why” solution. The “why” solution is at the end (solution #4); you can skip immediately to it if the discussion of 1–3 is too technical.

Solutions 1–3, while shedding no real light on *why* the answer is what it is, are not without interest. Each one involves different important techniques, although all of them stem from the same basic probability ideas:

Basic probability theory states that if you flip a coin n times, the probability of exactly r heads is

$$\binom{n}{r} p^r (1-p)^{n-r},$$

where p is the probability of getting heads on any given single flip. Then the desired probability is

$$\int_0^1 \binom{n}{r} p^r (1-p)^{n-r} dp,$$

and solutions 1–3 all show, in various ways, that this integral is equal to $1/(n+1)$.

(1) *Integration by Parts.* Consider

$$\int_0^1 p^r (1-p)^{n-r} dp.$$

The strategy is to use integration by parts to get rid of the powers of $(1-p)$ by repeated differentiation of these terms. Meanwhile the exponent of p will rise. To see how this works, let us look at a specific case, $n = 8, r = 5$. We will evaluate the integral

$$\int_0^1 p^5 (1-p)^3 dp.$$

Let

$$u = (1-p)^3, \quad dv = p^5 dp.$$

Then

$$du = -3(1-p)^2 dp, \quad v = p^6/6.$$

Thus

$$\int_0^1 p^5 (1-p)^3 dp = \left. \frac{(1-p)^3 p^6}{6} \right|_0^1 + \frac{3}{6} \int_0^1 p^6 (1-p)^2 dp.$$

Fortunately, the first term of the right-hand side of the above equation disappears, since it equals zero both when $p = 0$ and $p = 1$. Hence we have

$$\int_0^1 p^5 (1-p)^3 dp = \frac{3}{6} \int_0^1 p^6 (1-p)^2 dp.$$

Now we can apply the same integration-by-parts trick on the right-hand-side integral, getting

$$\int_0^1 p^6(1-p)^2 dp = \frac{2}{7} \int_0^1 p^7(1-p)^1 dp.$$

One final application of this trick yields

$$\int_0^1 p^7(1-p)^1 dp = \frac{1}{8} \int_0^1 p^8(1-p)^0 dp,$$

and now we have reduced the integral to one that we can evaluate easily! Combining this all together, we have

$$\begin{aligned} \int_0^1 p^5(1-p)^3 dp &= \frac{3}{6} \cdot \frac{2}{7} \cdot \frac{1}{8} \int_0^1 p^8 dp \\ &= \frac{3}{6} \cdot \frac{2}{7} \cdot \frac{1}{8} \cdot \frac{1}{9} \\ &= \frac{1}{9\binom{8}{3}}. \end{aligned}$$

Hence

$$\int_0^1 p^5(1-p)^3 dp = \frac{1}{9},$$

and this can clearly be generalized to any n, r . □

(2) *Binomial Sums and Induction.* We shall evaluate () using the Binomial theorem. We have

$$p^r(1-p)^{n-r} = p^r \left(1 - \binom{n-r}{1} p + \binom{n-r}{2} p^2 - \dots + (-1)^{n-r} p^{n-r} \right).$$

It is now easy to integrate the right-hand side from 0 to 1. $\int_0^1 p^k dp = 1/(k+1)$. We get

$$\frac{1}{r+1} - \frac{\binom{n-r}{1}}{r+2} + \frac{\binom{n-r}{2}}{r+3} - \dots (-1)^{n-r} \frac{1}{n+1}.$$

It is possible to show by induction that this quantity is $\frac{1}{\binom{n}{r}(n+1)}$, but it is just as easy to look at the more general sum

$$\sum_{k=0}^n (-1)^k \frac{\binom{n}{k}}{m+k} = \frac{\binom{n}{0}}{m} - \frac{\binom{n}{1}}{m+1} + \frac{\binom{n}{2}}{m+2} - \dots (-1)^n \frac{\binom{n}{n}}{m+n}$$

and show that it is equal to

$$\frac{(m-1)!n!}{(m+n)!}.$$

This induction is not hard; we leave it as an exercise. You will probably want to use the fundamental binomial identity

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

□

(3) *Generating Functions.* For $r = 0, 1, 2, \dots, n$, let

$$a_r = \binom{n}{r} p^r (1-p)^{n-r}.$$

Now define the polynomial function (the generating function for the binomial probability distribution)

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n.$$

Notice (by the Binomial Theorem) that

$$f(x) = (xp + (1-p))^n.$$

Now let us integrate $f(x)$ with respect to p from 0 to 1. It is an easy exercise to verify that (remember, p is the variable, while x is held constant)

$$\begin{aligned} \int_0^1 (xp + (1-p))^n dp &= \frac{1}{n+1} (x^{n+1} - 1x - 1) \\ &= \frac{1}{n+1} (1 + x + x^2 + \dots + x^n). \end{aligned}$$

Consequently

$$\begin{aligned} \int_0^1 f(x) dp &= \int_0^1 (a_0 + a_1x + a_2x^2 + \dots + a_nx^n) dp \\ &= \frac{1}{n+1} (1 + x + x^2 + \dots + x^n), \end{aligned}$$

so we conclude that for all r ,

$$\int_0^1 a_r dp = \int_0^1 \binom{n}{r} p^r (1-p)^{n-r} dp = \frac{1}{n+1}.$$

This solution used a sophisticated tool, generating functions. To learn more about them, consult the very readable *generatingfunctionology* by Herbert Wilf. □

(4) *Finally, the “Why” Solution.* The three solutions above magically produced the value $1/(n+1)$ with no hint of what is really going on. Here is a very simple argument which will convince you why this answer makes sense.

Consider a specific case, $n = 5$. We will simulate the experiment of selecting a p and then flipping the p -coin 5 times by using a random number generator which spits out “uniformly distributed” random numbers between 0 and 1.

- (a) First, spit out a random number. This is p .
- (b) Next, produce a new random number. If it is smaller than p , we declare that the coin turned up heads. If it is larger than p , we have tails. This makes sense, since the two random numbers are independent. The probability that the second one is less than p is, of course, p .
- (c) Likewise, produce 4 more random numbers, each simulating a coin toss.

Note that we used 6 random numbers in all; the first to decide on p , and the others (by comparing with the first) to decide whether the p -coin's toss was heads or tails.

If our simulation produced all heads, it would have been the case that the first random number (p) was larger than the next 5 numbers. Conversely, if we had all tails, then the first number would have been the smallest of the 6. What if the simulation produced exactly 1 head? In this case, exactly one of the 5 coin-toss numbers would be smaller than the first (p); in other words, the first number was the 2nd smallest number of the 6.

By now you may see the punchline: the number of heads produced by the simulation depends only on the relative rank of the first random number. If it is the largest of the 6, we produce 5 heads; if it the 2nd largest, 4 heads are produced; and so on.

But when you produce 6 uniformly distributed random numbers the relative rank of the first number is not special, it is no more likely for the first number to be the smallest, say as it is to be the 4th largest.

Thus, of the 6 possible ranks, all are equally likely. So each of the 6 possible coin simulations (0 heads, 1 head, \dots , 5 heads) are equally likely, each with probability $1/6$.

Certainly this generalizes to n coin tosses. The crux idea—the “why”—behind the solution—is the fact that when you select n uniformly distributed random numbers (between 0 and 1), the relative rank of any specific number is itself uniformly distributed (between 1 and n). This is a simple, intuitive property of uniformly distributed random numbers.

So now you should understand *why* the coin problem had the answer that it did. It is not unexpected, magical algebra. It is just simple, almost inevitable symmetry. \square