Topics in
Probability Theory and Stochastic Processes
Steven R. Dunbar

An Analytic Model for Coin Tossing

Rating
Mathematicians Only: prolonged scenes of intense rigor.
Section Starter Question

What are the axioms for a probability space? What are the assumptions about coin-flips or successive Bernoulli trials that make the process a probability space?

Key Concepts

1. The sequence of products

\[
\cos(x/2),
\cos(x/2) \cdot \cos(x/4),
\cos(x/2) \cdot \cos(x/4) \cdot \cos(x/8),
\ldots
\]

converges uniformly on bounded intervals to

\[
\prod_{k=1}^{\infty} \cos\left(\frac{x}{2^k}\right) = \frac{\sin x}{x}
\]

2. An unusual formula involving \(\pi\), due to Vieta, is

\[
\frac{2}{\pi} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2 + \sqrt{2}}}{2} \cdot \frac{\sqrt{2 + \sqrt{2 + \sqrt{2}}}}{2} \ldots
\]

3. Given \(t\) with \(0 \leq t \leq 1\) with unique binary expansion,

\[
t = \frac{\epsilon_1(t)}{2} + \frac{\epsilon_2(t)}{2^2} + \frac{\epsilon_3(t)}{2^3} + \ldots
\]

the Rademacher functions are \(r_k(t) = 1 - 2\epsilon_k(t)\).
4. For the Rademacher functions

\[ \int_{0}^{1} \prod_{k=1}^{\infty} \exp \left( ix \frac{r_k(t)}{2^k} \right) dt = \prod_{k=1}^{\infty} \int_{0}^{1} \exp \left( ix \frac{r_k(t)}{2^k} \right) dt \]

5. The correspondence of \( H \) with +1, \( T \) with −1, the \( k \)th toss with \( r_k(t) \), an event with the set of \( t \)'s in (0, 1), and the probability of an event with the measure of the corresponding set of \( t \)'s gives an analytic model for successive Bernoulli trials.

6. Almost every number \( t \) has asymptotically the same number of 0’s and 1’s in its binary expansion. That is, we say that the binary expansion is normal.

\[ \text{Vocabulary} \]

1. The function \( \frac{\sin x}{x} \) is often called the sinc function. The sinc function \( \text{sinc}(x) \), also called the “sampling function”, arises frequently in signal processing and the theory of Fourier transforms.

2. Every \( t \) with \( 0 \leq t \leq 1 \) has a unique binary expansion,

\[ t = \frac{\epsilon_1(t)}{2} + \frac{\epsilon_2(t)}{2^2} + \frac{\epsilon_3(t)}{2^3} + \ldots \]

3. The Rademacher functions are \( r_k(t) = 1 - 2\epsilon_k(t) \).

4. If \( t \) has asymptotically the same number of 0s and 1s in its binary expansion, the binary expansion is normal.
Mathematical Ideas

Vieta’s Formula from Trigonometry

The following is due both to Vieta in 1593 and also to Euler.

**Theorem 1.** The sequence of products

\[
\cos(x/2), \\
\cos(x/2) \cdot \cos(x/4), \\
\cos(x/2) \cdot \cos(x/4) \cdot \cos(x/8), \\
\ldots
\]

converges uniformly on bounded intervals to

\[
\prod_{k=1}^{\infty} \cos\left(\frac{x}{2^k}\right) = \frac{\sin x}{x}.
\]

*Remark.* The function \(\sin x/x\) is often called the **sinc** function. The cosine product definition shows that \(\text{sinc}(0) = 1\) while \(\sin x/x\) is undefined at \(x = 0\). However, \(\lim_{x \to 0} \sin(x)/x = 1\) so \(\text{sinc}(x)\) is continuous at 0. In the sequel the implicit assumption is that \(\sin x/x\) is 1 at \(x = 0\) even though technically the expression is undefined. The sinc function \(\text{sinc}(x)\), also called the “sampling function”, is a function that arises frequently in signal processing and the theory of Fourier transforms. The full name of the function is “sine cardinal”, but it is commonly referred to by its abbreviation, “sinc”.

*Remark.* The following proof is remarkably elementary for such an unusual identity. The proof combines repeated application of the half-angle identity for the cosine with an elementary limit of the sine function. The proof of uniform convergence uses the Cauchy criterion. The verification of the criterion uses an elementary bound on the cosine function from the series expansion.

*Proof.* 1. Start with repeated application of the half-angle formula from
trigonometry:

\[
\sin(x) = 2 \sin \left( \frac{x}{2} \right) \cos \left( \frac{x}{2} \right)
\]

\[= 2^2 \sin \left( \frac{x}{4} \right) \cos \left( \frac{x}{4} \right) \cos \left( \frac{x}{2} \right)\]

\[= 2^3 \sin \left( \frac{x}{8} \right) \cos \left( \frac{x}{8} \right) \cos \left( \frac{x}{4} \right) \cos \left( \frac{x}{2} \right)\]

\[\vdots\]

\[= 2^n \sin \left( \frac{x}{2^n} \right) \prod_{k=1}^{n} \cos \left( \frac{x}{2^k} \right).\]

2. A standard limit from elementary calculus shows that for \(x \neq 0\)

\[
1 = \lim_{n \to \infty} \frac{\sin(x/2^n)}{x/2^n} = \left. \frac{1}{x} \lim_{n \to \infty} 2^n \sin \left( \frac{x}{2^n} \right) \right|, \]

so

\[x = \lim_{n \to \infty} 2^n \sin \left( \frac{x}{2^n} \right).\]

3. Alternatively, use the series expansion for \(\sin x\)

\[
2^n \sin \left( \frac{x}{2^n} \right) = 2^n \sum_{k=1}^{\infty} \frac{x^{2k-1}/2^{n(2k-1)}}{(2k-1)!}.
\]

\[= x + \sum_{k=2}^{\infty} \frac{x^{2k-1}/2^{n(2k-1)}}{(2k-1)!},\]

and then let \(n \to \infty\). Details are left to the reader.

4. Putting the limits in step 1 and step 2 together

\[
\frac{\sin x}{x} = \prod_{k=1}^{\infty} \cos \left( \frac{x}{2^k} \right).\]

5. For the uniform convergence, consider first the bounded interval \([-\pi, \pi]\).

Then let \(f_n(x) = \prod_{k=1}^{n} \cos \left( \frac{x}{2^k} \right)\) and note that \(f_n(x) \geq 0\) on \([-\pi, \pi]\) with \(f_n(-\pi) = f_n(\pi) = 0\). See Figure 1.
6. Each \( f_n(x) \) is continuous, the sequence converges to the continuous function \( \frac{\sin(x)}{x} \) on the closed bounded (compact) interval \([-\pi, \pi]\), and \( f_n(x) \geq f_{n+1}(x) \) on \([-\pi, \pi]\) since \( \cos(x/2^{n+1}) \leq 1 \). By Theorem 7.13 of [2], the convergence is uniform for the monotone limit of continuous functions to a continuous function on a compact interval.

7. Note that \( f_n(-k\pi) = f_n(-(k-1)\pi) = f_n((k-1)\pi) = f_n(k\pi) = 0 \) for \( n \geq k \). Furthermore \((-1)^k f_n(x) \geq 0 \) for \( n \geq k \). See Figure 1 for an example with \( k = 1 \) and \( n = 2 \).

8. Then the previous step can be repeated on intervals of the form \([k\pi, (k+1)\pi]\) or \([-((k+1)\pi, -k\pi]\) to establish uniform convergence on those intervals. By symmetry, the threshold \( K(k, \epsilon) \) for uniform bound \( \epsilon \) on \([k\pi, (k+1)\pi]\) and \([-((k+1)\pi, k\pi]\) will be the same.

9. For any bounded interval with \(|x| \leq M\), choose \( k \) so large that \( M \leq k\pi \) and use the previous steps to find a uniform bound for uniform convergence on the finite collection of intervals \([-((j+1)\pi, -j\pi] \cup [j\pi, (j+1)\pi]\) for \( k = 1, \ldots, k \) together with \([-\pi, \pi]\).

\[ \cos \left( \frac{\pi}{2^{n+1}} \right) = \sqrt{2 + \sqrt{2 + \sqrt{2 + \cdots + \sqrt{2}}} \div 2} \]

**Proof.** 1. The proof is by mathematical induction.

---

Figure 1: Convergence of the products to \( \frac{\sin(x)}{x} \). The red curve is \( \cos(x/2) \) and the blue curve is \( \cos(x/2) \cdot \cos(x/4) \).

A special case is the basis for an unusual formula involving \( \pi \) due to Vieta in a book published in 1593. First, a simple trigonometric lemma.

**Lemma 2.**

\[ \cos \left( \frac{\pi}{2^{n+1}} \right) = \sqrt{2 + \sqrt{2 + \sqrt{2 + \cdots + \sqrt{2}}} \div 2} \]

**Proof.** 1. The proof is by mathematical induction.
2. The base case for \( n = 1 \) is \( \cos \left( \frac{\pi}{4} \right) = \frac{\sqrt{2}}{2} \), verifying the base case.

3. Given that the equality holds for \( n = k \), let

\[
\alpha = \sqrt{2 + \sqrt{2 + \sqrt{2 + \cdots + \sqrt{2}}}}
\]

so that the assumption is that

\[
\cos \left( \frac{\pi}{2^k+1} \right) = \frac{\alpha}{2}.
\]

4. Then for \( n = k + 1 \),

\[
\cos \left( \frac{\pi}{2^{k+2}} \right) = \cos \left( \frac{\pi/2^k}{2} \right)
= \sqrt{\frac{1 + \alpha/2}{2}}
= \sqrt{\frac{2 + \alpha}{4}}
= \sqrt{2 + \sqrt{2 + \sqrt{2 + \cdots + \sqrt{2}}}}
\]

establishing the induction.

\[\square\]

**Corollary 1.**

\[
\frac{2}{\pi} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2 + \sqrt{2}}}{2} \cdot \frac{\sqrt{2 + \sqrt{2 + \sqrt{2}}}}{2} \cdots
\]

**Proof.** Set \( x = \pi/2 \) and apply the lemma to the infinite product formula for \( \sin x/x \) from the theorem.

\[\square\]

A numerical evaluation of successive finite products is in Table [1] illustrating the convergence.
Table 1: Numerical example of Vieta’s formula.

<table>
<thead>
<tr>
<th>$\frac{2}{\pi}$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.6366198</td>
<td>0.7071068</td>
<td>0.6532815</td>
<td>0.6407289</td>
<td>0.6376436</td>
</tr>
</tbody>
</table>

Binary Expansions and Rademacher Functions

Every $t$ with $0 \leq t < 1$ has a unique binary expansion,

$$t = \frac{\epsilon_1(t)}{2} + \frac{\epsilon_2(t)}{2^2} + \frac{\epsilon_3(t)}{2^3} + \ldots$$

using the convention that terminating expansions have all digits equal to 0 from a certain point. For example, write

$$\frac{3}{4} = 1 \frac{1}{2} + 0 \frac{1}{2^2} + 0 \frac{1}{2^3} + 0 \frac{1}{2^4} + \ldots$$

rather than

$$\frac{3}{4} = 1 \frac{1}{2} + 0 \frac{1}{2^2} + 1 \frac{1}{2^3} + 1 \frac{1}{2^4} + \ldots$$

With the convention about terminating expansion, the graphs of $\epsilon_1(t)$, $\epsilon_2(t)$, $\epsilon_3(t)$, ... are as in Figure 2.

![Figure 2: The bit functions for binary expansions in [0, 1).](image)

It is more convenient to use the **Rademacher functions** defined by $r_k(t) = 1 - 2\epsilon_k(t)$, with graphs as in Figure 2. Note the similarity to the random variables $X_k = 0, 1$ and $Y_k = 1 - 2X_k = \pm 1$. 

8
In terms of the Rademacher functions, the binary expansion $[1]$ becomes

$$1 - 2t = \sum_{k=1}^{\infty} \frac{r(t)}{2^k}.$$  

**Remark.** Note that an alternative description of the Rademacher functions is

$$r_k(t) = \text{sgn}(\sin(2\pi 2^{k-1} t)).$$

In this sense, the Rademacher functions are a discretized version of the sine functions. The Rademacher function expansion follows directly from binary expansion and not from an orthogonal basis.

**Theorem 3.**

$$\int_0^1 \prod_{k=1}^{\infty} \exp \left( ix \frac{r_k(t)}{2^k} \right) \, dt = \prod_{k=1}^{\infty} \int_0^1 \exp \left( ix \frac{r_k(t)}{2^k} \right) \, dt$$

**Remark.** This is remarkable and unusual, since it says an integral of a product is a product of integrals. The proof combines the expansion of $1 - 2t$ in Rademacher functions, elementary integration, the expression of $\sin x$ and $\cos x$ in terms of complex exponentials, and Vieta’s formula.

**Remark.** On the other hand, from an advanced probability point of view, it may not be so remarkable after all. Consider $r_k(t)$ as a random variable over the probability space $[0,1)$. Although the left side is a product of
exponentials, it could be written as the exponential of a sum \( \sum_{k=1}^{\infty} r_k(t) \) of random variables. The left side is then the characteristic function (or re-scaled Fourier transform) of the sum of random variables. By a well-known correspondence, the characteristic function of a sum is equal to the product of individual characteristic functions. Characteristic functions transform questions about sums of random variables and convergence of random variables into analytic questions of products and pointwise convergence of functions. In this way, this theorem points the way to the analytic model of coin-flipping below.

**Proof.**  
1. Let \( u = 1 - 2t \), so \( du = -2 \, dt \) and

\[
\int_{0}^{1} e^{ix(1-2t)} \, dx = \int_{-1}^{1} e^{ixu} \frac{du}{2} = \frac{1}{2} \left( e^{ix} - e^{-ix} \right) = \frac{\sin x}{x}.
\]

2. On the other side

\[
\int_{0}^{1} \exp \left( ix \frac{r_k(t)}{2^k} \right) \, dt = 2^{k-1} \cdot \frac{e^{ix/2^k}}{2^k} + 2^{k-1} \cdot \frac{e^{-ix/2^k}}{2^k} = \cos \left( \frac{x}{2^k} \right).
\]

3. Then Vieta’s formula

\[
\frac{\sin x}{x} = \prod_{k=1}^{\infty} \cos \left( \frac{x}{2^k} \right).
\]

becomes

\[
\int_{0}^{1} e^{ix(1-2t)} \, dt = \prod_{k=1}^{\infty} \int_{0}^{1} \exp \left( ix \frac{r_k(t)}{2^k} \right) \, dt
\]
4. Then substituting
\[ 1 - 2t = \sum_{k=1}^{\infty} \frac{r(t)}{2^k}. \]
and expressing the sum in the exponent as a product of exponentials
\[ \int_0^1 \prod_{k=1}^\infty \exp \left( i r_k(t) \right) \, dt = \int_0^1 \prod_{k=1}^\infty \exp \left( i \frac{r_k(t)}{2^k} \right) \, dt. \]

Another proof of Vieta’s formula

Consider the function \( \sum_{k=1}^{n} c_k r_k(t) \). It is a step function that is constant on the intervals \( \left( \frac{j}{2^n}, \frac{j+1}{2^n} \right) \) for \( j = 0, 1, 2, \ldots, 2^n - 1 \). The values of the function are \( \pm c_1 \pm c_2 \pm c_3 \pm \cdots \pm c_n \). Every sequence of length \( n \) of +1s and −1s corresponds to exactly one interval \( \left( \frac{j}{2^n}, \frac{j+1}{2^n} \right) \). There are \( 2^n \) such intervals and the sequence of \( \pm 1 \) corresponds to the sequence of subintervals of length \( 2^{-j} \) with \( j \leq n \) that \( \left( \frac{j}{2^n}, \frac{j+1}{2^n} \right) \) is in. Then
\[ \int_0^1 \exp \left( i \sum_{k=1}^{n} c_k r_k(t) \right) \, dt = \frac{1}{2^n} \sum_{2^n} \exp \left( i \sum_{k=1}^{n} \pm c_k \right), \quad (2) \]
and the lower limit on the summation indicates that the summation over each of the \( 2^n \) possible subsequences of \( \pm 1 \). By turning a sum of powers into a product
\[ \frac{1}{2^n} \sum_{k=1}^{n} \exp \left( i \sum_{k=1}^{n} \pm c_k \right) = \prod_{k=1}^{n} \left( \frac{e^{ic_k} + e^{-ic_k}}{2} \right) = \prod_{k=1}^{n} \cos(c_k). \quad (3) \]
Putting together \( (2) \) and \( (3) \)
\[ \int_0^1 \exp \left( i \sum_{k=1}^{n} c_k r_k(t) \right) \, dt = \prod_{k=1}^{n} \cos(c_k) = \prod_{k=1}^{n} \int_0^1 \exp (ic_k r_k(t)) \, dt. \quad (4) \]
where the proof of the last equality is the same as in step \( 2 \) of the proof of Theorem 3.
Now set \( c_k = x/2^k \) to get

\[
\int_0^1 \exp \left( \sum_{k=1}^n ix \frac{r_k(t)}{2^k} \right) \, dt = \prod_{k=0}^n \cos \left( \frac{x}{2^k} \right).
\]

Since

\[
\lim_{n \to \infty} \sum_{k=1}^\infty \frac{r_k(t)}{2^k}
\]

converges uniformly in \((0, 1)\) to \(1 - 2t\) by the Weierstrass M-criterion. Therefore we can interchange limit and integral to see that

\[
\frac{\sin(x)}{x} = \int_0^1 e^{ix(1-2t)} \, dt = \lim_{n \to \infty} \int_0^1 \exp \left( ix \sum_{k=1}^n \frac{r_k(t)}{2^k} \right) \, dt = \lim_{n \to \infty} \prod_{k=1}^n \cos \left( \frac{x}{2^k} \right) = \prod_{k=1}^\infty \cos \left( \frac{x}{2^k} \right).
\]

This is a different proof of Vieta’s formula connecting the formula to binary representations of real numbers in \([0, 1)\).

**Measure from Rademacher Functions**

Let \( \mu(E) \) be the length, or more formally, the measure of the subset \( E \) of \([0, 1)\). If \( \delta_1, \delta_2, \ldots, \delta_n \) is a sequence of +1’s and −1’s then

\[
\mu[r_1 = \delta_1, r_2 = \delta_2, \ldots, r_n = \delta_n] = \mu[r_1 = \delta_1] \cdot \mu[r_2 = \delta_2] \cdots \mu[r_n = \delta_n].
\]

**Example.** Let \( \delta_1 = +1, \delta_2 = -1, \delta_3 = -1 \). Then \( \{t : \delta_1 = +1\} = [0, 1/2), \{t : \delta_2 = -1\} = [1/4, 1/2) \cup [3/4, 1), \{t : \delta_3 = -1\} = [1/8, 2/8) \cup [3/8, 4/8) \cup [5/8, 6/8) \cup [7/8, 8/8). \) Consider the set \( E = \{t : \delta_1 = +1, \delta_2 = -1, \delta_3 = -1\} = [3/8, 4/8). \) Then

\[
\frac{1}{8} = \mu[r_1 = \delta_1, r_2 = \delta_2, r_3 = \delta_3] = \mu[r_1 = \delta_1] \cdot \mu[r_2 = \delta_2] \cdot \mu[r_3 = \delta_3] = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}.
\]
Coin tossing | Analytic Model
---|---
Symbol $H$ | +1
Symbol $T$ | −1
$k$th toss | $r_k(t)$
Event | set of $t$’s in $(0, 1)$
Probability of an event | Measure of the corresponding set of $t$’s

Table 2: Correspondence of terms in coin tossing and the analytic model.

This gives another way to rewrite the proof of (4) that is basically the same as before:

$$\int_0^1 \exp \left( i \sum_{k=1}^n c_k r_k(t) \right) dt$$

$$= \sum_{\delta_1, \ldots, \delta_n} \exp \left( i \sum_{k=1}^n c_k \delta_k \right) \cdot \mu[r_1 = \delta_1] \cdot \mu[r_2 = \delta_2] \cdots \mu[r_n = \delta_n]$$

$$= \sum_{\delta_1, \ldots, \delta_n} \prod_{k=1}^n e^{ic_k \delta_k} \prod_{k=1}^n \mu[r_k = \delta_k]$$

$$= \sum_{\delta_1, \ldots, \delta_n} \prod_{k=1}^n e^{ic_k \delta_k} \mu[r_k = \delta_k]$$

$$= \prod_{k=1}^n \sum_{\delta_k} e^{ic_k \delta_k} \mu[r_k = \delta_k]$$

$$= \prod_{k=1}^\infty \int_0^1 \exp (ic_k r_k(t)) dt .$$

**Analytic Model of Coin Tossing**

The following correspondence gives an analytic model of successive Bernoulli trials. That is, consider the following probability scenario. A fair coin is successively tossed $n$ times, with independence of tosses, coming up Heads or Tails on each toss. Use Table 2 to interpret this physical probability experiment analytically.

*Example.* Consider an analytic model for the simplest probability problem
for Bernoulli trials: Find the probability that in \( n \) independent tosses of a fair coin, exactly \( l \) will be heads. Using the table to translate the problem to analytic terms, the problem becomes: Find the measure of the set of \( t \)'s such that exactly \( l \) of the \( n \) numbers \( r_1(t), r_2(t), \ldots, r_n(t) \) are equal to +1.

To solve this problem start by noticing that having exactly \( l \) of \( n \) of the \( r_k(t) \) being +1 means that \( n - l \) are -1, so that

\[
\sum_{k=1}^{n} r_k(t) = l - (n - l) = 2l - n. \tag{5}
\]

Second, notice that

\[
\frac{1}{2\pi} \int_{0}^{2\pi} e^{imx} \, dx = \frac{1}{2\pi} \left( \frac{\text{e}^{im2\pi} - \text{e}^{0}}{im} \right) = 0
\]

for \( m \neq 0 \) and so

\[
\frac{1}{2\pi} \int_{0}^{2\pi} e^{imx} \, dx = \begin{cases} 1 & m = 0 \\ 0 & m \neq 0 \end{cases}
\]

Therefore

\[
\phi(t) = \frac{1}{2\pi} \int_{0}^{2\pi} e^{ix(r_1(t)+r_2(t)+\cdots+r_n(t)-(2l-n))} \, dx
\]

will equal 1 on the set of \( t \)'s satisfying the condition \([5]\), and is equal to 0 otherwise. (Pay careful attention to the variable of integration.) Therefore

\[
\mu[\sum_{k=1}^{n} r_k(t)] = l - (n - l) = 2l - n
\]

\[
= \int_{0}^{1} \phi(t) \, dt
\]

\[
= \int_{0}^{1} \frac{1}{2\pi} \int_{0}^{2\pi} e^{ix(r_1(t)+r_2(t)+\cdots+r_n(t)-(2l-n))} \, dx \, dt
\]

\[
= \frac{1}{2\pi} \int_{0}^{2\pi} e^{-i(2l-n)x} \left( \int_{0}^{1} e^{ix(r_1(t)+r_2(t)+\cdots+r_n(t))} \, dt \right) \, dx.
\]

The interchange of order of integration is usually justified by Fubini’s Theorem, but that is not actually necessary here, since \( r_1(t)+r_2(t)+\cdots+r_n(t) \) is a
step function. Now use (3) on the inner integral with \( c_1 = c_2 = \cdots = c_n = x \) to obtain

\[
\mu[r_1(t) + r_2(t) + \cdots + r_n(t) = 2l - n] = \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \cos^n(x) \, dx.
\]

Evaluating this integral is in the problems, the result of the evaluation gives

\[
\mu[r_1(t) + r_2(t) + \cdots + r_n(t) = 2l - n] = \frac{1}{2^n} \binom{n}{l}.
\]

**Discussion of Independence of Events**

The natural mathematical modeling assumption for probability is that events that seem unrelated are probabilistically independent. That is, the joint probability of unrelated events is the product of the individual probabilities. The product rule is not a mathematical necessity. Rather it is a modeling rule based on experiment and practical experience, justifying the multiplication of probabilities. Thus, probabilistic independence is an intuitive notion with the sense that the multiplication rule is applicable and useful.

In a landmark 1909 paper, “Sur les probabilités dénombrables et leurs applications arithmétiques” E. Borel showed that binary digits, or equivalently the Rademacher functions, are independent in the sense that

\[
\int_0^1 \prod_{k=1}^\infty \exp \left( i x \frac{r_k(t)}{2^k} \right) \, dt = \prod_{k=1}^\infty \int_0^1 \exp \left( i x \frac{r_k(t)}{2^k} \right) \, dt.
\]

This observation gives well-defined mathematical objects to which the postulates of probability theory apply directly.

**Weak and Strong Laws**

The application of the postulates of probability to Rademacher functions eliminates casting probability in terms of coins and tosses. All of the previous proofs of theorems for coin-tossing have equivalents using Rademacher functions, or equivalently binary digits, with Table 2. The Weak Law of Large Numbers, the direct Borel-Cantelli Lemma and the Strong Law of Large Numbers serve as examples.
Lemma 4 (Orthogonality of Rademacher Functions). If $k_1 < k_2 < \cdots < k_n$,

$$
\int_0^1 r_{k_1}(t)r_{k_2}(t) \cdots r_{k_n}(t) \, dt = 0.
$$

Remark. Consider the example in Figure 4. Each of the integrals over subintervals $[0, 1/4]$, $[1/4, 1/2]$, $[1/2, 3/4]$ and $[3/4, 1]$ are like a rescaled version of $\pm \int_0^1 r_1(t) \, dt = 0$, so the entire integral is 0. The proof expands that idea using induction on the number of factors in the product.

Proof. 1. The proof is by induction on the number of factors $n$ in the product.

2. The base case for $n = 1$ is obvious by the symmetry and regularity of $r_k(t)$:

$$
\int_0^1 r_{k_1}(t) \, dt.
$$

3. Suppose that the conclusion is true for $n - 1$ factors,

$$
\int_0^1 r_{k_2}(t)r_{k_3}(t) \cdots r_{k_n}(t) \, dt = 0
$$

and consider

$$
\int_0^1 r_{k_1}(t)r_{k_2}(t) \cdots r_{k_n}(t) \, dt.
$$

4. Break the integral into $2^{k_1}$ integrals over the subintervals $[\ell/2^{k_1}, (\ell + 1)/2^{k_1}]$, $\ell = 0, \ldots, 2^k - 1$.

$$
\int_0^1 r_{k_1}(t)r_{k_2}(t) \cdots r_{k_n}(t) \, dt = \sum_{\ell=0}^{2^{k_1}-1} (-1)^\ell \int_{\ell/2^{k_1}}^{(\ell+1)/2^{k_1}} r_{k_2}(t)r_{k_3}(t) \cdots r_{k_n}(t) \, dt.
$$
Figure 4: Graph of the product of Rademacher functions.
5. For each integral, change variables \( t = \ell/2^k_1 + u_\ell/2^k_1 \) for \( u_\ell \in [0, 1] \).
Then the summation of integrals becomes
\[
\sum_{\ell=0}^{2^{k_1-1}} \frac{(-1)^\ell}{2^{k_1}} \left( \int_0^1 r_{k_2}(\frac{\ell}{2^{k_1}} + \frac{u_\ell}{2^{k_1}}) r_{k_3}(\frac{\ell}{2^{k_1}} + \frac{u_\ell}{2^{k_1}}) \cdots r_{k_n}(\frac{\ell}{2^{k_1}} + \frac{u_\ell}{2^{k_1}}) \right) du_\ell.
\]

6. Using the recursive relation that
\[
r_{k_i}(\frac{\ell}{2^{k_1}} + \frac{u_\ell}{2^{k_1}}) = r_{k_i-k_1}(u)
\]
the summation becomes
\[
\sum_{\ell=0}^{2^{k_1-1}} \frac{(-1)^\ell}{2^{k_1}} \left( \int_0^1 r_{k_2}(u_\ell) r_{k_3}(u_\ell) \cdots r_{k_n}(u_\ell) \right) du_\ell.
\]

7. Each summand is 0 by the induction hypothesis, so the sum is 0 and the induction step is complete.

\[\square\]

**Theorem 5** (Weak Law of Large Numbers). *For every \( \epsilon > 0 \),
\[
\lim_{n \to \infty} \mu[|r_1(t) + \cdots + r_n(t)| > \epsilon n] = 0
\]

*Remark.* The proof is essentially the standard one using Chebyshev’s inequality expressed directly in terms of the measure and integral of Rademacher functions.

**Proof.**
1. On the one hand,
\[
\int_0^1 (r_1(t) + \cdots + r_n(t))^2 \, dt
\]
\[
\geq \int_{|r_1(t) + \cdots + r_n(t)| > \epsilon n} (r_1(t) + \cdots + r_n(t))^2 \, dt
\]
\[
> \epsilon^2 n^2 \mu[|r_1(t) + \cdots + r_n(t)| > \epsilon n].
\]
2. On the other hand, using the Lemma on Orthogonality of Rademacher Functions,
\[
\int_0^1 (r_1(t) + \cdots + r_n(t))^2 \, dt = n.
\]

3. Combining the two previous points,
\[
\mu[|r_1(t) + \cdots + r_n(t)| > \epsilon n] < \frac{1}{\epsilon^2 n}
\]
so
\[
\lim_{n \to \infty} \mu[|r_1(t) + \cdots + r_n(t)| > \epsilon n] = 0.
\]

Lemma 6. If \( f_n(t) \) is a sequence of non-negative Lebesgue integrable functions, then convergence of \( \sum_{n=1}^{\infty} \int_0^1 f_n(t) \, dt \) implies
\[
\lim_{n \to \infty} f_n(t) = 0
\]
almost everywhere.

Remark (Borel-Cantelli Direct Half). The proof is a direct translation of the probabilistic proof of the direct Borel-Cantelli Lemma, using an analytic version of Markov’s inequality.

Proof. 1. Let an arbitrary \( a > 0 \) be given.

2. Since \( f_n(t) \geq 0 \)
\[
\int_0^1 f_n(t) \, dt = \int_{\{t : f_n(t) < a\}} f_n(t) \, dt + \int_{\{t : f_n(t) \geq a\}} f_n(t) \, dt \\
\geq a\mu[f_n(t) \geq a].
\]

3. Therefore \( \sum_{n=1}^{\infty} \int_0^1 f_n(t) \, dt < \infty \) implies \( \sum_{n=1}^{\infty} a\mu[f_n(t) \geq a] < \infty \).

4. In turn, that implies \( \mu[f_n(t) \geq a] \to 0 \) as \( n \to \infty \).
5. Since this is true for any \( a > 0 \),

\[
\lim_{n \to \infty} f_n(t) = 0
\]

almost everywhere.

\[\square\]

**Theorem 7** (Strong Law of Large Numbers). *For almost every \( t \) (that is, for all \( t \) except for a set of measure 0)*

\[
\lim_{n \to \infty} \frac{\sum_{i=1}^{n} r_i(t)}{n} = 0.
\]

**Remark.** This proof resembles the third proof of the Strong Law in [Strong Law of Large Numbers](#). The proof substitutes the analytic form of the direct half of the Borel-Cantelli Lemma to show the convergence almost everywhere.

**Proof.** 1. Set

\[
f_n(t) = \left( \frac{\sum_{i=1}^{n} r_i(t)}{n} \right)^4
\]

and consider

\[
\int_{0}^{1} f_n(t) \, dt = \int_{0}^{1} \left( \frac{\sum_{i=1}^{n} r_i(t)}{n} \right)^4 \, dt.
\]

2. Using the Orthogonality of Rademacher Functions

\[
\int_{0}^{1} \left( \frac{\sum_{i=1}^{n} r_i(t)}{n} \right)^4 \, dt = \frac{n + \binom{n}{2} \binom{n}{4}}{n^4} \leq \frac{C}{n^2}
\]

since the only terms that contribute positive values are the fourth powers and the pairs of squares.

3. Therefore

\[
\sum_{n=1}^{\infty} \int_{0}^{1} f_n(t) \, dt < \infty.
\]
4. The proof now uses the preceding lemma on the analytic form of the Borel-Cantelli Lemma. With this lemma
\[ \lim_{n \to \infty} \left( \frac{r_1(t) + \cdots + r_n(t)}{n} \right)^4 = 0 \]
almost everywhere, and so
\[ \lim_{n \to \infty} \frac{r_1(t) + \cdots + r_n(t)}{n} = 0. \]
almost everywhere.

Remark. Recall that \( r_k(t) = 1 - 2\epsilon_k(t) \) and that \( \epsilon_k(t) \) is the \( k \)th bit in the binary expansion of \( t \). Then the Strong Law of Large Numbers implies that for almost every \( t \)
\[ \lim_{n \to \infty} \frac{\epsilon_1(t) + \cdots + \epsilon_n(t)}{n} = \frac{1}{2}. \]
In other words, almost every number \( t \) has asymptotically the same number of 0’s and 1’s. That is, we say that the binary expansion is normal.

Sources
This section is adapted from: Statistical Independence in Probability, Analysis, and Number Theory by Mark Kac, 1959, Mathematical Association of America, pages 1–35. Problems 2 and 3 are from the same source. The Cauchy criterion for uniform convergence is Theorem 7.8, page 134 in Principles of Mathematical Analysis second edition by W. Rudin, McGraw-Hill, 1964. The remarks about the sinc function are from Mathworld: Sinc function.
Algorithms, Scripts, Simulations

Algorithm

Scripts

Problems to Work for Understanding

1. Show by direct integration that

\[
\frac{1}{2\pi} \int_{0}^{2\pi} e^{-i(0)x} \cos^0(x) \, dx = 1
\]

and

\[
\frac{1}{2\pi} \int_{0}^{2\pi} e^{-i(2\cdot0-1)x} \cos(x) \, dx = 1/2.
\]

and

\[
\frac{1}{2\pi} \int_{0}^{2\pi} e^{-i(2\cdot1-1)x} \cos(x) \, dx = 1/2.
\]

2. For \( n \geq 2 \) and \( l \leq n \), show that

\[
\frac{1}{2\pi} \int_{0}^{2\pi} e^{-i(2l-n)x} \cos^n(x) \, dx = \frac{1}{2\pi} \binom{n}{l}.
\]

3. Every \( t \) with \( 0 \leq t < 1 \) has a unique ternary expansion,

\[
t = \frac{\eta_1(t)}{2} + \frac{\eta_2(t)}{2^2} + \frac{\eta_3(t)}{2^3} + \ldots
\]

where \( \eta_k(t) = 0, 1 \) or 2. Prove that \( \eta_i(t) \) is independent of \( \eta_j(t) \).
4. Prove that
\[
\frac{\sin(x)}{x} = \prod_{n=1}^{\infty} \frac{1 + 2 \cos \left(\frac{(2x)/3^k}{3}\right)}{3}
\]

5. For \(0 < p < 1\) let
\[
T_p(t) = \begin{cases} 
  t/p, & 0 \leq t \leq p \\
  (t-p)/(1-p) & p < t \leq t 
\end{cases}
\]
and let
\[
\epsilon_p(t) = \begin{cases} 
  1 & 0 \leq t \leq p, \\
  0 & p < t \leq 1 
\end{cases}
\]
Plot the functions \(\epsilon_1^{(p)}(t) = \epsilon_p(t), \epsilon_2^{(p)}(t) = \epsilon_p(T_p(t)), \epsilon_3^{(p)}(t) = \epsilon_p(T_p(T_p(t)))\) and so on, and show that they are independent. This is the basis for an analytic model of the unfair coin. See the next problem.

6. Prove that the measure of the set on which
\[
\epsilon_1^{(p)} + \epsilon_2^{(p)} + \cdots + \epsilon_n^{(p)} = l
\]
where \(0 \leq l \leq n\) is
\[
\binom{n}{l} p^l (1-p)^{n-l}.
\]

Reading Suggestion:

References
