Chapter 1

Coin tossing process

Throughout this monograph, the *coin tossing process*[†1] plays a role of the model process of random number and pseudorandom number. This may sound very restrictive for applications, but it is not. Indeed, from a coin tossing process, any practical random variables and any stochastic processes can be constructed.

1.1 Borel's model of coin tossing process

To describe m coin tosses, we use a probability space $(\{0,1\}^m, 2^{\{0,1\}^m}, P_m)$, where 0 and 1 stand for Tails and Heads respectively, and P_m stands for the uniform probability measure on $\{0,1\}^m$;

$$P_m(B) := \frac{\#B}{2^m}, \quad B \subset \{0,1\}^m \, (B \in 2^{\{0,1\}^m}).$$

But each time m changes, we must take another probability space, which is not only boring but also inconvenient when we consider limit theorems. It is a good idea to construct an infinite many coin tosses all at once on a suitable probability space. Following Borel's idea, we construct them all on the Lebesgue probability space.

Definition 1.1

1. Let $\mathbb{T}^1$ be a 1-dimensional torus, i.e., an additive group consisting of the unit interval $[0,1)$ with addition $(x+y)$ mod 1. Let $\mathcal{B}$ be a σ-algebra on $\mathbb{T}^1 = [0,1)$ consisting of all the Borel measurable sets of it, $\mathbb{P}$ be the Lebesgue measure. The triplet $(\mathbb{T}^1, \mathcal{B}, \mathbb{P})$ is called the *Lebesgue probability space*.[†2] Let $(\mathbb{T}^k, \mathcal{B}^k, \mathbb{P}^k)$ denote the k-fold direct product of $(\mathbb{T}^1, \mathcal{B}, \mathbb{P})$, which is called the *k-dimensional Lebesgue probability space*.

2. Let $d_i(x) \in \{0,1\}$ denote the i-th digit of real $x \in \mathbb{T}^1$ in its dyadic expansion;

$$x = \sum_{i=1}^{\infty} d_i(x) 2^{-i}, \quad x \in \mathbb{T}^1, \tag{1.1}$$

[†1] We call the *fair* coin tossing process simply the coin tossing process.

[†2] We sometimes consider the completion of $\mathcal{B}$ by $\mathbb{P}$, i.e., σ-algebra of all the Lebesgue measurable sets. But for numerical calculations, $\mathcal{B}$ will do.

where

$$d_1(x) := \mathbf{1}_{[1/2,1)}(x), \quad d_i(x) := d_1(2^{i-1}x), \quad i \in \mathbb{N}^+, \quad x \in \mathbb{T}^1.$$

3. For each $m \in \mathbb{N}^+$, we define

$$D_m := \{i2^{-m} \,|\, i = 0, \ldots, 2^m - 1\} \subset \mathbb{T}^1. \tag{1.2}$$

Let $\mathcal{B}_m$ be the algebra generated by the collection of sets $\mathcal{I}_m := \{[a, b) | a, b \in D_m\}$. Namely, each element of $\mathcal{B}_m$ is a finite union of some elements of $\mathcal{I}_m$. Let $P_{(m)}$ be the uniform probability measure on D_m.

4. For each $m \in \mathbb{N}^+$ and each $x \in \mathbb{T}^1$, we define

$$\lfloor x \rfloor_m := \lfloor 2^m x \rfloor / 2^m \in D_m, \tag{1.3}$$
$$\lceil x \rceil_m := \lceil 2^m x \rceil / 2^m \in D_m, \tag{1.4}$$

and $\lfloor x \rfloor_\infty := x$.

Theorem 1.2 *([4]) The sequence of random variables $\{d_i\}_{i=1}^\infty$ defined on the Lebesgue probability space is a coin tossing process.*

Proof. For any $n \in \mathbb{N}^+$, any $\epsilon_1, \ldots \epsilon_n \in \{0, 1\}$, defining $t := \sum_{i=1}^n 2^{-i}\epsilon_i$, we see that

$$\left\{ x \in \mathbb{T}^1 \,\middle|\, d_i(x) = \epsilon_i,\ i = 1, \ldots, n \right\} = [t, t + 2^{-n}),$$

from which it follows that $\mathbb{P}(d_i = \epsilon_i,\ i = 1, \ldots, n) = \mathbb{P}([t, t + 2^{-n})) = 2^{-n}$. □

The dyadic expansion mapping $D_m \ni x \mapsto (d_1(x), \ldots, d_m(x)) \in \{0, 1\}^m$ is a bijection, and a mapping $\lfloor \bullet \rfloor_m : \mathbb{T}^1 \to D_m$ (or $\lceil \bullet \rceil_m : \mathbb{T}^1 \to D_m$) induces a bijection between $\mathcal{B}_m$ and 2^{D_m}. By these facts, the following three probability spaces are isomorphic to each other.

$$(\{0, 1\}^m, 2^{\{0,1\}^m}, P_m) \cong (D_m, 2^{D_m}, P_{(m)}) \cong (\mathbb{T}^1, \mathcal{B}_m, \mathbb{P}).$$

1.2 Construction of random variables from coin tossing process

Theorem 1.3 *([7] (1.2)Theorem) Let S be a real valued random variable defined on a probability space $(\Omega, \mathcal{F}, P)$. Then there exists a random variable f on the Lebesgue probability space such that f and S are identically distributed.*

Proof. Using the distribution function $F(S; t) := P(S \le x)$, $x \in \mathbb{R}$, of S, put

$$f(x) := \sup\{u \in \mathbb{R} \,|\, F(S; u) < x\}, \quad 0 < x < 1. \tag{1.5}$$

Then, f regarded as a random variable on the Lebesgue probability space is what we want. To show this, it is enough to prove that

$$\{\, 0 < x < 1 \,|\, f(x) \le t \,\} = \{\, 0 < x < 1 \,|\, x \le F(S; t) \,\}, \quad t \in \mathbb{R}, \tag{1.6}$$

because calculating the Lebesgue measures of the both hand sides, we see $\mathbb{P}(f \le t) = F(S;t)$. Let us show (1.6). First, since $x \le F(S;t)$ implies $t \notin \{u \in \mathbb{R} \,|\, F(S;u) < x\}$, we have $f(x) \le t$. On the other hand, because $F(S;\bullet)$ is right continuous, $x > F(S;t)$ implies that there exists an $\varepsilon > 0$ such that $x > F(S;t+\varepsilon)$. Hence $f(x) \ge t + \varepsilon > t$. From these facts, (1.6) follows. □

There are many f's that satisfy the condition of Theorem 1.3, and the f defined by (1.5) is merely one of them. The random variable f on the Lebesgue probability space can always be considered as a functional[†3] of the coin tossing process $\{d_i\}_{i=1}^{\infty}$ through the following formula;

$$f(x) = f\left(\sum_{i=1}^{\infty} d_i(x)2^{-i}\right), \quad x \in \mathbb{T}.$$

Theorem 1.3 therefore implies that for any random variable S, there exists a functional of a coin tossing process which has the same distribution as S.

Theorem 1.4 *([35]) Define a sequence of random variables $\{Z_n\}_{n=1}^{\infty}$ on the Lebesgue probability space by*

$$\begin{aligned}
Z_1 &= \frac{1}{2}d_1 + \frac{1}{2^2}d_3 + \frac{1}{2^3}d_6 + \frac{1}{2^4}d_{10} + \cdots \\
Z_2 &= \frac{1}{2}d_2 + \frac{1}{2^2}d_5 + \frac{1}{2^3}d_9 + \cdots \\
Z_3 &= \frac{1}{2}d_4 + \frac{1}{2^2}d_8 + \cdots \\
Z_4 &= \frac{1}{2}d_7 + \cdots \\
&\vdots
\end{aligned}$$

Then, $\{Z_n\}_{n=1}^{\infty}$ is a sequence of i.i.d.[†4] random variables, each Z_n being uniformly distributed on $\mathbb{T}^1$.

The proof of Theorem 1.4 is easy and hence it is omitted here.

Theorem 1.3 and Theorem 1.4 imply that any sequence of independent random variables of arbitrary distribution can be constructed from a coin tossing process. For instance, we can construct an i.i.d. sequence $\{\xi_n\}_{n=0}^{\infty}$ of $\mathcal{N}(0,1)$-variables from a coin tossing process. Using this sequence, Wiener constructed a Brownian motion process $\{B_t\}_{0\le t\le\pi}$ by

$$B_t := \frac{t}{\sqrt{\pi}}\xi_0 + \sqrt{\frac{2}{\pi}}\sum_{n=1}^{\infty}\frac{\sin nt}{n}\xi_n, \quad 0 \le t \le \pi.$$

For details see [11] p.21. Applying the procedure of Theorem 1.4 again, it is readily seen that we can even construct countably many independent Brownian motion processes from the coin tossing process $\{d_i\}_{i=1}^{\infty}$ defined on the Lebesgue probability space. As a matter of fact, except special cases (e.g., construction of uncountably many independent random variables), almost all random objects can be constructed from a coin tossing process. For details, see [30] Chapter 1.

[†3] A function of infinitely many variables is called a functional. Here, $f(x)$ can be regarded as a function of the infinitely many values of $d_i(x)$, $i = 1, 2, \ldots$.

[†4] i.i.d. stands for *i*ndependently *i*dentically *d*istributed.

1.3 Simulatable random variable

Theorem 1.3 does not say that any random variable S can always be constructed in practice from a coin tossing process. Indeed, the distribution function of S is usually hard to get explicitly, and hence we can seldom compute f of (1.5) in practice. Moreover, samples of coin tosses are not provided at once in Monte Carlo methods, but they are provided one by one successively from the first term. Consequently, S should be computed by a finite number of samples of coin tosses with probability 1.

When a functional f of the coin tossing process can be realized in practice, it is said to be *simulatable*. In this section, we consider a precise condition for f to be simulatable.[†5]

1.3.1 Stopping time and simulatable random variable

A random variable defined on the Lebesgue probability space which is $\mathcal{B}_m$-measurable for some $m \in \mathbb{N}^+$ is obviously simulatable.[†6] On the other hand, there exist functions which are not $\mathcal{B}_m$-measurable for any $m \in \mathbb{N}^+$ but are simulatable in practice. Look at the following example.

Example 1.5 (Hitting time) Consider a random variable

$$\sigma(x) := \inf\{\, n \in \mathbb{N}^+ \,|\, d_1(x) + d_2(x) + \cdots + d_n(x) = 5\}, \quad x \in \mathbb{T}^1,$$

defined on the Lebesgue probability space. σ is the first time when the total number of Heads becomes 5 in successive coin tosses. (Here we define $\inf \emptyset = \infty$.) Obviously, it is not $\mathcal{B}_m$-measurable for any $m \in \mathbb{N}^+$, but nevertheless when the 5-th Heads comes up, we can stop tossing the coin, and get the value of σ. Thus we can compute $\sigma(x)$ from finite coin tosses with probability 1.

Let us specify a general class of simulatable random variables that includes σ of Example 1.5.

Definition 1.6 A random variable $\tau : \mathbb{T}^1 \to \mathbb{N}^+ \cup \{\infty\}$ is called a $\{\mathcal{B}_m\}_m$-*stopping time* (cf. [1]) or simply a *stopping time* if it satisfies

$$\forall m \in \mathbb{N}^+, \quad \{\tau \le m\} := \{x \in \mathbb{T}^1 \,|\, \tau(x) \le m\} \in \mathcal{B}_m.$$

For a stopping time τ, we define a sub-σ-algebra

$$\mathcal{B}_\tau := \{A \in \mathcal{B} \,|\, \forall m \in \mathbb{N}^+,\, A \cap \{\tau \le m\} \in \mathcal{B}_m\}.$$

For simplicity, we use the term "τ-measurable" to mean "$\mathcal{B}_\tau$-measurable", and $L^p(\mathcal{B}_\tau)$ to mean $L^p(\mathbb{T}^1, \mathcal{B}_\tau, \mathbb{P})$.

[†5] The contents of § 1.3 will not be necessary until § 5.3, so the reader may skip this section at the first reading.

[†6] Of course, if m is an astronomical number, it would be impossible to deal with $\mathcal{B}_m$-measurable functions in practice. The simulatability here should be understood in a theoretical sense. More precisely, f is simulatable if there exists a Turing machine (cf. [6]) which computes f.

A constant time $\tau(x) \equiv m \in \mathbb{N}^+$ is a stopping time and $\mathcal{B}_\tau = \mathcal{B}_m$.

A function $f : \mathbb{T}^1 \to \mathbb{R} \cup \{\pm\infty\}$ is $\mathcal{B}_m$-measurable, if and only if $f(x) = f(\lfloor x \rfloor_m)$, $x \in \mathbb{T}^1$. As a generalization of this, we have the following.

Lemma 1.7 *Let τ be a stopping time. A function $f : \mathbb{T}^1 \to \mathbb{R} \cup \{\pm\infty\}$ is τ-measurable, if and only if*

$$f(x) = f\left(\lfloor x \rfloor_{\tau(x)}\right), \qquad x \in \mathbb{T}^1. \tag{1.7}$$

Proof. Necessity: Suppose that f is τ-measurable. Then, for each $m \in \mathbb{N}^+$ and each $t \in \mathbb{R}$, we have $\{\tau \le m\} \cap \{f \le t\} \in \mathcal{B}_m$. This means that $\tau(x) \le m$ implies $f(x) = f(\lfloor x \rfloor_m)$. Consequently,

$$\begin{aligned} f(x) &= \sum_{m\in\mathbb{N}^+} f(x)\mathbf{1}_{\{\tau=m\}}(x) + f(x)\mathbf{1}_{\{\tau=\infty\}}(x) \\ &= \sum_{m\in\mathbb{N}^+} f\left(\lfloor x \rfloor_m\right)\mathbf{1}_{\{\tau=m\}}(x) + f\left(\lfloor x \rfloor_\infty\right)\mathbf{1}_{\{\tau=\infty\}}(x) \\ &= \sum_{m\in\mathbb{N}^+} f\left(\lfloor x \rfloor_{\tau(x)}\right)\mathbf{1}_{\{\tau=m\}}(x) + f\left(\lfloor x \rfloor_{\tau(x)}\right)\mathbf{1}_{\{\tau=\infty\}}(x) \\ &= f\left(\lfloor x \rfloor_{\tau(x)}\right)\sum_{m\in\mathbb{N}^+} \mathbf{1}_{\{\tau=m\}}(x) + f\left(\lfloor x \rfloor_{\tau(x)}\right)\mathbf{1}_{\{\tau=\infty\}}(x) = f\left(\lfloor x \rfloor_{\tau(x)}\right). \end{aligned}$$

Sufficiency： Suppose that f satisfies (1.7). Then for each $m \in \mathbb{N}^+$ and each $t \in \mathbb{R}$, we have

$$\{f \le t\} \cap \{\tau \le m\} = \{f\left(\lfloor \bullet \rfloor_{\tau(\bullet)}\right) \le t\} \cap \{\tau \le m\} = \{f\left(\lfloor \bullet \rfloor_m\right) \le t\} \cap \{\tau \le m\} \in \mathcal{B}_m.$$

Thus f is τ-measurable. □

The random variable σ in Example 1.5 is a stopping time, and of course it is σ-measurable.

A function f which is τ-measurable for some stopping time τ that is finite with probability 1 is simulatable. Indeed, suppose that samples of the coin tosses $\{d_i(x)\}_{i=1}^\infty$ are provided one by one successively from the first term. Then the following algorithm computes f.

1. Set $m := 1$.
2. Set $t := \sum_{i=1}^m 2^{-i} d_i(x)\, (= \lfloor x \rfloor_m)$.
3. If $\tau(t) = m$, then output $f(t)$ and end.
4. If $\tau(t) > m$, then set $m := m + 1$ and go to 2.

Since τ is finite with probability 1, this algorithm ends in finite time with output $f(x)$ with probability 1.

Conversely, if f is simulatable, it must be computed by a finite number of coin tosses with probability 1. Namely, for $\mathbb{P}$-a.e. $x \in \mathbb{T}^1$, there exists an $m \in \mathbb{N}^+$ such that $f(x) =$

$f(\lfloor x \rfloor_m)$. Here m may depend on x and we write it as $\tau(x)$. Thus the function $\tau : \mathbb{T}^1 \to \mathbb{N}^+ \cup \{\infty\}$ satisfies $\mathbb{P}(\tau < \infty) = 1$. Suppose that τ is not a stopping time. Then it may happen that a sample sequence of the coin tossing process should be provided forever to compute $f(x)$, which means that f cannot be simulatable. Thus that τ is a stopping time is indispensable for f to be simulatable.

With these reasons, we define that *f is simulatable if it is τ-measurable for some stopping time τ that is finite with probability 1.*

Example 1.8 (Last exit time) A random variable

$$\tau'(x) := \sup\left(\left\{ n \in \mathbb{N}^+ \;\middle|\; d_1(x) + d_2(x) + \cdots + d_n(x) - \frac{n}{3} < 0 \right\} \cup \{1\}\right), \quad x \in \mathbb{T}^1,$$

is finite with probability 1 by the strong law of large numbers, but it is not a stopping time. Indeed, the value of $\tau'(x)$ can never be computed from a finite number of terms of $\{d_i(x)\}_{i=1}^{\infty}$. $\tau'(x)$ is not simulatable.

1.3.2 $\mathbb{T}^1$-valued uniform i.i.d. sequence as random source

In Monte Carlo methods, we usually take a $\mathbb{T}^1$-valued uniform i.i.d. sequence as the random source of simulations. In this context, the simulatability, or equivalently, the measurability with respect to stopping time is stated as follows.

Assumption 1.9 [†7] Suppose that f is a functional of a $\mathbb{T}^1$-valued uniform i.i.d. sequence $\{Z_l\}_{l=1}^{\infty}$, and that it requires only a finite number of $Z_1, \ldots, Z_T$ to be computed with probability 1. Here T is a random variable with the following property; for each $l \in \mathbb{N}^+$, whether the event $\{T \le l\}$ occurs or not can be judged by the values of $Z_1, \ldots, Z_l$ without any knowledge about $Z_{l'}$, $l' \ge l + 1$.

Example 1.10 In case f can always be computed from a constant number of Z_l's, i.e., T is a constant, it satisfies Assumption 1.9.

As is seen in the following example, we need not be aware of the stopping time T so explicitly in most of practical computations.

Example 1.11 (von Neumann's rejection method [28]) Let $p(x)$, $x \in [a, b]$, be a bounded probability density function. We consider an algorithm to generate a random variable f whose probability density function is p. Let $M > 0$ be an upper bound of the function p, and (ξ, η) be a random point which is uniformly distributed in $[a, b] \times [0, M]$. Then we have

$$\Pr(\, \xi \in [c, d] \,|\, p(\xi) \ge \eta \,) \;=\; \int_c^d p(x)dx, \quad a \le c < d \le b.$$

With this knowledge, we consider the following algorithm, which uses a $\mathbb{T}^1$-valued uniform i.i.d. sequence $\{Z_l\}_{l=1}^{\infty}$ as the random source.

1. Set $l := 1$.

[†7] A more precise formulation will be given in § 5.4.5.

2. Set $(\xi, \eta) := ((b-a)Z_{2l-1} + a, MZ_{2l})$.

3. If $p(\xi) \geq \eta$, then output $f := \xi$ and stop.

4. If $p(\xi) < \eta$, then set $l := l + 1$ and go to 2.

The output f of this algorithm, which obviously satisfies Assumption 1.9, is what we wish to get.

Example 1.12 (cf. Example 1.5) Define $\{Y_n\}_{n=1}^{\infty}$ by

$$Y_n := \begin{cases} 0 & (Z_n \in [0, 1/2)) \\ 1 & (Z_n \in [1/2, 1)) \end{cases} \qquad n = 1, 2, \ldots.$$

$\{Y_n\}_{n=1}^{\infty}$ is a coin tossing process. Then

$$f := \inf\{ n \in \mathbb{N}^+ \mid Y_1 + Y_2 + \cdots + Y_n = 5 \}$$

satisfies Assumption 1.9. f is the first time when the total number of Heads becomes 5 in successive coin tosses.

In practical computations, real numbers are treated in finite precision, say 2^{-K}. Accordingly, instead of $\{Z_l\}_{l=1}^{\infty}$, we use a D_K-valued uniform i.i.d. random variables $\{Z_l^{(K)}\}_{l=1}^{\infty}$, which is, for example, defined on $(\mathbb{T}^1, \mathcal{B}, \mathbb{P})$ by

$$Z_n^{(K)} := \sum_{i=1}^{K} 2^{-i} d_{(n-1)K+i}, \qquad n \in \mathbb{N}^+, \tag{1.8}$$

i.e.,

$$\begin{aligned} Z_1^{(K)} &= \frac{1}{2}d_1 + \frac{1}{2^2}d_2 + \cdots + \frac{1}{2^K}d_K \\ Z_2^{(K)} &= \frac{1}{2}d_{K+1} + \frac{1}{2^2}d_{K+2} + \cdots + \frac{1}{2^K}d_{2K} \\ Z_3^{(K)} &= \frac{1}{2}d_{2K+1} + \frac{1}{2^2}d_{2K+2} + \cdots + \frac{1}{2^K}d_{3K} \\ &\vdots \end{aligned}$$

If f satisfies Assumption 1.9, regarding it as a functional of $\{Z_l^{(K)}\}_{l=1}^{\infty}$, set

$$\tau(x) := \inf\{ lK \in l\,\mathbb{N}^+ \mid f(x) \text{ is computed from } Z_1^{(K)}(x), \ldots, Z_l^{(K)}(x) \}.$$

Then τ becomes a $\{\mathcal{B}_m\}_m$-stopping time, and f is τ-measurable.