Chapter 1

Basic Definitions: Indexed Collections and Random Functions

Section 1.1 introduces stochastic processes as indexed collections of random variables.

Section 1.2 builds the necessary machinery to consider random functions, especially the product σ-field and the notion of sample paths, and then re-defines stochastic processes as random functions whose sample paths lie in nice sets.

This first chapter begins at the beginning, by defining stochastic processes. Even if you have seen this definition before, it will be useful to review it.

We will flip back and forth between two ways of thinking about stochastic processes: as indexed collections of random variables, and as random functions.

As always, assume we have a nice base probability space $(\Omega, \mathcal{F}, P)$, which is rich enough that all the random variables we need exist.

1.1 So, What Is a Stochastic Process?

Definition 1 (A Stochastic Process Is a Collection of Random Variables) *A stochastic process $\{X_t\}_{t \in T}$ is a collection of random variables X_t, taking values in a common measure space $(\Xi, \mathcal{X})$, indexed by a set T.*

That is, for each $t \in T$, $X_t(\omega)$ is an $\mathcal{F}/\mathcal{X}$-measurable function from Ω to Ξ, which induces a probability measure on Ξ in the usual way.

It's sometimes more convenient to write $X(t)$ in place of X_t. Also, when $S \subset T$, X_s or $X(S)$ refers to that sub-collection of random variables.

Example 2 (Random variables) *Any single random variable is a (trivial) stochastic process. (Take $T = \{1\}$, say.)*

Example 3 (Random vector) *Let* $T = \{1, 2, \ldots k\}$ *and* $\Xi = \mathbb{R}$*. Then* $\{X_t\}_{t \in T}$ *is a random vector in* $\mathbb{R}^k$*.*

Example 4 (One-sided random sequences) *Let* $T = \{1, 2, \ldots\}$ *and* Ξ *be some finite set (or* $\mathbb{R}$ *or* $\mathbb{C}$ *or* $\mathbb{R}^k \ldots$*). Then* $\{X_t\}_{t \in T}$ *is a one-sided discrete (real, complex, vector-valued, ...) random sequence. Most of the stochastic processes you have encountered are probably of this sort: Markov chains, discrete-parameter martingales, etc. Figures 1.2, 1.3, 1.4 and 1.5 illustrate some one-sided random sequences.*

Example 5 (Two-sided random sequences) *Let* $T = \mathbb{Z}$ *and* Ξ *be as in Example 4. Then* $\{X_t\}_{t \in T}$ *is a* two-*sided random sequence.*

Example 6 (Spatially-discrete random fields) *Let* $T = \mathbb{Z}^d$ *and* Ξ *be as in Example 4. Then* $\{X_t\}_{t \in T}$ *is a* d*-dimensional spatially-discrete random field.*

Example 7 (Continuous-time random processes) *Let* $T = \mathbb{R}$ *and* $\Xi = \mathbb{R}$*. Then* $\{X_t\}_{t \in T}$ *is a real-valued, continuous-time random process (or random motion or random signal). Figures 1.6 and 1.7 illustrate some of the possibilities.*

Vector-valued processes are an obvious generalization.

Example 8 (Random set functions) *Let* $T = \mathcal{B}$*, the Borel field on the reals, and* $\Xi = \overline{\mathbb{R}}^+$*, the non-negative extended reals. Then* $\{X_t\}_{t \in T}$ *is a random set function on the reals.*

The definition of random set functions on $\mathbb{R}^d$ is entirely parallel. Notice that if we want not just a set function, but a measure or a probability measure, this will imply various forms of dependence among the random variables in the collection, e.g., a measure must respect countable additivity over disjoint sets. We will return to this topic in the next section.

Example 9 (One-sided random sequences of set functions) *Let* $T = \mathcal{B} \times \mathbb{N}$ *and* $\Xi = \overline{\mathbb{R}}^+$*. Then* $\{X_t\}_{t \in T}$ *is a one-sided random sequence of set functions.*

Example 10 (Empirical distributions) *Suppose* Z_i*,* $= 1, 2, \ldots$ *are independent, identically-distributed real-valued random variables. (We can see from Example 4 that this is a one-sided real-valued random sequence.) For each Borel set* B *and each* n*, define*

$$\hat{P}_n(B) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_B(Z_i)$$

i.e., the fraction of the samples up to time n which fall into that set. This is the empirical measure. $\hat{P}_n(B)$ is a one-sided random sequence of set functions — in fact, of probability measures. We would like to be able to say something about how it behaves. It would be very reassuring, for instance, to be able to show that it converges to the common distribution of the Z_i (Figure 1.8).

1.2 Random Functions

$X(t, \omega)$ has two arguments, t and ω. For each fixed value of t, $X_t(\omega)$ is straightforward random variable. For each fixed value of ω, however, $X(t)$ is a function from T to Ξ — a *random function*. The advantage of the random function perspective is that it lets us consider the realizations of stochastic processes as single objects, rather than large collections. This isn't just tidier; we will need to talk about *relations* among the variables in the collection or their realizations, rather than just properties of individual variables, and this will help us do so. In Example 10, it's important that we've got random *probability measures*, rather than just random *set functions*, so we need to require that, e.g., $\hat{P}_n(A \cup B) = \hat{P}_n(A) + \hat{P}_n(B)$ when A and B are disjoint Borel sets, and this is a relationship among the three random variables $\hat{P}_n(A)$, $\hat{P}_n(B)$ and $\hat{P}_n(A \cup B)$. Plainly, working out all the dependencies involved here is going to get rather tedious, so we'd like a way to talk about acceptable realizations of the whole stochastic process. This is what the random functions notion will let us do.

We'll make this more precise by defining a random function as a function-valued random variable. To do this, we need a measure space of functions, and a measurable mapping from $(\Omega, \mathcal{F}, P)$ to that function space. To get a measure space, we need a carrier set and a σ-field on it. The natural set to use is Ξ^T, the set of all functions from T to Ξ. (We'll see how to restrict this to just the functions we want presently.) Now, how about the σ-field?

Definition 11 (Cylinder Set) *Given an index set T and a collection of σ-fields $\mathcal{X}_t$ on spaces Ξ_t, $t \in T$. Pick any $t \in T$ and any $A_t \in \mathcal{X}_t$. Then $A_t \times \prod_{s \neq t} \Xi_s$ is a* one-dimensional cylinder set*.*

For any finite k, $k-$dimensional cylinder sets are defined similarly, and clearly are the intersections of k different one-dimensional cylinder sets. To see why they have this name, notice a cylinder, in Euclidean geometry, consists of all the points where the x and y coordinates fall into a certain set (the base), leaving the z coordinate unconstrained. Similarly, a cylinder set like $A_t \times \prod_{s \neq t} \Xi_s$ consists of all the functions in Ξ^T where $f(t) \in A_t$, and are otherwise unconstrained.

Definition 12 (Product σ-field) *The* product σ-field*, $\otimes_{t \in T} \mathcal{X}_t$, is the σ-field over Ξ^T generated by all the one-dimensional cylinder sets. If all the $\mathcal{X}_t$ are the same, $\mathcal{X}$, we write the product σ-field as $\mathcal{X}^T$.*

The product σ-field is enough to let us define a random function, and is going to prove to be *almost* enough for all our purposes.

Definition 13 (Random Function; Sample Path) *A Ξ-valued random function on T is a map $X : \Omega \mapsto \Xi^T$ which is $\mathcal{F}/\mathcal{X}^T$-measurable. The realizations of X are functions $x(t)$ taking values in Ξ, called its* sample paths*.*

N.B., it has become common to apply the term "sample path" or even just "path" even in situations where the geometric analogy it suggests may be somewhat misleading. For instance, for the empirical distributions of Example 10, the "sample path" is the *measure* $\hat{P}_n$, not the curves shown in Figure 1.8.

Definition 14 (Functional of the Sample Path) *Let $E, \mathcal{E}$ be a measure-space. A* functional of the sample path *is a mapping $f : \Xi^T \mapsto E$ which is $\mathcal{X}^T/\mathcal{E}$-measurable.*

Examples of useful and common functionals include maxima, minima, sample averages, etc. Notice that none of these are functions of any one random variable, and in fact their value cannot be determined from any part of the sample path smaller than the whole thing.

Definition 15 (Projection Operator, Coordinate Map) *A* projection operator *or* coordinate map *π_t is a map from Ξ^T to Ξ such that $\pi_t X = X(t)$.*

The projection operators are a convenient device for recovering the individual coordinates — the random variables in the collection — from the random function. Obviously, as t ranges over T, $\pi_t X$ gives us a collection of random variables, i.e., a stochastic process in the sense of our first definition. The following lemma lets us go back and forth between the collection-of-variables, coordinate view, and the entire-function, sample-path view.

Theorem 16 (Product σ-field-measurability is equvialent to measurability of all coordinates) *X is $\mathcal{F}/\otimes_{t \in T} \mathcal{X}_t$-measurable iff $\pi_t X$ is $\mathcal{F}/\mathcal{X}_t$-measurable for every t.*

Proof: This follows from the fact that the one-dimensional cylinder sets generate the product σ-field. □

We have said before that we will want to constrain our stochastic processes to have certain properties — to be probability measures, rather than just set functions, or to be continuous, or twice differentiable, etc. Write the set of all such functions in Ξ^T as U. Notice that U does not have to be an element of the product σ-field, and in general is *not*. (We will consider some of the reasons for this later.) As usual, by $U \cap \mathcal{X}^T$ we will mean the collection of all sets of the form $U \cap C$, where $C \in \mathcal{X}^T$. Notice that $(U, U \cap \mathcal{X}^T)$ is a measure space. What we want is to ensure that the sample path of our random function lies in U.

Definition 17 (A Stochastic Process Is a Random Function) *A Ξ-valued stochastic process on T with paths in U, $U \subseteq \Xi^T$, is a random function $X : \Omega \mapsto U$ which is $\mathcal{F}/U \cap \mathcal{X}^T$-measurable.*

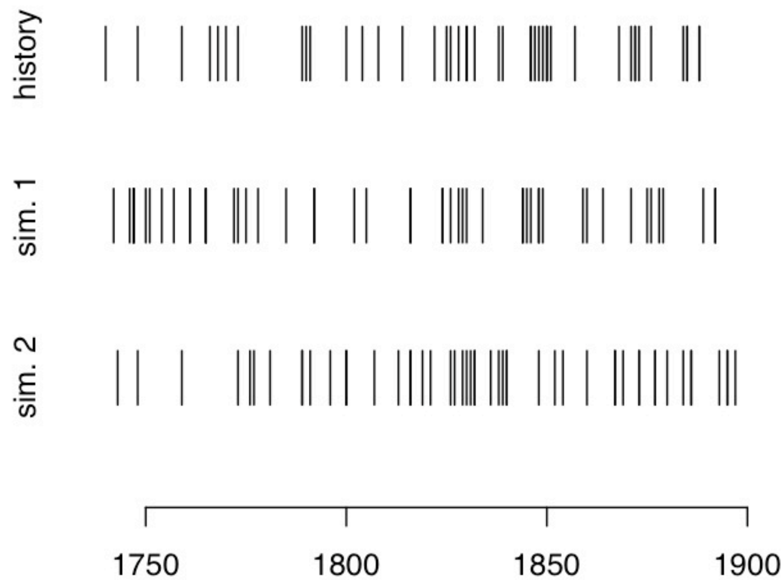

Figure 1.1: Examples of point processes. The top row shows the dates of appearances of 44 genres of English novels (data taken from Moretti (2005)). The bottom two rows show independent realizations of a Poisson process with the same mean time between arrivals as the actual history. The number of tick-marks falling within any measurable set on the horizontal axis determines an integer-valued set function, in fact a measure.

Corollary 18 (Measurability of constrained sample paths) *A function X from Ω to U is $\mathcal{F}/U \cap \mathcal{X}^T$-measurable iff X_t is $\mathcal{F}/\mathcal{X}$-measurable for all t.*

PROOF: Because $X(\omega) \in U$, $X(\omega)$ is $\mathcal{F}/U \cap \mathcal{X}^T$ iff it is $\mathcal{F}/\mathcal{X}^T$-measurable. Then apply Theorem 16. □

Example 19 (Random Measures) *Let $T = \mathcal{B}^d$, the Borel field on $\mathbb{R}^d$, and let $\Xi = \overline{\mathbb{R}}^+$, the non-negative extended reals. Then Ξ^T is the class of set functions on $\mathbb{R}^d$. Let M be the class of such set functions which are also measures (i.e., which are countably additive and give zero on the null set). Then a random set function X with realizations in M is a* random measure.

Example 20 (Point Processes) *Let X be a random measure, as in the previous example. If $X(B)$ is a finite integer for every bounded Borel set B, then X is a* point process. *If in addition $X(r) \leq 1$ for every $r \in \mathbb{R}^d$, then X is* simple. *The Poisson process is a simple point process. See Figure 1.1.*

Example 21 (Continuous random processes) *Let $T = \mathbb{R}^+$, $\Xi = \mathbb{R}^d$, and $\mathbf{C}(T)$ the class of continuous functions from T to Ξ (in the usual topology). Then a Ξ-valued random process on T with paths in $\mathbf{C}(T)$ is a* continuous random process*. The Wiener process, or Brownian motion, is an example. We will see that most sample paths in $\mathbf{C}(T)$ are not differentiable.*

1.3 Exercises

Exercise 1.1 (The product σ-field answers countable questions) *Let $\mathcal{D} = \bigcup_S \mathcal{X}^S$, where the union ranges over all countable subsets S of the index set T. For any event $D \in \mathcal{D}$, whether or not a sample path $x \in D$ depends on the value of x_t at only a countable number of indices t.*

(a) Show that $\mathcal{D}$ is a σ-field.

(b) Show that if $A \in \mathcal{X}^T$, then $A \in \mathcal{X}^S$ for some countable subset S of T.

```
AATGAAATAAAAAAAAACGAAAATAAAAAA
AAGGCCATTAAAGTTAAAATAATGAAAGGA
CAATGATTAGGACAATAACATACAAGTTAT
GGGGTTAATTAATGGTTAGGATGGGTTTTT
CCTTCAAAGTTAATGAAAAGTTAAAATTTA
TAAGTATTTGAAGCACAGCAACAACTAGGT
```

Figure 1.2: Examples of one-sided random sequences ($\Xi = \{\mathrm{A}, \mathrm{C}, \mathrm{G}, \mathrm{T}\}$). The top line shows the first thirty bases of the first chromosome of the cellular slime mold *Dictyostelium discoideum* (Eichinger *et al.*, 2005), downloaded from dictybase.org. The lower lines are independent samples from a simple Markov model fitted to the full chromosome.

```
111111111111110110011011000001
011110110111111011110110110110
011110001101101111111111110000
011011011011111101111000111100
011011111101101100001111110111
000111111111111001101100011011
```

Figure 1.3: Examples of one-sided random sequences. These binary sequences ($\Xi = \{0, 1\}$) are the first thirty steps from samples of a *sofic process*, one with a finite number of underlying states which is nonetheless not a Markov chain of any finite order. Can you figure out the rule specifying which sequences are allowed and which are forbidden? (Hint: these are all samples from the *even* process.)

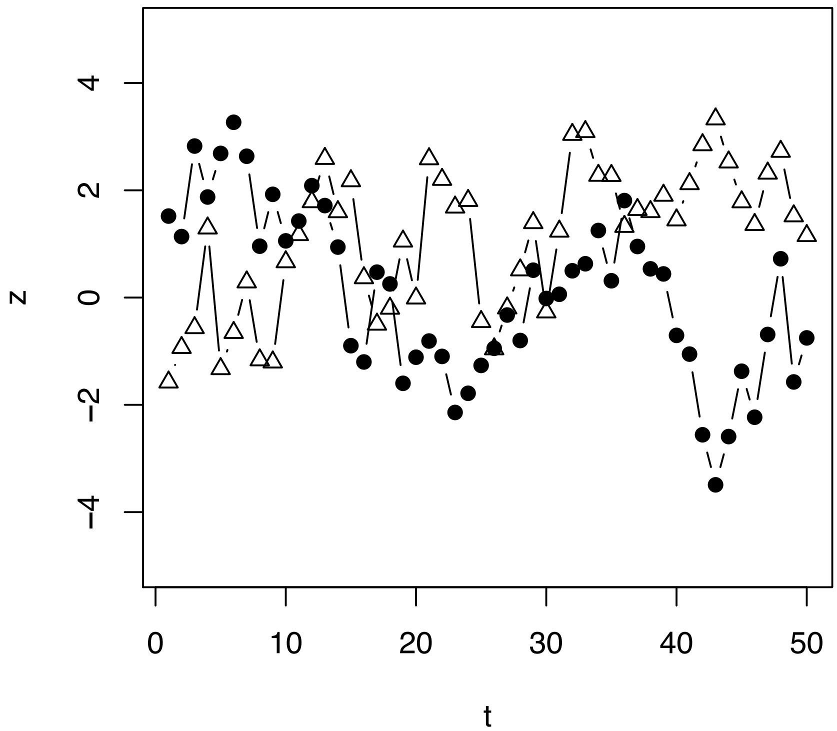

Figure 1.4: Examples of one-sided random sequences. These are linear Gaussian random sequences, $X_{t+1} = 0.8X_t + Z_{t+1}$, where the Z_t are all i.i.d. $\mathcal{N}(0, 1)$, and $X_1 = Z_0$. Different shapes of dots represent different independent samples of this *autoregressive* process. (The line segements are simply guides to the eye.) This is a Markov sequence, but one with a continuous state-space.

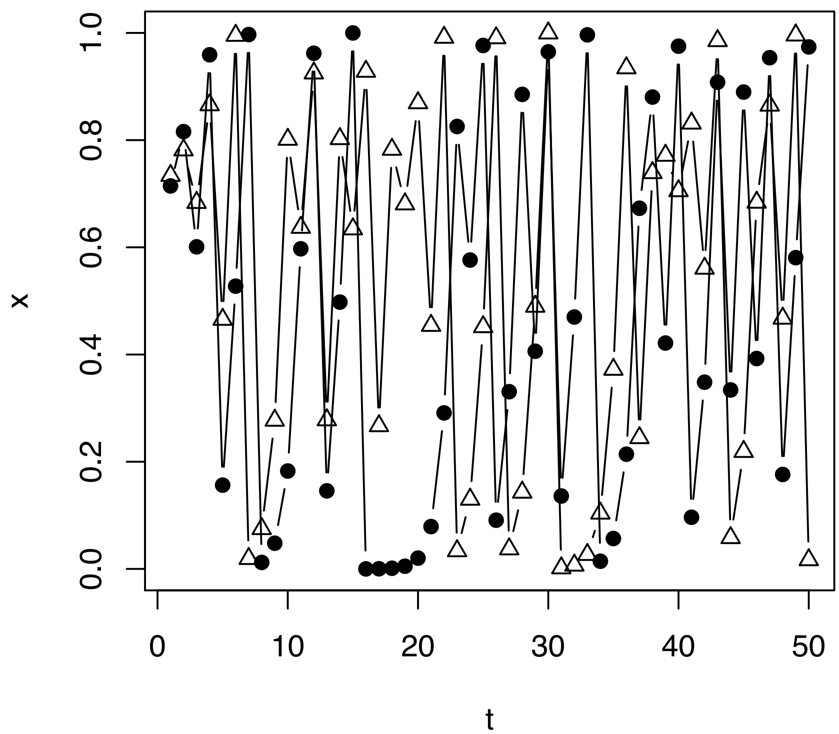

Figure 1.5: Nonlinear, non-Gaussian random sequences. Here $X_1 \sim U(0,1)$, i.e., uniformly distributed on the unit interval, and $X_{t+1} = 4X_t(1-X_t)$. Notice that while the two samples begin very close together, they rapidly separate; after a few time-steps their locations are, in fact, effectively independent. We will study both this approach to independence, known as *mixing*, and this example, known as the *logistic map*, in some detail.

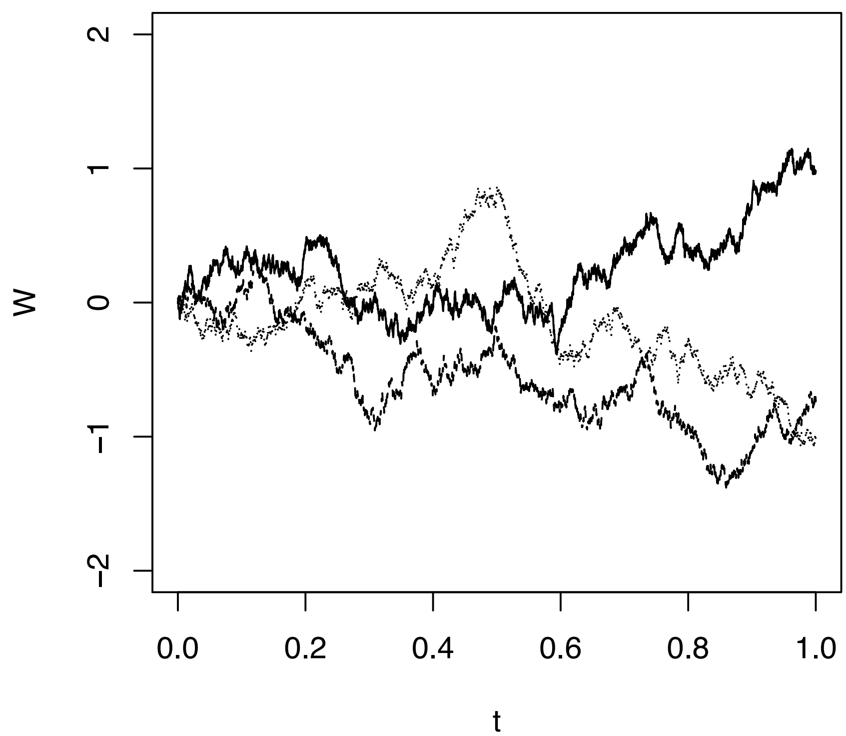

Figure 1.6: Continuous-time random processes. Shown are three samples from the standard Wiener process, also known as Brownian motion, a Gaussian process with independent increments and continuous trajectories. This is a central part of the course, and actually what forced probability to be re-defined in terms of measure theory.

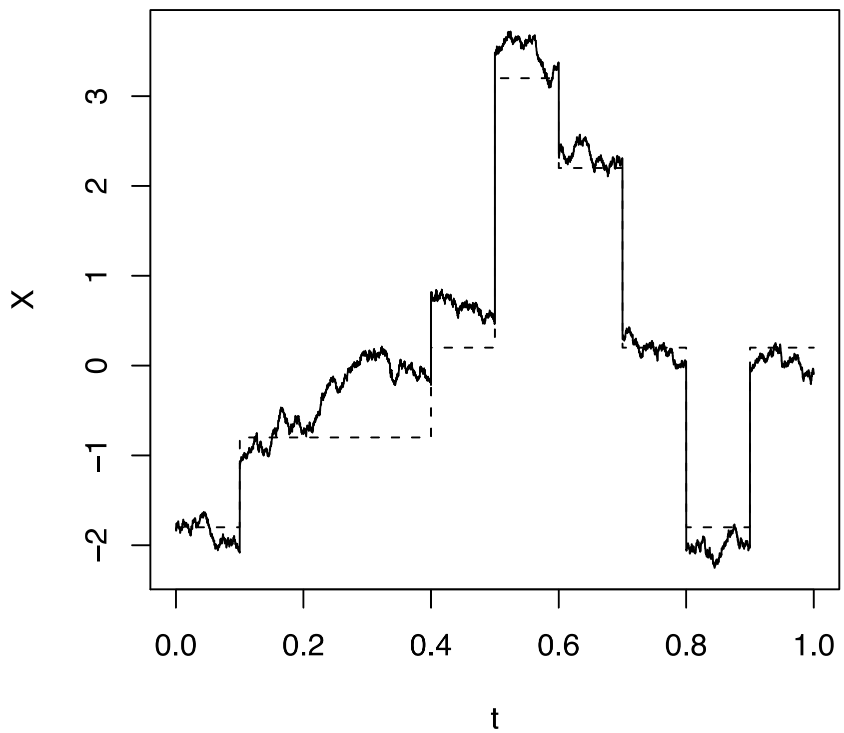

Figure 1.7: Continuous-time random processes can have discontinuous trajectories. Here $X_t = W_t + J_t$, where W_t is a standard Wiener process, and J_t is piecewise constant (shown by the dashed lines), changing at $t = 0.1, 0.2, 0.3, \ldots 1.0$. The trajectory is discontinuous at $t = 0.4$, but continuous from the right there, and there is a limit from the left. In fact, at every point the trajectory is continuous from the right and has a limit from the left. We will see many such *cadlag* processeses.

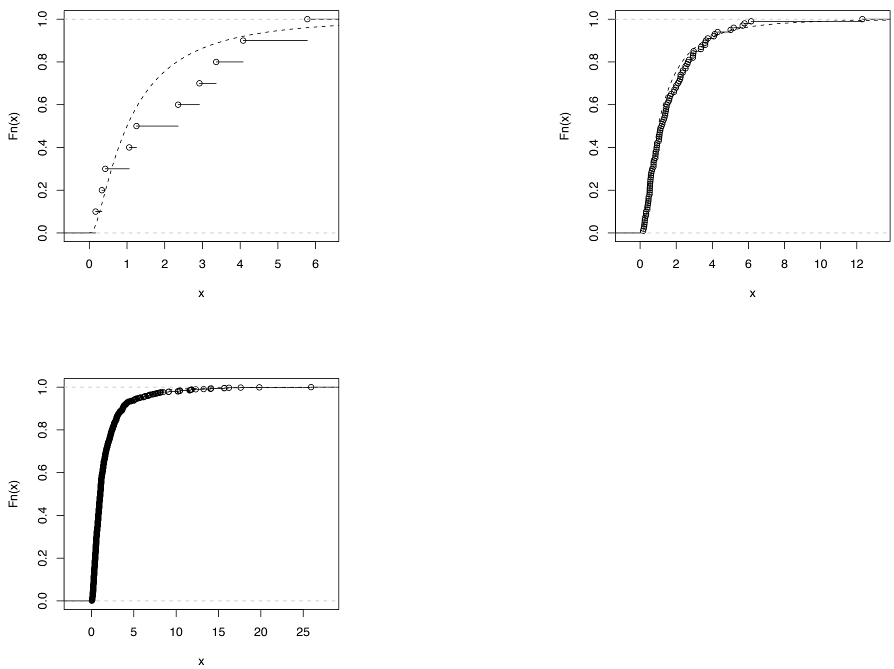

Figure 1.8: Empirical cumulative distribution functions for successively larger samples from the standard log-normal distribution (from left to right, $n = 10, 100, 1000$), with the theoretical CDF as the smooth dashed line. Because the intervals of the form $(-\infty, a]$ are a generating class for the Borel σ-field, the empirical C.D.F. suffices to represent the empirical distribution.