1

A few tools from probability theory

1 Introduction

In this book we do not need much probability theory. All that we need was pretty well known around 1930. For the moment, let us explain the Steinhaus approach to the theory.

For Steinhaus, the *probability space* is simply the interval $[0, 1]$ of the real line. A *random variable* is a function defined on $[0, 1]$. An *event* is a measurable set on $[0, 1]$. The *probability* of an event is its Lebesgue measure. The *expectation* of a random variable (when existing) is its Lebesgue integral.

Let us consider the binary expansion

$$\omega = \sum_{n=1}^{\infty} \beta_n 2^{-n},$$

where $\omega \in [0, 1]$, $\beta_n = \beta_n(\omega) = 0$ or 1, and $\sum_1^\infty \beta_n = \infty$ except if $\omega = 0$. In a natural sense, the β_n are mutually independent random variables, and each β_n takes the values 0 and 1 with the same probability $\frac{1}{2}$. Now let us define

$$\omega_j = \omega_j(\omega) = \sum_{n \in N} \beta_{m(n,j)} 2^{-n},$$

where $\mathbb{N} = \{1, 2, 3, \ldots\}$ and m is a one-to-one mapping from $\mathbb{N}^2$ into $\mathbb{N}$. In a natural sense again, the ω_j are mutually independent random variables. Now each of them is equidistributed on $[0, 1]$, i.e., given a subinterval I of $[0, 1]$, the probability of the ω-set where $\omega_j(\omega) \in I$ is the length of I.

The β_n are closely related to the Rademacher functions. If we write

$$\varepsilon_n = 1 - 2\beta_n,$$

$\varepsilon_n(\omega)$ is nothing but the nth Rademacher function except on a finite set of binary points.

The ω_j are often called the Steinhaus functions.

The reader may keep this model in mind, when we speak of a 'Rademacher sequence' or a 'Steinhaus sequence'. These objects will be defined later; but nothing is lost if we understand $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots$ instead of any Rademacher sequence, and $\omega_1, \omega_2, \ldots, \omega_j, \ldots$ instead of any Steinhaus sequence.

Nevertheless, it will be more convenient to introduce a slightly different model and to use the language of modern probability theory. From now on, our main references are the classical books of Saks [182] and Kolmogorov [131], together with Loève [141] and Meyer [157].

2 The basic notions

Let us define a probability space $(\Omega, \mathscr{A}, \mathscr{P})$. Ω is a set.

$\mathscr{A}$ is a σ-field on Ω. That is, $\mathscr{A}$ consists of subsets of Ω, it contains $\varnothing$ (the empty set), and is closed under the operations of complementation and of taking countable unions.

$\mathscr{P}$ is a probability on $(\Omega, \mathscr{A})$, in other words, a positive measure of total mass 1. That is, $\mathscr{P}(A)$ is defined for each $A \in \mathscr{A}$, $\mathscr{P}(A) \in [0, 1]$, $\mathscr{P}(\Omega) = 1$ and $\mathscr{P}(\bigcup A_n) = \sum \mathscr{P}(A_n)$ for any countable set of mutually disjoint $A_n (A_n \in \mathscr{A})$.

Moreover, $\mathscr{P}$ is complete. That is, $A \in \mathscr{A}$ whenever A is included in a set $B \in \mathscr{A}$ such that $\mathscr{P}(B) = 0$. Completeness is not always assumed in probability theory, but it is convenient for our purpose.

When these conditions are fulfilled, $(\Omega, \mathscr{A}, \mathscr{P})$ is called a *probability space*. Sometimes we say that Ω itself is a probability space (for example, when we speak of an element ω of the probability space), then it is always assumed that $\mathscr{A}$ and $\mathscr{P}$ have been defined. The elements of $\mathscr{A}$ are called *events*, and $\mathscr{P}(A)$ is the *probability* of A. If $\mathscr{P}(A) = 1$, A is said to be *almost sure*.

Suppose that we are given a probability space $(\Omega, \mathscr{A}, \mathscr{P})$, and a set E. A *random element of* E, or *random object*, is a mapping X from Ω into E. Then, for each $\omega \in \Omega$, we have $X(\omega) \in E$. The usual problem is to consider a subset B of E, and to ask what is the probability of the ω-set such that $X(\omega) \in B$. Of course, we must suppose (or prove, in a concrete case) that $X^{-1}(B) \in \mathscr{A}$. Then, instead of $X^{-1}(B)$, or $\{\omega : X(\omega) \in B\}$, we simply write $\{X \in B\}$, and we can speak of the event: 'X belongs to B'. When $X \in B$ is defined by a relation $\mathscr{R}(X)$, we write $\{\mathscr{R}(X)\}$ instead of $\{X \in B\}$, and speak of the event 'X satisfies $\mathscr{R}(X)$'. If this event is almost sure, we say 'X satisfies $\mathscr{R}(X)$ almost surely' and write: $\mathscr{R}(X)$ a. s.

As an example, let us take $E = \{0, 1\}$, $A \in \mathscr{A}$, and $X = 1_A$ the mapping

defined by

$$1_A(\omega)=1 \quad \text{if} \quad \omega \in A$$
$$1_A(\omega)=0 \quad \text{if} \quad \omega \notin A. \tag{1}$$

Then '$1_A=1$' is another way to write the event A.

3 Distribution and similarity

A very important case is when E is equal to $\mathbb{R}$, the real line. We say that a mapping X from Ω to $\mathbb{R}$ is a *real random variable* if $X^{-1}(I) \in \mathscr{A}$ for each interval I of the real line. More generally, if E is a topological space, we say that a mapping X from Ω to E is a *random variable in* E if $X^{-1}(G) \in \mathscr{A}$ for each open set G in E. Particular cases are real random variables, complex random variables, random variables in $\mathbb{R}^n$, and random vectors in a linear topological space. For each element B of the σ-field $\mathscr{B}$ generated by the open sets in E (the '*Borel field*' of E), $X^{-1}(B)$ belongs to $\mathscr{A}$. The mapping $B \to \mathscr{P}(X^{-1}(B))$ defines a positive measure μ_X of total mass 1 on $(E, \mathscr{B})$, which is called the *distribution* of X. In particular, the distribution of a real random variable X is the measure generated by the increasing function

$$\mu_X(x)=\mathscr{P}(X \in]-\infty, x]).$$

If μ_X is carried on an interval I and μ_X is proportional to Lebesgue measure on I, we say that X is *equidistributed* on I.

Two random variables in a topological space E which have the same distribution are said to be *equally distributed*.

In the general case, we are given a random object X, which maps Ω into E, a set without any topological structure. In a natural way, X carries the structure of a probability space from Ω over to E. Namely, $\mathscr{B}_X$ will consist of those subsets B of E such that $X^{-1}(B) \in \mathscr{A}$, and for each $B \in \mathscr{B}_X$ we define

$$\mu_X(B)=\mathscr{P}(X^{-1}(B)).$$

It is very easy to check that $(E, \mathscr{B}_X, \mu_X)$ is a probability space. Given two random objects X_1 and X_2, taking values in E and defined on two probability spaces (or on the same one), we say that X_1 and X_2 are *similar* if $(E, \mathscr{B}_{X_1}, \mu_{X_1})$ and $(E, \mathscr{B}_{X_2}, \mu_{X_2})$ are the same probability space.

If E is a topological space, the random object X is a random variable if and only if the Borel field $\mathscr{B}$ is contained in $\mathscr{B}_X$.

If E is a topological space and X_1 and X_2 are equally distributed random variables in E, they are similar. Conversely, if X_1 and X_2 are similar random variables, they are equally distributed.

From now on, we can speak of similar random vectors in a linear space, or similar random sequences, or similar random series, without assuming that any topology is given. As a typical example, the two series

$$\sum_{n=1}^{\infty} \varepsilon_n u_n, \quad \sum_{n=1}^{\infty} (-1)^n \varepsilon_n u_n,$$

where ε_n is defined as on p. 1, and the u_n are arbitrary vectors in a linear space, are similar random series.

4 Product probability space

Given two probability spaces $(\Omega_1, \mathscr{A}_1, \mathscr{P}_1)$ and $(\Omega_2, \mathscr{A}_2, \mathscr{P}_2)$, one can prove that there exists a smallest probability space $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2, \mathscr{P}_1 \times \mathscr{P}_2)$ with the following properties:

$\Omega_1 \times \Omega_2$ is the ordinary product set of Ω_1 and Ω_2,

$\mathscr{A}_1 \times \mathscr{A}_2$ is a σ-field on $\Omega_1 \times \Omega_2$ which contains the products $A_1 \times A_2$ ($A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$),

$\mathscr{P}_1 \times \mathscr{P}_2$ is a complete probability on $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2)$, such that $(\mathscr{P}_1 \times \mathscr{P}_2)(A_1 \times A_2) = \mathscr{P}_1(A_1)\mathscr{P}_2(A_2)$ ($A_1 \in \mathscr{A}_1$, $A_2 \in \mathscr{A}_2$).

$(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2, \mathscr{P}_1 \times \mathscr{P}_2)$ is called the *product probability space* of $(\Omega_1, \mathscr{A}_1, \mathscr{P}_1)$ and $(\Omega_2, \mathscr{A}_2, \mathscr{P}_2)$; it will be denoted often by $\Omega_1 \times \Omega_2$.

Given a finite or infinite sequence of probability spaces $(\Omega_n, \mathscr{A}_n, \mathscr{P}_n)$, one can prove again that there exists a smallest probability space $(\prod \Omega_n, \prod \mathscr{A}_n, \prod \mathscr{P}_n)$ with the following properties:

$\prod \Omega_n$ is the ordinary product set of the Ω_ns,

$\prod \mathscr{A}_n$ is a σ-field which contains all products $\prod A_n$ obtained by taking $A_n \in \mathscr{A}_n$ with $A_n = \Omega_n$ for all but a finite number of values of n,

$\prod \mathscr{P}_n$ is a complete probability on $(\prod \Omega_n, \prod \mathscr{A}_n)$ such that $(\prod \mathscr{P}_n)(\prod A_n) = \prod \mathscr{P}_n(A_n)$ for all products $\prod A_n$ of the form mentioned above.

We say that $(\prod \Omega_n, \prod \mathscr{A}_n, \prod \mathscr{P}_n)$ is the product of the probability spaces $(\Omega_n, \mathscr{A}_n, \mathscr{P}_n)$, or, simply, that $\prod \Omega_n$ is the *product probability space* of the Ω_n.

5 The standard model; independence; Steinhaus and Rademacher sequences

For our purpose the most convenient probability space will be the product space $\Omega = \prod_{n=1}^{\infty} \Omega_n$, where, for each n, Ω_n is the interval $[0, 1]$, $\mathscr{A}_n$ is the σ-field of Lebesgue measurable sets, and $\mathscr{P}_n$ is the Lebesgue measure. It will be called the *standard probability space*. An element of Ω

is denoted by $\omega=(\omega_1, \omega_2, \ldots, \omega_n, \ldots)$, where $0 \leqslant \omega_n \leqslant 1$ for all n. The sequence $\omega_1, \omega_2, \ldots, \omega_n, \ldots$, considered as a sequence of random variables defined on Ω, will be called the standard Steinhaus sequence. The sequence $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots$ defined by

$$\varepsilon_n(\omega) = \quad 1 \quad \text{if} \quad \omega_n \in [0, \tfrac{1}{2}[,$$
$$\varepsilon_n(\omega) = -1 \quad \text{if} \quad \omega_n \in [\tfrac{1}{2}, 1[$$

$(n=1, 2, \ldots)$ will be called the standard Rademacher sequence. More generally, let us consider an arbitrary partition of the integers into disjoint classes $N_1, N_2, \ldots, N_j, \ldots$ and let us write ω_{N_j} for $\{\omega_n\}_{n \in N_j}$. Any (finite or infinite) sequence of random objects $X_1, X_2, \ldots, X_j, \ldots$ which are defined on Ω and such that, for each j, X_j depends only on ω_{N_j}, is called a standard sequence of independent objects.

Given any other probability space Ω^* and a finite or infinite sequence of random objects Y_n defined on Ω^*, we shall say that the Y_n are *independent* (=mutually independent) if there exists a standard sequence of independent objects which is similar to the sequence Y_n. As a consequence, if $Y_1, Y_2, \ldots, Y_n, \ldots$ are independent random variables with values in topological spaces $E_1, E_2, \ldots, E_n, \ldots$, then

$$\mathscr{P}(Y_1 \in B_1, Y_2 \in B_2, \ldots, Y_n \in B_n)$$
$$= \mathscr{P}(Y_1 \in B_1)\mathscr{P}(Y_2 \in B_2) \ldots \mathscr{P}(Y_n \in B_n)$$

for all n and all $B_j \in \mathscr{B}_j$ (Borel field on E_j). Conversely, if $Y_1, Y_2, \ldots, Y_n$ are real or complex random variables and satisfy all the preceding equalities, they are independent.

A sequence similar to the standard Steinhaus – or Rademacher – sequence will be called a *Steinhaus sequence* or a *Rademacher sequence*. In other words, a Steinhaus sequence is a sequence of independent random variables equidistributed on $[0, 1]$, and a Rademacher sequence is a sequence of independent random variables taking the values $+1$ or -1 with the same probability $\frac{1}{2}$.

6 Integration: the main tools

Given a probability space $(\Omega, \mathscr{A}, \mathscr{P})$, it is possible to apply the usual theory of the Lebesgue integral to it. If X is a real random variable on Ω which is Lebesgue integrable with respect to $\mathscr{P}$, we write $X \in L^1(\Omega)$, and the integral

$$\mathscr{E}(X) = \int_\Omega X(\omega)\mathscr{P}(\mathrm{d}\omega)$$

is called the *expectation* of X. In particular when 1_A is defined as in §2,

$\mathscr{E}(1_A)=\mathscr{P}(A)$. Another expression for $\mathscr{E}(X)$ is the Stieltjes integral

$$\mathscr{E}(X)=\int_{-\infty}^{\infty} x\,\mathrm{d}\mu_X(x)=\int_{\mathbb{R}} x\mu_X(\mathrm{d}x).$$

$\mathscr{E}(X)$ does not change if we replace X by a similar random variable.

If $X\in L^p(\Omega)$, that is $X^p\in L^1(\Omega)$, we define the pth *moment* of X as $\mathscr{E}(X^p)$. If $X\in L^2(\Omega)$, X is called a real random variable *of the second order*; we define the *variance* of X as

$$V(X)=\mathscr{E}(|X-\mathscr{E}(X)|^2).$$

For every real random variable X,

$$\varphi_X(u)=\mathscr{E}(\mathrm{e}^{\mathrm{i}uX})=\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}ux}\,\mathrm{d}\mu_X(x)$$

exists for all real u. φ_X is the *characteristic function* of X. An elementary result from the theory of Fourier–Stieltjes transforms shows that $\varphi_X=\varphi_{X'}$ if and only if X and X' are similar.

The definitions of $L^p(\Omega)$, $\mathscr{E}(X)$, $V(X)$ extend to complex random variables.

Here are the main theorems on integration.

The Beppo–Levi theorem

Suppose $X_n\in L^1(\Omega)$, $X_n\geqslant 0$, *and* $\sum_1^\infty \mathscr{E}(X_n)<\infty$. *Then* $\mathscr{E}(\sum_1^\infty X_n)=\sum_1^\infty \mathscr{E}(X_n)$. *As a consequence,* $\sum_1^\infty X_n<\infty$ *a. s.*

The Fubini–Jessen theorem

Suppose $\Omega=\prod_1^\infty \Omega_n$ *and* $X\in L^1(\Omega)$. *Then*

$$\mathscr{E}(X)=\lim_{n\to\infty}\int_{\Omega_n}\cdots\int_{\Omega_2}\int_{\Omega_1} X(\omega)\mathscr{P}_1(\mathrm{d}\omega)\mathscr{P}_2(\mathrm{d}\omega)\ldots\mathscr{P}_n(\mathrm{d}\omega)\quad a.s.$$

In particular, if Ω *is the standard probability space,*

$$\mathscr{E}(X)=\lim_{n\to\infty}\int_0^1\cdots\int_0^1\int_0^1 X(\omega_1,\omega_2,\ldots,\omega_n,\omega_{n+1},\ldots)\times\mathrm{d}\omega_1\,\mathrm{d}\omega_2\ldots\mathrm{d}\omega_n\quad a.s.$$

In the case where X is a product of independent random variables, we have a simple formula.

If $X_1,\ldots,X_n,\ldots$ *is a finite sequence of independent random variables which belong to* $L^1(\Omega)$ *and if* $\prod|X_n|\in L^1(\Omega)$ *we have*

$$\mathscr{E}(\prod X_n)=\prod\mathscr{E}(X_n).$$

From these theorems we may derive two important results, usually called the Borel–Cantelli lemma and the zero-one law.

Let us consider an infinite sequence of events $A_1, A_2, \ldots, A_n, \ldots$. The event $\overline{\lim}\, A_n = \bigcap_k \bigcup_{n \geq k} A_n$ holds when infinitely many A_n hold.

The Borel–Cantelli lemma

If $\sum_1^\infty \mathscr{P}(A_n) < \infty$, *then* $\mathscr{P}(\overline{\lim}\, A_n) = 0$. *If the events* $A_1, A_2, \ldots, A_n, \ldots$ *are independent and if* $\sum_1^\infty \mathscr{P}(A_n) = \infty$, *then* $\mathscr{P}(\overline{\lim}\, A_n) = 1$.

Proof. If $\sum_1^\infty \mathscr{E}(1_{A_n}) < \infty$, then $\sum_1^\infty 1_{A_n} < \infty$ a. s. If the 1_{A_n} are independent and $\prod_N^\infty (1 - \mathscr{E}(1_{A_n})) = 0$ for each N, then $\prod_N^\infty (1 - 1_{A_n}) = 0$ for each N a. s.

Now we consider a sequence of independent random objects X_n and an event A which does not depend on the values of any finite number of X_n (in other words, A is an asymptotic property of the X_n). The zero-one law says that the probability of such an event is necessarily 0 or 1. It can be stated as follows.

The zero-one law

Let $\Omega = \prod \Omega_n$, *and let* X *be a real random variable defined on* Ω, *such that*

$$X(\omega_1, \omega_2, \ldots, \omega_n, \omega_{n+1}, \ldots) = X(\omega'_1, \omega'_2, \ldots, \omega'_n, \omega_{n+1}, \ldots)$$

whatever $\omega_1, \omega_2, \ldots, \omega_n, \omega'_1, \omega'_2, \ldots, \omega'_n, \omega_{n+1}, \ldots$ *may be. Then* X *is similar to a constant. In particular, if* $X = 1_A$, *then* $\mathscr{P}(A) = 0$ *or* $\mathscr{P}(A) = 1$.

Proof. Let us write

$$\mathscr{E}_n(X) = \int_{\Omega_n} \cdots \int_{\Omega_2} \int_{\Omega_1} X(\omega) \mathscr{P}_1(\mathrm{d}\omega) \mathscr{P}_2(\mathrm{d}\omega) \ldots \mathscr{P}_n(\mathrm{d}\omega).$$

In the particular case that $X = 1_A$, $\mathscr{E}_n(X)\mathscr{E}_n(1 - X) = 0$ for each n; therefore $\mathscr{E}(X)\mathscr{E}(1 - X) = \mathscr{P}(A)(1 - \mathscr{P}(A)) = 0$. In the general case, given a real number r, $\mathscr{P}(X > r)$ is 0 or 1 and therefore X is similar to a constant.

Let us give now two elementary inequalities. We denote by X a positive random variable. The first inequality is obvious.

Inequality I

If $X \in L^1(\Omega)$ *and* $a > 1$,

$$\mathscr{P}(X \geq a\mathscr{E}(X)) \leq \frac{1}{a}.$$

As an application, let Y be a random variable of the second order. Then

$$\mathscr{P}(|Y - \mathscr{E}(Y)| \geq a\sqrt{V(Y)}) \leq \frac{1}{a^2}$$

This is called the *Bienaymé inequality*.

Inequality II
If $X \in L^2(\Omega)$ *and* $0<\lambda<1$,

$$\mathscr{P}(X \geqslant \lambda \mathscr{E}(X)) \geqslant (1-\lambda)^2 \frac{\mathscr{E}^2(X)}{\mathscr{E}(X^2)}.$$

Proof. We define

$$X'(\omega)=X(\omega) \quad \text{if} \quad X(\omega)\geqslant \lambda\mathscr{E}(X)$$
$$X'(\omega)=0 \quad \text{if} \quad X(\omega)<\lambda\mathscr{E}(X).$$

By the Schwarz inequality we have

$$\mathscr{E}^2(X') \leqslant \mathscr{E}(X'^2)\mathscr{P}(X' \neq 0)$$
$$\leqslant \mathscr{E}(X^2)\mathscr{P}(X \geqslant \lambda\mathscr{E}(X)).$$

Moreover, $\mathscr{E}(X) \leqslant \mathscr{E}(X') + \lambda\mathscr{E}(X)$. Therefore

$$(1-\lambda)^2\mathscr{E}^2(X) \leqslant \mathscr{E}(X^2)\mathscr{P}(X \geqslant \lambda\mathscr{E}(X)).$$

The main application is a formula which will be proved in a more general form in chapter 3, namely

$$\mathscr{P}\left(\left|\sum_{n=1}^{N} \varepsilon_n a_n\right|^2 > \lambda \sum_{n=1}^{N} |a_n|^2\right) > \tfrac{1}{3}(1-\lambda)^2,$$

where the a_n are complex numbers and $\{\varepsilon_n\}$ is a Rademacher sequence. This formula plays a fundamental role in the work of Paley and Zygmund [166].

7 Symmetric random vectors

Let us take a random vector X in a linear space. We say that X is a *symmetric random vector* if $-X$ and X are similar. In particular, we shall speak of *symmetric random variables* with real or complex values. For example, a Rademacher variable $\varepsilon(\mathscr{P}(\varepsilon=1)=\frac{1}{2}, \mathscr{P}(\varepsilon=-1)=\frac{1}{2})$ is symmetric. If Ω is the standard probability space, $e^{2\pi i\omega_j}$ is symmetric for all j, because $-e^{2\pi i\omega_j} = e^{2\pi i(\omega_j+1/2)}$.

Very frequently we shall study sequences of independent symmetric random vectors $X_1, X_2, \ldots, X_n, \ldots$. Here are some preliminary remarks.

If $X_1, X_2, \ldots, X_n, \ldots$ *are independent symmetric random vectors and* $\varepsilon_1^*, \varepsilon_2^*, \ldots, \varepsilon_n^*, \ldots$ *is a fixed sequence with values* $+1$ *or* -1, *the random sequences* $\{X_n\}$ *and* $\{\varepsilon_n^* X_n\}$ *are similar.*

If $X_1, X_2, \ldots, X_n, \ldots$ *are independent symmetric random vectors and* $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots$ *a Rademacher sequence, independent from the* X_n, *the random sequences* $\{X_n\}$ *and* $\{\varepsilon_n X_n\}$ *are similar.*

As an application, suppose we consider a property P which can be satisfied or not by a sequence of vectors (such as convergence, sum-

mability, ...). *Suppose that, whenever* $\{\varepsilon_n\}$ *is a Rademacher sequence and* $\{x_n\}$ *is a constant sequence of vectors,* $\{\varepsilon_n x_n\}$ *satisfies* P *almost surely. Suppose, moreover, that '*$\{X_n\}$ *satisfies* P*' is an event. Then this event is almost sure.*

This last remark explains the importance of Rademacher sequences through this book. We call it the *principle of reduction.*

Given a random vector X defined on Ω, the random vector $Y(\omega, \omega') = X(\omega) - X(\omega')$ defined on the product space $\Omega \times \Omega$ is obviously a symmetric random vector. Sometimes we are able to prove that Y enjoys a given property a. s. (in $\Omega \times \Omega$); it follows that there exists $\omega' \in \Omega$ and a fixed vector $x = X(\omega')$ such that $X - x$ enjoys the same property a. s. (in Ω). This is the *device of symmetrization.*

8 Random functions and analytic sets

The general concept of a random function is the following. We are given a probability space Ω, a set E and a set F. *A random function defined on* E *with values in* F is a random element of the set F^E (which consists of all mappings from E to F). In the same way we can define a random measure, a random distribution, a random series of functions, and so on.

A random function Φ defined on E with values in F can be considered as a mapping from $E \times \Omega$ to F; $y = \Phi(x, \omega)$, say. For each $\omega \in \Omega$ we have a function from E to F, which we denote by Φ, and for $x \in E$ a random element in F, which we denote by $\Phi(x)$. The most interesting case is when E is a measure space, with a complete measure μ, because $E \times \Omega$ is again a measure space with the complete measure $\mu \times \mathscr{P}$. If P is a subset of $E \times \Omega$, it is equivalent to say that 'for almost every ω, we have $(x, \omega) \in P$ for almost every x' or 'for almost every x, we have $(x, \omega) \in P$ for almost every ω'; we can say without ambiguity that a given property holds almost surely almost everywhere, or that it holds almost everywhere almost surely. But it is not the same to say that 'for almost every ω, $(x, \omega) \in P$ for every x' or 'for every x, $(x, \omega) \in P$ for almost every ω'; therefore an expression such as 'almost surely everywhere' does not mean the same as 'everywhere almost surely'.

The above remarks apply to random real functions of a real variable, or to random series of functions. If we are interested in properties such as continuity, differentiability, non-differentiability, convergence, or divergence, the probability of such a property can be studied at a given point, or almost everywhere, or everywhere. If the property holds at every given point almost surely, a. s. again it holds almost everywhere, but not

necessarily everywhere. Actually it is generally much more difficult to study non-differentiability everywhere, or convergence everywhere, or divergence everywhere, than the corresponding questions almost everywhere.

Here we emphasize the role of *analytic sets* in the sense of Lusin. If Ω is the standard probability space, an analytic set on Ω is defined as the projection *on* Ω of a borelian set in $\Omega \times [0, 1]$. Such a set is known to be Lebesgue measurable. Suppose that a random function $f(\omega, t)$ is defined on $\Omega \times [0, 1]$ and the (ω, t)-set where it satisfies a given property (P) is a borelian set. Then the ω-set where $f(\omega, t)$ satisfies (P) for some t is analytic; therefore it is Lebesgue measurable, i.e., we may speak of the *event*: $\{f(\omega, t)$ satisfies (P) for some $t\}$. In the same way, we may speak of the event: $\{f(\omega, t)$ satisfies (P) for all $t\}$.