1
Review of Probability

In this chapter we shall recall some basic notions and facts from probability theory. Here is a short list of what needs to be reviewed:

1) Probability spaces, σ-fields and measures;
2) Random variables and their distributions;
3) Expectation and variance;
4) The σ-field generated by a random variable;
5) Independence, conditional probability.

The reader is advised to consult a book on probability for more information.

1.1 Events and Probability

Definition 1.1

Let Ω be a non-empty set. A *σ-field* $\mathcal{F}$ on Ω is a family of subsets of Ω such that

1) the empty set $\emptyset$ belongs to $\mathcal{F}$;
2) if A belongs to $\mathcal{F}$, then so does the complement $\Omega \setminus A$;

3) if $A_1, A_2, \ldots$ is a sequence of sets in $\mathcal{F}$, then their union $A_1 \cup A_2 \cup \cdots$ also belongs to $\mathcal{F}$.

Example 1.1

Throughout this course $\mathbb{R}$ will denote the set of real numbers. The family of *Borel sets* $\mathcal{F} = \mathcal{B}(\mathbb{R})$ is a σ-field on $\mathbb{R}$. We recall that $\mathcal{B}(\mathbb{R})$ is the smallest σ-field containing all intervals in $\mathbb{R}$.

Definition 1.2

Let $\mathcal{F}$ be a σ-field on Ω. A *probability measure* P is a function

$$P : \mathcal{F} \to [0,1]$$

such that

1) $P(\Omega) = 1$;

2) if $A_1, A_2, \ldots$ are pairwise disjoint sets (that is, $A_i \cap A_j = \emptyset$ for $i \neq j$) belonging to $\mathcal{F}$, then

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots .$$

The triple $(\Omega, \mathcal{F}, P)$ is called a *probability space*. The sets belonging to $\mathcal{F}$ are called *events*. An event A is said to occur *almost surely* (a.s.) whenever $P(A) = 1$.

Example 1.2

We take the unit interval $\Omega = [0,1]$ with the σ-field $\mathcal{F} = \mathcal{B}([0,1])$ of Borel sets $B \subset [0,1]$, and *Lebesgue measure* $P = \text{Leb}$ on $[0,1]$. Then $(\Omega, \mathcal{F}, P)$ is a probability space. Recall that Leb is the unique measure defined on Borel sets such that

$$\text{Leb}[a,b] = b - a$$

for any interval $[a,b]$. (In fact Leb can be extended to a larger σ-field, but we shall need Borel sets only.)

Exercise 1.1

Show that if $A_1, A_2, \ldots$ is an *expanding* sequence of events, that is,

$$A_1 \subset A_2 \subset \cdots ,$$

then

$$P(A_1 \cup A_2 \cup \cdots) = \lim_{n\to\infty} P(A_n).$$

Similarly, if $A_1, A_2, \ldots$ is a *contracting* sequence of events, that is,

$$A_1 \supset A_2 \supset \cdots ,$$

then

$$P(A_1 \cap A_2 \cap \cdots) = \lim_{n\to\infty} P(A_n).$$

Hint Write $A_1 \cup A_2 \cup \cdots$ as the union of a sequence of disjoint events: start with A_1, then add a disjoint set to obtain $A_1 \cup A_2$, then add a disjoint set again to obtain $A_1 \cup A_2 \cup A_3$, and so on. Now that you have a sequence of disjoint sets, you can use the definition of a probability measure. To deal with the product $A_1 \cap A_2 \cap \cdots$ write it as a union of some events with the aid of De Morgan's law.

Lemma 1.1 (Borel–Cantelli)

Let $A_1, A_2, \ldots$ be a sequence of events such that $P(A_1) + P(A_2) + \cdots < \infty$ and let $B_n = A_n \cup A_{n+1} \cup \cdots$. Then $P(B_1 \cap B_2 \cap \cdots) = 0$.

Exercise 1.2

Prove the Borel–Cantelli lemma above.

Hint $B_1, B_2, \ldots$ is a contracting sequence of events.

1.2 Random Variables

Definition 1.3

If $\mathcal{F}$ is a σ-field on Ω, then a function $\xi : \Omega \to \mathbb{R}$ is said to be *$\mathcal{F}$-measurable* if

$$\{\xi \in B\} \in \mathcal{F}$$

for every Borel set $B \in \mathcal{B}(\mathbb{R})$. If $(\Omega, \mathcal{F}, P)$ is a probability space, then such a function ξ is called a *random variable.*

Remark 1.1

A short-hand notation for events such as $\{\xi \in B\}$ will be used to avoid clutter. To be precise, we should write

$$\{\omega \in \Omega : \xi(\omega) \in B\}$$

in place of $\{\xi \in B\}$. Incidentally, $\{\xi \in B\}$ is just a convenient way of writing the inverse image $\xi^{-1}(B)$ of a set.

Definition 1.4

The σ-field $\sigma(\xi)$ *generated* by a random variable $\xi : \Omega \to \mathbb{R}$ consists of all sets of the form $\{\xi \in B\}$, where B is a Borel set in $\mathbb{R}$.

Definition 1.5

The σ-field $\sigma\{\xi_i : i \in I\}$ generated by a family $\{\xi_i : i \in I\}$ of random variables is defined to be the smallest σ-field containing all events of the form $\{\xi_i \in B\}$, where B is a Borel set in $\mathbb{R}$ and $i \in I$.

Exercise 1.3

We call $f : \mathbb{R} \to \mathbb{R}$ a *Borel function* if the inverse image $f^{-1}(B)$ of any Borel set B in $\mathbb{R}$ is a Borel set. Show that if f is a Borel function and ξ is a random variable, then the composition $f(\xi)$ is $\sigma(\xi)$-measurable.

Hint Consider the event $\{f(\xi) \in B\}$, where B is an arbitrary Borel set. Can this event be written as $\{\xi \in A\}$ for some Borel set A?

Lemma 1.2 (Doob–Dynkin)

Let ξ be a random variable. Then each $\sigma(\xi)$-measurable random variable η can be written as

$$\eta = f(\xi)$$

for some Borel function $f : \mathbb{R} \to \mathbb{R}$.

The proof of this highly non-trivial result will be omitted.

Definition 1.6

Every random variable $\xi : \Omega \to \mathbb{R}$ gives rise to a probability measure

$$P_\xi(B) = P\{\xi \in B\}$$

on $\mathbb{R}$ defined on the σ-field of Borel sets $B \in \mathcal{B}(\mathbb{R})$. We call P_ξ the *distribution* of ξ. The function $F_\xi : \mathbb{R} \to [0,1]$ defined by

$$F_\xi(x) = P\{\xi \le x\}$$

is called the *distribution function* of ξ.

Exercise 1.4

Show that the distribution function F_ξ is non-decreasing, right-continuous, and

$$\lim_{x\to-\infty} F_\xi(x) = 0, \quad \lim_{x\to+\infty} F_\xi(x) = 1.$$

Hint For example, to verify right-continuity show that $F_\xi(x_n) \to F_\xi(x)$ for any decreasing sequence x_n such that $x_n \to x$. You may find the results of Exercise 1.1 useful.

Definition 1.7

If there is a Borel function $f_\xi : \mathbb{R} \to \mathbb{R}$ such that for any Borel set $B \subset \mathbb{R}$

$$P\{\xi \in B\} = \int_B f_\xi(x)\, dx,$$

then ξ is said to be a random variable with ***absolutely continuous distribution*** and f_ξ is called the ***density*** of ξ. If there is a (finite or infinite) sequence of pairwise distinct real numbers $x_1, x_2, \ldots$ such that for any Borel set $B \subset \mathbb{R}$

$$P\{\xi \in B\} = \sum_{x_i \in B} P\{\xi = x_i\},$$

then ξ is said to have ***discrete distribution*** with values $x_1, x_2, \ldots$ and ***mass*** $P\{\xi = x_i\}$ at x_i.

Exercise 1.5

Suppose that ξ has continuous distribution with density f_ξ. Show that

$$\frac{d}{dx} F_\xi(x) = f_\xi(x)$$

if f_ξ is continuous at x.

Hint Express $F_\xi(x)$ as an integral of f_ξ.

Exercise 1.6

Show that if ξ has discrete distribution with values $x_1, x_2, \ldots$, then F_ξ is constant on each interval $(s, t]$ not containing any of the x_i's and has jumps of size $P\{\xi = x_i\}$ at each x_i.

Hint The increment $F_\xi(t) - F_\xi(s)$ is equal to the total mass of the x_i's that belong to the interval $[s, t)$.

Definition 1.8

The *joint distribution* of several random variables $\xi_1, \ldots, \xi_n$ is a probability measure $P_{\xi_1,\ldots,\xi_n}$ on $\mathbb{R}^n$ such that

$$P_{\xi_1,\ldots,\xi_n}(B) = P\{(\xi_1, \ldots, \xi_n) \in B\}$$

for any Borel set B in $\mathbb{R}^n$. If there is a Borel function $f_{\xi_1,\ldots,\xi_n} : \mathbb{R}^n \to \mathbb{R}$ such that

$$P\{(\xi_1, \ldots, \xi_n) \in B\} = \int_B f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n$$

for any Borel set B in $\mathbb{R}^n$, then $f_{\xi_1,\ldots,\xi_n}$ is called the *joint density* of $\xi_1, \ldots, \xi_n$.

Definition 1.9

A random variable $\xi : \Omega \to \mathbb{R}$ is said to be *integrable* if

$$\int_\Omega |\xi|\, dP < \infty.$$

Then

$$E(\xi) = \int_\Omega \xi\, dP$$

exists and is called the *expectation* of ξ. The family of integrable random variables $\xi : \Omega \to \mathbb{R}$ will be denoted by L^1 or, in case of possible ambiguity, by $L^1(\Omega, \mathcal{F}, P)$.

Example 1.3

The *indicator function* 1_A of a set A is equal to 1 on A and 0 on the complement $\Omega \setminus A$ of A. For any event A

$$E(1_A) = \int_\Omega 1_A\, dP = P(A).$$

We say that $\eta : \Omega \to \mathbb{R}$ is a *step function* if

$$\eta = \sum_{i=1}^{n} \eta_i 1_{A_i},$$

where $\eta_1, \ldots, \eta_n$ are real numbers and $A_1, \ldots, A_n$ are pairwise disjoint events. Then

$$E(\eta) = \int_\Omega \eta\, dP = \sum_{i=1}^{n} \eta_i \int_\Omega 1_{A_i}\, dP = \sum_{i=1}^{n} \eta_i P(A_i).$$

Exercise 1.7

Show that for any Borel function $h : \mathbb{R} \to \mathbb{R}$ such that $h(\xi)$ is integrable

$$E(h(\xi)) = \int_{\mathbb{R}} h(x)\, dP_{\xi}(x).$$

Hint First verify the equality for step functions $h : \mathbb{R} \to \mathbb{R}$, then for non-negative ones by approximating them by step functions, and finally for arbitrary Borel functions by splitting them into positive and negative parts.

In particular, Exercise 1.7 implies that if ξ has an absolutely continuous distribution with density f_{ξ}, then

$$E(h(\xi)) = \int_{-\infty}^{+\infty} h(x)\, f_{\xi}(x)\, dx.$$

If ξ has a discrete distribution with (finitely or infinitely many) pairwise distinct values $x_1, x_2, \ldots$, then

$$E(h(\xi)) = \sum_i h(x_i)\, P\{\xi = x_i\}.$$

Definition 1.10

A random variable $\xi : \Omega \to \mathbb{R}$ is called *square integrable* if

$$\int_{\Omega} |\xi|^2\, dP < \infty.$$

Then the *variance* of ξ can be defined by

$$\mathrm{var}(\xi) = \int_{\Omega} (\xi - E(\xi))^2\, dP.$$

The family of square integrable random variables $\xi : \Omega \to \mathbb{R}$ will be denoted by $L^2(\Omega, \mathcal{F}, P)$ or, if no ambiguity is possible, simply by L^2.

Remark 1.2

The result in Exercise 1.8 below shows that we may write $E(\xi)$ in the definition of variance.

Exercise 1.8

Show that if ξ is a square integrable random variable, then it is integrable.

Hint Use the Schwarz inequality

$$[E(\xi\eta)]^2 \le E(\xi^2)\, E(\eta^2) \tag{1.1}$$

with an appropriately chosen η.

Exercise 1.9

Show that if $\eta : \Omega \to [0,\infty)$ is a non-negative square integrable random variable, then

$$E(\eta^2) = 2\int_0^\infty tP(\eta > t)\, dt.$$

Hint Express $E(\eta^2)$ in terms of the distribution function $F_\eta(t)$ of η and then integrate by parts.

1.3 Conditional Probability and Independence

Definition 1.11

For any events $A, B \in \mathcal{F}$ such that $P(B) \neq 0$ the ***conditional probability*** of A given B is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Exercise 1.10

Prove the ***total probability formula***

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots$$

for any event $A \in \mathcal{F}$ and any sequence of pairwise disjoint events $B_1, B_2, \ldots \in \mathcal{F}$ such that $B_1 \cup B_2 \cup \cdots = \Omega$ and $P(B_n) \neq 0$ for any n.

Hint $A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots$.

Definition 1.12

Two events $A, B \in \mathcal{F}$ are called ***independent*** if

$$P(A \cap B) = P(A)P(B).$$

In general, we say that n events $A_1, \ldots, A_n \in \mathcal{F}$ are ***independent*** if

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$$

for any indices $1 \leq i_1 < i_2 < \cdots < i_k \leq n$.

Exercise 1.11

Let $P(B) \neq 0$. Show that A and B are independent events if and only if $P(A|B) = P(A)$.

Hint If $P(B) \neq 0$, then you can divide by it.

Definition 1.13

Two random variables ξ and η are called ***independent*** if for any Borel sets $A, B \in \mathcal{B}(\mathbb{R})$ the two events

$$\{\xi \in A\} \quad \text{and} \quad \{\eta \in B\}$$

are independent. We say that n random variables $\xi_1, \ldots, \xi_n$ are ***independent*** if for any Borel sets $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ the events

$$\{\xi_1 \in B_1\}, \ldots, \{\xi_n \in B_n\}$$

are independent. In general, a (finite or infinite) family of random variables is said to be ***independent*** if any finite number of random variables from this family are independent.

Proposition 1.1

If two integrable random variables $\xi, \eta : \Omega \to \mathbb{R}$ are independent, then they are *uncorrelated*, i.e.

$$E(\xi\eta) = E(\xi)E(\eta),$$

provided that the product $\xi\eta$ is also integrable. If $\xi_1, \ldots, \xi_n : \Omega \to \mathbb{R}$ are independent integrable random variables, then

$$E(\xi_1 \xi_2 \cdots \xi_n) = E(\xi_1) E(\xi_2) \cdots E(\xi_n),$$

provided that the product $\xi_1 \xi_2 \cdots \xi_n$ is also integrable.

Definition 1.14

Two σ-fields $\mathcal{G}$ and $\mathcal{H}$ contained in $\mathcal{F}$ are called ***independent*** if any two events

$$A \in \mathcal{G} \quad \text{and} \quad B \in \mathcal{H}$$

are independent. Similarly, any finite number of σ-fields $\mathcal{G}_1, \ldots, \mathcal{G}_n$ contained in $\mathcal{F}$ are *independent* if any n events

$$A_1 \in \mathcal{G}_1, \ldots, A_n \in \mathcal{G}_n$$

are independent. In general, a (finite or infinite) family of σ-fields is said to be *independent* if any finite number of them are independent.

Exercise 1.12

Show that two random variables ξ and η are independent if and only if the σ-fields $\sigma(\xi)$ and $\sigma(\eta)$ generated by them are independent.

Hint The events in $\sigma(\xi)$ and $\sigma(\eta)$ are of the form $\{\xi \in A\}$, and $\{\eta \in B\}$, where A and B are Borel sets.

Sometimes it is convenient to talk of independence for a combination of random variables and σ-fields.

Definition 1.15

We say that a random variable ξ is *independent* of a σ-field $\mathcal{G}$ if the σ-fields

$$\sigma(\xi) \quad \text{and} \quad \mathcal{G}$$

are independent. This can be extended to any (finite or infinite) family consisting of random variables or σ-fields or a combination of them both. Namely, such a family is called *independent* if for any finite number of random variables $\xi_1, \ldots, \xi_m$ and σ-fields $\mathcal{G}_1, \ldots, \mathcal{G}_n$ from this family the σ-fields

$$\sigma(\xi_1), \ldots, \sigma(\xi_m), \mathcal{G}_1, \ldots, \mathcal{G}_n$$

are independent.

1.4 Solutions

Solution 1.1

If $A_1 \subset A_2 \subset \cdots$, then

$$A_1 \cup A_2 \cup \cdots = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup \cdots,$$

where the sets $A_1, A_2 \setminus A_1, A_3 \setminus A_2, \ldots$ are pairwise disjoint. Therefore, by the definition of probability measure

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots) &= P(A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup \cdots) \\ &= P(A_1) + P(A_2 \setminus A_1) + P(A_3 \setminus A_2) + \cdots \\ &= \lim_{n\to\infty} P(A_n). \end{aligned}$$

The last equality holds because the partial sums in the series above are

$$\begin{aligned} P(A_1) + P(A_2 \setminus A_1) + \cdots + P(A_n \setminus A_{n-1}) &= P(A_1 \cup \cdots \cup A_n) \\ &= P(A_n). \end{aligned}$$

If $A_1 \supset A_2 \supset \cdots$, then the equality

$$P(A_1 \cap A_2 \cap \cdots) = \lim_{n\to\infty} P(A_n)$$

follows by taking the complements of A_n and applying De Morgan's law

$$\Omega \setminus (A_1 \cap A_2 \cap \cdots) = (\Omega \setminus A_1) \cup (\Omega \setminus A_2) \cup \cdots .$$

Solution 1.2

Since B_n is a contracting sequence of events, the results of Exercise 1.1 imply that

$$\begin{aligned} P(B_1 \cap B_2 \cap \cdots) &= \lim_{n\to\infty} P(B_n) \\ &= \lim_{n\to\infty} P(A_n \cup A_{n+1} \cup \cdots) \\ &\leq \lim_{n\to\infty} (P(A_n) + P(A_{n+1}) + \cdots) \\ &= 0. \end{aligned}$$

The last equality holds because the series $\sum_{n=1}^{\infty} P(A_n)$ is convergent. The inequality above holds by the subadditivity property

$$P(A_n \cup A_{n+1} \cup \cdots) \leq P(A_n) + P(A_{n+1}) + \cdots .$$

It follows that $P(B_1 \cap B_2 \cap \cdots) = 0$.

Solution 1.3

If B is a Borel set in $\mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ is a Borel function, then $f^{-1}(B)$ is also a Borel set. Therefore

$$\{f(\xi) \in B\} = \{\xi \in f^{-1}(B)\}$$

belongs to the σ-field $\sigma(\xi)$ generated by ξ. It follows that the composition $f(\xi)$ is $\sigma(\xi)$-measurable.

Solution 1.4

If $x \leq y$, then $\{\xi \leq x\} \subset \{\xi \leq y\}$, so

$$F_\xi(x) = P\{\xi \leq x\} \leq P\{\xi \leq y\} = F_\xi(y).$$

This means that F_ξ is non-decreasing.

Next, we take any sequence $x_1 \geq x_2 \geq \cdots$ and put

$$\lim_{n\to\infty} x_n = x.$$

Then the events

$$\{\xi \leq x_1\} \supset \{\xi \leq x_2\} \supset \cdots$$

form a contracting sequence with intersection

$$\{\xi \leq x\} = \{\xi \leq x_1\} \cap \{\xi \leq x_2\} \cap \cdots .$$

It follows by Exercise 1.1 that

$$F_\xi(x) = P\{\xi \leq x\} = \lim_{n\to\infty} P\{\xi \leq x_n\} = \lim_{n\to\infty} F_\xi(x_n).$$

This proves that F_ξ is right-continuous.

Since the events

$$\{\xi \leq -1\} \supset \{\xi \leq -2\} \supset \cdots$$

form a contracting sequence with intersection $\emptyset$ and

$$\{\xi \leq 1\} \subset \{\xi \leq 2\} \subset \cdots$$

form an expanding sequence with union Ω, it follows by Exercise 1.1 that

$$\lim_{x\to-\infty} F_\xi(x) = \lim_{n\to\infty} F_\xi(-n) = \lim_{n\to\infty} P\{\xi \leq -n\} = P(\emptyset) = 0,$$

$$\lim_{x\to\infty} F_\xi(x) = \lim_{n\to\infty} F_\xi(n) = \lim_{n\to\infty} P\{\xi \leq n\} = P(\Omega) = 1,$$

since F_ξ is non-decreasing.

Solution 1.5

If ξ has a density f_ξ, then the distribution function F_ξ can be written as

$$F_\xi(x) = P\{\xi \leq x\} = \int_{-\infty}^{x} f_\xi(y)\, dy.$$

Therefore, if f_ξ is continuous at x, then F_ξ is differentiable at x and

$$\frac{d}{dx} F_\xi(x) = f_\xi(x).$$

Solution 1.6

If $s < t$ are real numbers such that $x_i \notin (s,t]$ for any i, then

$$F_\xi(t) - F_\xi(s) = P\{\xi \le t\} - P\{\xi \le s\} = P\{\xi \in (s,t]\} = 0,$$

i.e. $F_\xi(s) = F_\xi(t)$. Because F_ξ is non-decreasing, this means that F_ξ is constant on $(s,t]$. To show that F_ξ has a jump of size $P\{\xi = x_i\}$ at each x_i, we compute

$$\begin{aligned}\lim_{t\searrow x_i} F_\xi(t) - \lim_{s\nearrow x_i} F_\xi(s) &= \lim_{t\searrow x_i} P\{\xi \le t\} - \lim_{s\nearrow x_i} P\{\xi \le s\} \\ &= P\{\xi \le x_i\} - P\{\xi < x_i\} = P\{\xi = x_i\}.\end{aligned}$$

Solution 1.7

If h is a step function,

$$h = \sum_{i=1}^{n} h_i 1_{A_i},$$

where $h_1, \ldots, h_n$ are real numbers and $A_1, \ldots, A_n$ are pairwise disjoint Borel sets covering $\mathbb{R}$, then

$$\begin{aligned}E(h(\xi)) &= \sum_{i=1}^{n} h_i E(1_{A_i}(\xi)) = \sum_{i=1}^{n} h_i P\{\xi \in A_i\} \\ &= \sum_{i=1}^{n} h_i P_\xi(A_i) = \sum_{i=1}^{n} \int_{A_i} h(x)\, dP_\xi(x) = \int_{\mathbb{R}} h(x)\, dP_\xi(x).\end{aligned}$$

Next, any non-negative Borel function h can be approximated by a non-decreasing sequence of step functions. For such an h the result follows by the monotone convergence of integrals. Finally, this implies the desired equality for all Borel functions h, since each can be split into its positive and negative parts, $h = h^+ - h^-$, where $h^+, h^- \ge 0$.

Solution 1.8

By the Schwarz inequality (1.1) with $\eta = 1$, if ξ is square integrable, then

$$[E(|\xi|)]^2 = [E(1\,|\xi|)]^2 \le E(1^2)\, E(\xi^2) = E(\xi^2) < \infty,$$

i.e. ξ is integrable.

Solution 1.9

Let $F(t) = P\{\eta \le t\}$ be the distribution function of η. Then

$$E(\eta^2) = \int_0^\infty t^2 dF(t).$$

Since $P(\eta > t) = 1 - F(t)$, we need to show that

$$\int_0^\infty t^2\, dF(t) = 2\int_0^\infty t\,(1 - F(t))\, dt \tag{1.2}$$

First, let us establish a version of (1.2) with ∞ replaced by a finite number a. Integrating by parts, we obtain

$$\begin{aligned} \int_0^a t^2 dF(t) &= \int_0^a t^2 d(F(t) - 1) \\ &= t^2(F(t) - 1)\big|_0^a - 2\int_0^a t(F(t) - 1)\, dt \\ &= -a^2(1 - F(a)) + 2\int_0^a t(1 - F(t))dt. \end{aligned} \tag{1.3}$$

We see that (1.2) follows from (1.3), provided that

$$a^2\,(1 - F(a)) \to 0, \quad \text{as } a \to \infty. \tag{1.4}$$

But

$$0 \le a^2\,(1 - F(a)) = a^2 P(\eta > a) \le (n+1)^2 P(\eta > n) \le 4n^2 P(\eta \ge n),$$

where n is the integer part of a, and

$$E(\eta^2) = \sum_{k=0}^{\infty} \int_{\{k \le \eta \le k+1\}} \eta^2 dP < \infty.$$

Hence,

$$n^2 P\,(\eta \ge n) \le \int_{\{\eta \ge n\}} \eta^2\, dP = \sum_{k=n}^{\infty} \int_{\{k \le \eta < k+1\}} \eta^2\, dP \to 0 \tag{1.5}$$

as $n \to \infty$, which proves (1.4).

Solution 1.10

Since $B_1 \cup B_2 \cup \cdots = \Omega$,

$$A = A \cap (B_1 \cup B_2 \cup \cdots) = (A \cap B_1) \cup (A \cap B_2) \cup \cdots\ ,$$

where

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \emptyset = \emptyset.$$

By countable additivity

$$\begin{aligned} P(A) &= P\,((A \cap B_1) \cup (A \cap B_2) \cup \cdots) \\ &= P\,(A \cap B_1) + P\,(A \cap B_2) + \cdots \\ &= P\,(A|B_1)\, P(B_1) + P\,(A|B_2)\, P(B_2) + \cdots\ . \end{aligned}$$

Solution 1.11

If $P(B) \neq 0$, then A and B are independent if and only if

$$P(A) = \frac{P(A \cap B)}{P(B)}.$$

In turn, this equality holds if and only if $P(A) = P(A|B)$.

Solution 1.12

The σ-fields $\sigma(\xi)$ and $\sigma(\eta)$ consist, respectively, of events of the form

$$\{\xi \in A\} \quad \text{and} \quad \{\eta \in B\},$$

where A and B are Borel sets in $\mathbb{R}$. Therefore, $\sigma(\xi)$ and $\sigma(\eta)$ are independent if and only if the events $\{\xi \in A\}$, and $\{\eta \in B\}$ are independent for any Borel sets A and B, which in turn is equivalent to ξ and η being independent.