Basics of Probability Theory

Daisuke Oyama

Faculty of Economics, University of Tokyo

oyama@e.u-tokyo.ac.jp

This version: February 17, 2014

1 Measurable Spaces and Random Variables

Let Ω be a set, and $\mathcal{F}$ a family of subsets of Ω. $\mathcal{F}$ is a *σ-algebra*, or *σ-field*, if

(1) $\Omega \in \mathcal{F}$,

(2) $A \in \Omega$ implies $A^c \in \mathcal{F}$, and

(3) $A_1, A_2, \ldots \in \mathcal{F}$ implies $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

A pair $(\Omega, \mathcal{F})$ of a set Ω and a σ-algebra $\mathcal{F}$ of subsets of Ω is called a *measurable space*. An element of $\mathcal{F}$ is called a *measurable set* or an *event*.

For a family $\mathcal{A}$ of subsets of Ω, the *σ-algebra generated by $\mathcal{A}$* is the σ-algebra given by

$$\sigma(\mathcal{A}) = \bigcap \{\mathcal{M} \mid \mathcal{M} \text{ is a } \sigma\text{-algebra containing } \mathcal{A}\},$$

which is the smallest σ-algebra containing $\mathcal{A}$. The *Borel algebra* for $\mathbb{R}$, which we denote $\mathcal{B}(\mathbb{R})$, is the σ-algebra generated by the family of all open sets in $\mathbb{R}$.

For a function $X \colon \Omega \to \mathbb{R}$, we write

$$\{X \in B\} = \{\omega \in \Omega \mid X(\omega) \in B\}$$

for $B \in \mathbb{R}$, and $\{X \geq a\} = \{\omega \in \Omega \mid X(\omega) \geq a\}$ for $a \in \mathbb{R}$, and so on. For a measurable space $(\Omega, \mathcal{F})$, a function $X \colon \Omega \to \mathbb{R}$ is a (real-valued) *random variable* on $(\Omega, \mathcal{F})$ if it is *$\mathcal{F}$-measurable*, i.e.,

$$\{X \in B\} \in \mathcal{F} \text{ for all } B \in \mathcal{B}(\mathbb{R}).$$

For a random variable X on $(\Omega, \mathcal{F})$, the σ-algebra generated by X is the σ-algebra given by

$$\sigma(X) = \{\{X \in B\} \in \mathcal{F} \mid B \in \mathcal{B}(\mathbb{R})\},$$

which is the smallest σ-field with respect to which X is measurable. Likewise, for a family $(X_\lambda)_{\lambda \in \Lambda}$ of random variables on $(\Omega, \mathcal{F})$, the σ-algebra generated by $(X_\lambda)_{\lambda \in \Lambda}$,

$\sigma((X_\lambda)_{\lambda\in\Lambda})$, is the σ-algebra generated by the family of sets $\{X_\lambda \in B\}$, $B \in \mathcal{B}(\mathbb{R})$, $\lambda \in \Lambda$, which is the smallest σ-algebra with respect to which all X_λ's are measurable.

For $A \subset \Omega$, we define the function $\mathbf{1}_A \colon \Omega \to \mathbb{R}$ by

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A, \end{cases}$$

which is called the indicator function of A. $\mathbf{1}_A$ is a random variable on $(\Omega, \mathcal{F})$ if and only if $A \in \mathcal{F}$. We say that a random variable X on $(\Omega, \mathcal{F})$ is simple if there are $A_1, \ldots, A_n \in \mathcal{F}$, $A_i \cap A_j = \emptyset$, $i \neq j$, such that $X = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$ with $a_1, \ldots, a_n \in \mathbb{R}$.

2 Probability Measures

For a measurable space $(\Omega, \mathcal{F})$, a function $P \colon \mathcal{F} \to [0,1]$ is a *probability measure* on $(\Omega, \mathcal{F})$ if

(1) $P(\emptyset) = 0$ and $P(\Omega) = 1$, and

(2) if $A_1, A_2, \ldots \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$ for $i \neq j$, then $P\left(\bigcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty P(A_i)$.

A tuple $(\Omega, \mathcal{F}, P)$ of a nonempty set Ω, a σ-algebra $\mathcal{F}$ of subsets of Ω, and a probability measure P on $(\Omega, \mathcal{F})$ is called a *probability space*.

Proposition 2.1.

(1) *If $A, B \in \mathcal{F}$ and $A \subset B$, then $P(A) \leq P(B)$.*

(2) *If $A_1, A_2, \ldots \in \mathcal{F}$ and $A_1 \subset A_2 \subset \cdots$, then $P(\bigcup_{n=1}^\infty A_n) = \lim_{n\to\infty} P(A_n)$.*

(3) *If $A_1, A_2, \ldots \in \mathcal{F}$ and $A_1 \supset A_2 \supset \cdots$, then $P(\bigcap_{n=1}^\infty A_n) = \lim_{n\to\infty} P(A_n)$.*

(4) *If $A_1, A_2, \ldots \in \mathcal{F}$, then $P(\bigcup_{n=1}^\infty A_n) \leq \sum_{n=1}^\infty P(A_n)$.*

The following is called the (first) Borel-Cantelli lemma.

Proposition 2.2. *If $A_1, A_2, \ldots \in \mathcal{F}$ and $\sum_{n=1}^\infty P(A_n) < \infty$, then $P\left(\bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k\right) = 0$.*

Proof. Let $B_m = \bigcup_{k=m}^\infty A_k$. Then $\bigcap_{n=1}^\infty B_n \subset B_m$ for any m. Thus

$$0 \leq P\left(\bigcap_{n=1}^\infty B_n\right) \leq P(B_m) \leq \sum_{k=m}^\infty P(A_k)$$

holds for any m, and the sum in the last term tends to 0 as $m \to \infty$ if $\sum_{n=1}^\infty P(A_n) < \infty$. ∎

If a property holds except on an event whose probability is zero, then this property is said to hold *almost surely*, abbreviated "a.s.".

3 Expectation

Let a probability space $(\Omega, \mathcal{F}, P)$ be given.

Definition 3.1. For a nonnegative simple random variable $X = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}$, the expectation of X is defined by

$$E[X] = \sum_{i=1}^{n} a_n P(A_n).$$

Lemma 3.1. *For any nonnegative random variable X, there exists a nondecreasing sequence of nonnegative simple random variables (X_n) such that $\lim_{n\to\infty} X_n = X$.*

If (X_n) and (Y_n) are nondecreasing sequences of nonnegative simple random variables, and $\lim_{n\to\infty} X_n = \lim_{n\to\infty} Y_n = X$, then $\lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} E[Y_n]$.

Definition 3.2. For a nonnegative random variable X, the expectation of X is defined by

$$E[X] = \lim_{n\to\infty} E[X_n],$$

where (X_n) is a nondecreasing sequence of nonnegative simple random variables such that $\lim_{n\to\infty} X_n = X$.

Note that this is well defined, i.e., the value does not depend on the choice of an approximating sequence.

A random variable X is said to be *integrable* if $E[|X|] < \infty$.

Write

$$X^+ = X \vee 0, \quad X^- = -(X \wedge 0).$$

Note that

$$X = X^+ - X^-, \quad |X| = X^+ + X^-,$$

and that if X is integrable, then X^+ and X^- are integrable.

Definition 3.3. For an integrable random variable X, the expectation of X is defined by

$$E[X] = E[X^+] - E[X^-].$$

For $A \in \mathcal{F}$, we write

$$E[X, A] = E[X\mathbf{1}_A].$$

Observe that if $E[X, A] = 0$ whenever $P(A) = 0$.

Proposition 3.2. *Let X, Y be integrable random variables.*

(1) *For* $a, b \in \mathbb{R}$*,* $aX + bY$ *is integrable, and*

$$E[aX + bY] = aE[X] + bE[Y].$$

(2) *If* $A, B \in \mathcal{F}$ *and* $A \cap B = \emptyset$*, then*

$$E[X, A \cup B] = E[X, A] + E[X, B].$$

(3) *If* $X \geq Y$ *a.s., then*

$$E[X] \geq E[Y].$$

Proposition 3.3 (Markov's Inequality)**.** *Let* X *be a random variable such that* $X \geq 0$ *a.s. Then for any* $\alpha \in \mathbb{R}$ *and any* $A \in \mathcal{F}$*,*

$$\alpha P(\{X \geq \alpha\} \cap A) \leq E[X, A].$$

Proof. Observe that

$$X \geq X\mathbf{1}_{\{X \geq \alpha\}} \geq \alpha \mathbf{1}_{\{X \geq \alpha\}} \text{ a.s.}$$

Thus we have $E[X, A] \geq E\left[\alpha \mathbf{1}_{\{X \geq \alpha\} \cap A}\right] = \alpha P(\{X \geq \alpha\} \cap A)$. ∎

Proposition 3.4 (Lebesgue's Convergence Theorem)**.** *Let* (X_n) *be a sequence of random variables, and suppose that there exists an integrable random variable* Y *such that for all* n*,* $|X_n| \leq Y$ *a.s. If* $\lim_{n \to \infty} X_n = X$ *a.s., then*

$$\lim_{n \to \infty} E[X_n] = E[X].$$

Proposition 3.5. *Suppose that* $f\colon I \times \Omega \to \mathbb{R}$*, where* $I \subset \mathbb{R}$ *is an open interval, satisfies the following conditions:*

(i) *for all* $t \in I$*,* $f(t, \cdot)$ *is integrable; and*

(ii) *for almost all* $\omega \in \Omega$*,* $f(\cdot, \omega)$ *is differentiable, and there exists an integrable function* $g\colon \Omega \to \mathbb{R}$ *such that for almost all* $\omega \in \Omega$*,* $|f_t(t, \omega)| \leq g(\omega)$ *for all* $t \in I$*.*

Then $E[f(t, \cdot)]$ *is differentiable in* t *on* I *with*

$$\frac{d}{dt} E[f(t, \cdot)] = E[f_t(t, \cdot)].$$

Proof. Fix any $t_0 \in I$. For any sequence (t_n) such that $t_n \neq t_0$ and $t_n \to t_0$, let

$$X_n(\omega) = \frac{f(t_n, \omega) - f(t_0, \omega)}{t_n - t_0}.$$

Thus

$$\frac{E[f(t_n, \cdot)] - E[f(t_0, \cdot)]}{t_n - t_0} = E[X_n].$$

We want to show that

$$\lim_{n\to\infty} E[X_n] = E[f_t(t_0,\cdot)].$$

Fix any $\omega \in \Omega$ as in condition (ii). First, $\lim_{n\to\infty} X_n(\omega) = f_t(t_0,\omega)$. Second, for each n, by the mean value theorem we have

$$\frac{f(t_n,\omega) - f(t_0,\omega)}{t_n - t_0} = f_t(s,\omega)$$

for some s between t_0 and t_n, but the right hand side is bounded in absolute value by $g(\omega)$, so that we have

$$|X_n(\omega)| \le g(\omega)$$

for all n. Hence, it follows from Lebesgue's convergence theorem that

$$\lim_{n\to\infty} E[X_n] = E\left[\lim_{n\to\infty} X_n\right] = E[f_t(t_0,\cdot)]$$

as desired. ∎

4 Independence

Sub-σ-algebras $(\mathcal{F}_\lambda)_{\lambda\in\Lambda}$ of $\mathcal{F}$ are *independent* if for any finite subfamily $\mathcal{F}_{\lambda_1},\dots,\mathcal{F}_{\lambda_n}$,

$$P(A_1 \cap \cdots \cap A_n) = P(A_1)\cdots P(A_n)$$

for all $A_i \in \mathcal{F}_{\lambda_i}$, $i = 1,\dots,n$. Random variables $(X_\lambda)_{\lambda\in\Lambda}$ are independent if $(\sigma(X_\lambda))_{\lambda\in\Lambda}$ are independent.

Proposition 4.1. *If $X_1,\dots,X_n$ are independent and integrable, then $X_1\cdots X_n$ is integrable and*

$$E[X_1\cdots X_n] = E[X_1]\cdots E[X_n].$$

In particular, if X_1 and X_2 are independent and integrable, then

$$E[X_1X_2, A] = E[X_1, A]E[X_2]$$

for $A \in \sigma(X_1)$. Indeed, if $A \in \sigma(X_1)$, then $X_1\mathbf{1}_A$ and X_2 are independent, and therefore by Proposition 4.1, $E[X_1X_2, A] = E[X_1\mathbf{1}_A \cdot X_2] = E[X_1\mathbf{1}_A]E[X_2] = E[X_1, A]E[X_2]$.

5 Martingales

For a thorough account, see e.g., Billingsley (2012, Section 35).

A sequence of random variables $(Z_k)_{k\in\mathbb{Z}_+}$ is a *supermartingale* if $E[|Z_k|] < \infty$ for all $k \in \mathbb{Z}_+$, and

$$E[Z_k, A] \le E[Z_{k-1}, A] \text{ for all } A \in \sigma(Z_0,\dots,Z_{k-1})$$

for all $k \ge 1$.

Lemma 5.1. *Let $(Z_k)_{k\in\mathbb{Z}_+}$ be a supermartingale. For $\alpha \in \mathbb{R}$, let*

$$\tau = \min\{k \in \mathbb{Z}_+ \mid Z_k \geq \alpha\}.$$

Then for any $n \in \mathbb{Z}_+$,

$$E[Z_{\tau\wedge n}] \leq E[Z_0].$$

Proof. Observe first that

$$\begin{aligned} &Z_{\tau\wedge n} - Z_0 \\ &= \left(Z_n \mathbf{1}_{\{\tau=n\}} + Z_{n-1}\mathbf{1}_{\{\tau=n-1\}} + \cdots + Z_1 \mathbf{1}_{\{\tau=1\}} + Z_0 \mathbf{1}_{\{\tau=0\}}\right) - Z_0 \\ &= (Z_n - Z_{n-1})\mathbf{1}_{\{\tau\geq n\}} + (Z_{n-1} - Z_{n-2})\mathbf{1}_{\{\tau\geq n-1\}} + \cdots + (Z_1 - Z_0)\mathbf{1}_{\{\tau\geq 1\}} \end{aligned}$$

(note that $\mathbf{1}_{\{\tau=k-1\}} = \mathbf{1}_{\{\tau\geq k-1\}} - \mathbf{1}_{\{\tau\geq k\}}$), so that

$$E[Z_{\tau\wedge n} - Z_0] = \sum_{k=1}^{n} E[Z_k - Z_{k-1}, \{\tau \geq k\}].$$

For each $k = 1, \ldots, n$, we have

$$\{\tau \geq k\} = \{Z_0 < \alpha\} \cap \{Z_1 < \alpha\} \cap \cdots \cap \{Z_{k-1} < \alpha\} \in \sigma(Z_0, \ldots, Z_{k-1}),$$

and therefore

$$E[Z_k - Z_{k-1}, \{\tau \geq k\}] \leq 0$$

by the assumption that $(Z_k)_{k\in\mathbb{Z}_+}$ is a supermartingale. Hence, we have $E[Z_{\tau\wedge n} - Z_0] \leq 0$. ∎

Proposition 5.2 (Doob's Supermartingale Inequality)**.** *Let $(Z_k)_{k\in\mathbb{Z}_+}$ be a supermartingale such that for all $k \in \mathbb{Z}_+$, $Z_k \geq 0$ a.s. Then for any $\alpha \in \mathbb{R}$ and for any $n \in \mathbb{Z}_+$,*

$$\alpha P\left(\max_{0\leq k\leq n} Z_k \geq \alpha\right) \leq E[Z_0].$$

Proof. Let $\tau = \min\{k \in \mathbb{Z}_+ \mid Z_k \geq \alpha\}$. Note that

$$\left\{\max_{0\leq k\leq n} Z_k \geq \alpha\right\} = \{Z_\tau \geq \alpha\} \cap \{\tau \leq n\}.$$

We have

$$\begin{aligned} E[Z_{\tau\wedge n}] &= E[Z_\tau, \{\tau \leq n\}] + E[Z_n, \{\tau > n\}] \\ &\geq E[Z_\tau, \{\tau \leq n\}] \\ &\geq \alpha P(\{Z_\tau \geq \alpha\} \cap \{\tau \leq n\}) = \alpha P\left(\max_{0\leq k\leq n} Z_k \geq \alpha\right), \end{aligned}$$

where the first inequality follows from the assumption that for all $k \in \mathbb{Z}_+$, $Z_k \geq 0$ a.s., while the second follows from Markov's inequality. But $E[Z_0] \geq E[Z_{\tau\wedge n}]$ by Lemma 5.1, and hence we have the desired inequality. ∎

References

BILLINGSLEY, P. (2012). *Probability and Measure*, John Wiley & Sons, Hoboken.