4. Random Variables

DASN, NTU https://personal.ntu.edu.sg/wptay/

Throughout this note, we consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

4.1 Measurable Functions

Definition 4.1. *A function $X : (\Omega, \mathcal{A}) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable if $\forall\, B \in \mathcal{B}(\mathbb{R})$, $X^{-1}(B) \triangleq \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$. In probability theory, X is called a random variable or random element.*

Lemma 4.1. *X is a random variable iff $\forall\, t \in \mathbb{R}$, $\{\omega \in \Omega : X(\omega) \le t\} \in \mathcal{A}$.*

Proof. '$\Rightarrow$': Suppose X is a random variable. Since $(-\infty, t] \in \mathcal{B}$, we have $\{\omega \in \Omega : X(\omega) \le t\} \in \mathcal{A}$.

'$\Leftarrow$': Let $\mathcal{D} = \{B \in \mathcal{B} : X^{-1}(B) \in \mathcal{A}\}$. Since

$$X^{-1}(B^c) = (X^{-1}(B))^c,$$
$$X^{-1}\left(\bigcup_{i \ge 1} B_i\right) = \bigcup_{i \ge 1} X^{-1}(B_i),$$

$\mathcal{D}$ is a σ-algebra. Since $(-\infty, t] \in \mathcal{D}$ and $\{(-\infty, t]\}$ generates $\mathcal{B}(\mathbb{R})$, we have $\mathcal{B}(\mathbb{R}) \subset \mathcal{D}$ and X is a random variable. $\square$

Lemma 4.2. *Suppose X is a random variable, then $\inf_{n \ge 1} X_n$, $\sup_{n \ge 1} X_n$, $\limsup_{n \to \infty} X_n$, $\liminf_{n \to \infty} X_n$ are random variables.*

Proof. We show that $\inf_{n \ge 1} X_n$ is a random variable. We have for any $t \in \mathbb{R}$,

$$\left\{\inf_{n \ge 1} X_n \le t\right\} = \bigcup_{n \ge 1} \{X_n \le t\} \in \mathcal{A},$$

since each X_n is a random variable. By Lemma 4.1, the result follows. The proofs for the other claims are similar. $\square$

The law or distribution of a random variable X is given by

$$\begin{aligned} \mathbb{P}_X(B) \triangleq \mathbb{P}(X \in B) &= \mathbb{P}(\{\omega : X(\omega) \in B\}) \\ &= \mathbb{P}\big(X^{-1}(B)\big) \\ &= \mathbb{P} \circ X^{-1}(B), \end{aligned}$$

and we write

$$\begin{aligned}\sigma(X) &= \sigma\text{-algebra generated by } X \\ &= \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}.\end{aligned}$$

Note that the RHS of the last equality is a σ-algebra.

Suppose $g : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable, i.e., $g^{-1}(B) \in \mathcal{B}(\mathbb{R})$ for all $B \in \mathcal{B}(\mathbb{R})$, then $g(X)$ is a random variable since

$$(g \circ X)^{-1}(B) = X^{-1}(g^{-1}(B)) \in \mathcal{A}.$$

4.2 Expectation

For $A \in \mathcal{A}$, the indicator function for A is given by

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

Definition 4.2. *$X(\omega)$ is a simple random variable if there is a discrete set of numbers $x_1 < x_2 < \ldots < x_n$, such that $X(\omega) = \sum_{i=1}^n x_i \mathbf{1}_{A_i}(\omega)$, where $A_i \in \mathcal{A}$, $A_i \cap A_j = \emptyset$ for $i \neq j$, and $\bigcup_{i=1}^n A_i = \Omega$.*

Definition 4.3. *The expectation (Lebesgue integral) of a simple random variable (measurable function) $X : \Omega \to \mathbb{R}$ is given by*

$$\int_\Omega X(\omega)\,\mathrm{d}\mathbb{P} = \int_\Omega X(\omega)\mathbb{P}(\,\mathrm{d}\omega) = \mathbb{E}[X] \triangleq \sum_{i=1}^n x_i \mathbb{P}(A_i).$$

Lemma 4.3. *For any measurable partition $B_1, B_2, \ldots, B_n$ of Ω (i.e., for $i \neq j$, $B_i \cap B_j = \emptyset$ and $\bigcup_{i=1}^n B_i = \Omega$), with $X(\omega) = \sum_{j=1}^m b_j \mathbf{1}_{B_j}(\omega)$, then $\mathbb{E}X = \sum_{j=1}^m b_j \mathbb{P}(B_j)$.*

Proof. For $\omega \in A_i \cap B_j$, we have $x_i = X(\omega) = b_j$. Therefore, we obtain

$$\begin{aligned}\sum_{i=1}^n x_i \mathbb{P}(A_i) &= \sum_{i=1}^n x_i \sum_{j=1}^m \mathbb{P}(A_i \cap B_j) \\ &= \sum_i \sum_j x_i \mathbb{P}(A_i \cap B_j) \\ &= \sum_i \sum_j b_j \mathbb{P}(A_i \cap B_j) \\ &= \sum_j b_j \mathbb{P}(B_j),\end{aligned}$$

where the interchange of summations is allowed because the number of terms in the summand is finite. $\square$

The following result follows from Definition 4.3 using similar computation as in Lemma 4.3.

Lemma 4.4. *Suppose X and Y are simple random variables. Then we have:*

(i) $X(\omega) \le Y(\omega) \implies \mathbb{E}X \le \mathbb{E}Y$.

(ii) For any $a \in \mathbb{R}$, aX is a simple random variable and $\mathbb{E}[aX] = a\mathbb{E}X$.

(iii) $X + Y$ is a simple random variable and $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$.

Definition 4.4. *If $X \ge 0$, we define*

$$\mathbb{E}X = \sup\{\mathbb{E}Z \, : \, Z \le X, \, Z \textit{ is simple}\}.$$

Let $X^+ = \max(X, 0) \ge 0$ and $X^- = -\min(X, 0) \ge 0$. We have $X = X^+ - X^-$. We can then define $\mathbb{E}X = \mathbb{E}X^+ - \mathbb{E}X^-$, if not both $\mathbb{E}X^+$ and $\mathbb{E}X^-$ are infinite. If $\mathbb{E}X^+$ and $\mathbb{E}X^-$ are both infinite, then the expectation of X is *undefined*.

If $\mathbb{E}|X| = \mathbb{E}X^+ + \mathbb{E}X^- < \infty$, we say that it is integrable. This implies that both $\mathbb{E}X^+$ and $\mathbb{E}X^-$ are finite, hence $\mathbb{E}X$ exists and is finite.

Lemma 4.5. *If $0 \le X \le Y$, then $\mathbb{E}X \le \mathbb{E}Y$.*

Proof. For any simple r.v. Z, since $\{Z \le X\} \subset \{Z \le Y\}$, $\mathbb{E}X \le \mathbb{E}Y$ follows from Definition 4.4. □

Our goal is to prove the following theorem, which shows that expectations are additive.

Theorem 4.1. *Suppose $X, Y \ge 0$, then $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$.*

We first show the following.

Lemma 4.6. *Suppose $X, Y \ge 0$, then $\mathbb{E}[X + Y] \ge \mathbb{E}X + \mathbb{E}Y$.*

Proof. $\forall \epsilon > 0$, $\exists$ simple random variables $Z_1 \le X$ and $Z_2 \le Y$ such that

$$\begin{aligned} \mathbb{E}X &\le \mathbb{E}Z_1 + \frac{\epsilon}{2}, \\ \mathbb{E}Y &\le \mathbb{E}Z_2 + \frac{\epsilon}{2}, \end{aligned}$$

so that

$$\begin{aligned} \mathbb{E}X + \mathbb{E}Y &\le \mathbb{E}Z_1 + \mathbb{E}Z_2 + \epsilon \\ &= \mathbb{E}[Z_1 + Z_2] + \epsilon \\ &\le \mathbb{E}[X + Y] + \epsilon, \end{aligned}$$

where the first equality follows from Lemma 4.4 and the last inequality from $Z_1 + Z_2 \le X + Y$ and Definition 4.4. Since ϵ can be arbitrarily small, the proof is complete. □

To complete the proof of Theorem 4.1, we need to show subadditivity: $\mathbb{E}[X+Y] \leq \mathbb{E}X + \mathbb{E}Y$. It turns out that this is highly non-trivial and in the process we learn several nice proof techniques and the very useful theorem below.

Theorem 4.2 (Monotone Convergence Theorem (MCT))**.** *Suppose* $0 \leq X_n \leq X_{n+1}$*, for* $n \geq 1$ *and* $X_n(\omega) \to X(\omega)$ *for all* $\omega \in \Omega$*. Then*

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

We first give an example why we cannot always interchange the order of integration and limit.

Example 4.1. *Consider the probability space* $((0,1), \mathcal{B}((0,1)), \lambda)$*, where* λ *is the Lebesgue measure, and let*

$$X_n(\omega) = \begin{cases} n, & \text{if } 0 < \omega \leq \dfrac{1}{n}, \\ 0, & \text{if } \dfrac{1}{n} < \omega < 1. \end{cases} \tag{2}$$

Then we have $\mathbb{E}X_n = 1$ *while* $\lim_{n\to\infty} X_n(\omega) = 0$ *for all* ω*.*

To prove the MCT, we first show the following.

Lemma 4.7. *Suppose that* $X \geq 0$ *is a simple random variable and* $B_n \subset B_{n+1}$ *for* $n \geq 1$*. Let* $B = \bigcup_{n\geq 1} B_n$*. We have*

$$\lim_{n\to\infty} \mathbb{E}[X\mathbf{1}_{B_n}] = \mathbb{E}[X\mathbf{1}_B].$$

Proof. Suppose

$$X = \sum_{i=1}^{m} x_i \mathbf{1}_{A_i},$$

then

$$X\mathbf{1}_{B_n} = \sum_{i=1}^{m} x_i \mathbf{1}_{A_i \cap B_n},$$

and from Lemma 4.3, we obtain

$$\mathbb{E}[X\mathbf{1}_{B_n}] = \sum_{i=1}^{m} x_i \mathbb{P}(A_i \cap B_n).$$

Since $\mathbb{E}[X\mathbf{1}_{B_n}] \leq \mathbb{E}\big[X\mathbf{1}_{\{B_{n+1}\}}\big]$ is an increasing sequence in $\mathbb{R}$, its limit exists (can be ∞). Taking limit as $n \to \infty$ on both sides, we obtain

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[X\mathbf{1}_{B_n}] &= \sum_{i=1}^{m} x_i \lim_{n\to\infty} \mathbb{P}(A_i \cap B_n) \\
&= \sum_{i=1}^{m} x_i \mathbb{P}\left(\bigcup_{n\geq 1} \{A_i \cap B_n\}\right) \\
&= \sum_{i=1}^{m} x_i \mathbb{P}(A_i \cap B) \\
&= \mathbb{E}[X\mathbf{1}_B].
\end{aligned}$$

□

Proof of MCT. Since $X_n \leq X$ for all $n \geq 1$, we have $\mathbb{E}X_n \leq \mathbb{E}X$ from Lemma 4.5. As $\mathbb{E}X_n \leq \mathbb{E}X_{n+1}$ is an increasing bounded sequence, its limit exists and

$$\lim_{n\to\infty} \mathbb{E}X_n \leq \mathbb{E}X.$$

Let Z be a simple random variable such that $0 \leq Z \leq X$. Let $0 < \rho < 1$, and define

$$B_n = \{\omega \,:\, X_n(\omega) \geq \rho Z(\omega)\}.$$

We have $B_n \subset B_{n+1}$. Furthermore, for n sufficiently large, we have $X_n(\omega) \geq \rho X(\omega) \geq \rho Z(\omega)$ $\forall\, \omega \in \Omega$ (i.e., $B_n = \Omega$), so that $\bigcup_{j\geq 1} B_j = \Omega$. For such n, we then have

$$\begin{aligned}
&\rho Z\mathbf{1}_{B_n} \leq X_n\mathbf{1}_{B_n} \leq X_m,\ \forall\, m \geq n, \\
&\rho\mathbb{E}[Z\mathbf{1}_{B_n}] \leq \mathbb{E}X_m,
\end{aligned}$$

from Definition 4.4 since $Z\mathbf{1}_{B_n}$ is simple. Letting $m \to \infty$,

$$\rho\mathbb{E}[Z\mathbf{1}_{B_n}] \leq \lim_{m\to\infty} \mathbb{E}X_m,$$

and $n \to \infty$, we have from Lemma 4.7,

$$\begin{aligned}
\rho\mathbb{E}\Big[Z\mathbf{1}_{\bigcup_{n\geq 1} B_n}\Big] &\leq \lim_{m\to\infty} \mathbb{E}X_m, \\
\rho\mathbb{E}[Z] &\leq \lim_{m\to\infty} \mathbb{E}X_m.
\end{aligned}$$

Taking sup over all simple $Z \leq X$, we have

$$\rho\mathbb{E}X \leq \lim_{n\to\infty} \mathbb{E}X_n,$$

and $\rho \to 1$ completes the proof. □

The following procedure gives an explicit construction of a sequence of increasing simple random variables that approximate $X \geq 0$. For each $k \geq 1$ and $0 \leq j < 2^{2k}$, let

$$\begin{aligned} B_{k,j} &= \left\{\omega \,:\, \frac{j}{2^k} < X(\omega) \leq \frac{j+1}{2^k}\right\}, \\ B_{k,2^{2k}} &= \left\{\omega \,:\, X(\omega) > 2^k\right\}, \end{aligned}$$

and

$$Z_k = \sum_{j=0}^{2^{2k}} \frac{j}{2^k} \mathbf{1}_{B_{k,j}}.$$

We have $Z_k(\omega) \leq Z_{k+1}(\omega)$ and

$$\begin{aligned} &0 \leq X(\omega) - Z_k(\omega) \leq 2^{-k}, \ \forall \omega \ \text{ s.t. } X(\omega) \leq 2^k, \\ &Z_k(\omega) = 2^k, \ X(\omega) > 2^k. \end{aligned}$$

Therefore $Z_k(\omega) \uparrow X(\omega)$ as $k \to \infty$. We finally have all the tools to prove Theorem 4.1.

Proof of Theorem 4.1. For any $X, Y \geq 0$, $\exists$ simple $X_n \uparrow X$ and simple $Y_n \uparrow Y$. We then have $X_n + Y_n \uparrow X + Y$ and $\mathbb{E}[X_n + Y_n] = \mathbb{E}X_n + \mathbb{E}Y_n$. From the MCT, taking $n \to \infty$, we obtain $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$. □

Note that additivity of expectations for general random variables follows from Theorem 4.1 since $\mathbb{E}X = \mathbb{E}X^+ - \mathbb{E}X^-$.

The steps in the proof of Theorem 4.1 are quite standard in measure theory:

(i) Prove for simple random variables.

(ii) Extend to non-negative random variables X by using simple random variables $\uparrow X$, then apply MCT.

(iii) Extend to general $X = X^+ - X^-$.

4.3 Fatou's Lemma and the DCT

Lemma 4.8 (Fatou's Lemma)**.** *Suppose that $X_n \geq 0$ for $n \geq 1$. We have*

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n\right] \leq \liminf_{n\to\infty} \mathbb{E}X_n.$$

Proof. Let $Y_k = \inf_{n \geq k} X_n \geq 0$. We have $Y_k \leq Y_{k+1}$ and $Y_k \leq X_n$, $\forall\, n \geq k$ so that

$$\begin{aligned} \mathbb{E}Y_k &\leq \inf_{n \geq k} \mathbb{E}X_n \\ \lim_{k\to\infty} \mathbb{E}Y_k &\leq \lim_{k\to\infty} \inf_{n\geq k} \mathbb{E}X_n = \liminf_{n\to\infty} \mathbb{E}X_n. \end{aligned}$$

From MCT and Lemma 1.8, we have

$$\begin{aligned} \lim_{k\to\infty} \mathbb{E}Y_k &= \mathbb{E}\Big[\lim_{k\to\infty} Y_k\Big] \\ &= \mathbb{E}\Big[\lim_{k\to\infty} \inf_{n\geq k} X_n\Big] \\ &= \mathbb{E}\Big[\liminf_{n\to\infty} X_n\Big], \end{aligned}$$

and we obtain

$$\mathbb{E}\Big[\liminf_{n\to\infty} X_n\Big] \leq \liminf_{n\to\infty} \mathbb{E}X_n.$$

$\square$

A similar proof as that for Lemma 4.8 shows that if $X_n \leq Y$ where $\mathbb{E}|Y| < \infty$, then

$$\mathbb{E}\Big[\limsup_{n\to\infty} X_n\Big] \geq \limsup_{n\to\infty} \mathbb{E}X_n.$$

Theorem 4.3 (Dominated Convergence Theorem (DCT))**.** *Suppose* $|X_n(\omega)| \leq Y(\omega)\ \forall\, \omega \in \Omega$, *and* $\mathbb{E}|Y| < \infty$. *If* $\lim_{n\to\infty} X_n = X$, *then* $\mathbb{E}X$ *exists,* $\mathbb{E}|X| \leq \mathbb{E}Y$ *and*

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

Proof. From Lemma 4.5, we have $\mathbb{E}|X_n| \leq \mathbb{E}Y$, and Fatou's Lemma (Lemma 4.8) yields $\mathbb{E}|X| \leq \mathbb{E}Y$. Therefore, $\mathbb{E}X$ exists. Since $Y + X_n \geq 0$, from Fatou's Lemma, we have

$$\mathbb{E}\Big[\liminf_{n\to\infty} (Y + X_n)\Big] \leq \liminf_{n\to\infty} \mathbb{E}[Y + X_n]. \tag{3}$$

We also have

$$\mathbb{E}\Big[\liminf_{n\to\infty} (Y + X_n)\Big] = \mathbb{E}[Y + X] = \mathbb{E}Y + \mathbb{E}X,$$

and

$$\liminf_{n\to\infty} \mathbb{E}[Y + X_n] = \mathbb{E}Y + \liminf_{n\to\infty} \mathbb{E}[X_n],$$

so that (3) yields

$$\mathbb{E}X \leq \liminf_{n\to\infty} \mathbb{E}X_n.$$

Similarly, we have $Y - X_n \geq 0$ and Fatou's Lemma implies that

$$\mathbb{E}\Big[\liminf_{n\to\infty}(Y - X_n)\Big] \leq \liminf_{n\to\infty}\mathbb{E}[Y - X_n],$$
$$\mathbb{E}X \geq \limsup_{n\to\infty}\mathbb{E}X_n.$$

Therefore, $\mathbb{E}X = \lim_{n\to\infty}\mathbb{E}X_n$. □

Lemma 4.9 (Scheffe)**.** *Suppose that $X_n \geq 0$ for $n \geq 1$, $\lim_{n\to\infty} X_n = X$, both X_n and X are integrable, and $\lim_{n\to\infty}\mathbb{E}X_n = \mathbb{E}X$. Then, $\lim_{n\to\infty}\mathbb{E}|X_n - X| = 0$.*

Proof. Since $X_n \geq 0$ and $X \geq 0$, we have

$$0 \leq (X_n - X)^- \leq X.$$

As $\mathbb{E}X < \infty$ and $(X_n - X)^- \to 0$ since $X_n \to X$, DCT yields $\mathbb{E}(X_n - X)^- \to 0$. From

$$\mathbb{E}(X_n - X) = \mathbb{E}(X_n - X)^+ - \mathbb{E}(X_n - X)^-,$$

and $\mathbb{E}(X_n - X) \to 0$, we obtain $\mathbb{E}(X_n - X)^+ \to 0$ and

$$\mathbb{E}|X_n - X| = \mathbb{E}(X_n - X)^+ + \mathbb{E}(X_n - X)^- \to 0.$$

□

Hence, any sequence of r.v.s satisfying the conditions of the DCT also converges in L^1.

4.4 Notes

Suppose $f : (\Omega, \mathcal{A}, \mu) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a non-negative measurable function, where μ is a general measure (i.e., does not have the restriction that $\mu(\Omega) = 1$), then the Lebesgue integral of f over a measurable set $A \in \mathcal{A}$ denoted by

$$\int_A f \, \mathrm{d}\mu$$

is defined in exactly the same way as in Definition 4.4. In fact, the proofs of MCT, Fatou's Lemma and DCT do not require that the underlying measure is a probability measure! Therefore, all the results we have seen so far are true for the general Lebesgue integral. If f is a.e. continuous, then its Lebesgue integral and Riemann integral give the same value, but their constructions are different.

For $p \geq 1$, we define

$$L^p(\Omega, \mathcal{A}, \mu) = \left\{ f : (\Omega, \mathcal{A}, \mu) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R})) \,\middle|\, \int_\Omega |f|^p \, \mathrm{d}\mu < \infty \right\}.$$

If obvious from the context, we usually shorten the notation to $L^p(\mu)$ or $L^p(\Omega)$. An integrable function means it is in $L^1(\mu)$. Using Minkowski's inequality, it can be shown that

$$\|f\|_p = \left(\int |f|^p \,\mathrm{d}\mu\right)^{1/p}$$

is a norm, hence $L^p(\mu)$ is a normed space.

If g is a measurable function, then from Definition 4.3, we have

$$\mathbb{E}g(X) = \int g(X(\omega)) \,\mathrm{d}\mathbb{P}(\omega).$$

Letting $x = X(\omega)$, we obtain

$$\begin{aligned}\mathbb{E}g(X) &= \int g(x) \,\mathrm{d}\mathbb{P} \circ X^{-1}(x) \\ &= \int g(x) \,\mathrm{d}\mathbb{P}_X(x). \end{aligned} \tag{4}$$

Note that $X^{-1}(B) = \{\omega : X(\omega) \in B\}$ is defined as the inverse map for Borel sets, and is different from a traditional inverse function, which requires bijectiveness. In particular, $X^{-1}(\cdot)$ is always well-defined.

For discrete random variables $X \in \{x_1, x_2, \ldots\}$, suppose $\mathbb{P}_X(\{x_i\}) = p_i$, a discrete measure. Then the expectation (4) reduces to a sum $\sum_i g(x_i)p_i$. For "continuous" random variables you are familiar with in undergrad probability courses, we need a definition.

Definition 4.5. *If $\mu(A) = 0 \implies \nu(A) = 0$ for all measurable sets A, then we say that ν is absolutely continuous w.r.t. μ and write $\nu \ll \mu$.*

A measure μ is σ-finite if $\Omega = \bigcup_{i\geq 1} \Omega_i$, where the Ω_i are disjoint and $\mu(\Omega_i) < \infty$.

Theorem 4.4 (Radon-Nikodym)**.** *Suppose the ν is finite and μ is σ-finite. If $\nu \ll \mu$, then there exists measurable $f \geq 0$ with $\int |f| \,\mathrm{d}\mu < \infty$ (i.e., $f \in L^1(\mu)$) such that*

$$\nu(A) = \int_A f \,\mathrm{d}\mu.$$

Notation-wise, we usually write $f = \dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$, which is called the Radon-Nikodym derivative.

Definition 4.6. *X is a continuous random variable if $\exists\, f \in L^1(\mathbb{R}, \lambda)$, where λ is the Lebesgue measure, i.e., $\lambda([a, b]) = b - a$, s.t.*

$$\mathbb{P}_X(A) = \int_A f \,\mathrm{d}\lambda = \int_A f(x) \,\mathrm{d}x$$

for all measurable A. The function f is called the probability density function (pdf) of X.

We also have $\mathbb{E}g(X) = \int g(x) \,\mathrm{d}\mathbb{P}_X(x) = \int g(x)f(x) \,\mathrm{d}x$.