# probability theory

Runi Malladi

March 7, 2023

## 1 unsorted

**Definition 1.1.** Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measure spaces. A *transition probability* from $E$ to $F$ is a function

$$\nu : E \times F \to [0, 1]$$

such that, for $x \in E$ and $A \in \mathcal{F}$,

1. $v_x(-)$ is a probability measure on $(F\mathcal{F})$ for all $x \in E$
2. $v_A(-)$ is $\mathcal{E}$-measurable for all $A \in \mathcal{F}$.

**Example 1.2.** Let $(F, \mathcal{F}, \mu)$ be $\sigma$-finite, and $f : E \times F \to \mathbb{R}_+$ be $\mathcal{E} \otimes \mathcal{F}$-measurable. If

$$\int_F f(x, y)\ \mu(dy) = 1$$

for all $x \in E$, then

$$v(x, A) = \int_A f(x, y)\ \mu(dy)$$

is a transition probability from $E$ to $F$. To verify the two conditions:

1. follows by assumption
2. by Fubini-Tonelli (Theorem 9.6 in *Measure Theory*, strengthed by Remark ).

do this remark

## 2 basic definitions

**Definition 2.1.** Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a measurable space $(E, \mathcal{E})$, an *$E$-valued random variable* is a measurable function $X : \Omega \to E$.

**Definition 2.2.** The *law* or *distribution* of $X$ is the pushforward of $\mathbb{P}$ along $X$, i.e. the measure on $(E, \mathcal{E})$ characterized by the equality

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$$

for all $B \in \mathcal{E}$.

**Definition 2.3.** The *expectation* of a random variable $X$ is defined as

$$\mathbb{E}[X] = \int_{\Omega} X \; d\mathbb{P},$$

provided this integral makes sense (i.e. either $X \geq 0$ or $\mathbb{E}[|X|] < \infty$).

**Proposition 2.4.** Let $X : \Omega \to [0, \infty]$ be a random variable. Then

$$\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X \geq x) dx.$$

Let $Y : \Omega \to \mathbb{Z}_+$ be a random variable. Then

$$\mathbb{E}[Y] = \sum_{k=0}^{\infty} k\mathbb{P}(Y = k) = \sum_{k=1}^{\infty} \mathbb{P}(Y \geq k).$$

*Proof.* Note

$$\int_0^{\infty} 1_{\{x \leq X(\omega)\}} dx = \lambda(\{0 \leq x \leq X(\omega)\}) = X(\omega),$$

and so

$$\mathbb{E}[X] = E\left[\int_0^{\infty} 1_{\{x \leq X\}} dx\right] = \int_0^{\infty} \int_{\Omega} 1_{\{x \leq X(\omega)\}} \mathbb{P}(d\omega) dx = \int_0^{\infty} \mathbb{P}(x \leq X) dx.$$

Swapping the integrals above was justified by Tonelli's theorem, since the function we are integrating is nonnegative.

For $Y$, observe

$$Y = \sum_{k=0}^{\infty} k 1_{\{Y(\omega)=k\}} = \sum_{k=1}^{\infty} 1_{\{Y(\omega) \geq k\}}.$$

Then the result follows by another application of Tonelli's theorem, e.g. by viewing the infinite sum as an integral with respect to counting measure. $\square$

# 3 conditional probability

There are several classes of objects we can take conditional probability with respect to, which can cause some confusion:

- ... of $A \in \mathcal{A}$ with respect to $B \in \mathcal{A}$, written $\mathbb{P}(A \mid B)$. Letting $A$ vary, this is a probability measure on $(\Omega, \mathcal{A})$ (Definition 3.1). Given a random variable $X$, We can also consider a conditional expectation $\mathbb{E}[X \mid B]$ (Definition 3.2), which is nothing but the expectation of $X$ with respect to the aforementioned conditional probability measure (Proposition 3.3).
- ... of a real random variable $X$ with respect to a random variable $Y$ taking values in a countable space (Definition 3.4).

**Definition 3.1.** Consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. For any $B \in \mathcal{A}$ with nonzero measure, we can define a new probability measure on $(\Omega, \mathcal{A})$ as follows:

$$\mathbb{P}(\cdot \mid B) \coloneqq \frac{\mathbb{P}(\cdot \cap B)}{\mathbb{P}(B)}.$$

This is the *conditional probability* given $B$.

*Proof this is a probability measure.* By (Measure Theory, Proposition 8.1), we see that $\mathbb{P}(\cdot \mid B)$ is an unsigned measure, and direct calculation shows

$$\mathbb{P}(\Omega \mid B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1.$$

□

**Definition 3.2.** As above, consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $B \in \mathcal{A}$ with nonzero measure. For a random variable $X$, we define the *conditional expectation*

$$\mathbb{E}[X \mid B] \coloneqq \frac{\mathbb{E}[X 1_B]}{\mathbb{P}(B)}.$$

**Proposition 3.3.** $\mathbb{E}[X \mid B]$ is the expectation of $X$ with respect to the probability measure $\mathbb{P}(\cdot \mid B)$.

*Proof.* By (Measure Theory, Proposition 8.1),

$$\int_\Omega X \, d\mathbb{P}(\cdot \mid B) = \frac{1}{\mathbb{P}(B)} \int_B X \, d\mu = \frac{\mathbb{E}[X 1_B]}{\mathbb{P}(B)}.$$

□

**Definition 3.4.** Let $(E, \mathcal{P}(E))$ be a countable space, and $Y : \Omega \to E$ a random variable. Let $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ be a real-valued random variable. Define

$$E' = \{y \in E : \mathbb{P}(Y = y) > 0\}$$

be the set of outcomes with nonzero probability. In particular, for all $y \in E'$, we can define the conditional expectation (Definition 3.2)

$$\mathbb{E}[X \mid Y = y] = \frac{\mathbb{E}[X 1_{Y=y}]}{\mathbb{P}(Y = y)}.$$

This induces a function $\phi_X : E \to \mathbb{R}$ by

$$\phi_X(y) = \begin{cases} \mathbb{E}[X \mid Y = y] & y \in E' \\ 0 & \text{otherwise} \end{cases}.$$

Finally, we define the *conditional expectation* of $X$ with respect to $Y$ to be the random variable $\Omega \to \mathbb{R}$

$$\mathbb{E}[X \mid Y](-) = \phi_X(Y(-)).$$

**Remark 3.5.** The countability of $E$ makes this definition canonical up to a set of measure zero. In particular, we can choose any value of $\phi_X(y)$ for $y \notin E'$, since

$$\mathbb{P}(Y \notin E') = \mathbb{P}\left(\bigcup_{y \notin E'} \{Y = y\}\right) = \sum_{y \notin E'} \mathbb{P}(Y = y) = 0,$$

where we have used the countable additivity of measure.

**Remark 3.6.** We may visualize Definition 3.4 as follows:

$$\begin{tikzcd}
(E, \mathcal{P}(E)) \arrow[r, "\phi_X"] & \mathbb{R} \arrow[dd, -, dotted, "?"] \\
(\Omega, \mathcal{A}, \mathbb{P}) \arrow[u, "Y"] \arrow[ur, "\mathbb{E}[X|Y]"'] \arrow[dr, "X"] & \\
& \mathbb{R}
\end{tikzcd}$$

The question now becomes how the random variables $\mathbb{E}[X \mid Y]$ and $X$ are related to each other.

**Definition 3.7.** Let $\mathcal{B} \subset \mathcal{A}$ be a sub-$\sigma$-algebra, and let $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ (i.e. $X$ is an absolutely integrable real random variable). Then the *conditional expectation* $\mathbb{E}[X|\mathcal{B}] \in L^1(\Omega, \mathcal{B}, \mathbb{P})$ is the unique random variable such that

$$\mathbb{E}[XZ] = \mathbb{E}[\mathbb{E}[X|\mathcal{B}]Z]$$

for all $\mathcal{B}$-measurable real random variables $Z$.

**Theorem 3.8.** $\mathbb{E}[X|\mathcal{B}]$ exists and is unique.

*Proof.* □ prove

## 4 stochastic processes

**Definition 4.1.** A *discrete random process* is a sequence $(X_n)_{n \in \mathbb{Z}_+}$ of random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ taking values on the same measurable space.

**Definition 4.2.** A *filtration* on $(\Omega, \mathcal{A}, \mathbb{P})$ is an increasing sequence of sub-$\sigma$-algebras $(\mathcal{F}_n)_{n \in \mathbb{Z}_+}$, i.e.

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{A}.$$

The probability space together with a filtration is called a *filtered probability space*.

**Definition 4.3.** A random process $(X_n)_{n \in \mathbb{Z}_+}$ is *adapted* to the filtration $(\mathcal{F}_n)_{n \in \mathbb{Z}_+}$ if $X_n$ is $\mathcal{F}_n$ measurable for all $n$.

**Definition 4.4.** Let $(X_n)_{n \in \mathbb{Z}_+} \subset L^1$ be a real-valued random process adapted to the filtration $(\mathcal{F}_n)_{n \in \mathbb{Z}_+}$. We say $(X_n)$ is *martingale* if $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$ for all $n$.

**Definition 4.5.** A (continuous) random variable $T : \Omega \to [0, \infty]$ is a *stopping time* (of the filtration $(\mathcal{F}_t)_{t\geq 0}$) if

$$\{T \leq t\} = \{\omega \in \Omega : T(\omega) \leq t\} \in \mathcal{F}_t$$

for all $t \geq 0$.

**Remark 4.6.** For intuition, consider the probability space to be the possible paths/evolution of a casino game. The time at which the player decides to leave is a stopping time, since it will only be influenced by events that occurred up to the moment they decide to leave, and not on future events.

# 5 Brownian motion

**Definition 5.1.** An $\mathbb{R}^d$-valued continuous random process $(B_t)_{t\geq 0}$ is called ($d$-dimensional) *Brownian motion* or a *Wiener process* if the following conditions hold:

BR1) $B_0 = 0$ almost surely.

BR2) (independent increments) future increments are independent of past increments, i.e. for all $u, t > 0$, it is the case that $B_{t+u} - B_t$ is independent of $B_s$ for $t > s$.

BR3) (Gaussian increments) $B_{t+u} - B_t \sim \mathcal{N}(0, u)$.

BR4) (continuous paths) $B_t$ is continuous in $t$.

**Theorem 5.2.** Brownian motion exists.

*Proof.* $\square$ prove

# References


[1] Jean-François LeGall. *Measure theory, probability, and stochastic processes*. Springer International Publishing, 2022.

[2] *Wiener process*. Feb. 2023. URL: https://en.wikipedia.org/wiki/Wiener_process.