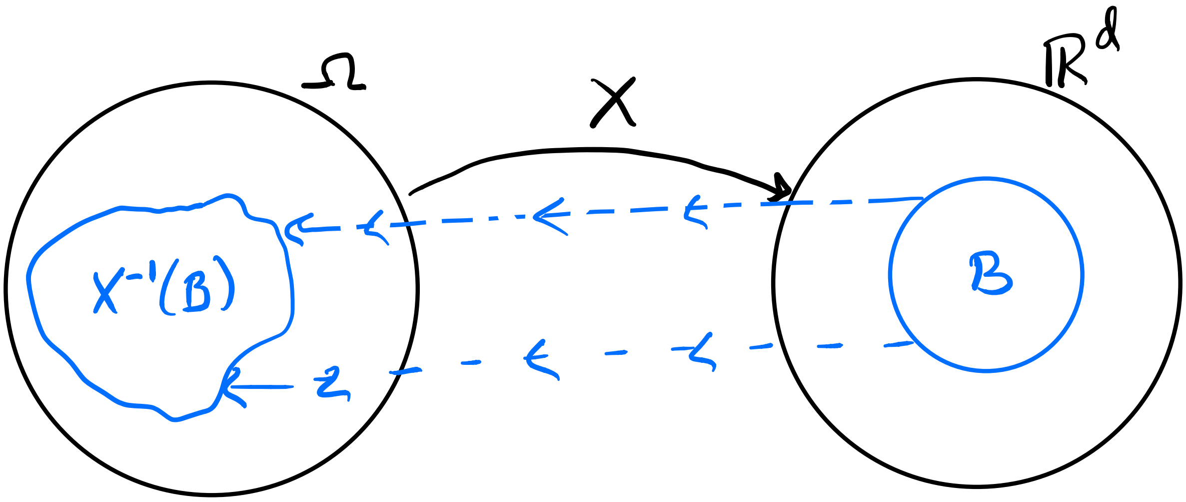
Recap

Random Variables: X is a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$ if $X : \Omega \to \mathbb{R}^d$ with $\underline{X^{-1}(B) \in \mathcal{F}}$ $\forall B \in \mathcal{B}(\mathbb{R}^d)$



$(\Omega, \mathcal{F}, \mathbb{P})$ $\qquad\qquad$ $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_X)$

where

$$\mathbb{P}_X(B) \triangleq \mathbb{P}(\underbrace{X^{-1}(B)}_{\in \mathcal{F}}) \qquad B \in \mathcal{B}(\mathbb{R}^d)$$

CDF

Define

$$F_X : \mathbb{R}^d \to [0,1]$$

by

$$F_X(a_1, \ldots, a_d) \triangleq \mathbb{P}_X\big((-\infty, a_1] \times \cdots \times (-\infty, a_d]\big)$$

Properties of CDFs

Let $F_X(x)$ be the cumulative distribution function of a real-valued random variable X (i.e. $X: \Omega \to \mathbb{R}, d=1$).

Then

(i) $\lim_{t\to\infty} F_X(t) = 1$

(ii) $\lim_{t\to-\infty} F_X(t) = 0$

(iii) F_X is monotonically non-decreasing and right continuous.

(iv) $\mathbb{P}_X((a,b]) = F_X(b) - F_X(a)$

Discrete vs. Continuous Random Variables

Definition (support): the support $\mathbb{P}_X$ denoted by $\mathrm{Supp}(\mathbb{P}_X)$ is the smallest Borel set $B \in \mathcal{B}(\mathbb{R}^d)$ s.t. $\mathbb{P}_X(B) = 1$.

1. Discrete Random Variables: X is discrete when $\mathrm{supp}(\mathbb{P}_X)$ is countable.

   Then we can define the pmf

$$p_X: \mathrm{supp}(\mathbb{P}_X) \to [0,1]$$

associated with X.

$$p_X(x) \triangleq \mathbb{P}_X(\{x\}) = \lim_{\varepsilon\to 0} F_X(x) - F_X(x-\varepsilon), \quad x \in \mathrm{supp}(\mathbb{P}_X)$$

(2) Continuous Random Variables: When $\mathrm{supp}(\mathbb{P}_X)$ is uncountable **AND** F_X is differentiable on the interior of the support, then X is a continuous random variable.

The probability density function (pdf) of X is

$$f_X(t) = \frac{dF_X(t)}{dt}$$

$$\Rightarrow \quad \mathbb{P}_{f_X} = \mathbb{P}_X$$

Expectation, Variance, Etc

Expectation: Let X be a rv on $(\Omega, \mathcal{F}, \mathbb{P})$. The expectation of X is

$$\mathbb{E}[X] \triangleq \int_{\mathbb{P}_d} x \, d\mathbb{P}_X(x)$$

where

(1) if X is discrete

$$\mathbb{E}[X] = \sum_{x \in \mathrm{supp}(\mathbb{P}_X)} x \cdot p_X(x)$$

② if X is continuous

$$\mathbb{E}[X] = \int_{\mathrm{Supp}(\mathbb{P}_X)} x \cdot f_X(x)\, dx$$

Proposition - Expectation of a Function

Let

$$X : \Omega \to \mathbb{R}^d$$

be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and

$$g : \mathbb{R}^d \to \mathbb{R}^{d'}$$

be a measurable function.

Then

$$\mathbb{E}[g(X)] \triangleq \int_{\mathbb{R}^d} g(x)\, d\mathbb{P}_X(x)$$

or $\mathrm{supp}(\mathbb{P}_X)$

Variance: Let X be a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$. The variance of X is

$$\mathrm{Var}(X) \triangleq \mathbb{E}[\underbrace{(X - \mathbb{E}[X])^2}_{\text{just think of this as } g(X)!}]$$

Law of Large Numbers

Let $X_1, X_2, \ldots$ be a sequence of $\mathbb{R}$-valued iid r.v.'s on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

1 Weak LLN

If $\mathbb{E}[|X_i|] < \infty \ \forall i$, then

$$\mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{\infty} X_i - \mathbb{E}[X_i] \right| > \varepsilon \right) \xrightarrow[n \to \infty]{} 0 \quad \forall \varepsilon > 0$$

2 Weak LLN for Functions

Let $f: \mathbb{R} \to \mathbb{R}$ be a measurable function such that $\mathbb{E}[f(X_i)] < \infty \ \forall i$, then

$$\mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{n} f(X_i) - \mathbb{E}[f(X_i)] \right| > \varepsilon \right) \xrightarrow[n \to \infty]{} 0 \quad \forall \varepsilon > 0$$

3 Uniform Weak LLN

Let $f_1, \ldots, f_k : \mathbb{R} \to \mathbb{R}$ be measurable functions with $\mathbb{E}[f_k(X_1)] < \infty \ \forall k \in \{1, \ldots, k\}$.

Then

$$\mathbb{P}\left(\bigcup_{l=1}^{k} \left\{ \left| \frac{1}{n} \sum_{i=1}^{n} f_l(X_i) - \mathbb{E}[f_l(X_i)] \right| \right\} > \varepsilon \right) \xrightarrow[n \to \infty]{} 0 \quad \forall \varepsilon > 0$$

Conditional Probability and Transition Kernels

Transition Kernel: Let $(\mathcal{X}, \mathcal{F})$ and $(\mathcal{Y}, \mathcal{G})$ be two measurable spaces. A transition kernel

$$K(\cdot \mid \cdot) : \mathcal{G} \times \mathcal{X} \longrightarrow \mathbb{R}$$

such that

set of ALL probability measures over target space $\mathcal{Y}$

① $K(\cdot \mid x) \in \mathcal{P}(\mathcal{Y})$, $\forall x \in \mathcal{X}$

② $K(B \mid \cdot) : \mathcal{X} \to \mathbb{R}$ is $\mathcal{F}$-measurable
(i.e. a random variable wrt $(\mathcal{X}, \mathcal{F})$), $\forall B \in \mathcal{G}$

Aside: Notation
We will typically use $P_{Y|X}$ for a transition kernel from $(\mathcal{X}, \mathcal{F})$ to $(\mathcal{Y}, \mathcal{G})$, to stress the underlying spaces.

Notation Switch: $\mathbb{P} \leftrightarrow P$

Comment: Transition kernels from $\mathcal{X}$ to $\mathcal{Y}$ should be thought of as random transformations of an $\mathcal{X}$-valued random variable $X \sim P_X$ into a $\mathcal{Y}$-valued random variable $Y \sim P_Y$.

Visually,

$$X \sim P_X \longrightarrow \boxed{P_{Y|X}} \longrightarrow Y \sim P_Y$$

? What we want

Indeed,

$$P_Y(B) \triangleq \mathbb{E}_{P_X}\left[K(B|X)\right] \qquad B \in \mathcal{G}$$

$$= \int_{\mathcal{X}} K(B|x)\, dP_X(x)$$

Proposition: P_Y as defined above is a probability measure on $(\mathcal{Y}, \mathcal{G})$

<u>Proof</u>

(i) $$P_Y(\mathcal{Y}) = \int_{\mathcal{X}} \underbrace{K(\mathcal{Y}|x)}_{1 \text{ for } \sum_{x\in\mathcal{X}}} dP_X(x)$$

$$= \int_{\mathcal{X}} dP_X(x) = 1$$

(ii)

$$\mathbb{P}_Y\left(\bigcup_{n=1}^{\infty} B_n\right) = \int_{\mathcal{X}} K\left(\bigcup_{n=1}^{\infty} B_n \,\middle|\, x\right) d\mathbb{P}_X(x)$$

$$= \int_{\mathcal{X}} \sum_{n=1}^{\infty} K(B_n | x)\, d\mathbb{P}_X(x)$$

Upper bounded by 1

(Fubini's Theorem)

$$= \sum_{n=1}^{\infty} \int_{\mathcal{X}} K(B_n | x)\, d\mathbb{P}_X(x)$$

$$= \sum_{n=1}^{\infty} \mathbb{P}_Y(B_n)$$

Conditional Expectation

$h(x) = \mathbb{E}[Y | X = x]$
$h(X) = \mathbb{E}[Y | X]$
} what is introduced in other courses

Definition - Conditional Expectation

Let $K(\cdot|\cdot)$ be a transition kernel from $\mathcal{X}$ to $\mathcal{Y}$.

Define

① $\mathbb{E}[Y | X = x] \triangleq \int_{\mathcal{Y}} y \, dK(y | x)$

② $\mathbb{E}[Y | X] \triangleq \int_{\mathcal{Y}} y \, dK(y | X)$

Theorem - Law of Total Expectation

$$\mathbb{E}[Y] = \mathbb{E}_{P_X}\left[\mathbb{E}[Y|X]\right]$$