

Announcements

① Midterm March 17th 7:30PM - 10 PM

② Lecture on Feb. 4th is cancelled

Lecture on Feb. 6th is moved to Friday
Feb. 7th from 2:55-4:10 @ 203 Phillips
Hall

Recap

- Probability Spaces
 - Discrete $(\Omega, 2^\Omega, P_p)$, p is pmf
 - Continuous $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), P_f)$, f is pdf
- Conditional Probability Spaces
 - Bayes' thm
- ⇒ - Law of Total Prob.

Comments on Borel σ -Algebra

• Impossibility Result: It is impossible to define the uniform (aka Lebesgue) measure on $\mathcal{F} = 2^{\mathbb{R}^d}$

⇒ Thus need to find a smaller σ -Algebra that precludes pathological (non-measurable) sets,

but contains the "interesting" ones.

What is an "interesting set"?

→ This is really up to us.

Provided an arbitrary collection \mathcal{E} of subsets of Ω , how do we expand \mathcal{E} into a σ -algebra?
(Preferably, a "small" one at that)

Theorem (Generated σ -Algebra)

Given an arbitrary collection \mathcal{E} of subsets of Ω , there exists a UNIQUE SMALLEST σ -Algebra that contains all the elements of \mathcal{E} , denoted by $\sigma(\mathcal{E})$.

That is, $\sigma(\mathcal{E})$ is the smallest if for any other σ -Algebra \mathcal{H} s.t. $\mathcal{E} \subseteq \mathcal{H}$, we have that $\sigma(\mathcal{E}) \subseteq \mathcal{H}$.

Proof] Let $\{\mathcal{F}_i\}_{i \in \mathbb{Z}}$ be the collection of all σ -Algebras over Ω that contain \mathcal{E}

Define

$$\sigma(\mathcal{E}) \triangleq \bigcap_{i \in \mathbb{Z}} \mathcal{F}_i$$

← intersection is unique

[Exercise: Show an arbitrary intersection of σ -Algebras is ALSO a σ -Algebra.]

To show smallness, let \mathcal{H} be a σ -algebra w/ $\mathcal{E} \subseteq \mathcal{H}$.
Indeed $\exists i \in \mathbb{Z}$ s.t. $\mathcal{F}_i = \mathcal{H} \Rightarrow \sigma(\mathcal{E}) \subseteq \mathcal{H}$.

Definition (Borel σ -Algebra)

Let $\Omega = \mathbb{R}^d$ and

$$\mathcal{E}_0 \triangleq \{(-\infty, a_1] \times \dots \times (-\infty, a_d] \mid a_1, \dots, a_d \in \mathbb{R}\}$$

The Borel σ -Algebra on \mathbb{R}^d is

$$\mathcal{B}(\mathbb{R}^d) \triangleq \sigma(\mathcal{E}_0).$$

Elements of $\mathcal{B}(\mathbb{R}^d)$ are called Borel measurable sets, or simply Borel sets.

Dedekind cuts?

It can be shown that $\mathcal{B}(\mathbb{R}^d) \subseteq 2^{\mathbb{R}^d}$

Conditional Probability Spaces

From Last Time

$$P(\cdot | A) : \mathcal{F} \rightarrow [0, 1]$$

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

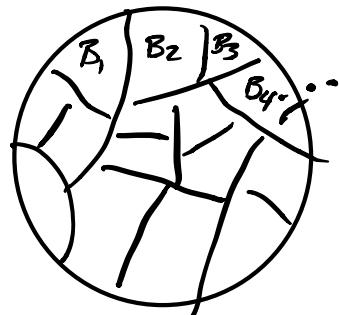
$$P(A|B) = \frac{P(B|A) P(A)}{P(B)} \quad (\text{Bayes Theorem})$$

Proposition: Law of Total Probability

Let (Ω, \mathcal{F}, P) be a probability space and $A, B_1, B_2, \dots, \in \mathcal{F}$ such that $\{B_n\}_{n=1}^{\infty}$ is a partition of Ω

$$(i) B_n \cap B_m = \emptyset$$

$$(ii) \bigcup_{n=1}^{\infty} B_n = \Omega$$



Then

$$P(A) = \sum_{n=1}^{\infty} P(B_n) P(A|B_n)$$

[Exercise: Prove the above

$$A = A \cap \Omega$$

Random Variables (aka Measurable Functions)

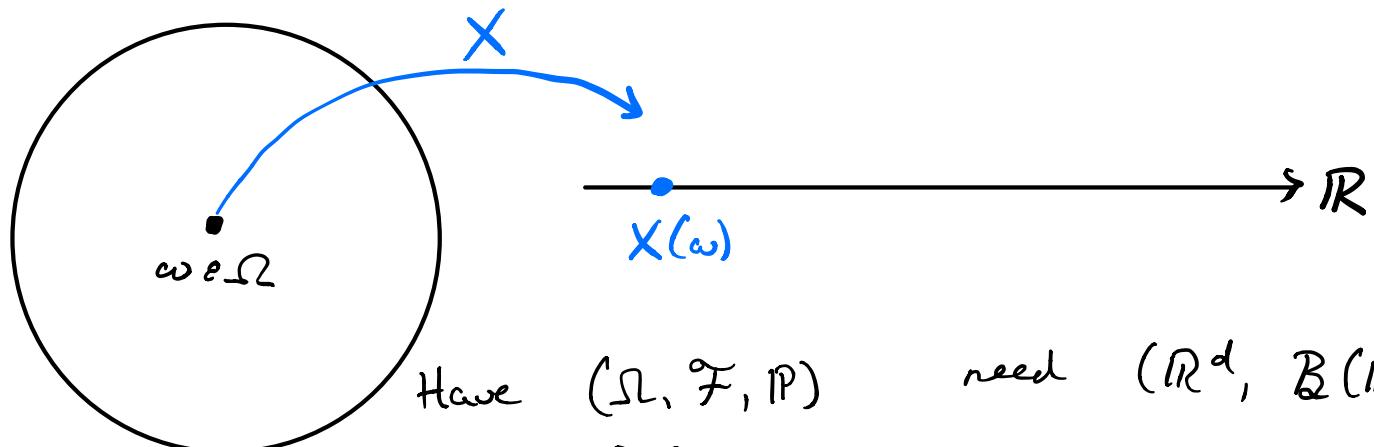
Let (Ω, \mathcal{F}, P) be a prob. space.

$$\Omega = \{ \square, \boxed{\cdot}, \boxed{\cdot\cdot}, \boxed{\cdot\cdot\cdot}, \boxed{\cdot\cdot\cdot\cdot}, \boxed{\cdot\cdot\cdot\cdot\cdot}, \boxed{\cdot\cdot\cdot\cdot\cdot\cdot} \}$$

\downarrow \downarrow \downarrow \downarrow \downarrow \downarrow
-2 -1 0 1 2 3

First and foremost, random variables are functions

$$X: \Omega \rightarrow \mathbb{R}^d$$
$$\omega \mapsto \mathbb{R} \times \dots \times \mathbb{R}$$



We would like to ask questions such as
 $P("gain") = P(X > 0)$, but P can only take elements
of \mathcal{F} as inputs.

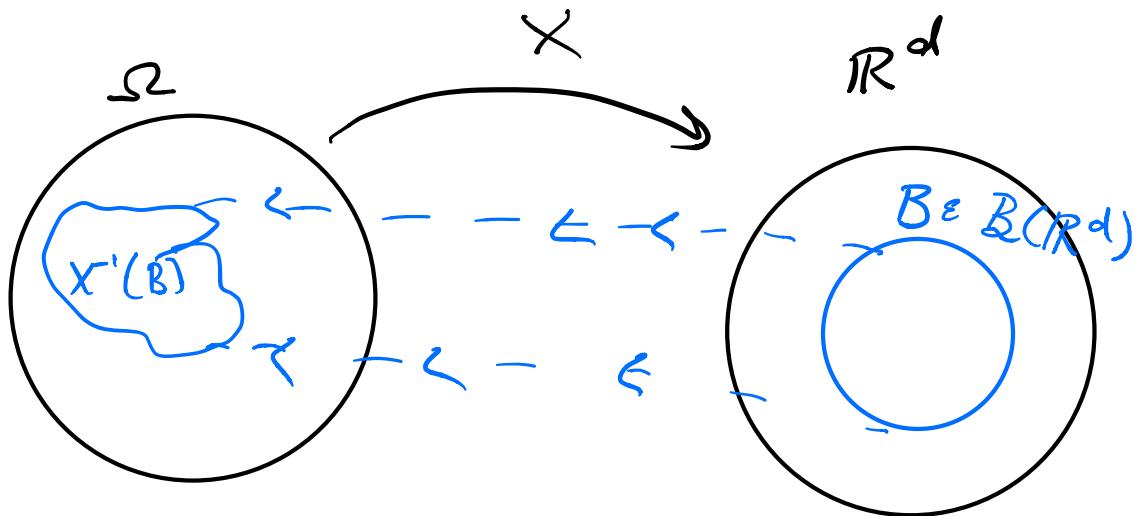
$$P("Gain") = P(\{\square, \boxed{\cdot}, \boxed{\cdot\cdot}\})$$

Generally, we want to account for quantities as $\mathbb{P}(X \in B)$, $B \in \mathcal{B}(\mathbb{R}^d)$ and so $\mathbb{P}(X \in B)$ is merely a notation for

$$\mathbb{P}(X \in B) \triangleq \mathbb{P}(X^{-1}(B))$$

where

$$X^{-1}(B) \triangleq \{\omega \in \Omega \mid X(\omega) \in B\}$$



Thus we need our random variable $X: \Omega \rightarrow \mathbb{R}^d$ to satisfy that $X^{-1}(B) \in \mathcal{F}$, $\forall B \in \mathcal{B}(\mathbb{R}^d)$

Definition (Random Variable)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a prob. space. X is a r.v. w.r.t $(\Omega, \mathcal{F}, \mathbb{P})$ if

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

satisfies

$$X^{-1}(B) \in \mathcal{F} \quad \forall B \in \mathcal{B}(\mathbb{R}^d)$$

and we define

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

as the probability law/distribution of X w.r.t P .

\Rightarrow The image probability space under a r.v. X is

$$(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), P_X)$$

Cumulative Distribution Functions

Given a r.v. X over (Ω, \mathcal{F}, P) the CDF of $X: \Omega \rightarrow \mathbb{R}^d$ is denoted by

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

and is defined as

$$F_X(a_1, \dots, a_d) = P_X((-\infty, a_1] \times \dots \times (-\infty, a_d])$$

It can be shown that F_X uniquely defines the law of X , P_X .

Discrete vs. Continuous Random Variables

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

① Discrete Random Variable: X is discrete when $\text{Supp}(P_X)$ is countable.

Then we can define the pmf

$$P_x: \text{Supp}(P_x) \rightarrow [0, 1]$$

associated with X .

$$p_x(x) \triangleq P_x(\{x\}) = \lim_{\varepsilon \rightarrow 0} F_x(x) - F_x(x-\varepsilon), \quad x \in \text{Supp}(P_x)$$

Note: One can check

$$R_{P_x} = P_x$$

Aside Showing $R_{P_x} = P_x$ (Done in OH)

First lets figure out what exactly we are even showing.

Two situations

① Have (Ω, \mathcal{F}) . Given a pmf P ,

$$P: \Omega \rightarrow [0, 1] \quad , \quad \sum_{\omega \in \Omega} = 1$$

can define

$$P_P(A) \triangleq \sum_{\omega \in A} P(\omega) \quad , \quad A \in \mathcal{F}$$

② Have (Ω, \mathcal{F}, P) and a rv $X: \Omega \rightarrow \mathbb{R}^d$

i.e $(\Omega, \mathcal{F}, P) \xrightarrow{X} (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), P_X)$

where

$$P_X(B) \triangleq P(X^{-1}(B)) , \quad B \in \mathcal{B}(\mathbb{R}^d)$$

and

$$X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\}$$

From here we motivated the definition of

$$S = \text{supp}(X)$$

and the pmf of X

$$p_X(x) \triangleq P_X(\{x\}) = P(X^{-1}\{x\}) = P(\{\omega \in \Omega \mid X(\omega) = x\})$$

Now note that in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), P_X)$ there exist zero probability events. To narrow this down we could take

$$(S, 2^S, ?)$$

In particular, take an event

$$A = B \cap \text{supp}(P_X) , \quad B \in \mathcal{B}(\mathbb{R}^d)$$

equivalent

Event A can be written as the union of the singleton events contained in A.

Well, then

$$P_{P_X}(A) = \sum_{a \in A} P_X(a)$$

and

$$P_X(A) = P_X\left(\bigcup_{i=1}^{\infty} \{a_i\}\right) \stackrel{\text{or } n; \text{ countable union}}{=} \sum_{i=1}^{\infty} P_X(\{a_i\})$$

But $P_X(\{a_i\})$ is precisely $P_X(a_i)$!

i.e.

$$P_X(A) = \sum_{i=1}^{\infty} P_X(a_i) = \sum_{a \in A} P_X(a)$$

So

$$P_{P_X}(A) = P_X(A)$$

That is, the push forward measure of P through X is the same as the probability measure induced by the pmf of X .