Announcements

① Midterm March 17th 7:30PM - 10PM

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 $(\Omega, 2^{\Omega}, \mathbb{P}_p)$, p is pmf
  - Continuous $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_f)$, f is pdf
- Conditional Probability Spaces
  - Bayes' thm

$\Rightarrow$ - Law of Total Prob.

Comments on Borel σ-Algebra

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

$\Rightarrow$ 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 $\mathcal{C}$ of subsets of Ω, how do we expand $\mathcal{C}$ into a σ-algebra? (Preferably, a "small" one at that)

Theorem (Generated σ-Algebra)

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

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

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

Define

intersection is unique

$$\sigma(\mathcal{C}) \triangleq \bigcap_{i \in \mathcal{I}} \mathcal{F}_i$$

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

To show smallness, let $\mathcal{H}$ be a σ-algebra w/ $\mathcal{C} \subseteq \mathcal{H}$.

Indeed $\exists\, i \in \mathcal{I}$ s.t. $\mathcal{F}_i = \mathcal{H} \Rightarrow \sigma(\mathcal{C}) \subseteq \mathcal{H}$.

∎

Definition (Borel σ-Algebra)

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

$$\mathcal{C}_0 \triangleq \{(-\infty, a_1] \times \cdots \times (-\infty, a_d] \mid a_1, \ldots, a_d \in \mathbb{R}\}$$

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

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

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

Dedekind cuts?

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

Conditional Probability Spaces

From Last Time

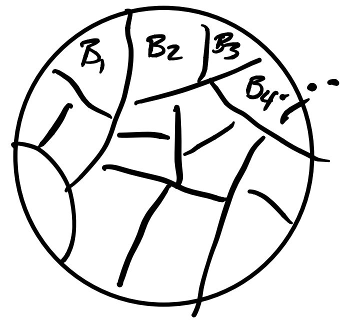
$$\mathbb{P}(\cdot \mid A) : \mathcal{F} \to [0,1]$$

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

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B)}$$ (Bayes Theorem)

Proposition: Law of Total Probability

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



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

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

Then

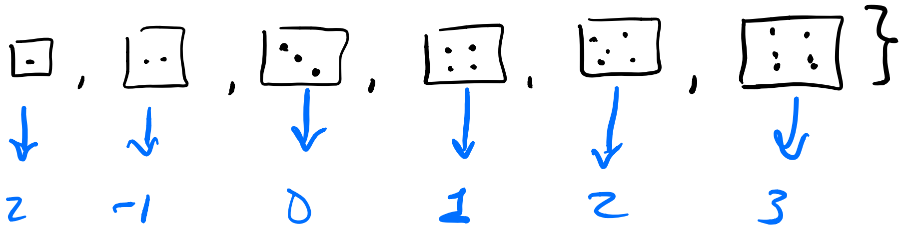
$$\mathbb{P}(A) = \sum_{n=1}^{\infty} \mathbb{P}(B_n)\, \mathbb{P}(A \mid B_n)$$

[Exercise: Prove the above

$A = A \cap \Omega$]

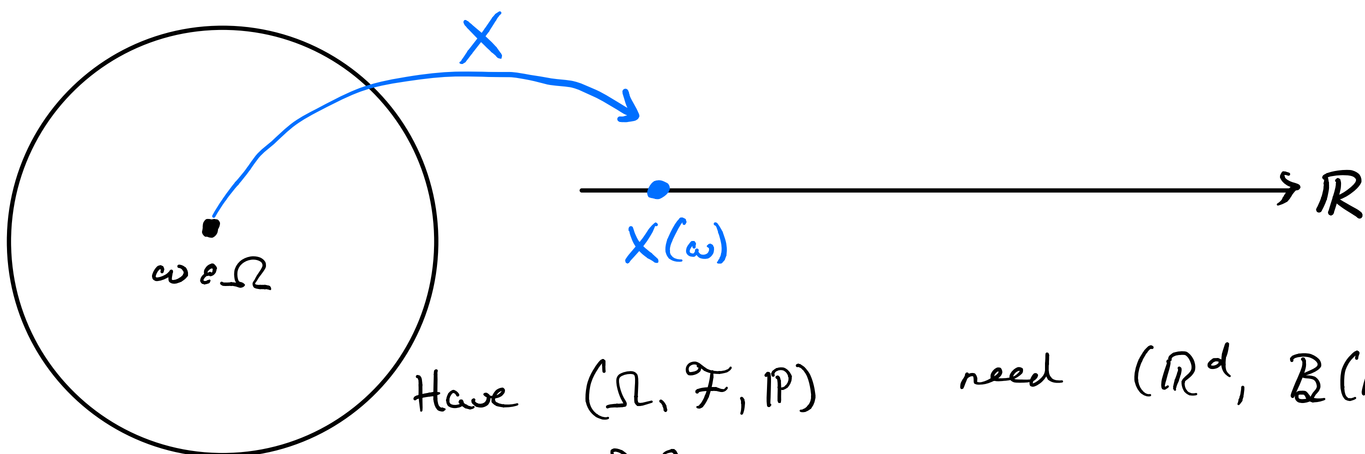
Random Variables (aka Measurable Functions)

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

$\Omega = \{$  $\}$

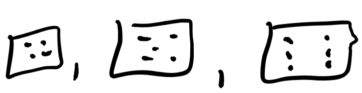
First and foremost, random variables are functions

$$X : \Omega \to \mathbb{R}^d$$
$$\omega \mapsto \mathbb{R} \times \cdots \times \mathbb{R}$$



Have $(\Omega, \mathcal{F}, \mathbb{P})$ need $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), ?)$
$\mathbb{P} : \mathcal{F} \to [0,1]$??

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

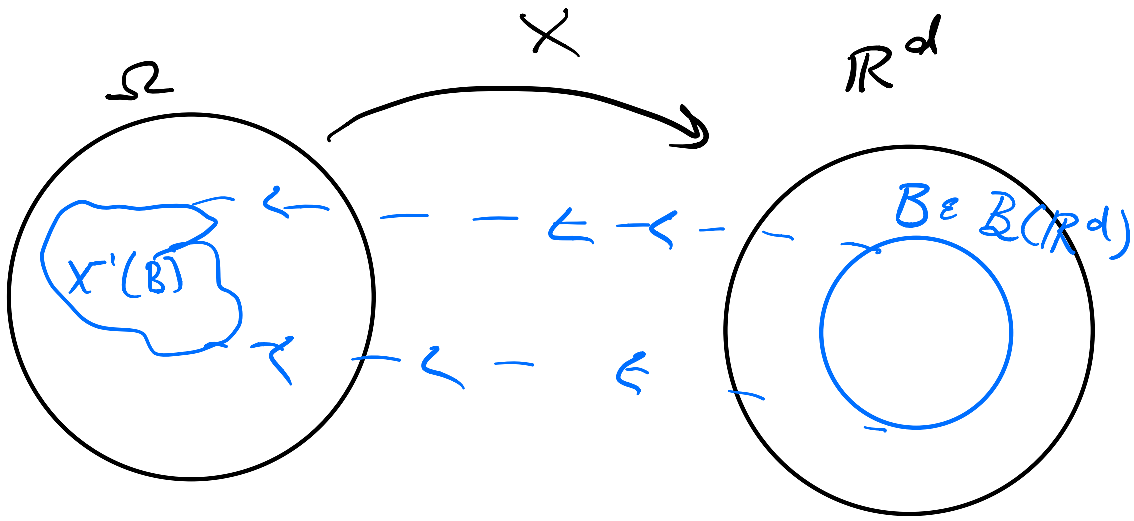
$\mathbb{P}(\text{"Gain"}) = \mathbb{P}(\{$  $\})$

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

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

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 $\mathbb{P}$.

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

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

Cumulative Distribution Functions

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

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

and is defines as

$$F_X(a_1, \dots, a_d) = \mathbb{P}_X\left((-\infty, a_1] \times \dots \times (-\infty, a_d]\right)$$

It can be shown that F_X uniquely defines the law of X, $\mathbb{P}_X$.

Discrete vs. Continuous Random Variables

Definition (support): the support $\mathbb{P}_X$ denoted by $\text{Supp}(\mathbb{P}_X)$ 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 $\text{supp}(\mathbb{P}_X)$ is countable.

Then we can define the pmf

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

associated with X.

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

Note: One can check

$$\mathbb{P}_{p_X} = \mathbb{P}_X$$

Aside Showing $\mathbb{P}_{p_X} = \mathbb{P}_X$ (Done in OH)

First lets figure out what exactly we are even showing.

Two situations

① Have $(\Omega, \mathcal{F})$. Given a pmf p,

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

can define

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

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

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

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

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

From here we motivated the definition of

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

and the pmf of X

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

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

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

equivalent

In particular, take an event

$$A = B \cap \operatorname{supp}(\mathbb{P}_X) \quad , \quad B \in \mathcal{B}(\mathbb{R}^d)$$

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

Well, then

$$\mathbb{P}_{p_X}(A) = \sum_{a \in A} p_X(A)$$

and

or n; countable union

$$\mathbb{P}_X(A) = \mathbb{P}_X\left(\bigcup_{i=1}^{\infty} \{a_i\}\right) = \sum_{i=1}^{\infty} \mathbb{P}_X(\{a_i\})$$

But $\mathbb{P}_X(\{a_i\})$ is precisely $p_X(a_i)$!

i.e.

$$\mathbb{P}_X(A) = \sum_{i=1}^{\infty} p_X(a_i) = \sum_{a \in A} p_X(a)$$

So

$$\mathbb{P}_{p_X}(A) = \mathbb{P}_X(A)$$

That is, the push forward measure of $\mathbb{P}$ through X is the same as the probability measure induced by the pmf of X.