$\boxed{\text{Statistics, Estimation theory, Inference, etc.}} \longrightarrow \boxed{\text{Probability theory}} \longrightarrow \boxed{\text{Set theory}}$

$\boxed{D1}$ Sample space $\longrightarrow$ Set, S, of all possible outcomes of a particular experiment.

Examples:

1. $S = \{H, T\}$ $\rightarrow$ tossing a coin
2. $S = \{C, D, H, S\}$ $\rightarrow$ selecting a card

* S $\rightarrow$ Countable / uncountable

$\boxed{D2}$ <u>Event</u> $\rightarrow$ A, an event is a subset of S

* Event is any collection of possible outcomes of an experiment

Example: $A_1 = \{C, D\}$ $A_2 = \{D, H, S\}$ (card experiment)

Set theory

$A \subset B \iff x \in A \implies x \in B$ (containment)

$A = B \iff A \subset B$ and $B \subset A$ (equality)

Union → $A \cup B = \{x : x \in A$ or $x \in B\}$

Intersection → $A \cap B = \{x : x \in A$ and $x \in B\}$

Complementation $A^c = \{x : x \notin A\}$

* Let A, B, C be events on Sample space S

1. Commutative $A \cup B = B \cup A$; $A \cap B = B \cap A$
2. Associative $A \cup (B \cup C) = (A \cup B) \cup C$; $A \cap (B \cap C) = (A \cap B) \cap C$
3. Distributive $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
   $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
4. DeMorgan's Law $(A \cup B)^c = A^c \cap B^c$
   $(A \cap B)^c = A^c \cup B^c$

Proof → homework

* What is probability? → frequency of outcome

→ belief in the chance of an event occuring.

$\downarrow$ Objetive definition
Satisfying some axioms

for $A \subset S$ we ~~A~~ want $P(A)$

and $0 \leq P(A) \leq 1$

what is the domain of P?

$\boxed{D3}$ Sigma Algebra $\rightarrow$ a collection of subsets of S denoted by $\mathcal{B}$ (Borel field) which satisfies

a. $\emptyset \in \mathcal{B}$

b. if $A \in \mathcal{B}$ then $A^c \in \mathcal{B}$

c. if $A_1, A_2 \cdots \in \mathcal{B}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$

Example $S = \{1, 2, 3\}$

$\mathcal{B}$ is the collection of $2^3 = 8$ sets

$\{1\}$ $\{1,2\}$ $\{1,2,3\}$ $\{2\}$ $\{1,3\}$ $\{3\}$ $\{2,3\}$ $\emptyset$

$\boxed{D4}$. **Probability function**

given S and an associated Sigma Algebra $\mathcal{B}$, 'probability function' is a function P with domain $\mathcal{B}$ that satisfies

1. $P(A) \geq 0$ for all $A \in \mathcal{B}$
2. $P(S) = 1$
3. if $A_1, A_2 \ldots \in \mathcal{B}$ are pairwise disjoint, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Axioms of Probability or Kolmogorov Axioms

Example: Fair coin toss

$$S = \{H, T\}$$

(1) $P(\{H\}) = P(\{T\})$ [not from axioms)

- Since $S = \{H\} \cup \{T\}$, using axiom 2

  we have $P(\{H\} \cup \{T\}) = 1$

- As $\{H\}$ and $\{T\}$ are disjoint

  $P(\{H\} \cup \{T\}) = P(\{H\}) + P(\{T\})$ and

  $P(\{H\}) + P(\{T\}) = 1$

How to define legitimate probability functions?

Let $S = \{s_1, \dots s_n\}$ be a finite set,

$\mathcal{B} \rightarrow$ any Sigma Algebra of S,

$p_1 \cdots p_n$ non-negative and sum to 1.

For any $A \in \mathcal{B}$ define $P(A)$ as

$$P(A) = \sum_{(i : s_i \in A)} p_i$$

$\boxed{\text{Proof} \quad - \quad \text{homework}}$

Calculus of Probabilities (Self Study)

* if P is a probability function and $A \subset B$

   a. $P(\emptyset) = 0$

   b. $P(A) \leq 1$

   c. $P(A^c) = 1 - P(A)$

* P – probability function A and B are in $\mathcal{B}$ then

   a. $P(B \cap A^c) = P(B) - P(A \cap B)$

   b. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

   c. if $A \subset B$ then $P(A) \leq P(B)$

Conditional probabilities and Independence

$\boxed{D5}$ If A, B are events in S and $P(B) > 0$ then the conditional probability of A given B, is

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

Prisoner example - understanding conditional probabilities

Prisoner A, B, C $\rightarrow$ one will be pardoned.

Warden tells A that B is being executed.

$P(A) = P(B) = P(C) = \frac{1}{3}$; let w be an event that warden says B dies

$$P(A|W) = \frac{P(A \cap W)}{P(W)}$$

$P(W) = P($warden says B dies and A is pardoned$) + P(\ldots$ B dies and C is pardoned$)$

$+ P(\ldots$ B dies and B is pardoned$)$

$= \frac{1}{6} + \frac{1}{3} + 0$

Bayes' Rule

① $P(A \cap B) = P(A|B)\,P(B)$

② $P(A \cap B) = P(B|A)\,P(A)$

$$\Downarrow$$

$$\boxed{P(A|B) = P(B|A)\,\frac{P(A)}{P(B)}}$$

General form

$$P(A_i|B) = \frac{P(B|A_i)\,P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)\,P(A_j)}$$

A_i's $\to$ partition of Space and B any set.

Bayes' rule example : Morse code.

$$P(\text{dot sent}) = \frac{3}{7} \quad \& \quad P(\text{dash sent}) = \frac{4}{7}$$

A_1 – dot sent $\qquad$ A_2 – dash sent

B_1 – dot received $\qquad$ B_2 – dash received

with 1/8 probability a dot is mistakenly received as dash

$$P(B_1 | A_1) = 7/8$$

Question – $P(\text{dot sent} \mid \text{dot received})$?

$\downarrow$ Bayes' rule.

$$P(A_1 | B_1) = P(B_1 | A_1) \frac{P(A_1)}{P(B_1)}$$

$$P(B_1) = P(B_1 \cap A_1) + P(B_1 \cap A_2) = \ldots$$ ~~Leave~~ Homework

Independence

$$P(A|B) = P(A) \qquad \boxed{B \text{ has no effect on } A}$$

$$P(B|A) = P(A|B)\frac{P(B)}{P(A)} = P(A)\frac{P(B)}{P(A)} = P(B)$$

as $P(B|A)\,P(A) = P(A \cap B)$

$$\boxed{P(A \cap B) = P(A)\,P(B)}$$

Statistical Independence

Independence

* $A_1 \ldots A_n$ – $\boxed{\text{mutually independent}}$ if for $\boxed{\text{any}}$ subcollection $A_{i_1} \cdots A_{i_k}$

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

↑
Strong definition.

Homework:

* Tossing a coin 3 times
  e.g. sample point HHT
  Show that the occurance of a head on any toss has no effect on any other tosses

Random Variable

'Summary variable' rather than the original probability structure

$$\text{Sample Space} \xrightarrow[\text{(function/mapping)}]{\text{Random variable}} \text{real numbers}$$

(new sample space)

Random Variable examples

Experiment	Random variable
① Toss 2 dice	X = Sum of numbers
② Toss a fair coin 3 times	X = number of heads obtained

S	HHH	HHT	HTH	THH	TTH	THT	HTT	TTT
$X(s)$	3	2	2	2	1	1	1	0

Random Variable & Probability functions.

$S = \{s_1 \cdots s_n\} \longrightarrow$ Probability function P

$X \longrightarrow \mathcal{X} = \{x_1 \cdots x_n\} \longleftarrow P_X$

Observe $X = x_i \iff$ Outcome is $s_j \in S$ such that

$X(s_j) = x_i$

Thus

$$P_X(X = x_i) = P\left(\{s_j \in S : X(s_j) = x_i\}\right)$$

Random Variable & Probability functions

Example: Toss a fair coin 3 times

$$P(X=2) = ?$$

$$= P(\{HHT, HTH, THH\})$$

$$= 3/8$$

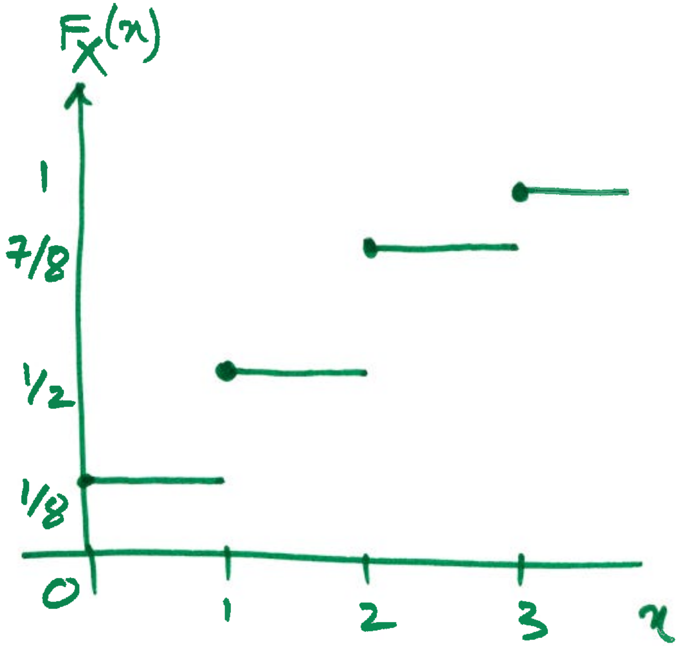
Distribution functions

$\boxed{\text{D6}}$ Cumulative distribution function (CDF) of X

$$F_X(x) = P_X(X \le x), \text{ for all } x.$$

Example (Same as before: 3 Coin toss)

$$F_X(x) = \begin{cases} 0 & \text{if } -\infty \le x < 0 \\ 1/8 & \text{if } 0 \le x < 1 \\ 1/2 & \text{if } 1 \le x < 2 \\ 7/8 & \text{if } 2 \le x < 3 \\ 1 & \text{if } 3 \le x < \infty \end{cases}$$



CDF

$F(x)$ is a CDF if and only if

a. $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to +\infty} F(x) = 1$

b. $F(x)$ is non decreasing function of x

c. $F(x)$ is right continuous

↑

Consequance of definition

Homework

Tossing for a head

$p \to$ probability of a head on any given toss

X = number of tosses done to get a head.

$P(X = x) = ?$ $\qquad F_X(x) = ?$

CDF

Continuous CDF example.

$$F_X(x) = \frac{1}{1+e^{-x}}$$

$\boxed{D7}$

$X \rightarrow$ continuous if $F_X(x)$ is a continuous function of x.

discrete $\longrightarrow$ step function.

Identically distributed

R.v. X, Y are I.D. if

for an every set $A \in \mathcal{B}'$

$$P(X \in A) = P(Y \in A)$$

Does not imply $X = Y$

eg: number of heads obs. vs. number of tails observed.

Probability mass and density functions (PMF & PDF)

$\boxed{D8}$ PMF

$$f_X(x) = P(X = x) \text{ for all } x$$

* $$P(a \leq X \leq b) = \sum_{k=a}^{k=b} f_X(k)$$

PMF & PDF

$\boxed{D9}$ PDF $f_X(x)$ of X is such that

$$F_X(x) = \int_{-\infty}^{x} f_x(t)\, dt \quad \text{for all } x.$$

$$\frac{d}{dx} F_X(x) = f_X(x)$$

$f_X(x) \rightarrow$ PDF or PMF if and only if

a. $f_X(x) \geq 0$ for all x

b. $\sum_x f_X(x) = 1$ or $\int_{-\infty}^{\infty} f_X(x) = 1$