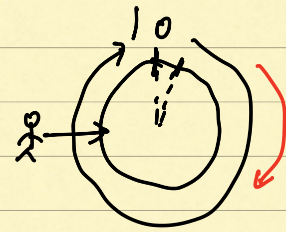
Lecture 1: Probability Review (Sections 1.1 – 1.4.1 of Gallager)

Ω: sample space, i.e., the set of all sample points of a random expt.

Eg: i) Coin Toss $\Omega = \{H, T\}$
ii) Dice Throw $\Omega = \{1, 2, 3, \cdots, 6\}$.
iii) Rolling a wheel $\Omega = [0, 1)$



i) & ii) sample space Ω is finite
iii) sample space Ω is uncountable

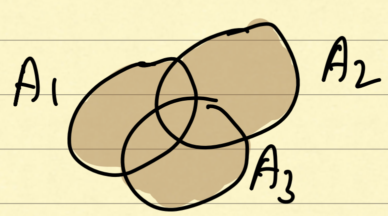
iv) # of buses you to wait before your bus comes
$\Omega = \{0, 1, 2, \cdots\}$ countably infinite.

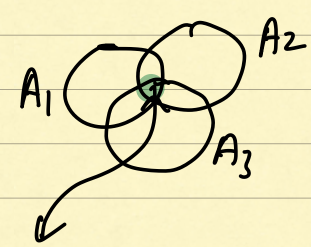
Events: (Legitimate) subsets of Ω

Eg. ii) $E = \{2, 4, 6\} \subseteq \Omega$ even outcomes
iii) $\mathbb{Q} \cap [0, 1) \subseteq \Omega$ rational #s in $[0, 1)$.

If we have events (subsets of Ω) $A_1, A_2, \cdots, A_n \subset \Omega$ their union is denoted as

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n.$$





Intersection $\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n$

$= \{x \in \Omega \mid x \text{ belongs to } A_i \text{ for all } 1 \le i \le n\}$

Complement of A in Ω is $A^c = \Omega \setminus A$.

<u>Countable set</u> is a set in which the elements of the set can be placed in one-to-one correspondence with $\mathbb{N} = \{1, 2, 3, \cdots\}$.

$E = \{0, 2, 4, \cdots\}$ is countable

$$\begin{array}{ccc} 0 & 2 & 4 \\ \updownarrow & \updownarrow & \updownarrow \\ 1 & 2 & 3 \end{array}$$

$\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, 3, \cdots\}$ integers.

$\mathbb{Q} = \{\frac{m}{n} : m, n \in \mathbb{Z}, n \neq 0\}$ countable.

$[0,1]$: not countable.

<u>Axioms for events.</u> Ω: sample space

Def: The class of all subsets of Ω that constitute the set of (legitimate) events is called a σ-algebra (σ-field).

Def: A <u>σ-algebra</u> $\mathcal{F}$ of Ω is a family $\{A_i\}$ of subsets of Ω (i.e., $A_i \subset \Omega$) s.t.

i) $\Omega \in \mathcal{F}$

ii) $\forall A_1, A_2, \cdots \in \mathcal{F}$, their countable union $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.
(i.e., σ-algebra is closed under countable union)

iii) $\forall A \in \mathcal{F}$, $A^c = \Omega \setminus A \in \mathcal{F}$.
(i.e., σ-algebra is closed under complementation)

The class of all subsets that satisfy (i) – (iii) is called a σ-algebra.

The sets in a σ-algebra are called events (legitimate subsets).

Eg: $\Omega = \{H, T\}$

Claim: $\mathcal{F} = \{\phi, \Omega\}$ is a σ-algebra (trivial σ-algebra)

Claim: $\mathcal{F} = \{\phi, \{H\}, \{T\}, \underbrace{\{H, T\}}_{\Omega}\}$ is a σ-algebra. (Ex).

Claim: $\mathcal{F}' = \{\phi, \{H\}, \underbrace{\{H, T\}}_{\Omega}\}$ is not a σ-alg. (Ex)

Claim: $\Omega = [0,1]$. Consider the family of all subset of Ω, $\mathcal{F} = 2^{\Omega}$. This is a σ-algebra, but will not work for probability purposes.

Fact: $\phi \in \mathcal{F}$ (i.e., ϕ is a legit event)

Pf: $\phi = \Omega^c$ & Ω is an event ($\Omega \in \mathcal{F}$). By closedness under complementation, ϕ is also an event.

Fact: If $A_1, A_2, \ldots, A_n$ is a finite collection of events (i.e., $A_i \in \mathcal{F}$ for all $1 \leq i \leq n$), then $\bigcup_{i=1}^{n} A_i$ is an event.

Pf: Axiom (ii) says that if $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ so is $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Now from the previous fact, we proved that ϕ is an event.

Now set $A_{n+1} = A_{n+2} = \cdots = \phi$. The sequence $A_1, A_2, \ldots$ is a sequence of events.

Applying (ii) we get $\mathcal{F} \ni \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} A_i$

Fact: Every finite or countably infinite intersection of events is still an event.

Pf: De Morgan's laws

$$\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c$$

$B_i = A_i^c$ $\iff$ $$\overbrace{\left(\underbrace{\bigcup_i \underbrace{B_i^c}_{\text{event (iii)}}}_{\text{event (ii)}}\right)^c}^{\text{event}} = \underbrace{\bigcap_i B_i}_{\text{event}}.$$

$B_i \in \mathcal{F}$ are events $\Rightarrow$ B_i^c is also an event

$\Rightarrow$ $\bigcup_i B_i^c$ is also an event (σ-alg. closed under countable $\cup$)

$\Rightarrow$ $\left(\bigcup_i B_i^c\right)^c$ is also an event (σ-alg. closed under comp.)

<u>Axioms of Probability</u>

Ω: sample space
$\mathcal{F}$: σ-algebra defined on Ω ($\mathcal{F}$ satisfies 3 axioms)

Def: Probability rule $P_r : \mathcal{F} \to [0,1]$.

i) $P_r(\Omega) = 1$

ii) $\forall A \in \mathcal{F}$ $P_r(A) \geq 0$

iii) (Countable Additivity) $\forall A_1, A_2, \ldots \in \mathcal{F}$ such that A_i's are disjoint ($A_i \cap A_j = \phi$ $\forall i \neq j$), then

$$P_r\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P_r(A_i) = \lim_{m\to\infty} \sum_{i=1}^{m} P_r(A_i)$$

Fact i) $P_r(\phi) = 0$

Pf: Let $A_i = \phi \ \forall i \in \mathbb{N}$. The A_i's are disjoint (i.e., $A_i \cap A_j = \phi \cap \phi = \phi \ \forall i \neq j$).

$$Pr(\phi) = Pr\left(\bigcup_{i=1}^{\infty} A_i\right) \overset{(iii)}{=} \lim_{m\to\infty} \sum_{i=1}^{m} Pr(A_i)$$

$$= \lim_{m\to\infty} \sum_{i=1}^{m} Pr(\phi)$$

$$= \lim_{m\to\infty} m\,Pr(\phi)$$

Two cases: i) $\underbrace{Pr(\phi) > 0}_{\delta > 0 \text{ (not possible)}}$ ii) $\underline{Pr(\phi) = 0.}$ ✓

Since $Pr(\phi)$ is non-negative, $Pr(\phi) = 0$.

Fact ii) If $A_1, \dots, A_n \in \mathcal{F}$ are disjoint $Pr\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} Pr(A_i)$.

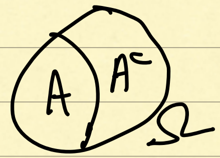
Pf: Apply axiom (iii) of prob. rule to the sequence of events $A_1, A_2, \dots, A_n$ & $A_{n+1} = A_{n+2} = \cdots = \phi$

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} A_i$$

$$Pr\left(\bigcup_{i=1}^{n} A_i\right) = Pr\left(\bigcup_{i=1}^{\infty} A_i\right) \overset{(iii)}{=} \sum_{i=1}^{\infty} Pr(A_i) = \sum_{i=1}^{n} Pr(A_i) + \sum_{i=n+1}^{\infty} Pr(A_i)$$

$$= \sum_{i=1}^{n} Pr(A_i) + \sum_{i=n+1}^{\infty} 0 = \sum_{i=1}^{n} Pr(A_i) \qquad ////.$$

Fact (iii) : $Pr(A^c) = 1 - Pr(A) \ \forall A \in \mathcal{F}$.

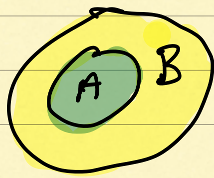


$A \in \mathcal{F}$, $A^c \in \mathcal{F}$

$A \sqcup A^c = \Omega$ (A and A^c are disjoint)

$$1 = \Pr(\Omega) \underset{\substack{\uparrow \\ \text{disjoint}}}{=} \Pr(A) + \Pr(A^c) \quad ////$$

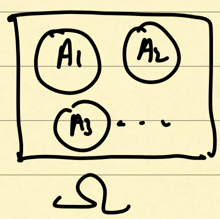
Fact (iv): $\forall A, B \in \mathcal{F}$ $A \subseteq B$, $\Pr(A) \leq \Pr(B)$.



$B = A \sqcup (B \setminus A)$ (disjoint union)

$$\Pr(B) \overset{\text{disjoint}}{=} \Pr(A) + \Pr(B \setminus A)$$
$$\geq \Pr(A) \quad ////$$

Fact (v): Let $(A_i)_{i=1}^{\infty} \subset \mathcal{F}$ be a sequence of disjoint events



$$\sum_{i=1}^{\infty} \Pr(A_i) \leq 1$$

$$1 = \Pr(\Omega) \underset{\substack{\uparrow \\ \text{previous fact.}}}{\geq} \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) \overset{\text{disjoint}}{\underset{\substack{\uparrow \\ \text{countable additivity.}}}{=}} \sum_{i=1}^{\infty} \Pr(A_i)$$

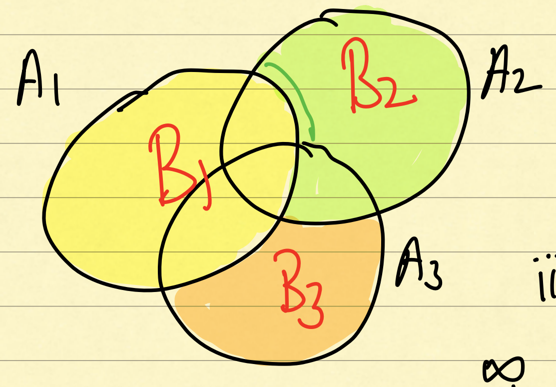
Fact (vi) [Continuity of Measure]

For any sequence of not necessarily disjoint events $(A_i)_{i=1}^{\infty}$ (i.e., $A_i \in \mathcal{F}$),

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{m \to \infty} \Pr\left(\bigcup_{i=1}^{m} A_i\right)$$

Idea: Create disjoint events $(B_i)_{i=1}^{\infty}$

$B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus (A_1 \cup A_2)$, ..., $B_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j$



iii) $\bigcup_{i=1}^{m} A_i = \bigsqcup_{i=1}^{m} B_i$

i) The B_i's are disjoint. ii) $\bigcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{\infty} B_i$

Given (ii) $\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \Pr\left(\bigsqcup_{i=1}^{\infty} B_i\right)$

$$= \sum_{i=1}^{\infty} \Pr(B_i) = \lim_{m \to \infty} \left(\sum_{i=1}^{m} \Pr(B_i) \right) \quad -(1)$$

Note that

$$\sum_{i=1}^{m} \Pr(B_i) = \Pr\left(\bigsqcup_{i=1}^{m} B_i\right) = \Pr\left(\bigcup_{i=1}^{m} A_i\right) \quad -(2).$$

Combining eqns (1) & (2)

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{m \to \infty} \Pr\left(\bigcup_{i=1}^{m} A_i\right) \qquad ////.$$

<u>Fact</u> $(A_n)_{n=1}^{\infty} \subset \mathcal{F}$ not necessarily disjoint.

$$\Pr\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \Pr(A_n)$$ (union-of-events bound)

Pf: $\Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{m\to\infty} \Pr\left(\bigcup_{n=1}^{m} A_n\right)$ — (*)

Note $\Pr(A_1 \cup A_2) = \Pr(A_1) + \Pr(A_2) - \Pr(A_1 \cap A_2)$
$\leq \Pr(A_1) + \Pr(A_2)$

By induction, $\Pr\left(\bigcup_{n=1}^{m} A_n\right) \leq \sum_{n=1}^{m} \Pr(A_n)$.

(*) — $\Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{m\to\infty} \Pr\left(\bigcup_{n=1}^{m} A_n\right)$

$\leq \lim_{m\to\infty} \sum_{n=1}^{m} \Pr(A_n) = \sum_{n=1}^{\infty} \Pr(A_n)$

///.

Rmk: $\Pr\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \min\left\{ \sum_{n=1}^{\infty} \Pr(A_n), 1 \right\}$

Rmk: $\Pr\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \left(\sum_{n=1}^{\infty} \Pr(A_n)\right)^{\rho}$, $0 \leq \rho \leq 1$.

Probability Review. $(\Omega, \mathcal{F})$

Def: $\forall\, A, B \in \mathcal{F}$ under a probability rule $\Pr$ defined on $(\Omega, \mathcal{F})$ the conditional probability of A given B is

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} \quad \text{if} \quad \Pr(B) > 0$$

Bayes law: $\Pr(A|B)\Pr(B) = \Pr(B|A)\Pr(A)$

Def: Two events $A, B \in \mathcal{F}$ are independent under $\Pr$ if

$$\Pr(A \cap B) = \Pr(A)\Pr(B)$$

If $\Pr(B) > 0$, this is equivalent to $\Pr(A|B) = \Pr(A)$.

Def: $A, B, C \in \mathcal{F}$.
A and B are conditionally independent given C if

$$\Pr(A \cap B | C) = \Pr(A|C)\Pr(B|C)$$

Def: The events $A_1, A_2, \dots, A_n$ $(n \geq 2)$ are mutually independent if for all $S \subseteq \{1, \dots, n\}$, $|S| \geq 2$

$$\Pr\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} \Pr(A_i) \quad \text{—} (*)$$

Rmk: $\overset{(*)}{\text{This}}$ includes the entire collection $S = \{1, \dots, n\}$

i.e., $\Pr\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \Pr(A_i)$ — (+)

However (+) does not imply mutual indep among the A_i's according to the true definition.

Ex:

	Sample Points	A_1	A_2	A_3
1/8	1	1	1	1
1/8	2	1	1	0
1/8	3	1	0	1
.	4	1	0	0
.	5	0	1	0
.	6	0	1	0
.	7	0	0	1

1/8 8 0 0 1

$A_2 = \{1,2,5,6\}$ $A_2 \cap A_3 = \{1\}$
$A_3 = \{1,3,7,8\}$

Throw a 8-face dice
$A_1 = \{1,2,3,4\}$, ...

$Pr(A_i) = 4/8 = 1/2.$
$i = 1,2,3$

$Pr(A_1 \cap A_2 \cap A_3) = 1/8$
$= Pr(A_1)Pr(A_2)Pr(A_3) = 1/8$

However, $Pr(A_2 \cap A_3) = 1/8 \neq Pr(A_2)Pr(A_3) = 1/4.$
$\Rightarrow A_2$ & A_3 are dependent.

$Pr(A_1 \cap A_2) = Pr(A_1)Pr(A_2)$ ✓
$Pr(A_1 \cap A_2) = Pr(A_1)Pr(A_2)$ ✓

Pairwise independence doesn't imply mutual indep

2 tosses of a fair coin.

$H_1 = \{1st \text{ toss} = H\}$, $H_2 = \{2nd \text{ toss} = H\}$
$D = \{2 \text{ tosses are different}\}$.

$Pr(H_1) = Pr(H_2) = Pr(D) = \frac{1}{2}.$

$H_1 \perp\!\!\!\perp H_2$

$$Pr(D|H_1) = \frac{Pr(H_1 \cap D)}{Pr(H_1)} = \frac{1/4}{1/2} = \frac{1}{2} = Pr(D)$$

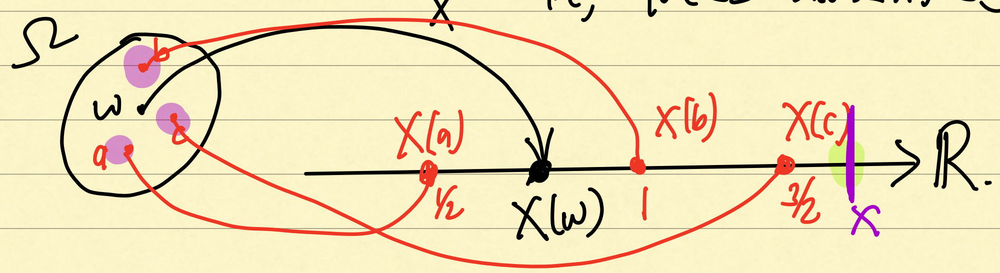
$D \perp\!\!\!\perp H_1$, $D \perp\!\!\!\perp H_2.$

$Pr(H_1 \cap H_2 \cap D) = 0 \neq Pr(H_1)Pr(H_2)Pr(D) = \frac{1}{8}.$

Random Variables.

Def: A rv X is a function that maps from the sample space Ω to the real line (i.e., $X: \Omega \to \mathbb{R}$) st.

$\{w \in \Omega : X(w) \le x\}$ is an event for all $x \in \mathbb{R}$.
i.e., $\{w \in \Omega : X(w) \le x\} \in \mathcal{F}$.



Eg: $\Omega = \{a, b, c\}$ $\quad X(a) = \frac{1}{2}, \; X(b) = 1, \; X(c) = \frac{3}{2}$.

$\mathcal{F} = \{\phi, \{a\}, \{b\}, \{c\}, \{a,b\}, \{b,c\}, \{a,c\}, \Omega\}$

Let $x = 1.1 \in \mathbb{R}$ Look at

$\{w \in \Omega : X(w) \le 1.1\} = \{a, b\} \in \mathcal{F}$

Let $x = 0.4 \in \mathbb{R}$.

$\{w \in \Omega : X(w) \le 0.4\} = \phi \in \mathcal{F}$

Rmk: $\forall x \in \mathbb{R}. \quad \{w \in \Omega : X(w) \le x\} \in \mathcal{F}$

Question: $\forall x \in \mathbb{R} \quad \{w \in \Omega : X(w) < x\} \in \mathcal{F}$?? True ✓

Question: $\forall x \in \mathbb{R} \quad \{w \in \Omega : X(w) = x\} \in \mathcal{F}$?? True ✓

Def: The cumulative distribution function (cdf) of X is

$$F_X(x) = \Pr(\{w \in \Omega : X(w) \le x\}) =: \Pr(X \le x)$$

Upper case

Lower case

Fact: i) $x \in \mathbb{R} \mapsto F_X(x)$ is non-decreasing.
$\forall x, y \in \mathbb{R}$ s.t. $x \le y$, $F_X(x) \le F_X(y)$.

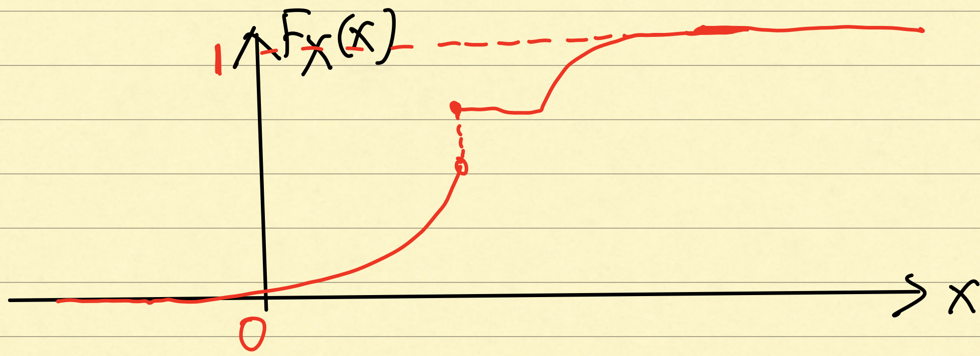
$$\{\omega \in \Omega : X(\omega) \le x\} \subseteq \{\omega \in \Omega : X(\omega) \le y\}$$

$$A \subseteq B \Rightarrow \Pr(A) \le \Pr(B)$$

ii) $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to \infty} F_X(x) = 1$

iii) $x \mapsto F_X(x)$ is right-continuous.

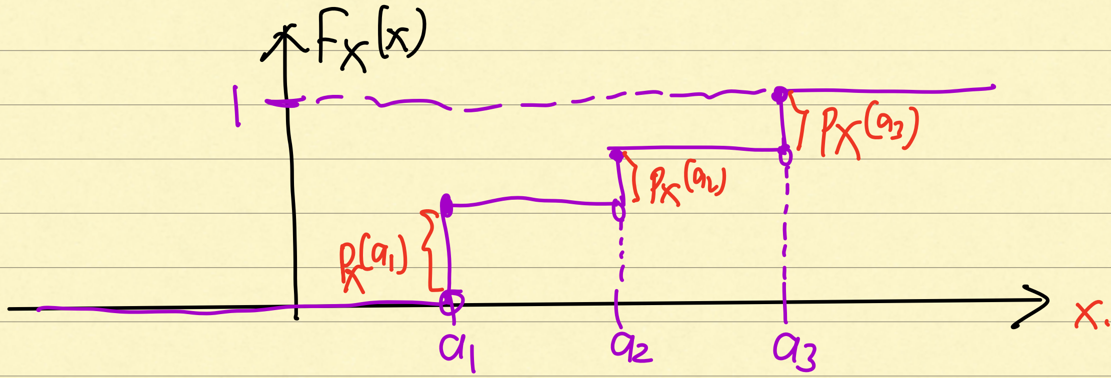
$$\lim_{\varepsilon \to 0^+} F_X(x+\varepsilon) = F_X(x) \quad \forall x \in \mathbb{R}, \quad (\text{Ex})$$



Special case: If X has a finite or countably infinite # of possible values, say $a_1, a_2, a_3, \cdots$, the prob. $\Pr(X = a_i)$ of each sample a_i is called the probability mass function of X.

$$P_X(a_i) = \Pr(X = a_i).$$
$$= \Pr(\{\omega \in \Omega : X(\omega) = a_i\})$$

This need not belong to $\mathcal{F}$ at first sight but this is indeed an event if X is r.v.

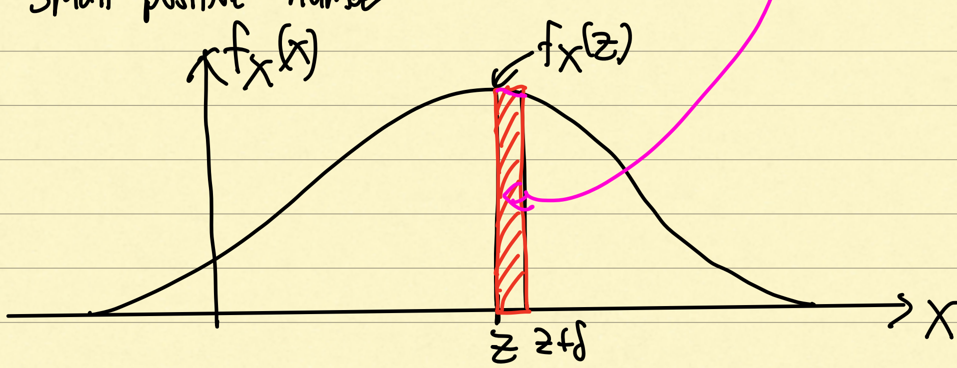


If F_X has a derivative at x, then

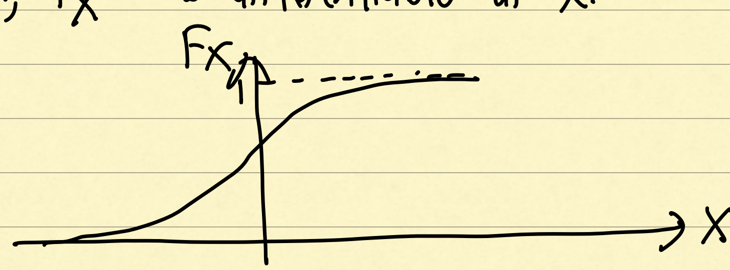
$$f_X(x) = F_X'(x) = \frac{d}{dx}F_X(x)$$

is called the probability density function of X at x.

$$\underbrace{f_X(z)\,\delta}_{\text{small positive number}} \approx \Pr(z < X \le z+\delta) = \int_z^{z+\delta} f_X(u)\,du$$



Def: A rv X is <u>continuous</u> if $\forall x \in \mathbb{R}$, $f_X(x)$ exists, i.e., F_X is differentiable at x.



Multiple rvs $X_1, \dots, X_n$

Joint cdf $F_{X_1,\dots,X_n}(x_1,\dots,x_n) = \Pr(X_1 \le x_1, \dots, X_n \le x_n)$.

Given $F_{X_1,\dots,X_n}$, how do get the cdf of a single rv.

$1 \le i \le n$ $\quad F_{X_i}(x_i) = F_{X_1,\dots,X_n}(\infty,\dots,\infty, x_i, \infty, \dots, \infty)$

Joint pmf: $p_{X_1,\dots,X_n}(x_1,\dots,x_n) = \Pr(X_1 = x_1, \dots, X_n = x_n)$.

Independence of rvs

Two rvs are indep. if

$$F_{XY}(x,y) = F_X(x) F_Y(y) \quad \text{for all } x, y \in \mathbb{R}.$$

If rvs X & Y are discrete,

$$p_{XY}(x_i, y_j) = p_X(x_i) p_Y(y_j) \quad \forall x_i, \forall y_j$$

Example of a stochastic press (Bernoulli process)

Def: A stochastic process (random process) is an infinite collection of rvs $\{X_i\}_{i=1}^{\infty}$ defined on a common sample space Ω.

The rvs are usually indexed by an Integer $n \in \mathbb{N}$ or a real-valued parameter $t \in \mathbb{R}$

X_n: discrete-time $\quad X_t$: continuous-time.

Def: A Bernoulli process is a sequence of i.i.d. binary (Bernoulli) rvs $Z_1, Z_2, Z_3, \ldots$

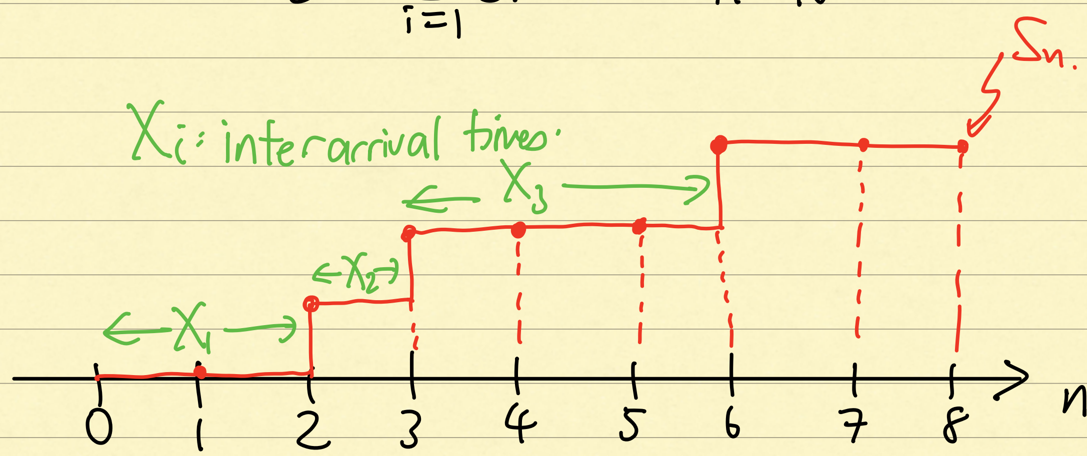
$$p := \Pr(Z_1 = 1) \qquad q = 1-p = \Pr(Z_1 = 0).$$

$\{Z_i = 1\} \Leftrightarrow$ a customer walked into our shop at time $i \in \mathbb{N}$

$\{Z_i = 0\} \Leftrightarrow$ no customer walked into our shop at time $i \in \mathbb{N}$

Based on $\{Z_i\}_{i=1}^{\infty}$, define another sequence of rvs

$$S_n := \sum_{i=1}^{n} Z_i \qquad n \in \mathbb{N}.$$



Z_i	0	1	1	0	0	1	0	0
S_n	0	1	2	2	2	3	3	3

Consider X_1: the first interarrival time. What's its pmf?

$X_1 = 1$ iff. $Z_1 = 1$ $\quad p_{X_1}(1) = p$

$X_1 = 2$ iff $Z_1 = 0, Z_2 = 1$ $\quad p_{X_1}(2) = qp = (1-p)p$

$X_1 = 3$ iff $Z_1 = Z_2 = 0, Z_3 = 1$

$p_{X_1}(3) = (1-p)^2 p$

Geometric r.v. $P_{X_1}(j) = (1-p)^{j-1} p, \quad j \in \mathbb{N} = \{1, 2, 3, \dots\}$

Claim: P_{X_k} is the same P_{X_1} for all $k \geq 1$.

Claim: All the interarrival times $\{X_k\}_{k=1}^{\infty}$ are mutually independent.

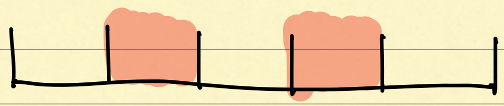
What's the distribution of $S_n = \sum_{j=1}^{n} Z_j$

Each S_n is the # of arrivals up to & including n.

$Pr(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$

$0 \leq k \leq n$

$k = 2$



$n = 5$