EE5137: Stochastic Processes.

Lecturer: A/Prof. Vincent Tan (vtan@nus.edu.sg)

Grader: Haiyun He (haiyun.he@u.nus.edu)

Grading: i) Homeworks (10%)

ii) Quiz 1 (12%)

iii) Quiz 2 (18%)

v) Exam (60%).

Main Textbook: "Stochastic Processes" by Gallager.

Reading Section 1.1 - 1.4.1 of Gallager's book.

Ω: sample space, i.e., set of all sample points of an expt.

Eg: Coin toss $\Omega = \{H, T\}$.

Dice Throw $\Omega = \{1, 2, 3, 4, 5, 6\}$.

The sample space is finite.

Infinite sample space $\Omega = \mathbb{N}$, $\Omega = [0,1]$.

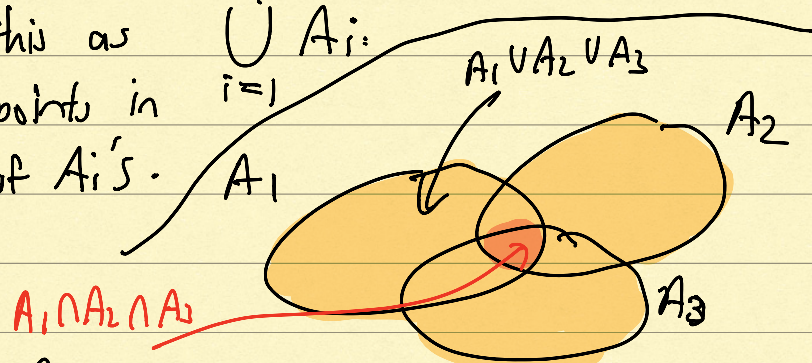
Events : (Legitimate) subsets of Ω

Eg: Dice Throw $E = \{2, 4, 6\} \subseteq \Omega$ even outcomes

$O = \Omega \setminus E$ odd outcomes

If we have events (subsets of Ω) $A_1, A_2, \dots, A_n \subseteq \Omega$ their union is denoted as $A_1 \cup A_2 \cup \dots \cup A_n$; we also write this as $\bigcup_{i=1}^{n} A_i$.

Consists of all points in at least one of A_i's.



Intersection $\bigcap_{i=1}^{n} A_i$ or $A_1 \cap A_2 \cap \dots \cap A_n$ is the set of points in Ω that are contained in ALL the A_i's.

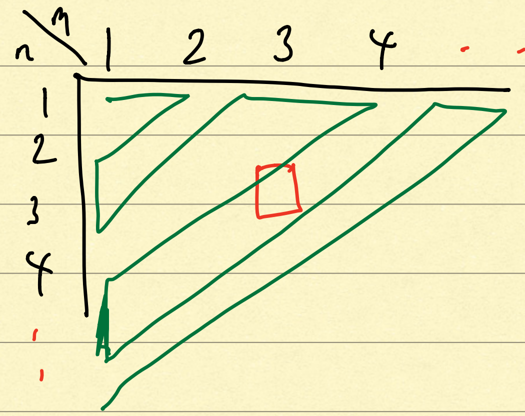
Complement of A (in Ω) is $A^c = \Omega \setminus A$.

Countable set is one in which the objects/elements of the set can be placed in 1-to-1 correspondence with $\mathbb{N} = \{1, 2, 3, 4, 5, \dots\}$.

Eg: Set of non-negative even numbers is countable $E = \{0, 2, 4, \dots\}$

Set of all integers $\mathbb{Z} = \{\cdots, -2, -1, \ , 1, 2, \ldots\}$

Set of all non-negative rational numbers $\mathbb{Q} = \{\frac{m}{n} : m, n \in \mathbb{Z}, \frac{m}{n} \geq 0\}$



Axioms for events : Ω sample space.

Def: The class of all subsets of Ω that constitute the set of (legitimite) events is called a <u>σ-algebra</u>, and a σ-algebra $\{A_i\}$ (where $A_i \subset \Omega$) satisfies

(i) Ω is an event;

(ii) $\forall A_1, A_2, \cdots$ events, their union $\bigcup_{i=1}^{\infty} A_i$ is also an event (closed under infinite <u>union</u>)

(iii) $\forall A$ events, $A^c = \Omega \setminus A$ is an event;
(closed under complementation).

The class of all subsets that satisfy (i)—(iii) is called a <u>σ-algebra</u>.

Fact : $\underset{}{\phi}^{=\Omega^c}$ (the empty set) is an event.

Pf: $\phi = \Omega^c$. Ω is an event by (i). Complements of events are events by (iii). Hence ϕ is an event.

Fact:
If $A_1, A_2, \dots, A_n$ is a finite collection of events, then $\bigcup_{i=1}^{n} A_i$ is an event.

Pf: Use Axiom (ii) & let $A_{n+1} = A_{n+2} = \cdots = \phi$.
The sequence $A_1, A_2, A_3, \dots, A_n, A_{n+1}, A_{n+2}, \dots$ is a sequence of events. By (ii) $\bigcup_{i=1}^{\infty} A_i$ is an event.

But $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} A_i$, hence an event.

Fact: Every finite or countable intersection of events is also an event.

Pf: $\Big(\underbrace{\bigcup_i A_i}_{\text{event}}\Big)^c = \bigcap_i A_i^c$ (de Morgan's rule)

$$\Longleftrightarrow \overbrace{\Big(\underbrace{\bigcup_i \underbrace{B_i^c}_{\text{(iii)}}}_{\text{event}}\Big)^c} = \Big(\bigcap_i B_i\Big) \leftarrow \text{event.} \qquad (B_i = A_i^c)$$

$(B_i)_{i=1}^{\infty}$

If we know that B_i's are events, then $\bigcap_i B_i$ an event?

Rmk: The class of all subsets of Ω for Ω uncountable (e.g. $\Omega = [0,1]$) does not allow for probability axioms

to be satisfied in a "sensible" way.
But this is not needed for our module (Real Analysis, Measure & Integration / Vitali set Wikipedia).

__Axioms of probability__ Ω: sample space. ; $\mathcal{E}$ a σ-alg defined on Ω ($\mathcal{E}$ satisfies the 3 axioms of a σ-algebra.)

Def:

A __probability rule__ $P_r : \mathcal{E} \to [0,1]$

i) $P_r(\Omega) = 1$;
ii) $\forall A \in \mathcal{E}$, $P_r(A) \geq 0$.
iii) $\forall A_1, A_2, \ldots \in \mathcal{E}$ disjoint ($A_i \cap A_j = \phi$ if $i \neq j$), then

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

__Fact i)__ $P_r(\phi) = 0$.

Pf: Let $A_n = \phi$ $\forall n \in \mathbb{N}$. The A_n's are disjoint. ($A_n \cap A_m = \phi$ $\forall n \neq m$). $\boxed{\bigcup_{n=1}^{\infty} A_n = \phi.}$

$$P_r(\phi) = P_r\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{m\to\infty} \sum_{n=1}^{m} P_r(A_n) = \lim_{m\to\infty} \sum_{n=1}^{m} P_r(\phi)$$
$$= \lim_{m\to\infty} m\, P_r(\phi)$$

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

a) $\cancel{Pr(\phi) > 0}$ b) $Pr(\phi) = 0$

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

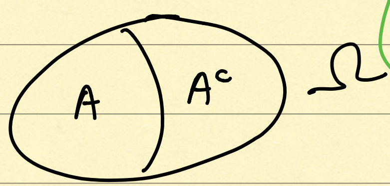
Fact (ii) If $A_1, \dots, A_m$ disjoint, $Pr\left(\bigcup_{n=1}^{m} A_n\right) = \sum_{n=1}^{m} Pr(A_n)$

Pf: Apply axiom (iii) to the sequence $A_1, A_2, \dots, A_m, A_{m+1}, A_{m+2} = \dots = \phi$. $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{m} A_n$.

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

Pf: $\Omega = A \sqcup A^c$ (A & A^c are disjoint)

disjoint union

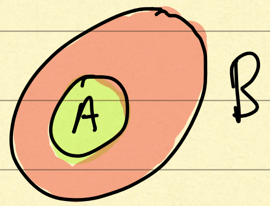


$$Pr(A \sqcup A^c) \overset{(ii)}{=} Pr(A) + Pr(A^c)$$
$$\|$$
$$Pr(\Omega) \overset{\text{(axiom i)}}{=} 1$$

$$\underline{Pr(A) = 1 - Pr(A^c)}$$

Fact (iv) $Pr(A) \le Pr(B)$ if $A, B \in \mathcal{E}$, $A \subseteq B$.

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



$$\Pr(B) = \Pr(A) + \underbrace{\Pr(B \setminus A)}_{\geq 0.}$$

$$\geq \Pr(A) \quad ////$$

Fact (v): $\Pr(A) \leq 1$ for all $A \in \mathcal{E}$

Pf: From Fact (iv), set $B = \Omega \in \mathcal{E}$. $A \subset B = \Omega$.

$\Pr(B) = \Pr(\Omega) = 1$.

$$\Pr(A) \leq \Pr(B) = \Pr(\Omega) = 1.$$

Fact (vi) $\sum_n \Pr(A_n) \leq 1$ $\forall$ sequence of disjoint events $(A_i)_{i=1}^{\infty}$.

Pf: Let A in Fact (v) be $A = \bigcup_{n=1}^{\infty} A_n$. Then use Fact (ii).

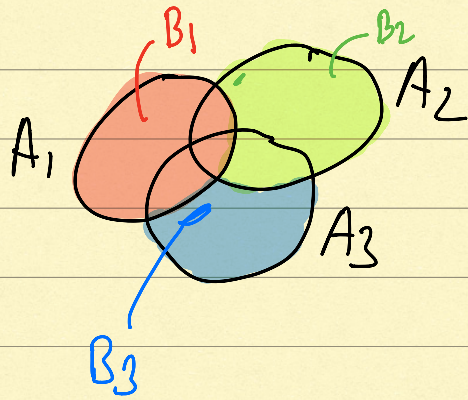
$$1 \overset{(Fv)}{\geq} \Pr\left(\bigsqcup_n A_n\right) \overset{(Fii)}{=} \sum_n \Pr(A_n)$$

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

[Continuity of measure]

for any sequence of not necessarily disjoint events $(A_i)_{i=1}^{\infty}$ $A_i \in \mathcal{E}$.

Idea:



Pf: Create $(B_i)_{i=1}^{\infty}$ that are disjoint &

$$\bigcup_{n=1}^{\infty} A_n = \bigsqcup_{n=1}^{\infty} B_n, \quad \bigcup_{n=1}^{m} A_n = \bigsqcup_{n=1}^{m} B_n \quad \boxed{B_n \subseteq A_n}$$

$$B_1 = A_1, \quad B_2 = A_2 \setminus A_1, \ldots, B_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j$$

Hence $Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = Pr\left(\bigsqcup_{n=1}^{\infty} B_n\right)$

$$= \sum_{n=1}^{\infty} Pr(B_n) = \lim_{m \to \infty} \sum_{n=1}^{m} Pr(B_n) \quad \text{— (1)}$$

$$\sum_{n=1}^{m} Pr(B_n) = Pr\left(\bigsqcup_{n=1}^{m} B_n\right) = Pr\left(\bigcup_{n=1}^{m} A_n\right) \text{ — (2)}$$

$$(1) \cup (2) \Rightarrow \quad Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{m \to \infty} Pr\left(\bigcup_{n=1}^{m} A_n\right)$$

Fact (viii) $(A_n)_{n=1}^{\infty} \subset \mathcal{E}$ not necessarily disjoint

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

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

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

$$= \sum_{n=1}^{\infty} \Pr(A_n)$$

Union-of-events bound (Union bound)

$(*)$ $\Pr(A_1 \cup A_2) \leq \Pr(A_1) + \Pr(A_2)$

$$\Pr(A_1 \cup A_2) = \Pr(A_1 \sqcup (A_2 \setminus A_1))$$

$$= \Pr(A_1) + \Pr(A_2 \setminus A_1) \leq \Pr(A_1) + \Pr(A_2)$$

$$\underset{A_2}{\cap}$$

Probability Review

Def: For any 2 events $A, B \in \mathcal{E}$ under a prob. model $\Pr(\cdot)$ the conditional prob. of A conditioned on B is defined if $\Pr(B) > 0$ by

$A \cap B \in \mathcal{E}$

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

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

Def: Two events A, B are <u>independent</u> if

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

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

Rmk: A & B are <u>conditionally independent</u> of each other given C if

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

Def: The events $A_1, A_2, \cdots, A_n$ $(n \geq 2)$ are <u>mutually independent</u> if $\forall S \subseteq \{1, \cdots, n\}$ $|S| \geq 2$

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

Rmk: This includes the entire collection, i.e., $S = \{1, \ldots, n\}$ but the statement

$$\Pr\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \Pr(A_i)$$

does not imply mutual indep. among A_i's.

	Sample points Ω	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

$\Pr(A_1 \cap A_2 \cap A_3) = 1/8 \qquad \Pr(A_i) = 1/2, \quad i \in \{1,2,3\}$

$$\Pr(A_1 \cap A_2 \cap A_3) = \prod_{i=1}^{3} \Pr(A_i) = 1/8$$

However $\Pr(A_2 \cap A_3) = 1/8 \neq \Pr(A_2)\Pr(A_3) = 1/4$

A_2 & A_3 are dependent.

$$\Pr(A_1^c \cap A_2^c \cap A_3^c) = 0 \neq (1-\Pr(A_1))(1-\Pr(A_2))(1-\Pr(A_3)) = 1/8$$

You can check

$\Pr(A_1 \cap A_2) = \Pr(A_1)\Pr(A_2) = \frac{1}{4}$

$\Pr(A_1 \cap A_3) = \Pr(A_1)\Pr(A_3) = 1/4.$

$\Downarrow$

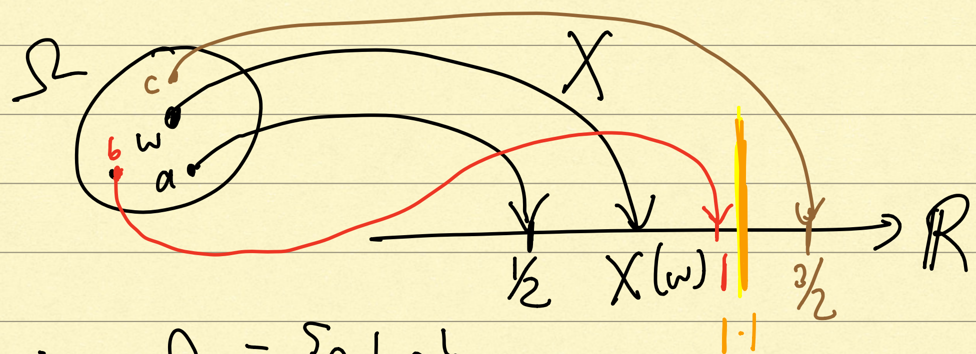
$A_1 \perp\!\!\!\perp A_2$

$A_1 \perp\!\!\!\perp A_3.$

Random Variable. (rv)

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

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{E} \text{ is an event for all } x \in \mathbb{R}.$$



Eg: $\Omega = \{a, b, c\}$

$\underline{X(a) = \frac{1}{2}}$, $\underline{X(b) = 1}$, $\underline{X(c) = 3/2}$.

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

$$\{\omega \in \Omega : X(\omega) \leq x = 1.1\} = \{a, b\}$$

If $\{a, b\} \in \mathcal{E}$

Let $x = 0.4 \in \mathbb{R}$. $\{\omega \in \Omega : X(\omega) \leq x = 0.4\} = \phi \in \mathcal{E}$

$\Omega = \{a, b, c\}$. If $\mathcal{E} = \{\phi, \Omega\}$ ← trivial σ-alg. is the chosen σ-alg. assoc. to Ω, then X above is not a rv.

$\because \{a, b\} \notin \mathcal{E}$.

$\Omega = \{a, b, c\}$, $\mathcal{E} = 2^{\Omega} = \{\phi, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}$

$\{a,c\}, \{a,b,c\} = \Omega\}$

then X above IS a rv.

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

$$F_X(x) = \Pr\left(\{\omega \in \Omega : X(\omega) \le x\}\right) \quad , x \in \mathbb{R}.$$

upper small $\in \mathcal{E}$ $= \Pr(X \le x)$

Fact: i) $x \mapsto F_X(x)$ is non-decreasing.
i.e., $\forall x, y$ s.t. $x \le y$, $F_X(x) \le F_X(y)$.

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

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

iii) $F_X(\cdot)$ is right-continuous.

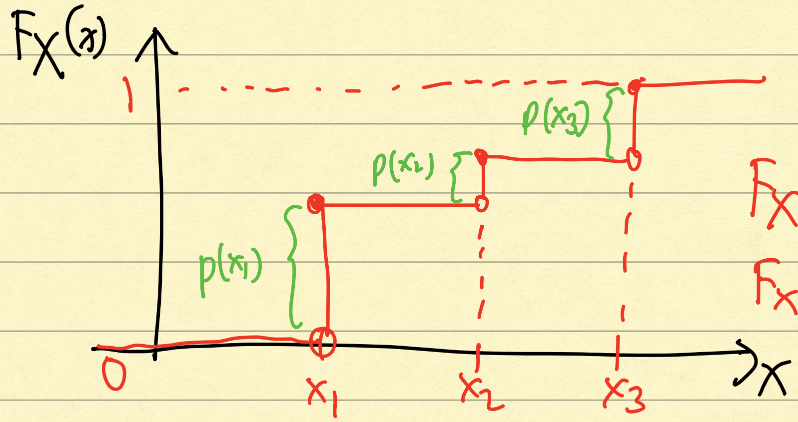
$$\lim_{\varepsilon \to 0^+} F_X(x+\varepsilon) = F_X(x) \qquad \forall x \in \mathbb{R}. \quad (\text{Ex}).$$

Special case: If X has a finite or countable number of possible values, say $x_1, x_2, x_3, \ldots$, the prob $\boxed{\Pr(X = x_i)}$ of each sample x_i is called the <u>probability mass function</u> evaluated at x_i.

$$p(x_i) = p_X(x_i) = \Pr(X = x_i). \qquad i = 1, 2, \ldots$$

$$\Pr\left(\{\omega \in \Omega : X(\omega) = x_i\}\right) = \Pr(X = x_i)$$

This need not belong to $\mathcal{E}$ at first sight but it is indeed an event & so we can write $\Pr(X = x_i)$



$F_X(x_3) = 1$

$F_X(x_3^-) = 1 - p(x_3)$

If F_X has a derivative at x, it is called the probability density function (pdf) evaluated at $x \in \mathbb{R}$, written as $f_X(x)$.

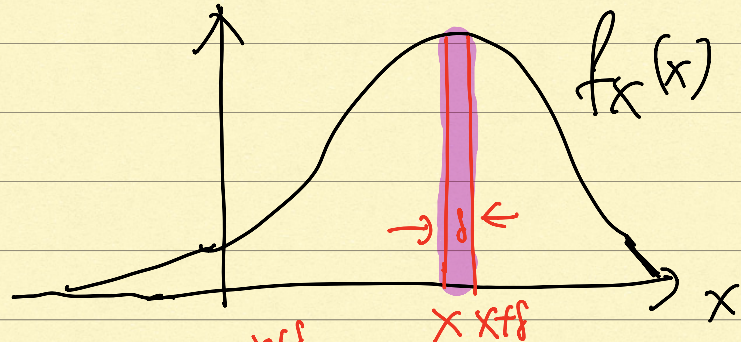
upper lower.

$$f_X(x) = F_X'(x) = \frac{d}{dx} F_X(x) \qquad x \in \mathbb{R}$$

$$f_X(x)\,\delta \simeq \Pr(x \le X \le x + \delta) = \int_x^{x+\delta} f_X(x)\,dx$$

small positive number

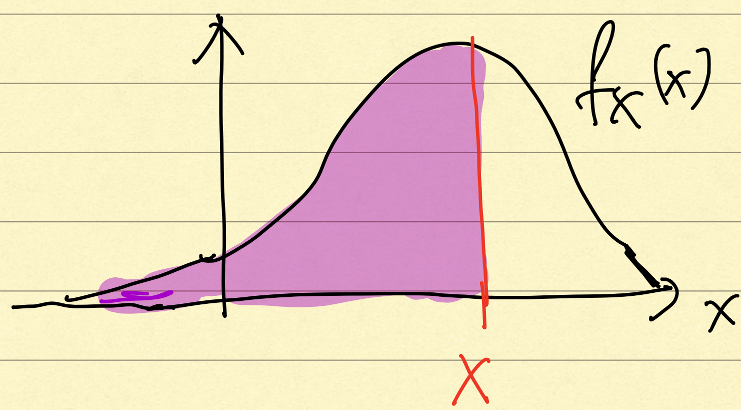
$$\Pr(x \le X \le x + \delta) = F_X(x+\delta) - F_X(x)$$



$$\int_x^{x+\delta} f_X(t)\,dt \approx \delta f_X(x)$$

Def: A rv X is continuous if $\forall x \in \mathbb{R}$, $\exists f_X(x)$. In other words $F_X(x)$ is diff$\underline{^{ble}}$ at every $x \in \mathbb{R}$.

$$Pr(X \le x) = F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$



<u>Multiple rvs</u> $\quad X_1 \cdots X_n$

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

$$= \Pr\left(\{\omega \in \Omega : X_1(\omega) \le x_1, \ldots, X_n(\omega) \le x_n\}\right)$$

Given $F_{X_1 \ldots X_n}$ how do we get the cdf of a single rv (marginal cdf)

$1 \le i \le n \qquad F_{X_i}(x_i) = F_{X_1 \cdots X_n}(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty)$

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

Independence of 2 rvs.

Two rvs are indep if

$$F_{XY}(x, y) = F_X(x) F_Y(y) \quad \forall x, y \in \mathbb{R}$$

If X, Y are discrete, this is equiv to

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

Example of a stochastic process (Bernoulli process)

Def: A stochastic process is an infinite collection of rvs defined on a common sample space Ω.

The rvs are usually indexed by an integer n

(representing discrete time) or a real-valued parameter t (representing continuous time).

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

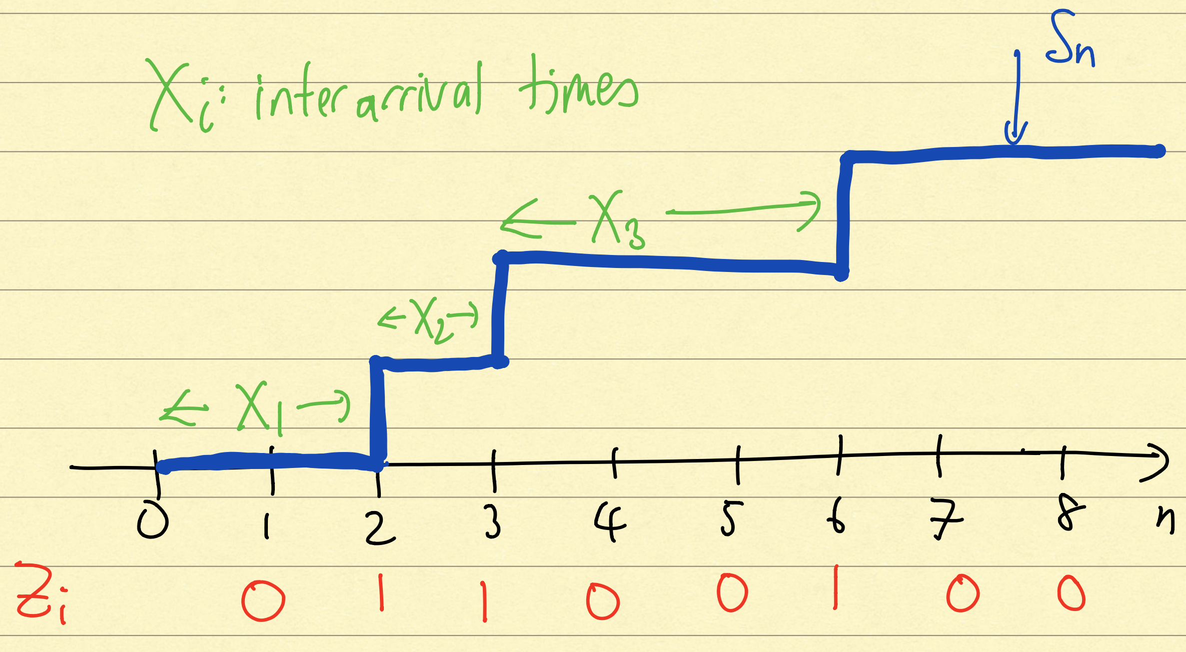
$$p = \Pr(Z_1 = 1) \qquad q = 1-p = \Pr(Z_1 = 0)$$

$\{Z_i = 1\} \Rightarrow$ a customer arrived at time i

$\{Z_i = 0\} \Rightarrow$ a customer did not arrive at time i

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

$$S_n = \sum_{i=1}^{n} Z_i$$



$S_n \quad 0 \quad 1 \quad 2 \quad 2 \quad 2 \quad 3 \quad 3 \quad 3$

Consider X_1, the first interarrival time. What is its dist$^{\underline{n}}$?

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

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

$X_1 = 3$ iff $Z_1 = Z_2 = 0, Z_3 = 1$, $\quad p_{X_1}(3) = (1-p)^2 p$.

This is geometric r.v.

$$p_{X_1}(j) = p(1-p)^{j-1}, \quad j \geq 1.$$

Claim: p_{X_k} is the same as p_{X_1} $\forall k \in \mathbb{N}$.

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

What's the dist$^{\underline{n}}$ of the partial sums $S_n = \sum_{i=1}^{n} Z_i$.

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

$$\Pr(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \leq k \leq n$$

This is called the Binomial distribution.