Lecture 1: EE5137.

Subject :

Date:

Reading: Sections 1.1 - 1.4.1 of Gallager's Book.

Ω: sample space, i.e, the set of all sample points for a given experiment.
Eg: Coin toss $\Omega = \{H, T\}$ Dice throw $\Omega = \{1,2,3,4,5,6\}$.

"Events": Subsets (Legitimate) of the sample space.
Eg: Odd outcomes $E = \{1, 3, 5\} \subset \Omega$.

If we have events $A_1, A_2, \ldots A_n \subset \Omega$, their union is denoted as $\bigcup_{i=1}^{n} A_i$ or $A_1 \cup \cdots \cup A_n$; this consists of all pts in at least one of the events $A_1, A_2, \ldots A_n$.

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

The complement of a set A, denoted as A^c, is the set of points not in A. $A^c = \Omega \setminus A$ or $A^c = \Omega \cap A^c$.

A countable set is one in which the objects can be placed in one-to-one correspondence with the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$.
Eg: The set of even numbers is countable.
The set of all integers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$ is countable.
The set of all rational numbers $\mathbb{Q} = \{m/n : m, n \in \mathbb{Z}\}$ is countable.

Axioms for Events.
Ω: sample space. The class of all subsets (of Ω) that constitute the set of (legitimate) events (i.e, the σ-algebra) satisfies
i) Ω is an event
ii) $\forall A_1, A_2, \ldots$ event, the union $\bigcup_{i=1}^{\infty} A_i$ is an event
iii) $\forall A$ event, $A^c = \Omega \setminus A$ is an event.

Subject : Date:

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

Fact: ϕ, the empty set, is an event.
Why? $\phi = \Omega^c$. Now use (i) & (iii)

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

Why? Use (ii) & let $A_{n+1} = A_{n+2} = \dots = \phi$.

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

Pf: $\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c \iff \underbrace{\left(\underbrace{\bigcup_i \underbrace{A_i^c}_{\text{event}}}_{\text{event}}\right)^c}_{\text{event}} = \underbrace{\bigcap_i A_i}_{\text{event}}$

Fact: The class of all subsets of Ω (if Ω is uncountable) does not allow for probability axioms to be satisfied in a "sensible" way.
But this is not needed for our module.

<u>Axioms of probability</u>

For any sample space Ω, and any class of legitimate events $\mathcal{E}$ (a σ-algebra which satisfies the axioms of events), a probability rule $Pr\{\dots\} : \mathcal{E} \to [0,1]$ s.t.

i) $Pr(\Omega) = 1$
ii) $\forall A \in \mathcal{E}$, $Pr(A) \geq 0$
iii) $\forall A_1, A_2, \dots$ disjoint events $\in \mathcal{E}$, $Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} Pr(A_n) = \lim_{m\to\infty} \sum_{n=1}^{m} Pr(A_n)$

This leads to the following facts:

Subject :   Date:

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

Pf: Let $A_n=\phi$ $\forall n\in\mathbb{N}$. The A_n's are disjoint. $\bigcup_{n=1}^{\infty} A_n=\phi$.

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

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

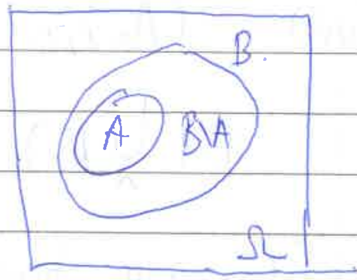
Fact ii) If $A_1,\dots,A_m$ disjoint, $Pr\left(\bigcup_{n=1}^{m} A_n\right)=\sum_{n=1}^{m} Pr(A_n)$

Apply axiom (iii) to the sequence $A_1, A_2, \dots, A_m, A_{m+1}=A_{m+2}=\dots=\phi$.

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

$\Omega=A\cup A^c$ & A & A^c are disjoint

Now use (i) & Fact (ii).



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

$B=A\cup(B\setminus A)$ A & $B\setminus A$ are disjoint.

$$Pr(B)=Pr(A)+Pr(B\setminus A)\ge Pr(A)$$

Fact (v) $Pr(A)\le 1$. Let $B=\Omega$.

Fact (vi) $\sum_n Pr(A_n)\le 1$ $\forall$ sequence of disjoint events $A_1, A_2, \dots$

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

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 $A_1, A_2, \dots \in\mathcal{E}$.

Subject: Date:

Pf: $\bigcup_{n=1}^{\infty} A_n = \bigsqcup_{n=1}^{\infty} B_n$ where $B_1 = A_1$

$B_2 = A_2 \setminus A_1$

$\vdots$

$B_n = A_n \setminus \bigcup_{i=1}^{n-1} A_i$

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

$$= \sum_{n=1}^{\infty} \Pr(B_n) = \lim_{k\to\infty} \sum_{n=1}^{k} \Pr(B_n)$$

But $\sum_{n=1}^{k} \Pr(B_n) = \Pr\left(\bigsqcup_{n=1}^{k} B_n\right) = \Pr\left(\bigcup_{n=1}^{k} A_n\right)$

$$\Rightarrow \Pr\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{k\to\infty} \Pr\left(\bigcup_{n=1}^{k} A_n\right).$$

Fact (viii) $(A_n)_{n\geq 1}$ not necessarily disjoint

$$\Pr\left(\bigcup_n A_n\right) \leq \sum_n \Pr(A_n) \qquad \text{(union-of-events bound)}$$

[Can you tighten the union bound?]

Probability Review.

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

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

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

Def: Two events are independent 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 conditionally indep. of each other given C if
$\Pr(A\cap B|C) = \Pr(A|C)\Pr(B|C)$.

Subject : Date:

Def: The events $A_1, A_2, \dots, A_n$ $(n \geq 2)$ are independent if $\forall S \subseteq \{1, \dots, n\}$ s.t. $|S| \geq 2$,

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

Rmk: This includes the entire collection, i.e., $S = \{1, 2, \dots, n\}$ but the statement

$$Pr\left(\bigcap_{i \in [1:n]} A_i\right) = \prod_{i \in [n]} Pr(A_i)$$

does not imply statistical independence among the A_i's.

Eg:

Sample point Ω	A_1	A_2	A_3
1	1	1	1
2	1	1	0
3	1	0	1
4	1	0	0
5	0	1	0
6	0	1	0
7	0	0	1
8	0	0	1

$Pr(A_1 \cap A_2 \cap A_3) = \frac{1}{8}$ $\quad Pr(A_i) = \frac{1}{2}$ $\forall i = 1, 2, 3.$

However, $Pr(A_2 \cap A_3) = \frac{1}{8} \neq Pr(A_2) Pr(A_3) \implies A_2$ & A_3 dependent!
$Pr(A_1^c \cap A_2^c \cap A_3^c) = 0 \neq Pr(A_1^c) Pr(A_2^c) Pr(A_3^c) = 1/8$

Note that $Pr(A_1 \cap A_2) = \frac{1}{4} = Pr(A_1) Pr(A_2)$ ✓ $\quad A_1 \perp\!\!\!\perp A_2$
$Pr(A_1 \cap A_3) = \frac{1}{4} = Pr(A_1) Pr(A_3).$ ✓ $\quad A_1 \perp\!\!\!\perp A_3.$

Subject: Date:

Random Variables

Def: A r.v. X is a function that maps from the sample space Ω to the real line $\mathbb{R}$ s.t.

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

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

$X(a)=1/2$, $X(b)=1$, $X(c)=3/2$.

let $x=1.1$, Then $\{\omega\in\Omega: X(\omega)\le x\}=\{a,b\}$ is an event.

let $x=0.4$. Then $\{\omega\in\Omega: X(\omega)\le x\}=\phi$ is also an event

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

$$F_X(x)=\Pr\left(\{\omega\in\Omega: X(\omega)\le x\}\right)=\Pr(X\le x)$$

Fact: i) $x\mapsto F_X(x)$ is non-decreasing

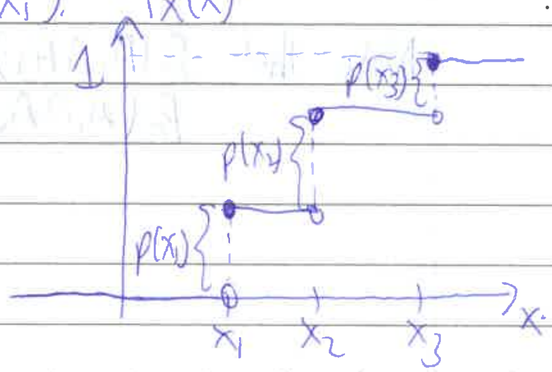
ii) $\lim_{x\to-\infty}F_X(x)=0$, $\lim_{x\to+\infty}F_X(x)=1$.

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

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

Special case: If X has a finite or countable # of possible values, say $x_1,x_2,x_3,\dots$, the prob. $\Pr(X=x_i)$ of each sample x_i is called the probability mass function evaluated at x_i.

$$p(x_i)=p_X(x_i)=\Pr(X=x_i).$$



If $F_X(x)$ has a derivative at x, the derivative is called the probability density fn or (pdf) of X.

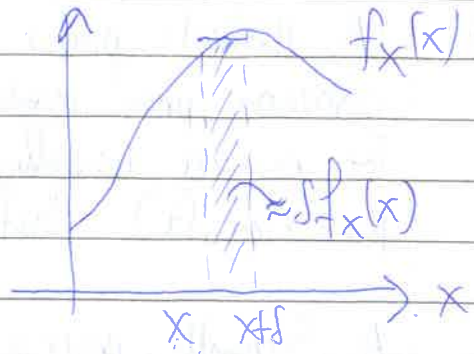
$$f_X(x)=\frac{d}{dx}F_X(x).$$

Subject: Date:

δ small

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

$$= F_X(x+\delta) - F_X(x)$$



A rv is continuous if $\exists f_X(x)\ \forall x \in \mathbb{R}$ in which case

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy$$

Multiple random variables:

Joint cdf $F_{X_1 \cdots X_n}(x_1, \ldots, x_n) = \Pr(X_1 \le x_1, \ldots, X_n \le x_n)$
$= \Pr(\{w \in \Omega : X_1(w) \le x_1, \ldots, X_n(w) \le x_n\})$

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

ith position

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

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

Independence of 2 rvs. Two rvs are independent if

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

If X, Y are discrete, this is equivalent to

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

where $\mathcal{X}$ (resp. $\mathcal{Y}$) is the set of values that X (resp. Y) takes on.

Subject : Date:

Example of a stochastic process (the Bernoulli process)

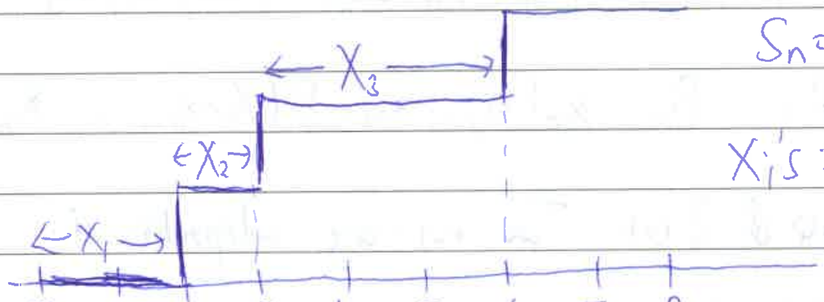
Def: A stochastic process is an infinite collection of rvs defined on a common prob. model
The rvs are usually indexed by an integer (n) or a real-valued parameter (t) that is interpreted as time.

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

$$p = P_r(Z_i = 1) \qquad q = 1-p \quad \forall i \in \mathbb{N}.$$

$\{Z_i = 1\} \Rightarrow$ customer arrives at time i.
$\{Z_i = 0\} \Rightarrow$ no customer arrives at time i.

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



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

X_i's : interarrival time

i	0	1	2	3	4	5	6	7	8
Z_i		0	1	1	0	0	1	0	0
S_i		0	1	2	2	2	3	3	3

Consider X_1: first interarrival time. What is its dist^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 $Z_1 = Z_2 = 0, Z_3 = 1$ $\quad p_{X_1}(3) = (1-p)^2 p$

Subject: Date:

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

Geometric dist$^{\underline{n}}$

In fact, the dist$^{\underline{n}}$ ~~of~~ P_{X_k}, $k=1,2,3,\ldots$ is the same as P_{X_1}, a geometric distribution.

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 an including time n.

Binomial distribution $Bin(n, p)$

$$Pr(S_n = k) = p_{S_n}(k) = \text{prob. that } k \text{ out of } n\ Z_i\text{'s equal } 1.$$

$$= \binom{n}{k} p^k (1-p)^{n-k} \qquad 0 \leq k \leq n.$$

Often, we want to study $p_{S_n}(k)$ for larger n & $k = \lfloor \alpha n \rfloor$ for some ~~$0 \leq \alpha \leq n$.~~ $\alpha \in [0,1]$. See book for details.