

28/9: Lecture 1 - MOTIVATION & HISTORY

.svg/.svg
format
styling labels

Historically - Nala - Rituparna (links on Zulip).

Great Gambler - "Sakuni".

"Modern-day history" - Luca Pacioli (1494) - Problem of points (Wikipedia).

Game of rounds by two players - Equal chance of winning a round.

First player to win 'n' rounds is the winner. Say $n=100$.

1st player has won n_1 rounds & 2nd player has won n_2 rounds. Game stops!

Prize Money = P . How does one split P ?

Find prob. of Player 1 winning, say p_1 . $p_2 = 1 - p_1$.

Player 1 gets $P \times p_1$ & Player 2 gets $P \times p_2$.

Pacioli proposed $p_1 = \frac{n_1}{n_1 + n_2}$; Niccolo Tartaglia in 1600's
Error! $n_1=2, n_2=0, n=100 \Rightarrow p_1=1$
1600's Luigi Cardano wrote a book on "Games of chance" - not known until 1600's.

NT proposed p_1 based on $n_1 - n_2$.

Error! $n=100$; $n_1=75, n_2=65 \Rightarrow p_1$ based on $\frac{1}{10}$.
 $n_1=99, n_2=89 \Rightarrow p_1$ based on $\frac{1}{10}$.

Resolved later by Fermat & Pascal.

beginning of Prob. theory.

Gambler's ruin: (17th Century) 2 players play. A winner of a round gets 1 INR
- loser loses 1 INR.
Total money is n INR = (sum of money with both players)

Game ends when a player gets all n INR.

Players start with k_0 INR & $n - k_0$ INR.

They stop abruptly when they have " k " INR & " $n - k$ " INR.

How to split the money? Pascal & Fermat solved it.

Laplace made a "prediction" of prob. that the sun will rise tomorrow.
Various attempts to formalize — von Mises ^(1900's) using a frequentist / empirical approach.
still no axiomatic approach like Euclidean geometry.

Kolmogorov laid down axiomatic framework in 1930's.

[Ref: Wikipedia or Scheffe's book]

"DATA SCIENCE"

This course — "Intuitive / Heuristic / computational" approach.

Any mathematical discipline has

- (1) Formal logical content. (basics in this course)
- (2) Intuition & Ideas (lots in ")
- (3) Applications (good many)

Read Ch. 1 of Feller Vol. 1 or more on philosophy & motivation.

Caution: Unlike many math., prob. can be described in everyday English (toss a coin, roll a die et al.) — advantageous but also scope for confusion.

Current-day applications:

Statistical physics: Exact positions of particles not fixed but random.

Biology: Genes mutate, evolve & spins in a magnet are random.
Populations evolve randomly; combine randomly.

Data Science / Machine Learning:

You go to Google — you type ISI, which ISI should Google show?
"Predict under uncertainty"

Sally Clark case:

Applications within mathematics -

- Probabilistic method (pioneered by Erdős).

Prob. ("object satisfies a constraint") $> 0 \Rightarrow \exists$ one object satisfying the constraint



- Two polynomials $P(x)$, $Q(x)$.

$$P(x) = \sum_{i=0}^n c_i x^i; \quad Q(x) = \prod_{i=1}^n (x - a_i) = \sum_{i=0}^n b_i x^i$$

qf $P(x) = Q(x)$?

Take $\boxed{x_1, \dots, x_{n+1}}$, suppose it is equal at $x_1, \dots, x_n$ & not the rest $\Rightarrow P \neq Q$.

Randomized algorithms - suppose equal at $x_1, \dots, x_{n+1} \Rightarrow P = Q$.
choose x_i "randomly" $P(P(x_i) \neq Q(x_i))$.

Eg: Matrices A, B, C $n \times n$ matrices qf $\underline{AB} = \underline{C}$?

Casino's - Machines shuffle cards - qf the shuffling good?
Do long enough - shuffling is good.

Information theory: n bits $x = 01 \dots 1 \dots 0 \dots$ sent over a "noisy" network.
Bits are corrupted randomly. You receive $y = 10 \dots 10$

Can we code " x " s.t. you can recover x from y ?
"adding more bits"

28/9 LECTURE - 2.

Basics of set theory & Combinatorial analysis. - Refer to D'Angelo & West; Ross - Ch. 1 & parts of Ch. 2; K.L. Chung - Ch. 1 ("Elementary Prob.")

(TA will revise some on Friday)

Say we perform an "Experiment". Assume there are finitely many possible outcomes. Our goal in probability is to mathematically model the "experiment" and make "prediction" about outcomes.

The bigger focus here - How to "understand" the model?
 (Indicates there is uncertainty about outcomes)

(Finite) **SAMPLE SPACE:** Ω - a finite set ["Set of all possible outcomes" for the experiment]

(later) **Event:** An event is a subset of Ω . ["collection of outcomes"]

Denote events by A, B, C & so on.

Notation: In general small letters (say x, y, w, w_i, x_i & so on) for elements of Ω & capitals for subsets or events.

Fig. 1: $\Omega = \{0, 1\}$ [Toss a coin; $H=1$ & $T=0$ ^{EXPERIMENT} possible outcomes]
 $A = \emptyset, \{0\}, \{1\}, \Omega$. [Play a game; $W=1$ & $L=0$]

Fig 2: $\Omega = \{1, 2, \dots, 6\}$ [Roll a dice - 6 possible outcomes - $1, \dots, 6$]

Fig. A = Set of all even nos = $\{2, 4, 6\}$.

Fig. 3: $\Omega = \{0, 1\}^n$; $n \in \mathbb{N} = \{1, 2, 3, \dots\}$ - natural numbers.

$\{(a_1, \dots, a_n) : a_i \in \{0, 1\}\}$ [Toss a coin 'n' times; Play a game of

Ex: Compute cardinalities of sample spaces & some events ^{'n' rounds}
($|\Omega|$) ($|A|$)

Fig. 4: $\Omega = \{0, 1, \dots, n\}$ - [what is experiment? ^{cardinality.}]

Fig 5: $\Omega = [52] = \{1, \dots, 52\}$ ($[n] := \{1, 2, \dots, n\}$)

(selecting a card from a pack)
 A = set of spades. [label cards by $1, \dots, 52$] $|A| = ?$

Convention: $\{1, \dots, 13\}$ $\{14, \dots, 26\}$ $\{27, \dots, 39\}$ $\{40, \dots, 52\}$
spade club ♦ Hearts.

Examples from statistical physics.

Particles - Electrons/Protons/Neutrons.

Particles occupy cells/sites. always labeled.

Assume there are ' n ' cells & ' r ' particles. $n, r \in \mathbb{N}$.

Two possible cases - labelled particles (i.e., we can distinguish particles)
 - unlabelled particles (we know no: of particles in a cell)

Fig 6: Ω = possible positions of labelled particles.
 $= \{ (p_1, \dots, p_r) \mid \text{such that or a constraint } 1 \leq p_i \leq n \text{ for } 1 \leq i \leq r \} \mid p_i \in \mathbb{N}$
 (p_i - position of the i th particle).

$= \{1, \dots, n\}^r =: [n]^r$
↳ [defn. of a "quantity"]

A = At least one cell is empty

A_i^c = i th cell is empty $= \{ (p_1, \dots, p_r) \in \Omega : p_k \neq i \text{ for all } 1 \leq k \leq r \}$
 ($\subseteq \Omega$) ↳ belongs to ↳ for all
↳ such that

So $A = \bigcup_{i=1}^n A_i^c \subseteq \Omega$ (cos $A_i^c \subseteq \Omega$)

if $n > r$ then $A = \Omega$. [L*]

Ex: Describe Ω when two ^{lab.} particles can't occupy the same cell. (Ω here is a subset of $[n]^r$)
 (n cells & r particles)

Eg: Cells labelled $1, \dots, n$ — Particles cannot occupy successive cells.
(spectrum frequencies) (successive frequencies can't be allocated)

Unlabelled Particles

[see Feller I.2 for other contexts]

Eg: $\Omega_{r,n}$ = possible configurations of unlabelled particles
 $= \{ (r_1, \dots, r_n) : r_i \in \{0, \dots, r\} \ \forall \ 1 \leq i \leq n$
 $\& \ r_1 + \dots + r_n = r. \}$
 $r_i = \#$ of particles in cell i .

interesting Events $A = ?$

Eg: $\Omega_{r,n}^*$ = possible config of unlab. particles such that no two particles at same site.
 $= \{ (r_1, \dots, r_n) : r_i \in \{0, 1\} \ \forall \ 1 \leq i \leq n$
 $\& \ r_1 + \dots + r_n = r \}$

[Note: $\Omega_{r,n}^* \subseteq \Omega_{r,n}$; $\Omega_{r,n}^* \subseteq \{0, 1\}^n$]

Experiment is "interesting" if only outcomes ^{can} differ every time you perform the experiment.
 Natural to consider "the probability" of an outcome.

(means "certainty with which outcome occurs").
 We need to define $P(w)$ $\forall \ w \in \Omega$.

von Mises (Frequentist approach) : $P(w) \approx$ "proportion of occurrence of w when experiment is repeated many times".
Problem: Experiment has to be repeated many times — say for n large.

Also say you toss a coin 1000 times, 6% heads $\Rightarrow$ " $P(1) = 0.6$ "
 number I toss the coin further 9000 times, total 55% heads $\Rightarrow$ " $P(1) = 0.55$ "
 $\frac{\# \text{ of occurrence of } w \text{ in 'n' experiments}}{n} \longrightarrow ? ?$
 as $n \rightarrow \infty$

Defn: Let Ω be a finite set. Let $p: \Omega \rightarrow [0,1]$ be a function such that (s.t.) $\sum_{\omega \in \Omega} p(\omega) = 1$. Then (Ω, p) is a finite prob. space. p is called prob. mass function (pmf).

Define $P: 2^\Omega \rightarrow [0,1]$ as $P(A) = \sum_{\omega \in A} p(\omega)$, $A \subseteq \Omega$.
ii
Power-set of Ω Then P is called prob. distribution.

$\{A: A \subseteq \Omega\}$

[Note: $|2^\Omega| = 2^{|\Omega|}$]

Find interesting (Ω, p) & an event A and compute $P(A)$.

30/9 LECTURE 3: EXAMPLES OF PROB. SPACES.

Defn 3.1 Let Ω be a finite set. Let $p: \Omega \rightarrow [0,1]$ be a function such that (s.t.) $\sum_{\omega \in \Omega} p(\omega) = 1$. We call (Ω, p) a finite prob. space.

p is called prob. mass function (pmf). subset of

Define $P: 2^\Omega \rightarrow [0,1]$ as $P(A) := \sum_{\omega \in A} p(\omega)$, $A \subseteq \Omega$, $P(\emptyset) = 0$.

2^Ω Power-set of Ω Then P is called prob. distribution.

$\{A: A \subseteq \Omega\}$

[Note: $|2^\Omega| = 2^{|\Omega|}$]

Find interesting (Ω, p) & an event A and compute $P(A)$.

PROPN 3.2: Let (Ω, p) be a finite prob. space and P as defined above. Then P satisfies the following properties.

[PROPOSITION]

(i) $P(A) \geq 0 \quad \forall A \subseteq \Omega$ (ii) $P(\Omega) = 1$

(iii) $P(A_1 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i)$ if $A_1, \dots, A_n \subseteq \Omega$ & are pairwise disjoint ($A_i \cap A_j = \emptyset \quad \forall i \neq j$)
[finite additivity]

Proof: (i) Since $p(\omega) \geq 0$ then so is $P(A) = \sum_{\omega \in A} p(\omega)$, $A \subseteq \Omega$.

(ii) $P(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$.
(by defn of P) (by defn. of pmf)

(iii) We shall prove by induction on n .

finite additivity is trivially true for $n=1$ ($P(A_1) = P(A_1)$)

Suppose it is true for $n-1$.

Let us consider f.d. for n .

Say $A_1, \dots, A_n$ are pairwise disjoint.

Define $B := \bigcup_{i=1}^n A_i$

Since $A_1, \dots, A_n$ are pairwise disjoint then $A_n \cap B = \emptyset$.

$$P\left(\bigcup_{i=1}^n A_i\right) = P(A_n \cup B) \quad (\text{by defn of } B)$$

$$(A_1 \cup \dots \cup A_n) = \sum_{\omega \in A_n \cup B} p(\omega) \quad (\text{defn of } P)$$

$$\left(\sum_{\substack{i \in I \\ I \cap J = \emptyset}} a_i = \sum_{i \in I} a_i + \sum_{i \in J} a_i \right) = \sum_{\omega \in A_n} p(\omega) + \sum_{\omega \in B} p(\omega) \quad (A_n \cap B = \emptyset)$$

$$= P(A_n) + P(B) \quad (\text{defn of } P)$$

$$= P(A_n) + \sum_{i=1}^{n-1} P(A_i) \quad (\text{since f.a. holds for } n-1)$$

$$= \sum_{i=1}^n P(A_i). \quad \blacksquare - \text{Q.E.D.}$$

Ex. 3.3 Suppose $P: 2^\Omega \rightarrow \mathbb{R}$ satisfying (i), (ii) & (iii) in PROP 3.2. Show that $\exists$ a $p: \Omega \rightarrow [0,1]$ s.t. (Ω, p) is a finite prob. space &

P is the prob. distribution corresponding to pmf p .
i.e., $P(A) = \sum_{\omega \in A} p(\omega) \quad \forall A \subseteq \Omega$.

Given (Ω, p) , we get P satisfying (i), (ii) & (iii). [Prop 3.2]

Given Ω & P , we get that $\exists p \ni (\Omega, p)$ is a prob. space & P is the prob. distribution. [Ex. 3.3]

[Satisfying (i), (ii) & (iii)]

Ex 3.4: $p(\omega) = \frac{1}{|\Omega|}$

Ω finite set, $\Omega \neq \emptyset$.

[Uniform prob. distribution on Ω]
(PD)

Easily check p is a pmf on Ω .

We'll consider uniform PD on the examples of Ω 's we have discussed.

Eg 3.5: $\Omega = \{0, 1\}$. Define $p(0) = \frac{1}{2} = p(1)$.

[Tossing a fair coin]
Since prob. of H = prob. of T

Easy to compute $P(A) \quad \forall A \subseteq \Omega$

$$A = \emptyset \Rightarrow P(A) = 0; \quad A = \{0\}, P(A) = p(0) = 1/2$$

$$A = \{1\} \Rightarrow P(A) = p(1) = 1/2; \quad A = \Omega \Rightarrow P(\Omega) = p(0) + p(1)$$

$$= 1.$$

Eg 3.6 $\Omega = \{0, 1\}^2 = \{(0,0), (0,1), (1,0), (1,1)\}$

[Tossing two coins]
← [both coins are fair]

$$P((0,0)) = P((0,1)) = P((1,0)) = P((1,1)) = \frac{1}{4}.$$

P - uniform PD.

$$A = \text{first toss is a head} = \{(1,0), (1,1)\}$$

$$P(A) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

$$B = \text{Two tosses are not the same} = \{(0,1), (1,0)\}$$

$$P(B) = 1/2.$$

Ex 3.7: (Ω, p) is prob. space with uniform PD. Then show that

$$P(A) = \frac{|A|}{|\Omega|}.$$

Eg 3.8: $\Omega = \{1, \dots, 6\}$

[Roll a dice]

$$p(\omega) = \frac{1}{6}$$

[fair dice]

$$A = \text{Dice shows an even number} = \{2, 4, 6\}$$

$$\text{So } P(A) = \frac{|A|}{|\Omega|} = \frac{1}{2}$$

(By Ex.)

$$A = \text{Dice shows a multiple of 3} = \{3, 6\}$$

$$P(A) = \frac{1}{3}$$

Eg 3.9: $\Omega = \{0, 1\}^n = \{(a_1, \dots, a_n) : a_i \in \{0, 1\} \quad \forall 1 \leq i \leq n\}$ [bit string or n coin tosses]

$$p(w) = \frac{1}{2^n} \quad w = (w_1, \dots, w_n) \in \Omega. \quad \left[\begin{array}{l} \text{Toss coins randomly or} \\ \text{message is chosen randomly} \end{array} \right]$$

$$B_k = \text{set of bit strings with exactly } k \text{ 1's (or 1's)} \\ = \{w : \sum_{i=1}^n w_i = k\}, \quad k = 0, \dots, n.$$

$$P(B_k) = \frac{|B_k|}{|\Omega|} = \frac{|B_k|}{2^n} = \binom{n}{k} \frac{1}{2^n} \quad \left[|B_k| = \binom{n}{k} \right] \\ \text{Ex. } \binom{n}{k}$$

Eg 3.10: $\Omega = [52]$; $p(w) = \frac{1}{52}, w \in \Omega.$ [Pick a card uniformly at random (uar)]

$A_s = \text{set of spades}$

$$P(A_s) = \frac{|A_s|}{|\Omega|} = \frac{1}{4}.$$

$$A_1 = \text{set of 1's} \Rightarrow P(A_1) = \frac{|A_1|}{|\Omega|} = \frac{1}{13}.$$

Space of all outcomes — Ω

how an outcome is chosen in the experiment — p or P

Event — colln. of outcomes, $A \subseteq \Omega.$

In prev. example, Prob. a card selected uar is a spade $= P(A_s).$

Eg. 3.11: [BOSE-EINSTEIN PROB. DISTRIBUTION]

$$\Omega_{r,n} = \{ (r_1, \dots, r_n) : r_i \geq 0 \forall i \text{ \& } \sum_{i=1}^n r_i = r \} \quad \left[\begin{array}{l} \text{Configuration} \\ \text{of unlabelled} \\ \text{particles} \end{array} \right]$$

$$p(w) = \frac{1}{|\Omega_{r,n}|}, \quad w \in \Omega_{r,n}.$$

$$A_i^c = \text{ith cell is empty. } P(A_i^c) = ?$$

Eg 3.12: [Fermi-Dirac PD]

[Configuration of electrons/unlabelled & not particles in the same site]

$$\Omega_{r,n}^* = \{ (r_1, \dots, r_n) : r_i \in \{0,1\} \forall i \text{ \& } \sum_{i=1}^n r_i = r \}$$

$$p(\omega) = \frac{1}{|\Omega_{r,n}^*|}$$

$A_i = i^{\text{th}}$ cell is empty. $P(A_i) = ?$

ASIDES

Perform an experiment n times. Observe the outcomes

& see how many outcomes are in $A_1, \dots, A_k$ ($A_1, \dots, A_k \subseteq \Omega$)

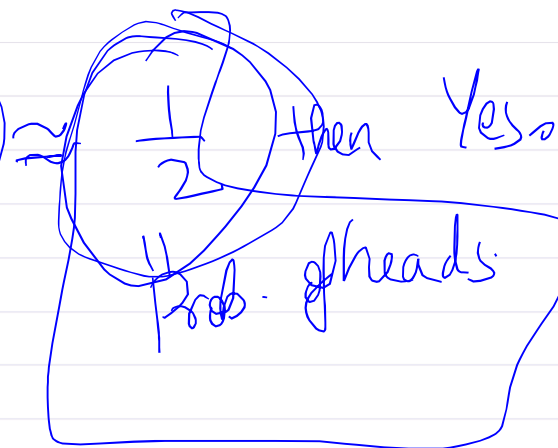
if $\forall i$ $P(A_i) \approx \frac{\# \text{ of times outcome in } A_i}{n}$ then the prob. model is "good".

Is the coin really fair?

Toss a coin n times.

if

$$\frac{\# \text{ of heads}}{n}$$



For eg. $\frac{\# \text{ of heads}}{n} = \underline{0.55}$
 $n=100$

5/10 LECTURE 4: MORE EXAMPLES & CALCULATIONS

→ we'll refer to prob. spaces as $(\Omega, \mathcal{P})$ & $\mathcal{P}(\Omega, \mathcal{P})$ as specifying pmf or PD are equivalent.

→ if $P: 2^\Omega \rightarrow [0, \infty)$ satisfies $P(A \cup B) = P(A) + P(B)$ for $A \cap B = \emptyset$ then it also satisfies finite additivity. [Ex.] $\left[2^\Omega = \{A: A \subseteq \Omega\} \right]$
 $\mathcal{P}(\Omega)$

THM 5.1 Let $(\Omega, \mathcal{P})$ be a fin. prob. space. Then

(i) $P(A) \leq P(B)$ $\forall A \subseteq B \subseteq \Omega$. [Monotonicity]

(ii) $P(A) \leq 1$ $\forall A \subseteq \Omega$

(iii) $P(\emptyset) = 0$

(iv) $P(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^n P(A_i)$ $\forall A_1, \dots, A_n \subseteq \Omega$. [finite subadditivity]

(v) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ [Inclusion-Exclusion]

Proof later.

Remarks: (i) $\Rightarrow$ (ii) is easy. Take $B = \Omega$ & use $P(\Omega) = 1$.

(iii) follows because $P(A \cup \emptyset) = P(A) + P(\emptyset)$ for any $A \subseteq \Omega$.

$\Rightarrow P(A) = P(A) + P(\emptyset)$ as $A \cap \emptyset = \emptyset$.

$\Rightarrow P(\emptyset) = 0$.

Ex 5.2 Let $(\Omega, \mathcal{P})$ be unif. prob. space i.e., $P(A) = \frac{|A|}{|\Omega|}$.

Then (i)-(v) can be re-stated respectively as

(i) $|A| \leq |B|$ for $A \subseteq B \subseteq \Omega$

(ii) $|A| \leq |\Omega|$ for $A \subseteq \Omega$.

(iii) $|\emptyset| = 0$

(iv) $|A_1 \cup \dots \cup A_n| \leq \sum_{i=1}^n |A_i|$ $\forall A_1, \dots, A_n \subseteq \Omega$ [- - -]

(v) $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$ - Proof via Venn diagrams.

Equivalent to
 $\left[\Leftrightarrow P(A) \leq P(B) \right]$
 $\left[\Leftrightarrow P(A) \leq 1 \right]$
 $\left[\Leftrightarrow P(\emptyset) = 0 \right]$

Ex 5.3: (Bose-Einstein prob. distribution) n cells, r unlabelled particles.

$$\Omega_{r,n} = \{ (r_1, \dots, r_n) : r_i \geq 0 \forall i \text{ \& } \sum_{i=1}^n r_i = r \}$$

$$A_i = \text{\textit{i}}^{\text{th}} \text{ cell is empty} = \{ (r_1, \dots, r_n) \in \Omega_{r,n} : r_i = 0 \}$$

P - uniform PD i.e.,

$$P(A) = \frac{|A|}{|\Omega_{r,n}|}$$

Compute $P(A_i)$.

By defn of P , $P(A_i) = \frac{|A_i|}{|\Omega_{r,n}|}$

$$\begin{aligned} A_i &= \{ (r_1, \dots, r_n) \in \Omega_{r,n} : r_i = 0 \} \\ &= \{ (r_1, \dots, r_n) : r_j \geq 0 \forall j, \sum_{j=1}^n r_j = r, r_i = 0 \} \\ &= \{ (r_1, \dots, r_n) : r_j \geq 0 \forall j, \sum_{\substack{j=1 \\ j \neq i}}^n r_j = r, r_i = 0 \} \\ &= \{ (r_1, \dots, r_n) \in \Omega_{r,n} : \sum_{\substack{j=1 \\ j \neq i}}^n r_j = r \} \end{aligned}$$

$$B = \{ (r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_n) : r_j \geq 0 \forall j \neq i, \sum_{\substack{j=1 \\ j \neq i}}^n r_j = r \}$$

$$|B| = |\Omega_{r,n+1}| \leftarrow \begin{matrix} \downarrow \\ r \text{ unlabelled particles} \\ \text{in } (n+1) \text{ cells} \end{matrix}$$

(Ex.) $\exists$ a bijection from A_i to B

$$\left[\text{Hint: } (r_1, \dots, r_n) \in A_i \xrightarrow{\text{maps to}} (r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_n) \in B \right]$$

$$\text{So } |A_i| = |B| = |\Omega_{r,n+1}|$$

$$\Rightarrow P(A_i) = \frac{|\Omega_{r,n+1}|}{|\Omega_{r,n}|} \quad \left[|\Omega_{r,n+1}| \text{ in Assignment} \right]$$

substitute & find the value.

Suppose A = at least one cell is empty.

Then $A = \bigcup_{i=1}^n A_i$. But A_i 's aren't disjoint.

$$\text{finite sub-add} \Rightarrow P(A) \leq \sum_{i=1}^n P(A_i) = \frac{n |\Omega_{r,n+1}|}{|\Omega_{r,n}|}$$

if $n \frac{|\Omega_{r,n-1}|}{|\Omega_{r,n}|} \geq 1$ then the upper bound is "useless" as we know that $P(A) \leq 1$ $\square$

PROPERTY of P .

LEMMA 5.4: Let (Ω, P) be prob. space. Then $\forall A \subseteq B$

$$P(B \setminus A) = P(B) - P(A).$$

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

Proof: $B = A \sqcup (B \setminus A) \rightarrow$ disjoint union. i.e., $A \cap (B \setminus A) = \emptyset$
(union of disjoint sets)

finite add. $\Rightarrow P(B) = P(A) + P(B \setminus A)$

$$\Rightarrow P(B \setminus A) = P(B) - P(A).$$

if $B = \Omega$ then $B \setminus A = A^c$ & $P(B) = 1.$

This gives $P(A^c) = 1 - P(A).$ $\square$

Back to Bose-Einstein PD.

$A =$ at least 1 cell is empty $= \bigcup_{i=1}^n A_i^c$

$$A^c = \bigcap_{i=1}^n A_i = \{ (r_1, \dots, r_n) \in \Omega_{r,n} : r_i \geq 1 \ \forall i \}$$

$$= \{ (r_1, \dots, r_n) : r_i \geq 1 \ \forall i, \sum_{i=1}^n r_i = r \}$$

Trivially $A^c = \emptyset$ if $r < n$

$$\Rightarrow P(A) = 1 \text{ if } r < n.$$



Assume $r \geq n$. $A^c = \{ (r_1, \dots, r_n) : r_i \geq 1 \ \forall i, \sum_{i=1}^n r_i = r \}$

[Ex: show this is a bijection]

$\downarrow$ - maps to

$$\{ (r_1-1, \dots, r_n-1) : r_i-1 \geq 0 \ \forall i, \sum_{i=1}^n (r_i-1) = r-n \}$$

|| (by putting $s_i = r_i-1$)

$A^c \rightarrow \Omega_{r,n}$ is a bijection.

$$\Omega_{r,n} = \{ (s_1, \dots, s_n) : s_i \geq 0 \ \forall i, \sum_{i=1}^n s_i = r-n \}$$

$r-n$ unlabelled particles into n cells.

So $|A^c| = |\Omega_{r-n,n}|$.

$$P(A) = 1 - P(A^c) = 1 - \frac{|A^c|}{|\Omega_{r,n}|} = 1 - \frac{|\Omega_{r-n,n}|}{|\Omega_{r,n}|},$$

Substitute values for $|\Omega_{r-n,n}|$ & $|\Omega_{r,n}|$ from

Assignment.

Ex 6.9: $A =$ At least one cell has at least 2 particles. $(P(A) = ?)$

Ex 6.9: (Fermi-Dirac distribution). Assume $r \leq n$.

$$\Omega_{r,n}^* = \{ (r_1, \dots, r_n) : r_i \in \{0,1\} \ \forall i \ \& \ \sum_{i=1}^n r_i = r \}$$

$A_i^c =$ i^{th} cell is empty.

Compute as before $P(A_i^c) = \frac{|A_i^c|}{|\Omega_{r,n}^*|} = \begin{cases} 0 & \text{if } r > n \\ > 0 & \text{as } r \leq n \end{cases}$

Show like in BE PD, $|A_i^c| = |\Omega_{r,n+1}^*|$

i.e., if i^{th} cell is empty, we are distributing r particles in $(n+1)$ -cells & no two particles in a cell.

$$\Rightarrow P(A_i^c) = \frac{|\Omega_{r,n+1}^*|}{|\Omega_{r,n}^*|} \rightarrow (> 0 \text{ because } r \leq n) \quad \text{(substitute values from Assignment)}$$

[note $r = n \Rightarrow \Omega_{r,n+1}^* = \emptyset \Rightarrow P(A_i^c) = 0$ - Easy]

$A = \bigcup_{i=1}^n A_i^c =$ at least one cell is empty.

$$P(A) = \begin{cases} 1 & r < n \\ 0 & r = n \end{cases}$$

7/10 LECTURE 5: MORE EXAMPLES ON "EQUALLY LIKELY OUTCOMES".

For us prob. space, finite set Ω & pmf p or PD P .

"In school" — finite set Ω & $P(A) = \frac{|A|}{|\Omega|} = \frac{\# \text{ favourable outcomes}}{\# \text{ outcomes}}$
 Prob. of A

Such a P as above is what we call as uniform PD or "Equally likely outcomes".

For eg. by saying Ω is a prob. space with equally likely outcomes means that the PD P on Ω is uniform PD.

Why go from "Eq. lik. outcomes" to more gen. prob. spaces?

— In reality, not all outcomes are equally likely.

Eg: $\Omega = \{W, L\}$ of "eq. lik. out." then $p(W) = p(L) = 1/2$.

Eg 5.2 [sampling with replacement]

A box contains ' n ' ^{labelled} objects. We pick an object, note down the label & put it back. Then we repeat this process $(k+)$ times more.

— choosing k objects out of n (with repetition/replacement).

sample space $\Omega =$ set of k -tuples from $[n] := \{1, \dots, n\}$

$$= \{(r_1, \dots, r_k) : r_i \in [n] \forall i\}$$

[k labelled particles into ' n ' cells]

$$= [n] \times [n] \times \dots \times [n] \quad (A \times B \text{ — Cartesian product})$$

$$[A \times B = \{(a, b) : a \in A, b \in B\}, \text{ } A \text{ \& B are sets}]$$

$$=: [n]^k$$

$$(A^k = A \times \dots \times A \text{ } k\text{-times})$$

$$A^k = \{(a_1, \dots, a_k) : a_i \in A \forall i\}$$

Equally likely outcomes $\rightarrow p(w) = \frac{1}{|\Omega|} = \frac{1}{n^k}$, $w \in \Omega$
 $w = (w_1, \dots, w_k)$
 (Ω, p) or (Ω, P) is said as "sampling with replacement from $[n]$ k -many times".

Ex 3: [hashing] There are k individuals & each individual is assigned a code. There are n distinct codes/labels.

For proper identification, codes of k ind. "should" be distinct.

Keeping track of prev. codes isn't easy.

Say a 'lazy' person assigns (hash) codes randomly.

Sample space = set of all ^{possible assignment of} (hash) codes for k individuals.

$$= \{ (h_1, \dots, h_k) : h_i \in [n] \forall i \}$$

$$= [n]^k \quad h_i - \text{hash code of } i^{\text{th}} \text{ person.}$$

All assignments are equally likely.

$$\Rightarrow p(w) = \frac{1}{n^k} \quad w = (w_1, \dots, w_k) \in [n]^k$$

(Error) $E = \{h_1, \dots, h_k \text{ are not distinct}\}$

(Good) $G = E^c = \{h_i \text{'s are distinct}\}$

$$= \{ (h_1, \dots, h_k) \in \Omega : h_i \neq h_j \forall i \neq j \}$$

Prob. (lazy hash code is good) $= P(G)$ (defn of G)

$$= \frac{|G|}{|\Omega|} \quad (\text{Eq. likely outcomes})$$

$$= \frac{|G|}{n^k}$$

$$|G| = n(n-1) \cdot \dots \cdot (n-k+1) = \prod_{i=0}^{k-1} (n-i)$$

$$P(G) = \prod_{i=0}^{k-1} \left(1 - \frac{i}{n}\right)$$

Eg 5.4 (Birthday problem) — if there are k people in a room
 what is the chance a pair shares same birthday? [Von Mises & others?]
 label days as $1, \dots, 365$; Set $n = 365$.

b_i = birthday of i^{th} person.

Ω = all possible birthdays of k people

$$= \{ (b_1, \dots, b_k) : b_i \in [n] \forall i \}$$

(Same as sampling with repl. or hashing)

Assume any ordering of b_i 's is equally likely.

$$P(w) = \frac{1}{n^k} \quad w = (w_1, \dots, w_k) \in \Omega$$

$$IP(\exists \text{ at least one pair sharing same birthday}) = IP(\{ (b_1, \dots, b_k) : b_i = b_j \text{ for some } i \neq j \})$$

(from computation of $P(B)$ in Eg 5.3)

$$= 1 - \prod_{i=0}^{k-1} \left(1 - \frac{i}{n}\right) = c(k).$$

$n = 365$

$c(2) = 0.002$, $c(5) = 0.027$, $c(10) = 0.116$, $c(23) = 0.507$
 if there are 23 people, $IP(\text{at least 2 people share same birthday}) > 1/2$.

$$c(40) > 0.8 \quad c(50) \approx 1.$$

$IP(\text{person 1 shares a birthday with one another})$

$$= 1 - IP(\text{person 1 doesn't share a birthday with anybody})$$

$$= 1 - IP(\{ (b_1, \dots, b_k) : b_j \neq b_1 \forall j \neq 1 \})$$

$$= 1 - \frac{|\{ (b_1, \dots, b_k) : b_j \neq b_1 \forall j \neq 1 \}|}{n^k}$$

$$= 1 - \frac{\underbrace{(n-1) \cdot (n-1)^{k-1}}_{b_1}}{n^k} = 1 - \left(1 - \frac{1}{n}\right)^{k-1} = b(k)$$

$$b(40) \leq 0.11, \quad c(40) \geq 0.89.$$

$$\approx \frac{40}{365}$$

THM 4.1 Let $(\Omega, \mathcal{P})$ be a fin. prob. space. Then

(i) $P(A) \leq P(B) \quad \forall A \subseteq B \subseteq \Omega$. [Monotonicity]

(ii) $P(A) \leq 1 \quad \forall A \subseteq \Omega$

(iii) $P(\emptyset) = 0$

(iv) $P(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^n P(A_i) \quad \forall A_1, \dots, A_n \subseteq \Omega$. [finite subadditivity]

(v) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ [Inclusion-Exclusion]

Proof: (i) $P(A) = \sum_{\omega \in A} p(\omega) \leq \sum_{\omega \in B} p(\omega) = P(B) \quad (A \subseteq B \text{ \& } p(\omega) \geq 0 \quad \forall \omega)$
= $P(B)$ (defn)

(ii) Set $B = \Omega$. (i) $\Rightarrow P(A) \leq P(\Omega) = 1$
 $\hookrightarrow$ (prop. of P)

(iii) $P(A \cup \emptyset) = P(A) + P(\emptyset) \quad \text{as } A \cap \emptyset = \emptyset$.
 $\Rightarrow P(A) = P(A) + P(\emptyset) \quad \text{as } A \cup \emptyset = A$.
 so $P(\emptyset) = 0$.

(iv) we'll prove by induction.

$n=1$ is trivially true. $P(A_1) \leq P(A_1) \quad \checkmark$

Let $n=2$. Then $A_1 \cup A_2 = A_1 \sqcup (A_2 \setminus A_1)$

$P(A_1 \cup A_2) = P(A_1 \sqcup (A_2 \setminus A_1))$
 $= P(A_1) + P(A_2 \setminus A_1) \quad \left(\begin{array}{l} \text{finite} \\ \text{additivity of } P \end{array} \right) \quad \left(\begin{array}{l} \text{disjoint union} \\ (A_1 \cap (A_2 \setminus A_1) = \emptyset) \end{array} \right)$
 $\leq P(A_1) + P(A_2)$
(monotonicity as $A_2 \setminus A_1 \subseteq A_2$)

Assume finite subadd. holds for $n-1 \geq 2$.

Consider $A_1, \dots, A_n$. We've to prove fin. subadd for n .

$$A_1 \cup \dots \cup A_n = A_1 \cup (A_2 \cup \dots \cup A_n)$$

$$\begin{aligned} P(A_1 \cup \dots \cup A_n) &= P(A_1) + P((A_2 \cup \dots \cup A_n) / A_1) \quad (\text{fin. add.}) \\ &\leq P(A_1) + P(A_2 \cup \dots \cup A_n) \quad (\text{mon.}) \\ &\leq P(A_1) + P(A_2) + \dots + P(A_n) \quad (\text{fin. subadd. holds for } n-1) \end{aligned}$$

$$(v) \quad A \cup B = A \cup (B \setminus A)$$

$$\begin{aligned} P(A \cup B) &= P(A) + P(B \setminus A) \quad (\text{fin. add.}) \\ &= P(A) + P(B) - P(A \cap B) \quad (A \cap B \subseteq B \text{ \& use Lemma 4.4}) \end{aligned}$$

□

12/0 LECTURE 6 - MORE EXAMPLES.

Eg 6.1 (Sampling without replacement) $k \leq n$

Choose k objects from n objects without repetition/replacement done by one.

Ω = set of all ^{ordered} k -tuples from $[n]$ with no repetition

$$= \{ \omega = (w_1, \dots, w_k) \in [n]^k : w_i \neq w_j \ \forall i \neq j \}$$

$$|\Omega| = n(n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!}$$

w_i - type 1 if $w_i \leq m$, $1 \leq m \leq n$.

w_i - type 0 if $w_i > m$

A = $\exists$ are exactly l objects of type 1

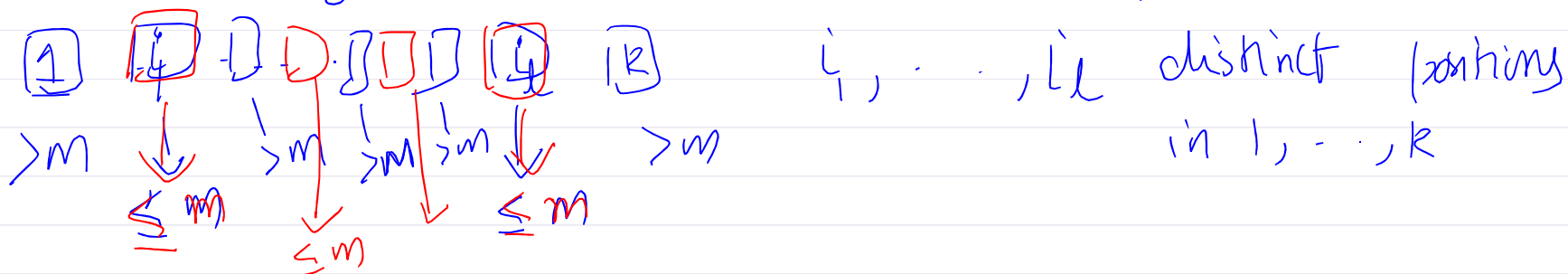
$$= \{ \omega \in \Omega : \# \{ i : w_i \leq m \} = l \}, \quad l \leq k, l \leq m$$

Assume k objects are chosen UAR. What is $P(A) = ?$

Concrete eg: $1, \dots, n$ are ^{prices} of objects in a shop.

k customers come & buy one object each UAR.

A = Exactly l customers bought objects of price $\leq m$.



There are l choices in $1, \dots, k$ for $w_i \leq m$ & remaining

$k-l$ choices, $w_i > m$. Choose $i_1, \dots, i_l \in [k]$ & distinct.

$$A(i_1, \dots, i_l) = \{ \underline{\omega} \in \Omega : w_{i_1}, \dots, w_{i_l} \leq m \text{ \& } w_j > m \ \forall j \neq i_1, \dots, i_l \}$$

$$A = \bigcup_{\substack{i_1, \dots, i_l \in [k] \\ \text{distinct}}} A(i_1, \dots, i_l)$$

$$P(A) = \sum_{\substack{i_1, \dots, i_k \\ \text{distinct}}} P(A(i_1, \dots, i_k)) \quad (\text{fin-add-})$$

$$= \sum_{\substack{i_1, \dots, i_k \\ \text{distinct}}} \frac{|A(i_1, \dots, i_k)|}{|\Omega|} \quad (\text{UAR})$$

$$|A(i_1, \dots, i_k)| = \underbrace{m(m-1) \dots (m-l+1)}_{\substack{\downarrow \\ \text{fill } l \text{ positions with} \\ 1, \dots, m \text{ without} \\ \text{repetition}}} \underbrace{(n-m)(n-m-1) \dots (n-m-(k-l)+1)}_{\substack{\downarrow \\ \text{fill } (k-l) \text{ positions} \\ \text{with } m+1, \dots, n \\ \text{without repetition}}}$$

$$\text{so } P(A) = \binom{k}{l} \frac{m!}{(m-l)!} \frac{(n-m)!}{(n-m-(k-l))!} / \frac{n!}{(n-k)!}$$

$$= \frac{\binom{m}{l} \binom{n-m}{k-l}}{\binom{n}{k}}$$

Ex 6.2 Choose $S \subseteq [n]$, $|S| = m$. $A_S = \{w \in \Omega : \# \{i : w_i \in S\} = l\}$
 Show that $P(A_S) = \frac{\binom{m}{l} \binom{n-m}{k-l}}{\binom{n}{k}}$ n, m, k, l same as in Ex 6.1

Ex 6.3 Assume k (unordered) objects are chosen UAR from n objects.
 Compute $P(A_S)$ for the same.

There are m objects of type 1 in n objects, k obj. are chosen uniformly at random from n objects in order.

$$P(\text{exactly } l \text{ objects of type 1 in } k \text{ objects}) = \frac{\binom{m}{l} \binom{n-m}{k-l}}{\binom{n}{k}}$$

Eg 6.4 A factory has m defective watches out of n .
 k watches are selected UAR & each watch is checked.
 l of them are found defective. Estimate m .

$$p(k, l) = P(\text{exactly } l \text{ defective watches in } k \text{ selected}) \stackrel{(\text{from Eg 6.1})}{=} \frac{\binom{m}{l} \binom{n-m}{k-l}}{\binom{n}{k}} \rightarrow$$

Estimate "statistically" $p(k, l)$ & then find m .

↓
 "Repeat experiment (select k UAR) & compute proportion of experiments with exactly l defective watches."

Eg 6.5: There are n fish in a lake (unknown)

Ecologists catch m fish & then mark them RED & put them back.
 After some time, you catch r fish & k red fish are observed.
 UAR

$$P(k \text{ red fish in } r \text{ fish}) = \frac{\binom{m}{k} \binom{n-m}{r-k}}{\binom{n}{r}}$$

- Technique for estimating population size 'n'.

Why not take $\left(\frac{k}{r} = \frac{m}{n} \right)$?

14/10

LECTURE 7 : MORE EXAMPLES — NON-UNIFORM PROB. SPACES.

Ex 7.1 : A standard deck of cards is "well-shuffled" (= cards are arranged UAR).
What is Prob. first 10 cards have a set of 4 aces?

Sample space Ω = possible arrangements of 52 cards

$$= \{ (a_1, \dots, a_{52}) : a_i \in [52], a_i \neq a_j \forall i \neq j \}$$

(a_i — the card in the i^{th} position i.e., i^{th} card)

(Sampling without replacement 52 objects from 52 objects)

$k=n$ in sampling without replacement is $\binom{n}{k} = \binom{n}{n}$
 $S_n :=$ set of all permutations on $[n] = \{1, \dots, n\}$

so above $\Omega = S_{52}$; $|\Omega| = 52!$ (from Ex 6.1)

A = first 10 cards have a set of 4 aces.

Total # of aces = 4. Let (Ω, P) be unif. prob. space.

$$P(A) = \frac{|A|}{|\Omega|} \quad (\text{UAR by assumption})$$

$$A = \{ (a_1, \dots, a_{52}) \in \Omega : 1, 14, 27, 40 \in \{a_1, \dots, a_{10}\} \}$$

(1, 13 — spade ace, 14 — clover ace, 27 — hearts ace, 40 — diamond ace)

$$|A| = \binom{10}{4} 4! 48!$$

of positions
for aces in first 10 slots

ordering of aces.

remaining 48 cards however
you like.

$$P(A) = \frac{10! 48!}{52! 6!} = \frac{10 \times 9 \times 8 \times 7}{52 \times 51 \times 50 \times 49} = \frac{9 \times 8}{52 \times 51 \times 35}$$

of we ask prob. of 4 aces in first 20 positions?

$$P(A) = \frac{\binom{20}{4}}{\binom{52}{4}}$$

AWAY FROM UNIFORM PROB. SPACES

Ex 2: Take simplest, $\Omega = \{0, 1\}$.

A pmf p on Ω is $p(0), p(1) \geq 0$ & $p(0) + p(1) = 1$

Set $p(1) = p \in [0, 1]$ (abuse of notation = $p: \Omega \rightarrow [0, 1]$
 $p \in [0, 1]$)

Then $p(0) = 1 - p$.

The PD P is called Bernoulli (p) ^{prob.} distribution.

Ex. let $(\Omega_i, p_i), i=1, \dots, k$ be finite prob. spaces.

A 2.7 $\Omega = \prod_{i=1}^k \Omega_i$ is the cartesian product.

Define $p((w_1, \dots, w_k)) = \prod_{i=1}^k p_i(w_i) \quad \forall (w_1, \dots, w_k) \in \Omega$.

Show that (Ω, p) is a finite prob. space

$$\& P(A) = \prod_{i=1}^k P_i(A_i) \quad \forall A = A_1 \times \dots \times A_k \subseteq \Omega.$$

$$A = \{(w_1, \dots, w_k) : w_i \in A_i \forall i\}$$

"Product of prob. spaces with product. pmf is a prob. space"

Ex 7.3

$\Omega = \{0, 1\}^n$; Consider $(\{0, 1\}, P_i)$ to be $\text{Ber}(p)$ prob. distribn

$$p((w_1, \dots, w_n)) = \prod_{i=1}^n p_i(w_i) = \prod_{i=1}^n p^{w_i} (1-p)^{1-w_i} \quad \forall i=1, \dots, n$$

$$\left[p_i(w_i) = p^{w_i} (1-p)^{1-w_i} = \begin{cases} p & \text{if } w_i = 1 \\ 1-p & \text{if } w_i = 0 \end{cases} \right] \left[\begin{matrix} (\Omega, p) \\ \parallel \\ (\Omega, P) \end{matrix} \right]$$

$$\rightarrow = p^{\sum_{i=1}^n w_i} (1-p)^{n - \sum_{i=1}^n w_i}$$

[pmf on Ω depends only on $\sum_{i=1}^n w_i$]

[product $\text{Ber}(p)$ distribution on $\{0, 1\}^n$]

if $p = 1/2$, this uniform PD on $\{0, 1\}^n$ i.e., $p((w_1, \dots, w_n)) = \frac{1}{2^n}$.

Ex 7.4 $\Omega = \{0, \dots, n\}$. Define $P(k) := \binom{n}{k} p^k (1-p)^{n-k}$, $p \in [0, 1]$.
 $\parallel$
 $P(k)$

P is called the Binomial (n, p) distribution.

$$\text{check } P(k) \geq 0, \quad \sum_{k=0}^n P(k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p+1-p)^n = 1.$$

Eg 7.5:

Let P be product $\text{Ber}(p)$ distribution on $\{0, 1\}^n$.

Define $f: \{0, 1\}^n \rightarrow \{0, \dots, n\}$ as $f(w_1, \dots, w_n) = \sum_{i=1}^n w_i$.

$$A_k = \{(w_1, \dots, w_n) : f(w_1, \dots, w_n) = k\} \quad 0 \leq k \leq n.$$

$$\subseteq \{0, 1\}^n.$$

$$P(A_k) = \sum_{w \in A_k} P(w)$$

$$(P(w) = P(\{w\}) = p(w))$$

$$= \sum_{w \in A_k} p^{\sum w_i} (1-p)^{n-\sum w_i}$$

(by defn of P in Eg 7.3)

$$= p^k (1-p)^{n-k} |A_k| \quad (\sum w_i = k \quad \forall w \in A_k)$$

$$\Omega = \bigsqcup_{k=0}^n A_k \quad \text{as } f(w) \in \{0, \dots, n\}$$

$$= \binom{n}{k} p^k (1-p)^{n-k} \quad (|A_k| = \binom{n}{k} \text{ Exactly } k \text{ 1's in } n \text{ slots})$$

$$1 = P(\Omega) = \sum_{k=0}^n P(A_k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k}.$$

(fin. add.)

Eg 7.6

Consider $\Omega = \{0, 1\}^n$ as above with product $\text{Ber}(p)$ distribn.

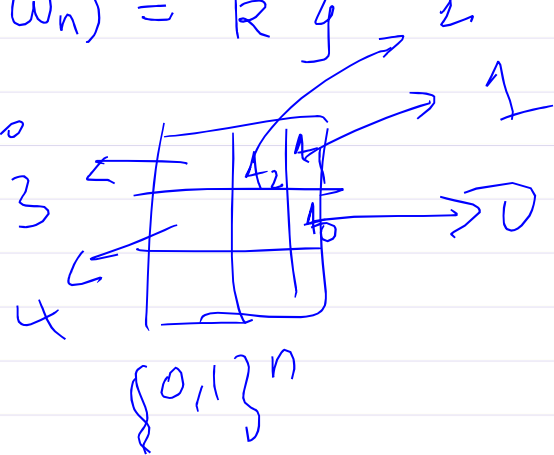
Let $f: \{0, 1\}^n \rightarrow \{0, \dots, n\}$ be a function.

$0 \leq k \leq n$. Define $A_k = \{(w_1, \dots, w_n) : f(w_1, \dots, w_n) = k\}$.

A_k 's are pairwise disjoint for $k=0, \dots, n$.

$$\Omega = \bigsqcup_{k=0}^n A_k$$

$$\text{So } 1 = P(\Omega) = \sum_{k=0}^n P(A_k) \quad (\text{fin. add.})$$



Suppose we define $\tilde{p}: \{0, \dots, n\} \Rightarrow [0, 1]$
as $\tilde{p}(k) = P(A_k)$.

Check $\tilde{p}$ is a pmf on $\{0, \dots, n\}$. Why? $\tilde{p}(k) \geq 0 \forall k$
so $(\{0, \dots, n\}, \tilde{p})$ is a finite prob. space.
 $\sum_{k=0}^n \tilde{p}(k) = 1$.
 $\Rightarrow \tilde{P}(A) = \sum_{k \in A} \tilde{p}(k) \quad \forall A \subseteq \{0, \dots, n\}$.

So using (Ω, P) & $f: \Omega \rightarrow \{0, \dots, n\}$
we have constructed a new prob. space $(\{0, \dots, n\}, \tilde{p})$.
sample space.

LEMMA 7.7 (Induced prob. lemma).

Let (Ω, P) be a finite prob. space & $\tilde{\Omega}$ is a finite set.
Let $f: \Omega \rightarrow \tilde{\Omega}$ be a function.

Define $\tilde{p}(\tilde{\omega}) := P(\{\omega \in \Omega : f(\omega) = \tilde{\omega}\})$, $\tilde{\omega} \in \tilde{\Omega}$.
 $= P(f^{-1}(\tilde{\omega}))$

$$[f^{-1}(\tilde{\omega}) = \{\omega \in \Omega : f(\omega) = \tilde{\omega}\}]$$

Then we have that $(\tilde{\Omega}, \tilde{p})$ is a fin. prob. space.

Proof: To show $(\tilde{\Omega}, \tilde{p})$ is a fin. prob. space
we have to show $\tilde{p}$ is a pmf on $\tilde{\Omega}$.
i.e., $\tilde{p}(\tilde{\omega}) \geq 0$, $\sum_{\tilde{\omega} \in \tilde{\Omega}} \tilde{p}(\tilde{\omega}) = 1$.

EXTRA: Given $a_0, \dots, a_n \geq 0$. To check $\sum_{i=0}^n a_i = 1$

Construct (Ω, P) & events $A_0, \dots, A_n$
 $\Rightarrow \Omega = \bigsqcup_{k=0}^n A_k$ & $a_k = P(A_k) \quad \forall k=0, \dots, n$.

Qn: Given $a_0, \dots, a_n \geq 0$ & $\sum_{i=0}^n a_i = 1$.

$\Omega = \{0,1\}^n$ & P - product $\text{Ber}(p)$ distribution.

~~Can~~ Can we always find p & $A_0, \dots, A_n$ s.t.
 $a_k = P(A_k)$?

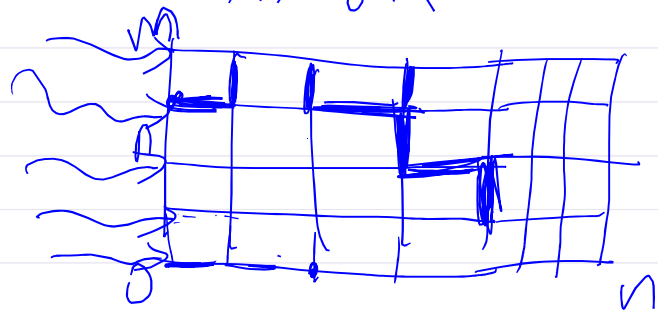
Try: $p = 1/2 \Rightarrow P$ - Uniform PD on Ω .

$$P(A_k) = \frac{|A_k|}{2^n}.$$

Can I find A_k 's s.t. $a_k = \frac{|A_k|}{2^n}$? True if $2^n a_k \in \mathbb{N} \cup \{0\}$.

[OPEN PROBLEM] Take a $n \times n$ grid

[Percolation theory]



$$= \left\{ (i,j) : \begin{array}{l} 0 \leq i \leq n \\ 0 \leq j \leq m \end{array} \right\}$$

E = set of pipes = $\left\{ (i,j) \rightarrow (i,j+1), (i,j) \rightarrow (i+1,j) \right\}$

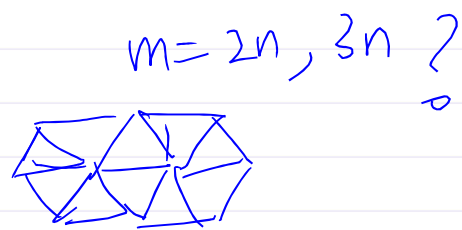
$\Omega = 2^E$ = collection of subset of pipes. $i \neq j$

$$P(\omega) = \frac{1}{2^{|E|}}, \quad \omega \subseteq E \quad N = |E| = n(m+1) + m(n+1)$$

prob. ω is not destroyed

$P(\text{Water flows from Left to right}) = ??$

S. Smirnov solved it in Triangular lattice



26/10

LECTURE 8.

Recall $(\Omega, \mathcal{F})$ on $(\Omega, \mathcal{P})$ is a finite prob. space.

$\downarrow$ pmf $\downarrow$ PD

obvious example $p(\omega) = \frac{1}{|\Omega|} \quad \omega \in \Omega$ (Eq. likely outcomes / UAR).

Job: Find $A \subseteq \Omega$ & compute $P(A)$.

Eg. $\text{Ber}(p)$ $p(0) = 1 - p = 1 - p(1)$ $\Omega = \{0, 1\}$. [coin toss head prob. p]

Eg. $\text{Bin}(n, p)$ $\Omega = \{0, \dots, n\}$ $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$ [# of heads in n biased coin tosses]

Eg 8.1: $\text{Hyg}(\underline{n}, \underline{m}, \underline{r})$ $\underline{\Omega} = [n] \cup \{0\}$; $r \leq m \leq n$.

(A3) (Hypergeometric)
 $k \in \Omega$.

$$p(k) = \begin{cases} \frac{\binom{m}{k} \binom{n-m}{r-k}}{\binom{n}{r}} & k \leq r \text{ \& } k \leq m \\ 0 & \text{else.} \end{cases}$$

Context: Choose r objects out of n . There are ^{exactly} m 'good' ones
 [see computation in sampling without replacement] Prob. (one has chosen k good objects) $= p(k)$

Show (Ω, p) is a prob. space.

To show (Ω, p) (in general) is a prob. space

— show p is a pmf

— show P is a PD.

— use induced prob. lemma.

— use "product" of prob. spaces [A 2.7]

— use "projection" $(\omega_1, \omega_2) \mapsto \omega_1$ [A 2.8]

$[\Omega_1 \times \Omega_2 \xrightarrow{\uparrow} \Omega_1]$
 projection map

$\downarrow$
 [check: std. case of Ind. prob. lemma]

For eg. $\text{Bin}(n, p)$ — verify $\text{Ber}(p)$ is a PD/pmf

induced prob. lemma. $\nwarrow$ construct $\Downarrow$ product space

Also directly check $\text{Bin}(n, p)$ is a PD.

Eg 8.2 (Maxwell-Boltzmann distribution)

(A3)

$$\Omega_{r,n} = \{ (r_1, \dots, r_n) : r_i \geq 0, \sum_{i=1}^n r_i = r \}$$

r unlabelled particles into n labelled cells

Define $p(w) := \frac{r!}{r_1! \dots r_n!} \frac{1}{n^r}$ $w = (r_1, \dots, r_n)$

show $(\Omega_{r,n}, p)$ is a PS.

COMPARISON:

BE $p(w) = \frac{1}{|\Omega_{r,n}|}$ $w \in \Omega_{r,n}$.

FD $p(w) = \frac{1}{|\Omega_{r,n}^*|}$ $w \in \Omega_{r,n}^* (0 \text{ else})$

If you solve A3, you will understand why MB came first & very natural mathematically.

Photons "follow" BE & Electrons "follow" FD.

Eg 8.3 [Boltzmann-Gibbs distribution] Ω - finite set.

Let $H: \Omega \rightarrow \mathbb{R}$ be a function. $\beta \geq 0$

Define $Z_\beta := \sum_{\omega \in \Omega} e^{-\beta H(\omega)}$

Define $p(\omega) := \frac{e^{-\beta H(\omega)}}{Z_\beta} \quad \omega \in \Omega.$

Check (Ω, p) is a PS. ($p(\omega) \geq 0$ & $\sum_{\omega \in \Omega} p(\omega) = 1$)

Interpretation: ω - a state

$H(\omega)$ - energy of the state.

$$\beta = \frac{1}{kT} \quad \begin{array}{l} T - \text{temp} \\ T \geq 0 \end{array}$$

low temp $T=0 \iff \beta = \infty \rightarrow$ system VAR on minimal energy state

high temp $T=\infty \iff \beta = 0 \rightarrow$ system chooses a state VAR.

If $\beta = 0$ then $Z_\beta = |\Omega|$ & $p(\omega) = \frac{1}{|\Omega|}$ (UAR)

Show
(Analysis)
ex:

$$\lim_{\beta \rightarrow \infty} p(w) = \frac{\mathbb{1}[w \in \Omega_*]}{|\Omega_*|} \quad \text{where} \quad \Omega_* := \left\{ w \in \Omega : H(w) = \min_{s \in \Omega} H(s) \right\}$$

$$\mathbb{1}[w \in \Omega_*] = \begin{cases} 1 & \text{if } w \in \Omega_* \\ 0 & \text{else} \end{cases}$$

Indicator function

At any temperature $\beta \in (0, \infty)$, system likes to be in minimal energy states but may not be.

Penalise a non min-energy state at small temperatures (ie, T close to 0)

Ex. Let (Ω, p) be a PS $\exists p(w) > 0 \quad \forall w \in \Omega$.

8.4 Show that $\exists H: \Omega \rightarrow \mathbb{R} \quad \exists \beta \geq 0 \quad p(w) = \frac{e^{-\beta H(w)}}{Z_\beta}$.

[Giorgio Parisi Nobel (2021) in Physics for spin systems.]

$\Omega = \{-1, +1\}^n \quad H: \Omega \rightarrow \mathbb{R} \quad \text{is random}$

Qn: How do compute $P(A)$ given $(\Omega, \mathcal{P})$?

— we know some examples.

— what are our tools?

Recall

THM 4.1 Let $(\Omega, \mathcal{P})$ be a fin. prob. space. Then

(i) $P(A) \leq P(B) \quad \forall A \subseteq B \subseteq \Omega$.
the following hold.

[Monotonicity]

(i') $P(A) \leq 1 \quad \forall A \subseteq \Omega$

(ii') $P(\emptyset) = 0$

(iv) $P(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^n P(A_i) \quad \forall A_1, \dots, A_n \subseteq \Omega$. [finite subadditivity]

$$(v) \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

+ finite additivity.

[Inclusion-Exclusion]

THM 8.9: Let (Ω, P) be prob. space. Let $A_1, \dots, A_n$ be events.

[I-E formula]

$$P\left(\bigcup_{i=1}^n A_i\right) = S_1 - S_2 + S_3 - \dots + (-1)^{n+1} S_n$$

$$S_k \equiv \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k})$$

Proof:

$$S_1 = \sum_{i=1}^n P(A_i) \quad S_2 = \sum_{1 \leq i < j \leq n} P(A_i \cap A_j)$$

$$S_n = P(A_1 \cap \dots \cap A_n)$$

we'll prove - by induction.

I-E True for $n=1, 2$.

Assume I-E holds upto $(n-1)$ for $n \geq 3$.

We'll prove for n .

Write $B = A_2 \cup \dots \cup A_n$.

$$P\left(\bigcup_{i=1}^n A_i\right) = P(A_1 \cup B) = P(A_1) + P(B) - P(A_1 \cap B) \quad (1)$$

$$P(B) = P(A_2 \cup \dots \cup A_n)$$

$$= S_1' - S_2' + S_3' - \dots + (-1)^{n-2} S_{n-1}' \quad (\text{I-E for } n-1) \quad (2)$$

$$S_k' = \sum_{2 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k})$$

$$A_1 \cap B = \bigcup_{i=2}^n (A_1 \cap A_i)$$

Use I-E on $n-1$ on $A_1 \cap B$

$$P(A_1 \cap B) = \tilde{S}_1 - \tilde{S}_2 + \dots + (-1)^{n-2} \tilde{S}_{n-1} \quad (3)$$

$$\tilde{S}_k = \sum_{2 \leq i_1 < \dots < i_k \leq n} P(A_1 \cap A_{i_1} \cap \dots \cap A_{i_k})$$

From (1), (2) & (3)

$$P\left(\bigcup_{i=1}^n A_i\right) = P(A_1) + S_1' - S_2' + S_3' - \dots + (-1)^{n-2} S_{n-1}' + 0$$

$\swarrow \quad \swarrow \quad \swarrow \quad \swarrow$
 $0 - \tilde{S}_1 + \tilde{S}_2 - \tilde{S}_3 + \dots + (-1)^{n-1} \tilde{S}_{n-1}$

check

$$P(A_1) + S_1 = P(A_1) + \sum_{i=2}^n P(A_i^c) = S_1$$

$$-S_2 - \tilde{S}_1 = -S_2$$

$$\vdots$$

$$(-1)^{n-2} S_{n-1} + (-1)^{n-2} \tilde{S}_{n-2} = S_{n-1} (-1)^{n-2}$$

$$(-1)^{n+1} \tilde{S}_{n-1} = (-1)^{n+1} S_n$$

$$[\tilde{S}_{n+1} = S_n]$$

$$S_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k})$$

$$= \sum_{2 \leq i_2 < i_3 < \dots < i_k \leq n} P(A_1 \cap A_{i_2} \cap \dots \cap A_{i_k}) + \sum_{2 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k})$$

(Fix $i_1 = 1$) (Take $i_1 > 1$)

(by
sym)

$$= \tilde{S}_{k-1} + S_k$$

$$(-1)^{k+1} S_k = (-1)^{k+1} S_k - (-1)^{k-2} \tilde{S}_{k-1}$$

□

Eg:
 $n=3$

$$P(A \cup B \cup C) = P(A) + P(B \cup C) - P(A \cap (B \cup C)) \quad (\text{IE for } n=2)$$

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

$$= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C) \dots$$

Exo (Bonferroni Inequalities). Assume as in I-E.

then

$$P\left(\bigcup_{i=1}^n A_i^c\right) \leq \sum_{k=1}^m (-1)^{k+1} S_k \quad \forall \text{ odd } m.$$

$$P\left(\bigcap_{i=1}^n A_i^c\right) \geq \sum_{k=1}^m (-1)^{k+1} S_k \quad \forall \text{ even } m.$$

28/10 LECTURE 9: Applications of I-E & Bonferroni

THM 8.5: Let (Ω, P) be prob. space. Let $A_1, \dots, A_n$ be events.

[I-E formula]

$$P\left(\bigcup_{i=1}^n A_i\right) = S_1 - S_2 + S_3 - \dots + (-1)^{n+1} S_n$$

$$S_k \equiv \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k})$$

Eg 9.1: r labelled balls are thrown randomly into n labelled cells.

$$p_m(r, n) = P(\text{exactly } m \text{ cells are empty}), \quad m = 0, \dots, n-1$$

$$\Omega = \{(\omega_1, \dots, \omega_r) : \omega_i \in [n]\}, \quad |\Omega| = n^r$$

↳ cell of first ball.

$$p(\omega) = \frac{1}{n^r} \quad (\text{eq. likely})$$

Take $m = 0$: $A_j = j^{\text{th}} \text{ cell is empty} = \{\omega : \omega_i \neq j \text{ for } i = 1, \dots, r\}$

$$\text{no cell is empty} = \bigcap_{j=1}^n A_j^c = \left(\bigcup_{j=1}^n A_j\right)^c$$

$$\begin{aligned}
 P_0(r, n) &= \text{IP}(\text{no cell is empty}) = \text{IP}\left(\bigcap_{j=1}^n A_j^c\right) \\
 &= \text{IP}\left(\left(\bigcup_{j=1}^n A_j\right)^c\right) = 1 - \text{P}\left(\bigcup_{j=1}^n A_j\right) \cdot \left(\text{IP}(A) + \text{IP}(A^c) = 1\right) \quad \text{--- (1)}
 \end{aligned}$$

To compute $\text{IP}\left(\bigcup_{j=1}^n A_j\right)$ we'll use I.F.O.

$$\text{IP}(A_1 \cap \dots \cap A_k) = \frac{|A_1 \cap \dots \cap A_k|}{n^r} \quad (\text{UAR})$$

$$|A_1 \cap \dots \cap A_k| = \left| \left\{ \omega : \omega_i \neq 1, \dots, k \quad \forall i=1, \dots, r \right\} \right|$$

$$= (n-k)^r$$

$$\text{IP}(A_1 \cap \dots \cap A_k) = \left(1 - \frac{k}{n}\right)^r$$

Same way by def of S_k , $S_k = \binom{n}{k} \left(1 - \frac{k}{n}\right)^r$ — (2)

$1 \leq i_1 < \dots < i_k \leq n.$

$$P(A_1 \cup \dots \cup A_n) \stackrel{(IE)}{=} \sum_{k=1}^n (-1)^{k-1} S_k$$

$$= \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^r \quad (\text{using (2)})$$

$$p_0(r, n) = 1 - \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^r \quad (\text{using (1)})$$

$$|\{w: \text{no cell is empty}\}| = n^r p_0(r, n) \quad (\text{by defn of } p_0 \text{ \& UAR})$$

$m \geq 1$

$$p_m(r, n) = P(\text{exactly } m \text{ cells are empty})$$

Exactly m cells are empty

→ Choose exactly m cells to be empty. $\leftarrow \binom{n}{m}$ ways.

→ r balls are to be thrown into remaining $n-m$ cells such that none of them are empty. $\leftarrow (n-m)^r p_0(r, n-m)$ ways.

$$p_m(r, n) = \frac{\binom{n}{m} (n-m)^r p_0(r, n-m)}{n^r}$$

$$= \binom{n}{m} \sum_{k=0}^{n-m} (-1)^k \frac{(n-m)^r}{n^r} \binom{n-m}{k} \left(1 - \frac{k}{n-m}\right)^r \quad \left(\begin{array}{l} \text{substitute } p_0(r, n-m) \\ \text{value} \end{array} \right)$$

$$= \binom{n}{m} \sum_{k=0}^{n-m} (-1)^k \binom{n-m}{k} \left(1 - \frac{m+k}{n}\right)^r$$

Ex* Fix m
 Find $\lim_{n \rightarrow \infty} p_m(r, n) = ?$
 Q.2

Case 1: Fix r : Case 2: let $r \rightarrow \infty \Rightarrow \frac{r}{n} \rightarrow 0$

Case 3: let $r \rightarrow \infty$ & $ne^{-r/n} = \lambda \in (0, \infty)$
 fixed

$n=365$. $p_m(r, n) = P(\text{Exactly } m \text{ days have no birthdays among } r \text{ people})$

THM 9.3 (Bonferroni's inequalities) Let $A_1, \dots, A_n$ be events in a prob. space $(\Omega, \mathcal{P})$ & $A = \bigcup_{i=1}^n A_i$. Then

$$P(A) \leq \sum_{k=1}^m (-1)^{k-1} S_k \quad \text{for odd } m$$

$$P(A) \geq \sum_{k=1}^m (-1)^{k-1} S_k \quad \text{for even } m.$$

Eg 9.4 (Contd. of Eg 9.2) Find bounds $P(\text{at least one cell is empty})$
 $A_j = j^{\text{th}} \text{ cell is empty.}$

$$A = \bigcup_{j=1}^n A_j = \text{at least one cell is empty.}$$

From Bonferroni's ineq.

$$S_1 - S_2 \leq P(A) \leq S_1$$

$$S_1 = n \left(1 - \frac{1}{n}\right)^r \quad S_2 = \binom{n}{2} \left(1 - \frac{2}{n}\right)^r$$

$$n=10; \quad r=40;$$

$$0.1418 \leq P_1(r, n) \leq 0.1478$$

Ex* Find out for what range of r, n $S_1 - S_2 \approx S_1$ are close

Proof of THM 9.3: (sketch)

We'll prove for $m=1, 2$ & rest is exercise.

For $m=1$ $P(A) \leq S_1$ (finite subadditivity / Union bound)

$$m=2 \therefore P(A) = \sum_{w \in A} p(w) \quad (\text{defn})$$

$$= \sum_{w \in \Omega} p(w) \mathbb{1}_A(w)$$

$$\left[\begin{array}{l} \mathbb{1}_A: \Omega \rightarrow \{0,1\} \\ w \mapsto \begin{cases} 1 & \text{if } w \in A \\ 0 & \text{else} \end{cases} \end{array} \right]$$

Indicator.

Monotonicity

Suppose we show that $\mathbb{1}_A(w) \geq f(w) \forall w \in \Omega$

for some function $f: \Omega \rightarrow \mathbb{R}$

then since $p(w) \geq 0 \forall w \in \Omega$, we have $\sum_{w \in \Omega} p(w) \mathbb{1}_A(w) \geq \sum_{w \in \Omega} f(w) p(w)$

Assume $\mathbb{1}_A(\omega) \geq \sum_{i=1}^n \mathbb{1}_{A_i^c}(\omega) - \sum_{1 \leq i_1 < i_2 \leq n} \mathbb{1}_{A_{i_1}}(\omega) \mathbb{1}_{A_{i_2}}(\omega) \quad \text{--- (4)}$

Note that $\mathbb{1}_{A_{i_1}}(\omega) \mathbb{1}_{A_{i_2}}(\omega) = \mathbb{1}_{A_{i_1} \cap A_{i_2}}(\omega) \quad \text{--- } f(\omega)$

If (4) holds, from monotonicity we have that

$$P(A) = \sum_{\omega \in \Omega} p(\omega) \mathbb{1}_A(\omega)$$

$$\geq \sum_{\omega \in \Omega} p(\omega) \left(\sum_{i=1}^n \mathbb{1}_{A_i^c}(\omega) - \sum_{1 \leq i_1 < i_2 \leq n} \mathbb{1}_{A_{i_1} \cap A_{i_2}}(\omega) \right)$$

$$= \sum_{\omega \in \Omega} p(\omega) \sum_{i=1}^n \mathbb{1}_{A_i^c}(\omega) - \sum_{\omega \in \Omega} p(\omega) \sum_{1 \leq i_1 < i_2 \leq n} \mathbb{1}_{A_{i_1} \cap A_{i_2}}(\omega)$$

(interchange sums)

$$= \sum_{i=1}^n \sum_{\omega \in \Omega} \mathbb{1}_{A_i^c}(\omega) p(\omega) - \sum_{1 \leq i_1 < i_2 \leq n} \sum_{\omega \in \Omega} p(\omega) \mathbb{1}_{A_{i_1} \cap A_{i_2}}(\omega)$$

↪ (justified as we are dealing with finitely many terms)

$$= \sum_{i=1}^n P(A_i^c) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1}^c \cap A_{i_2}^c)$$

$$= S_1 - S_2.$$

Proof for $n=2$ is complete assuming (4)

Let us prove (4).

Case 1: $\omega \notin A$. $\mathbb{1}_A(\omega) = 0$; $\mathbb{1}_{A_i^c}(\omega) = 0$ as $A_i^c \subseteq A$.

Case 2: $\omega \in A$. Assume ω is in exactly r many A_i^c 's for some $r \geq 1$.

Without loss of generality (WLOG), assume $\omega \in A_1^c \cap \dots \cap A_r^c$
& $\omega \notin A_i^c \forall i > r$.

$$\mathbb{1}_A(\omega) = 1; \quad \sum_{i=1}^n \mathbb{1}_{A_i^c}(\omega) = r, \quad \sum_{1 \leq i_1 < i_2 \leq n} \mathbb{1}_{A_{i_1}^c \cap A_{i_2}^c}(\omega) = \binom{r}{2}$$

$$\text{So } f(\omega) = r - \binom{r}{2} = r \left(1 - \frac{r-1}{2}\right) \leq 1 \text{ if } r \geq 1$$

$$\text{So } 1_A(w) \geq f(w) \quad \forall w \in A.$$

Case 1 + Case 2 $\Rightarrow$ (4) is true.

So proof is complete for $m = 2$.

General m — Ex. (A4). (P)

Ex. (A4) Prove I-E using above approach.

See Feller (Ch 5) & Venkatesh (Ch IV) for more applications.

$$\text{Ex. (A4)} \quad P(\text{exactly } m \text{ of } A_1, \dots, A_m \text{ occur}) = \sum_{k=0}^{n-m} (-1)^k \binom{m+k}{m} S_{m+k}$$

$$\text{Ex. (A4)} \quad P(\text{at least } m \text{ of } A_1, \dots, A_n \text{ occur}) = \dots$$

2/11 LECTURE 10 - CONDITIONAL PROBABILITY.

1980/81 Wimbledon final - Prob. (Björk winning) $\approx \frac{2}{3}$.
"Survey"

Prob (Björk loses 1st set & loses the match) $\approx \frac{1}{2}$.

> Prob (loses the match) = $\frac{1}{3}$.

Wrong! Most people thought

"Prob. Björk loses the match given he loses the first set"?

Prob. of Kohli winning the toss against Afghanistan given he lost both the previous tosses. = ??

$\Omega = \{0, 1\}^3$
↑ ↑
lose win
the toss the toss

$A = \{(0, 0, c) : c \in \{0, 1\}\}$ — lost first two tosses.

$B = \{(a, b, 1) : a, b \in \{0, 1\}\}$ — win 3rd toss.

Assume eq. likely outcomes. — $P(B) = \frac{1}{2}$; $P(A) = \frac{1}{4}$.

$$P(B \text{ given } A) =: P(B|A)$$

Since A is given, not all outcomes in Ω are valid anymore.

So "new sample space is A "! May be $\underline{P(B|A)} \neq P(B \cap A)$

2 problems — $P(B|A) \nmid B \subseteq A$ is not a prob. dist. on $A \rightarrow$

$$(P(A|A) = P(A) \leq 1)$$

— $P(B|A) < P(A), P(B)$ — doesn't go with heuristic understanding.

DEFN 10.1 | (Ω, P) be a P-space. Let $B \subseteq \Omega$ be $\exists P(B) > 0$.

Then $P(A|B) := \frac{P(A \cap B)}{P(B)}$ — conditional prob. of A given B
or
prob of A given B .

prob. of A conditioned on B .

Ex. 10.2: Consider eg. of Kohli winning the toss.

[Gambler's fallacy]

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{|A \cap B| / |\Omega|}{|A| / |\Omega|} = \frac{1}{2}$$

win 3rd toss lose first two tosses

(Prod. of Bernoulli PD)

Ex*: $\Omega = \{0,1\}^5$ $P_p(\omega) = \prod_{i=1}^5 p^{w_i} (1-p)^{1-w_i}$ $\omega = (w_1, \dots, w_5) \in \Omega$

10.3 $A_i = \{\omega : w_i = 1\}$ $\leftarrow p \in [0,1]$

check: $P_p(A_i) = p$ (prob. of being winning a set.)

$B =$ winning the match. (wins at least 3 &ts)

$P_p(B^c | A_i^c) > P_p(B^c)$ possible for some p ?

$$[P_p(B^c \cap A_i^c) \leq P_p(B^c) \quad \forall p]$$

Ex 10.4: Let $B \subseteq \Omega$ & $P(B) > 0$. Define $P_B(A) := P(A|B) \quad \forall A \subseteq \Omega$.

(A) Show that (B, P_B) is a P.Space i.e.,

(i) $P_B(A) \geq 0 \quad \forall A \subseteq B$ (ii) $P_B(B) = 1$ (iii) $P_B(A \cup C) = P_B(A) + P_B(C)$
(i.e., $A \cap C = \emptyset$).

Ex.
10.5

(Ω, P_B) is a P. space where $P_B(A) = P(A|B) \quad \forall A \subseteq \Omega$.

Ex
10.6

Let (Ω, P) be P.S. with eq. likely outcomes in Ω .

i.e., $P(A) = \frac{|A|}{|\Omega|}$

Assume $|B| > 0$ i.e., $P(B) > 0$.

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

$$\Rightarrow P_B(A) = \frac{|A|}{|B|} \quad \text{if } A \subseteq B.$$

$\Rightarrow (B, P_B)$ is a P.S. with eq. likely outcomes in B .

Eg
107

A house with 2 children is chosen at random & you are told that there is a boy. What is the prob. that the other child is a boy?

Soln 1: $\Omega = \{bb, bg, gb, gg\}$; $p(w) = \frac{1}{4} \quad \forall w \in \Omega$.

$$B = \{\exists \text{ a boy}\} = \{bb, bg, gb\} \quad P(B) = \frac{3}{4}$$

$$A = \{\text{both are boys}\} = \{bb\}$$

$$P(A|B) = \frac{|A|}{|B|} = \frac{1}{3}.$$

Soln 2:

$$\Omega = \{bb, \underset{ob}{bg}, gg\} \quad p(w) = \frac{1}{3} \quad \forall w \in \Omega.$$

$$B = \{bb, bg\} \quad A = \{bb\} \quad P(A|B) = \frac{1}{2}.$$

Soln 3:

$$\Omega = \{bb, bg, gg\} \quad p(bb) = p(gg) = \frac{1}{4}; \quad p(bg) = \frac{1}{2}.$$

$$P(B) = \frac{3}{4} ; \quad P(A \cap B) = \frac{1}{4}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{3} \quad (\text{Same as first soln})$$

$$p(gg) = p, \quad p(bb) = q, \quad p+q \leq 1. \quad \text{Solve the problem.}$$

Ex*

10.8

Ex

10.9

Same as above. You are told the eldest child is a boy.

What is Prob. both are boys?

$$\Omega = \{bb, bg, gb, gg\} \quad p(w) = \frac{1}{4} \quad \forall w \in \Omega.$$

$$B = \{bb, bg\} \quad \text{eldest child is a boy.}$$

$$A = \{bb\} \quad \text{both are boys.} \quad P(A|B) = \frac{1}{2}.$$

Ex

10.10

Same as above. You call a house with 2 children, a boy picks the phone & puts it down.
Prob. both are boys?

$$\Omega = \underbrace{\{bb, gb, bg, gg\}}_{\text{children in house}} \times \underbrace{\{b, g\}}_{\text{who picks up the phone}}$$

$$P((gg, b)) = 0 = P((bb, g))$$

$$P((gg, g)) = \frac{1}{4} = P((bb, b))$$

$$P((gb, *)) = \frac{1}{8} = P((bg, *)) \quad * \in \{b, g\}$$

B = two children & boy picks up the phone

$$= \{(bb, b), (gb, b), (bg, b)\} \quad P(B) = \frac{1}{2}$$

A = both are boys = $\{(bb, b)\}$

$$P(A \cap B) = P(A) = \frac{1}{4} \quad P(A|B) = \frac{1}{2} \quad \square$$

Thm: Let $A_1, \dots, A_n$ be events in Ω s.t. $P(A_1 \cap \dots \cap A_n) > 0$.
10.11 Then $P(A_1 \cap \dots \cap A_n) = P(A_1 | A_2 \cap \dots \cap A_n) \times P(A_2 | A_3 \cap \dots \cap A_n)$
 $\times \dots \times P(A_{n-1} | A_n) \times P(A_n)$
 (chain rule for condnl. prob.)

Proof: By induction. For $n=1$ $P(A_1) = P(A_1)$
 For $n=2$ $P(A_1 \cap A_2) = P(A_1 | A_2) \times P(A_2)$
 (defn of $P(A_1 | A_2)$)

Suppose holds for $n-1$.

Since $P(A_1 \cap \dots \cap A_n) > 0$, all condnl. prob. are well-defined.

$$B = A_2 \cap \dots \cap A_n$$

$$P(A_1 \cap \dots \cap A_n) = P(A_1 \cap B) \stackrel{\text{by defn}}{=} P(A_1 | B) P(B)$$

$$\text{(by indn)} \quad = P(A_1 | B) P(A_2 | A_3 \cap \dots \cap A_n) \times \dots \times P(A_{n-1} | A_n) \times P(A_n)$$

THM 10.12. Let $\Omega = \bigcup_{i=1}^n A_i$ & $P(A_i) > 0 \quad \forall i$.
Then $\forall B \subseteq \Omega$ $P(B) = \sum_{i=1}^n P(B|A_i) P(A_i)$. [Law of Total Prob.]

Try Ex 9.

LECTURE 11: 09/11 - Law of total prob., Bayes Formula & applications.

THM
10.11 Let $A_1, \dots, A_n$ be events in Ω s.t. $P(A_1 \cap \dots \cap A_n) > 0$.
Then $P(A_1 \cap \dots \cap A_n) = P(A_1 | A_2 \cap \dots \cap A_n) \times P(A_2 | A_3 \cap \dots \cap A_n)$
 $\times \dots \times P(A_{n-1} | A_n) \times P(A_n)$.
(chain rule for condn. prob.)

THM
10.12 Let $\Omega = \bigsqcup_{i=1}^n A_i$ & $P(A_i) > 0 \quad \forall i$.
Then $\forall B \subseteq \Omega \quad P(B) = \sum_{i=1}^n P(B | A_i) P(A_i)$. [Law of Total prob.]

Proof: $n=2$. Assume $\Omega = A_1 \sqcup A_2$.
$$P(B) = P(B \cap A_1) + P(B \cap A_2)$$

$$= P(B | A_1) P(A_1) + P(B | A_2) P(A_2)$$

(finite additivity & $B = B \cap A_1 \sqcup B \cap A_2$)
(defn of $P(\cdot | A_i)$)
Ex. prove the general case.

Medical test:

if a patient has a disease, the test returns positive with 95% prob.

Eg 11.1

if a patient is healthy, - - - with 10% prob.

5% of population has the disease.

What is the prob. someone is healthy given they test positive?

Solution:

$$\Omega = \{H, D\} \times \{P, N\}$$

$$\bar{H} = H \times \{P, N\}, \bar{D}$$

Given information: $P(\bar{H}) = 0.95$, $P(\bar{D}) = 0.05$

$$\bar{P} = \{H, D\} \times P; \quad P(\bar{P} | \bar{H}) = 0.10 \quad P(\bar{P} | \bar{D}) = 0.95$$

+ve

Question

$P(\bar{H} | \bar{P}) = \text{Prob. (healthy | test is positive)}$

$$P(\bar{H} | \bar{P}) = \frac{P(\bar{H} \cap \bar{P})}{P(\bar{P})} =$$

$[H - \text{Person is healthy}, \bar{D} - \text{Person tests positive, } N -]$

$$= \frac{P(\bar{P}|\bar{H}) P(\bar{H})}{P(\bar{P}|\bar{H}) P(\bar{H}) + P(\bar{P}|\bar{D}) P(\bar{D})} \quad \leftarrow \begin{array}{l} \text{using defn} \\ \text{\& LTP} \end{array}$$

[We can use LTP b'cos $\Omega = \bar{H} \sqcup \bar{D}$ & $P(\bar{H}), P(\bar{D}) > 0$]

$$= \frac{0.10 \times 0.95}{0.10 \times 0.95 + 0.95 \times 0.05} = \frac{2}{3}$$

Test is bad. Think why? Most people are healthy but also test works badly on 6% healthy!

Ex. What should $P(\bar{P}|\bar{D})$, $P(\bar{P}|\bar{H})$ be so as to make error < 0.01

More serious error is $P(\bar{D}|\bar{N})$?

Find a "cheap test" to minimize $P(\bar{D}|\bar{N})$ even if $P(\bar{H}|\bar{P}) \geq 0.1$.
very small
reasonable

THM 11.2 (Bayes Formula): let $\Omega = \bigcup_{i=1}^n A_i^o$, $P(A_i) > 0 \ \forall i$.

Then
$$P(A_i^o | B) = \frac{P(B | A_i^o) P(A_i^o)}{\sum_{j=1}^n P(B | A_j^o) P(A_j^o)} \quad \text{for } P(B) > 0.$$

Proof:

$$P(A_i^o | B) \stackrel{(\text{defn})}{=} \frac{P(B \cap A_i^o)}{P(B)} = \frac{P(B | A_i^o) P(A_i^o) \text{ --- defn}}{\sum_{j=1}^n P(B | A_j^o) P(A_j^o) \text{ --- LTP.}} \quad \square$$

Eg 11.3 Consider medical test as before.

$$P(\bar{D} | \bar{H}) = 0.02 \quad \text{i.e., 2\% healthy test positive.}$$

$$\underbrace{P(\bar{D} | \bar{N})}_{\text{disease negative}} = \frac{P(\bar{N} | \bar{D}) P(\bar{D})}{P(\bar{N} | \bar{D}) P(\bar{D}) + P(\bar{N} | \bar{H}) P(\bar{H})}$$

$$= \frac{0.05 \times 0.05}{0.05 \times 0.05 + 0.98 \times 0.95}$$

$$\approx 0.0026.$$

$$(P(\bar{N}|\bar{D}) = 1 - P(\bar{P}|\bar{D}))$$

Ideal test: $P(\bar{N}|\bar{D}) = 0$; $P(\bar{N}|H) = 1.$

|||

Fig 120

What is prob. plane is in region i given search in

Region 1 is unsuccessful?

$$\Omega = \{1, 2, 3\} \times \{F, S\}$$

region where plane is there.

$$R_i = \{i\} \times \{F, S\} \quad \text{--- plane is in region } i$$

$\overline{F} = [3] \times \{F\}$ - search in region 1 is unsuccessful.

$w \in \Omega \Rightarrow w = (l, *)$

$\xrightarrow{\text{missing region}}$

$\xrightarrow{\text{F/S after search in region 1}}$

$P(R_i | \bar{F}) =$ Prob. plane is in region i if search in region 1 is unsuccessful.

$$P(R_1 | \bar{F}) = \frac{P(\bar{F} | R_1) P(R_1)}{\sum_{i=1}^3 P(\bar{F} | R_i) P(R_i)}$$

$$= \frac{\beta_1}{\beta_1 + 2}$$

$$\left\{ \begin{array}{l} P(\bar{F} | R_1) = \beta_1 \\ P(\bar{F} | R_i) = 1 \quad i=2,3. \\ P(R_i) = 1/3 \quad i=1,2,3. \end{array} \right.$$

~~$i=2,3$~~ $P(R_i | \bar{F}) = \frac{1}{\beta_1 + 2}$

$$P(\bar{F}) = \frac{\beta_1 + 2}{3} = \sum_{i=1}^3 P(\bar{F} | R_i) P(R_i)$$

Ex 12.2: Choose a permutation from S_n U.A.R.

What is the prob. of exactly k matches? $\sigma(i) = i$ exactly for k indices

Soln:

$$P_n(k) = \text{Prob. of exactly } k \text{ matches.} \quad (\text{using, I-f})$$
$$P_n(0) = \sum_{i=0}^n \frac{(-1)^i}{i!} \quad (A4 / \text{class})$$

Fix k indices. $E(i_1, \dots, i_k) = \{ \sigma(i_j) = j \text{ in the given } k \text{ indices} \}$
 $F(i_1, \dots, i_k) = \{ \text{no matches in the remaining } n-k \text{ indices} \}$

$$P_n(k) = \sum_{i_1, \dots, i_k}^{\neq} P(E(i_1, \dots, i_k) \cap F(i_1, \dots, i_k))$$

(in. additivity)

Fix $i_1, \dots, i_k$ & let $E = E(i_1, \dots, i_k)$
& $F = F(i_1, \dots, i_k)$

$$P(E \cap F) = P(F|E) P(E)$$

$$(*) \quad P(E) = \frac{1}{N} \cdot \frac{1}{(N-1)} \cdots \frac{1}{(N-k+1)} = \frac{(N-k)!}{N!}$$

(Ex.) $P(F|E) = \text{Prob. there are no matches in}$
 remaining $N-k$ in dices

$$P_N(k) = \frac{1}{k!} P_{N-k}(0) = \frac{1}{k!} \sum_{i=0}^{N-k} \frac{(-1)^i}{i!}$$

Ex 12.3

Monty hall problem. - There are 3 closed doors. Prize behind one of the doors. Zonks (0 value items) behind other two.

Randomly arranged. Contestant picks one door at random.

Host (Mr. H.) opens one of the other two doors & shows a zonk.

Mr. H. asks if you want to change your door?

Should you change or not?

Soln:

$$\Omega = \left\{ \overset{\text{prize} = P}{P} \overset{\text{zinks} = \alpha}{\alpha} \overset{\text{zinks} = \alpha}{\alpha}, \overset{\text{zinks} = \alpha}{\alpha} \overset{\text{prize} = P}{P} \overset{\text{zinks} = \alpha}{\alpha}, \overset{\text{zinks} = \alpha}{\alpha} \overset{\text{zinks} = \alpha}{\alpha} \overset{\text{prize} = P}{P} \right\} \times \{1, 2, 3\}$$

arrangement of prize.

↪ door contestant chooses

$$p(w) = \frac{1}{9} \quad \forall w \in \Omega.$$

P_i - Prize is in door i ; D_i - Cont. picks door i

$$P(P_1 | D_1) = \frac{P(D_1 \cap P_1)}{P(D_1)} = \frac{1/9}{1/3} = \frac{1}{3}$$

prize is in door 1. contestant picks door 1

$$P(P_2 \cup P_3 | D_1) = \frac{2}{3} \quad \left(\text{Prob. prize is in 1 one of the other doors.} \right)$$

Answer is same $\forall i=1, 2, 3$

Prob. of winning if you change = $\frac{2}{3}$. ~~1/3~~

Eg 12.4: Shuffle a deck of cards i.e., pick a permutation in S_{52} UAR.

You are shown cards one by one. You've to guess A spade.

Before first card, you have to guess if it is a A sp.

If you guess correctly, you win. Repeat it.

Soln: Suppose you guess up-front - i.e., choose some card i

$\sum_{i=1}^{52} p(i) = 1$. $p(\cdot)$ is a pmf on $[52]$ with $p(i)$ - prob of i

$$P(\text{winning}) = \sum_{i=1}^{52} P(\text{A sp is in } i^{\text{th}} \text{ position} \mid \text{you've chosen } i)$$

$$\begin{array}{l} \text{LTP} \swarrow \\ = \sum_{i=1}^{52} \frac{1}{52} p(i) = \frac{1}{52} \end{array} \quad \begin{array}{l} P(\downarrow) \end{array}$$

is there a better strategy?

16/11 LECTURE 13 - INDEPENDENCE.

Let A & B be 2 events. ^{chance of} Outcome in A isn't affected by outcome being in B & vice-versa.
We say A & B are "independent", how to express mathematically?

Eg 13.1 Pick a card at random from a standard deck.

A = card is a spade

$$P(A) = \frac{1}{4}$$

B = card is $A, 2, \dots, 10$.

$$P(B) = \frac{10}{13}$$

$$P(A|B) = \frac{1}{4} = P(A)$$

$$P(B|A) = \frac{10}{13} = P(B)$$

$$\Rightarrow P(A \cap B) = P(A) P(B) = \frac{1}{4} \cdot \frac{10}{13} = \frac{10}{52}$$

Eg 13.2 Shuffle a standard deck of cards.

A = first card is a spade.

$$P(A) = \frac{1}{4}$$

B = second card is a spade.

$$P(B) = \frac{1}{4}$$

$$P(B|A) = \frac{12}{51} < \frac{1}{4} = P(B)$$

Knowing A makes B less likely!

Reasonable to say A & B are "indep" if $P(A|B) = P(A)$ & $P(B|A) = P(B)$.

$$\text{If } P(A), P(B) > 0 \text{ then } P(A|B) = P(A) \Leftrightarrow P(B|A) = P(B) \\ \Leftrightarrow P(A \cap B) = P(A)P(B)$$

Def 13.3: $A, B \subseteq \Omega$ are independent if $P(A \cap B) = P(A)P(B)$.

Eg 13.4: Let $A \cap B = \emptyset$ then A & B are indep. iff $P(A) = 0$ or $P(B) = 0$.

Eg 13.5: If $P(A) = 0$ & $B \subseteq \Omega$ then A & B are independent.

Eg 13.6: If $A = \Omega$ then A & B are indep. $\forall B \subseteq \Omega$.

Ex 13.7: If A & B are indep. then so are A^c & B , A & B^c , A^c & B^c .

Proof: $P(A^c \cap B) = P(B/A \cap B) = P(B) - P(A \cap B) \quad (A \cap B \subseteq B)$
 $\quad (\text{indep}) = P(B) - P(A)P(B) = P(B)(1 - P(A))$
 $\quad = P(B)P(A^c).$

Others can be proved similarly. $\square$

Def 13.8 $A_1, \dots, A_n$ are pairwise indep. if $P(A_i \cap A_j) = P(A_i)P(A_j) \quad \forall i \neq j$.
 $A_1, \dots, A_n$ are (mutually) indep. if $P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i) \quad \text{--- (Ind)}$
 $\quad \quad \quad \forall \emptyset \neq J \subseteq [n].$

(trivially) indep $\Rightarrow$ pairwise indep.

Ex 13.9. Construct an example of 3 events that are pairwise indep but not indep.

Lemma 13.10 If $A_1, \dots, A_n$ are independent then so are $A_1^*, \dots, A_n^*$
where $A_i^* = A_i$ or A_i^c .

Proof: One case. $A_1^* = A_1, \dots, A_{n-1}^* = A_{n-1}$ & $A_n^* = A_n^c$.

$$\begin{aligned} P(A_1 \cap \dots \cap A_n^c) &= P(A_1 \cap \dots \cap A_{n-1} / A_1 \cap \dots \cap A_n) \\ &= P(A_1 \cap \dots \cap A_{n-1}) - P(A_1 \cap \dots \cap A_n) \\ (\text{indep}) &= \prod_{i=1}^{n-1} P(A_i) - \prod_{i=1}^n P(A_i) \end{aligned}$$

$$(\text{Ind}) \quad = \prod_{i=1}^{n-1} P(A_i) P(A_n^c).$$

shows for $J = [n]$. Verify $\forall J \subseteq [n], J \neq \emptyset$ & also other cases. $\square$

LEMMA

13.11

Let $(\Omega, P) = (\Omega_1 \times \dots \times \Omega_n, P_1 \times \dots \times P_n)$ i.e., (Prod. prob. space)

$$P(\omega) = P_1(\omega_1) \dots P_n(\omega_n) \text{ for } \omega = (\omega_1, \dots, \omega_n)$$

Let $A_i \subseteq \Omega_i, i=1, \dots, n$. $\bar{A}_i = \Omega_1 \times \dots \times A_i \times \dots \times \Omega_n \subseteq \Omega$.

Then $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$ are indep. [Abuse of notation - we say $A_1, \dots, A_n$ are indep.]

Proof: Let me check (Ind) condition for $J = [n]$ & rest is an exercise.

$$\bigcap_{j=1}^n \bar{A}_j = A_1 \times \cdots \times A_n.$$

Prob. of prod. prob is $P(\bar{A}_i) = P(A_i^c) \quad \forall i$ [A2]

$$P(A_1 \times \cdots \times A_n) = \prod_{i=1}^n P(A_i)$$

$$\text{So } P\left(\bigcap_{j=1}^n \bar{A}_j\right) = P(A_1 \times \cdots \times A_n) = \prod_{j=1}^n P(A_j) = \prod_{j=1}^n P(\bar{A}_j). \quad \square$$

Terminological $(\Omega_i, \mathcal{P}_i)$ — Experiment ' i '. $i = 1, \dots, n$.

If we say, Experiment 1, $\dots$, n are performed independently

then we refer to the prod. prob. space $(\Omega, \mathcal{P}) = (\Omega_1 \times \cdots \times \Omega_n, \mathcal{P}_1 \times \cdots \times \mathcal{P}_n)$.

Eg 13.12

$$\Omega_i = \{0, 1\} \quad P_i(1) = p_i \quad 1 \leq i \leq n.$$

Experiment i is Coin toss with heads prob ' p_i '

Let us assume the n coins are tossed independently.

Then the prob. space is $\Omega = \Omega_1 \times \dots \times \Omega_n = \{0, 1\}^n$

$$P(w) = \prod_{i=1}^n p_i^{w_i} (1-p_i)^{1-w_i} \quad w = (w_1, \dots, w_n) \in \{0, 1\}^n.$$

Eg 13.13 (1) n coins (all with heads prob. p) are tossed independently.

Prob. space is ?

(2) Toss a fair coin, roll a dice & pick a card independently.

(3) Shuffle two decks of cards independently.

(4) Pick ' r ' cells from ' n ' cells independently i.e., pick a cell from ' n ' cells w.o.r & repeat this exp. ' r ' times indep.!

18/11

LECTURE 14 - INDEPENDENCE - EXAMPLES.

Eg 14.1: Pick 'r' cells from 'n' cells independently i.e., pick a cell from 'n' cells U.A.R. & repeat this exp. 'r' times indep.!

$$\Omega = [n]^r \quad p(\omega) = \frac{1}{n^r} \quad \omega = (\omega_1, \dots, \omega_r) \in \Omega.$$

Eg 14.2: Special cases of Eg 14.1

- (1) $n=2$ $\rightarrow$ r indep. fair coin tosses (r indep. bits)
- (2) $n=6$ $\rightarrow$ r indep. die rolls.
- (3) $n \geq 1$ $\rightarrow$ r persons indep. pick a hash-code from $1, \dots, n$.

Eg 14.3: n persons pick independently a no: whereby i^{th} person picks a no: in $[i]$ U.A.R.

$$\Omega = \prod_{i=1}^n [i] \quad ; \quad p(w) = \prod_{i=1}^n \frac{1}{i} = \frac{1}{n!} \quad w = (w_1, \dots, w_n) \in \Omega$$

$$= \prod_{i=1}^n \{1, \dots, i\} = \{1\} \times \{1, 2\} \times \{1, 2, 3\} \times \dots \times \{1, 2, \dots, n\}.$$

Ex 14.4: Consider (Ω, p) as above.

Let $w = (w_1, \dots, w_n) \in \Omega$.

Permutations

$w_1 = 1$	$\rightarrow$	$\underline{1}$	$\rightarrow$	$\sigma^1 \in S_1$
$w_2 = 2$	$\rightarrow$	$\underline{1} \ 2$	$\rightarrow$	$\sigma^2 \in S_2$
$w_3 = 2$	$\rightarrow$	$\underline{1} \ 3 \ 2$	$\rightarrow$	$\sigma^3 \in S_3$
$w_4 = 3$	$\rightarrow$	$\underline{1} \ 3 \ 4 \ 2$	$\rightarrow$	$\sigma^4 \in S_4$
$\vdots$				

$w_n =$ position of n among existing $(n+1)$ numbers $\rightarrow \sigma^n \in S_n$.

Reflexively, in σ^{i-1} , we put i in the w_i^{th} position &

get σ^i . $\sigma^1 = 1$.

$$\Omega \ni \omega \mapsto \sigma^\omega \in S_n$$

$P(\sigma^\omega = \sigma) = ?$ for a fixed $\sigma \in S_n$.
(i.e., induced prob. on S_n)

observe $\omega \mapsto \sigma^i \in S_i$ $i=1, \dots, n$.

TO PROVE THAT $P(\sigma^i = \sigma) = \frac{1}{i!} \quad \forall i=1, \dots, n \text{ \& } \sigma \in S_i$.

— ①

INDUCTION: ① holds for $i=1$ trivially.

Assume ① holds for $n-1$.

Let π be $\sigma \in S_n$ restricted to S_{n-1} . (remove n from σ).

i.e., if $n=4$ & $\sigma = 1342$ then $\pi = 132$.

$$P(\sigma^n = \sigma) = P(\sigma^n = \sigma \mid \sigma^{n+} = \pi) P(\sigma^{n+} = \pi)$$

Let $j \in [n]$ be the position of n in σ . (Above eg. position of 4 is 3 in σ)

$$= P(W_n = j \mid \sigma^{n+1} = \pi) P(\sigma^{n+1} = \pi)$$

σ^{n+1} depends on $(w_1, \dots, w_{n+1})$. $\pi \in S_{n+1}$

$$C = \{ \sigma^{n+1} = \pi \} = B \times [n], \quad B \subseteq [1] \times [2] \times \dots \times [n+1]$$

$$D = \{ w_n = j \} = [1] \times [2] \times \dots \times [n+1] \times A \quad \text{where } A = \{j\}.$$

Since our prob. space is prod. space, C & D are indep. [check]

So

$$P(W_n = j \mid \sigma^{n+1} = \pi) = P(D \mid C) = P(D) = \frac{1}{n}.$$

So

$$P(\sigma^n = \sigma) = \frac{1}{n} P(\sigma^{n+1} = \pi) = \frac{1}{n!} \quad (\text{by induction step})$$

$\downarrow$
rand. permutation in S_{n+1} .

$\Rightarrow \sigma^n$ is a rand. permutation in S_n . $\blacksquare$

LEMMA 14.4 If A, B, C are independent events then so are $A \cup B$ & C
and $A \cap B$ & C .

Proof:

$$\begin{aligned} P((A \cup B) \cap C) &= P((A \cap C) \cup (B \cap C)) \\ &= P(A \cap C) + P(B \cap C) - P(A \cap B \cap C) \quad (\text{IE}) \\ &= P(A)P(C) + P(B)P(C) - P(A \cap B)P(C) \quad (\text{indep.}) \\ &= P(C) (P(A) + P(B) - P(A \cap B)) \\ &= P(C) P(A \cup B) \quad (\text{IE}) \end{aligned}$$

||| $\Rightarrow$ for $A \cap B$ & C .

Ex. 14.5. Generalize above to more events. For eg. A, B, C, D all indep.
then so are $A \cup B$ & $C \cap D$.

Ex 14.6 There are n types of coupons. A coupon of type i is selected
with prob. p_i ($p_i \geq 0$ & $\sum_{i=1}^n p_i = 1$).
 k indep. coupons are selected.

Find prob. $\exists$ at least one type i coupon & $\exists$ at least one type i or
at least one type j .

Soln:

$$\Omega = [n]^k; \quad P(\omega) = \prod_{l=1}^k p_{\omega_l} \quad \omega_l \in [n]; \quad l=1, \dots, n.$$

$A_i^o = \exists$ at least one type i coupon.

$$P(A_i^o) = 1 - P(A_i^c) = 1 - P(\text{no coupon of type } i)$$

So $A_i^c = ([n] \setminus \{i\})^k$ $P(A_i^c) = (1 - p_i)^k$ (Since A_i^c is a prod. set & using indep.)

$$P(A_i^o) = 1 - (1 - p_i)^k.$$

$$P(A_i^o \cup A_j^o) = P(A_i^o) + P(A_j^o) - P(A_i^o \cap A_j^o)$$

$$\underbrace{\quad}_{=} 1 - P((A_i^o \cup A_j^o)^c) = 1 - P(A_i^c \cap A_j^c)$$

$$A_i^c \cap A_j^c = ([n] \setminus \{i, j\})^k \quad P(A_i^c \cap A_j^c) = (1 - p_i - p_j)^k$$

$$P(A_i^o \cup A_j^o) = 1 - (1 - p_i - p_j)^k$$

□

$$P\left(\bigcap_{i=1}^n A_i^o\right) = ? \quad \text{if } k < n \quad P\left(\bigcap_{i=1}^n A_i^o\right) = 0.$$

Eg 14.7 (Gambler's ruin problem).

Two gamblers A & B. Keep tossing a p -biased coin independently.

If H, B pays 1\$ to A & if T, A pays 1\$ to B.

Initially A has i \$ & B has $n-i$ \$.

Game over if A or B have 0\$. A winner if A has n \$

Solution:

p -biased coin

$$P(H) = p$$

$$P(T) = 1-p = q$$

$$P_i = P(\text{A wins starting with } i \text{ \$}) = P(A_i)$$

H_1 = event first ^{" A_i "}toss is heads. $T_1 = \dots$ is tails.

$$P_i = P(A_i) = P(A_i | H_1) P(H_1) + P(A_i | T_1) P(T_1)$$

Given H_1 , A has $i+1$ \$ ^(LTP) & B has $n-i-1$ \$ & game continues if $i+1 < n$.

Since the next tosses are indep. of first toss, this is same as starting the game with $(i+1)\$$ for A .

$$\text{So } P(A_i^o | H_1) = P(A_{i+1}^o) = P_{i+1}^o; \quad P(A_i^o | T_1) = P_{i-1}^o$$

$$\Rightarrow P_i^o = p P_{i+1}^o + q P_{i-1}^o$$

$$p+q=1 \Rightarrow p P_i^o + q P_i^o = p P_{i+1}^o + q P_{i-1}^o$$

$$\Rightarrow P_{i+1}^o - P_i^o = \frac{q}{p} (P_i^o - P_{i-1}^o)$$

$$P_n^o = 1, \quad P_0^o = 0$$

$$\text{use this } P_i^o - P_{i-1}^o = \left(\frac{q}{p}\right)^{i-1} P_1^o$$

$$\begin{array}{l} \text{Add } P_i - P_{i-1} \\ + P_{i-1} - P_{i-2} \\ + \dots + P_2 - P_1 \end{array} \Rightarrow P_i - P_1 = P_1 \left[\left(\frac{q}{p}\right) + \dots + \left(\frac{q}{p}\right)^{i-1} \right]$$

$$1 = P_{\text{win}} = P_1 \left[1 + \left(\frac{q}{p}\right) + \dots + \left(\frac{q}{p}\right)^{n-1} \right]$$

$$= P_1 \frac{1 - \left(\frac{q}{p}\right)^n}{1 - \frac{q}{p}}, \quad \frac{q}{p} \neq 1.$$

$$\Rightarrow P_1 = \frac{1 - \frac{q}{p}}{1 - \left(\frac{q}{p}\right)^n} \Rightarrow P_i = \frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^n}, \quad \frac{q}{p} \neq 1$$

$$\text{If } \frac{q}{p} = 1 \ (\Leftrightarrow p = 1/2) \quad P_i = \frac{i}{n}.$$

$$Q_i = P(\text{B winning starting with } n-i \text{ f}) = \frac{1 - \left(\frac{p}{q}\right)^{n-i}}{1 - \left(\frac{p}{q}\right)^n}$$

check $P_i + Q_i = 1$. $\forall i = 0, \dots, n$. 

23/11 LECTURE 15 - THE RISE OF URNS.

Ex 15.1 Let $(\Omega_i, \mathbb{P}_i), i=1, \dots, n$ be prob. spaces. Let $\tilde{p}$ be a pmf on $[n]$. We select i at random according to $\tilde{p}$ & then choose an element of Ω_i at random according to $\mathbb{P}_i$. What is the prob. space?

Ex 15.2 . There are $(m+1)$ urns labelled $0, \dots, m$. Urn k contains k red balls & $m-k$ white balls. An urn is chosen at random & n balls are sampled from the urn with replacement.

$n \geq 1$ Given that n balls are red, what is the prob. $(n+1)^{\text{th}}$ ball is also red?

Soln. $A_n = \{ \text{all } 'n' \text{ balls are red} \}$ $A_{n+1} = \{ \dots \}$

Qn- is $P(A_{n+1} | A_n) = ?$

$$P(A_{n+1} | A_n) = \frac{P(A_{n+1} \cap A_n)}{P(A_n)} = \frac{P(A_{n+1})}{P(A_n)}$$

$B_k = \{ \text{Urn 'k' is chosen} \}$; By assumption $P(B_k) = \frac{1}{m+1}$.
 $P(A_n | B_k) = \left(\frac{k}{m}\right)^n$ (prob. of choosing red ball in k^{th} urn is $\frac{k}{m}$)

$$\text{LTP} \Rightarrow P(A_n) = \sum_{k=1}^m P(A_n | B_k) P(B_k) \\ = \frac{1}{m+1} \sum_{k=1}^m \left(\frac{k}{m}\right)^n \xrightarrow[m \rightarrow \infty]{\text{Ex.}} \frac{1}{n+1}$$

so $P(A_{n+1} | A_n) \approx \frac{n+1}{n+2}$ (Laplace's rule of succession)

Laplace Sun-rise prediction - Each urn is a universe. In k^{th} universe prob. of sun rising is $\frac{k}{m}$. Prob. (sun will rise tmrw | it has risen so far) = ...
 $n = 1,826,213$.

Fig 15.3 URN MODELS :

There is an urn that contains $r \geq 0$ red balls & $b \geq 0$ black balls.
A ball is selected at random & put back after noting the colour.
- c ^{new} balls of the same colour are added. $c \in \mathbb{Z}$ } fixed.
- d ^{new} balls of the opposite colour are added. $d \in \mathbb{Z}$ }

The urn contains $r+b+c+d$ balls, if $r+c \geq 0, b+d \geq 0$.
for n steps as long as # of balls ≥ 0 .
we repeat this procedure

$c < 0 \Rightarrow -c$ balls are removed. Same for d .

Very hard to analyze except special cases. For more, Antar B (ISI-D)
& Krishanu M. (ISI-K).

Case 1: $c = -1, d = 0$. - At each step a ball is removed at random.

Process will go on until $r+b$ steps.

B_i - black ball is chosen at i^{th} step ; R_i - Red ball is chosen at i^{th} step

$$P(B_1) = \frac{b}{r+b} = 1 - P(R_1)$$

$$P(B_2 | B_1) = \frac{b-1}{r+b-1}$$

$$P(B_2 | R_1) = \frac{b}{r+b-1}$$

use LTP
to compute
(P(B₂)).

$$P(B_3 | B_2 \cap B_1) = \frac{b-2}{r+b-2}$$

$$P(B_3 | R_1 \cap B_2) = \frac{b-1}{r+b-2}$$

$$P(B_3 | R_1 \cap R_2) = \frac{b}{r+b-2}$$

$$P(B_3 | B_1 \cap R_2) = \frac{b-1}{r+b-2}$$

$$P(B_3 \cap R_2 \cap B_1) = \frac{b-1}{r+b-2} P(R_2 \cap B_1)$$

$$= \frac{b-1}{r+b-2} \cdot \frac{r}{r+b-1} \cdot \frac{b}{r+b} = \frac{b(b-1)r}{(r+b)(r+b-1)(r+b-2)}$$

$$P(B_3 \cap B_2 \cap R_1) = \frac{b-1}{r+b-2} \cdot \frac{b}{r+b-1} \cdot \frac{r}{r+b} = \downarrow$$

$\Rightarrow$ order doesn't seem to matter. Can you prove it? EX.

Ex. 15.4. Suppose $r \geq b$. What is the prob. that # of red balls removed up to step $k \geq$ # of black balls removed up to step k . $\forall k=1, \dots, r+b$.
[Ballot problem]

Eg 15.5 (POLYA'S URN) $d=0$; $c>0$.

At each stage, more balls of chosen colour are added. [Reinforcement model]

Analyse as before

$$P(B_2 \cap B_1) = \frac{b(b+c)}{(b+r)(b+r+c)}$$

$$P(R_2 \cap R_1) = \frac{r(r+c)}{(b+r)(b+r+c)}$$

$$P(B_2 \cap R_1) = \frac{br}{(b+r)(b+r+c)}$$

$$P(R_2 \cap B_1) = \frac{br}{(b+r)(b+r+c)}$$

$n \geq 3$ $S \subseteq [n]$ $|S|=n_1$, $n-n_1=n_2$: $A_S = \bigcap_{i \in S} R_i \cap \bigcap_{j \in S^c} B_j$

red balls^{chosen} at indices in S & Black balls^{chosen} at indices in S^c .

$n \geq 3$

$P(A_S) =$

$$\left\{ \begin{array}{ll} \frac{b(b+c) \cdots (b+n_1c-c)}{(b+r)(b+r+c) \cdots (b+r+(n-1)c)} & n_1 = 0 \\ \frac{r(r+c) \cdots (r+(n-1)c)}{(b+r) \cdots (b+r+(n-1)c)} & n_2 = 0 \\ \frac{b(b+c) \cdots (b+(n_2-1)c) r(r+c) \cdots (r+(n-1)c)}{(b+r) \cdots (b+r+(n-1)c)} & n_1 > 0, n_2 > 0 \end{array} \right.$$

Proof By induction on n .
True for $n=1, 2$.

WLOG, let

Assume

$n_1 > 0, n_2 > 0$

[Formula is true for $n_1=0$ or $n_2=0$]

$n \in S^c$.

Set $B =$

$$\bigcap_{i \in S} R_i \cap \bigcap_{j \in S^c \setminus n} B_j$$

Since $A_S = B \cap B_n$

$$P(A_S) = P(B_n | B) P(B).$$

$$= \frac{b + (n_2 - 1)c}{b + r + (n - 1)c} P(B)$$

If B happens, we have chosen n_1 red balls & $n_2 - 1$ black balls in the first $(n - 1)$ steps $\Rightarrow$ $n_1 c$ new red balls are added & $(n_2 - 1)c$ new black balls are added.

$$P(B_n | B) = \frac{\# \text{ black balls given } B}{\# \text{ balls given } B} = \frac{b + (n_2 - 1)c}{b + r + (n - 1)c}.$$

By induction hypothesis,

$$P(B) = \frac{b(b+c) \cdots (b + (n_2 - 2)c) r(r+c) \cdots (r + (n_1 - 1)c)}{(b+r) \cdots (b+r + (n - 2)c)}$$

substitute formulas for $P(B_n | B)$ & $P(B)$ to derive formula for $P(A_n)$.
 Wdy prove for $n \in \mathbb{S}$. □

Eg 15.6 (Ehrenfest urn model). Two chambers & n particles in total.
 A particle is chosen at random & moved to the other chamber.
 Colours: particles in chamber I in Red; chamber II in black.

Whenever a particle is chosen, change its colour.

This is urn model with $c=-1, d=1$. # of balls = $r+b=N$

$$NR_{k,l} = \{ \# \text{ of red balls after } l \text{ rounds} = k \} \quad 0 \leq k \leq N$$

$$= NB_{N-k,l}$$

$$P(R_{l+1} | NR_{k,l}) = \frac{k}{N} = 1 - P(B_{l+1} | NR_{k,l})$$

$$(LTP) \quad P(NR_{k,l}) = \sum_{j=0}^N P(NR_{k,l} | NR_{j,l-1}) P(NR_{j,l-1})$$

At each step # of red balls $\nearrow +1$ $P(NR_{k,l} | NR_{j,l-1}) = 0$
 $\searrow -1$ $j \neq k+1, k-1$

$$= P(NR_{k,l} | NR_{k+1,l+1}) P(NR_{k+1,l+1}) \\ + P(NR_{k,l} | NR_{k+1,l}) P(NR_{k+1,l})$$

$$P(NR_{k,l}) = \frac{k+1}{N} P(NR_{k+1,l+1}) + \frac{N-k+1}{N} P(NR_{k+1,l})$$

$$\left[\begin{aligned} \text{Since } P(NR_{k,l} | NR_{k+1,l+1}) &= P(R_l | NR_{k+1,l+1}) = \frac{k+1}{N} \\ P(NR_{k,l} | NR_{k+1,l}) &= P(B_l | NR_{k+1,l}) = \frac{N-k+1}{N} \end{aligned} \right]$$

Assume $P(NR_{k,l}) \rightarrow u_k$ as $l \rightarrow \infty \quad \forall 0 \leq k \leq N$.

Then we get that $u_k = \frac{k+1}{N} u_{k+1} + \frac{N-k+1}{N} u_{k+1} \quad ; \quad u_{-1} = u_{N+1} = 0$

Solve this i.e., compute u_k .

26/11 LECTURE 16 - RANDOM VARIABLES.

Not always do we observe outcome of an experiment but may be some 'observable' only.

For eg; experiment is 'n' coin tosses. We record/observe only # of heads.
of heads is an 'observable'.

For eg; experiment is a roll of 2 dice. An observable is sum of the dice.

For eg; " is a shuffle of deck of 'n' cards.
of matches is an observable.

Defn 16.1 (Ω, P) be a finite PS. A fn. $X: \Omega \rightarrow \mathbb{R}$ is called a random variable. We denote it by $X: (\Omega, P) \rightarrow \mathbb{R}$.
(r.v.) (Ω, p) .

PROP 16.2: Let $X: (\Omega, P) \rightarrow E \subset \mathbb{R}$ ($E \subset \mathbb{R}$, E is a finite subset of $\mathbb{R}$)
be a r.v.
(i.e. X takes values in E)

Define $p_x(x) = P\{X=x\} = P(\{\omega \in \Omega : X(\omega)=x\})$, $x \in E$.

Then (E, p_x) is a PS i.e., p_x is a pmf on E .

Proof:

(1) $p_x(x) \geq 0 \quad \forall x \in E$.

(2) $\{X=x\} = X^{-1}(x)$, $X^{-1}(x), x \in E$ are ^{pairwise} disjoint sets.

$$\sum_{x \in E} p_x(x) \stackrel{\substack{\uparrow \\ \text{def of } p_x}}{=} \sum_{x \in E} P(X^{-1}(x)) \stackrel{\substack{\downarrow \\ \text{fin-add}}}{=} P\left(\bigsqcup_{x \in E} X^{-1}(x)\right) = P(\Omega) = 1.$$

This proves p_x is a pmf on E . 

defn 16.3

Given a r.v. $X: (\Omega, P) \rightarrow \mathbb{R}$, we call the function

$p_x: \mathbb{R} \rightarrow [0,1]$ defined as $p_x(x) := P\{X=x\} = P(X^{-1}(x))$, $x \in \mathbb{R}$

where $p_x(x) = 0$ if $x \notin X(\Omega)$. to be the pmf of X .

Define $\forall A \subseteq \mathbb{R}$, $P_x(A) := P(X \in A) = P(X^{-1}(A))$
 $[\{X \in A\} := X^{-1}(A)]$

P_X is called the prob. distribution of X .

Ex
1604.

$$(1) \Omega = \{0, 1\} \quad P(1) = p.$$

$$X: \Omega \rightarrow \mathbb{R} \quad (\text{identity fn.}) \\ \omega \mapsto \omega$$

$$P_X(1) = p, \quad P_X(0) = 1-p, \quad P_X(x) = 0 \quad \forall x \notin \Omega.$$

$$(2) \Omega = \{0, 1\}^n \quad P_p - \text{product Bernoulli distribution. } (n \text{ indep. biased coins})$$

$$X: \Omega \rightarrow \mathbb{R} \\ \omega = (\omega_1, \dots, \omega_n) \mapsto \sum_{i=1}^n \omega_i \quad (= \# \text{ of heads in } n \text{ tosses})$$

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & k = 0, \dots, n \\ 0 & k \notin \{0, \dots, n\} \end{cases}$$

$\parallel$
 $P(X^{-1}(k))$

$$(3) \Omega = [n], \quad P - \text{uniform prob. distribn.}$$

$$X: \Omega \rightarrow \mathbb{R} \quad (\text{identity}) \\ \omega \mapsto \omega$$

$$p_X(x) = \begin{cases} \frac{1}{n} & \text{if } x \in [n] \\ 0 & \text{else.} \end{cases}$$

(4) $\Omega = S_n$, P - unif. Prob. distribution.

$$X: \Omega \rightarrow \mathbb{R} \\ \omega \mapsto \sum_{i=1}^n \mathbb{1}[\omega_i = i] \quad (= \# \text{ of matches in a permutation})$$

$$p_X(x) = \begin{cases} P(\exists \text{ exactly } x \text{ matches}) & \text{if } x \in [n] \\ 0 & x \notin [n] \end{cases}$$

$\overset{''}{P(X=x)}$

Defn 16.6 let $X: (\Omega, P) \rightarrow \mathbb{R}$ be a r.v. Expectation of a r.v. X is defined as $\mathbb{E}[X] := \sum_{\omega \in \Omega} X(\omega) p(\omega)$.

Suppose $X(\Omega) \subseteq E \subset \mathbb{R}$ then

$$\begin{aligned}
E[X] &= \sum_{\omega \in \Omega} X(\omega) p(\omega) = \sum_{x \in E} \sum_{\omega \in X^{-1}(x)} X(\omega) p(\omega) \left(\Omega = \bigsqcup_{x \in E} X^{-1}(x) \right) \\
&= \sum_{x \in E} x \sum_{\omega \in X^{-1}(x)} p(\omega) \\
&= \sum_{x \in E} x \underbrace{P(X^{-1}(x))}_{= P_X(x)} = \sum_{x \in E} x p_X(x) \quad \text{--- (1)}
\end{aligned}$$

if $f: \mathbb{R} \rightarrow \mathbb{R}$ $\exists f(x) \neq 0$ for only finitely many values x
 (i.e., $\mathbb{R} \setminus f^{-1}(0) \subset \subset \mathbb{R}$) then

$$\sum_x f(x) = \sum_{x \in \mathbb{R}} f(x) = \sum_{x: f(x) \neq 0} f(x) .$$

Using this defn. & above observation, we have that $E[X] = \sum_x x p_X(x)$
 (1) as $P_X(x) = 0 \quad \forall x \notin E$.

Eg 16.6 (Revisit Eg 16.4) (1) $X: \{0,1\} \rightarrow \mathbb{R} \quad \omega \mapsto \omega \cdot p(1) = p$.

$$E[X] = \sum_x x p_X(x) = 0 \cdot (1-p) + 1 \cdot p = p.$$

"Avg. of observable".

"IP(X=1)"

(2) $X = \#$ of heads in 'n' indep. ^{biased} coin tosses. ($\Omega = \{0,1\}^n$)

$$E[X] = \sum_x x p_X(x) \quad (P_p -)$$

Avg. # of heads in 'n' coin tosses.

$$= \sum_{k=0}^n k p_X(k) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = np. \text{ [check]}$$

(3) $X =$ Position of a randomly picked card among 'n' cards.

($\Omega = [n]$; P - Unif., $X: \Omega \rightarrow \Omega \quad w \mapsto w$)

$$E[X] = \sum_{k=1}^n k P_X(k) = \sum_{k=1}^n \frac{k}{n} = \frac{n+1}{2}.$$

average position

(4) $X = \#$ of matches in a random permutation.

$$E[X] = \sum_{k=0}^n k p_X(k) = \sum_{k=1}^n k P(\exists \text{ exactly 'k' matches})$$
$$= ??? \quad (\text{check})$$

(5) $X: (\Omega, \mathcal{P}) \rightarrow \mathbb{R}$, $A \subseteq \Omega$.
 $\omega \mapsto \mathbb{1}[\omega \in A]$.

[A is set of good outcomes
 $X = \mathbb{1}[\text{good outcome}]$]

$$P_X(x) = \begin{cases} P(A^c) & x=0 \\ P(A) & x=1 \\ 0 & x \notin \{0,1\} \end{cases}$$

" $P(X^{-1}(x))$

" $P(\{\omega: X(\omega)=x\})$

$$E[X] = P(A). \quad [\text{Every prob. of an event is } E[X]! \text{ for some } X.]$$

J.D Take a prob. space, find an interesting event A & compute $P(A)$.

EJD " , define a R.V. & compute $E[X]$.

EX. (1) Toss a fair coin. If heads I get 1 INR
if tails I lose 1 INR.

$X =$ Net profit. $E[X] = ?$

(2) Toss 'n' fair coins indep. Same as above for heads & tails.

$X =$ Net profit after 'n' coin tosses. $E[X] = ?$

(3) $\Omega = \{0, 1, 2\}$ $p(2) = 0$, $p(1) = p = 1 - p(0)$.

$X: \Omega \rightarrow \mathbb{R}$ identity fn. $E[X] = ?$

26/11 LECTURE 17 - EXPECTATION

Defn
16.3

Given a r.v. $X: (\Omega, \mathcal{P}) \rightarrow \mathbb{R}$, we call the function

$p_X: \mathbb{R} \rightarrow [0, 1]$ defined as $p_X(x) := \mathbb{P}(X = x) = \mathbb{P}(X^{-1}(x))$, $x \in \mathbb{R}$

where $p_X(x) = 0$ if $x \notin X(\Omega)$. to be the pmf of X .

Define $\forall A \subseteq \mathbb{R}$, $P_X(A) := \mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A))$ (Prob. distribn. of X).
 $[\{X \in A\} := X^{-1}(A)]$

$$\mathbb{E}[X] := \sum_{\omega \in \Omega} X(\omega) p(\omega) = \sum_x x p_X(x).$$

$\forall A \subseteq \Omega$, $X(\cdot) = \mathbb{1}_A(\cdot)$ then $\mathbb{E}[X] = \mathbb{P}(A)$

LEMMA
17.1

$X: (\Omega, \mathcal{P}) \rightarrow \mathbb{R}$ be a r.v. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a fn.

Then $f(X)$ is a r.v. & $\mathbb{E}[f(X)] = \sum_x f(x) p_X(x)$.

Proof: $f(X): (\Omega, \mathcal{P}) \rightarrow \mathbb{R}$ is a composition of f & X & so a r.v.

$$\begin{aligned}
\mathbb{E}[f(X)] &= \sum_{\omega \in \Omega} f(X(\omega)) p(\omega) \\
&= \sum_{x \in E} \sum_{\omega \in X^{-1}(x)} f(x) p(\omega) \quad (\text{let } E = X(\Omega)) \\
&= \sum_{x \in E} f(x) p_X(x) = \sum_x f(x) p_X(x) \quad \square
\end{aligned}$$

Defn 17.2 Two r.v. X & Y are said to be equal in distribution if $p_X \equiv p_Y$ as fns in $\mathbb{R}$ i.e., $p_X(x) = p_Y(x) \forall x \in \mathbb{R}$.
 Denoted by $X \stackrel{d}{=} Y$.

FACT: $X \stackrel{d}{=} Y \implies \mathbb{E}[f(X)] = \mathbb{E}[f(Y)] \quad \forall f: \mathbb{R} \rightarrow \mathbb{R}.$

Cor 17.3 Let $X: (\Omega, \mathcal{P}) \rightarrow \mathbb{R}$ be a r.v. & ^{let} $X(\Omega) \subseteq E \subset \subset \mathbb{R}$ hold.
 Let $Y: (E, p_X) \rightarrow \mathbb{R}$ be identity fn.

Then $\mathbb{E}[f(x)] = \mathbb{E}[f(y)] \quad \forall f: \mathbb{R} \rightarrow \mathbb{R}.$

Proof:

Exercise.

For eg. Choose $E = \{x \in \mathbb{R} : p_x(x) > 0\} \subset \mathbb{R}$
 $\hookrightarrow \text{Yes } \Omega \text{ is finite.}$

Terminology: (1) X is Bernoulli(p) r.v. if $p_x(0) = 1-p = 1-p_x(1).$

(2) X is Binomial(n, p) r.v. if $p_x(k) = \binom{n}{k} p^k (1-p)^{n-k}$

(3) X is Hypergeometric(n, m, r) r.v. if $P(X=k) = \frac{\binom{m}{k} \binom{n-m}{n-k}}{\binom{n}{r}}$

$X \stackrel{d}{=} \text{Ber}(p) \Rightarrow X \text{ is Bernoulli}(p) \text{ r.v.}$

$X \stackrel{d}{=} \text{Bin}(n, p) \Rightarrow X \text{}$

$X \stackrel{d}{=} \text{HyperG}(n, m, r) \Rightarrow \text{}$

$X \stackrel{d}{=} \text{Unif}([n]) \Rightarrow P(X=k) = \frac{1}{n}, k=1, \dots, n.$

PROPOSITION: (1) X_1, X_2 are r.v. then $E[aX_1 + bX_2] = aE[X_1] + bE[X_2]$
17.4 (Linearity of Expectation) $\forall a, b \in \mathbb{R}$.

(2) X_1, X_2 are r.v. $\ni P_{X_1} \leq P_{X_2}$ then $E[X_1] \leq E[X_2]$.

(3) Let $X_1, X_2: (\Omega, \mathcal{P}) \rightarrow \mathbb{R}$ be r.v. $\ni X_1(\omega) \leq X_2(\omega)$ (Monotonicity of E)
then $E[X_1] \leq E[X_2]$.

Proof: (1) Assume $X_1, X_2: (\Omega, \mathcal{P}) \rightarrow \mathbb{R}$ [Think what happens if not?]

then $aX_1 + bX_2$ is also a r.v.

By defn $E[aX_1 + bX_2] = \sum_{\omega \in \Omega} (aX_1(\omega) + bX_2(\omega)) p(\omega)$
(defn)

(split the sum as it is fin.) $= \sum_{\omega \in \Omega} aX_1(\omega)p(\omega) + \sum_{\omega \in \Omega} bX_2(\omega)p(\omega)$

(defn) $= aE[X_1] + bE[X_2]$

$$(2) \quad E[X_1] = \sum_x x p_{X_1}(x) \stackrel{\substack{\uparrow \\ \text{assumption}}}{\leq} \sum_x x p_{X_2}(x) = E[X_2]$$

$$(3) \quad E[X_1] = \sum_{\omega} X_1(\omega) P(\omega) \stackrel{\leq}{\leq} \sum_{\omega} X_2(\omega) P(\omega) = E[X_2].$$

Ex 17.5

Let (Ω, P) be a PS. & $A_1, \dots, A_n$ be events.

Define $X = \mathbb{1}_{A_1 \cup \dots \cup A_n}$ ($\omega \mapsto \mathbb{1}[\omega \in A_1 \cup \dots \cup A_n]$)

$$X_i = \mathbb{1}_{A_i} \quad i=1, \dots, n.$$

Trivially $X \leq \sum_{i=1}^n X_i$.

By mon. & linearity

$$E[X] \leq \sum_{i=1}^n E[X_i]$$

$$\parallel \quad P(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^n P(A_i) \quad \parallel \quad \text{[Union bound]}$$

Ex 17.6

Try proving I-E & Bonferroni ineq. using lin. & mon.

Ex 17.7

$$X \stackrel{d}{=} \text{Bin}(n, p) \quad \mathbb{E}[X] = \sum_{k=0}^n k p_X(k)$$

Let $\Omega = \{0, 1\}^n$, P_p - prod. Bernoulli distribution.

$Y_i: \Omega \rightarrow \mathbb{R}; \omega \mapsto \omega_i$. (Heads or tails in the i^{th} toss)

$$P_{Y_i}(0) = 1 - p = 1 - P_{Y_i}(1) \Rightarrow \mathbb{E}[Y_i] = p.$$

$\uparrow$
 $(P(Y_i=0))$

$$Y = Y_1 + \dots + Y_n \quad (= \# \text{ heads in } n \text{ tosses})$$

check $Y \stackrel{d}{=} X \quad (P_Y(k) = P_X(k) \quad \forall k=0, \dots, n)$

so by om fact $\mathbb{E}[X] = \mathbb{E}[Y] = \mathbb{E}[Y_1 + \dots + Y_n]$

$$\stackrel{(\text{linearity})}{=} \sum_{i=1}^n \mathbb{E}[Y_i] = np. \quad \square$$

Ex

17.8

$X = \#$ matches in a random permutation.

$$X: S_n \rightarrow \mathbb{R} \quad X(w) = \sum_{i=1}^n \mathbb{1}[w(i) = i]$$

$\mathbb{P}$ -v.p.d.

$$\mathbb{E}[X] = \sum_{k=1}^n k p_X(k)$$

Define $X_i: S_n \rightarrow \mathbb{R} \quad X_i(w) = \mathbb{1}[w(i) = i]$

$$\begin{aligned} \mathbb{E}[X_i] &= \mathbb{P}[X_i(w) = 1] = \mathbb{P}(\{w: w(i) = i\}) \\ &= \frac{1}{n} \quad \text{"prob. } i^{\text{th}} \text{ coordinate is fixed} \end{aligned}$$

By linearity $\mathbb{E}[X] = \mathbb{E}[X_1 + \dots + X_n] \quad (X = X_1 + \dots + X_n)$.

$$= n \cdot \frac{1}{n} = 1.$$

$$\mathbb{E}[\# \text{ matches}] = 1.$$

EX.

Work out # of empty bins in FD, MB, BE distri

$$E[X^2], \quad E[X^3], \quad \dots$$

30/11 LECTURE 18: MORE ON RANDOM VARIABLES.

Eg 18.1: (Coupon Collector problem) There are n ^{distinct} coupons. Each day a coupon is selected indep. & UAR. All coupons are available for selection every day. (Sampling with replacement).

T = # of days needed to get all n ^{distinct} coupons. (r.v.)

Fix $m \geq 1$. $P(T > m)$ = Prob. that in m days, I haven't got

$P(T = m)$ = Prob. that I collect the m^{th} distinct coupon on

$$\{T = m\} = \{T > m-1, T \leq m\} = \{T > m-1\} \cap \{T > m\}^c$$

$$P(T = m) = P(T > m-1) - P(T > m)$$

$T > m$ - depends on the coupons received on the first m' days.

A_j = j^{th} coupon isn't selected at all on the first m' days.

$\{T > m\}$ = at least one coupon isn't selected at all - - - m' days

$$= \bigcup_{j=1}^n A_j^c$$

$$P(T > m) = P\left(\bigcup_{j=1}^n A_j^c\right) = S_1 - S_2 + S_3 - \dots + (-1)^{n+1} S_n$$

$$S_k = \sum_{i_1, \dots, i_k}^+ P(A_{i_1}^c \cap \dots \cap A_{i_k}^c)$$

$$P(A_i^c) = P(i^{\text{th}} \text{ coupon isn't selected on day 1, } \dots, m)$$

$$\text{(indep. of selections)} = \prod_{l=1}^m P(i^{\text{th}} \text{ on day } l)$$

$$= \prod_{l=1}^m \left(\frac{n-1}{n}\right) = \left(\frac{n-1}{n}\right)^m$$

$$P(A_i^c \cap A_j^c) = P(i^{\text{th}} \& j^{\text{th}} \text{ coupon aren't } \dots \text{ on days } 1, \dots, m)$$

$$= \left(\frac{n-2}{n}\right)^m$$

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = \left(\frac{n-k}{n}\right)^m$$

$$\Rightarrow S_k = \binom{n}{k} \left(\frac{n-k}{n}\right)^m$$

$$P(T > m) = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} \left(\frac{n-k}{n}\right)^m.$$



Job of prob - Compute prob. $\rightarrow$ Compute $E[f(x)]$.

Propn
18.2

(MARKOV'S INEQ.) Let X be a r.v. & $X \geq 0$ (i.e., $P_X(x) = 0 \forall x < 0$).

then $P(X \geq t) \leq \frac{EX}{t} \quad \forall t > 0$.

[Check if $X(\omega) \geq 0 \forall \omega \in \Omega$ then $P_X(x) = 0 \forall x < 0$].

Proof:

T.S.T.

$$t P(X \geq t) \leq E[X]$$

$$t E[\mathbb{1}_{[X \geq t]}] \leq E[X] \quad \square$$

linearity

$$\xrightarrow{\uparrow} E[t \mathbb{1}_{[X \geq t]}] \leq E[X]$$

Observe that $t \mathbb{1}[X \geq t] \leq X \quad \forall t > 0$.

$$\left[\begin{array}{l} X(\omega) \geq t \Rightarrow \text{LHS} = t \quad ; \quad \text{RHS} = X(\omega) \text{ \& LHS} \leq \text{RHS} . \\ X(\omega) < t \Rightarrow \text{LHS} = 0 \quad ; \quad \text{RHS} = X(\omega) \geq 0 \text{ \& LHS} \leq \text{RHS} \end{array} \right]$$

So by monotonicity of $\mathbb{E}$, we have that

$$\mathbb{E}[t \mathbb{1}[X \geq t]] \leq \mathbb{E}[X]$$

$$\& \text{ so } P(X \geq t) \leq \frac{\mathbb{E}[X]}{t} \quad \forall t > 0. \quad \square$$

Cor 18.3 Let X be a r.v. & let $f: \mathbb{R} \rightarrow \mathbb{R}_+ = (0, \infty)$ & f is non-decreasing.
Then $\forall t \in \mathbb{R}$

$$P(X \geq t) \leq \frac{\mathbb{E} f(X)}{f(t)} \quad \forall t \in \mathbb{R}.$$

Proof: If $X \geq t$ then $f(X) \geq f(t)$ as f is non-decreasing.
So $\{X \geq t\} = \{\omega: X(\omega) \geq t\} \subseteq \{f(X) \geq f(t)\}$

By monotonicity, $P(X \geq t) \leq P(f(X) \geq f(t))$
 (M. Ineq. as $f(X) \geq 0$) $\leq \frac{E[f(X)]}{f(t)}$ $\square$

Ex 18.4 let $f(x) = e^{sx}$, $s > 0$, $x \in \mathbb{R}$, $f(x) > 0$ & non-decreasing.

By above cor $P(X \geq t) \leq e^{-st} E[e^{sx}] \quad \forall t \in \mathbb{R}.$
 (Chernoff's bound).

Ex 18.5 let X be a r.v. $P(|X| \geq t) \leq \frac{E[X^2]}{t^2} \quad \forall t > 0.$

Ex 18.6: $\Omega = S_n$, $P(\sigma) = \frac{1}{n!}$ $\sigma \in S_n$.

Fix $k \in \{2, \dots, n\}$

$P(\exists \text{ a increasing subseq. of length } k) = ?$

Increasing subseq. of length k is $i_1 < i_2 < \dots < i_k$
 $\exists \sigma(i_1) < \sigma(i_2) < \dots < \sigma(i_k)$

Ex. $n=3, k=2$. $\sigma = (\sigma(1), \sigma(2), \sigma(3)) = (2, 3, 1)$.

only one incr. subseq. of length 2, $\sigma(1) < \sigma(2)$.

$\sigma = (1, 3, 2)$, 2 incr. subseq. of length 2 $\sigma(1) < \sigma(2)$

$\sigma = (1, 2, 3)$, 3 - - - of 2 $\sigma(1) < \sigma(3)$.

$X_k(\sigma) = \#$ of increasing subsequences of length k . (r.v.)

$P(X_k \geq m) = ?$ i.e., Prob. $\exists$ m increasing subsequence of length k .

$L(\sigma) =$ Length of longest increasing subsequence. (r.v.)
 $= \max \{ k : \exists \text{ an incr. subsequence of length } k \}$

$$L: S_n \rightarrow [n] \quad ; \quad X_k: S_n \rightarrow \{0\} \cup \left[\binom{n}{k} \right]$$

(check $X_k(\sigma) \leq \binom{n}{k} \forall \sigma \in S_n$)

$$L(\sigma) = \max \{k : X_k(\sigma) \geq 1\}$$

$$\{L(\sigma) \geq k\} = \{X_k(\sigma) \geq 1\}.$$

$$\{L \geq k\} := \{\sigma : L(\sigma) \geq k\} \quad \{X_k \geq 1\} = \{\sigma : X_k(\sigma) \geq 1\}$$

$$P(L \geq k) = P(X_k \geq 1)$$

$$\text{(Markov's)} \rightarrow \leq E[X_k].$$

ineq.

$$X_k(\sigma) = \# \{ (i_1, \dots, i_k) : i_1 < \dots < i_k \text{ \& \> } \sigma(i_1) < \dots < \sigma(i_k) \}$$

$$= \sum_{i_1 < \dots < i_k} \mathbb{1}[\sigma(i_1) < \dots < \sigma(i_k)]$$

$$X_{i_1, \dots, i_k}(\sigma) = \mathbb{1}[\sigma(i_1) < \dots < \sigma(i_k)]$$

Use linearity

$$\mathbb{E}[X_k] = \sum_{i_1 < \dots < i_k} \mathbb{E}[X_{i_1, \dots, i_k}]$$

(check)

$$= \sum_{i_1 < \dots < i_k} \mathbb{P}(\{\sigma : \sigma(i_1) < \dots < \sigma(i_k)\})$$

$$= \sum_{i_1 < \dots < i_k} \frac{|\{\sigma : \sigma(i_1) < \dots < \sigma(i_k)\}|}{n!}$$

(check)

$$= \sum_{i_1 < \dots < i_k} \frac{1}{k!} = \binom{n}{k} \frac{1}{k!}$$

$$\mathbb{P}(L \geq k) \leq \binom{n}{k} \frac{1}{k!} \quad \text{Find seq. } k_n \Rightarrow \text{RHS} \rightarrow 0 \text{ as } n \rightarrow \infty$$

Dan Romik — "The beautiful mathematics of LIS"

Exo 1.50 union bound on combato. $\mathbb{P}(V_n > 1)$.

Defn
18.7

Define variance of X ($\text{VAR}[X]$) by

$$\text{VAR}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2]$$

(STD. DEV.) $\text{SD}[X] := \sqrt{\text{VAR}[X]}$

Obs:

$$\text{VAR}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

$$\text{VAR}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\mathbb{E}[X] + (\mathbb{E}[X])^2]$$

$$\stackrel{\text{(lin. of exp.)}}{=} \mathbb{E}[X^2] - 2\mathbb{E}[X\mathbb{E}[X]] + \mathbb{E}[(\mathbb{E}[X])^2]$$

$$\stackrel{\downarrow \text{linearity}}{=} \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + (\mathbb{E}[X])^2$$

$$= \mathbb{E}[X^2] - (\mathbb{E}[X])^2. \quad \square$$

Exo If X is const r.v. (i.e. $P_X(a) = 1$ for some $a \in \mathbb{R}$) then $\mathbb{E}[X] = a$.

2/12 LECTURE 19: E, VAR & SUMMARY OF DATA.

X r.v. with pmf p_X . $\mu_X = E[X] = \sum_x x p_X(x)$ ($\mu = \mu_X$)
 $E[f(X)] = \sum_x f(x) p_X(x)$ $f: \mathbb{R} \rightarrow \mathbb{R}$ [Expectation/mean]
 $VAR[X] := E[(X - \mu)^2]$ (Variance of X).

$$= E[X^2 - 2\mu X + \mu^2]$$

$$= E[X^2] - \mu^2 \quad (\text{linearity})$$

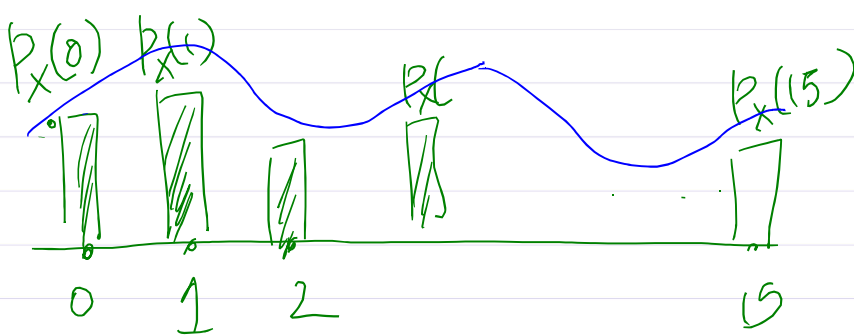
$$VAR[X] \geq 0$$

Alternative $E[(X - \mu)^2] = \sum_x (x - \mu)^2 p_X(x)$ (Take $f(x) = (x - \mu)^2$)

$$= \sum_x x^2 p_X(x) - \mu^2 = E[X^2] - \mu^2.$$

Ex 19.1 Assume X r.v. with pmf $p_X(x) = 0$ if $x \notin \{0, \dots, 15\}$.
 You are given $p_X(0), p_X(1), \dots, p_X(15)$.

Pictorially
 p_X is an
 histogram.



Can you summarise the histogram / r.v.?

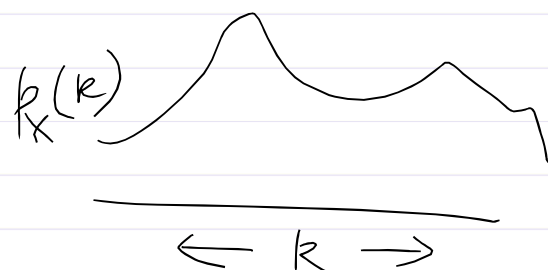
First summary is mean — $\mu_X = E[X] = \sum_x x p_X(x)$

Ex 19.2 Suppose I have scores of students in Quiz 1.

Define $p_X(k) = \frac{\# \text{ of scores} = k}{\# \text{ of students}}$. $\sum_{k=0}^{15} p_X(k) = 1$, $p_X(x) \geq 0$.

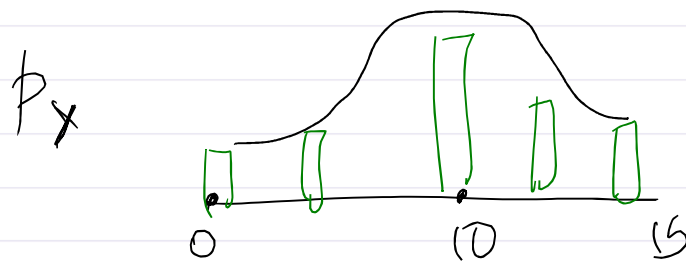
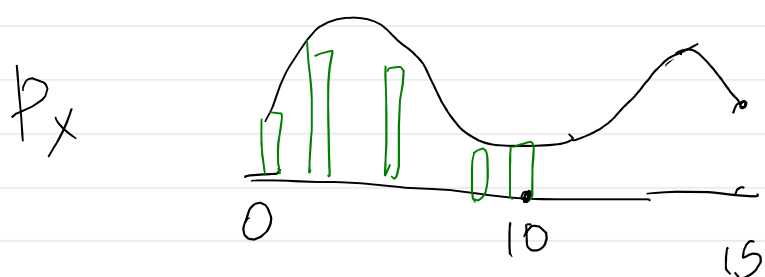
$\Rightarrow p_X$ is a pmf of some r.v. X .

(mean) $\mu_X = \sum_x x p_X(x)$ ($= 9.6\dots$)



Qn*: Find explicitly $(\Omega, \mathcal{P})$ & $X: \Omega \rightarrow \mathbb{R}$ $\exists$ $P(X=k) = p_X(k)$.

Qn: Suppose I have two pmfs p_x & p_y



Assume $\sum_x x p_x(x) = \mu_x = \mu_y = \sum_y y p_y(y)$.

$\text{VAR}[X] = \sum_x (x - \mu)^2 p_x(x) \Rightarrow$ larger variance $\Rightarrow$ larger "disparity"

LEMMA 19.3 (1) $\text{VAR}[X] = 0$ iff $p_x(\mu) = 1$ (i.e., $X \equiv \mu$)

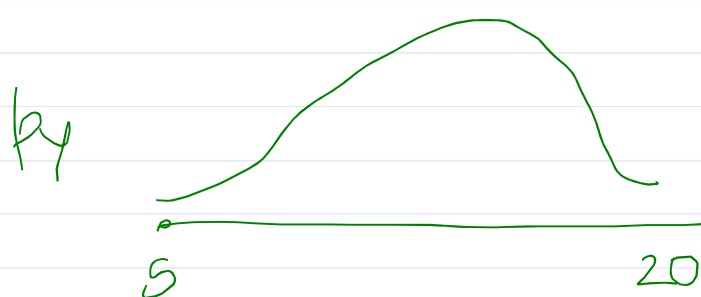
(2) $\text{VAR}[aX + b] = a^2 \text{VAR}[X] \quad \forall a, b \in \mathbb{R}$

(3) $\mu_{aX+b} = a\mu_x + b$. (recall)

[why not $\sum_x (x - \mu) p_x(x)$??]

Ex 19.4 (Again ex. of shift series). Suppose I add 5 points to all.

Then the new r.v. $Y = X + 5$. $p_y(x) = p_x(x - 5)$, $x \in \mathbb{R}$

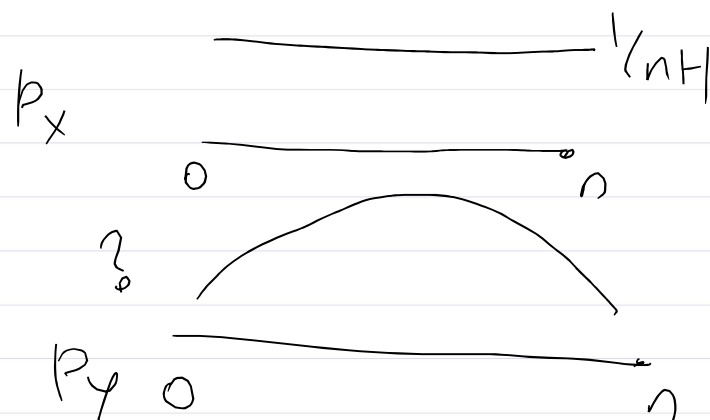


$\mu_y = \mu_x + 5$, $\text{VAR}[Y] = \text{VAR}[X]$

Ex 19.5 (A) $X \stackrel{d}{=} \text{Unif}\{0, 1, 2, \dots, n\}$. $\mathbb{E}[X] = \frac{n}{2}$

$Y \stackrel{d}{=} \text{Bin}(n, 1/2)$

$\mathbb{E}[Y] = \frac{n}{2}$



Compute $\text{VAR}(X)$, $\text{VAR}(Y)$.

[& what can you say about disparity?]

Further summaries - $E[X]$, $VAR[X]$, $E[X^k]$, $k \geq 1$.

For eg. if $X \stackrel{d}{=} \text{Ber}(p)$ then $X^k \stackrel{d}{=} \text{Ber}(p) \quad \forall k \geq 1$.

$$[P_X(1) = p = 1 - P_X(0)] \quad [\text{i.e., } P_{X^k}(1) = p = 1 - P_{X^k}(0)]$$

$$\Rightarrow E[X^k] = p \quad \forall k \geq 1.$$

EX
19.7

Let X be a r.v. $f: \mathbb{R} \rightarrow \mathbb{R}$ & define $Y = f(X)$.

Then Y is a r.v. with pmf $P_Y(\cdot) = ??$ (write it in terms of P_X).

EX
19.8

Sampling with replacement. Total population = N .

There are two types. # of type 1 = $Np \in \mathbb{N}$, $p \in [0, 1]$

of type 2 = $N(1-p) \in \mathbb{N}$.

Suppose you choose a sample of size n from the population at random & with replacement.

Let $X = \#$ of samples of type 1 in the n randomly chosen samples.

$$\begin{aligned} P_X(k) = P(X=k) &= \underset{(\text{check})}{\binom{n}{k}} \left(\frac{Np}{N}\right)^k \left(\frac{N(1-p)}{N}\right)^{n-k} \\ &= \binom{n}{k} p^k (1-p)^{n-k}. \end{aligned}$$

So X is a $\text{Bin}(n, p)$ r.v.

Qn. How to find p ?

Define $Y = \frac{X}{n}$. $P_Y\left(\frac{k}{n}\right) = \binom{n}{k} p^k (1-p)^{n-k} \quad k=0, \dots, n.$

$$E[Y] = \underset{(\text{linearity})^n}{E[X]} = p. \quad (\text{mean of } Y)$$

$$\underset{(\text{by LEMMA 19.6})}{VAR[Y]} = \underset{(A)}{\frac{VAR[X]}{n^2}} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

$$\text{VAR}[Y] = \sum_x (y - p)^2 p_y(y) = \frac{p(1-p)}{n} \leq \frac{1}{4n}$$

$$(p(1-p) \leq \frac{1}{4})$$

$$\text{VAR}[Y] \leq \frac{1}{4n}$$

$$\text{If } n = 2500, \text{ VAR}[Y] \leq 10^{-4}$$

Intuitively, " $|Y - p| \leq 10^{-4}$ ";
 If p is 10^{-6} then this is bad; if " p is small or large", $\frac{p(1-p)}{n}$ is very small

Ex
19.4
(A)

Repeat the above exercise without replacement.

Find p_y , $E[Y]$ & $\text{VAR}[Y]$.

$$\mu_y = E[Y]; \quad \sigma_y = \text{SD}[Y] := \sqrt{\text{VAR}[Y]}$$

LEMMA
19.10

(Chebyshev's Ineq.)

X r.v.

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2} \quad \forall t > 0.$$

Proof:

$$|X - \mu| \geq t \Rightarrow (X - \mu)^2 \geq t^2 \quad \forall t > 0$$

$$\text{So } \{ |X - \mu| \geq t \} \subseteq \{ (X - \mu)^2 \geq t^2 \}$$

$$\Rightarrow P(|X - \mu| \geq t) \leq P((X - \mu)^2 \geq t^2)$$

$$\begin{aligned} & \text{(M. Ineq.)} \\ & \text{as } (X - \mu)^2 \geq 0 \end{aligned} \quad \leq \frac{E[(X - \mu)^2]}{t^2} = \frac{\sigma_X^2}{t^2}$$

So Chebyshev's ineq. $\Rightarrow$

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \quad \forall k \geq 1$$

Apply to Eg. of sampling with replacement.

$$Y = \frac{X}{n}; \quad \mu_Y = p; \quad \sigma_Y = \sqrt{\frac{p(1-p)}{n}}$$

$$P(|Y - p| \geq K \sigma_Y) \leq \frac{1}{K^2}$$

"Prob. of Error in $|Y - p|$ is larger than $\frac{K \sqrt{p(1-p)}}{\sqrt{n}}$ is at most $\frac{1}{K^2}$!"

If $K=10$, then $|Y - p| \leq \frac{K \sqrt{p(1-p)}}{\sqrt{n}}$ with 99% accuracy.

Ex. Do this for sampling without replacement.

Find t s.t. $P(|Y - a| \geq t) \leq \frac{1}{100}$. ??

04/12 LECTURE - 20. DISCRETE PROB. SPACES.

Suppose Ω is a countably infinite sample space. $\left[\begin{array}{l} \Omega - \text{discrete} \\ \text{if } \Omega \text{ is count.} \end{array} \right]$
 Can we define prob. (i.e., pmf) on it?
 Ω is countable if $\exists$ a bijection from Ω to $A \subseteq \mathbb{N}$.
 (or $\exists$ an injection from Ω to $\mathbb{N}$).

If p is pmf on Ω then $p(\omega) \geq 0 \quad \forall \omega \in \Omega$
 & $\sum_{\omega \in \Omega} p(\omega) = 1$.

For eg. $\Omega = \mathbb{N}$. Need to define $p: \mathbb{N} \rightarrow [0,1] \rightarrow \sum p(n) = 1$.

Suppose $a_n := p(n)$. We want a seq. $a_n \geq 0 \rightarrow \sum_{n=1}^{\infty} a_n = 1$.
 More so, we only need $b_n \geq 0 \rightarrow \sum_{n=1}^{\infty} b_n < \infty$.

If such a b_n exists set $a_n = \frac{b_n}{b}$ where $b = \sum_{n=1}^{\infty} b_n$.

So $a_n \geq 0$ & $\sum_{n=1}^{\infty} a_n = 1$. Thus $p(n) = a_n$ is a valid pmf.
 When we write $\sum_{n=1}^{\infty} a_n$, there is an implicit order.

For any seq. (a_n) $\sum_{n=1}^{\infty} a_n := \lim_{m \rightarrow \infty} \sum_{n=1}^m a_n$ (if it exists)

Fact: (1) If $a_n \geq 0$. then $\sum_{n=1}^{\infty} a_n = \infty$ or $\sum_{n=1}^{\infty} a_n < \infty$.
 (2) Let $\sigma: \mathbb{N} \rightarrow \mathbb{N}$ be a bijection, $a_n \geq 0$ & $\sum_{n=1}^{\infty} a_n < \infty$.
 then $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\sigma(n)}$. — (1)

In this case we set $\sum_n a_n = \sum_{n=1}^{\infty} a_n$.

(3) Let $\sigma: \mathbb{N} \rightarrow \mathbb{N}$ be a bijection; a_n any sequence such that $\sum_{n=1}^{\infty} |a_n| < \infty$. Then (1) holds
 & $\sum_n a_n = \sum_{n=1}^{\infty} a_n$.

(4) Given a seq. a_n , define $a_n^+ := \max\{a_n, 0\} \geq 0$
 observe $a_n = a_n^+ - a_n^-$; $|a_n| = a_n^+ + a_n^-$ $\left\{ \begin{array}{l} a_n^+ := \max\{a_n, 0\} \geq 0 \\ a_n^- := \max\{-a_n, 0\} \geq 0 \end{array} \right.$
 & $a_n^+ a_n^- = 0$.

Check: (1) $\sum_{n=1}^{\infty} |a_n| < \infty$ iff $\sum_{n=1}^{\infty} a_n^+ < \infty$ & $\sum_{n=1}^{\infty} a_n^- < \infty$.
 (2) If $\sum_{n=1}^{\infty} |a_n| < \infty$ then $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_n^+ - \sum_{n=1}^{\infty} a_n^-$.

Defn 20.1 Let Ω be countable. Let $f: \Omega \rightarrow \mathbb{R}$.
 We say that $\sum_{\omega \in \Omega} f(\omega) = \sum_n f(\omega)$ is well-defined if
 $\sum_{\omega \in \Omega} |f(\omega)| := \sum_{n=1}^{\infty} |f(\sigma^{-1}(n))| < \infty$ where $\sigma: \Omega \rightarrow \mathbb{N}$ is an injection
 & set $\sum_{\omega} f(\omega) = \sum_{n=1}^{\infty} f(\sigma^{-1}(n))$. [$\sigma^{-1}(n)$ isn't defined $f(\sigma^{-1}(n)) = 0$]

From above facts the well-defined-ness is indep. of σ .

Defn 20.2 Given a sequence $\{a_n; n \geq 0\}$ we say that $\sum_n a_n$ is well-defined if $\sum_{n=1}^{\infty} |a_n| < \infty$
 & we set $\sum_n a_n := \sum_{n=1}^{\infty} a_n$.

Defn 20.3 p is a pmf on $\mathbb{N}^*$ if $p(n) \geq 0 \quad \forall n \geq 0$ & $\sum_n p(n) = 1$.
 ($\mathbb{N}^* := \mathbb{N} \cup \{0\}$) (i.e., $\sum_n p(n)$ is well-defined)

Eg 20.4 If $\Omega \subset \mathbb{N}^*$ & (Ω, p) is a PS. then p is a pmf on $\mathbb{N}^*$ (set $p(n) = 0 \quad \forall n \notin \Omega$).

Ex 20.5 (1) $p(n) = \frac{1}{2^n}; n \geq 1 \quad p(0) = 0$.

(2) $p(n) = p(1-p)^{n-1}, n \geq 1 \quad p \in (0, 1]$

$$(3) \quad p(n) = p(1-p)^n; \quad n \geq 0 \quad ; \quad p \in (0,1]$$

$$(4) \quad p(n) = \frac{e^{-\lambda} \lambda^n}{n!}; \quad n \geq 0 \quad ; \quad \lambda \geq 0. \quad (0^0 = 1).$$

Check all of the above are pmf on $\mathbb{N}^*$.

$$(5) \quad p(n) = \binom{n-1}{r-1} p^r (1-p)^{n-r} \quad n \geq 1 \quad ; \quad p \in (0,1], r \geq 1.$$

EXTRA EXTRA: If $f: \Omega \rightarrow \mathbb{R}_+ = [0, \infty)$

then define $\sum_{\omega} f(\omega) := \sup \left\{ \sum_{\omega \in A} f(\omega) : A \subset \Omega \right\}$

show that $\sum_{\omega} f(\omega) < \infty$ iff $\sum_{\omega} f(\omega) < \infty$ as in Defn 20.1.
(i.e., via injection to $\mathbb{N}$).

For more see Ch-6 of Manjunath Krishnapur's notes.

9/12 LECTURE 21 — DISCRETE PD.

p is a pmf on $\mathbb{N}^*$ if $p(n) \geq 0$ & $\sum_n p(n) = 1$.
 $[(\mathbb{N}^*, p) \text{ is a PS.}]$

Define $P(A) := \sum_{w \in A} p(w)$ $A \subseteq \mathbb{N}^*$ ①

Ex.
21.1

Check (i) $P(A) \geq 0$ $\forall A \subseteq \mathbb{N}^*$

(ii) $P(\mathbb{N}^*) = 1$.

(iii) $P(A_1 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i)$ (finite add.)

($A_1, \dots, A_n$ can also be ∞ sets)

P defined as in ① is called PD. (prob. distribution).

Qn: Suppose P satisfies (i), (ii) & (iii) then is there a pmf $p: \mathbb{N}^* \rightarrow [0,1]$ $\ni P(A) = \sum_{w \in A} p(w)$?

[we know ans is Yes if P corresponds to a finite prob. space
 i.e., $P(\Omega) = 1$ for some $\Omega \subset \subset \mathbb{N}^*$]

Approach to solution: Define $p(w) = P(\{w\})$.

Easily $p(w) \geq 0$.

Remains to check $\sum_w p(w) = 1 \iff \sum_{w \in \mathbb{N}^*} P(\{w\}) = 1$

(iii) $\Rightarrow \sum_{w=1}^n P(\{w\}) = P(\{1, \dots, n\}) \quad \forall n \geq 0$.

Not enough to show $\sum_{w \in \mathbb{N}^*} P(\{w\}) = 1$. X.

Defn 21.2 A fn $P: 2^{\mathbb{N}^*} \rightarrow \mathbb{R}$ is called a PD if

(i) $P(A) \geq 0 \quad \forall A \subseteq \mathbb{N}^*$

(ii) $P(\mathbb{N}^*) = 1$

(iii) $P(\bigsqcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ (count. add.)

$[A_i]'s \text{ are pairwise disjoint}]$.

Ex 21.3 check if P is a PD on N^* then $p(w) = P(\{w\})$ is a pmf on N^* .

Also check count add. $\Rightarrow$ fin. add.

PROP 21.4

Let (N^*, P) or (N^*, p) be a PS.

(1) $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$ (count. subadd; union bd; Boole's inv)

(2) $P(A) \subseteq P(B) \quad \forall A \subseteq B \subseteq N^*$ (Mon.)

(3) $P(B \setminus A) = P(B) - P(A) \quad \forall A \subseteq B \subseteq N^*$

(4) I-E holds for $P(A_1 \cup \dots \cup A_n) \quad \forall n \geq 1$

(5) Bonferroni ineq. hold for $P(A_1 \cup \dots \cup A_n) \quad \forall n \geq 1$.

Proof: Exo!! if the statement involves only finitely "many" sets, same proof as for fin. prob. spaces.

Else try using count. add.

LEMMA 21.5 Let $A_n \uparrow A$ i.e., $A_n \subseteq A_{n+1}$ & $\bigcup_{n=1}^{\infty} A_n = A$. (A_n increases to A)

Then $\lim_{n \rightarrow \infty} P(A_n) = P(A)$.

Holds also for $A_n \downarrow A$ i.e., $A_n \supseteq A_{n+1}$ & $\bigcap_{n=1}^{\infty} A_n = A$.

Proof: Assume $A_n \uparrow A \Rightarrow A_n \subseteq A_{n+1}$ & $A_n \subseteq A$

& so $P(A_n) \leq P(A_{n+1}) \leq P(A)$. (monotonicity of P)

$P(A_n)$ is an increasing seq. & bdd by $P(A)$.

$\Rightarrow \lim_{n \rightarrow \infty} P(A_n)$ exists by MCT. (monotone convergence thm)

Exo show that $\sup_{n \geq 1} P(A_n) = P(A)$. [use count. add.]

or show $P(A|A_n) \rightarrow 0$ as $n \rightarrow \infty$. ["]

Suppose $A_n \downarrow A$. Then $A_n^c \uparrow A^c$.

$$\Rightarrow \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} (1 - P(A_n^c))$$

$$= 1 - \lim_{n \rightarrow \infty} P(A_n^c) = 1 - P(A^c) = P(A)$$

$\Leftrightarrow (A_n^c \uparrow A^c)$.

RMK 21.5 Indep. of events, Condnl. prob. are all defined as before i.e.,

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

Events $A_1, \dots, A_n$ are indep. if

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j) \quad \forall J \subseteq [n].$$

Defn 21.6 X is a discrete r.v. (in $\mathbb{N}^*$) if $P(X \in \mathbb{N}^*) = 1$.
i.e., X takes values only in $\mathbb{N}^*$.

EQUIVALENTLY, $\exists$ a fn $p_X: \mathbb{R} \rightarrow [0,1]$ $\exists \sum_{n \in \mathbb{N}^*} p_X(n) = 1$.
& interpreted as $p_X(n) = P(X=n)$.

p_X is called as pmf of X .

$$\text{define } P(X \in A) = \sum_{x \in A} p_X(x) = \sum_{x \in A \cap \mathbb{N}^*} p_X(x)$$

$\hookrightarrow$ (as $p_X(x) = 0$ if $x \notin \mathbb{N}^*$)

Again when we say X is a r.v., we are also given a pmf p_X !

Ex 21.7

A p -biased coin is tossed repeatedly & independently until we get heads. ($p \in (0,1]$)
Let $X = \#$ of tosses.

$$p_X(n) = P(X=n) = P(\text{first } n-1 \text{ tosses are tails \& } n^{\text{th}} \text{ toss is heads}) \\ = (1-p)^{n-1} p \quad (\text{the tosses are indep. \& prob-heads} = p)$$

check $p_X(n) \geq 0 \forall n \geq 1$ & $\sum_n p_X(n) = 1$.

X with pmf $p_X(n) = (1-p)^{n-1} p, n \geq 1$ is called GEOMETRIC RV.
(Geom(p)).

check $\sum_n p_X(n) = \sum_n P(n-1 \text{ tails \& } n^{\text{th}} \text{ heads})$

(count-add) $= P\left(\bigcup_n \{n-1 \text{ tails \& } n^{\text{th}} \text{ heads}\}\right)$

$= P(\text{eventually there is an head})$

$P(X \geq n) = \sum_{k \geq n} p_X(k) = \dots$

$P(X \geq n) = P(\text{first } n-1 \text{ tosses are tails}) = (1-p)^{n-1}, n \geq 1.$

Eg 21.8 A p -biased coin is tossed repeatedly & indep. until we get ' r ' heads. ($p \in (0,1]$)

$X = \#$ of tosses

$p_X(n) = P(X=n)$

$= P(r^{\text{th}} \text{ head in } n^{\text{th}} \text{ toss})$

(since tosses are indep.) $= P(r-1 \text{ heads in } n-1 \text{ tosses \& } n^{\text{th}} \text{ toss is a head})$
 $= P(r-1 \text{ heads in } n-1 \text{ tosses}) P(n^{\text{th}} \text{ toss is a head})$

$= \binom{n-1}{r-1} p^{r-1} (1-p)^{n-r} \times p$

$= \binom{n-1}{r-1} p^r (1-p)^{n-r}$ → Binomial prob.

Check $\sum_{n \geq 1} p_X(n) = 1$. (X, p_X) is a NEG. BIN. RV
[NBIN(r, p)]

$$NBin(1, p) = \text{Geom}(p):$$

Eg
21.09

Let X be a $\text{Geom}(p)$ r.v. i.e., $P_X(n) = (1-p)^{n-1} p$, $\forall n \geq 1$.

$$P(X > k+l | X > k) = ?$$

$$\{X > k\} = \{\text{no heads until } k\}$$

$$P(X > k+l | X > k) = P(\text{no heads until } k+l | \text{no heads until } k)$$

$$P(X > k+l | X > k) \stackrel{\text{def. of CP}}{=} \frac{P(\{X > k+l\} \cap \{X > k\})}{P(X > k)}$$

$$= \frac{P(X > k+l)}{P(X > k)} \stackrel{\text{Geom}(p)}{=} \frac{(1-p)^{k+l}}{(1-p)^k} = (1-p)^l$$

$$\Rightarrow P(X > k+l | X > k) = P(X > l) \text{ [Memoryless ness of Geom}(p) \text{ r.v.]}$$

$$\Leftrightarrow P(X > k+l) = P(X > k) P(X > l)$$

Eg
21.10

A p -biased coin is tossed repeatedly until we get $(0,1)$ or $(1,0)$,
if we stop with (T,H) , set $Y = 0$ & indep. T, H or H, T .

(H,T) , set $Y = 1$, [Y is a new r.v.]

What is $P_Y(\cdot) = ?$

$$P_Y(0) = P(\text{ends with } (T,H)) \xrightarrow{B}$$

(count. add.)

$$= \sum_{n=2}^{\infty} P(\text{we end at } n \text{ with } (T,H)) \xrightarrow{B_n}$$

$$= \sum_{n=2}^{\infty} P(n-1 \text{ Tails \& } n^{\text{th}} \text{ heads}) \stackrel{\text{(indep.)}}{=} p \sum_{n=2}^{\infty} (1-p)^{n-1}$$

$$= (1-p).$$

$X = \#$ of tosses; compute $P_X(n)$.

10/12 LECTURE 22. DISCRETE RV & EXPECTATIONS

Eg 22.1

I toss two indep. coins (p -biased) repeatedly until I get HT or TH.

Two coins have same value. we toss again. if they have diff. value, we stop.

$Y = 0$ if HT

$Y = 1$ if TH.

$P_Y(0) = P(\text{we end with HT})$

(count. add.) $= \sum_{n=1}^{\infty} P(\text{HT or TH in the first } n-1 \text{ tosses} \text{ \& HT in the } n^{\text{th}} \text{ toss})$

$= \sum_{n=1}^{\infty} P(\text{HT or TH})^{n-1} P(\text{HT})$ (by indep. of tosses)

(by indep.) $= \sum_{n=1}^{\infty} (1 - 2pq)^{n-1} pq$ $q = 1 - p$
($p^2 + q^2 + 2pq = 1$)

$P(\text{HT}) = P(\text{coin 1 is H \& coin 2 is T}) = pq$ (indep.)
 $P(\text{HT or TH in } n \text{ tosses}) = P(\text{coin 1 = coin 2 in the first } n-1 \text{ tosses})$

so $P_Y(0) = \frac{pq}{2pq} = \frac{1}{2}$.

so $P_Y(1) = \frac{1}{2} = 1 - P_Y(0)$.

$Y \stackrel{d}{=} \text{Unif}\{0, 1\} \stackrel{d}{=} \text{Ber}(1/2)$ (unbiased coin!!)

Ex* Using p -biased coin, can you define $\text{Unif}\{1, \dots, K\}$ r.v.s?

$X = \# \text{ tosses before HT or TH}$

$P_X(m) = 2pq (1 - 2pq)^{m-1}$. $X \stackrel{d}{=} \text{Geom}(2pq)$.

Defn 22.2. Let (X, P_X) be a discrete r.v with values in $\mathbb{N}^*$.

Let $f: \mathbb{N}^* \rightarrow \mathbb{R}$.

Define $E[f(X)] := \sum_{x \in \mathbb{N}^*} f(x) P_X(x)$ if $\sum_{x \in \mathbb{N}^*} |f(x)| P_X(x) < \infty$

(Exp. of $f(x)$ or mean of $f(x)$ or Avg. of $f(x)$)

else $\vdash[H(x)]$ is undefined.

Recall if $\sum_{x \in \mathbb{N}^*} |g(x)| < \infty$ then $\sum_{x \in \mathbb{N}^*} g(x)$ is well-defined

$$\sum_{n=0}^{\infty} |g(n)|$$

$$\sum_{x \in \mathbb{N}^*} g(x) = \sum_{n=0}^{\infty} g(n)$$

Again given discrete r.v. X , important to compute

$E[X]$ - mean, $VAR[X] = E[X^2] - (E[X])^2$ - VARIANCE
 R^{th} moment of X - $E[X^R]$

we can describe a r.v. via its moments !!!

So $E[X]$ is well-defined if $\sum_x |x| p_X(x) = \sum_x x p_X(x) = \sum_{n=0}^{\infty} n p_X(n) < \infty$,
 $\text{VAR}[X]$ is if $\sum_x x^2 p_X(x) = \sum_{n=0}^{\infty} n^2 p_X(n) < \infty$

so if $E[X^k]$ is well-defined if $\sum_{n=1}^{\infty} n^k p_X(n) < \infty$.

Ex 22.3 If $E[X^m]$ is well-defined for some $m \in \mathbb{N}$, (X is $\mathbb{N}^*$ -valued) show that $E[X^k]$ is well-defined $\forall k \leq m$.

Eg
22.4

$X \stackrel{d}{=} \text{Geom}(p)$. $p_X(k) = p(1-p)^{k-1}$ $(k \geq 1)$. $p > 0$

$E[X]$ is well-defined if $\sum_{n=1}^{\infty} n p_X(n) < \infty$

$$\sum_{n=1}^{\infty} n p_X(n) = p \sum_{n=1}^{\infty} n (1-p)^{n-1} \stackrel{\text{check}}{=} \frac{1}{p} < \infty.$$

so $E[X]$ is well-defined & $E[X] = \sum_{n=1}^{\infty} n p_X(n) = \frac{1}{p}$.

$$\sum_{n=1}^{\infty} n(1-p)^{n-1} = \frac{d}{dp} \left(\sum_{n=0}^{\infty} (1-p)^n \right)$$

Compute $\text{VAR}[X]$, $\mathbb{E}[X^k]$, $k \geq 2$

Ex 22.5 $X \stackrel{d}{=} \text{NBin}(r, p)$ $P_X(m) = \binom{m-1}{r-1} p^r (1-p)^{m-r}$
 — r # of tosses until r heads

$$\begin{aligned}
 \sum_{n=1}^{\infty} n \binom{n-1}{r-1} p^r (1-p)^{n-r} &= p^r \sum_{n=1}^{\infty} n \binom{n-1}{r-1} (1-p)^{n-r} \\
 &= r p^r \sum_{n=1}^{\infty} \binom{n}{r} (1-p)^{n-r} \\
 &= \frac{r}{p} \sum_{n=1}^{\infty} \binom{n}{r} p^{r+1} (1-p)^{n-r} \\
 &= \frac{r}{p} < \infty. \quad \left(\begin{array}{l} \text{pmf of} \\ \text{another NBin} \end{array} \right)
 \end{aligned}$$

And so $E[X] = \frac{r}{p}$ (Avg. waiting for r heads

$\text{VAR}[X], E[X^k] = ? \quad k \geq 2.$ = Avg. waiting for 1st head + ... + ... (rth head)

Eg 22.6

Let us take a queue. (Q at a bus-stop, railway station) or internet servers

Customers arrive in the Q at rate λ , $\lambda \in (0, \infty)$

(1) Avg. # of customers arriving btw $[t, t+s] = \lambda s$

(2) No two customers arrive close to each other.

↳ # of customers arriving in $[t, t+s] \in \{0, 1\}$ if s is very small

(3) Customers arrive indep.

i.e., # in $[a, b]$ & # in $[c, d]$ are indep. if $[a, b] \cap [c, d] = \emptyset$

$P(\# \text{ in } [a, b] = k, \# \text{ in } [c, d] = j) = P(\# \text{ in } [a, b] = k) P(\# \text{ in } [c, d] = j) \quad \forall k, j$

For eg. $a = 9 \text{ AM}$, $b = 10 \text{ AM}$, $c = 11 \text{ AM}$, $d = 12 \text{ AM}$

in 9-10 AM doesn't affect # in 11-12 AM.

Fix to $N(t) = \# \text{ customers arriving by time } t$
 $= \# \text{ in } [0, t].$

So $N(t)$ is a R.V. taking values in $\mathbb{N}^*$

What is a good pmf for $N(t)$? $P_X(k) = ? \quad X = N(t)$

let n be large.

$$0 \quad \lambda_n \quad \lambda_n \quad \lambda_n \quad \lambda_n \quad \dots \quad \lambda_n t \quad ; \quad t \in \mathbb{N}$$

$Y_i = \#$ of customers arriving in interval $[\frac{i-1}{n}, \frac{i}{n}]$ $i \leq tn$

$\#$ of intervals $= tn$. since $t \in \mathbb{N}$.

All intervals are eq. length.

By (2), $Y_i \in \{0, 1\}$ two customers can't arrive in $[\frac{i-1}{n}, \frac{i}{n}]$.

By (4), $E[Y_i] = \frac{\lambda}{n} \Rightarrow Y_i \stackrel{d}{=} \text{Ber}(\frac{\lambda}{n})$. } in any i

As if each interval, we are tossing a coin with prob. $\frac{\lambda}{n}$.

(5) $\Rightarrow$ coin tosses are indep.

$N(t) \approx \#$ of heads in tn many coin-tosses
with prob. $\frac{\lambda}{n}$

Fix k
 $P(N(t) = k) \approx P(X_n = k)$ [think what happens if $t \notin \mathbb{N}$]

$$= \binom{nt}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{nt-k} = a_n$$

$$\lim_{n \rightarrow \infty} a_n = \frac{(nt)!}{(nt-k)!} \frac{1}{n^k} \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_\downarrow e^{-\lambda} \frac{\lambda^k}{k!} \frac{1}{\underbrace{\left(1 - \frac{\lambda}{n}\right)^k}_{\rightarrow 1 \text{ as } n \rightarrow \infty}}$$

$$\approx \frac{\lambda^k}{k!} e^{-\lambda t}$$

$$\frac{(nt)!}{(nt-k)!} \frac{1}{n^k}$$

$$\approx \frac{\lambda^k}{k!} e^{-\lambda t}$$

$$\frac{nt (nt-1) \dots (nt-k+1)}{n \cdot n \cdot \dots \cdot n}$$

$$\approx \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

$$\frac{1}{\lambda t} \left(1 - \frac{1}{nt}\right) \left(1 - \frac{2}{nt}\right) \dots \left(1 - \frac{k-1}{nt}\right)$$

$$\rightarrow \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

$$P(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k \geq 0$$

$$p_X(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad \text{Check this is a pmf}$$

& X is called Poisson (λt) r.v.. Ex. $E[X]$, $\text{VAR}[X]$.

14/12 LECTURE 24 : CDF & PROPERTIES.

Eg 4.1 $N = \#$ of customers arriving by time 1 i.e., $\#$ in $[0,1]$. [Poisson RV]

Reasonable to assume $N \stackrel{d}{=} \text{Poi}(\lambda)$ i.e., $P_N(k) = \frac{\lambda^k e^{-\lambda}}{k!}$, $k \geq 0$, $\lambda > 0$.
(Poisson)

$$\text{Check } \sum_{k=0}^{\infty} P_N(k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{\lambda} e^{-\lambda} = 1.$$

$$\begin{aligned} E[N] &= \sum_{n=0}^{\infty} n P_N(n) = \sum_{n=0}^{\infty} n \frac{\lambda^n}{n!} e^{-\lambda} = \lambda \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda} \\ &= \lambda \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} = \lambda. \end{aligned}$$

compute $\text{VAR}[N]$; $E[N^k]$, $k \geq 1$.

Eg 4.2 N is same as above. Customers are happy w.p. p & {indep of
not happy w.p. $(1-p)$ & {others.
 $H = \#$ of happy customers. $P_H(\cdot) = ?$
 $E[H] = \lambda p$. (guess)

$$P_H(k) = P(H=k)$$

$$= \sum_{n=0}^{\infty} P(\{H=k\} \cap \{N=n\}) \quad (\text{law of total prob.})$$

$$(\Omega = \bigsqcup_{n=0}^{\infty} \{N=n\})$$

$$P(\{H=k\} \cap \{N=n\}) = P(H=k | N=n) P(N=n) \quad (\text{conditional prob formula})$$

$$\underbrace{\# \text{ h-cust}}_{=k} \quad \underbrace{\# \text{ cust}}_{=n} = P(\# \text{ h-cust} = k \mid \text{Total \# of cust} = n) P_N(n).$$

$$= P(\text{Bin}(n, p) = k) P_N(n)$$

$$= \binom{n}{k} p^k (1-p)^{n-k} e^{-\lambda} \frac{\lambda^n}{n!} \quad (\text{pmfs of Bin \& Poi})$$

$$= \frac{e^{-\lambda p} (\lambda p)^k}{k!} \frac{(\lambda(1-p))^{n-k} e^{-\lambda(1-p)}}{(n-k)!} \quad \text{--- (1)}$$

Note that $P(\{H=k\} \cap \{N=n\}) = 0$ if $n < k$.

$$\begin{aligned}
\text{So } P_H(k) &= \sum_{n=k}^{\infty} P(\{H=k\} \cap \{N=n\}) \\
&= \frac{e^{-\lambda p} (\lambda p)^k}{k!} \sum_{n=k}^{\infty} \frac{(\lambda(1-p))^{n-k}}{(n-k)!} e^{-\lambda(1-p)} \\
&= \frac{e^{-\lambda p} (\lambda p)^k}{k!} \quad H \stackrel{d}{=} \text{Poi}(\lambda p) \\
&\quad \Downarrow \\
&\quad E[H] = \lambda p.
\end{aligned}$$

$H' = N - H$ = # not happy customers.

$$H' \stackrel{d}{=} \text{Poi}(\lambda(1-p))$$

Fix k, j

$$\begin{aligned}
P(H=k, H'=j) &= P(\{H=k\} \cap \{H'=j\}) \\
&= P(\{H=k\} \cap \{N=k+j\}) \quad (N=H+H') \\
&= \frac{(\lambda p)^k}{k!} e^{-\lambda p} \frac{(\lambda(1-p))^j}{j!} e^{-\lambda(1-p)} \quad (\text{from (1)})
\end{aligned}$$

$$\begin{aligned}
&= P(H=k) P(H'=j). \quad [\text{Thinning prop.}] \\
&\text{If } N \stackrel{d}{=} \text{Poi}(\lambda), \text{ then } \forall k, j \text{ \& } p \in [0,1] \\
&P(H=k, H'=j) = P(H=k) P(H'=j). \quad \text{--- (2)}
\end{aligned}$$

Ex** Suppose N is a r.v. with values in $\mathbb{N}^*$
 $\forall k, j$ & $p \in [0,1]$, N satisfies (2). Is N a Poisson r.v.?

Defn: [Cumulative Distribution Function (CDF)]

Let X be a r.v. with pmf p_X . CDF of X is a function $F_X: \mathbb{R} \rightarrow [0,1]$ & defined as

$$F_X(x) := P(X \leq x) = \sum_{y: y \leq x} p_X(y)$$

[Note: Given a pmf $p_X: \mathbb{R} \rightarrow [0,1]$, CDF exists].

($\sum_x p_x(x) = 1 \Rightarrow p_x \neq 0$ on at most countably many values)

For eg. $p_x(x) = \frac{1}{x}, x > 0$. $p_x(x) \rightarrow 0$ as $x \rightarrow \infty$.

But $\sum_x p_x(x)$ isn't defined!

If we write $\sum_x p_x(x)$, it implies^x that $p_x \neq 0$ on at most^a countable sets

EG 24.4 (1) $X \stackrel{d}{=} \text{Unif}(n)$ $p_x(k) = \frac{1}{n}, k \in \{1, \dots, n\}$

$$F_x(x) = \begin{cases} 0 & -\infty < x < 1 \\ 1/n & 1 \leq x < 2 \\ 2/n & 2 \leq x < 3 \\ \vdots & \\ (n-1)/n & (n-1) \leq x < n \\ 1 & n \leq x < \infty \end{cases}$$

(2) $X \stackrel{d}{=} \text{Bern}(p)$ $p_x(0) = 1-p = 1-p_x(1)$.

$$F_x(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

THM 24.5 (Prop. of CDF). Let X be ar.v. with CDF F_x .

(1) $F_x(x) \leq F_x(y) \quad \forall \quad x \leq y$ (Monotonicity)

(2) $\lim_{x \rightarrow \infty} F_x(x) = 1$; $\lim_{x \rightarrow -\infty} F_x(x) = 0$.

(3) $F_x(x+) = \lim_{y \downarrow x} F_x(y) = F_x(x)$. $\forall x \in \mathbb{R}$ (F_x is R.C.)

(4) $p_x(x) = P(X=x) = F_x(x) - F_x(x-)$
 $= F_x(x) - \lim_{y \uparrow x} F_x(y) \quad \forall x \in \mathbb{R}$

$$(5) \quad P(x < X \leq y) = F_X(y) - F_X(x), \quad x < y.$$

Proof

$$(1) \quad F_X(x) = \sum_{z: z \leq x} p_X(z) \leq \sum_{z: z \leq y} p_X(z) = F_X(y).$$

$\downarrow$ (defn of CDF) $\uparrow$ ($p_X \geq 0$) $\uparrow$ (defn of CDF)

$$(2) \quad F_X(x) = P(X \leq x).$$

Take $x_n \rightarrow \infty$ as $n \rightarrow \infty$. ETPT $\lim_{n \rightarrow \infty} F_X(x_n) = 1$.

$$\Rightarrow \lim_{x \rightarrow \infty} F_X(x) = 1$$

Given $\{x_n\}$, pick subsequence $y_n \nearrow y_n \uparrow \infty$.

$$F_X(y_n) = P(X \leq y_n) = P(X \in (-\infty, y_n]) \leq P(X \in (-\infty, y_{n+1}])$$

Let $A_n = \{X \in (-\infty, y_n]\}$. Then since $y_n \uparrow \infty$

$$A_n \uparrow A = \bigcup_{n \geq 1} A_n = \bigcup_{n \geq 1} \{X \in (-\infty, y_n]\} = \{X \in \bigcup_{n \geq 1} (-\infty, y_n]\}$$

$$[\text{Proven before in class}] = \{X \in \mathbb{R}\} \text{ as } y_n \uparrow \infty$$

By C. add. $P(A_n) \uparrow P(A) = P(X \in \mathbb{R}) = 1$.

Let $Y = -X$.

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(-Y \leq x) \\ &= P(Y \geq -x) \leq P(Y > -x) \\ &= 1 - P(Y \leq -x) \\ &= 1 - F_Y(-x) \end{aligned}$$

$\hookrightarrow$ monotonicity

So $F_X(x) \leq 1 - F_Y(-x)$

$$0 \leq \lim_{x \rightarrow \infty} F_X(x) \leq 1 - \lim_{x \rightarrow \infty} F_Y(-x) = 1 - \lim_{y \rightarrow \infty} F_Y(y) = 0.$$

$\nearrow$ (since limit exists.)

Since $\lim_{x \rightarrow -\infty} F_X(x) = 0$ & $F_X \geq 0$,
 $\lim_{x \rightarrow -\infty} F_X(x) = 0$.

by prev. part.

(3) T.S.T. $\lim_{y \downarrow x} F_X(y) = F_X(x)$.

Let $y_n \downarrow x$ as $n \rightarrow \infty$. $A_n = \{X \leq y_n\} \downarrow A = \bigcap_{n \geq 1} \{X \leq y_n\}$

Again we have shown that $P(A_n) \downarrow P(A)$ as $n \rightarrow \infty$. $\left\{ \begin{array}{l} = \{X \leq y_n, n \geq 1\} \\ = \{X \leq x\} \end{array} \right.$

$\Rightarrow F_X(y_n) = \underset{\text{defn}}{P(A_n)} \underset{\text{proven above}}{\downarrow} \underset{\text{defn}}{P(A)} = F_X(x)$.

(4) T.S.T. $P_X(x) = F_X(x) - F_X(x-)$

Let $y_n \uparrow x$ as $n \rightarrow \infty$.
 $(y_n < x, \forall n \geq 1)$ $A_n = \{X \leq y_n\} \uparrow A = \bigcup_{n \geq 1} \{X \leq y_n\}$
 $= \{X \leq y_n \text{ for some } n \geq 1\} = \{X < x\}$

Again $F_X(y_n) = P(A_n) \uparrow P(A) = P(X < x)$.

$\Rightarrow \lim_{y \uparrow x} F_X(y) = F_X(x-) = P(X < x)$.

So $F_X(x) - F_X(x-) = P(X \leq x) - P(X < x)$
 $= P(\{X \leq x\} \setminus \{X < x\})$
 $= P(X = x) = P_X(x)$.

[observe F_X is cts at x iff $P_X(x) = 0$]

(5) $P(x < X \leq y) = P(X \in (x, y])$, $x < y$.

$$\begin{aligned}
 &= P(\{X \in (-\infty, y]\} \setminus \{X \in (-\infty, x]\}) \\
 &= P(X \in (-\infty, y]) - P(X \in (-\infty, x]) \\
 &\quad \quad \quad ((-\infty, x] \subseteq (-\infty, y]) \\
 \text{(defn of CDF)} &= F_X(y) - F_X(x).
 \end{aligned}$$

□

Check:

$$\begin{aligned}
 P(x \leq X \leq y) &= F_X(y) - F_X(x) + p_X(x) = F_X(y) - F_X(x-) \\
 P(x \leq X < y) &= F_X(y) - F_X(x) + p_X(x) - p_X(y) \\
 P(x < X < y) &= F_X(y) - F_X(x) - p_X(y) = F_X(y-) - F_X(x)
 \end{aligned}$$

Cor. 21.6 Let X be a ^{discrete} r.v. with CDF F_X . Then the pmf p_X is given by $p_X(x) = F_X(x) - F_X(x-) \quad \forall x \in \mathbb{R}$.

Defn. 21.7. A fn. $F: \mathbb{R} \rightarrow \mathbb{R}$ satisfying the following 3 conditions is called a DF (distribn. fn.)

- (1) $F(x) \leq F(y) \quad \forall x \leq y$ (Monotonicity)
- (2) $\lim_{x \rightarrow \infty} F(x) = 1; \quad \lim_{x \rightarrow -\infty} F(x) = 0$ (∞ & $-\infty$ limits)
- (3) $F(x+) = \lim_{y \downarrow x} F(y) = F(x)$. (Right cty)

• CDF F_X of a discrete r.v. X is a DF!

Ex: If X is a discrete r.v. with values in $\mathbb{N}^*$ then

$$F_X(x-) = \sum_{y: y < x} p_X(y) = \begin{cases} F_X(x-1) & \text{if } x \in \mathbb{N}^* \\ F(1, x, 1) & x \notin \mathbb{N}^* \end{cases}$$

$\lfloor x \rfloor$ (L.S.) $\rightarrow$ floor of x .

16/12 LECTURE 25 - CTS RV - INTRO.

Discrete prob. space Ω - countable set. Define $p(w) \geq 0, w \in \Omega$
 $\exists \sum_w p(w) = 1.$

If Ω is uncountable what is $\sum_w p(w)$?

Eg
25.1

$\Omega = \{0,1\}^{\mathbb{N}}$ - ∞ 'ly many coin tosses. $w \in \Omega$ $w = (w_1, w_2, \dots)$

WANT TO DEFINE P - corresponding to ∞ 'ly many fair coin tosses.

$$A_n = \{w \in \{0,1\}^{\mathbb{N}} : w_1 = \dots = w_n = 1\}$$

$$P(A_n) \stackrel{?}{=} 2^{-n}$$

But is there a $p \exists P(A_n) = \sum_{w \in A_n} p(w)$?

If $\exists p$, $p(w) \stackrel{?}{=} P(\{w\}) \quad \forall w \in \Omega$

[For eg. take $w = (1, \dots, 1, \dots)$
 $P(\{w\}) \leq P(A_n) \quad \forall n$]

But $P(A_n) \neq 0$ $\Rightarrow 0$ (intuitive guess).

Eg
25.2

$\Omega = [0,1]$. Pick a random point from Ω .

Let X be the random point

$$P(X \in [a,b]) \stackrel{?}{=} b-a \quad 0 \leq a < b \leq 1.$$

$$\text{as } P(X \in [a,b]) = \int_a^b P(X=x) dx \quad ?$$

If we take $P(X=x) \stackrel{?}{=} 1 \quad \forall x$, then ok?? [Not Possible]

Intuition $\Rightarrow P(X=x) \leq P(X \in (x-\epsilon, x+\epsilon)) \stackrel{?}{=} 2\epsilon, \quad \forall \epsilon > 0$
 $\Rightarrow P(X=x) = 0.$

$$\text{So } P(X \in [a,b]) = \int_a^b f(x) dx \quad \text{but } P(X=x) \neq f(x) \quad ?$$

$\rightarrow$ We'll not define prob. on uncountable spaces but just focus on RV's on $\mathbb{R}$ taking uncountably many values.

ET to define RV's & not Prob. spaces. (so not a restriction)

RESTRICTION — $\Omega = \mathbb{R}!$

Why? we know properties of functions on $\mathbb{R}$ — Ctx, diff & integration

Defn 25.3 A DF is a fn $F: \mathbb{R} \rightarrow [0,1]$ s.t.

(i) $F(s) \leq F(t) \quad \forall s \leq t$ (increasing)

(ii) $\lim_{s \rightarrow -\infty} F(s) = 0 = 1 - \lim_{s \rightarrow \infty} F(s)$ (0-1 limits)

(iii) $F(t+) = F(t)$ (R.C.)

Cor 25.4. If X is a discrete r.v. taking values in $\mathbb{R}$ (i.e. $p_X(x) \geq 0, \sum_x p_X(x) = 1$) then the CDF of X , F_X is a DF.

$$(F_X(x) := \sum_{y \leq x} p_X(y))$$

(done in last class)

$$- p_X(x) = F_X(x) - F_X(x-)$$

Cor 25.5 Let F be a DF & assume that $\{x: F(x) - F(x-) \neq 0\}$ is countable. Let $-\infty < x_{-2} < x_{-1} < 0 < x_0 < x_1 < x_2 < \dots$

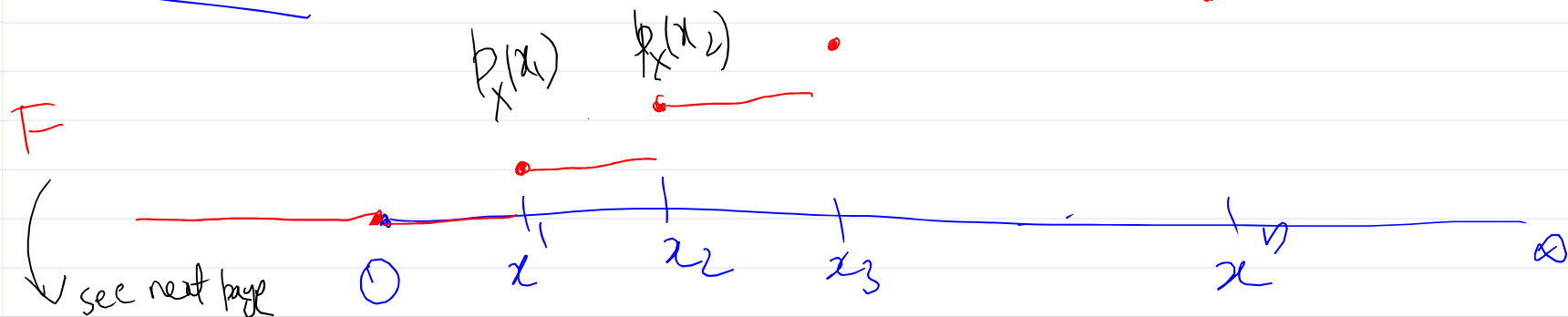
We also assume that $F(x) = F(y)$ if $x, y \in (x_k, x_{k+1}) \quad \forall k \in \mathbb{Z}$.

(i.e., F is a step function with jumps at $x_k, k \in \mathbb{Z}$)

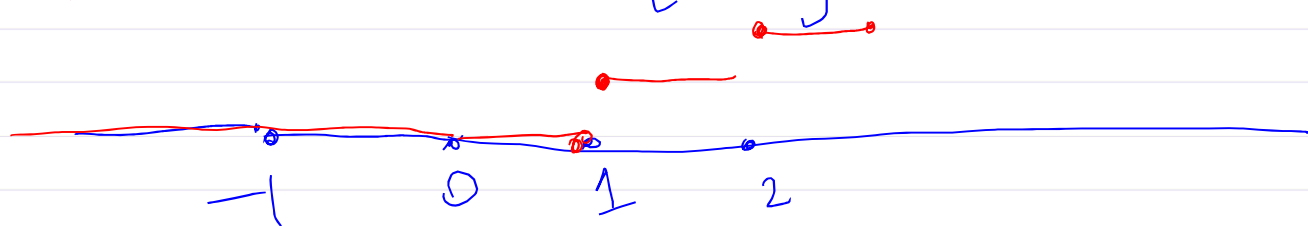
Then $\exists$ a discrete r.v. X s.t. $F_X = F$.

For simplicity assume there are no jumps in $(-\infty, 0)$

Proof in pictures: Let $x_1, \dots, x_n, \dots$ be $F(x_i) - F(x_i-) \neq 0$.



Eg. $F(x) = \max\{\lfloor x \rfloor, 0\}$ floor of x .



is a step fn
with countably
many jumps.

Contd. $P_X(x_i) := F(x_i) - F(x_{i-1}^-)$.

$F \uparrow \Rightarrow P_X(x_i) \geq 0$.

We need to check $\sum_{x_i} P_X(x_i) = 1$. $[EX_0]$

S.T. $\sum_{i=1}^{\infty} P_X(x_i) = 1 - F(0) ; \sum_{i=1}^{\infty} P_X(x_{-i}) = F(0)$

$\Rightarrow$ DF is a step function with countably many jumps
then it is CDF of a discrete r.v. !

We can use DF's to define discrete r.v. !
as any DF, CDF of a discrete r.v. ?

Eg.
25.6

$$F(s) = \begin{cases} 0 & s < 0 \\ s & 0 \leq s \leq 1 \\ 1 & s \geq 1 \end{cases}$$

Check F is a DF. But $\nexists$ no discrete r.v. $X \ni F_X = F$!

Defn
25.7

Let $f: \mathbb{R} \rightarrow [0, \infty)$ be a piece-wise continuous function

(i.e., discontinuous at finitely many points)

$\ni \int_{-\infty}^{\infty} f(x) dx = 1$. Such a f is called a

prob. density function (pdf).

Say f is disks at $a_1, a_2, \dots, a_k$, $a_1 < a_2 < \dots < a_k$.

then $\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{a_1} f(x) dx + \int_{a_1}^{a_2} f(x) dx + \dots + \int_{a_k}^{\infty} f(x) dx$.

on each (a_i, a_{i+1}) , f is cts & so $\int_{a_i}^{a_{i+1}} f(x) dx$ is well-defined!

Ex 25.8 (1) $f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases}$; f is a pdf

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^{\infty} f(x) dx$$

$\parallel$ $\parallel$ $\parallel$
 0 1 0

$\Rightarrow f$ is a pdf.

(2) $f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$, $\lambda > 0$.

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^{\infty} \lambda e^{-\lambda x} dx$$

$$= 0 + \lambda \cdot \frac{1}{\lambda} = 1$$

$\Rightarrow f$ is a pdf.

LEMMA 25.9 If f is a pdf, then $F(x) := \int_{-\infty}^x f(t) dt$, $x \in \mathbb{R}$ is a DF.

Proof (i) Let $x \leq y$. $F(x) = \int_{-\infty}^x f(t) dt \leq \int_{-\infty}^y f(t) dt = F(y)$

$\uparrow$ $\uparrow$ $\uparrow$
 def $f \geq 0$ def

(ii) $\lim_{x \rightarrow \infty} F(x) = \lim_{x \rightarrow \infty} \int_{-\infty}^x f(t) dt = \int_{-\infty}^{\infty} f(t) dt = 1$.

$\uparrow$ $\uparrow$ $\uparrow$
 def $f \geq 0$ & def of $\int$ f is pdf.

$$(ii) \lim_{x \rightarrow -\infty} F(x) = 0.$$

$$\begin{aligned}
 (iii) \quad F(x+) &= \lim_{h \downarrow 0} F(x+h) = \lim_{h \downarrow 0} \int_{-\infty}^{x+h} f(t) dt \\
 &= \lim_{h \downarrow 0} \left(\int_{-\infty}^x f(t) dt + \int_x^{x+h} f(t) dt \right) \\
 &= \int_{-\infty}^x f(t) dt + \lim_{h \downarrow 0} \int_x^{x+h} f(t) dt \\
 &= \int_{-\infty}^x f(t) dt + 0
 \end{aligned}$$

$\left[\begin{aligned} F(x+) &\stackrel{\text{def}}{=} \lim_{y \downarrow x} F(y) \\ &= \lim_{y \downarrow x} \int_{-\infty}^y f(t) dt \\ &= \lim_{h \downarrow 0} \int_{-\infty}^{x+h} f(t) dt \end{aligned} \right]$

Check $F(x-) = F(x)$. i.e., left-cts & so cts.

By Fund. thm of calculus, $F'(x) = f(x)$ except when x is a pt of disc'ty of f .

THM Let f be a pdf & F the corresponding DF.

25.10 Then $\exists$ a r.v. $X \ni F(t) = P(X \leq t) \quad \forall t \in \mathbb{R}$.

[WE WILL NOT SEE PROOF ! ! !]

X as above is called a CTS RV. & F is its CDF & f is its pdf.

Discrete r.v.'s $X \iff$ pmf $p_x \iff$ step fn DF's F_x

CTS r.v.'s $X \iff$ pdf $f_x \iff F'_x = f_x$ DF's F_x
 $\hookrightarrow$ except at finitely many pts.

RULES & TOOLS TO WORK WITH CTS RV's.

$$(1) \quad P(X \in (a, b]) = F(b) - F(a)$$

$$P(X \in (a, b)) = P(X \in [a, b]) = P(X \in [a, b))$$

$$(2) \quad \text{If } I = \bigcup_{j=1}^{\infty} (a_j, b_j] \quad (I \text{ is a union of countably many disjoint intervals})$$

$$\text{then } P(X \in I) = \sum_{j=1}^{\infty} F(b_j) - F(a_j).$$

Rule (2) not APPLICABLE TO GENERAL SETS !!!

$P(X \in A)$ NOT DEFINED $\forall A$!!!

17/12 LECTURE 25

X - r.v. - $X \in \mathbb{R}$ - takes values in $\mathbb{R}$.
cts r.v.

Discrete r.v.

(1) X takes countably many values.

(2) Formal description - pmf $P_X(x) = P(X=x)$

(3) CDF: $F_X(x) = P(X \leq x)$

$$P_X(x) = F_X(x) - F_X(x-)$$

CDF gives pmf everywhere.

(1) X takes uncountably many values

(2) pdf - f_X - piecewise cts

$$P(X \in (a,b)) = \int_a^b f_X(x) dx$$

(3) $F_X(x) = P(X \leq x)$

$$F_X'(x) = f_X(x) \text{ if } f_X \text{ is cts at } x.$$

CDF gives pdf except at finitely many points.

X cts r.v. $\Leftrightarrow \exists$ a pdf f_X ($f_X: \mathbb{R} \rightarrow [0, \infty)$ p.cts, $\int_{-\infty}^{\infty} f_X(x) dx = 1$)

$\Leftrightarrow \exists$ a CDF $F_X: \mathbb{R} \rightarrow [0,1]$ $\exists F_X'$ exists except at fin. many pts & F_X' is a pdf.

• Two pdfs f & g are equal if $f=g$ except at finitely many points.
($f=g$)

• Two cts r.v. X & Y are equal in distribution ($X \stackrel{d}{=} Y$)
if $f_X = f_Y$.

• $f_X = f_Y \Rightarrow F_X(x) = F_Y(x) \quad \forall x \in \mathbb{R}$.

X - choose a point at random in $[0,1]$.

So $P(X \in (a,b)) = b-a \quad 0 \leq a < b \leq 1$.

prob. "random point is in (a,b) "

So $f_X(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{else} \end{cases}$ pdf

X as above is called UNIFORM $(0,1]$ r.v. ($X \stackrel{d}{=} \text{Unif}[0,1]$)

Eg
25.1

$$C) F \quad F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1. \end{cases}$$

Eg 2b.2 X - Choose a random point ^{uniformly} in (a, b) , $a < b$, $a, b \in \mathbb{R}$.

$$\text{So } P(X \leq c) = \begin{cases} 0 & c \leq a \\ 1 & c \geq b \end{cases}$$

$$\text{Prob. random pt } \leq c \begin{cases} 1 & c \geq b \end{cases}$$

$$a < c < b, \quad P(X \leq c) \propto c - a \quad \text{proportional to}, \quad P(X \leq a) = 0 = 1 - P(X \leq b)$$

$$\text{So } P(X \leq c) = \frac{c - a}{b - a}$$

$$\text{check. } F_X(x) = \begin{cases} 0 & x < a \\ \frac{x - a}{b - a} & a \leq x < b \\ 1 & x \geq b \end{cases} \text{ is a DF.}$$

F_X is CDF of X . X is called UNIFORM(a, b) r.v.

$$\text{pdf } f_X(x) = \begin{cases} \frac{1}{b - a} & a < x < b \\ 0 & \text{elsewhere.} \end{cases} \quad X \stackrel{d}{=} \text{UNIF}(a, b) \stackrel{d}{=} U(a, b).$$

Eg 2b.3 Q of customers arriving. T - time of arrival of ^{the} first customer.

$N(t) = \#$ of customers upto time $t \stackrel{d}{=} \text{Poi}(\lambda t)$

$$P(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, \dots$$

$$\{T > t\} = \text{No customer upto time } t. \quad t \geq 0.$$

$$= \{N(t) = 0\}$$

$$P(T \leq t) = 1 - P(T > t) = 1 - P(N(t) = 0)$$

$$= 1 - e^{-\lambda t}, \quad t \geq 0$$

so CDF of T is $F_T(t) = 1 - e^{-\lambda t}$, $t \geq 0$

pdf of T is $f_T(t) = \lambda e^{-\lambda t}$, $t \geq 0$.

T is called EXPONENTIAL (r.v.) ($T \stackrel{d}{=} \text{EXP}(\lambda)$).

T_k - Time of arrival of k^{th} customer. $F_{T_k}(\cdot)$ & $f_{T_k}(\cdot)$

Eg
25.4

X - Standard Normal r.v. / Gaussian r.v. ($X \stackrel{d}{=} N(0,1)$)

i.f $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $x \in \mathbb{R}$.

To check this is a pdf.

$$I = \int_{-\infty}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = 1 ?$$

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{(x^2+y^2)}{2}} dx dy \quad \begin{matrix} \text{Polar} \\ \text{coordinates} \end{matrix} \quad \left(\begin{matrix} \text{set } x = r \cos \theta \\ y = r \sin \theta \end{matrix} \right)$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \int_0^{\infty} r e^{-r^2/2} dr d\theta$$

$$= \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta \right) \times \left(\int_0^{\infty} r e^{-r^2/2} dr \right) = 1.$$

RE-VISIT RULES:

$$(1) \quad P(X \in (a, b]) = \overset{\text{by def of CDF}}{F_X(b) - F_X(a)} = \overset{\text{FTC}}{\int_a^b f_X(x) dx}$$

$$(2) \quad P(X \in \bigcup_{i=1}^{\infty} (a_i, b_i]) = \sum_{i=1}^{\infty} (F_X(b_i) - F_X(a_i))$$

$$= \sum_{i=1}^{\infty} \int_{a_i}^{b_i} f_X(x) dx \quad \begin{matrix} \text{(countable} \\ \text{additivity)} \\ \text{Assume} \end{matrix}$$

$$(3) \quad P(X \in \bigcup_{i=1}^{\infty} (a_i, b_i]) \leq \sum_{i=1}^{\infty} P(X \in (a_i, b_i])$$

$$= \sum_{i=1}^{\infty} F_X(b_i) - F_X(a_i) \quad \begin{matrix} \text{(count} \\ \text{subadd)} \end{matrix}$$

(4) $P(X \in A) \leq P(X \in B)$ when A, B are intervals
& $A \subseteq B$.

(5) $P(X \in A \cup B) = P(X \in A) + P(X \in B) - P(X \in A \cap B)$ A, B are intervals

(6) $\mathbb{I}$ & Bonferroni all hold for intervals!!!

$\mathcal{I} = \text{Intervals} = \{ (a, b], [a, b), (a, b), [a, b] : -\infty \leq a < b \leq \infty \}$

Eg: Let $X \stackrel{d}{=} U(0,1)$. Define $Y = aX + b$, $a > 0$, $b \in \mathbb{R}$.

25.5 What is Y ?

$$\begin{aligned} F_Y(x) &:= P(Y \leq x) = P(aX + b \leq x) \\ &= P(aX \leq x - b) = P\left(X \leq \frac{x-b}{a}\right) \end{aligned}$$

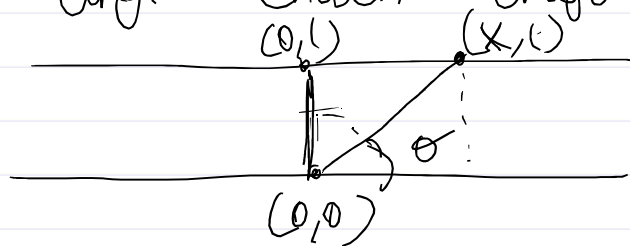
$$= \begin{cases} 0 & x-b < 0 \Leftrightarrow x < b \\ \frac{x-b}{a} & 0 \leq x-b < a \Leftrightarrow b \leq x < b+a \\ 1 & x-b \geq a \Leftrightarrow x \geq b+a \end{cases}$$

$$f_Y(x) = F_Y'(x) = \begin{cases} 1/a & b \leq x \leq b+a \\ 0 & \text{else} \end{cases}$$

$$\Rightarrow Y \stackrel{d}{=} \text{Unif}(b, b+a)$$

Ex: Let $X \stackrel{d}{=} N(0,1)$. What is $Y = aX + b$? $a > 0$, $b \in \mathbb{R}$.

Eg: 20.6 Let θ be an angle chosen uniformly at random in $(0, \pi)$.



$$X = r \cdot V.$$

$$\theta \stackrel{d}{=} \text{Unif}(0, \pi) \quad X = \tan\left(\frac{\pi}{2} - \theta\right)$$

$$\theta \in (0, \pi) \Rightarrow \frac{\pi}{2} - \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

$\Rightarrow \tan\left(\frac{\pi}{2} - \theta\right)$ is strictly decreasing on $(0, \pi)$

$\Rightarrow$ inverse $\arctan(x) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

$$\begin{aligned} F_X(x) &= P(X \leq x) = P\left(\tan\left(\frac{\pi}{2} - \theta\right) \leq x\right) \\ &= P\left(\frac{\pi}{2} - \theta \leq \arctan(x)\right) \quad \text{as } \tan\left(\frac{\pi}{2} - \theta\right) \text{ is strictly } \downarrow \\ &= P\left(\theta \geq \frac{\pi}{2} - \arctan(x)\right) \\ &= 1 - P\left(\theta < \frac{\pi}{2} - \arctan(x)\right) \\ &= 1 - F_\theta\left(\frac{\pi}{2} - \arctan(x)\right) \quad \text{since } \theta \text{ is CB} \\ &\quad \quad \quad F_\theta \text{ is CB} \end{aligned}$$

$$= 1 - \frac{1}{\pi} \left(\frac{\pi}{2} - \arctan(x)\right)$$

$$= \frac{1}{2} + \frac{1}{\pi} \arctan(x), \quad x \in \mathbb{R}$$

What is the pdf? $f_X(x) = F'_X(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R}.$

X is called Cauchy r.v.

$Y = aX + b$?

————— *

PREVIEW: (1) Given $h: \mathbb{R} \rightarrow \mathbb{R}$ diff'ble & nice.

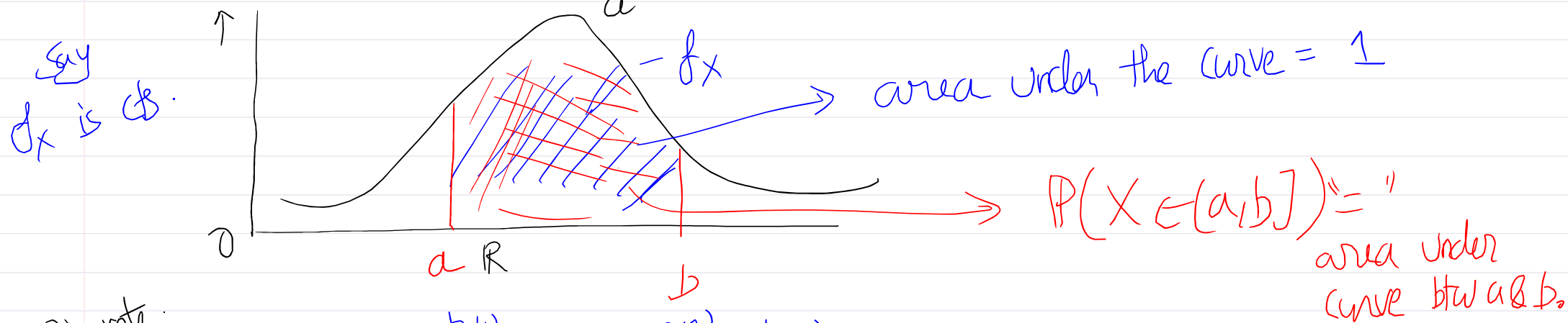
What is pdf/CDF of $h(X)$?

(2) $h: \mathbb{R} \rightarrow \mathbb{R}$; $E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx$ if well-defined.

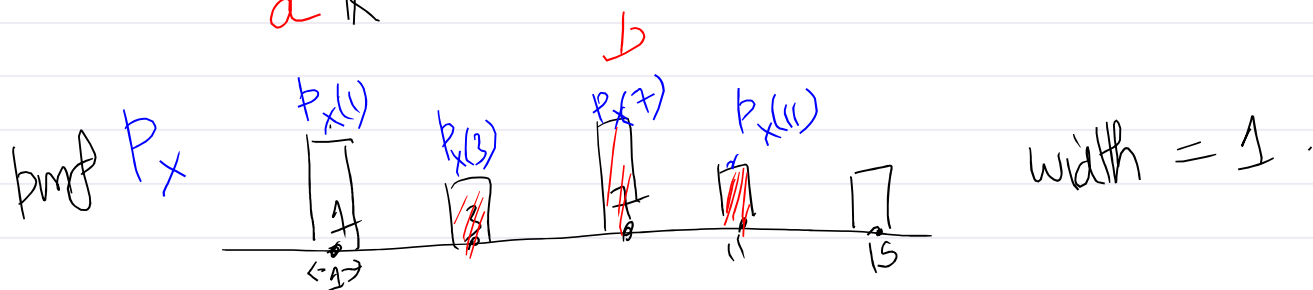
2/12 LECTURE 26 : TRANSFORMATIONS OF R.V.

X is r.v. $\Leftrightarrow \exists$ a pdf f_X ($f_X: \mathbb{R} \rightarrow [0, \infty)$ p.c.f., $\int_{-\infty}^{\infty} f_X(x) dx = 1$)
 $\Leftrightarrow \exists$ a ^{CB} CDF $F_X: \mathbb{R} \rightarrow [0, 1]$ $\exists F'_X$ exists except at fin. many pts & F'_X is a pdf.

$$P(X \in (a, b]) = \int_a^b f_X(x) dx = F_X(b) - F_X(a) \quad -\infty \leq a < b < \infty$$



Discrete
X-

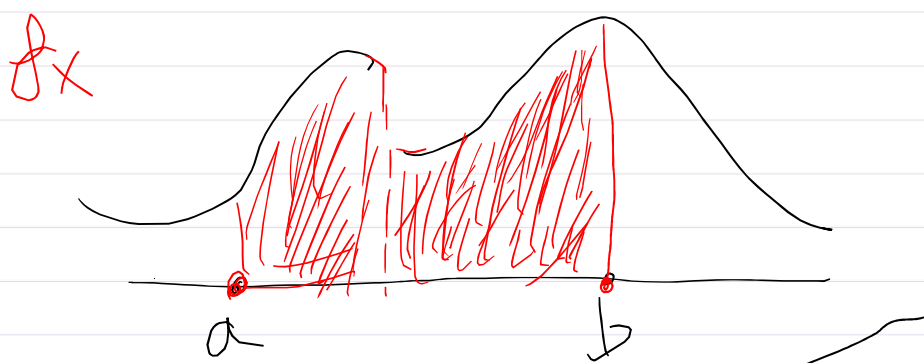


area of bar at $i = P_X(i)$

$$P(X \in (1, 1.5]) = \sum_{1 < i \leq 1.5} P_X(i)$$

area of "shaded region." = Total area of bars btw 1 & 1.5 (incl.)

X is r.v.
 f_X p.c.f.



$$P(X \in (a, b]) = \text{area of shaded region.}$$

Plot CDF's
 F_X



Fig 2b.1

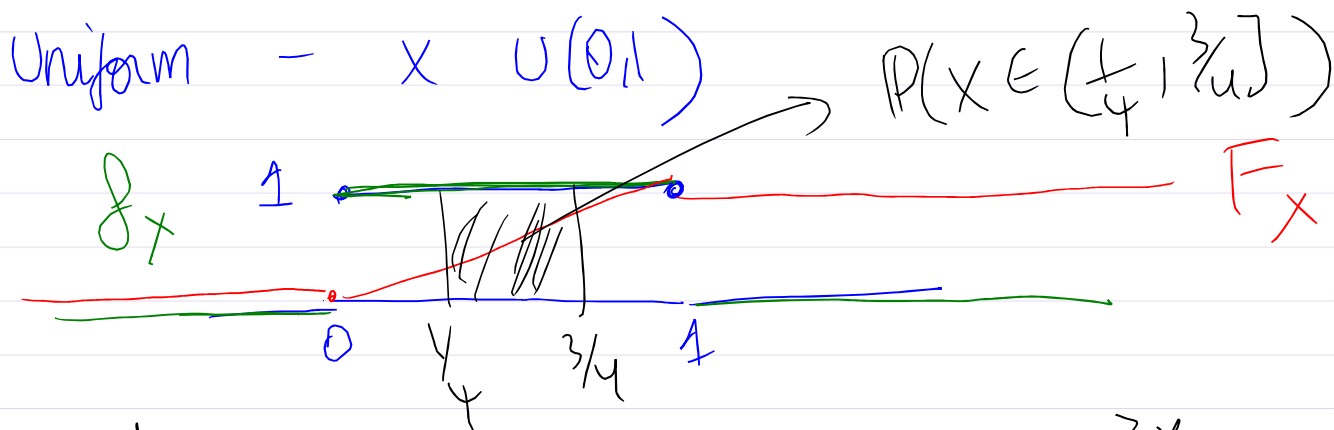
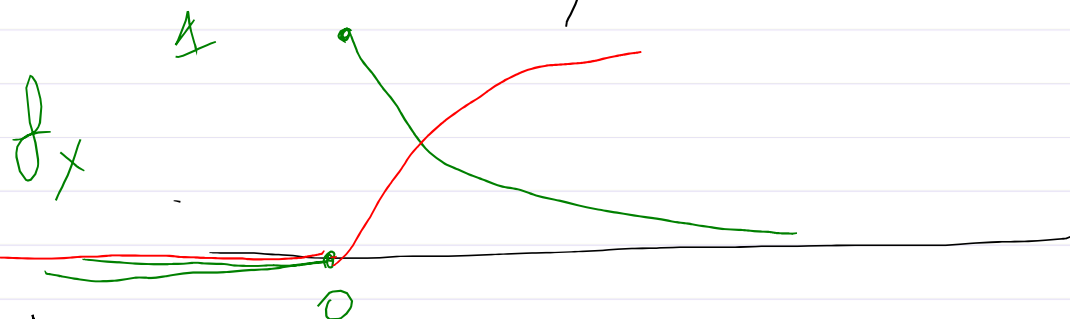


Fig 2b.2

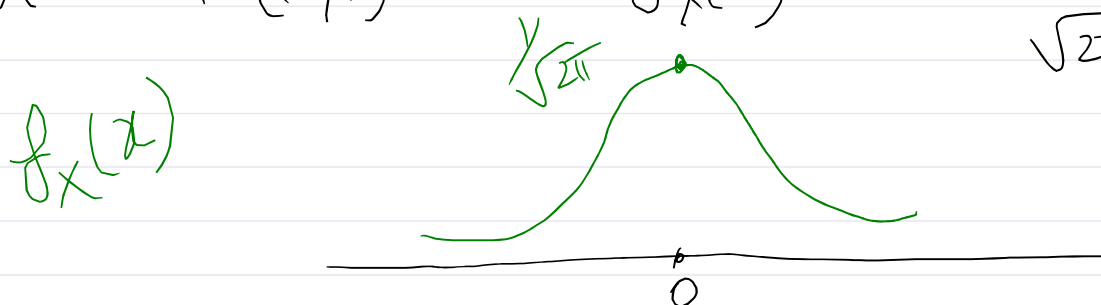
$X \stackrel{d}{=} \text{EXP}(\lambda)$ $f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$



$F_X(x) = 1 - e^{-\lambda x}, \quad x \geq 0$

Fig 2b.3

$X \stackrel{d}{=} N(0,1)$ $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}$



Bell curve
Gaussian curve

Qn: $h: \mathbb{R} \rightarrow \mathbb{R}$ is a cts fn. X is a cts r.v.

Let $Y = h(X)$. Is Y cts? What is the pdf?

LEMMA 2b.4 Let $Y = aX + b$, $a > 0$ & $b \in \mathbb{R}$. X is a cts r.v.

Then Y is a cts r.v. & $f_Y(y) = \frac{1}{a} f_X(\frac{y-b}{a})$.

Proof:

$F_Y(y) = P(Y \leq y)$

$= P(aX + b \leq y) = P(X \leq \frac{y-b}{a})$

$= F_X(\frac{y-b}{a})$ (Check F_Y is a CDF).

Since F_X is cts, so is F_Y .

If $F_X'(\frac{y-b}{a})$ exists,

then by chain rule

$F_Y'(y) = \frac{1}{a} F_X'(\frac{y-b}{a}) = \frac{1}{a} f_X(\frac{y-b}{a})$
(FTC)

since F_x isn't diff'ble only at fin. many pts ($\because f_x$ is p.cts)
 so is F_y . Hence $f_y(y) = F_y'(y) = \frac{1}{a} f_x\left(\frac{y-b}{a}\right)$
 whenever $\frac{y-b}{a}$ isn't a pt. of disc'ty of f_x .

$[F_x'(x) \text{ exists iff } x \text{ is a cty point of } f_x]$

Clearly f_y is p.cts. (f_y is cts at $y \Leftrightarrow f_x$ is cts at $\frac{y-b}{a}$)

$$\frac{1}{a} \int_{-\infty}^{\infty} f_x\left(\frac{y-b}{a}\right) dy \stackrel{\text{cov.}}{=} \frac{1}{a} \int_{-\infty}^{\infty} a f_x(x) dx \quad \text{set } x = \frac{y-b}{a} \\ dy = a dx$$

$$= \int_{-\infty}^{\infty} f_x(x) dx = 1$$

so $F_y' = f_y$ is a pdf of Y . ▀

Alternate - guess f_y & check $P(Y \leq y) = \int_{-\infty}^y f_y(y) dy$!

Eg 2605

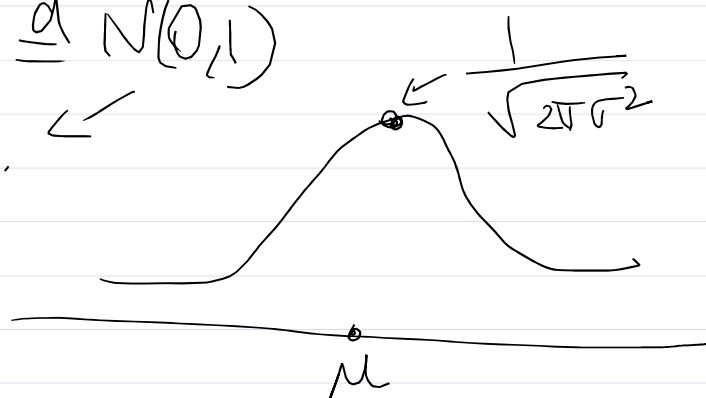
(1) $X \triangleq N(0,1)$ $f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$Y = \sigma X + \mu$, $\mu \in \mathbb{R}$, $\sigma > 0$. $f_y(y) = \frac{1}{\sigma} f_x\left(\frac{y-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$, $y \in \mathbb{R}$

We say Y has ^{Normal} distribution $N(\mu, \sigma^2)$ ($Y \triangleq N(\mu, \sigma^2)$)

Trivially $X \triangleq N(0,1)$

Standard Normal.



= Bell

$$(2) \quad X \stackrel{d}{=} \text{EXP}(\lambda) \quad f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0.$$

$$Y = aX, \quad a > 0.$$

$$f_Y(y) = \frac{\lambda}{a} e^{-\frac{\lambda}{a}y}, \quad y \geq 0$$

$$Y \stackrel{d}{=} \text{EXP}\left(\frac{\lambda}{a}\right).$$

Ex. Uniform & Cauchy. Try!

LEMMA 2.6. Let X be a cts r.v. with pdf f & CDF F .

Define $Y = \frac{1}{a}(X-b)^2$ $a > 0, b \in \mathbb{R}$.

Then Y is a cts r.v.

$$f_Y(y) = \frac{\sqrt{a}}{2\sqrt{y}} (f_X(b+\sqrt{ay}) + f_X(b-\sqrt{ay})), \quad y > 0.$$

& at pts of discty on RHS.

Proof:

$$F_Y(y) = P(Y \leq y), \quad y > 0$$

$$= P((X-b)^2 \leq ay)$$

$$= P(-\sqrt{ay} \leq X-b \leq \sqrt{ay})$$

$$= P(b-\sqrt{ay} \leq X \leq b+\sqrt{ay})$$

$$= F_X(b+\sqrt{ay}) - F_X(b-\sqrt{ay}) \quad (\text{since } X \text{ is cts})$$

F_Y is a CDF & it is diff'ble except at fin. many points.
 of $b+\sqrt{ay}$ & $b-\sqrt{ay}$ are cty points of f_X

(i.e., F_X is diff'ble at $b+\sqrt{ay}, b-\sqrt{ay}$)

then $f_Y(y) = F_Y'(y) = \frac{d}{dy} (F_X(b+\sqrt{ay}) - F_X(b-\sqrt{ay}))$

(use chain rule & inv.)

$$= \frac{1}{2} \sqrt{\frac{a}{y}} (f_X(b+\sqrt{ay}) + f_X(b-\sqrt{ay})).$$

Define $f_Y(y) = \frac{\sqrt{a}}{2\sqrt{y}} (f_X(b+\sqrt{ay}) + f_X(b-\sqrt{ay}))$, $y > 0$. O.E.C.

Easily f_Y is p-cts & $\int_{-\infty}^{\infty} f_Y(y) dy = 1$.

$$\& F_Y(y) = \int_{-\infty}^y f_Y(y) dy \implies$$

So f_Y is a pdf for Y . ▮

Gen. n - see THM 7.1 of Ross.

Eg
26.7

$$X \stackrel{d}{=} N(0,1)$$

$$Y = X^2$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

f_X is even

$$f_Y(y) = \frac{1}{2\sqrt{y}} (f_X(\sqrt{y}) + f_X(-\sqrt{y})) = \frac{f_X(\sqrt{y})}{\sqrt{y}}$$

$$= \frac{1}{\sqrt{2\pi}y} e^{-y/2} = \frac{y^{-1/2} e^{-y/2}}{\sqrt{2\pi}}$$

Y has Gamma($1/2, 1/2$) distribution. i.e., $Y \stackrel{d}{=} \Gamma(1/2, 1/2)$.

Since f_Y is a pdf

$$\int_0^{\infty} y^{-1/2} e^{-y/2} dy = \sqrt{2\pi}.$$

$$\frac{1}{\sqrt{2}} \int_0^{\infty} y^{-1/2} e^{-y/2} dy = \sqrt{\pi}.$$

Eg
26.8

[Gamma(ν, λ) distribution]

$$\Gamma(\nu, \lambda), \quad \nu > 0, \lambda > 0$$

ν - shape parameter,

λ - scale parameter.

$$X \stackrel{d}{=} \Gamma(\nu, \lambda) \text{ if}$$

$$f_X(t) =$$

$$\begin{cases} \frac{\lambda^\nu t^{\nu-1} e^{-\lambda t}}{\Gamma(\nu)} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

where $\Gamma(\gamma) := \lambda^\gamma \int_0^\infty t^{\gamma-1} e^{-\lambda t} dt$. , Γ - Gamma
 γ - nu

check: f_x is cts except at $t=0$.

$\Gamma(\gamma)$ - Gamma integral;

$$\Gamma(\gamma) = \lambda^\gamma \int_0^\infty t^{\gamma-1} e^{-\lambda t} dt = \int_0^\infty u^{\gamma-1} e^{-u} du. \quad \leftarrow u = \lambda t, du = \lambda dt$$

$\Gamma(\gamma)$ doesn't depend on λ .

f_x is well-defined if $0 < \Gamma(\gamma) < \infty$

Since $\gamma > 0$, $\Gamma(\gamma) > 0$.

Is $\Gamma(\gamma) < \infty$?

$$\Gamma(\gamma+1) = \int_0^\infty u^\gamma e^{-u} du = \left[-e^{-u} u^\gamma \right]_0^\infty + \gamma \int_0^\infty u^{\gamma-1} e^{-u} du$$

Integration by parts.

$$= 0 + \gamma \Gamma(\gamma)$$

(since $\gamma > 0$)

$$\left(\lim_{u \rightarrow \infty} u^\gamma e^{-u} = 0 \right)$$

$$\text{so } \Gamma(\gamma+1) = \gamma \Gamma(\gamma).$$

so if $\Gamma(\gamma) < \infty \quad \forall \quad \gamma \in (0, 1]$ then

$$\Gamma(\gamma) < \infty \quad \forall \quad \gamma > 0.$$

$$\Gamma(1) = \int_0^\infty e^{-u} du = 1.$$

$$\Rightarrow \Gamma(n+1) = n! \quad \forall \quad n \in \mathbb{N}_0.$$

$$\Gamma(1/2) = \int_0^\infty u^{-1/2} e^{-u} du = \frac{1}{\sqrt{2}} \int_0^\infty u^{-1/2} e^{-u/2} du = \sqrt{\pi}$$

$$\Rightarrow \Gamma(n+1/2) < \infty \quad \forall \quad n \in \mathbb{N}_0.$$

Check $\Gamma(\gamma) < \infty \quad \forall \quad \gamma \in (0, 1]$.

X cts r.v. $\Leftrightarrow \exists$ a pdf f_X $f_X: \mathbb{R} \rightarrow [0, \infty)$ p.c.f.s, $\int_{-\infty}^{\infty} f_X(x) dx = 1$
 $\Leftrightarrow \exists$ a ^{cts} DF $F_X: \mathbb{R} \rightarrow [0, 1]$ $\exists F_X'$ exists except at fin. many pts.

$$P(X \in (a, b]) = \int_a^b f_X(x) dx = F_X(b) - F_X(a) \quad -\infty \leq a < b < \infty$$

THM 27.1 Let X be a cts r.v. $\Rightarrow F_X$ is a strictly $\uparrow$ cts fn.
 Then (1) $F_X^{-1}(U) \stackrel{d}{=} X$ ($U \stackrel{d}{=} U(0,1)$ & F_X^{-1} is well-defined on $(0,1)$)
 (2) $F_X(X) \stackrel{d}{=} U$.

Proof: (1) F_X is strictly $\uparrow$ & so F_X^{-1} is well-defined on $(0,1)$
 & also strictly increasing (ex.)
 $x \in \mathbb{R}$, $P(F_X^{-1}(U) \leq x) \stackrel{\downarrow}{=} P(U \leq F_X(x)) = F_X(x)$

so X & $F_X^{-1}(U)$ have same CDFs i.e., $X \stackrel{d}{=} F_X^{-1}(U)$
 (equal in distribution)

(2) Similarly, $x \in (0,1)$

$$P(F_X(X) \leq x) \stackrel{\uparrow}{=} P(X \leq F_X^{-1}(x)) = F_X(F_X^{-1}(x)) \stackrel{\leftarrow}{=} x = P(U \leq x)$$

F_X strictly $\uparrow$ F_X is strictly $\uparrow$

$$\Rightarrow F_X(X) \stackrel{d}{=} U$$

Applications: output of RAND is U . To generate X , you take $F_X^{-1}(\text{RAND})$.

Qn: Suppose X is a discrete r.v. (For eg. X has pmf p on $\{1, \dots, n\}$)
 so there $h: \mathbb{R} \rightarrow \mathbb{R} \Rightarrow h(U) \stackrel{d}{=} X$?

Defn
27.2

X be cts r.v. We define mean / Expectation of X as

$$E[X] := \int_{-\infty}^{\infty} x f_X(x) dx \quad \text{if} \quad \int_{-\infty}^{\infty} |x| f_X(x) dx < \infty.$$

Else undefined or $E[X]$ doesn't exist.

$$[X \text{ Discrete}, E[X] = \sum_x x p_X(x) \quad \text{if} \quad \sum_x |x| p_X(x) < \infty]$$

Note $x \mapsto x f_X(x)$ is p.cts.

Ex
27.3

$$(1) \quad U \triangleq U(0,1). \quad E[X] = \int_0^1 x dx = \frac{1}{2}.$$

$$[\text{check } \int_{-\infty}^{\infty} |x| f_X(x) dx = \int_0^1 x dx < \infty]$$

$$(2) \quad X \triangleq \text{EXP}(\lambda) \quad E[X] = \lambda \int_0^{\infty} x e^{-\lambda x} dx \stackrel{\text{IBP}}{=} \frac{1}{\lambda}.$$

check $E[X]$ exists

$$(3) \quad X \triangleq N(0,1)$$

$$\text{Check } \int_{-\infty}^{\infty} |x| f_X(x) dx < \infty.$$

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

$$\left[\underset{f_X \text{ even}}{=} 2 \int_0^{\infty} x f_X(x) dx \quad ; \quad f_X(x) \leq \frac{C}{x^3} \quad \forall x > 0 \right]$$

$$= \int_0^{\infty} x f_X(x) dx + \int_{-\infty}^0 x f_X(x) dx$$

$$= \int_0^{\infty} x f_X(x) dx + \int_0^{\infty} x f_X(-x) dx \quad \xrightarrow{\text{Cov } x \mapsto -x}$$

$$= \int_0^{\infty} x f_X(x) dx - \int_0^{\infty} x f_X(x) dx$$

$$= \int_0^{\infty} x f_X(x) dx - \int_0^{\infty} x f_X(x) dx$$

$$= 0$$

(since f_X is even i.e., $f_X(x) = f_X(-x)$)

• If f_X is even & $\int_{-\infty}^{\infty} x f_X(x) dx < \infty$ then $E[X] = 0$.
(Same proof as above)

• If $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$ then $\lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b x f_X(x) dx = \int_{-\infty}^{\infty} x f_X(x) dx$

If $\sum_{n=1}^{\infty} |a_n| < \infty$ then $\lim_{n \rightarrow \infty} \sum_{l=1}^n a_l$ exists $(a_n = (-1)^n, n \geq 0)$

Ex
27.4

$X \stackrel{d}{=} \text{Cauchy}$

$$f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$$

f_X is even. $\int_{-\infty}^{\infty} |x| f_X(x) dx = \frac{2}{\pi} \int_0^{\infty} \frac{x}{1+x^2} dx = \infty!$
check.

So $E[X]$ doesn't exist.

(Compare with $\sum_{n=1}^{\infty} \frac{1}{n^2}$)

Def
27.5

Let X be cts r.v. & $h: \mathbb{R} \rightarrow \mathbb{R}$ p-cts.

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx \quad \text{if } \int_{-\infty}^{\infty} |h(x)| f_X(x) dx < \infty.$$

else undefined / $E[h(X)]$ doesn't exist.

VARIANCE. $\text{VAR}[X] := E[(X - \mu_X)^2]$ if $E[X^2]$ exists

where $\mu_X := E[X]$.

k^{th} MOMENT - $E[X^k]$, $k=1, 2, \dots$ if it exists.

Thm
27.6

(1) If $E[X]$ exists then $\forall a, b \in \mathbb{R}$

$$E[ax + b] = aE[X] + b$$

(2) If $h_1 \leq h_2$ and $E[h_i(X)]$ exists

then $E[h_1(X)] \leq E[h_2(X)]$ (Monotonicity)

(3) if $E[X^2]$ exists

$$\text{VAR}[X] = E[X^2] - \mu_X^2$$

(4) if $E[X^2]$ exists, then $\forall a, b \in \mathbb{R}$.

$$\text{VAR}[aX+b] = a^2 \text{VAR}[X]$$

Proof: (1) To check $\int_{-\infty}^{\infty} |ax+b| f_X(x) dx < \infty$

observe $|ax+b| \leq |a||x| + |b|$. ('triangle ineq')

$$\begin{aligned} \int_{-\infty}^{\infty} |ax+b| f_X(x) dx &\leq \int_{-\infty}^{\infty} (|a||x| + |b|) f_X(x) dx \\ &\quad \text{(mon. for non-neg. integrals)} \leftarrow \\ &\quad \text{(linearity of integrals)} \leq |a| \int_{-\infty}^{\infty} |x| f_X(x) dx + |b| \int_{-\infty}^{\infty} f_X(x) dx \end{aligned}$$

$$\begin{aligned} \text{Since } E[X] \text{ exists } &\leftarrow < \infty & + |b| < \infty. \\ & & \rightarrow f_X \text{ is pdf} \end{aligned}$$

so $E[aX+b]$ exists.

$$\text{by defn } E[aX+b] = \int_{-\infty}^{\infty} (ax+b) f_X(x) dx$$

$$\text{(lin. of integrals)} = a \int_{-\infty}^{\infty} x f_X(x) dx + b \int_{-\infty}^{\infty} f_X(x) dx$$

$$\text{(defn of } E[X] \text{ \& } f_X \text{ is a pdf)} = a E[X] + b$$

(2) Ex.

$$(3) \text{VAR}[X] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \quad (\text{defn})$$

$$\begin{aligned} &= \int_{-\infty}^{\infty} x^2 f_X(x) dx - 2\mu_X \int_{-\infty}^{\infty} x f_X(x) dx + \mu_X^2 \int_{-\infty}^{\infty} f_X(x) dx \\ &\quad \text{(linearity of integral)} \end{aligned}$$

$$= E[X^2] - 2\mu_X^2 + \mu_X^2 \quad (\text{by defns})$$

$$= E[X^2] - \mu_x^2.$$

(4) Let $Y = aX + b$.

From (1), $\mu_Y = a\mu_X + b$

$$\int_{-\infty}^{\infty} (ax+b)^2 f_X(x) dx \leq a^2 \int_{-\infty}^{\infty} x^2 f_X(x) dx + 2|b| \int_{-\infty}^{\infty} |x| f_X(x) dx + b^2$$

$< \infty$ as $E[X^2], E[X]$ exist.

so $E[(aX+b)^2]$ exists.

$$\Rightarrow \text{VAR}[aX+b] = \int_{-\infty}^{\infty} (ax+b - a\mu_X - b)^2 f_X(x) dx$$

(defn & $\mu_Y = a\mu_X + b$)

$$= a^2 \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx = a^2 \text{VAR}[X]. \quad \square$$

Note:

$$Y = aX + b$$

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

(1) of thm $\rightarrow$

$$E[aX+b] = E[Y]$$

$$= \int_{-\infty}^{\infty} y f_Y(y) dy = \frac{1}{a} \int_{-\infty}^{\infty} y f_X\left(\frac{y-b}{a}\right) dy$$

$$= \int_{-\infty}^{\infty} (ax+b) f_X(x) dx$$

(change $\frac{y-b}{a} \rightarrow x$; $dy = a dx$)

So defn. of $E[h(X)]$ is consistent in the simple case of $h(X) = aX + b$!!!

[for more gen. case. say $Y = h(X)$
then $E[Y] = E[h(X)]$ follows by a gen. Cov]

Eg
27.7

$$(1) \quad U \quad E[U] = \frac{1}{2}$$

$$k \geq 1 \quad E[U^k] = \int_0^1 x^k dx = \frac{1}{k+1}$$

$$VAR[U] = E[U^2] - (E[U])^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

$$(2) \quad X \stackrel{d}{=} U(b, b+a) \quad f_X(x) = \begin{cases} \frac{1}{a} & x \in (b, b+a) \\ 0 & \text{elsewhere} \end{cases} \quad \begin{matrix} a > 0 \\ b \in \mathbb{R} \end{matrix}$$

From previous class.

$$X \stackrel{d}{=} Y, \quad Y = \underline{a}U + b$$

$$E[X] = E[Y] \stackrel{\text{Thm 27.6(4)}}{=} \underset{\text{Thm 27.6(1)}}{a} E[U] + b = \frac{a}{2} + b$$

$$VAR[X] \stackrel{\text{Thm 27.6(4)}}{=} a^2 VAR[U] = \frac{a^2}{12}$$

$$(3) \quad X \stackrel{d}{=} \text{EXP}(\lambda) \quad E[X] = \frac{1}{\lambda}$$

$$E[X^2] = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$$VAR[X] = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

$$(4) \quad Z \stackrel{d}{=} N(0,1) \quad X \stackrel{d}{=} N(\mu, \sigma) \quad \sigma > 0$$

Last class, if $Y = \sigma Z + \mu$, then $Y \stackrel{d}{=} X$.

$$E[X] = E[Y] = \sigma E[Z] + \mu = \mu$$

$$VAR[X] = \sigma^2 VAR[Z]$$

$$VAR[Z] = E[Z^2] \quad (E[Z] = 0)$$

$$= \int_{-\infty}^{\infty} z^2 f_Z(z) dz$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z^2 e^{-z^2/2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} x^{1/2} e^{-x/2} dx \quad \leftarrow x = z^2 \quad dz = \frac{dx}{2\sqrt{x}}$$

$$= \frac{1}{\sqrt{2\pi}} \Gamma(3/2) = 1.$$

$$\text{So } \text{VAR}[x] = \sigma^2 \text{VAR}[z] = \sigma^2.$$

$$X \stackrel{d}{=} N(\mu, \sigma) ; \quad \mu = E[X], \quad \sigma^2 = \text{VAR}[X]$$

$$\sigma = \text{SD}[X] = \sqrt{\text{VAR}[X]}.$$

Exercises Compute $E[X^k]$ in exp. $N(\mu, \sigma)$, . .

Beob $X_n \stackrel{d}{=} \text{Bin}(n, \frac{\lambda}{n})$ $\lambda > 0$ $E X_n = \lambda$, $\text{Var}[X_n] = \lambda(1 - \frac{\lambda}{n})$

$$k=0,1, \dots \quad P(X_n = k) \xrightarrow{1} P(X = k) \quad X \stackrel{d}{=} \text{Poi}(\lambda)$$

↳ Poisson limit theorem / law of small numbers

"David Hand!"

$$S_n \stackrel{d}{=} \text{Bin}(n, p) \quad p \in (0, 1]$$

$$\mathbb{E}[S_n] = np \quad \text{VAR}[S_n] = np(1-p).$$

$a, b \in \mathbb{R}$

$$\mathbb{P}\left(a < \frac{S_n - np}{\sqrt{np(1-p)}} < b\right) \xrightarrow{d} \mathbb{P}(N \in (a, b))$$

De Moivre Laplace Central Limit theorem!

Ans:

(1) Given a seq r.v. X_n , $E[X_n] \rightarrow \lambda$, $VAR[X_n] \rightarrow \lambda$
Can one show Poisson limit theorem?

(2) Given a seq r.v. S_n , $E[S_n] \rightarrow \infty$, $VAR[S_n] \rightarrow \infty$

is it true $P(a < \frac{S_n - E[S_n]}{\sqrt{VAR[S_n]}} < b) \rightarrow P(N \in (a, b))$
(CLT) ?

Try to show U !

THANK U