PROBABILITY

Axioms

① $P(A) \geq 0$ ② $P(\Omega) = 1$

③ $P(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_k) = \sum_{i=1}^{K} P(A_i)$

If all are disjoint.

Consequences

- If $A \subset B$ then $P(A) \leq P(B)$

  Proof: $B = A + O$ $\{O \equiv \text{other}\}$

  $P(B) = P(A+O)$ $\{A \cap O = \phi\}$

  $P(B) = P(A) + P(O)$

  $P(B) - P(A) = P(O)$

  $P(O) \geq 0$

  $\therefore P(B) \geq P(A)$

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

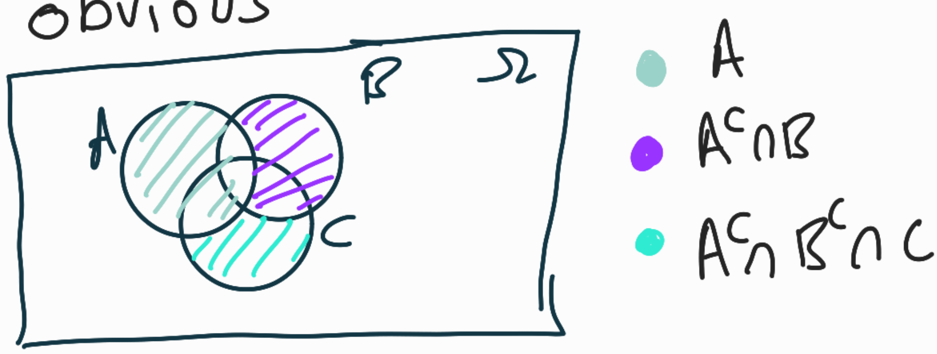
  Proof: Venn diagram / Inclusion-exclusion Principle

- Union Bound

  $P(A \cup B) \leq P(A) + P(B)$

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

<u>Proof</u>: obvious

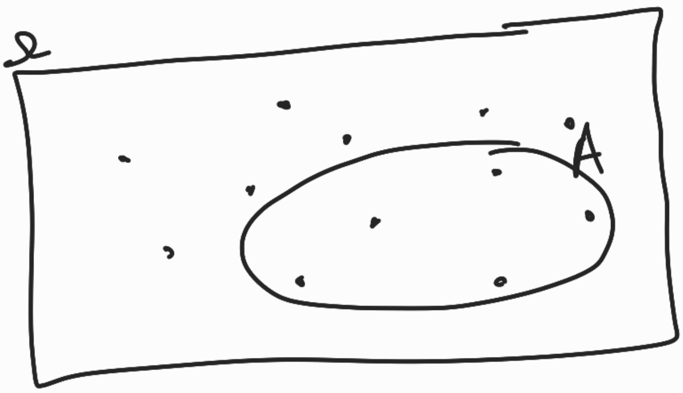


<u>Alt</u>: $P(A \cup B \cup C) = \sum_{A,B,C} P(A) - \sum_{A,B,C} P(A \cap B) + P(A \cap B \cap C)$

Discrete Uniform Law

Assume Ω has n equally likely events

Assume A has K elements



then, $P(A) = \dfrac{K}{n}$ //

Countable Additivity axiom

If $A_1, A_2, A_3, \ldots$ is an infinite sequence of disjoint events, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

{Can't be applied to a non-countable scenario}

Bonferroni Inequality

$$P(A \cap B) \geq P(A) + P(B) - 1$$

Proof: $P(A \cap B)^c = P(A^c \cup B^c) \leq P(A^c) + P(B^c)$

$$\Rightarrow 1 - P(A \cap B) \leq 1 - P(A) + 1 - P(B)$$

$$\Rightarrow P(A \cap B) \geq P(A) + P(B) - 1$$

General:

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$$

Conditional Probabilities

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad (P(B) > 0)$$

Properties:

- $P(A|B) > 0$
- $P(\Omega|B) = 1 = P(B|B)$
- If $A \cap C = \phi$ then $P(A \cup C|B) = P(A|B) + P(C|B)$

<u>multiplication Rule</u>:

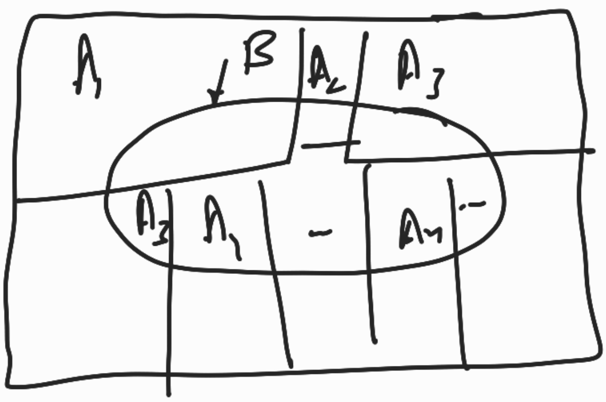
$$P(A \cap B) = P(B)\,P(A|B) = P(A)P(B|A)$$

$$P(A \cap B^c \cap C) = P(A)\,P(B^c|A)\,P(C|A \cap B^c)$$

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \prod_{i=2}^{n} P(A_i | A_1 \cap A_2 \ldots \cap A_{i-1})$$

<u>Total Probability</u>:

'n' Partitions of Ω



$\therefore$ Total Probability of B is given by

$$P(B) = \sum P(B \cap A_i)$$

$$P(B) = \sum P(A_i)\, P(B|A_i)$$

Bayes's Rule

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(A_i|B)}{\sum_{k=1}^{n} P(A_k)P(B|A_k)}$$

Independence

- when new Info about an event doesn't change its probability.

$$\Rightarrow P(B|A) = P(B)$$

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

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

Actual definition

Conditional Independence

- when Inpendence holds after some event has happened

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

Conditional Inpendence $\not\Leftrightarrow$ Normal Inpendence

Independence for collection of events

- $P(A_1 \cap A_2 \cap \ldots \cap A_n) = \bigwedge_{i=1}^{n} A_i$

Random Variable: Definition

- It associates a value to every possible outcome.
- Its a function from sample space to real numbers / Integers
- Discrete or Continuous.

Notation:

(r.v.) random variable X

numerical value x

- We can combine r.v.s.

PMF for a discrete R.v.

- Probability law of X

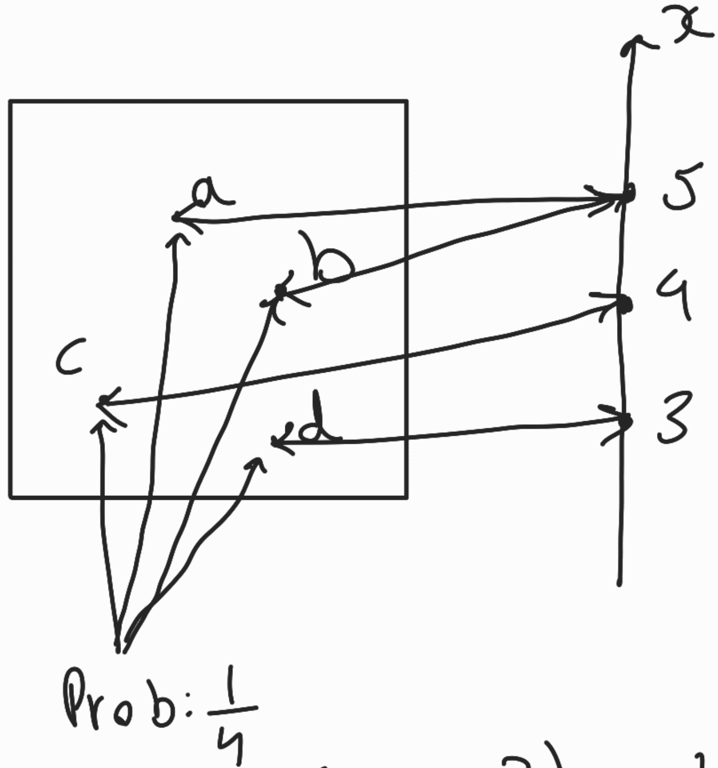
$$p_X(x) = P(X=x) = P(\{\omega \in \Omega : X(\omega) = x\})$$

Probability that r.v. X takes value x

those outcomes in Ω that yield value x

- $\displaystyle\sum_x p_X(x) = 1$

ex:



$$p_X(5) = P(\{a, b\}) = \frac{1}{2}$$

$P_X(4) = P(\{c\}) = \frac{1}{4}$

$P_X(3) = P(\{d\}) = \frac{1}{4}$

$$\sum_{i=3,4,5} P_X(i) = 1$$

ex: Tetrahedral die rolled for 2 times

Let every outcomes have $P = \frac{1}{16}$

$$\Omega = \left\{ \begin{matrix} (1,1), (1,2), (1,3), (1,4), \\ (2,1), (2,2), (2,3), (2,4), \\ (3,1), (3,2), (3,3), (3,4), \\ (4,1), (4,2), (4,3), (4,4) \end{matrix} \right\}$$

X: r.v. = no. on 1^{st} roll

Y: r.v. = no. on 2^{nd} roll

$Z = X + Y$

we need PMF $P_Z(z)$

$\therefore$ for all z

We see $z \in \{2, 3, 4, 5, 6, 7, 8\}$

$$\therefore P_Z(2) = P(\{(1,1)\}) = \frac{1}{16}$$

$$P_Z(3) = P(\{(2,3),(3,2)\}) = \frac{2}{16}$$

$$P_Z(4) = P(\{(2,2),(3,1),(1,3)\}) = \frac{3}{16}$$

Similarly $P_Z(5) = \frac{4}{16}$

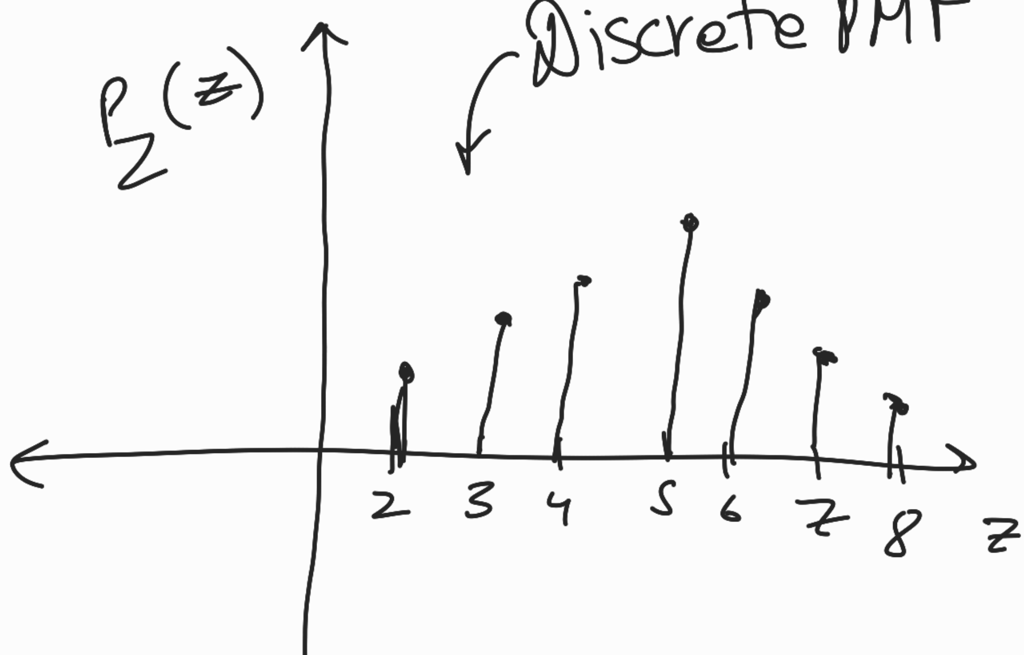
$$P_Z(6) = \frac{3}{16}$$

$$P_Z(7) = \frac{2}{16}$$

$$P_Z(8) = \frac{1}{16}$$

$\therefore$ PMF $\quad P_Z(z)$ $\quad$ Discrete r.v.

Discrete PMF



Bernoulli r.v.

$$X = \begin{cases} 1 & ; p \\ 0 & ; 1-p \end{cases}$$

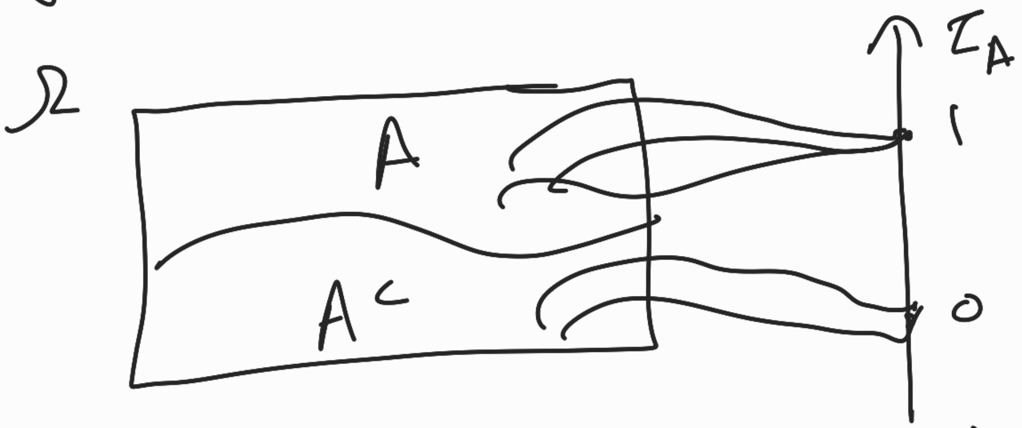
$$P_X(0) = 1-p$$

$$P_X(1) = p$$

Uses:→

- model for success/Failure or Head/Tails.

- Indicator r.v. of event A



$\therefore P_{I_A}(1) = P(I_A = 1) = P(A)$

{ Linking r.v. with another event A }

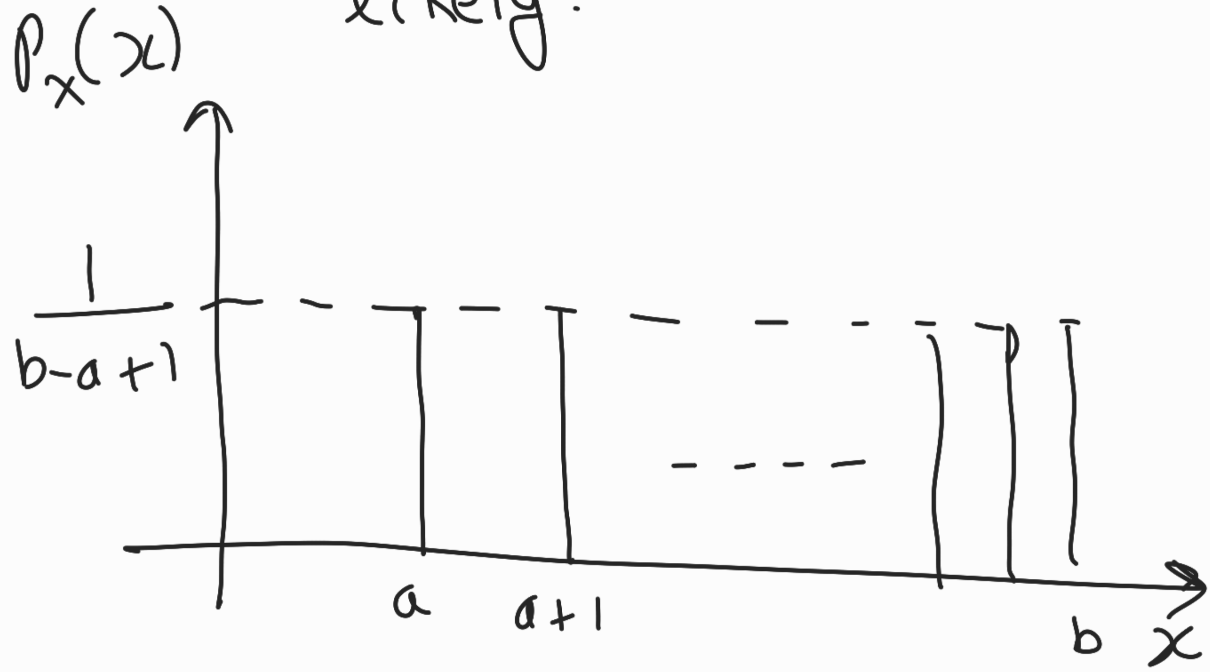
Discrete uniform r.v. ; parameter a, b

Int: a, b $(a \leq b)$

$\Omega = \{a, a+1, \ldots, b\}$

$X : X(\omega) = \omega \quad \{\omega \in \Omega\}$

all $b-a+1$ outcome equally likely.



Binomial r.v. ; parameter $p \in [0,1]$.

Exp: n Independent tosses of coin

$P(\text{Head}) = p$

$P(\text{tail}) = 1-p$

Ω = Sequence of H & T of length n.

X = number of Heads observed.

$P_X(x) = P(X=x) = P(x \text{ heads observed})$

$$P_X(x) = {}^nC_x\, p^x (1-p)^{n-x}$$

<u>Geometric r.v.</u>; parameter $p \in (0,1]$

Exp: Infinitely many tosses of coin

$P(\text{Head}) = p$

Ω = Infinite sequences of H & T

X = number of tosses until a head

PMF calculation:

$P_X(1) = P(\text{Head in } 1^{st} \text{ toss}) = p$

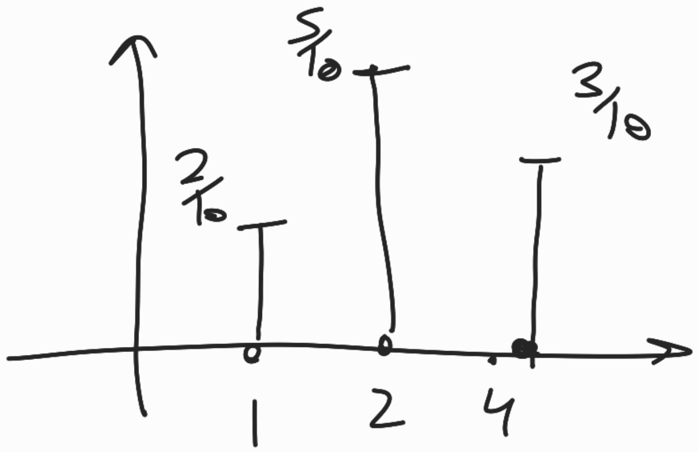
$P_X(2) = P(TH) = (1-p)\,p$

$$P_X(3) = P(TTH) = (1-p)^2 p$$

$$\text{PMF} \rightarrow \quad P_X(x) = (1-p)^{x-1} p$$

Expection / mean of r.v.

$$X = \begin{cases} 1 & \text{with prob } 2/10 \\ 2 & \text{w.p. } 5/10 \\ 4 & \text{w.p. } 3/10 \end{cases}$$



$$E(X) = \sum_{x} x\, P_X(x)$$

$$\therefore E(X) = 1 \cdot \frac{2}{10} + 2 \cdot \frac{5}{10} + 4 \cdot \frac{3}{10}$$

$E(X) = 2.4$

- $E(x)$ For bernoulli r.v.

$$X = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases}$$

$$E(X) = 1 \cdot p + 0 \cdot (1-p) = p$$

- $E(X)$ For uniform r.v.

$$X = \begin{cases} 0, 1, 2, \ldots, n & \text{w.p. } \frac{1}{n+1} \end{cases}$$

$$E(X) = \frac{1}{n+1} \sum_{i=0}^{n} i$$

$$= \frac{1}{n+1} \frac{n(n+1)}{2} = \frac{n}{2}$$

- $E(X)$ of Binomial r.v.

$$X = \begin{cases} x & \text{w.p. } {}^{n}C_{x}\, p^{x} (1-p)^{n-x} \end{cases}$$

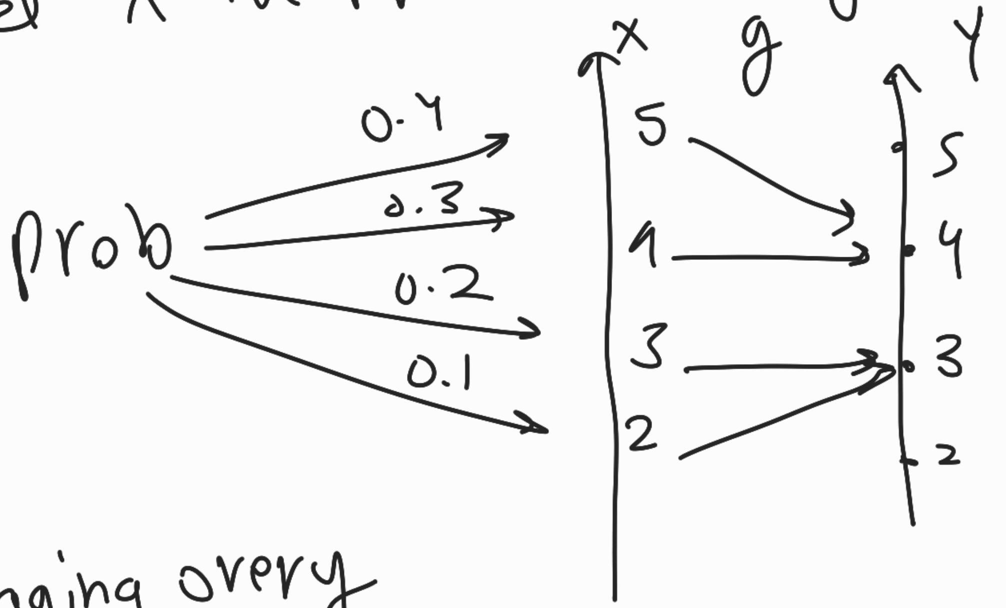
$E(X) = np$ (after calc)

Properties

- If $X \geq 0$ then $E(X) \geq 0$
- If $a \leq X \leq b$ then $a \leq E(X) \leq b$

Expected value rule

- Let X be r.v. & let $Y = g(X)$



averaging over y

$$E[Y] = \sum_y y\, P_Y(y)$$

$$E(Y) = 3(0.1 + 0.2) + 4(0.4 + 0.3)$$

averaging over x: $3.0.1 + 3.02 + 4.03 + 4.05$

averaging over x.

$$\boxed{\therefore E[Y] = E[g(X)] = \sum_x g(x)\, p_X(x)}$$

Linearity of Expectations

$$E[aX+b] = aE[X]+b$$

Poisson r.v. →

$$f(x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad x \in \{1, 2, 3, \ldots\}$$

It is the number of events taking place in the given time period.

$$E[X] = \lambda$$

$$\text{Variance} = \lambda$$

Continuous Random Variables

→ Their domain is continuous

→ If $P(X=x)=0$, still the event may occur

We also consider cumulative distribution function (cdf) :→

$$F_x(x) = P\{X \leq x\}$$

↳ can be used to compute Probability that X lies in range $[a,b]$

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

Probability density function (pdf)

pdf is derivative of cdf

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

$$P(x \in A) = \int_A f_X(x)\,dx$$

Properties:

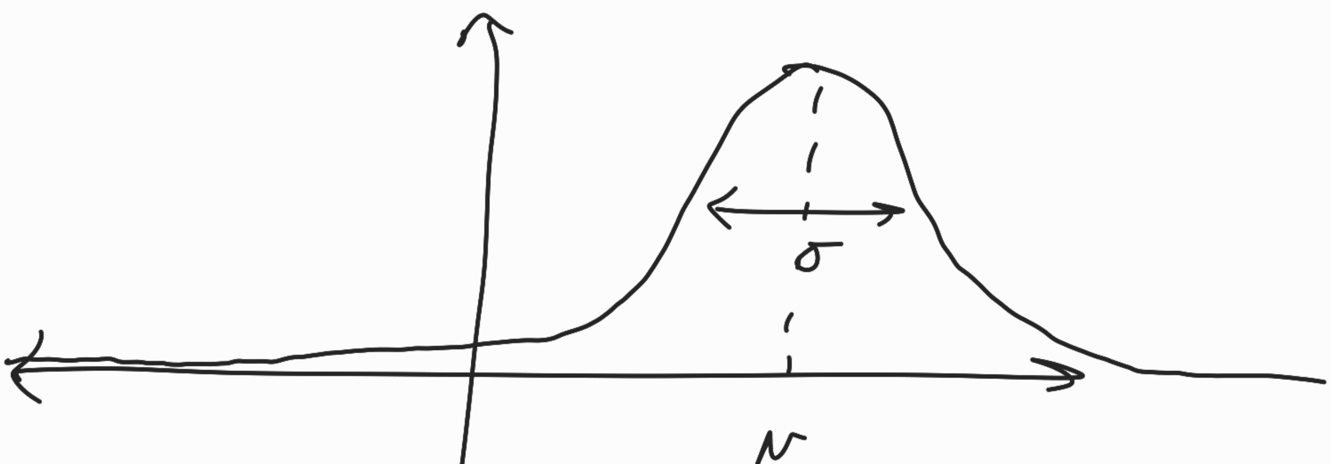
$$\int_R f_X(x)\,dx = 1 \quad \text{if } R = (-\infty, \infty)$$

$$P(a \le X \le b) = \int_{(a,b)} f_X(x)\,dx$$

$$= F_X(b) - F_X(a)$$

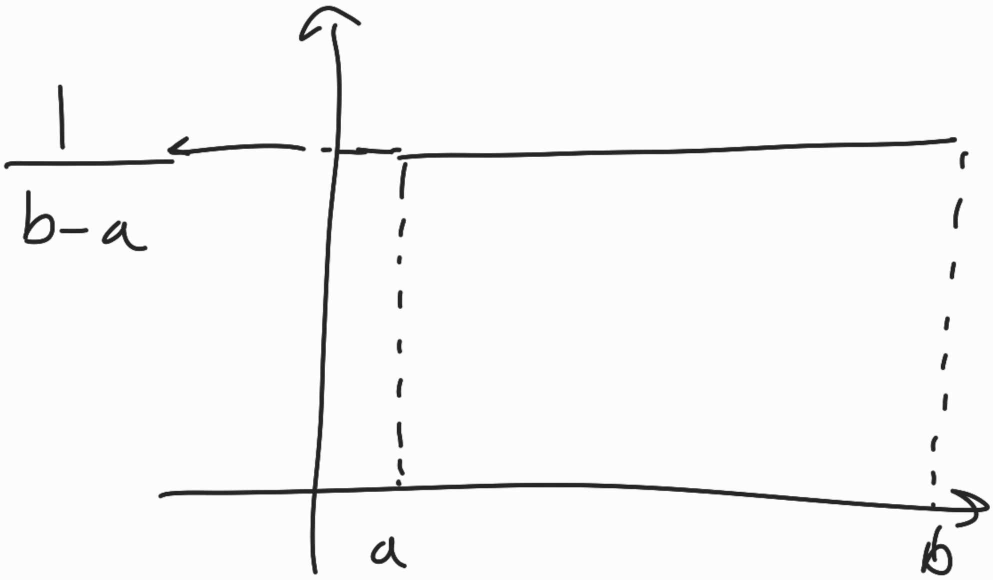
Gaussian pdf

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



Uniform pdf

$$f_X(x) = (b-a)^{-1}, \quad a \leq x \leq b$$



Expected value of Continuous rv

$$E(X) = \int_{R} x f_X(x)\, dx$$

Countable Set

if $\exists$ a mapping (one-one onto)

subset of $\mathbb{N}$ to the set

From a subset of N to the set

Countable sets maybe finite or infinite.

$A = \mathbb{N}, \mathbb{Z}, \mathbb{Q}$

Uncountable → not Countable

$B = \mathbb{R}, (0,1), \mathbb{Q}^c$

Binomial exp →

$$(x+y)^n = \sum_{k=0}^{n} {}^nC_k \, x^k y^{n-k}$$

${}^nC_k$ = Binomial Coefficient.

multinomial Coeff :→

given a population of n elements

, let $n_1, n_2, n_3, \dots, n_k$ be positive integers such that $n_1 + n_2 + n_3 + \dots + n_K = n$. Then,

there are,

$$N = \frac{n!}{n_1! \times n_2! \times n_3! \times n_4! \times \dots \times n_K!}$$

ways to partition the population into K subgroups of sizes $n_1, n_2, n_3, \dots, n_K$

multinomial theorem

$$\left(\sum_{i=0}^{r} x_i\right)^n = \sum_{(n_1, n_2, \dots, n_r):\, n_1+n_2+\dots+n_r=n} \binom{n}{n_1, n_2, n_3, \dots n_r} x_1^{n_1} x_2^{n_2} \dots x_r^{n_r}$$

$$\binom{n}{n_1, n_2, n_3 \dots n_r} = \frac{n!}{n_1!\, n_2!\, n_3! \dots n_r!}$$

stirling's Approx

$$\boxed{n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n = \sqrt{2\pi n}\, n^n e^{-n}}$$

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

Probability: Axiomatic Definition :→

Samplespace → S

event → A

Prob. of A → $P(A)$

- Axiom 1: $0 \leq P(A) \leq 1$
- Axiom 2: $P(S) = 1$
- Axiom 3: For any Sequence of mutually exclusive ($A_i \cap A_j = \phi$), events $A_1, A_2, ---, A_n$ defined on sample space, $P\left(\bigcup_{n=i}^{\infty} A_i\right) = \sum_{i=1}^{n} P(A_i)$

Field & σ Field

Let $\mathcal{F}$ is a collection of subsets of set S. Then $\mathcal{F}$ is called a field if a $S \in \mathcal{F}$ & $\mathcal{F}$ is closed under compliment & finite union ie.

union, i.e.

1. $S \in \mathcal{F}$

2. If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$

3. If A_i $(i \to 1$ to $n) \in \mathcal{F}$ is countable set of sets then $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$

<u>σ-Field</u> [If countable union]

1. Largest σ-field is collection of all subsets of S

2. The smallest σ field $= \{\phi, S\}$

3. If A is a non-empty subset of S then smallest σ-F having A $= \{\phi, S, A, A^c\}$

★ Smallest σ-field having A & B

$= \{\phi, S, A, B, A^c, B^c, A \cup B, A^c \cap B^c, A^c \cup B, B^c \cup A, A \cap B^c, B \cap A^c,$ 4 more $\}$

<u>Measure</u> :-

A measure is a set function $\mu : \underset{\sigma\text{-field}}{\mathcal{F}} \to R$

with properties :-

1.. $\mu(A) \geq \mu(\phi) = 0 \; \forall A \in F$, &

2. If A_i is a sequence of disjoint sets then

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$$

Probability is a special measure

with $\mu(S) = 1$. & $\mu: F \to (0,1)$

Probability Space

Its a triplet (S, F, P)

set of events

Probability

Sample space

& $P: F \to [0,1]$ that assigns

probability to an event.

σ- field properties

Let F & G be two sigma field on Ω

. Prove that $F \cap G$ is also a sigma-field

on Ω..

1. $\phi \in F, \phi \in G \Rightarrow \phi \in F \cap G$

$S \in F, S \in G \Rightarrow S \in F \cap G$

2. $A \in F \cap G \Rightarrow A \Rightarrow F, A \in G$

$\Rightarrow A^c \in F, A^c \in G$

$\Rightarrow A^c \in F \cap G$

3. $A_i \in F \cap G \Rightarrow A_i \in F$ & $A_i \in G$

$\bigcup_{i=1}^{n} A_i \in F$ & $\bigcup_{i=1}^{n} A_i \in G$

$\Rightarrow \bigcup_{i=1}^{n} A_i \in F \cap G.$

Q Show that $F \cup G$ need not be a σ-field.

$F = \{\phi, \{1\}, \{3\,3\}, \{1, 4\,3\}\}$

$G = \{\phi, \{2\}, \{1, 3\}, \{1, 4, 3\}\}$

$F \cup G = \{\phi, \{1\}, \{2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$

$\{1\} \cup \{2\} \notin F \cup G$

$\{1\} \cup \{2\}$

Discrete Probability space

Let S be countable set, thats

finite or countably ∞.

Let $\mathcal{F}$ = set of subsets of S.

Let

$$P(A) = \sum_{k \in A} P(k) \quad \text{where } p(k) \geq 0$$

[events valid to us]

$$\& \sum_{k \in S} P(k) = 1$$

then S is called discrete Prob space.

Probability space satisfies

self-duality,

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

Inclusion-exclusion Principle

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum^{n} P(A_i)$$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1} P(A_i)$$

$$- \sum_{i>1}^{n} P(A_{i_1} \cap A_{i_2})$$

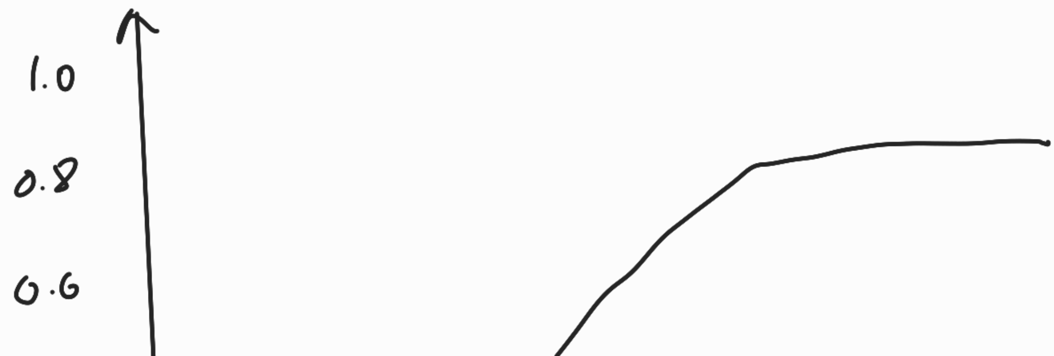
$$+ (-1)^{r+1} \sum_{i_1 < i_2 < \dots i_r} P(A_{i_1} \cap A_{i_2} \dots \cap A_{i_r}) + \dots$$

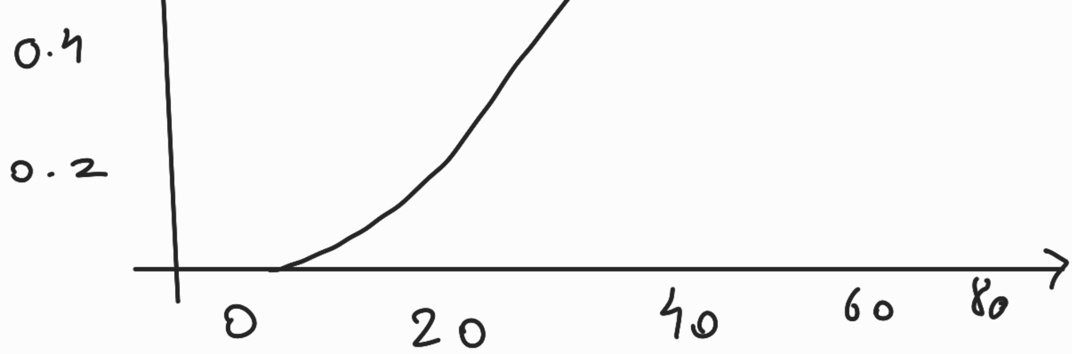
$$+ (-1)^{n+1} P(A_1 \cap A_2 \dots \cap A_n)$$

Birthday Paradox

What's the probability that atleast one pair have their birthday on same day. People $= n$

$$P(E) = 1 - \frac{{}^{365}P_n}{(365)^n}$$





Independence ⇒

when occurrence or non-occurrence of one event does not change our view on the probability of those events.

$$P(A \cap B) = P(B)P(A) \quad \{P(A/B)$$

If A & B are disjoint $P(A \cap B) = 0$

– If A & B Independent,

→ $P(A \cap B) = \phi$]

→ Two mutually exclusive event A & B are never

Independent:

mutually Independent :→

$$P\left(\bigcap_{i\in j} A_i\right) = \prod_{i\in j} P(A_i)$$

pairwise Independent :→

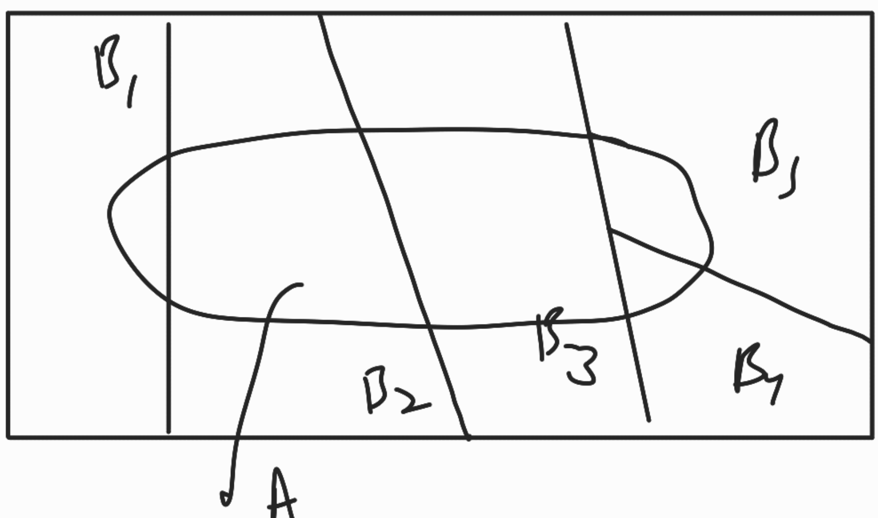
$A_i A_j \in S$

$$P(A_i \cap A_j) = P(A_i)\, P(A_j)$$

Pairwise Independent $\not\Leftrightarrow$ mutually Independent

Law of total Probability :→

B_i → mutually exclusive & exhaustive



$$A = (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \cup \ldots \ldots$$

$$\& \quad (A \cap B_1) \cap (A \cap B_2) \cap (A \cap B_3) \cap \ldots = \phi$$

$$\therefore P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_n)$$

Baye's Theorem :→

Likelihood

B being true given A is true

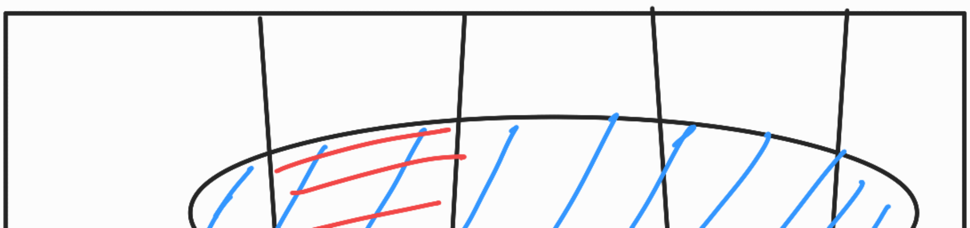
Prior

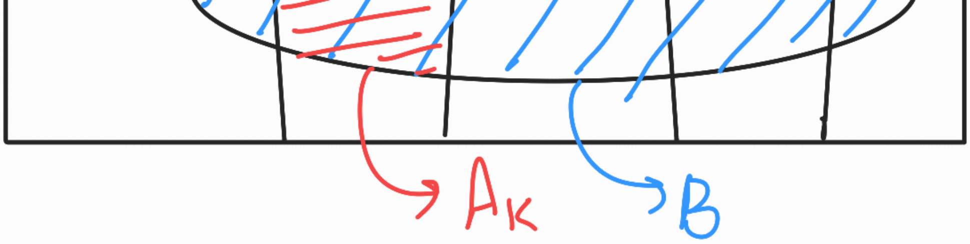
Probability of A being true before the new info

$$P(A/B) = \frac{P(B/A) \times P(A)}{P(B)}$$

Probability that B is true

$$P(A_k/B) = \frac{P(B/A_k)\,P(A_k)}{\sum_{i=0}^{n} P(B/A_i)\,P(A_i)}$$





Central Tendency

- Its a statistical measures that identifies a single value as representative of entire distribution.

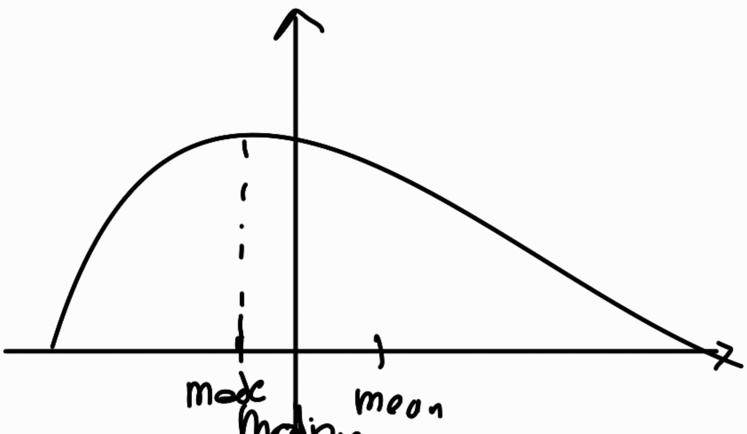
numerical
↓
Mean v/s Average

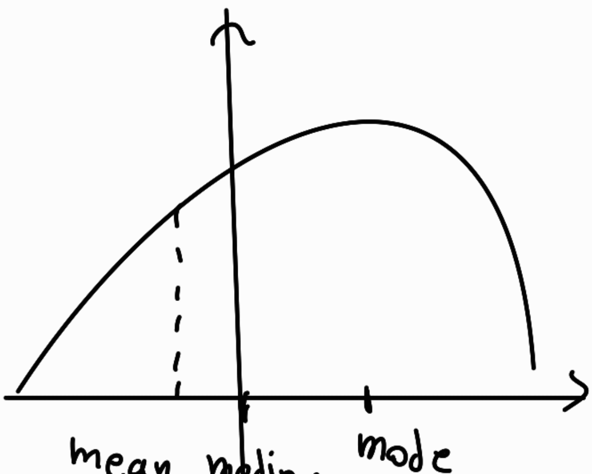
Mean
↓
may be arthimetical, geometric, Harmonic

Average
↓
only one way to calculate

positional
↓
Median
Mode

- mean is skewed by the outlier while median is not.





Skewed Left   Skewed Right

Measure of Dispersion :→

- Range
- Mean Deviation
- Variance & Std. deviation

Random Variable

- We map events to real number so as to express probability in a better way.
- Thus Random Variable is a function from Sample space to Real line.

$$X : S \to R$$

Another definition :→

A random variable X is a function $X : S \to R$ such that $\{w \in S : X(w) \leq x\} \in F$

$X: S \to R$ such that $\{$

for each $x \in R$

$X: S \to R$ is F-measurable if pre-Image of every set of the form $(-\infty, x]$ is an event

$$X^{-1}(-\infty, x) = \{w \in S : X(w) \leq x\} \in F$$

- Just as not all subsets are necessarily events, not all functions from S to R are random variables..

- To Find the probability, we define an event for the random variable. The event $X = a$ is equivalent to the set.

$$A = \{x \in S : X(x) = a\}$$

Thus, $P[X=a]$ means probability of

event for which $X=a$.

or, $X^{-1}(a) = \{w \in S \mid X(w) = a\}$ gives event A.

$$\therefore A = X^{-1}[P(X=a)]$$

$$X \leq a = \{x \in S : X(x) \leq a\}$$

$$X > a = \{x \in S : X(x) > a\}$$

$$a < X < b = \{x \in S : a < X(x) < b\}$$

<u>Discrete R.v.</u>

→ takes only countable number of values.

→ The probabilities are summarized by PMF.

<u>PMF</u> :→

The probability mass function of a random variable X is a function $f_X : R \to [0,1]$ which specifies the probability of obtaining

... denote a PMF as.

a number $X(w) = x$ denote

$$f_X(x) = P(X = x)$$

also, $f_X(x) \geq 0$

& $\sum_x f_X(x) = 1$

Continuous r.v. :⇒

- Can't have certainity at one single point. Thus $P(x = X) = 0$ for a Continuous - r.v..

PDF :⇒

The function $f_X(x)$ is a probability density function (PDF) for the continuous r.v. X, defined over the set of real numbers, if.

- $f(x) \geq 0 \quad \forall x \in R$

- $$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

- $$\boxed{P[a \le X \le b] = \int_{a}^{b} f_X(x)\,dx}$$

Cumulative distribution

$$F_X(x) = P(X \le x)$$

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \sum_{\{i \,:\, i < x \,\&\, i \in S\}} f_X(x)$$

- probability upto that point

- If ∞ in domain,

$$\boxed{F_X(\infty) = 1}$$

- Since $f(x) \ge 0$

Since

$\therefore$ $F_X(x)$ is increasing.

- CDF for discrete r.v. is a step function. (Right continuous)

- $P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$

- $P_X(x > x_0) = F_X(x) - F_X(x_0)$

CDF $\Rightarrow$ PMF

- discrete:

$$\text{PMF} \rightarrow f_X(x_k) = F_X(x_k) - F_X(x_{k-1})$$

difference in consecutive values of x

- Continuous:

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

Expectation

$$E[X] = \sum_{x \in X} x F_X(x)$$

$$= \int x F_X(x) \, dx$$

Indicator function :→

$$I : A \to R$$

$$I(\omega) = \begin{cases} 0 & x \in A^c \\ 1 & x \in A \end{cases}$$

$$E[X] = 1 \cdot P(A) + 0 \cdot P(A^c)$$

$$E[X] = P(A)$$

Linearity of Expectation :→

$$E[aX+b] = aE[X]+b$$

<u>Proof</u> :- $$E[aX+b] = \sum_x (ax+b) f_X(x)$$

$$= a\sum x f_X(x) + b\sum f_X(x)$$

$$= aE[X] + b$$

(Same for Continuous)

<u>Properties of expectation</u> :-

· when a r.v. X is a discrete random Variable with PMF $F_X(x)$ & $g: R \to R$ is a function of X, then

$$E[g(X)] = \sum_x g(x)\, f_X(x)$$

If Continuous

$$E[g(X)] = \int^{\infty} g(x)\, F(x)\, dx$$

$$E[g(X)] = \int_{-\infty}$$

<u>Proof</u> :- (discrete)

$$Y = g(X)$$

$$E[Y] = \sum_y y \cdot P(Y = y)$$

$$= \sum_y y \cdot P\left(X = x ; g(X) = y\right)$$

$$= \sum_y y \cdot \left(\sum_{x : g(x) = y} f_X(x) \right)$$

$$= \sum_y \sum_{x : g(x) = y} y \cdot f_X(x)$$

$$= \sum_x y \cdot f_X(x)$$

- $E[X_1 + X_2] = E[X_1] + E[X_2]$

$E[X_1 + X_2] = E[X_1] + E[X_2]$

- $E[g(x)] \neq g(E[x])$
- $E[X]E[Y] \neq E[X] + E[y]$
- $E E[X^2] \neq E[X^2]$
- If r.v. X & Y are Independent,

$$E[X.Y] = E[X] + E[Y]$$

n^{th} moment

$$E[X^n] = \sum_i x_i^n f_X(x_i)$$

$$= \int_{-\infty}^{\infty} x^n f(x)$$

n^{th} Central moment $\Rightarrow$

$$E[(X - \mu(x))^n] = \sum_i (x_i - \mu(x))^n f_X(x)$$

$$= \int (x_i - \mu(X))^n F_X(x_i)$$

If $n = 2 \Rightarrow E[(X-\mu(X))^2] = \int (x - \mu(X))^2 F_X(x) dx$

Variance :→

- Useful when expectation of r.v. are similar.

- 2nd order central moment,

$$\sigma_X^2 = E[(X-\mu)^2]$$

$$= E[X^2 - 2X\mu - \mu^2] \quad [\mu = E[X]]$$

$$\boxed{\sigma_X^2 = E[X^2] - \mu^2}$$

- If X & Y are Independent,

$Var(X) + Var(Y)$

$Var(X+Y) = Var(X) + Var(Y)$

- $Var(aX+b) = a^2\, Var(X)$

Standard deviation

$$s.d. = \sqrt{Var(x)} \quad \text{(same units as } \underline{\underline{x}}\text{)}$$

- $\sigma_{g(x)} = E\left[(g(X) - \mu_{g(x)})^2\right]$

$$= \sum_x \left(g(X) - \mu_{g(x)}\right)^2 f_X(x)$$