Class 1 : 7^{th} Jan 2019

Random Experiment

An experiment whose outcome is unknown.

eg.- Tossing a coin.

Sample Space

Given a random experiment, set of all possible outcomes in a sample surface

eg.- $\Omega = \{H, T\}$ tossing a coin

$\Omega = \{1, 2, 3, 4, 5, 6\}$ rolling a die.

$\Omega = [0, \infty)$ waiting in a queue.

Event

Subset of sample space.

discrete

Continuous.

Assignment of probabilities to events:

Case 1. Discrete / finite / all events are equally likely.

Let A be an event

$$P(A) = \frac{\# A}{\# \Omega}$$ Relative freq. assignment.

Case 2. Discrete / finite / all events are not equally likely.

- Run the expt. 'n' no. of times.
- By f_A denote the relative frequency.

$$f_A = \frac{\text{no. of times event A occured in } n}{n}$$

$$P(A) = \lim_{n \to \infty} f_A$$

Observation

Only for certain key events, we need to assign probabilities. For instance, we need to assign only P(1), P(2) ... P(5), P(6) in case of rolling a die. Then the probability of all other events are automatically assigned.

- Because we intutively assume certain laws of probability.
- We will formalise these laws/intutions into axioms of probability.

why do we need axiomatic definition?
- It works in all possible scenes.
- It matches with the ideas sketched.

Event

A subset of Ω. Union/intersection of an event is also an event.

Definition (σ-field)

A collection of sets $\mathcal{A}$ is called σ-algebra/σ-field if

1. $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$
2. $A_1, A_2 \dots A_n \in \mathcal{A}$ then $\bigcup_{i=1}^{n} A_i$ and $\bigcap_{i=1}^{n} A_i \in \mathcal{A}$

example $\Omega = \{H, T\}$

$$\mathcal{A} = 2^{\Omega} = \{\phi, \{H,T\}, \{H\}, \{T\}\}$$

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

$\mathcal{A} = 2^{\Omega}$: will contain 2^6 elements.

Definition

Let Ω be the sample space.
Let $\mathcal{A}$ be a σ-algebra associated with Ω

Then a probability measure 'P' is a function of $\mathcal{A}$ to $\mathbb{R}$

i.e. $P : \mathcal{A} \longrightarrow \mathbb{R}$

such that:

1. $P(A) \geq 0 \quad \forall A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. If A_i are mutually disjoint for $i = 1, 2 \dots n$ (i.e. $A_i \cap A_j = \phi$ for $\forall i \neq j$) then

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

PROPERTIES OF PROBABILITY :

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

$\Rightarrow$ $A \cup A^c = \Omega$

$P(A \cup A^c) = P(\Omega) = 1$

$P(A) + P(A^c) = P(A \cup A^c)$

$\because$ both are disjoint.

2. $P(\phi) = 0$

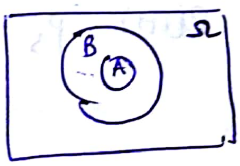
take $A = \Omega$ in previous.

3. if $A < B$ then $P(A) < P(B)$

$B = A \cup B \cap A^c$

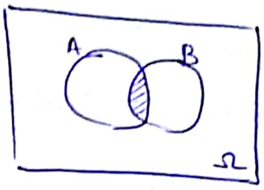
$P(B) = P(A) + P(B \cap A^c)$

always +ve.



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

$$A \cup B = (A \cap B^c) \cup (A^c \cap B) \cup (A \cap B)$$



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

notice that

$$A = (A \cap B^c) \cup (A \cap B)$$

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

similarly $P(B) = P(A^c \cap B) + P(A \cap B)$

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

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

Similarly using induction

$P(A_1 \cup A_2 \cup A_3 \ldots \cup A_n) =$

Example

1. Experiment : toss pair of coin.

$$\Omega = \{(H,T), (H,H), (T,H), (T,T)\}$$

$$\mathcal{A} = 2^{\Omega}$$

Assign/Assume $P(HH) = P(TT) = P(HT) = P(TH) = \frac{1}{4}$

$P(HH) = P_1$

$P(HT) = P_2$

$P(TH) = P_3$

$P(TT) = 1 - (P_1 + P_2 + P_3)$

$0 \leq P_1, P_2, P_3, P_4 \leq 1$

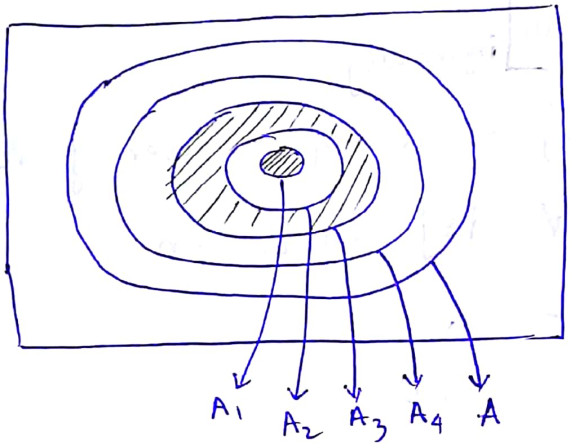
$0 \leq P_4 \leq 1$

Theorem 1

Let $A_1 \subset A_2 \subset A_3$ ---- (Sequence of increasing sets.)

and $A = \bigcup_{i=1}^{n} A_i$ then

$$P(A) = \lim_{n \to \infty} P(A_n)$$



Let $B_1 = A_1$

$B_2 = A_2 \cap A_1^c$

$B_3 = A_3 \cap A_2^c$

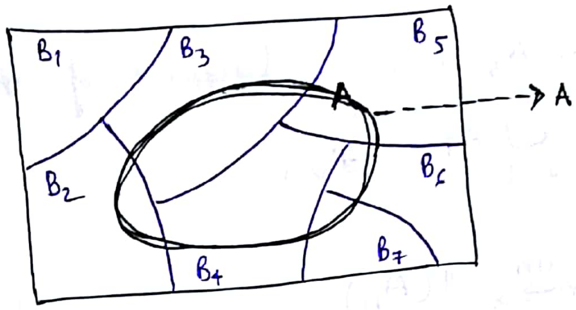
$B_4 = \vdots$

$\vdots$

$B_n = A_n \cap A_{n-1}^c$

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

$$\therefore P(A_n) = p\left(\bigcup_{i=1}^{n} B_i\right) = \sum_{i=1}^{n} P(B_i)$$



[union of B = Ω
intersection of B = null] informally.

Let $\{B_1, B_2 \, --- \, B_n\}$ be a disjoint cover of Ω

$B_i \cap B_j = \phi \quad \forall \quad i \neq j = 1, 2, --- n$

and $\bigcup_{i=1}^{n} B_i = \Omega$.

Then for any event $A \in \mathcal{A}$

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

x

Class 2 : 8^{th} Jan 2019.

CONDITIONAL PROBABILITY

Let A and $B \in \mathcal{A}$ such that $P(B) > 0$.
conditional prob. of A given B has occured

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

example. R = rolling a die.

A = 2 appears on top face of die.
B = even no. appears on top face of die.

$A = \{2\}$ $\quad$ $B = \{2, 4, 6\}$

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

$$P(A|B) = \frac{\#(A\cap B)/\#\Omega}{\# B/\#\Omega}$$ motivation for the formula.

Product rule of probability.

$$P(A\cap B) = P(A|B)\cdot P(B) \quad \text{--- } ①$$

If $P(A) > 0$

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

$$P(A\cap B) = P(B|A)\cdot P(A) \quad \text{--- } ②$$

from ① and ②

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

$$\left.\begin{aligned} P(A|B) &= \frac{P(B|A) \cdot P(A)}{P(B)} \\ P(B|A) &= \frac{P(A|B) \cdot P(B)}{P(A)} \end{aligned}\right\} \rightarrow \boxed{\text{Baye's Rule.}}$$

* Consider $\{B_1, B_2, \ldots B_n\}$ disjoint cover of Ω such that

$$B_i \cap B_j = \phi \quad \text{for } i \neq j$$

and $\bigcup_{i=1}^{n} B_i = \Omega$

for $A \in \mathcal{A}$

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

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

$$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \ldots$$
$$\ldots + P(A|B_n) \cdot P(B_n)$$

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i) \quad \Big\} \text{ total probability law.}$$

$$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)} \quad \Big\} \text{ Modified Baye's Law.}$$

example

A = event of getting A grade.
B_i = ~~event~~ Attendance record.

or.

A = lung cancer.
B_i = Smoker / Non-smoker.

Class 3

14th Jan 2019

Principle independence of events :-

$$P(A \cap B) = P(A) \cdot P(B)$$

It is clear that A, B are pair wise and

Independent Events

* Let A, B be 2 events from $(\Omega, \mathcal{A}, P)$ are called pairwise indep. if

$$P(A \cap B) = P(A) \cdot P(B)$$

* Let A, B, C are 3 events from $(\Omega, \mathcal{A}, P)$ we call A, B, C mutually indep. if

$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(A \cap C) = P(A) \cdot P(C)$$
$$P(B \cap C) = P(B) \cdot P(C)$$
$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$

* We can generalise def^n of mutual independence to the events

$A_1 \; A_2 \; A_3 \; --- \; A_n$ in an inductive way.

- Any sub collection of $A_1, A_2 --- A_n$ containing atleast 2 events and atmost any events is mutually indep.
- $P(A_1 \cap A_2 \cap --- \cap A_n) = P(A_1) \cdot P(A_2) \; --- \; P(A_n)$

* Setting of probability space in any of expected random experiment (with any 2 possible outcomes)

Experiment

Tossing a die $\qquad \Omega = \{0, 1\}$

$$\mathcal{A} = \{ \phi, \Omega, \{0\}, \{1\} \}$$

$$P(\{1\}) = p$$

$$P(\{0\}) = 1 - p$$

Set up probability space when we repeat the expt. 'n' number of times within the assumptions that expectations are independent of each other

For this compound experiment :-

$\Omega = \{ y_1, y_2, \cdots y_n \}$

$\mathcal{A} = 2^{\Omega}$

prob. of R successes in an experiment of specified conditions, consider an element $w \in \Omega$

$$w = y_1 \; y_2 \cdots y_n$$

$$w = (\underbrace{1 \cdot 1 \cdot 1 \cdot \cdots 1}_{R \text{ times.}} \cdot \underbrace{0 \cdot 0 \cdot 0 \cdot 0}_{n-R \text{ times}})$$

A_i = success at i^{th} for $i = 1, 2, \cdots$

$$P(R) = P(y_1 \cdots y_n)$$

$$= P(\underbrace{1 \cdot 1 \cdot 1 \cdot \cdots 1}_{R \text{ times.}} \cdot \underbrace{0 \, 0 \cdot 0 \cdots}_{n-R \text{ times.}})$$

$$= P(A_1 \cap A_2 \cap \cdots A_R \cap A_{R+1}^c \cap A_{R+2}^c \cap \cdots \cap A_n^c)$$

$$= P(A_1) \cdot P(A_2) \cdot \cdots P(A_R) \cdot P(A_{R+1}^c) \cdots$$

$$= {}^nC_R \; p^R \cdot (1-p)^{n-R}$$

$\uparrow$

for exactly R successes.

Scanned by CamScanner

Example A box containing 10 balls
6 red 60%
4 blue. 40%

Case 1 Pick a ball and replace it in the box and pick a ball again.

A_1 = red ball is picked in 1^{st} trial
A_2 = Blue ball is picked in 2^{nd} trial.

$$P(A_2 | A_1) = P(A_2) = \frac{4}{10} = 0.4$$

This process is called sampling with replacement.

Case 2 Pick a ball and do not replace it. Pick a second ball.

A_1 = red in 1^{st} trial
A_2 = blue in 2^{nd} trial.

$$P(A_2 | A_1) = \frac{4}{9} \neq P(A_2) = \frac{6}{10} \cdot \frac{4}{9} + \frac{4}{10} \cdot \frac{3}{9} = \frac{2}{5}.$$

then A_1 and A_2 are not indep.

* Increase 'n' and keep population of red and blue balls const. A_1 & A_2 start behaving more independent.

Uniform probability spaces :-

Expt. picking up a no. randomly from a given set.

1.] Given set is finite and discrete

$\{1, 2, ---- n\}$

probability of two subsets of Ω with an equal size ~~can~~ be equal. (This is space prob. principle)

Here we associate size for cardinality of set.
As a result, we can see that the singletont set are assigned prob. of $1/n$.
(These spaces are called discrete uniform prob. space)

2.] The given set Ω is finite and an interval of $\mathbb{R}$ of the type $\Omega = [0, 1]$

principle of uniformity :-

- pick two subsets of Ω with equal 'size' then they should have same probability.
- Take an interval $(a, b) \leq [0, 1]$
- Size of (a,b) is associated with length of (a,b) which is $b-a$.

* probability assignment by applying principle of uniformity.

Any two interval of same length are assigned with prob. equal to their length.

$P([a,b]) = (b-a)$ for any $(a, b) \leq [0, 1]$

The σ-algebra $\mathcal{A}$ is the set of subsets of $[0,1]$ which are generated as countable unions and intersection of intervals.

This is continuous uniform prob. space.

Remarks.—

1. Any singleton set $\{a\}$ has zero probability.
2. Consider set of rational no.s in $[0,1]$
   denote $\rightarrow \mathbb{Q} \cap [0,1]$

They form (1) and hence

$$P(\mathbb{Q} \cap [0,1]) = 0$$

Scanned by CamScanner

15th Jan 2019

Ques A machine contains 4 components in parallel with 0.1, 0.2, 0.3, 0.4 as their prob. of failure respectively. The machine fails if all components fail simultaneously. Note that failure of machine is independent. What is the prob. that once the machine has been started, will not fail?

Solⁿ→ Let A_i = component i fails

$A_1 \cap A_2 \cap A_3 \cap A_4$: machine fails.

$$P(A_1 \cap A_2 \cap A_3 \cap A_4)^c = 1 - P(A_1 \cup A_2 \cup A_3 \cup A_4)$$

$$= 1 - P(A_1) P(A_2) \cdot P(A_3) \cdot P(A_4)$$

$$= 1 - 0.1 (0.2) (0.3) (0.4)$$

$$= 1 - 0.0024$$

$$= 0.9976.$$

A rocket engine fails if one key component fails. The prob. of failure of this key comp. is 5%. In order to increase the success prob. of the rocket engine, assembly of this key component in parallel is proposed so that the engine fails if all these key components fail simultaneously. What are the min no. of key comp.s in parallel so that engine has 99% of success prob.?

Solⁿ.

$$1 - (0.05)^n \geq 0.99$$

$$0.01 \geq (0.05)^n$$

$$n \log_{10}(0.05) \leq 2$$

$$n \leq \frac{2}{\log_{10}(0.05)}$$

A course on prob. is very famous & students are allowed to register after the teacher consents. It is observed that 20% times after obtaining consent, students fail to register. There are ~~xxx~~ 100 seats available in class. If a teacher consents 102 students, what ~~are~~ is prob. that all students will be ~~xxx~~ accomodated in class

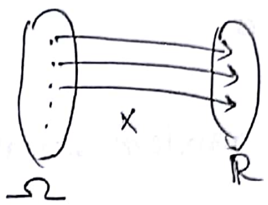
Sol →
$$1 - (0.8)^{102} - 102 \times (0.8)^{101} \times 0.2$$

RANDOM VARIABLE

Random Experiment.

$(\Omega, \mathcal{A}, P)$

A random ~~expt.~~ variable X is a function from Ω to $\mathbb{R}$ if $\{w : X(\Omega) \in \mathbb{R}\} \in \mathcal{A}$



$\forall x \in \mathbb{R}$, $\{w : X(w) = x\} = X^{-1}(x)$

inverse image of X

$\{w \mid X(w) = x\} \equiv (X = x)$

$P(X = x) = P(\{w \mid X(w) = x\})$

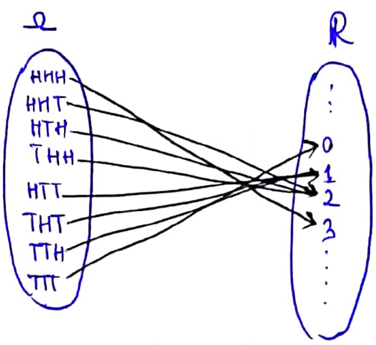
Expt. 3 fair coins are tossed indept.

$\Omega = \{$ HHH, HHT, HTH, THH, TTH, THT, TTH, TTT $\}$

$\mathcal{A} = 2^{\Omega}$ this is non numeric.

$P(HHH) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

X: no. of heads in 3 tosses.



$x = 2$, $X^{-1}(2) = \{w \in \Omega \mid X(w) = 2\}$

$$P(X=2) = P(\{w \in \Omega \mid X(w) = 2\})$$
$$= P(HHT, HTH, THH)$$
$$= \frac{3}{8}$$

$x = 8.7$

$$P(X=x) = P(X=8.7) = 0.$$

countably finite, countably infinite $\rightarrow$ discrete random variable.

* Range of X : $\{x \in \mathbb{R} : X(x) > 0\}$

* A random variable is called a discrete random variable if the range of X is finite or countably infinite.

In previous example,

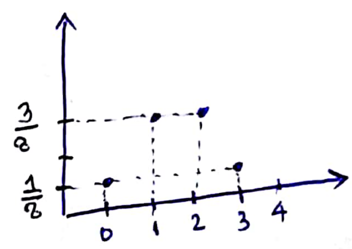
$R_X = \{0, 1, 2, 3\}$ and hence X is a discrete random variable.

DEFINITION $\rightarrow$ A real valued function $f(x) = P(X=x)$ is called a discrete density or probability mass function (pmf.) of the discrete random variable X.

$$f(x) = \frac{1}{8}, \quad x=0$$
$$= \frac{3}{8}, \quad x=1$$
$$= \frac{3}{8}, \quad x=2$$
$$= \frac{1}{8}, \quad x=3$$

$f(x) = 0 \ \forall$ other values of x.

May 1



Way 2

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

Example 1 Bernouli Random Variable.

expt $\rightarrow$ Tossing a coin.

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

$P(H) = p$ ... unfair coin. $X(\text{Heads}) = 1$

$P(T) = 1-p$ $X(\text{Tails}) = 0$

$P(X=1) = p$

$P(X=0) = 1-p$

$$f(x) = \begin{cases} 1-p & ; x=0 \\ p & ; x=1 \\ 0 & ; \forall x \text{ other than these} \end{cases} \quad \text{---> pmf.}$$

$x \sim \text{Bernoulli}(p)$

"x follows bernouli, with parameter of the distribution : p"

$p \rightarrow$ prob. of success.

$R_X = \{0, 1\}$

Example 2 Binomial.

Expt $\rightarrow$ n independent bernouli trials are performed.

X: no. of successes in 'n' trials.

$R_X = \{0, 1, 2, --- n\}$

$$f(x) = P(X=x) = \begin{cases} {}^nC_x \, p^x (1-p)^{n-x} & ; x = 0, 1, 2 \ldots n \\ 0 & ; \text{otherwise.} \end{cases}$$

$x \sim \text{Binomial}(n, p)$

Example 3 Geometric.

expt → perform indep. bernouli trials (p) until the 1st success.

$$\Omega = \{S, FS, FFS, \ldots\}$$
$$\downarrow \quad \downarrow \quad \downarrow$$
$$p \quad p(1-p) \quad (1-p)^2 p \ \ldots$$

X: no. of failures before the 1st success.

$R_X = \{0, 1, 2 \ldots\}$ (discrete random variable)

$$f(x) = \begin{cases} p(1-p)^x & ; \ x = 0, 1, 2 \ldots \\ 0 & ; \ \text{otherwise.} \end{cases}$$

for any random expt, $P(\Omega) = 1$

let X be a R.V. defined in Ω. Let R_X be the range of this random variable.

$$R_X = \{x_1, x_2, \ldots\}$$

$$\sum_{x_i \in R_X} f(x_i) = \sum_{x_i \in R_X} P(X = x_i)$$

$$= \sum_{x_i \in R_X} P(\{w \mid X(w) = x_i\})$$

$$= 1$$

Random expt → $\Omega, \mathcal{A}, P$ → X → pmf.

PROPERTIES OF PMF

1. $f(x) \geq 0 \quad \forall\ x \in \mathbb{R}$

2. $\{x : f(x) > 0\} = \{x_1, x_2, \ldots\}$ is a finite or countable set.

3. $\sum_{x_i} f(x_i) = 1$

Any function f which satisfies these properties is called a discrete density or probability mass function.

Example 1 Uniform density pmf.

$$f(x) = \begin{cases} \frac{1}{S} & ; x \in \{x_1, x_2, \ldots x_S\} \\ 0 & ; \text{otherwise.} \end{cases}$$

Example 2 Geometric pmf.

$$f(x) = \begin{cases} p(1-p)^x & ; x = 0, 1, 2, \ldots \\ 0 & ; \text{otherwise.} \end{cases}$$

geometric series will converge to one.

Example 3 Negative Binomial pmf (α, p)

$$f(x) = \begin{cases} p^\alpha \binom{\alpha+x-1}{x} (1-p)^x & ; x = 0, 1, 2, \ldots \\ 0 & ; \text{otherwise.} \end{cases}$$

"To get first 'α' successes"

→ generalization of geometric pmf.

[performing indp. bernoulli trials until 'α' no. of successes are observed] eg. for $\alpha = 2$

$\Omega = \{ HH, HTH, THH, \text{-----} \}$

Example 4 Poisson density (λ) $\lambda > 0$

$$f(x) = \begin{cases} \dfrac{e^{-\lambda} \lambda^x}{x!} & ; x = 0, 1, 2, \ldots \\ 0 & ; \text{otherwise}, \end{cases}$$

$\Sigma \dfrac{\lambda^x}{x!} \rightarrow e^{\lambda}$ $\therefore$ $e^{-\lambda}$ is a normalizing const.

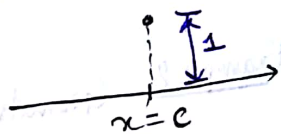
SPECIAL EXAMPLES OF RANDOM VARIAIBLES.

Example 1 Const. Rand. Var.

$X(w) = c \quad \forall \; w \in \Omega$

$P(X = c) = 1$

$P(X \neq c) = 0$



Example 2 Indicator Rand. Var.

let $A \in \mathcal{A}$

$$\chi_A(w) = 1 \quad \forall \; w \in A$$
$$= 0 \quad \forall \; w \in A^c$$

indicates when event A has occured.

* Hyper Geometric Distribution

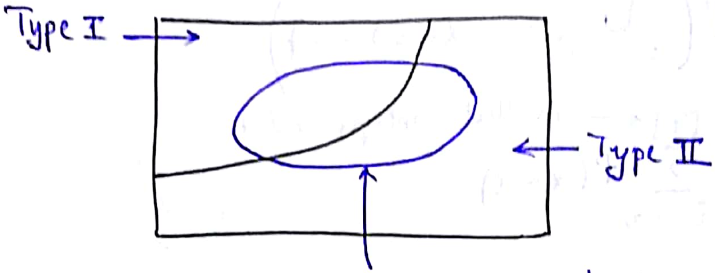
Population of 'r' objects

Type I objects $\rightarrow r_1$

Type II objects $\rightarrow r - r_1 = r_2$

Let sample of size 'n' is chosen from this population. $(n \leq r)$

X = no. of objects of type 1 in the random space.



$$P(X = x) = \frac{{}^{r_1}C_x \; {}^{r-r_1}C_{n-x}}{{}^{r}C_n} \qquad x = 0, 1, \ldots$$

$$= 0 \quad \text{otherwise.}$$

Scanned by CamScanner

COMPUTATIONS WITH PMF

$(\Omega, \mathcal{A}, P)$

X: discrete random variable with pmf $f(x)$

Event: $X \cancel{x} \leq t$ for som real no. t.

$$= \{w : X(w) \leq t\}$$

$$= \bigcup_{i=-\infty}^{t} \{w : X(w) = i\}$$

$$P(X \leq t) = P\left(\bigcup_{i=-\infty}^{t} \{w : X(w) = i\}\right)$$

$[t] \leftarrow$ smallest integer value.

$$= \sum_{i=-\infty}^{[t]} P(X = i)$$

$\Rightarrow$ cumulative distribution function tells prob. upto a certain value. i.e. probability adds up.

CUMULATIVE DISTRIBUTION FUNCTION

Discrete random variable X with pmf $f(x)$; for every real number $t \in \mathbb{R}$ define

$$\left.\begin{aligned} F(t) &= P(X \leq t) \\ &= \sum_{x \leq t} f(x) \end{aligned}\right\} \text{cumulative distribution function.}$$

example $x \sim$ uniform $(s = 10)$

$$f(x) = \frac{1}{10} \qquad x = 0, 1, 2, \ldots, \cancel{10}\ 9$$

$$= 0 \qquad \text{otherwise.}$$

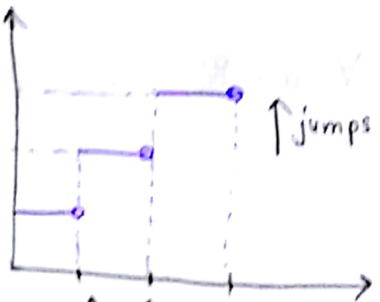
clearly, for $t < 0$, $F(t) = 0$

$F(0) = f(0) = \frac{1}{10}$

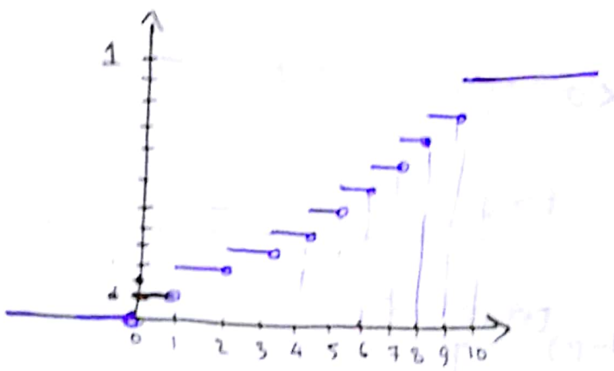
for any $t \in [0,1)$; $F(t) = \frac{1}{10}$

$F(1) = f(0) + f(1) = \frac{2}{10}$

for any $t \in [1,2)$; $F(t) = \frac{2}{10}$



↑ jumps = prob. at that point.



<u>cumulative dist. function.</u>

→ Always atleast non-decreasing
→ non-zero function.
→ bounded function [0, 1]
→ step function (discrete dist.)
→ Right continuous.

<u>CDF</u>: Cumulative distribution function.

Let X be a discrete random variable with pmf $f(x)$

For any $t \in \mathbb{R}$, we define

$$F(t) = \sum_{x \leq t} f(x)$$

1. Clearly F is a function from $\mathbb{R}$ to $\mathbb{R}$. F is a non-decreasing function

for $t_1 < t_2 \Rightarrow F(t_2) \leq F(t_1)$

$\lim_{t \to -\infty} F(t) = 0$; $\lim_{t \to \infty} F(t) = 1$.

The graph of $F(t)$ is like a staircase, jumps occuring exactly at the points which are in the range of X.

Right continuity of "$F(x)$"

$$\lim_{x \to a^+} F(x) = F(a) \quad \forall\ a \in \mathbb{R}.$$

Example. $X \sim$ geometric (p)

$$f(x) = p(1-p)^x \qquad x = 0, 1, 2 \dots$$
$$= 0 \qquad \text{otherwise.}$$

$$F(t) = 0 \qquad \forall\ t < 0$$

$$F(t) = \sum_{x=0}^{[t]} p(1-p)^x \qquad t \geq 0$$

$$= p\left[\frac{1-(1-p)^{[t]+1}}{1-(1-p)}\right]$$

$$= 1 - (1-p)^{[t]+1}$$

for this geometric ~~varia~~ distribution ;

$$P(X \geq x) = 1 - P(X < x)$$
$$= 1 - P(X \leq x-1)$$
$$= 1 - F(x-1)$$
$$= 1 - \{1 - (1-p)^{[x-1]+1}\}$$
$$= (1-p)^{[x]}$$

events.
$X \geq x$ and $X < x$ are disjoint.
$X \geq x \cup X < x = \Omega$

lim

bution

prove. $P(X > n+m \mid X > n) = P(X > m)$ [Very Very Imp.]

Memory-less property.

$$\frac{P(X > n+m \cap X > n)}{P(X > n)} = \frac{P(X > n+m)}{P(X > n)}$$

Physical Significance,

H, TH, TTH, TTTH, ----

$$\boxed{P(X > n+m \mid X > n) = P(X > m)}$$ Very Imp. Property.

at least $n+m$ failures. n failures have occured. at least m failures.

No memory of the previous 'n' failures.

Rand. Expt. $\longrightarrow \Omega, \mathcal{A}, P \longrightarrow$ discrete Rand. var $\longrightarrow$ pmf $f(x)$ $\longrightarrow$ cdf $F(x)$.

CONTINUOUS RANDOM VARIABLE

X is a random variable $(\Omega, \mathcal{A}, P)$. Then probability of the event $X \leq x$ is rep. by

$$F(x) = P(X \leq x) \qquad \forall\, x \in \mathbb{R}.$$

This function F is called the cumulative distribution function $F(x)$ of r.v. X aka. $F_X(x)$

Properties of CDF.

1. $0 \leq F_X(x) \leq 1 \qquad \forall\, x \in \mathbb{R}.$

2. $\lim_{x \to +\infty} F_X(x) = 1$

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

The function $F_X(x)$ is non. decreasing.

3. i.e. for $x_1 < x_2$

$\Rightarrow F_X(x_1) \leq F_X(x_2)$

Right continuity of the function.

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

OR

for every $a \in \mathbb{R}$ and $\delta > 0$

$$\lim_{\delta \to 0} \left(F_X(a+\delta) - F_X(a) \right) = 0$$

If any function has these properties, it will give rise to a probability space.

Lebe-sgue Decomposition Theorem

If a function $F_X(x)$ satisfies the properties stated above, then this function can be represented as a sum of two functions say $G_X(x)$ and $H_X(x)$

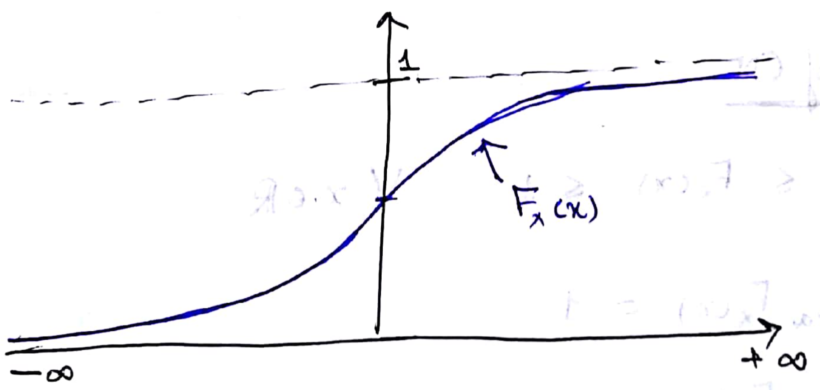
$$F_X(x) = G_X(x) + H_X(x)$$

where $G_X(x)$ is continuous and

$H_X(x)$ is a right continuous function with jumps coinsiding with those of $F_X(x)$ and $H_X(-\infty) = 0$

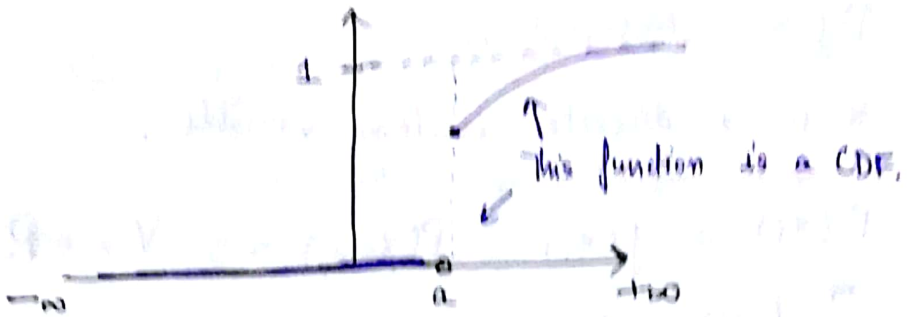
i.e. $\lim_{x \to -\infty} H_X(x) = 0$

eg.

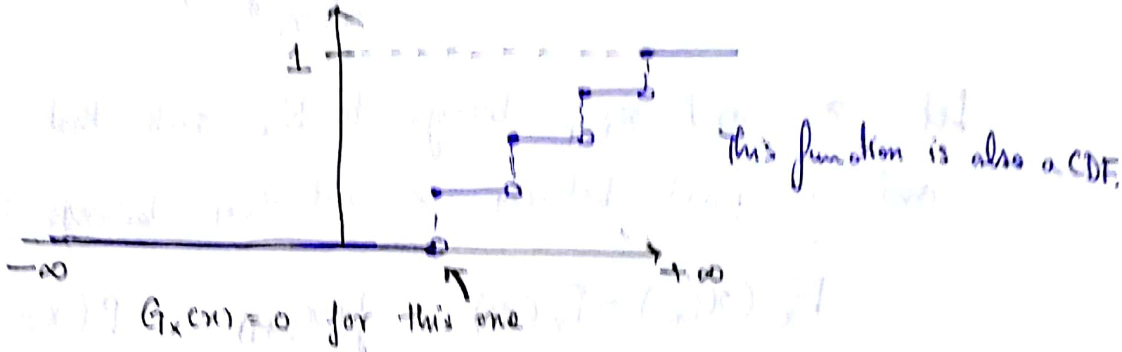


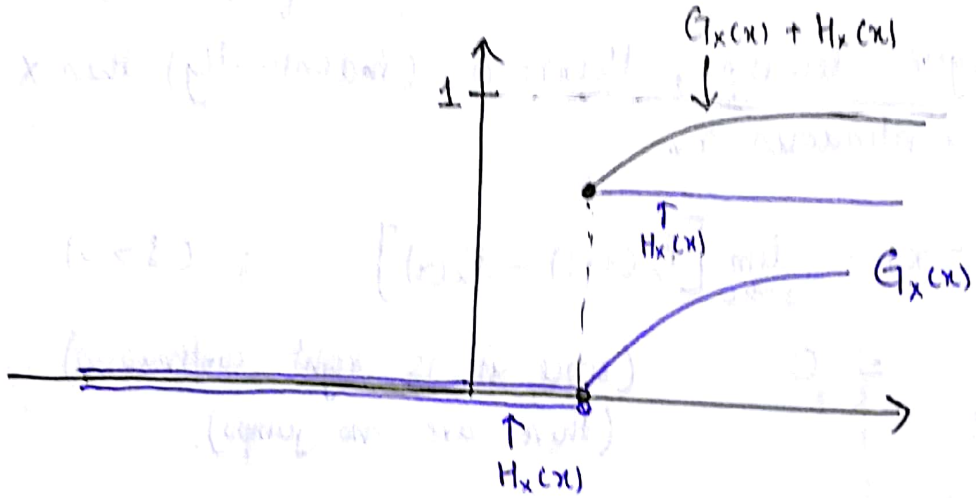
follows all the properties $\Rightarrow F_X(x)$ is a CDF.

Ex. →



Ex. →





In Lebesgue Decomposition, if $G_x(x) = 0$ (identically) then the random variable is called discrete random variable.

In Lebesgue Decomposition, if $H_x(x) = 0$ (identically) then the random variable is called continuous r.v.

If neither of the functions are identically zero, we call X a mixed variable.

Scanned by CamScanner

Def$\xrightarrow{n}$ Observations.

If X is a discrete random variable,

1. $P(x_i) = f(x_i) = P(X = x_i) > 0 \quad \forall x \in R_X$
2. $\sum_{x_i} f(x_i) = 1$
3. $F_X(x_i) = P(X \leq x_i) = \sum_{x \leq x_i} f(x)$

Let x_i and x_{i+1} belongs to R_X such that $x_i < x_{i+1}$ and no point between x_i and x_{i+1} belongs to R_X

$$F_X(x_{i+1}) - F_X(x_i) = f(x_{i+1}) = P(X = x_{i+1})$$

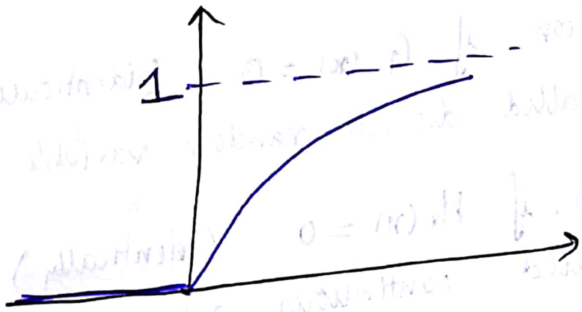
Continuous Random Variable.

means for all.

In Lebesgue decomp., $H_X(x) = 0$ (identically) then X is a continuous r.v.

$$P(X = x) = \lim_{\delta \to 0} \left[F_X(x+\delta) - F_X(x) \right] \quad ; \ (\delta > 0)$$

$= 0$ (since it is right continuous)
(there are no jumps).



can't take derivative here.

prob. mass function $\sim$ prob. density function.

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

$\uparrow$ always exists except for a few points.

(in the above example, $f_X(0)$ does not exist.)

$f_X(x)$: prob. density function.

28th Jan 2019

Example

Expt. $\rightarrow$ Throw dart on this circular board.

$(\Omega, \mathcal{A}, P)$

uniform prob. space.

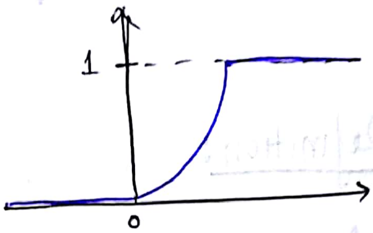
X: distance of dart from center 'o' of the board. [bull's eye]

$\Rightarrow$ Let $F_X(x)$ denote CDF of X.

$$F_X(x) = P(X \leq x) = \frac{\pi x^2}{\pi r^2} \quad \swarrow \text{uniform probability principle.}$$

$$= \frac{x^2}{r^2}$$

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ x^2/R^2 & , x \in [0, R] \\ 1 & , x > R. \end{cases}$$



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

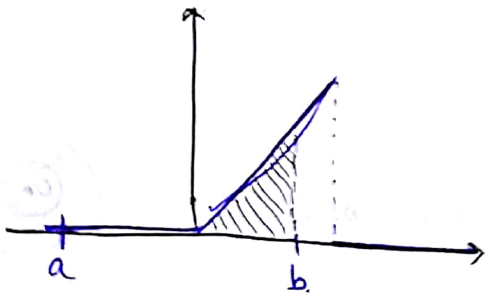
$$= \begin{cases} 0 & , x < 0 \\ 2x/R^2 & , 0 \leq x \leq R. \\ 0 & , x > R \end{cases}$$

$$= \begin{cases} \frac{2x}{R^2} & , 0 \leq x \leq R \\ 0 & , \text{otherwise.} \end{cases}$$

$$P(a \leq x \leq b) = F_X(b) - F_X(a)$$

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

$$= \int_{a}^{b} f_X(x)\,dx$$



→ Let X be a continuous random variable, with CDF $F_X(x)$ and $f_X(x) = \frac{d}{dx}(F_X(x))$ is its Probability density function (Pdf). Then the range of X is set of $R_X \leq R$ such that $\forall x \in R_X$, $f(x) > 0$

what is this?

<u>Definition:</u>

A density function or Pdf is a non-negative function such that

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

then obviously,

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

Satisfy all properties of CDF.

$$P(a \leq x \leq b) = \int_a^b f_x(x)\,dx = F_x(b) - F_x(a)$$

Example → For what value of 'k'

$$f(x) = \begin{cases} kx^3 & , \quad 0 \leq x \leq 1 \\ 0 & , \quad \text{otherwise} \end{cases}$$

Sol^n →

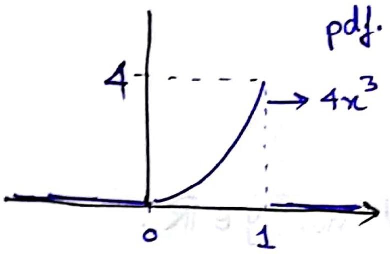
a) $k > 0$

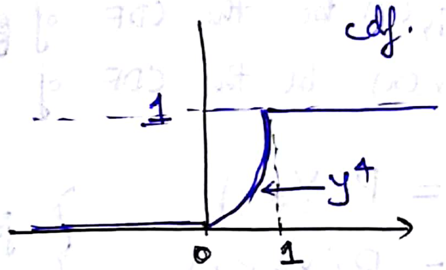
b) $\int_0^1 kx^3\,dx = 1 \Rightarrow k\frac{x^4}{4}\Big|_0^1 = 1$

$\Rightarrow \boxed{k = 4}$

c) for $y \in (0, 1)$

$$F_x(y) = \int_0^y 4x^3\,dx = y^4$$





Example →

$$f_x(x) = \begin{cases} kx(x+1) & ; \quad 0 \leq x \leq 1 \\ 0 & ; \quad \text{otherwise.} \end{cases}$$

a) $k > 0$

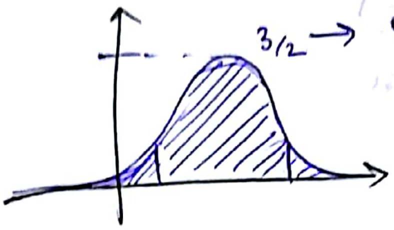
b) $\int_0^1 kx(x+1)\,dx = \int_0^1 kx^2 + kx\,dx = 1$

$$= \frac{k}{3} + \frac{k}{2} = 1$$

$\Rightarrow \frac{5}{6}k = 1$ or $\boxed{k = \frac{6}{5}}$

Example → $f_X(x) = \begin{cases} 6x(1-x) & ; 0 \le x \le 1 \\ 0 & ; \text{otherwise.} \end{cases}$

Sol^n → $\int_0^1 6x(1-x)\,dx = -\frac{6}{3} + \frac{6}{2} = -2+3 = 1$



3/2 → value can be > 1

pdf

CHANGE OF VARIABLE FORMULA

Let X be a continuous rand. var. with pdf. $f_X(x)$.
Find the density of rand var. $Y = X^2$

Let $G_Y(y)$ be the CDF of Y
$F_X(x)$ be the CDF of X

$$\left.\begin{aligned} G(y) &= P(Y \le y) \\ &= P(X^2 \le y) \end{aligned}\right\} \text{ for any real no. } y \in \mathbb{R}.$$

$$= P(-\sqrt{y} \le X \le \sqrt{y})$$

$$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

In order to find density of Y, say $g(y)$ we take derivative of $G_Y(y)$

$$g(y) = \frac{d}{dy} G_Y(y)$$

$$= \frac{d}{dy}\left[F_X(\sqrt{y}) - F_X(-\sqrt{y})\right]$$

$$= \frac{1}{2\sqrt{y}}\left[F_X'(\sqrt{y}) + F_X'(-\sqrt{y})\right]$$

$$= \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right] \quad ; y > 0$$

Example → $f_X(x) = \begin{cases} \frac{2x}{R^2} & ; 0 \le x \le R \\ 0 & ; \text{otherwise} \end{cases}$

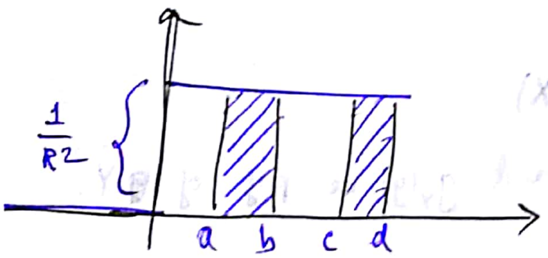
find density of $Y = X^2$

Sol? → let $g_Y(y)$ be the density of Y

$$g_Y(y) = \frac{1}{2\sqrt{y}} \left[f_X(\sqrt{y}) + f_X(-\sqrt{y}) \right]$$

$$= \frac{1}{2\sqrt{y}} \left[\frac{2\sqrt{y}}{R^2} \right] \quad 0 \le y \le R^2$$

$$\therefore g_Y(y) = \begin{cases} \frac{1}{R^2} & , 0 \le y \le R^2 \\ 0 & , \text{otherwise.} \end{cases}$$



Uniform prob. principle.

CONTINUOUS UNIFORM RANDOM VARIABLE.

Let X be a CRV with pdf defined as :-

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise.} \end{cases}$$

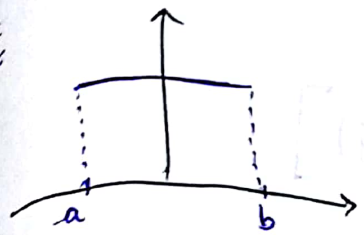
for any two $a, b \in \mathbb{R}$ such that $a < b$, then X is said to follow uniform density with parameters a & b.

$$X \sim U(a,b)$$

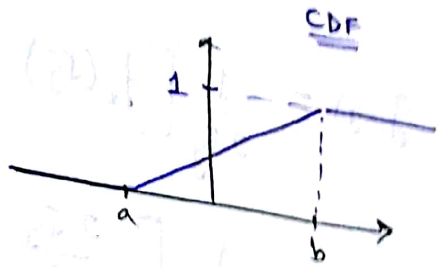
$$F_X(x) = \int_{-\infty}^{\infty} f_X(x)\,dx = \frac{x-a}{b-a} \qquad a \le x \le b$$

$$F_X(x) = \begin{cases} 0 & ; \quad x \le a \\ \frac{x-a}{b-a} & ; \quad a < x < b \\ 1 & ; \quad x \ge b \end{cases}$$

PDF



CDF



Example → Let $X \sim U(0,1)$

$$f_X(x) = \begin{cases} 1 & ; \quad 0 \le x \le 1 \\ 0 & , \quad \text{otherwise.} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & , \quad x \le 0 \\ x & ; \quad 0 < x < 1 \\ 1 & , \quad x \ge 1 \end{cases}$$

Consider the transformation

$$Y = -\frac{1}{\lambda} \log(1-X)$$

Let $G_Y(y)$ be the CDF of y and $g_Y(y)$ be pdf of Y.

$$G_Y(y) = P(Y \le y)$$

$$= P\left(-\frac{1}{\lambda}\log(1-X) \le y\right)$$

$$= P\left(\log(1-X) \ge -\lambda y\right)$$

$$= P\left(1-X \ge e^{-\lambda y}\right)$$

$$= P\left(X \le 1-e^{-\lambda y}\right)$$

$$\therefore \quad G_Y(y) = 1-e^{-\lambda y} \quad ; \quad y > 0$$

$$g_Y(y) = \frac{d}{dy} G_Y(y)$$

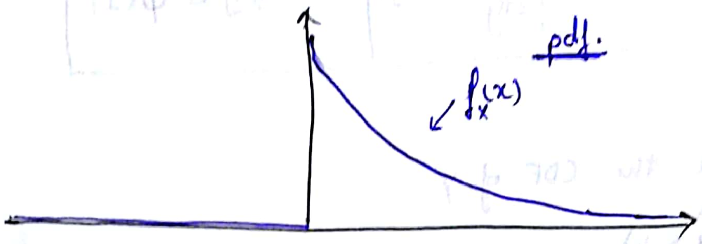
Scanned by CamScanner

$$g_Y(y) = \begin{cases} \lambda e^{-\lambda y} & , \ y > 0 \\ 0 & , \ y \leq 0. \end{cases}$$

Exponential Density with parameter λ $(\lambda > 0)$

Let $X \sim \exp(\lambda)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & , \ x > 0 \\ 0 & , \ \text{otherwise}. \end{cases}$$



CDF $\quad F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$

for any real number $x, y > 0$

$$\begin{aligned} P(X > x+y) &= 1 - F_X(x+y) \\ &= 1 - \left(1 - e^{-\lambda(x+y)}\right) \\ &= e^{-\lambda(x+y)} \\ &= e^{-\lambda x} \cdot e^{-\lambda y} \\ &= P(X > x) \cdot P(X > y) \end{aligned}$$

$$\frac{P(X > x+y)}{P(X > y)} = P(X > x) \qquad \text{memory less property.}$$

$$P(X > x+y \mid X > y) = P(X > x)$$

Theorem : Let ϕ be a differentiable function which is strictly increasing or strictly decreasing on an interval I_1

Let $\phi(I)$ denote the range of I.

ϕ^{-1} be the inverse of ϕ on I.

Let X be a continuous rand. var. having density $f_X(x)$ such that $f_X(x) \neq 0$ in I.

Then $y = \phi(x)$ whose density is given by

$$\boxed{g(y) = f(\phi^{-1}(y)) \left| \frac{d}{dy} \phi^{-1}(y) \right| \quad ; y \in \phi(I)}$$ V. Imp.!

Proof. let $G_Y(y)$ be the CDF of Y

$$G_Y(y) = P(Y \leq y)$$
$$= P(\phi(x) \leq y)$$
$$= P(X \leq \phi^{-1}(y))$$
$$= F_X(\phi^{-1}(y))$$

$$g_Y(y) = F_X'(\phi^{-1}(y)) \cdot \frac{d}{dy} \phi^{-1}(y)$$

note : $$\left. g_Y(y) = f_X(\phi^{-1}(y)) \cdot \left| \frac{d}{dy} \phi^{-1}(y) \right| \right\} \phi \text{ is } \uparrow$$

$$G_Y(y) = P(Y \leq y) = P(\phi(x) \leq y)$$
$$= P(X \geq \phi^{-1}(y))$$

$$G_Y(y) = 1 - f_X(\phi^{-1}(y))$$

note :

SYMMETRIC DENSITIES

$f(x)$ is a symmetric density if $f(x) = f(-x) \quad \forall (x) \in \mathbb{R}$.

eg. → $U(-a, a)$ is symmetric.

A rand. var. is symmetric if its pdf $f_X(x)$ is a symmetric density.

ex. → If X is symmetric r.v. with cdf $F_X(x)$ then P.T. $F_X(0) = \frac{1}{2}$

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

$$= 1 - F_X(x).$$

OR

$$F_X(-x) = 1 - F_X(x)$$

$$\Rightarrow \boxed{F_X(x) + F_X(-x) = 1} \quad \forall x \in \mathbb{R}.$$

Ex. $g(x) = \frac{1}{1+x^2}$

$-\infty < x < \infty$ Is $g(x)$ a pdf?

Sol? $$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \tan^{-1}(x) \Big|_{-\infty}^{\infty} = \pi$$

$\therefore$ it is not a pdf.

but $g_2(x) = \frac{1}{\pi(1+x^2)}$ is a pdf.

Cauchy Density

since $g(x) = g(-x)$, cauchy density is symmetric.

$$g(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty \quad \text{---> cauchy density.}$$

Example 2:

$$g(x) = e^{-x^2/2}$$

$$c = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx$$

$$c^2 = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy$$

$$c^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{(x^2+y^2)}{2}} dy.dx$$

polar coordinates substitution: $(r\cos\theta, r\sin\theta)$

$$c^2 = \int_{-\infty}^{\infty}\int_{0}^{2\pi} r e^{-r^2/2} dr\, d\theta.$$

$$= 2\pi$$

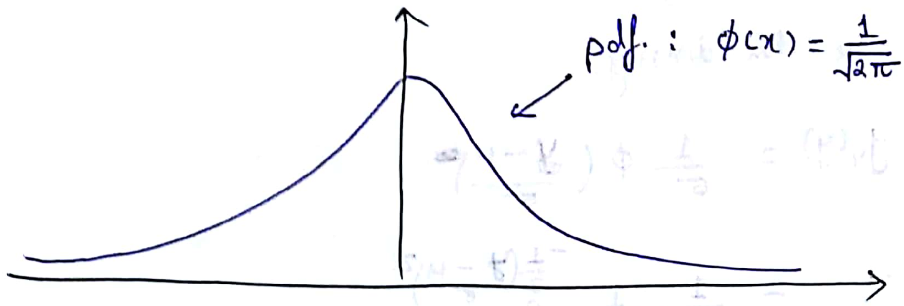
$$\therefore \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$$

Define.

$$\boxed{\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad -\infty < x < \infty}$$

Clearly this function ($\phi(x)$) is a pdf. This density is called standard normal density. The random variable associated with this density is called the standard normal random variable.

Standard normal density $\phi(x)$ is clearly symmetric.



pdf : $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$R_X = \mathbb{R}$ as $\phi(x) > 0 \quad \forall x \in \mathbb{R}$.

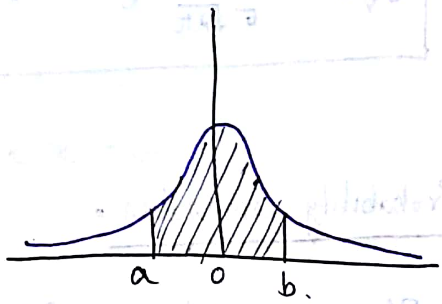
* $P(a \leq x \leq b) = F_X(b) - F_X(a)$

let $\Phi(x)$ denote the CDF of the standard normal random variable.

$$P(a \leq x \leq b) = \Phi_X(b) - \Phi_X(a)$$

$$= \int_a^b \phi(x)\,dx$$

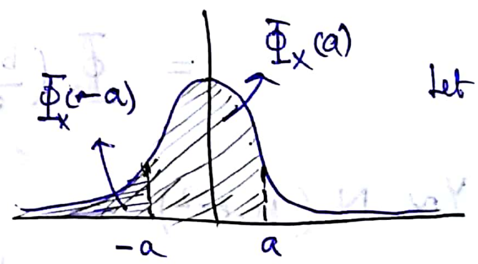
$$= \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$$



It seems that it is enough to know

$\Phi_X(a) \quad \forall a \in \mathbb{R}$

$$\Phi_X(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$$



Let $a > 0$

$\therefore \quad \Phi_X(-a) = 1 - \Phi_X(a)$ (Table)

Qu. $P(-3 < x < 3) = \Phi(3) - \Phi(-3)$

$$= 2\Phi(3) - 1$$

$$= 2(0.99865) - 1$$

$$= 0.9973.$$

Scanned by CamScanner

* Let X be a standard normal random variable. Define:

$$Y = \mu + \sigma X \qquad \text{where } \sigma > 0$$

Let $g_Y(y)$ be the density.

$$g_Y(y) = \frac{1}{\sigma}\,\phi\left(\frac{y-\mu}{\sigma}\right)$$

$$= \frac{1}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \qquad -\infty < y < \infty$$

this y is said to follow normal density within two parameters μ and σ^2

$$Y \sim N(\mu, \sigma^2)$$

$$\boxed{g_Y(y) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\frac{(y-\mu)^2}{\sigma^2}}} \qquad -\infty < y < \infty$$

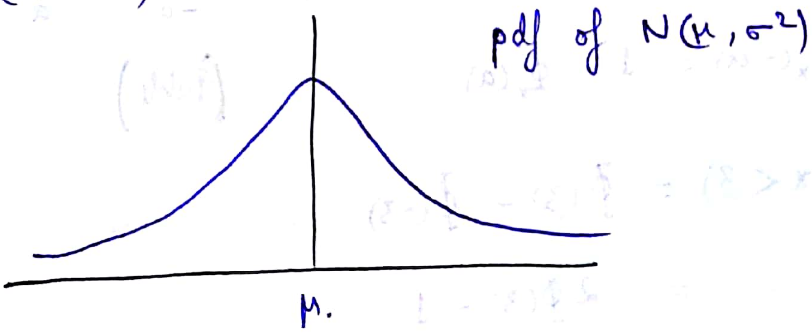
<u>Probability Computation:</u>

$$P(a < y < b) = P(a < \mu + \sigma x < b)$$

$$= P\left(\frac{a-\mu}{\sigma} < x < \frac{b-\mu}{\sigma}\right)$$

$$= \Phi_X\left(\frac{b-\mu}{\sigma}\right) - \Phi_X\left(\frac{a-\mu}{\sigma}\right)$$

$Y \sim N(\mu, \sigma^2)$



Scanned by CamScanner

Example. →

$P(\mu - 3\sigma < y < \mu + 3\sigma)$

$= P(-\frac{3\sigma}{\sigma} < x < \frac{3\sigma}{\sigma})$

$= P(-3 < x < 3)$

$= \Phi_x(3) - \Phi_x(-3)$

$= 0.9973.$

GAMMA DENSITIES

Let $X \sim N(0, \sigma^2)$

$Y = X^2$

find the density of Y

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} \qquad -\infty < x < \infty,$$

Let $g_Y(y)$ be density of Y

$$g_Y(y) = \frac{1}{2\sqrt{y}}\left(f(\sqrt{y}) + f(-\sqrt{y})\right) \qquad y > 0$$

$$g_Y(y) = \frac{1}{\sqrt{2\pi y}\cdot\sigma} e^{-y/2\sigma^2} \qquad y > 0$$

Note. $g_Y(y)$ is a gamma density.

* Gamma functions :-

$$\int_0^\infty x^{\alpha-1} \cdot e^{\lambda\alpha} \cdot dx = \frac{\Gamma(x)}{\lambda\alpha}$$

gamma pdf :-

$$\Gamma(x;\alpha;\lambda) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \cdot e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

parameters in this density are α, λ.

take $\alpha = \frac{1}{2}$, $\lambda = \frac{1}{2\sigma^2}$

EXPECTATIONS OF A RANDOM VARIABLE

Definition: Let X be discrete r.v. with R_X as its range.

$$E(X) = \sum_{x \in R_X} x \cdot P(X=x)$$ } provided that sum is convergent.

example -

let X be a discrete uniform r.v. on $\{x_1, x_2 \dots x_n\}$

$$P(X=x_i) = \begin{cases} \frac{1}{n} & ; X \in R_X \\ 0 & ; \text{otherwise} \end{cases}$$

$$E(X=x_i) = \sum x_i P(X=x_i)$$

$$= \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{(arithmetic mean)}$$

In case of uniform discrete r.v. E(x) is nothing but AM of R_X.

In general, we can understand E(x) is weighted average.

Scanned by CamScanner

1. $X \sim$ Bernoulli (p)

x	0	1
$p(x)$	$(1-p)$	p

pmf of X in tabular form

$$E(X) = 0(1-p) + 1(p) = p.$$

2. $X \sim$ Binomial (n, p)

$$P(X=x) = \begin{cases} {}^{n}C_{x}\, p^{x} (1-p)^{n-x} & ; x = 0, 1, 2, \ldots n. \\ 0 & ; \text{otherwise} \end{cases}$$

$$E(X) = \sum_{j=0}^{n} j \binom{n}{j} p^{j} (1-p)^{n-j}$$

Note: $\circ$ $j\binom{n}{j} = j \times {}^{n}C_{j} = \dfrac{j \times n!}{(n-j)!\, j!} = n \times {}^{n-1}C_{j-1}$

$$E(X) = np.$$

3. $X \sim$ poission(λ)

$$P(X=x) = \begin{cases} \dfrac{e^{-\lambda} \cdot \lambda^{x}}{x!} & x = 0, 1, 2 \ldots \\ 0 & \text{otherwise.} \end{cases}$$

$$E(X) = \sum_{j=0}^{\infty} j \frac{e^{-\lambda} \cdot \lambda^{j}}{j!} = \lambda$$

4. $X \sim$ geometric (p)

$$P(X=x) = \begin{cases} p(1-p)^{x} & , x = 0, 1, 2 \ldots \\ 0 & , \text{otherwise} \end{cases}$$

$$E(X) = \sum_{j=0}^{\infty} j \cdot p(1-p)^{j}$$

$$= p \sum_{j=0}^{\infty} j (1-p)^{j}$$

Scanned by CamScanner

$$E(X) = p(1-p)\sum_{j=0}^{\infty} j(1-p)^{j-1}$$

$$= -p(1-p)\sum_{j=0}^{\infty}\frac{d}{dp}(1-p)^{j-1}$$

Interchanging Σ and $\frac{d}{dp}$ because the series is absolutely summable.

$$E(X) = -p(1-p)\cdot\frac{d}{dp}\sum_{j=0}^{\infty}(1-p)^{j-1}$$

$$= \frac{(1-p)}{p}$$

<u>Example</u>. Let X be a discrete r.v. with pmf

$$f(x) = \begin{cases} \frac{1}{x(x+1)} & , x = 1, 2, --- \\ 0 & , \text{otherwise.} \end{cases}$$

clearly, $f(x) \geq 0 \quad \forall\ x \in \mathbb{R}$.

and $\sum_{x=1}^{\infty} f(x) = \sum_{x=1}^{\infty}\frac{1}{x(1+x)}$

$$= \sum_{x=1}^{\infty}\left(\frac{1}{x} - \frac{1}{x+1}\right) \quad \text{Telescoping series.}$$

$$= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + --- \infty$$

$$= 1$$

$\therefore\ E(X) = \sum_{x=1}^{\infty} x f(x) = \sum_{x=1}^{\infty}\frac{1}{1+x}$

DOES NOT EXIST.

$\because$ infinite series sum where sum does not converge.

<u>PROPERTIES OF EXPECTATION :-</u>

Let $Z = \phi(X)$ be a function of discrete random variable X.

$$E(Z) = E(\phi(X)) = \sum_{x}\phi(x)\cdot P(X=x)$$

provided that the expectation exists.

example. 'X' is a discrete r.v. with pmf →

X.	−2	−1	0	1	2
P(x)	1/5	1/5	1/5	1/5	1/5

$E(X)$, $E(X^2)$

Ans. $E(X) = \left(-2\times\frac{1}{5}\right) + \left(\frac{-1}{5}\right) + 0 + \frac{1}{5} + \frac{2}{5} = 0.$

$E(X^2) = \cancel{2x} \ \frac{1}{5}\times[4+1+0+1+4] = 2$

Let X and Y be discrete r.v. with finite expectation

1. for an $\alpha \in \mathbb{R}$ such that $P(X=\alpha) = 1$

$$E(X) = \alpha$$

2. for some const. $c \in \mathbb{R}$

$$E(cX) = cE(X)$$

3. $E(X+Y) = E(X) + E(Y)$

4. If $P(X \geq Y) = 1$ then $E(X) \geq E(Y)$

moreover, $E(X) = E(Y)$ if and only if $P(X=Y) = 1$.

5. $|E(X)| \leq E(|X|)$

Exercise

⟨ Hypergeometric ⟩

population (r) → Type I (r_1)
→ Type II $(r_2 = r - r_1)$

A sample of size n is drawn from the population.

S_n = number of objects of type 1 in this sample.

Define : X_i = i^{th} indicator random variable indicating whether i^{th} object in the sample is of Type I.

$$S_n = X_1 + X_2 + \cdots X_n$$

$$E(X_i) = \frac{r_1}{r}$$

$$\text{so } E(S_n) = n \cdot \left(\frac{r_1}{r}\right)$$

Moments :

$$E(X^r) = \sum_x x^r P(X=x) \quad \text{for } r = 1, 2 \ldots$$

these are called raw moments of X.
$r = 1$ is $E(X)$ is called the mean of X.

$$E(X-a)^r = \sum_x (x-a)^r P(X=x) \quad \text{for some fixed } a \in R$$

central moments of X.

for $a = 0$; central moments are same as raw moments.

Scanned by CamScanner

In particular, take $a = \mu = E(X)$ and for $r = 2$

$$E(X-\mu)^2 = E(X-E(X))^2 = \text{Variance of } X = \sigma^2$$

The positive square root of variance is called the standard deviation of X

Interpretations of $\sigma^2 = var(X)$

Measuring squared variability of X around $a = E(X-a)^2$

Interested in minimising $E(X-a)^2$ wrt. a.

(Approximating a r.v. by a constant 'a' and error in approximation is $E(X-a)^2$)

$$E(X-a)^2 = E(X^2 - 2aX + a^2)$$

$$= E(X^2) - 2aE(X) + a^2.$$

diff. wrt $a. \rightarrow$ $\quad -2E(X) + 2a = 0$

$$\Rightarrow \quad a = E(X) = \mu.$$

Another interpretation of variance :—

Given a random variable X and a number 'a' $\in \mathbb{R}$

$$(X-a)^2 = [(X-\mu) + (\mu-a)]^2 \qquad \text{where } \mu = E(X)$$

$$= (X-\mu)^2 + 2(X-\mu)(\mu-a) + (\mu-a)^2$$

$$E(X-a)^2 = E(X-\mu)^2 + E(\mu-a)^2 + 2E(\mu-a)\cancel{(X-\mu)}^{0}$$

$$\boxed{E(X-a)^2 = var(X) + (\mu-a)^2}$$

Chebyshev's Inequality →

Let X be a non-negative rand. var. with finite expectation.
$t > 0$ (any +ve real no.)

Define a new rand. var. —

$$Y = 0 \quad \text{if} \quad X < t$$
$$Y = t \quad \text{if} \quad X \geq t$$

$$E(Y) = 0 \times P(Y=0) + t \times P(Y=t)$$

$$E(Y) = t\, P(X \geq t)$$

note $X \geq Y$

$$E(X) \geq E(Y) = t\, P(X \geq t)$$

$$\Rightarrow \quad \boxed{P(X \geq t) \leq \frac{E(X)}{t}}$$

Chebyshev's Inequality

X be a rand. var. with mean μ and variance σ^2
for any real no. $t > 0$,

$$\boxed{P(|X-\mu| > t) \leq \frac{\sigma^2}{t^2}}$$

Proof: consider $(x-\mu)^2 \geq 0$,

$$P((X-\mu)^2 \geq t^2) \leq \frac{E(X-\mu)^2}{t^2}$$

Moment Generating function :-

$M_X(t) = E(e^{tx})$ for a given random variable X

In case of a discrete r.v. X with pmf

$$P(X = x_i) = f(x_i) \quad \text{for } (x_i) \in \mathbb{R}$$

$$M_X(t) = E(e^{tx})$$

$$= \sum_{x_i} e^{tx_i} \cdot P(X = x_i)$$

Ex. Let $X \sim$ Bernoulli (p)

$$M_X(t) = e^{t.0}(1-p) + e^{t.1}(p)$$

$$= e^t p + (1-p) = q + p.e^t$$

Ex Let $X \sim$ Binomial (n, p)

$$M_X(t) = \sum_{x=0}^{n} e^{tx} \, {}^nC_x \, p^x \cdot (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} {}^nC_x \, (e^t p)^x (1-p)^{n-x}$$

$$= (q + pe^t)^n = (q + pe^t)^n$$

Ex. Let $X \sim$ Poisson (λ)

$$M_X(t) = \sum_{x=0}^{n} e^{tx} \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{n} \frac{(e^t \lambda)^x}{x!}$$

$$= e^{-\lambda} e^{\lambda e^t}$$

Consider. (i) $X \sim$ Bernoulli (p)

$$M_X(t) = (1-p) + pe^t$$

$$\frac{d}{dt}(M_X(t)) = pe^t \quad \text{if } t=0 \text{ ; } = p.$$

(ii) $X \sim$ Binomial (n, p)

$$M_X(t) = ((1-p)+pe^t)^n$$

$$\frac{d}{dt}[M_X(t)]\Big|_{t=0} = np.$$

$$E(e^{tX}) = E\left(1 + tx + \frac{t^2x^2}{2!} + \cdots\right)$$

$$= E(1) + E(tx) + E\left(\frac{t^2x^2}{2!}\right) + \cdots$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \cdots$$

$$\frac{d}{dt}[M_X(t)] = E(X) + \frac{2t}{2!}E(X^2) + \cdots$$

$$\frac{d}{dt}[M_X(t)]_{t=0} = E(X)$$

$$\frac{d^2}{dt^2}(M_X(t))\Big|_{t=0} = E(X^2)$$

* $M_X(t=0)$, $M_X(t)$ is always defined for $t=0$

* $M_X(t)$ as a function of t is defined only around $t=0$

* Let X be r.v. with MGF $M_X(t)$
Then find MGF of r.v. $(aX+b)$ for some $a, b \in \mathbb{R}$, $a \neq 0$

$$M_{aX+b}(t) = E(e^{(aX+b)t})$$

$$= E(e^{aXt} \cdot e^{bt})$$

$$= e^{bt} E(e^{aXt})$$

$$= e^{bt} \cdot M_X(at)$$

$$\left. \frac{d}{dt}(M_{aX+b}(t)) \right|_{t=0} = ? \quad = b e^{bt} M_X(at) + a e^{bt} M_X'(at)$$

$$\cancel{(a+b)}$$

$$= a^2 E(X) + ba\, E(X)$$

Let X be a r.v. with MGF $M_X(t)$. Find the MGF of the r.v. $aX+b$ for some $a,b \in \mathbb{R}$, $a \neq 0$

$$M_{aX+b}(t) =$$

$$= e^{bt} M_X(at)$$

Expectation of a continuous r.v.

Let Expectation of a continuous r.v. be X

Ex. $X \sim U(a,b)$

$$E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_a^b x \cdot \frac{1}{b-a}\, dx$$

$X \sim (U)(a,b)$ & $f_X(x) = \begin{cases} \frac{1}{b-a} & , a \le x \le b) \\ 0 & , \text{otherwise} \end{cases}$

$\Rightarrow \boxed{E(X) = \frac{a+b}{2}}$ mean of a uniformly distributed r.v.

Ex. $X \sim \text{Gamma}(\alpha, \lambda)$

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \cdot e^{-\lambda x} & ; x > 0 \\ \text{otherwise} & 0 \end{cases}$$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^{\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot x^{\alpha-1} e^{-\lambda x}\, dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^\alpha e^{-\lambda x}\, dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}}$$

$$\boxed{E(X) = \frac{\alpha}{\lambda}}$$

* remember properties of Γ function.

Scanned by CamScanner

Ex. $X \sim \text{Gamma}(1, \lambda)$; $\lambda > 0$

$$f_X(x) = \begin{cases} \frac{\lambda}{\Gamma(1)} \cdot x^{1-1} \cdot e^{-\lambda x} & ; x > 0 \\ 0 & ; \text{otherwise}; \end{cases}$$

$$= \begin{cases} \lambda e^{-\lambda x} & ; x > 0 \\ 0 & ; \text{otherwise}. \end{cases}$$

$\Rightarrow$ exponential density

$\text{Gamma}(1, \lambda) = \exp(\lambda)$; $\lambda > 0$

$\boxed{E(X) = \frac{1}{\lambda}}$ Mean of exponential distribution.

Ex. Let $f_X(x)$ denote the density of Cauchy r.v.

$$f_X(x) = \frac{1}{\pi(x^2+1)} \qquad -\infty < x < \infty$$

$$E(|x|) = \int_{-\infty}^{\infty} \frac{|x|\,dx}{\pi(x^2+1)} = \int_0^{\infty} \frac{2x\,dx}{\pi(1+x^2)}$$

$$= \frac{1}{\pi} \lim_{k\to\infty} \int_0^k \frac{2x}{1+x^2}\,dx$$

$$= \frac{1}{\pi} \lim_{k\to\infty} \log(1+k^2)$$

$\therefore$ $E(|x|) \longrightarrow \infty$

$\Rightarrow$ $E(x)$ does not exist!

Cauchy r.v. has no finite mean.

Moments of a continuous r.v.

1. $E(X^m) = \int_{-\infty}^{\infty} x^m f_X(x)\, dx$: m^{th} raw moment

2. $E(X-a)^m = \int_{-\infty}^{\infty} (x-a)^m f_X(x)\, dx$: m^{th} central moment.

$m = 1, 2, \cdots$

Variance of $X = \sigma^2 = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx$

where $\mu = E(X)$

Ex $X \sim \text{gamma}(\alpha, \lambda)$

$$E(X^m) = \int_0^{\infty} x^m \cdot \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot x^{\alpha-1} \cdot e^{-\lambda x}\, dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{m+\alpha-1} e^{-\lambda x}\, dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+m)}{\lambda^{\alpha+m}}$$

$$\boxed{E(X^m) = \frac{\alpha(\alpha+1)(\alpha+2)\cdots(\alpha+m-1)}{\lambda^m}}$$

Variance of $X = \sigma^2 = E(X^2) - (E(X))^2$

$$= \frac{(\alpha+1)\alpha}{\lambda^2} - \frac{\alpha^2}{\lambda^2}$$

$$\boxed{V(X) = \frac{\alpha}{\lambda^2}} = \sigma^2$$

Ex. If $X \sim \exp(\lambda)$

$$\boxed{E(X^m) = \frac{m!}{\lambda^m}}$$

$\because \exp(\lambda) = \text{Gamma}(1, \lambda)$

$$V(X) = \frac{1}{\lambda^2}$$

Ex. $X \sim U(a,b)$

$$E(X^2) = \int_a^b x^2 \frac{1}{b-a}\,dx = \frac{b^3-a^3}{3} \cdot \frac{1}{b-a}.$$

$$Var(X) = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{a^2-2ab-b^2}{12}$$

Advantage of symmetry of a r.v. X

Let X be a symmetric r.v.

$\Rightarrow$ X and $-X$ have same density.

For any integer $m \in \mathbb{N}$

X^m and $(-X)^m$ will also have same density.

for m : odd $\rightarrow$ (X^m) and $(-X^m)$ have same density.

for m : even $\rightarrow$ X^m is the density.

$\Rightarrow E(X^m) = E(-X^m) = -E(X^m)$ for m : odd.

$\Rightarrow \boxed{E(X^m) = 0}$

in particular, mean of all symmetric densities is zero.

eg. $N(0, \sigma^2)$

Moment Generating Function for continuous r.v.

$$M_X(t) = E(e^{tX}) \quad \text{if it exists.}$$

Ex. $X \sim N(\mu, \sigma^2)$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad ; -\infty < x < \infty$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-1}{2\sigma^2}\cdot (x-\mu)^2\right\} \quad \text{another notation.}$$

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} dx$$

put $y = x - \mu$ ---> transform to make this a symmetric distribution about $y=0$

$$M_X(t) = \int_{-\infty}^{\infty} e^{(y+\mu)t} \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2\sigma^2}y^2} dy$$

$$= e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\left(ty - \frac{y^2}{2\sigma^2}\right)} dy$$

[Notice: $ty - \frac{y^2}{2\sigma^2} = -\frac{(y-\sigma^2 t)^2}{2\sigma^2} + \frac{\sigma^2 t^2}{2}$] Very imp. trick in stats. Completing the square.

$$M_X(t) = e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\sigma^2 t)^2}{2\sigma^2}} \cdot e^{\frac{\sigma^2 t^2}{2}} dy$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2}} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\sigma^2 t)^2}{2\sigma^2}} dy}_{1}$$ ---> density of a normal r.v.

$\because$ pdf of $N(\sigma^2 t, \sigma^2)$

$$\boxed{M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}}$$

$$\therefore E(X) = \frac{d}{dt} M_X(t)\Big|_{t=0} = \left(\mu + \frac{2}{2}\sigma^2 t\right) e^{\mu t + \frac{\sigma^2 t^2}{2}}\Big|_{t=0} = \mu$$

$$\therefore \boxed{E(X) = \mu}$$

$$E(X^2) = \frac{d^2}{dt^2} M_t(X) = \frac{d}{dt}\left(\mu e^{\mu t + \frac{1}{2}\sigma^2 t^2}\right) + \frac{d}{dt}\left(\sigma^2 t\, e^{\mu t + \frac{1}{2}\sigma^2 t^2}\right)\Bigg|_{t=0}$$

$$\boxed{E(X^2) = \sigma^2 + \mu^2}$$

$$\therefore \quad V(X) = (\sigma^2 + \mu^2) - (\mu)^2$$

$$\boxed{V(X) = \sigma^2}$$

Ex. Let ~~Xkm~~ $Y \sim N(0,1)$

$$M_Y(t) = e^{\frac{t^2}{2}}$$

Define $X = \mu + \sigma Y \qquad \sigma > 0$

$$M_X(t) = M_{\mu+\sigma Y}(t)$$

$$= E\left(e^{(\mu+\sigma Y)t}\right)$$

$$= e^{\mu t} M_Y(\sigma t)$$

$$= e^{\mu t} \cdot e^{\frac{\sigma^2 t^2}{2}}$$

$$= e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

mean of r.v.   variance of r.v.

[coeff. of t = mean.
coeff. of $\frac{t^2}{2}$ = variance.]

$\Rightarrow X \sim N(\mu, \sigma^2)$

In general, if $X \sim N(\text{mean}, \text{variance})$

$$M_X(t) = e^{(\text{mean})t + (\text{variance})\frac{t^2}{2}}$$

Very Very Important.

Ex. $X \sim N(\mu, \sigma^2)$

Define: $Y = \frac{(X-\mu)}{\sigma} = \underbrace{\frac{-\mu}{\sigma}}_{a} + \underbrace{\frac{1}{\sigma}}_{b}X$

$M_{aX+b}(t) = M_Y(t) = e^{t^2/2}$ follows standard normal.

Ex. $X \sim gamma\ (\alpha, \lambda)$

$$M_X(t) = \int_0^\infty e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot x^{\alpha-1} \cdot e^{-\lambda x} \cdot dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-(\lambda-t)x} \cdot x^{\alpha-1} dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha)}{(\lambda-t)^\alpha}$$

$$\Rightarrow M_X(t) = \left(\frac{\lambda}{\lambda-t}\right)^\alpha \quad ; \quad -\infty < t < \lambda$$

* MGF of exp. $(\lambda) = \left(\frac{\lambda}{\lambda-t}\right)$ $\because \alpha = 1$; $-\infty < t < \lambda$

QUARTILE:

A point Q_1 in $\mathbb{R}$ is 1^{st} Quartile when $P(X \leq Q_1) = \left(\frac{1}{4}\right)$

similarly $q_2 \in \mathbb{R}$ is 2^{nd} Quartile/Median when $P(X \leq q_2) = \frac{1}{2}$

similarly $q_3 \in \mathbb{R}$ is 3^{rd} Quartile when $P(X \leq q_3) = \frac{3}{4}$

Mode → point which is highest probable (discrete)

X a r.v. with density $f_X(x)$

$M = \arg\max_x f_X(x)$

Ex. Let $X \sim \text{Binomial}\left(3, \frac{1}{2}\right)$

Compute median & mode.

Sol^n: mean $= \frac{3}{2}$

$$f_X(x) = \begin{cases} {}^3C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{3-x} & ; x = 0,1,2,3 \\ 0 & ; \text{otherwise} \end{cases}$$

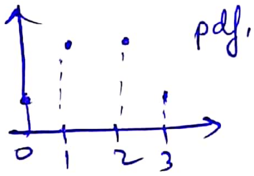
x	0	1	2	3
$f_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

median $= 1$

mode $= 1$ and 2.

$(1,2)$

the highly likely X.



Ex. $X \sim N(\mu, \sigma^2)$

mean $= \mu$

median $= \mu$.

$P(X \le m) = 0.5$.

$\Rightarrow P\left(\frac{X-\mu}{\sigma} \le \frac{m-\mu}{\sigma}\right) = 0.5$

$\Rightarrow \Phi\left(\frac{m-\mu}{\sigma}\right) = 0.5$

$\Rightarrow \frac{m-\mu}{\sigma} = 0 \Rightarrow \boxed{m = \mu}$

mode $(x) = \mu$.

$\boxed{\text{mean} = \text{mode} = \text{median} = \mu}$ for normal density.

Ex. $X \sim U(a,b)$

$\Rightarrow$ mean $= \frac{a+b}{2}$.

$F(m) = 0.5 \quad \Rightarrow \quad \frac{m-a}{b-a} = \frac{1}{2} \quad \Rightarrow \quad \boxed{m = \frac{a+b}{2}}$ median.

Ex. $X \sim \exp(\lambda)$

$F(m) = \frac{1}{2} \quad \Rightarrow \quad 1 - e^{-\lambda m} = \frac{1}{2}$

$\Rightarrow \quad e^{-\lambda m} = \frac{1}{2}$

$\Rightarrow \quad \boxed{m = +\frac{1}{\lambda}(\log 2)}$ median.

$E(X) = \frac{1}{\lambda}$ mean

Ex. A point is chosen randomly between the interval $[-10, 10]$ (by uniform probability principle). Let X be a r.v. defined in such a way that X denotes the coordinate of the chosen point if the point belongs to $[-5, 5]$.

X takes values -5 if the point $\in [-10, -5]$ and X takes value 5 if point belongs to $[5, 10]$. Compute CDF of X.

Ex. X is a cts. r.v. with pdf

$$f_X(x) = \frac{1}{2} e^{-|x|} \qquad -\infty < x < \infty$$

(2 sided exponential).

compute $P(1 \leq |x| \leq 2)$

RANDOM VECTORS

$(\Omega, \mathcal{A}, P)$

Let $X_1, X_2, \ldots, X_r$ be 'r' discrete random variables on $(\Omega, \mathcal{A}, P)$

Define: $X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{bmatrix}$ is an 'r' dimensional vector.

For an $w \in \Omega$

$$X(w) = \begin{bmatrix} X_1(w) \\ X_2(w) \\ \vdots \\ X_r(w) \end{bmatrix} \in \mathbb{R}^r.$$

Suppose, $X_1(w) = x_1, \quad X_2(w) = x_2 \quad \ldots \quad X_r(w) = x_r$

then $X(w) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix} \in \mathbb{R}^r$

Definition Let $X = \begin{bmatrix} X_1 \\ \vdots \\ X_r \end{bmatrix}$ be a vector where X_i is a random variable on $(\Omega, \mathcal{A}, P)$. For every $\underline{x} \in \mathbb{R}^r$

vector (understand from context)

The set $\{w \in \Omega \mid \underline{X}(w) = \underline{x}\} \in \mathcal{A}$

Then $X : \Omega \longrightarrow \mathbb{R}^r$ is called as an r-dimensional random vector.

Now we are interested in $P(\underline{X} = w) = \underline{x}$

Let $\underline{X} = \begin{bmatrix} X_1(w) \\ \vdots \\ X_r(w) \end{bmatrix}$ be a discrete random ~~variable~~ vector

If $\underline{x}$ denotes the value assumed by the random vector $\underline{X}$

then $\{\underline{x} : (P(\underline{X} = w) = \underline{x}) > 0\}$ is finite or countably infinite.

Definition → The discrete density or discrete joint pmf. of the random vector $\underline{X}$ is defined as:

$$f(x_1, x_2, \dots x_r) = \text{Prob}(X_1 = x_1, X_2 = x_2 \dots, X_r = x_r)$$

where $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{bmatrix}$ and $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix}$

In vector notation:

$$f(\underline{x}) = P(\underline{X} = \underline{x}) \quad \forall \ \underline{x} \in R^r$$

For a subset $A \subseteq R^r$

$$P(\underline{X} \in A) = \sum_{\underline{x} \in A} f(\underline{x})$$

Definition → The function 'f' is called as discrete joint pmf. if

1.) $f(\underline{x}) \geq 0 \quad \forall \ \underline{x} \in R^r$

2.) $\{\underline{x} : f(\underline{x}) \neq 0\}$ is finite or countably infinite. We denote the elements of this set as $\underline{x}_1, \underline{x}_2, \underline{x}_3 \dots$

3.) $\sum_i f(\underline{x}_i) = 1$

example →

$X_1 \backslash X_2$	1	2	3	4
1	1/4	1/8	1/16	1/16
2	1/16	1/16	1/4	1/8

→ this table is a discrete joint pmf

→ sum of all probs is 1.

$\underline{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$

$$P(\underline{X} = \underline{x}) = P(X_1 = x_1 ; X_2 = x_2)$$
$$= f(x_1, x_2)$$
$$= f(\underline{x})$$

Scanned by CamScanner

$$P(X_1 \geq X_2) = P(X_1=1, X_2=1) + P(X_1=2, X_2=1) + P(X_1=2, X_2=2)$$

$$= \frac{1}{4} + \frac{1}{16} + \frac{1}{16} = \frac{3}{8}$$

$$P(X_1=1) = P(X_1=1, X_2=1,2,3,4)$$

$$= \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16} = \frac{1}{2}.$$

$$\text{Prob}(X_1=1) = \sum_{x_2} f(1, x_2) = \frac{1}{2}$$

marginal pdf.

$$P(X_1=2) = \sum_{x_2} f(2, x_2) = \frac{1}{2}$$

Marginal pmf. of X_1

$$\boxed{f_{X_1}(x_1) = \sum_{x_2} f(x_1, x_2) \quad \text{for } x_1 \in R_{X_1}}$$

Independent r.v.s

Let X_1 X_2 ---- X_r be r discret random variables with pmfs f_1 f_2 ---- f_r respectively. The rand. var. $X_1, X_2, ---X_r$ are called mutually independent if their joint pmf f is given by:

$$f(x_1, x_2 --- x_r) = f_1(x_1) \cdot f_2(x_2) --- f_r(x_r)$$

NOTATION

$$f(x_1, x_2, --- x_r) = \text{Prob}(X_1 = x_1, X_2 = x_2 --- X_r = x_r)$$

Define $A_1 \subseteq \Omega$ such that $A_1 = [X_1 = x_1]$

$$A_1 = \{w : X_1(w) = x_1\}$$

$$\text{Prob}(A_1 \cap A_2 \cap --- \cap A_r)$$

$$= \text{Prob}(A_1) \cdot \text{Prob}(A_2) --- \text{Prob}(A_r)$$

$$= \text{Prob}(X_1 = x_1) \cdot \text{Prob}(X_2 = x_2) --- \text{Prob}(X_r = x_r)$$

$$= f_1(x_1) \cdot f_2(x_2) ---- f_r(x_r)$$

for our previous example ;

$X_1 \backslash X_2$	1	2	3	4	
1	$1/4$	$1/8$	$1/16$	$1/16$	$1/2$
2	$1/16$	$1/16$	$1/4$	$1/8$	$1/2$
	$\frac{5}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{3}{16}$	

$\therefore$ these two random variables are not independent.

$$f(1,1) \neq f_1(1) \cdot f_2(1)$$
$$\left(\frac{1}{4}\right) \neq \left(\frac{1}{2}\right) \times \left(\frac{5}{16}\right)$$

example →

X_1: 1^{st} toss; X_2: 2^{nd} toss.

$X_1 \backslash X_2$	0	1	
0	$1/4$	$1/4$	$1/2$
1	$1/4$	$1/4$	$1/2$.
	$1/2$	$1/2$	

one coin tossed 2 times.

These r.v.s are indep. of each other.

Ex Let X_1 and X_2 be 2 independent r.v.s. each with geometric distribution, parameter p.
Find the distribution of min (X_1, X_2)

Sol^n. Prob $(\min(X_1, X_2) \geq z)$

$= \text{Prob}(X_1 \geq z \,;\, X_2 \geq z)$ because of independence of X_1 and X_2

$= \text{Prob}(X_1 \geq z) \cdot P(X_2 \geq z)$

$= (1-p)^z (1-p)^z$

$= (1-p)^{2z}$