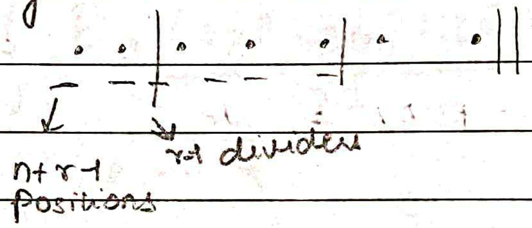
1) n balls & r boxes → all distinguishable
How many ways can we put n balls in r boxes
→ r^n → Maxwell-Boltzmann statistics

2) n balls and they are numbered.
How many ways can we choose r balls out of these n balls at a time/without replacement
→ ${}^nC_r$ or $\binom{n}{r}$ → Fermi-Dirac Statistics

3) n identical balls in r numbered boxes. How many possible arrangengements are there?
→ $\binom{n+r-1}{r-1}$



→ Bose-Einstein Statistics

4) n identical balls are placed in r boxes such that no box remains empty. How many possible arrangements?
→ $\binom{n-1}{r-1}$ → 1 ball in r boxes, $n-r$ balls left
→ Divide $n-r$ balls like 3, $\binom{n-r+r-1}{r-1} = \binom{n-1}{r-1}$

5) $a_1 + a_2 + \dots a_r = n$, $r, n \in N$
$a_i \in N \cup \{0\}$ $r < n$
P(solution of this Equation will be from positive integers)
→ $\dfrac{\binom{n-1}{r-1}}{\binom{n+r-1}{r-1}}$ → positive integral solution
→ non-negative integral solution

DOMS Page No.
Date / /

Random experiment :- Experiment where
1) more than one outcomes
2) we don't know the outcome of a particular trial
3) we can re-do the experiment in identical condition countably many times (practically not possible → assumed)

Sample space :- collection of all possible outcomes of a random experiment (Ω or S)

Probability :- If the outcome (of sample space) of a random experiment has finite cardinality
i.e $|\Omega|$ is finite,
Then the the probability of a subset of Ω is defined as

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

Q What is the probability that a randomly chosen natural number will be an even number?

→ $1/2$

Frequency definition of probability

If (Ω) is sample space and $A \subseteq \Omega$ then the $P(A)$ is defined as

$P(A) = \lim_{n\to\infty} \frac{\lvert A_n\rvert}{\lvert\Omega_n\rvert}$	$\lim_{n\to\infty} A_n = A$
	$\lim_{n\to\infty} \Omega_n = \Omega$
if the limit exists	$A_i \subseteq \Omega_i$ always.

Eg:- random natural number even/odd?

→ n_0, ~~leave~~ Fix n_0, even numbers less than or equal to $n_0 = \left[\frac{n_0}{2}\right]$

$[x]$ denotes the largest integer value lesser or equal to x

$$\lim_{n_0 \to \infty} \frac{[n_0/2]}{n_0} = \frac{1}{2}$$

Eg:- What is the probability that a random chosen number from $\mathbb{N}$ will be a ~~eve~~ a k-digit number?

→ $$\lim_{n_0 \to \infty} \frac{10^k}{n_0} = 0$$

But through total probability we get 1.

Algebra

A collection '$\mathcal{A}$' of the subsets of Ω is called algebra if

i) $\Omega \in \mathcal{A}$

ii) any $A \subseteq \Omega$ & $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$

[closed under complementation]

iii) any $A, B \subseteq \Omega$ & $A, B \in \mathcal{A}$ implies $A \cup B \in \mathcal{A}$

[closed under finite union]

σ-algebra / σ-field

An algebra '$\mathcal{A}$' of the subsets of Ω is called an σ-algebra/field if $\{A_i\} \subseteq \Omega$ and $\{A_i\}^I \in \mathcal{A}$ implies $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ [closed countable union]

DOMS Page No.
Date / /

Example of σ- algebra

Ω given

1) 2^{Ω} = power set of Ω = collection of all subsets of $\Omega = \mathcal{A}$
2) $\{\phi, \Omega\} = \mathcal{A}_1$
3) $\{\{A, A^c, \phi, \Omega\}\} = \mathcal{A}$

Q. $A, B \subset \Omega$
construct the σ- algebra with the subsets associated with it.

→ ~~$\{A, B, A\cup B, A^c, B^c, \phi, \Omega\}$~~ $\{\Omega, \phi, A, B, A^c, B^c, A\cup B, A^c\cap B^c, A^c\cap B, A\cap B^c, A\cup B^c, A^c\cup B, A^c\cup B^c, A\cap B\}$

Axiomatic Definition of Probability (Kolmogorov)

If $\mathcal{A}$ is σ-algebra of the subsets of ~~the~~ a non-empty set Ω, then probability (P) is defined to be a function $P: \mathcal{A} \to [0,1]$ which satisfies

1) $P(\Omega) = 1$
2) $P(A) \geq 0$ for any $A \in \mathcal{A}$
3) $\{A_i\} \in A$ implies $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ if $A_i \cap A_j = \phi$ $\forall i \neq j$

Q) What is the probability that a randomly chosen number from $\Omega = [0,1]$ (a) will be rational? (b) will be less than 0.4?

→ a) P(obtaining a rational no.) $= P(\cup_{k=1}^{\infty}$ k^{th} rational no. is chosen$)$
$= \sum_{k=1}^{\infty} P($ " " " " "$)$
$= \sum_{k=1}^{\infty} 0 = 0$

(rational numbers ~~are~~ can be listed)

b) 0.4 | $0 \leq a \leq b \leq 1$
$P([a,b]) = P((a,b]) = P([a,b)) = P((a,b)) = b-a$ in this case

c) irrational no. is chosen? $\longrightarrow P(\_\_) = 1$

Note:

$\{\Omega, A, A^c, \phi\} = \mathcal{A}$

$P(\Omega) = 1$, $P(A) = 0.3$, $P(C) = 0.7$, $P(\phi) = 0$

$P(B) = ?$ $B \notin \mathcal{A}$

$B \subseteq \Omega$

we cannot answer it as $B \notin$ the sigma algebra

Eg: A coin is tossed until we get the first head 'H'.

P(we get first head 'H' in an even number of tries)

$\rightarrow$

H

TH

TTH

TTTH

TTTTH

$= P($ First head is coming in $2k^{th}$ trial$)$ $k \in \mathbb{N}$

$= P\left(\bigcup_{k=1}^{\infty} \text{First head is coming in } 2k^{th} \text{ trial}\right)$

$= \sum_{k=1}^{\infty} P($ " " " " " " $)$

Note:

Probability $= 0$ doesn't mean event never occurs

" $= 1$ " " " always "

Probability Space

$(\Omega, \mathcal{A}, P)$ ~~is known~~

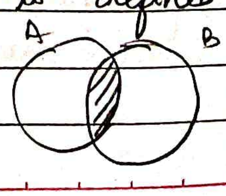
Event

For a given Probability space $(\Omega, \mathcal{A}, P)$ if $A \subseteq \Omega$ and $A \in \mathcal{A}$ then A is called an event

Conditional ~~Events~~ Probability :- Given $(\Omega, \mathcal{A}, P)$

If A & B are two events such that $P(B) > 0$ then conditional probability of A given B is defined as

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



Independent Event :- Given $(\Omega, \mathcal{A}, P)$

$P(A|B) = P(A) \implies$ A & B are independent events
or
$P(A \cap B) = P(A)\,P(B)$

Pairwise Independence

$P(A_i \cap A_j) = P(A_i)\,P(A_j)$, $\forall\, i \neq j$ $\Rightarrow$ all sets are Pairwise Independent.

Mutual Independence

$P(\cap \Sigma_{i=1}^{k} A_i) = \pi_{i=1}^{k} P(A_i)$ $\forall\, i$ are different,
↳ $\forall\, A_i$ are mutually independent.

Q. Pairwise independendence $\neq$ mutual independence.
Show. ?

Mutually exclusive events

$A_i \cap A_j = \phi \quad \forall\, i \neq j$

→ do not depend on Probability

Mutually exhaustive event

$\cup_{i=1}^{\infty} A_i = \Omega$

Partition

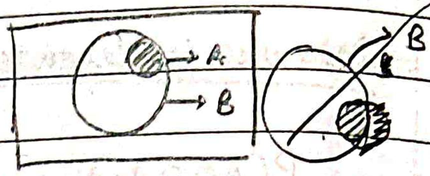
$\{A_i\}$ when the sets are mutually exclusive & exhaustive

Results

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

ii) $P(\phi) = 0$

iii) If $A \subseteq B$ then $P(A) \leq P(B)$



Proof :-

$$P(B) = P(A \cup (A^c \cap B))$$
$$= P(A) + P(A^c \cap B)$$
$$\geq P(A) \quad \text{as } P(A^c \cap B) \geq 0$$

iv) $1 - P(\cup_{i=1}^{\infty} A_i) = P(\cap_{i=1}^{\infty} A_i^c)$

Proof :-

$$1 - P(\cup_{i=1}^{\infty} A_i) = P\left[(\cup_{i=1}^{\infty} A_i)^c\right]$$
$$= P(\cap_{i=1}^{\infty} A_i^c)$$

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

vi) $P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$

Proof :-

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$\leq P(A) + P(B) \quad (\text{as } P(A \cap B) \geq 0)$$

by induction we can prove this result

vii) $P(\cup_{i=1}^{n} A_i) = \sum_{k=1}^{n} (-1)^{k-1} S_k$ where

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

Proof :-

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) \rightarrow 3 \text{ terms}$$
$$- P(A \cap B) - P(B \cap C) - P(C \cap A) \rightarrow 3 \text{ terms}$$
$$+ P(A \cap B \cap C) \rightarrow 1 \text{ term}$$

$$P(A \cup B \cup C \cup D) = \binom{4}{1} - \binom{4}{2} + \binom{4}{3} - \binom{4}{4} \text{ terms}$$

Page No.
Date / /

8) Suppose n letters are put in n envelops distinct by addresses. What is the prob. that no letter will reach to the correct address. What is the limiting probability as $n \uparrow \infty$?

→ A_i = i^{th} letter is put in correct envelop.

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

$$= 1 - \left(\sum_{k=1}^{n} (-1)^{k-1} S_k\right)$$

$$= 1 - \sum_{k=1}^{n}\left((-1)^{k-1}\binom{n}{k}\frac{(n-k)!}{n!}\right)$$

$$\frac{1\,2\ldots k}{\uparrow\uparrow\uparrow\uparrow\uparrow} \Big| \frac{(n-k)!}{}$$

$$= 1 - \left(\sum_{k=1}^{n} (-1)^{k-1}\frac{1}{k!}\right)$$

$$= 1 - 1 + \frac{1}{2} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} \cdots + (-1)^n \frac{1}{n!}$$

$$= e^{-1} \text{ as } n \uparrow \infty$$

Bayes Theorem

Let $A_1, A_2, \ldots A_k$ is a partition (mutually exclusive & exhaustive) of Ω and probability space $= (\Omega, \mathcal{A}, P)$ with $P(A_i) > 0 \ \forall i$ and $P(B) > 0$ for some $B \subseteq \Omega$, Then

$$P(A_i|B) = \frac{P(B|A_i)\,P(A_i)}{\sum_{i=1}^{k} P(B|A_i)P(A_i)}$$

Proof :- $P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)\,P(A_i)}{P(B)}$

$$P(B) = P(B \cap \Omega) = P\left(B \cap \left(\bigcup_{i=1}^{k} A_i\right)\right) = P\left(\bigcup_{i=1}^{k} (B \cap A_i)\right)$$

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

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

Q) In a city 40% Male, 60% Female | Among Male 50% smoker | Among female 30% smoker.

P(A randomly chosen smoker is a male)

→ A_1 = Male B = smoker
A_2 = Female

$P(A_1) = 0.4$, $P(A_2) = 0.6$, $P(B|A_1) = 0.5$, $P(B|A_2) = 0.3$

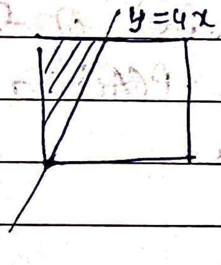
$$P(A_1|B) = \frac{P(B|A_1).P(A_1)}{P(B|A_1).P(A_1) + P(B|A_2).P(A_2)}$$

$$= \frac{0.5 \times 0.4}{0.5 \times 0.4 + 0.6 \times 0.3}$$

$$= \frac{0.2}{0.2 + 0.18} = \frac{\cancel{2}\ 1}{\cancel{3.8}\ 19} = \frac{10}{19}$$

$P(A_1|B) = 10/19$

Q) Consider the quadratic equation $u^2 - \sqrt{4u} + X = 0$ where (x, y) is a random point chosen uniformly from a unit square. What is the probability that the equation will have a real root?

→ 

$4 - 4x > 0$

$$P = \frac{\text{area of } \Delta}{\text{area of } \square} = \frac{\frac{1}{2} \times 1 \times \frac{1}{4}}{1 \times 1} = \frac{1}{8}$$

Q Give a randomized algorithm to approximate value of
a) π b) e

→ monte carlo simulation

Note :- Random Variable is not Random,
it is deterministic func

Random Variable

$(\Omega, \mathcal{A}, P) \longrightarrow$ Probability space

Eg :-

A coin is tossed 3 times

Ω	Total H	H−T	
{ HHH	3	3	
HHT	2	1	
HTH	2	1	
HTT	1	−1	
THH	2	1	
THT	1	−1	
TTH	1	−1	
TTT }	0	−3	
	$\downarrow$ u	$\downarrow$ v	

$\begin{cases} H \longrightarrow 1 \\ T \longrightarrow 0 \end{cases}$ $x(\{H\}) = 1$, $x(\{0\}) = 0$

$u(\{H, H, H\}) = 3$

$u(\{H, T, H\}) = 2$

$u = x_1 + x_2 + x_3$

$Y = \dfrac{X - 0.5}{0.5} = \begin{cases} 1 & \text{if } x = 1 \\ -1 & \text{if } x = 0 \end{cases}$

$V = Y_1 + Y_2 + Y_3$

- We can assign real numbers to the elements of Ω
- " " " vectors to the elements of Ω
- Some 'nice' function of previous function from $\Omega \to \mathbb{R}$
- All are deterministic → law is well defined when domain & Range is fixed

Let $(\Omega, \mathcal{A}, P)$ is a probability space

$X : \Omega \longrightarrow \mathbb{R}$ i.e. $X(\omega) \in \mathbb{R}$ when $\omega \in \Omega$

$X^{-1}((-\infty, x]) = \{\omega \mid X(\omega) \leq x\} \in \mathcal{A} \ \forall \ x \in \mathbb{R}$

Then we call X is a random variable

→ We will be able to calculate the probability of $X \leq x$

$\Omega = \{HHH, \ldots\ldots\ldots TTT\}$

$\mathcal{A} =$ power set

$P(H) = p$ $P(T) = q$ s.t. $p + q = 1$

$P(\{H, H, H\}) = p^3$

$P(\{H, T, H\}) = p^2 q$

Page No.
Date / /

$P(u \le -\infty) = 0$

$P(u \le -1) = 0$

$P(u \le 0) = q^3$

$P(u \le 1) = q^3 + 3q^2p$

$P(u \le 2) = q^3 + 3q^2p + 3qp^2$

$P(u \le 3) = q^3 + 3q^2p + 3qp^2 + p^3 = (q+p)^3 = 1$

$P(u \le \infty) = 1$

Consider a function $g: \mathbb{R} \to \mathbb{R}$ and X is a random variable on $(\Omega, \mathcal{A}, P)$.
Then $Y = g(X)$ is also a random variable.
It implies that $X^{-1}(g^{-1}(-\infty, x]) \in \mathcal{A}$ for any $x \in \mathbb{R}$.
It means $Q((-\infty, x]) = P(Y \in (-\infty, x]) = P(X \in g^{-1}((-\infty, x]))$
which is known as **PUSH FORWARD** probability.

$$P(V \le -3) = P(Y_1 + Y_2 + Y_3 \le -3)$$
$$= P(X_1 + X_2 + X_3 \le 0)$$
$$= P\{\omega \mid X_1(\omega) + X_2(\omega) + X_3(\omega) \le 0\}$$
$$= P(\{T, T, T\}) = q^3$$

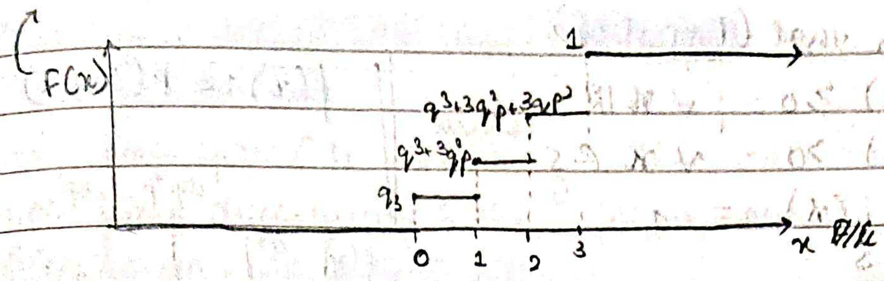
Cumulative distribution function (c.d.f)

c.d.f of a random varible X is a function $F: \mathbb{R} \to [0,1]$ defined as

$$F(x) = P(X \le x)$$
$$= P(X^{-1}(-\infty, x])$$
$$= P(\{\omega \mid X(\omega) \in (-\infty, x]\}) \quad \forall x \in \mathbb{R}$$

c.d.f uniquely identifies a random variable

- c.d.f can ~~fr~~ have finite / countable no. of jump discontinuities.
- The jumps have to finally sum to '1'. (total probability)



Properties.

→ ~~P(~~ $F(-\infty) = \lim_{n\downarrow -\infty} F(n) = 0$

→ $F(+\infty) = \lim_{n\uparrow \infty} F(n) = 1$

→ $F(x) \leq F(y) \quad \forall\ x \leq y \in \mathbb{R}$ (Non-decending)

→ $F(a) = \lim_{x\downarrow a} F(x) \quad \forall\ x \in \mathbb{R}$ (Right continuity)

Discrete valued Random variable.

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

a Random variable X is said to be a discrete valued random variable if $S = \{X(\omega) \mid \omega \in \Omega\}$ is a finite or countably ~~infity~~ infinite ~~ste~~ set and $X^{-1}(S_i) \in \mathcal{A} \quad \forall\ S_i \in S$

$u \in \{0, 1, 2, 3\} \subset \mathbb{R}$
$v \in \{-3, -1, 1, 3\} \subset \mathbb{R}$

Need not be integer valued

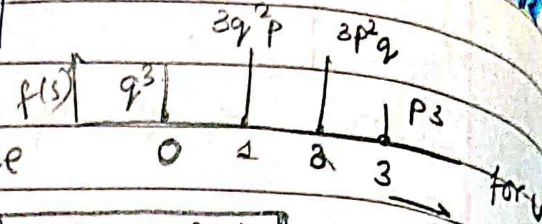
Probability mass function (p.m.f).

Let X be a discrete valued random variable and there exist a prob. space $(\Omega, \mathcal{A}, P)$ then the non negative function $f(x): \mathbb{R} \rightarrow [0,1]$

such that (Properties)

- $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
- $f(x) > 0 \quad \forall x \in S$ $\Rightarrow$ when $X(\Omega) = S$
- $\sum_{s \in S} f(s) = 1$
- $S = \{x \mid f(x) > 0\}$ is finite/countable set

$f(x) := P(X = x), \; x \in \mathbb{R}$



for v

c.d.f interms of p.m.f $\quad \boxed{F(x) = \sum_{s \leq x} f(x)}$

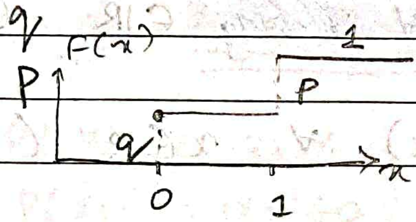
$F(x) = P(X \leq x) \quad , x \in \mathbb{R}$

$\downarrow \qquad = \sum_{s \leq x} f(s) \quad , s \in S$

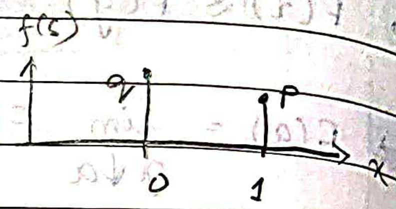
c.d.f $\qquad\qquad\quad \downarrow$ p.m.f

for discrete valued random variable.

$X = 0$ with prob q
$X = 1$ " " p



$F(x)$



$f(s)$

For the coin toss eg:-

$F(2) = f(0) + f(1) + f(2) = \sum_{s \leq 2} f(s) \quad s \in X(\Omega)$

$P(x \leq 2) = P(x=0) + P(x=1) + P(x=2) = \sum_{x \leq 2} f(x)$

$\hookrightarrow$ Discrete Valued Random variable

Note that :- • We can find cdf $F(x)$ for $x \notin S$

• In $\sum_{s \leq x} f(s)$, we need to consider only points where there is a jump in a discrete setup

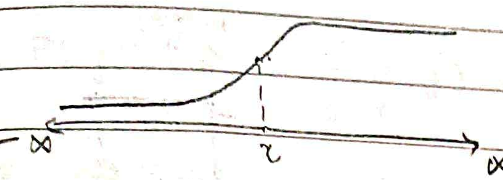
Vector valued Random variable?

\# Note: If X & Y are random variables

- $aX + bY$, $\max\{X, Y\}$, $\min\{X, Y\}$, XY are random variables $\forall a, b \in \mathbb{R}$
- If $P(Y(\omega) = 0) = 0$, then X/Y is also a random variable.

Continuous valued Random Variable

For prob. space $(\Omega, \mathcal{A}, P)$ if the cdf has no-jump discontinuity & the property such that
$P(X=x)=0 \quad \forall x \in \mathbb{R}$, $X : \Omega \to \mathbb{R}$



(ABSOLUTE CONTINUOUS)

$F(x) - F(x^-) = 0$

$F(x^+) - F(x) = 0$

Probability density function

$f : \mathbb{R} \longrightarrow [0, \infty)$

$P(X \in A) = \int 1_{(x \in A)} f(x)\, dx \quad A \in \mathcal{A}$

$= \int_{-\infty}^{\infty} 1_{(x \in A)} f(x)\, dx$

$1_A = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$

If $A = (-\infty, x]$

$P(X \in (-\infty, x]) = F(x) = \int_{-\infty}^{x} f(t)\, dt = \int_{-\infty}^{\infty} 1_{(x \in A)} f(x)\, dx$

$\forall x \in \mathbb{R}$

$F(x) \to$ ~~pdf~~ cdf

? $\underbrace{f(x)}_{\cancel{\text{cdf}}\ \text{pdf}} = \frac{d}{dx} F(x) = \frac{d}{dx} \int_{-\infty}^{x} f(t)\, dt$

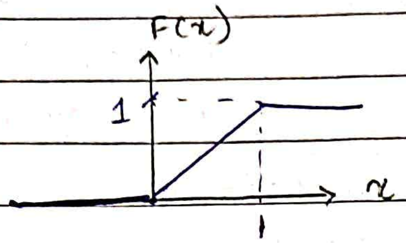
$\underbrace{f(x)}_{\text{pdf}} = \frac{d}{dx} \underbrace{F(x)}_{\text{cdf}} = \frac{d}{dx} \int_{-\infty}^{x} f(t)\, dt$

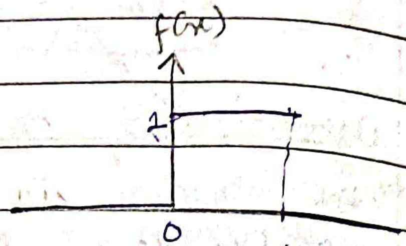
Probability Density function (pdf):- If X is a continuous valued random variable in a given probability space $(\Omega, \mathcal{A}, P)$ with cdf $F(\cdot)$ then a non negative funⁿ $f : \mathbb{R} \to [0, \infty)$ is pdf of X if $P(X \in A) = \int f(x)\, 1_{(x \in A)}\, dx$

Page No.
Date / /

Eg1:- X is said to follow $U[0,1]$ distribution
If the cdf is $F(x) = \begin{cases} 0 & , x<0 \\ x & , 0 \le x \le 1 \\ 1 & , x>1 \end{cases}$

$\therefore$ pdf of the function is $f(x) = \begin{cases} 1 & \text{if } x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$





if $0 \le a < b \le 1$

$P(x \in (a,b)) = P(x \in (a,b]) = P(x \in [a,b)) = P(x \in [a,b])$
$= b-a$

$F(x) = P(X \le x) = P(0 \le X \le x) \quad \forall x \in [0,1]$

Q) $Y = 0.5X$ when $X \sim U[0,1]$
Find the pdf & cdf of Y

$\implies$ $Y \in (0, 0.5)$

$F(y) = P(Y \le y)$
$= P(0.5x \le y)$
$= P(x \le 2y)$
$= 2y \quad \forall y \in [0, 0.5]$

cdf $\rightarrow$ $F(y) = \begin{cases} 0 & , y<0 \\ 2y & , y \in [0,0.5] \\ 1 & , y>0.5 \end{cases}$

pdf $\rightarrow$ $f(y) = \begin{cases} 2 & , y \in [0,0.5] \\ 0 & , \text{otherwise} \end{cases}$

we must define pdf & cdf on entire real line

Discrete vs continuous

x is discrete, $x \in [0, \infty)$ with pmf $f(x)$

$P(x \leq 5) = f(0) + f(1) \dots + f(5) = \sum f(n)$

* y is continuous, $y \in [0, \infty)$ with pdf $g(y)$

$P(y \leq 5) = \int_0^5 g(y)\, dy$

pmf runs from $[0, 1]$ → for discrete
pdf " " $[0, \infty)$ → for continuous
↳ area under pdf graph = 1

Error

Set of data :- $\{d_1, d_2, \dots d_n\}$
Mean/AM / Median / Mode :- A
Squared error : $e = \sum_{i=1}^{n} (d_i - A)^2$

mean minimizes the squared error loss

Absolute error :- $e = \sum_{i=1}^{n} |d_i - A|$

median minimizes the absolute error loss

Mean

Theoretical mean of a random variable X / Expectation of X :-
Expectation of X $E(X) = \int x \, dF(x)$

$$= \begin{cases} \sum_x x f(x), & \text{if } \sum_x |x| f(x) < \infty \\ \int_x x f(x) dx, & \text{if } \int_x |x| f(x) < \infty \end{cases}$$

Absolute summability of a series

here $F(x)$ is cdf, Σ for discrete & $\int$ for continuous

Eg:- $f(x) = \frac{1}{|x|(1+|x|)}$, $x \in S = \{(-1)^n n \mid n \in \mathbb{N}\}$

$$\sum_{x=1}^{\infty} x f(x) = \frac{-1}{1\cdot 2} + \frac{2}{2.3} + \frac{-3}{3.4} + \dots$$

$$= \left(+1 \frac{-1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6}\right)_{-1} = \log_e 2 - 1$$

$$\sum_x |x| f(x) = \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} \dots$$ harmonic series

$$> \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \dots$$

$$= \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots + \infty$$

diverging.

Thus we cannot say that $\log_e 2 - 1$ is the expectation value of X exist.

Q. $f(x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x \in \{0, 1, 2, \dots\}$

Find $E(X) = ?$

→

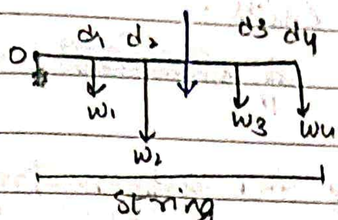
Moment Generating function (m.g.f)

The moment generating function of a random variable X is defined as

$$M_X(t) = E(e^{tX}) \text{ if } E(e^{tX}) < \infty \ \forall t \in (-e, e)$$

for some $e > 0$

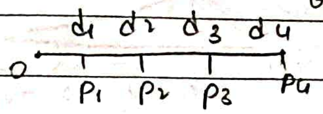
1/C.d.f Eg:



O is fixed, w_i is force at d_i

Resultant will act at $\frac{\sum w_i d_i}{\sum w_i}$ (centre of mass)

⇓ Similarly for probabilities on a number line



$$\frac{\sum d_i p_i}{\sum p_i} \Rightarrow \boxed{\frac{\sum d_i p_i}{1}}$$

First order Raw moment

(or)

→ theoritical expectation (if exists)

(or)

→ mean (if exists)

(measures centre of data)

$$\mu \text{ or } \mu_1' = \begin{cases} \sum x_i f(x_i) \\ \text{or} \\ \sum x_i p_i \\ \int_x x f(x)\,dx \end{cases} = E(X)$$

r-th order Raw moment

$$E(X^r) = \mu_r' = \begin{cases} \sum f(x_i) x_i^r \\ \int_x x^r f(x)\,dx \end{cases} \quad \text{if exists} \Leftrightarrow E(|x|^r) < \infty$$

r-th order central moment is defined as

$[\mu_1' = \mu]$

$$\mu_r = E\left((X-\mu)^r\right) \quad \text{if exists}$$

Now

$\mu_1 = 0$ non informative

$\rightarrow$ if exists

$\mu_2 = E\left((X-\mu)^2\right) = \mu_2' - \mu^2 = E(x^2) - \left(E(X)\right)^2$

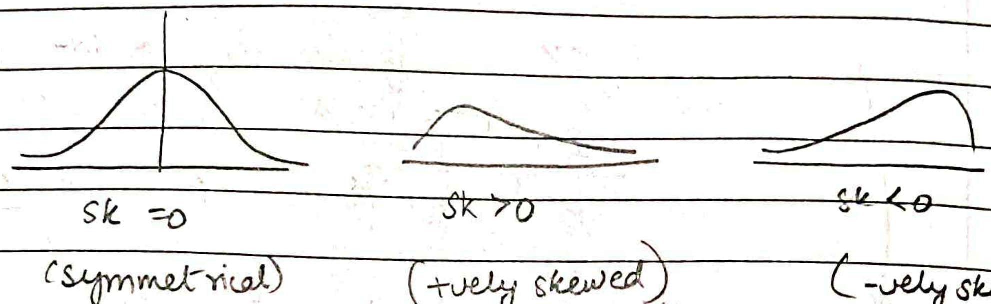
$\hookrightarrow$ Variance of X (σ^2)

Expected squared deviation of from its mean $E(x) = \mu$

$\mu_3 = E\left((X-\mu)^3\right) \rightarrow$ if exists.

measure of of skewness $= \dfrac{\mu_3}{\sigma^3} = \dfrac{\mu_3}{\mu_2^{3/2}}$

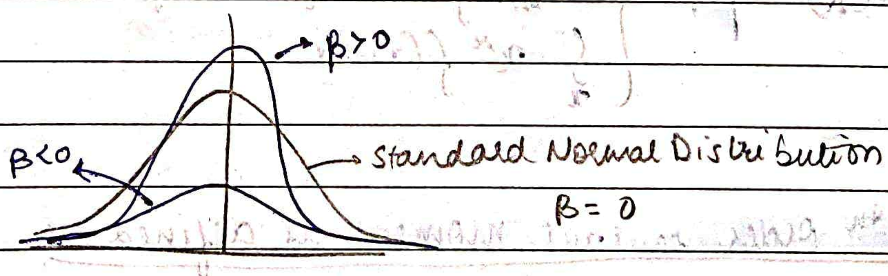
(It measures the assymetry of the distribution)



Note:- To make skewness unit free, σ^3 is used in the denominator as it is always a +ve value and doesnt affect the sign or the value of skewness

$\mu_4 = E\left((X-\mu)^4\right) \rightarrow$ if exists

β = measure of Kurtosis $= \dfrac{\mu_4}{\mu_2^2} - 3 = \dfrac{\mu^4}{\sigma^4} - 3$



It measures the peak of the distribution.

Sample Version

$\mu = \mu_1' \Rightarrow m_1' = m = \frac{1}{n} \sum x_i$

$\mu_2' = \frac{1}{n} \sum x_i^2$, $\mu_2 = \frac{1}{n} \sum (x_i - \mu)^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$

$\mu_r' \hat{=} \frac{1}{n} \sum x_i^r = m_r'$, $\mu_r \hat{=} \frac{1}{n} \sum (x_i - \mu)^r = m_r$

$$S^2 = \underbrace{\frac{1}{n} \sum (x_i - \bar{x})^2}_{\text{interprets how data moves away from mean}} = \underbrace{\frac{1}{2n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - x_j)^2}_{\text{interprets scattering of data}}$$

Proof:-

$$RHS = \frac{1}{2n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - \bar{x} + \bar{x} - x_j)^2$$

$$= \frac{1}{2n^2} \sum_i \sum_j \left\{ (x_i - \bar{x})^2 + (x_j - \bar{x})^2 - 2(x_i - \bar{x})(x_j - \bar{x}) \right\}$$

$$= \frac{1}{2n^2} \left(n \sum_i (x_i - \bar{x})^2 + n \sum_j (x_j - \bar{x})^2 - 2 \underbrace{\cancel{\sum_i (x_i - \bar{x})}}_{0} \underbrace{\cancel{\sum_j (x_j - \bar{x})}}_{0} \right)$$

$$= \frac{1}{n} \sum_i (x_i - \bar{x})^2 = LHS$$

$\sqrt{\text{Variance}}$ = Standard Deviation = s.d

Distribution can be measured in many ways, variance is just one of them.

Moment generating function

A random variable X is said to have moment generating function (m.g.f) $M_X(t)$ if $E(e^{tX}) < \infty$

then $M_X(t) = E(e^{tX})$ $\quad -\epsilon < t < \epsilon$, $\epsilon > 0$

Note if mgf is same for 2 distributions, then c.d.f & (p.d.f) is all same for both

$$M_X(t) = E(e^{tX})$$
$$= E\left(\sum_{k=0}^{\infty} \frac{(tX)^k}{k!}\right)$$
$$= \sum_{k=0}^{\infty} \frac{t^k E(X^k)}{k!} \text{ , if finite}$$

Now $\left.\frac{\partial^r}{\partial t^r} M_X(t)\right|_{t=0}$

$$= \frac{r!\, E(X^r)}{r!} = \boxed{E(X^r) = \mu_r'}$$

terms $k<r$ & $k>r$ vanish

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k \mu_k'}{k!}$$

Even ~~thou~~ though all moments exist, mgf may not exist

Note: Two random variables have same moments but different distributions (pdf, cdf) etc. → Not in the scope of this course

Markov Inequality

$P(X \geq t) \leq \frac{E(X)}{t}$, $E(X)$ exists , $P(X \geq 0) = 1$, $t > 0$ (Random variable has to be non negative)

$$E(X) = \int_0^{\infty} x f(x)\, dx$$
$$= \underbrace{\int_0^{t} x f(x)\, dx}_{0 \text{ or } +ve} + \int_t^{\infty} x f(x)\, dx$$
$$\geq \int_t^{\infty} x f(x)\, dx$$
$$\geq \int_t^{\infty} t f(x)\, dx \quad (t \text{ is constant}, t > 0)$$

$$= t\int_{t}^{\infty} f(x)\,dx$$

$$= t\,P(x>t)$$

$$\boxed{E(X) \geq t\,P(X>t)}$$ when $P(X\geq 0)=1$, $t>0$

Chebyshev's Inequity

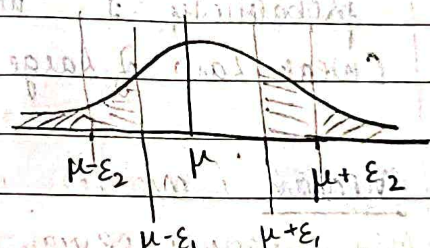
X is a random variable with finite expectation. $E(X)=\mu$ and finite variance

$$\boxed{P(|X-\mu|>\epsilon) \leq \frac{E[(X-\mu)^2]}{\epsilon^2}}$$

→ Finite expectation

→ $\epsilon>0$

→ $\sigma^2 = E(X-\mu)^2 < \infty$



$P(|X-\mu|>\epsilon_1) > P(|X-\mu|>\epsilon_2)$

$\epsilon_1 < \epsilon_2$

Upperbound given by $\frac{E(X-\mu)^2}{\epsilon^2}$

Proof :-

X is a random variable with finite mean & variance

Let $Y = |(X-\mu)|$, hence $Y\geq 0$

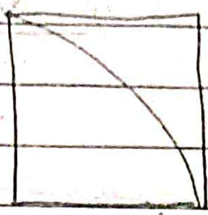
$P(|X-\mu|>\epsilon)$

$= P((X-\mu)^2 > \epsilon^2) \leq \frac{E(X-\mu)^2}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}$ using Markov Inequality

$P(|X-\mu|>\epsilon) \leq E[(X-\mu)^2]/\epsilon^2$

$$\boxed{P(|X-\mu|<\epsilon) \geq 1 - E[(X-\mu)^2]/\epsilon^2}$$

Q) Estimation of value of π



$\Rightarrow$ P(within the circle) $= p = \pi/4$

$E(X) = p = \mu$

$V(X) = E(X-\mu)^2 = p(1-p)$

$E(\bar{X}) = p = \mu = \pi/4$

$V(\bar{X}) = p(1-p)/n$

$$P(|\bar{X} - p| < \varepsilon) \geq 1 - \frac{E(\bar{X}-p)^2}{\varepsilon^2}$$

$$= 1 - \frac{p(1-p)}{n\,\varepsilon^2}$$

Sample mean $(\bar{X})$ converges to the population mean with probability 1 when variance is finite.
(Weak Law of Large Numbers)

Median (middle element of the distribution when sorted)
(Mean Absolute Deviation – MAD)

$g(a) = \frac{1}{n} \sum |x_i - a|$

→ $g(a)$ is minimized when 'a' is the median

→ median need not be unique (Eg:- when there are even no. of observation)

theoretically, let m = median

$P(X \geq m) \geq \frac{1}{2}$ & $P(X \leq m) \geq \frac{1}{2}$ → discrete

$P(X \geq m) = \frac{1}{2} = P(X \leq m)$ → continuous

Mode :- x for which $f(x)$ is maximum
$\downarrow$
pdf/pmf

Symmetric distribution

median = mean = mode

Asymmetric distribution

Right skewed

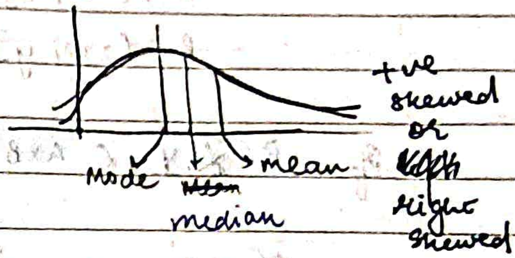
+ve skewed 

mean > median > mode

left skewed

-ve skewed 

mean < median < mode



Quatile

$$P(X \leq Q_1) = \frac{1}{4}$$

$$P(X \leq Q_2) = \frac{1}{2} \longrightarrow \text{median}$$

$$P(X \leq Q_3) = 3/4$$

$$\text{skewness} = \frac{|Q_3 - Q_2| - |Q_2 - Q_1|}{|Q_3 - Q_1|}$$

Note:- median, mode & quatile will always exist, Mean may not always exist

Quatile helps to measure skewness when mean doesn't exist

Modeling with Random Variables

Independent Random Variables:

$A : (-\infty, x]$, $x \in \mathbb{R}$
$B : (-\infty, y]$, $y \in \mathbb{R}$

If $P((X,Y) \in A \times B) = P(X \in A)\, P(Y \in B)$

AND

$P(X \le x, Y \le y) = P(X \le x)\, P(Y \le y)$
$\forall\, x, y \in \mathbb{R}$

Then we say that X & Y are independently distributed.

↳ Also implies that $f(x,y) = f_X(x)\, f_Y(y)$?

Identically Distributed Random Variable

Consider X & Y are two random variables and if
$P(X \le t) = P(Y \le t) \quad \forall t \in \mathbb{R}$
then we say that X & Y are identically distributed

i.i.d $\Leftrightarrow$ independent & identical distribution

Uniform [0, 1]   U[0, 1]

X is said to follow U[0, 1] if A is an interval in [0, 1] then $P(X \in A) = \dfrac{\text{length}(A)}{\text{length}\,[0,1]}$
$= \text{length } A$

$P((a,b)) = P((a,b]) = P([a,b)) = P([a,b]) = b - a$
$(0 \le a \le b \le 1)$

$$f(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & \text{otherwise} \end{cases} \qquad F(x) = \begin{cases} 0, & x < 0 \\ x, & x \in [0,1) \\ 1, & x > 1 \end{cases}$$

↳ for uniform $[0,1]$

Discrete Uniform

$0 \quad a_1 \quad a_2 \quad \dots \quad a_{k-1} \quad a_k$

$a_0 = 0, \ a_1 = 1/k, \ a_2 = 2/k, \ \dots \ a_k = k/k = 1$

Let $U \sim U[0,1]$ if $a_i \le u \le a_{i+1}$ then $X = i+1$

$P(X=j) = 1/k$, $j = 1, 2, \dots k$

→ $X \sim$ discrete uniform with $P(X=j) = 1/k$ when $j = 1 \dots k$

$Y = s_i$ when $a_i < x \le a_{i+1}$

$$\boxed{P(Y=s_i) = 1/k}$$

→ $Y \sim$ discrete uniform on $\{s_1, s_2 \dots s_k\}$
↳ need not be integers

☆

Bernoulli (p)

$X \sim B(p)$ — $p \to$ probability of success or $x=1$

if pdf is / pmf is

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

$U \sim U[0,1]$

$0 \quad p \quad 1$

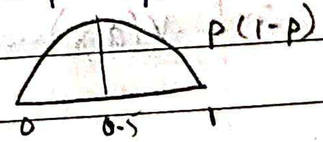
$$X = \begin{cases} 1 & \text{if } 0 \le U \le p \\ 0 & \text{if } p < U \le 1 \end{cases}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ 1-p, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

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

$V(X) = E(X^2) - [E(X)]^2$
$= p - p^2$
$= p(1-p)$

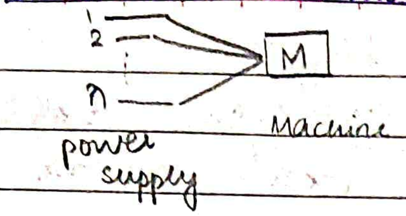
Variance is maximum if p is 0.5



Eg:-
— Tossing a coin
→ Switch of machine (on/off)

Binomial Distribution (sum of n iid Bernoulli(p))

Eg



→ 'M' works when exactly 'k' out of 'n' connections are on

→ each connection can be on with probability 'p'

→ they can be on/off indently

$$P(\text{'M' will work}) = \binom{n}{k} p^k (1-p)^{n-k}$$

let $X_1, X_2, \ldots X_n \overset{iid}{\sim}$ Bernoulli (p)

Y = no. of functioning power supply $= \sum_{i=1}^{n} x_i \sim$ binomial (n,p)

$$f(y) = P(Y=y) = \begin{cases} \binom{n}{y} p^y (1-p)^{n-y} & \text{if } y = 0,1 \cdots n \\ 0 & \text{otherwise} \end{cases}$$

pmf of Y

$Y \sim$ binomial (n, p)

$E(Y) = E(\sum x_i) = \sum E(x_i) = np$

$V(Y) = np(1-p) = n\,V(x_i)$

↳ independent

Note :-

Bernoulli (p) = binomial (1, p)

True facts for all Random variables

- X and Y are two random variables then $E(X+Y) = E(X) + E(Y)$

- X & Y " " " independent random variables then $E(X \cdot Y) = E(X) \cdot E(Y)$

  & $V(X+Y) = V(X) + V(Y)$

- $E(aX) = a\,E(X)$
- $V(aX) = a^2\,V(X)$

Geometric Distribution

Eg: → A coin has success probability p

→ the coin is tossed untill the first success

→ Y := number of failures preceding to the first success

→ all trials are independ.

$$\Longrightarrow f(y) = \begin{cases} (1-p)^y p & \text{for } y = 0, 1, 2 \ldots \infty \\ 0 & \text{otherwise} \end{cases}$$

S → $y=0$ p

FS → $y=1$ $(1-p)p$

FFS → $y=2$ $(1-p)^2 p$

$$E(Y) = \sum_{y=0}^{\infty} y f(y)$$

Waiting time distribution in discrete setup ⇒ Geometric Distribution

Calculating $E(Y)$

Let $q = 1-p$

$$\left.\begin{array}{l} 0 f(0) \\ +1 f(1) \\ +2 f(2) \\ \vdots \end{array}\right\} \begin{array}{l} 0 \\ +1 f(1) \\ + f(2) + f(2) \to q^2 p \\ + f(3) + f(3) + f(3) \end{array}$$

$$E(Y) = \sum_{y=0}^{\infty} y f(y) = q + q^2 + q^3 + \ldots = \frac{q}{1-q} = \frac{q}{p} = \frac{1-p}{p}$$

?|| If z is a non-negative random variable then

$$E(z) = \sum_{k=1}^{\infty} P(z \geq k)$$

↳ true in general for discrete valued random variable

Now $E(Y^2)$

$$\begin{array}{l} 0 f(0) \\ +1 f(1) \to f(1) \\ +4 f(2) \to f(2) \quad f(2) \quad f(2) \quad f(2) \\ +9 f(3) \to f(3) \quad f(3) \quad f(3) \quad f(3) \quad \ldots \quad 9 \text{ times} \end{array}$$

(Arithmetico Geometric Series) $E(Y^2) = \underbrace{q}_{1 \text{ time}} \quad \underbrace{q^2 \quad q^2 \quad q^2}_{3 \text{ time}} \quad q^3 \text{ (5 times)} \quad q^4 \text{ (7 times)}$

H.W

$Y \sim geo(p)$
$E(Y) = q/p$
Prove that $V(Y) = q/p^2$

$\Rightarrow E(Y^2) = q + 3q^2 + 5q^3 + 7q^4 + \dots$
$= q(1 + 3q + 5q^2 + \dots) = q\left(\frac{1}{p} + \frac{2q}{p^2}\right)$

now $(E(Y))^2 = \frac{q^2}{p^2}$

$V(Y) = E(Y^2) - (E(Y))^2 = \frac{q}{p} + \frac{2q^2}{p^2} - \frac{q^2}{p^2}$
$= \frac{q}{p}\left(1 + \frac{2q}{p} - \frac{q}{p}\right) = \frac{q}{p^2}$

Note:-
some books define geometric dist. as

Z: no. of trials to get first succ.

$Z = Y + 1$

$g(z) = \begin{cases} p(1-p)^{z-1} & \text{for } 1, 2, \dots \\ 0 & \end{cases}$

$E(Z) = 1 + \frac{q}{p} = \frac{1}{p}$

$V(Z) = q/p^2$

(Variance is location invariant
Mean is " equivalent)

Negative Binomial

Eg:- W: number of failures preceeding to the r^{th} success

$W = \sum_{i=1}^{r} Y_i$. $Y_i \overset{iid}{\sim} geo(p)$ $W \sim NB(r, p)$

$f(w) = \begin{cases} \binom{w+r-1}{w} q^w p^r & w = 0, 1, \dots \\ 0 & \text{otherwise} \end{cases}$

$E(W) = r\, q/p$
$V(W) = r\, q/p^2$

Discrete Valued Random Variables

$$\begin{array}{ccccc} U[0,1] & \rightarrow & Bernoulli(p) & \xrightarrow{iid} & Binomial(n,p) \\ \downarrow & \nearrow & \downarrow & & \\ & & geometric(p) & \xrightarrow{iid} & negative\ binomial(r,p) \\ \text{Discrete uniform}(k) & & & & \end{array}$$

Hypergeometric Distribution (Urn Model Design / Biased coin Design)

Suppose there are 'n' number of items in a population out of them : $n_1 \rightarrow$ Catagory 1 $n_1 + n_2 = n$
$n_2 \rightarrow$ " 2

If we draw 'r' samples from the population without replacement then let 'x' denote no. of items from category 1 in the sample

$\underbrace{|n_1|n_2|}_{n} \longrightarrow$ draw r samples | $r = r_1 + r_2$

X = no. of n_1 in r samples

$$P(X=r_1) = \begin{cases} \dfrac{\binom{n_1}{r_1}\binom{n_2}{r_2}}{\binom{n}{r}} & , r_1 = 0, 1, 2, \ldots \min\{r, n_1\} \\ 0 & , \text{otherwise} \end{cases}$$

To find E(X)

Let
$X = z_1 + z_2 + \cdots + z_r$ $z_i = \begin{cases} 1 & \text{if sample } i \text{ from category 1} \\ 0 & \text{otherwise} \end{cases}$

$E(z_i) = \dfrac{n_1}{n}$

$\boxed{E(X) = \dfrac{r n_1}{n}}$ # $z_1 \ldots z_r$ are dependent but still holds due to sum law of Expectation

$\boxed{V(X) = r\left(\dfrac{n_1}{n}\right)\left(1 - \dfrac{n_1}{n}\right)\left(1 - \dfrac{r-1}{n-1}\right)}$ → cannot be proven now, need more knowledge

Page No.
Date / /

Poisson Random variable :- (limiting case values of bin(n,p) when $n\uparrow\infty$, $p\downarrow 0$ & $np=\lambda$)

$X \sim$ poisson (λ) , $\lambda>0$

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

(sums to 1 → proper pmf)

1 → Number of defects in a lot
2 → " " misprint " " page
3 → " " events in a given interval of time/space.
4 → " " spam mails in your account entering in an hour.

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

$\boxed{E(X) = Var(X) = E(X-\lambda)^3 = \lambda}$ → only for poisson (λ)

MGF of X i.e $M_X(t) = E(e^{tx})$

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

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

$$\boxed{M_X(t) = e^{\lambda(e^t-1)}}$$

for such events → probability of ~~occ~~ occurance is less
→ no. of events is more.

A lot has n items. Any item can be defected with probability p. Y is the total number of defected items in the lot

Y ~ binomial (n, p)

MGF of Y i.e $M_Y(t) = \sum_{t=0} e^{ty}\binom{n}{y} p^y (1-p)^{n-y}$

$$\underbrace{M_Y(t)}_{\text{binomial}(n,p)} = (pe^t + 1-p)^n$$

Assume p is a function of n, say p_n such that
$$n(p_n) \longrightarrow \lambda > 0 \quad , \text{ as } n \uparrow \infty$$

$Z_n \sim \text{bin}(n, p_n)$

$$M_{Z_n}(t) = (p_n e^t + 1 - p_n)^n \quad , \; n \uparrow \infty, \; p_n \downarrow 0 \; \& \; np_n \to \lambda$$

Let us consider
$$\lim_{n\uparrow\infty} M_{Z_n}(t) = \lim_{n\to\infty} (p_n e^t + 1 - p_n)^n$$
$$= \lim_{n\to\infty} \left(\frac{\lambda}{n} e^t + 1 - \frac{\lambda}{n}\right)^n$$
$$= \lim_{n\to\infty} \left(1 + \frac{\lambda}{n}(e^t - 1)\right)^n$$
$$\boxed{M_{Z_n}(t) = e^{\lambda(e^t-1)}} = M_X(t)$$

$\rightarrow$ $\lim\limits_{n\uparrow\infty} \text{bin}(n, p_n) = \text{poisson}(\lambda)$ as the limiting MGF is same
(here $\lambda = np_n$)

Theorem
Let $\{X_n\}$ be a sequence of Random variables with MGFs as $M_{X_n}(t)$ such that $\lim\limits_{n\uparrow\infty} M_{X_n}(t) = M_Y(t)$ for some random variable Y.
then $\lim\limits_{n\uparrow\infty} F_{X_n}(a) = F_Y(a)$ $\forall$ 'a' where F_Y is continuous

same cdf

$X_i \overset{iid}{\sim} \text{pois}(1)$ | $\rightarrow$ additive property

$$S_n = \sum_{i=1}^{n} X_i \sim \text{pois}(n)$$

Didn't understand

Poisson Process ?

A process or a stochastic process is a collection of random variables over time or space

Example :- (of stochastic process)

→ Number of customers standing in a queue at time t
$\{X_t \mid t \in T\}$

Counting process → type of stochastic process
A collection of random variables $\{N(t) \mid t \geq 0\}$ which represents the total number of ~~ev~~ events upto time 't' is called counting process

Properties of counting process
→ $N(t) \geq 0$ always
→ $s < t \Rightarrow N(s) \leq N(t)$
→ $s < t \Rightarrow N(t) - N(s)$ stands for the number of events in $(s, t]$

Poisson process :- A counting process is said to be poisson process with rate $\lambda (>0)$ if
→ $N(0) = 0$
→ the process has an independent increment
→ the number of events occurring in an interval with length t has poisson distribution with parameter (λt)

$E(N(t)) = \lambda t$
$V(N(t)) = \lambda t$

$$\boxed{P\left(N(t+s) - N(s) = k\right) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}} \rightarrow \text{free from } s$$

The distribution of the number of events in intervals of same length are identical

Continuous Distributions

1) Uniform (a, b)

$X \sim U[0,1]$

$Y = a + (b-a)X \quad , \quad b > a$

$F_Y(y) = P(Y \leq y) \quad , \quad y \in [a,b]$

$= P(a + (b-a)X \leq y)$

$= P\left(X \leq \frac{y-a}{b-a}\right)$

$$\boxed{F_Y(y) = \frac{y-a}{b-a}}$$

$$F_Y(y) = \begin{cases} 0 & \text{if } y < a \\ \frac{y-a}{b-a} & \text{if } y \in [a,b] \\ 1 & \text{if } y > b \end{cases}$$

$$f_Y(y) = \begin{cases} 0 & \text{if } y < a \text{ or } y \geq b \\ 1/b-a & \text{if } y \in [a,b) \end{cases}$$

$$E(Y) = \frac{1}{b-a} \frac{y^2}{2}\Big|_a^b = \frac{\cancel{(b-a)}(b+a)}{\cancel{(b-a)}2} = \frac{a+b}{2}$$

$$V(Y) = \frac{(b-a)^2}{12}$$

2) Exponential Distribution

$X \sim U(0,1)$

$Y = \frac{-1}{\lambda} \log_e (1-X)$, $\lambda > 0$

$Y \sim ?$

$F_Y(y) = P(Y \leq y)$ $\quad y \in (0, \infty)$

$= P(\frac{-1}{\lambda} \log_e (1-X) \leq y)$

$= P(X \leq 1-e^{-\lambda y})$

$= 1-e^{-\lambda y}$

$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0 \\ 1-e^{-\lambda y} & \text{if } y \geq 0 \end{cases}$$

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y} & \text{if } y \in (0, \infty) \\ 0 & \text{otherwise} \end{cases}$$

$E(Y) = \int_0^\infty y f_Y(y)\, dy = \int_0^\infty P(X > t)\, dt \rightarrow$ only for non-negative valued random variables

$\boxed{E(Y) = 1/\lambda}$ $\qquad \lambda \rightarrow$ rate parameter

$\boxed{V(Y) = 1/\lambda^2}$

→ Waiting time

→ Interarrival time of similar events

• Memory less property

$Y \sim \exp(\lambda)$

$P(Y > s+t \mid Y > t)$

The event has not occurred till time 't' and it will not occur till time 't+s' $t, s > 0$

$\cancel{= \frac{P(Y > s+t) \cap P(Y > t)}{P(Y > t)}}$

$$= \frac{P(\{Y>s+t\} \cap \{Y>t\})}{P(\{Y>t\})}$$

$$= \frac{P(Y>s+t)}{P(Y>t)}$$

$$P(\{Y>s+t \mid Y>t\}) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(Y>s)$$

Note:- • memoryless property is seen in ~~geo~~ geometric distribution
• exponential & geometric distributions are used as waiting time distribution.
• Both have memoryless property

Discretising exponential distribution

$Y \sim exp(\lambda)$ — Z is discretized form of Y (Z → geometric, Y → exponential)

$Y \in (0, \infty)$

$Z = [Y]$ → greatest integer function

$Z \sim ?$

$Z \in \{0, 1, 2 \dots\}$

$P(Z=0) = P(Y \in [0,1)) = 1 - e^{-\lambda}$

$P(Z=1) = P(Y \in [1,2)) = e^{-\lambda} - e^{-2\lambda}$

$\vdots$

$P(Z=k) = P(Y \in [k, k+1)) = e^{-\lambda k} - e^{-(k+1)\lambda}$

$$P(Z=k) = \begin{cases} (e^{-\lambda})^k (1-e^{-\lambda}) & , k \in \{0, 1, 2 \dots\} \\ 0 & , \text{otherwise} \end{cases}$$

$Z \sim geo(1-e^{-\lambda})$

5) Gamma distribution

$X_1, X_2, \ldots X_n \overset{iid}{\sim} \exp(\lambda)$

$Y = \sum_{i=1}^{n} X_i$ If X_i's are inter arrival time, then Y is the arrival time of the n^{th} event.

$Y \sim \text{Gamma}(n, \lambda) \longrightarrow E(Y) = n/\lambda$, $V(Y) = n/\lambda^2$

shape parameter (n), rate parameter (λ)

$Y \sim \text{Gamma}(\alpha, \lambda)$ if it has pdf

$$f_Y(y) = \begin{cases} \dfrac{e^{-\lambda y} y^{\alpha-1} \lambda^{\alpha}}{\Gamma\alpha}, & y>0,\ \lambda, \alpha > 0 \\ 0 & \text{, otherwise} \end{cases}$$

Note

- $\boxed{\text{Gamma}(1, \lambda) \equiv \exp(\lambda)}$ when $\alpha = 1$

- If $\left.\begin{array}{l} Z_1 \sim \text{Gamma}(\alpha_1, \lambda) \\ Z_2 \sim \text{Gamma}(\alpha_2, \lambda) \end{array}\right\}$ independent

  then $Z_1 + Z_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \lambda)$ → Additive property of gamma

$$\boxed{E(Y) = \frac{\alpha}{\lambda}}$$ Mgf = ?

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

cdf of gamma → if α is a +ve integer

$$F_Y(y) = e^{-\lambda y} \sum_{i=\alpha}^{\infty} \frac{(\lambda y)^i}{i!} = 1 - \sum_{i=0}^{\alpha-1} \frac{(\lambda y)^i}{i!} e^{-\lambda y}$$

4) **Beta(α_1, α_2)**

$Z_1 \sim G(\alpha_1, \lambda)$
$Z_2 \sim G(\alpha_2, \lambda)$ — Independent

Gamma function

$Y = \dfrac{Z_1}{Z_1 + Z_2}$, $Y \in (0,1)$

$Y \sim \text{Beta}(\alpha_1, \alpha_2)$

$$f_Y(y) = \begin{cases} \dfrac{y^{\alpha_1 - 1}(1-y)^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2)} & , y \in [0,1) \\ 0 & , \text{otherwise} \end{cases}$$

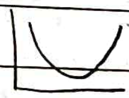
$B(\alpha_1, \alpha_2) = \dfrac{\Gamma\alpha_1 \, \Gamma\alpha_2}{\Gamma(\alpha_1 + \alpha_2)}$

$\Gamma\alpha = (\alpha - 1)\Gamma(\alpha - 1)$
$\Gamma n = (n-1)!$ if ~~isn~~ n is integer

$E(Y) = \dfrac{\alpha_1}{\alpha_1 + \alpha_2}$

$V(Y) = \dfrac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2 (\alpha_1 + \alpha_2 + 1)}$ ← check if correct

→ Plot $B(1/2, 1/2)$



Page No.
Date / /

4) Normal Distribution

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

mean (location parameter) — μ

variance (scale parameter) — σ^2

Let $Z \sim N(0,1)$ ← Standard Normal

$\mu + Z \sim N(\mu, 1)$

$X = \mu + \sigma Z \sim N(\mu, \sigma^2)$

$$\boxed{f(z) = \phi(z) = \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}, \quad z \in \mathbb{R}}$$

pdf of standard Normal $(N(0,1))$

proof of $\int_{-\infty}^{\infty} f(z)\,dz = 1$

Let us prove that

$$\sqrt{2\pi} = \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2}\,dz$$

$$c = \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\,dz \qquad \text{Let}$$

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

$x = r\cos\theta \qquad 0 < r$

$y = r\sin\theta \qquad 0 < \theta < 2\pi$

$$c^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2/2}\, r\,dr\,d\theta = \int_0^{\infty} e^{-r^2/2}\, r\,dr \int_0^{2\pi} d\theta$$

$c^2 = 1 \cdot 2\pi$

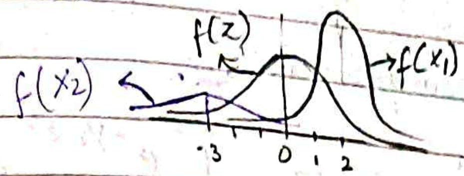
$c = \sqrt{2\pi}$

LHS = RHS, hence proved

→ $Z \sim N(0,1)$

$X_1 = 2 + 0.5Z$

$X_2 = -3 + 2Z$



$$N(\mu, \sigma^2) = \mu + \sigma N(0,1)$$

$\mu \in \mathbb{R}$

$\sigma \in \mathbb{R}^+$, non-negative

If $E(X)=0$ & $V(X)=0$, then $P(X=0)=1$

→ $X_1 \sim N(\mu_1, \sigma_1^2)$ > independent | → need to be independent

$X_2 \sim N(\mu_2, \sigma_2^2)$

$$a_1X_1 + a_2X_2 \sim N(a_1\mu_1 + a_2\mu_2, \ a_1^2\sigma_1^2 + a_2^2\sigma_2^2)$$

→ Iff $X_1, X_2 \dots X_n$ are normally distributed random variables then any linear combination are also Normally distributed

↳ need not be independent

$$\sum_{i=1}^{n} a_i x_i \sim N(\ ,\ )$$

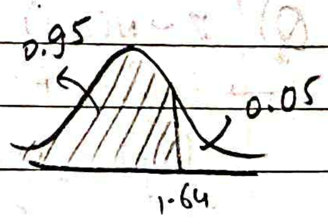
$E(\sum a_i x_i)$ → $V(\sum a_i x_i)$

holds only for normal distribution

→ $Z \sim N(0,1)$

$P(Z \leq 1.64) \approx 0.95$ || important results

$P(Z \leq 1.96) \approx 0.975$



Note:-

→ $Z \sim N(0,1)$

$\phi(z) = \phi(-z)$ → pdf of $N(0,1)$ (Standard Notation) | $\therefore P(Z \leq -1.64) = 0.05$

even function

$$\Phi(t) = P(Z \leq t) = \int_{-\infty}^{t} \phi(z)\,dz = \int_{-\infty}^{t} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\,dz$$

cdf of $N(0,1)$ (Standard Notation)

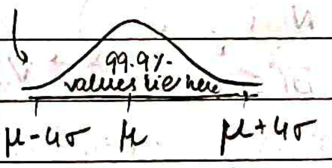
$$\Phi(t) = 1 - \Phi(-t)$$

→ $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} = \frac{1}{\sigma}\,\phi\left(\frac{x-\mu}{\sigma}\right) \qquad \mu \in \mathbb{R}, \ \sigma \in \mathbb{R}^+$$

→ Error distribution
→ Central limit theorem (CLT)
Normal approximation for large number
→ Height / Weight distribution

} practical uses



effective range

Q) $Z \sim N(0,1)$ and $Y = Z^2$, Find the distribution of Y

→ $F(y) = P(Y \leq y)$, $0 < y$

$$F(y) = P(Z^2 \leq y) = P(-\sqrt{y} \leq Z \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\,dz$$

$$F(y) = 2\int_{0}^{\sqrt{y}} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\,dz \implies f(y) = \frac{d}{dy}F(y) = \begin{cases} \dfrac{\sqrt{\frac{1}{2}}\, e^{-y/2}\, y^{\frac{1}{2}-1}}{\Gamma\frac{1}{2}}, & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

$Y \sim \text{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right)$ $\boxed{\frac{1}{\sqrt{2\pi}} = \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{\pi}} = \frac{\sqrt{\frac{1}{2}}}{\Gamma\frac{1}{2}}}$

$\rightarrow$ If $Z \sim N(0,1)$, $Z^2 \sim Gamma(1/2, 1/2)$

5) Chi-Square Distribution

If $Z_1, Z_2, Z_3, \ldots Z_n$ iid $N(0,1)$ then $\boxed{\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n}$

$\downarrow$ chi-square distribution of n-degrees of freedom

$Z^2 \sim \chi^2 \equiv G(1/2, 1/2)$
(α, λ)

$\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n \equiv G(n/2, 1/2)$

$E(\chi^2_n) = n \longrightarrow E(\Sigma Z_i^2) = \sum_{i=1}^{n} E(Z_i^2) = \sum_{i=1}^{n} 1 = n$

$V(\chi^2_n) = 2n$

$V(Z_i) = E(Z_i^2) - (E(Z_i))^2$, with $V(Z_i) = 1$, $E(Z_i) = 0$

(can also be proved through Gamma)

Note:- $\sum_{i=1}^{n} Z_i^2$ represents square of Euclidean Distance of $Z = (Z_1, Z_2 \ldots Z_n)$ from the origin
(Physical Representation)

6) T-distribution

If $Z \sim N(0,1)$ and $Y \sim \chi^2_k$ and both are INDEPENDENT

$\boxed{T = \frac{Z}{\sqrt{Y/k}} \sim t_k}$ T-distribution with k degrees of freedom

7) **F-distribution**

If $Y_1 \sim \chi^2_{k_1}$, $Y_2 \sim \chi^2_{k_2}$ } independent

$$\boxed{F = \frac{Y_1/k_1}{Y_2/k_2} \sim F_{k_1, k_2}}$$ F-distribution with k_1, k_2 degrees of freedom

Theorem

Let g be a strictly monotone function on $I = (a, b)$ with range $g(I)$ and differentiable.
Also assume g^{-1} is ~~di~~ also differentiable on $g(I)$.
Now let X be a continuous random variable with pdf $f(x) = 0$ if $x \notin I$ and
$Y = g(X)$ has pdf $h(y)$ on $g(I)$
then

$$h(y) = \underbrace{f(\underbrace{g^{-1}(y)}_{\text{old variable in terms of new}})}_{\text{old pdf}} \underbrace{\left|\frac{d}{dy} g^{-1}(y)\right|}_{\text{rate of change}}$$

Proof

Case I :-

- g is increasing (strictly)

$Y = g(x)$

$$F_Y(y) = P(Y \le y)$$
$$= P(g(X) \le y)$$
$$= P(X \le g^{-1}(y))$$
$$= \int_a^{g^{-1}(y)} f(t)\,dt$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = f(g^{-1}(y)) \frac{d}{dy}\left(g^{-1}(y)\right)$$

Case 2

- g is strictly decreasing

$$F_Y(y) = P(Y \leq y)$$
$$= P(g(X) \leq y)$$
$$= P(X \geq g^{-1}(y))$$
$$= 1 - P(X \leq g^{-1}(y))$$
$$= 1 - \int_a^{g^{-1}(y)} f(t)\,dt$$

$$f_Y(y) = f(g^{-1}(y))\left[(-1)\frac{d}{dy}g^{-1}(y)\right]$$
$$= f(g^{-1}(y))\left|\frac{d}{dy}(g^{-1}(y))\right|$$

Q) $Z \sim N(0,1)$

$Y = \mu + \sigma Z$ $\quad \mu \in \mathbb{R}$ $\quad f(z) = \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}$

$= g(z)$ $\quad \sigma > 0$

pdf of $Y = h_Y(y) = f(g^{-1}(y))\left|\frac{d}{dy}(g^{-1}(y))\right|$

now

$$Z = \frac{Y-\mu}{\sigma} = g^{-1}(y)$$

$$h_Y(y) = f\left(\frac{y-\mu}{\sigma}\right)\left|\frac{d}{dy}\left(\frac{y-\mu}{\sigma}\right)\right| = \frac{e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

- → $X \sim N(\mu, \sigma^2)$. Find the pdf of $Y = e^X$

log-Normal distribution

(waiting time distribution)

- $\left(\frac{X}{\lambda}\right)^k \sim \exp(1)$. then find the distribution of X.

X is said to follow **Weibull distribution** (waiting time distribution)

Pareto Distribution (Income distribution)

$$F(x) = \begin{cases} 1-\left(\frac{\lambda}{x}\right)^k & k>0,\ x>\lambda \\ 0 & \text{otherwise} \end{cases}$$

$X \sim F$ continuous valued random variable

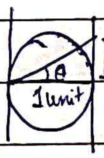
$Y = F(X) \in [0,1]$

$$P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y$$

$$G(y) = \begin{cases} 0 & y<0 \\ y & y \in (0,1) \\ 1 & y>1 \end{cases}$$

$Y \sim U[0,1]$

Cauchy (0,1): $C(0,1)$ → Stock Market Data



What will be the distribution of Y?

$\theta \sim U(-\pi/2, \pi/2)$

$Y = \tan\theta \in \mathbb{R}$

cdf $F_Y(y) = P(Y \leq y)$

$= P(\tan\theta \leq y)$

$= P(\theta \leq \tan^{-1} y)$

$$F_Y(y) = \frac{\tan^{-1}y - (-\pi/2)}{\pi/2 - (-\pi/2)}$$

$$\boxed{F_Y(y) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1} y}$$

$$\boxed{f_Y(y) = \frac{1}{\pi(1+y^2)}} \quad y \in \mathbb{R}$$

Note :- $E(X) \to$ does not exist for cauchy distributions as $E(|X|)$ does not converge.

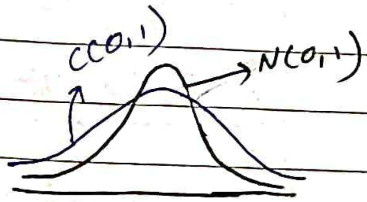
General form :-

Cauchy (μ, σ)

μ ↓ location parameter (MEDIAN)

σ ↓ Scale parameter

distribution is symmetric around μ



Let $Y \sim C(0,1)$

$Z = \mu + \sigma Y$, find the p.d.f of Z

→ $f_Y(y) = \frac{1}{\pi(1+y^2)}$, $y \in \mathbb{R}$

$Z = \mu + \sigma Y = g(y)$

$Y = \frac{Z-\mu}{\sigma} = g^{-1}(z)$

$$h_Z(z) = f\left(\frac{z-\mu}{\sigma}\right)\left|\frac{d}{dz}\left(\frac{z-\mu}{\sigma}\right)\right| = \frac{1}{\sigma\pi}\frac{1}{\left(1+\left(\frac{z-\mu}{\sigma}\right)^2\right)} \equiv \text{Cauchy}(\mu,\sigma)$$

Note :- $X_1, X_2 \overset{iid}{\sim} N(0,1)$

$\frac{X_1}{X_2} \sim C(0,1)$ | $\frac{X_1}{|X_2|} \sim C(0,1)$ | $t_1 \equiv C(0,1)$

Page No.
Date / /

Q) Let $x_1, x_2 \dots x_n \overset{iid}{\sim} F(x)$ where F is continuous and F' exists. Let $F'(x) = f(x)$ (say)

$$Y = \max\{X_1, X_2 \dots X_n\} = X_{(n)}$$

Find the cdf & pdf of Y

$\rightarrow$ $G(y) = P(Y \leq y)$

$$= P(X_{(n)} \leq y)$$

$$= P(x_1 \leq y, x_2 \leq y, \dots, x_n \leq y)$$

$$= \prod_{i=1}^{n} P(X_i \leq y) = \prod_{i=1}^{n} F(y)$$

$$G(y) = (F(y))^n$$

$$g(y) = G'(y) = \frac{d}{dy} G(y) = n\, F(y)^{n-1} f(y)$$

<u>Special case:-</u>

$$X_i \overset{iid}{\sim} U(0,1)$$

$$\text{Beta}(n,1) \equiv g(y) = \begin{cases} n y^{n-1}, & y \in [0,1] \\ 0, & \text{otherwise} \end{cases} \quad \Big| \; G(y) = y^n$$

y/B

H.W

$$E(Y) = \int_{-\infty}^{\infty} y(g(y))\, dy = n \int_0^1 y^n\, dy = \frac{n y^{n+1}}{n+1}\Big|_0^1 = \frac{n}{n+1}$$

$$V(Y) = E(Y^2) - (E(Y))^2 = \frac{n y^{n+2}}{n+2}\Big|_0^1 - \frac{n^2 y^{2n+2}}{(n+1)^2}\Big|_0^1$$

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

$$V(Y) = \frac{n - n^3 - n^2}{(n+1)^2(n+2)}$$

DOMS Page No.
Date / /

Q) Let $X_1, X_2 \dots X_n \overset{iid}{\sim} F(x)$ where F is continuous with $F'(x) = f(x)$
$Z = \min\{X_1, X_2 \dots X_n\}$ find cdf & pdf of Z

→ $H(z) = P(Z \leq z)$

$= P(\min\{X_1, X_2 \dots X_n\} \leq z)$

$= 1 - P(\min\{X_1, X_2 \dots\} > z)$

$= 1 - P(X_1 > z, X_2 > z, \dots, X_n > z)$

$= 1 - \prod_{i=1}^{n} P(X_i > z)$

$= 1 - \prod_{i=1}^{n} (1 - F(z))$

$= 1 - (1 - F(z))^n$

$h(z) = \frac{d}{dz}\left(1 - (1 - F(z))^n\right) = -n(1 - (F(z)))^{n-1}(-f(z))$
$= n(1 - F(z))^{n-1} f(z)$

If $X_i \overset{iid}{\sim} U(0,1)$

Beta(1, n) $H(z) = 1 - (1 - z)^n$ $z \in [0,1]$

$\equiv$ $h(z) = \begin{cases} n(1-z)^{n-1} & z \in [0,1] \\ 0 & \text{otherwise} \end{cases}$

$E(z) = \int_0^1 z n (1-z)^{n-1} dz$

$V(z) = E(z^2) - (E(z))^2$

Q) $X_1 \ldots X_n \overset{iid}{\sim} F(x)$, $F'(x) = f(x)$, F is continuous

$X(r) =$ r^{th} order statistic

$X(1) < X(2) \ldots < X(r) \ldots < X(n) = \max\{X_1 \ldots X_n\}$

→ Let $X(r) = W$

$$G(w) = P(W \leq w)$$
$$= P(X(r) \leq w)$$
$$= P(\text{atleast } r \text{ observations from } X_1, \ldots X_n \text{ are} \leq w)$$
$$= \sum_{k=r}^{n} \binom{n}{k} (F(w))^k (1 - F(w))^{n-k}$$

$$g(w) = G'(w) = \frac{n!\,(F(w))^{r-1} f(w) (1 - F(w))^{n-r}}{(r-1)!\,1!\,(n-r)!}$$

$\xleftarrow{\ (r-1)\ }$ w $w+\Delta w$ $\xleftarrow{\ n-r\ }$

$$G'(w)\,\Delta w = g(w)\,\Delta w \approx \frac{n!\,(F(w))^{r-1} f(w) (\Delta w) (1 - F(w))^{n-r}}{(r-1)!\,1!\,(n-r)!}$$

* Special Case :- $X_1, X_2, \ldots \overset{iid}{\sim} U(0,1)$

$$G(w) = \sum_{k=r}^{n} \binom{n}{r} w^k (1-w)^{n-k}$$

$$g(w) = \begin{cases} \dfrac{n!\, w^{(r-1)} (1-w)^{n-r}}{(r-1)!\,1!\,(n-r)!}, & w \in (0,1) \\ \text{otherwise} \end{cases} = \text{Beta}(r, n-r+1)$$

Q) $E(X(r)) = ?$

$V(X(r)) = ?$

pdf of $W = \dfrac{w^{(k-1)}(1-w)^{n-r}}{B(r, n-k+1)}$ & $G(w) = \int_0^w g(t)\,dt$

$G(p) = P(X(r) \leq p) \qquad p \in (0,1)$

$= \int_0^p \dfrac{t^{r-1}(1-t)^{n-r}}{B(r, n-r+1)}\,dt \rightarrow$ cdf of beta$(r, n-r+1)$ for $y = p$

$= \sum_{k=r}^{n} \binom{n}{k} p^k (1-p)^{n-k} \longrightarrow$ 1 - cdf of Binomial(n,p) for $y = k-1$

Q) $X_i \sim \exp(\lambda_i)$ independent

Find distribution of $Y = \min\{X_1, X_2 \dots X_n\}$

$\rightarrow$ $G(y) = P(Y \leq y)$

$= 1 - P(Y > y)$

$= 1 - P(X_1 > y, X_2 > y \dots X_n > y)$

$= 1 - \prod_{i=1}^{n} P(X_i > y)$

$= 1 - \prod_{i=1}^{n} 1 - P(X_i \leq y)$

$= 1 - \prod_{i=1}^{n} 1 - F_{\lambda_i}(y)$

$= 1 - \prod_{i=1}^{n} 1 - (1 - e^{-\lambda_i y})$

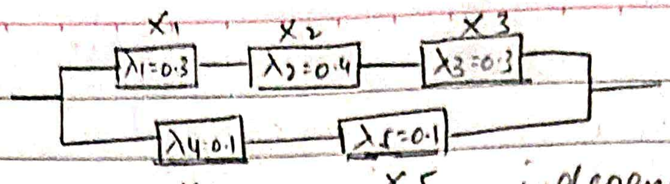
$= 1 - \prod_{i=1}^{n} e^{-\lambda_i y}$

$= 1 - e^{-y(\lambda_1 + \lambda_2 + \dots \lambda_n)}$

$= 1 - e^{-y \sum \lambda_i}$

$Y \sim \exp\left(\sum_{i=1}^{n} \lambda_i\right)$

Q



independent

$X_i \sim \exp(\lambda_i)$

Find the expected life time of the system

$S_1 = \min(X_1, X_2, X_3)$
$S_2 = \min(X_4, X_5)$
$L = \max(S_1, S_2)$

$\rightarrow$ $S_1 \sim \exp(\lambda_1+\lambda_2+\lambda_3)$ | $S_2 \sim \exp(\lambda_4+\lambda_5)$

$\sim \exp(1)$ | $\sim \exp(0.2)$

$L = \max(S_1, S_2)$

$G(l) = P(L \leq l)$

$= P(S_1 \leq l, S_2 \leq l)$

$= P(S_1 \leq l) \cdot P(S_2 \leq l)$

$= (1-e^{-l})(1-e^{-0.2l}) \qquad l \geq 0$

$= 1 - e^{-l} - e^{-0.2l} + e^{-1.2l}$

$$g(l) = \begin{cases} e^{-l} + 0.2e^{-0.2l} - 1.2e^{-1.2l}, & l \in (0,\infty) \\ 0 & \text{otherwise} \end{cases}$$

$$E(L) = \int_0^\infty l\, g(l)\, dl = \frac{1}{1} + \frac{1}{0.2} - \frac{1}{1.2} = 31/6$$

DOMS Page No.
Date / /

Q) $Z \sim N(0,1)$

$\rightarrow$ $M_Z(t) = E(e^{tz})$

$$= \int_{-\infty}^{\infty} \frac{e^{tz} e^{-z^2/2}}{\sqrt{2\pi}} dz = \int_{-\infty}^{\infty} \frac{e^{-1/2(z^2-2tz+t^2)+t^2/2}}{\sqrt{2\pi}} dz$$

$$= e^{t^2/2} \int_{-\infty}^{\infty} \frac{e^{-1/2 \frac{(z-t)^2}{1}}}{1\sqrt{2\pi}} dz$$

$\rightarrow$ pdf of $N(z,1)$

$\boxed{M_Z(t) = e^{t^2/2}}$ $\rightarrow$ for $N(0,1)$

Let

$X = \mu + \sigma Z$

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

$$M_X(t) = E(e^{t(\mu+\sigma z)})$$
$$= e^{t\mu} E(e^{t\sigma z})$$
$$= e^{t\mu} e^{t^2\sigma^2/2}$$

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

$Y \sim \text{lognormal}(\mu, \sigma^2)$

$X \sim \text{Normal}(\mu, \sigma^2)$

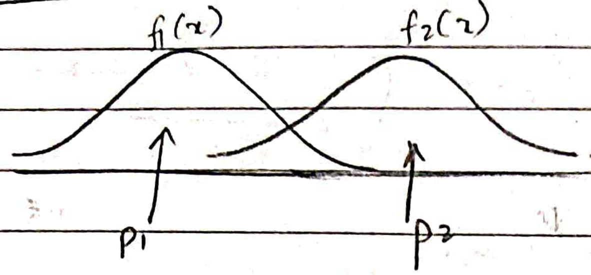
$Y = e^X \sim \text{log Normal}(\mu, \sigma^2)$

$E(Y) = E(e^X)$
$= e^{\mu + \sigma^2/2}$

Note

$\boxed{E(Y) = e^{\mu + \sigma^2/2} \text{ when } Y \sim \text{lognormal}(\mu, \sigma^2)}$

Mixture Distribution



$p_1 + p_2 = 1$

Correct:-
- draw a sample from $f_1(x) \longrightarrow$ say X_1
- " " " " $f_2(x) \longrightarrow$ " X_2
- Put $Z = \begin{cases} X_1 & \text{with probability } p_1 \\ X_2 & \text{" " } p_2 = 1 - p_1 \end{cases}$

$Z = \{ X_1 \; X_1 \; X_2 \; X_1 \; X_2 \; X_1 \; \dots \}$

$Z \sim$ ~~a~~ mixture distribution

Incorrect:- $\| \; Z = p_1 X_1 + p_2 X_2 \; \|$

only true :- $E(Z) = p_1 E(X_1) + p_2 E(X_2)$

pdf $= g(z) = p_1 f_1(x) + p_2 f_2(x)$

① Uniform (a,b) $X \sim U[0,1]$. $b>a$

Tranform $Y = a+(b-a)X$

$$F_Y(y) = \begin{cases} 0 & y<a \\ \frac{y-a}{b-a} & a<y<b \\ 1 & y>b \end{cases} \qquad f_Y(y) = \begin{cases} \left(\frac{1}{b-a}\right); & a\le y<1 \\ 0, & \text{otherwise.} \end{cases}$$

$$\boxed{\begin{aligned} E(Y) &= \frac{a+b}{2} \\ V(Y) &= \frac{(b-a)^2}{12} \end{aligned}}$$

② Exponential distribution.

Exp in Continous $\equiv$ Geo in discrete. (Memoryless)

$X \sim U[0,1]$

Tranform $Y = -\frac{1}{\lambda}\log_e(1-x)$.

$$F_Y(y) \begin{cases} 1-e^{-\lambda y} & y\geq 0 \\ 0 & y<0 \end{cases} \qquad f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y\geq 0 \\ 0; & y<0 \end{cases}$$

$$\boxed{\begin{aligned} E(Y) &= 1/\lambda \\ V(Y) &= 1/\lambda^2 \end{aligned}}$$

③ Gamma distribution

$X_1\ X_2\ X_3 \dots X_\alpha \overset{iid}{\sim} Exp(\lambda)$.

$Y = \sum_{i=1}^{\alpha} X_i$.

$$\boxed{\begin{aligned} E(Y) &= \alpha/\lambda \\ V(Y) &= \alpha/\lambda^2 \end{aligned}}$$

$Y \sim Gamma(\alpha,\lambda)$.

$$f_Y(y) \begin{cases} \lambda^\alpha \frac{e^{-\lambda y} y^{\alpha-1}}{\Gamma\alpha} & y>0,\ \lambda,\alpha>0 \\ 0 & \text{else.} \end{cases}$$

$Y_1 \sim Gamma(\alpha_1,\lambda)$
$Y_2 \sim Gamma(\alpha_2,\lambda)$
$Y_1+Y_2 \sim Gamma(\alpha_1+\alpha_2,\lambda)$.

$Z \sim N(0,1)$
$Y = Z^2$
$Y \sim G(1/2, 1/2)$.

$\Gamma\alpha = (\alpha-1)\Gamma\alpha-1$

$\Gamma n = n-1!$

$B(\alpha_1,\alpha_2) = \frac{\Gamma\alpha_1\,\Gamma\alpha_2}{\Gamma\alpha_1+\alpha_2}$

Theorem.

$$h(y) = f(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right|$$

④ Beta distribution $Z_1, Z_2 \sim Gamma$

$B(\alpha_1,\alpha_2) \left(\frac{\Gamma\alpha_1\,\Gamma\alpha_2}{\Gamma\alpha_1+\alpha_2}\right)$.

$Y = \frac{Z_1}{Z_1+Z_2} \in (0,1)$.

$Y \sim Beta(\alpha_1,\alpha_2)$.

$$\boxed{\begin{aligned} E(Y) &= \frac{\alpha_1}{\alpha_1+\alpha_2} \\ V(Y) &= \frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)} \end{aligned}}$$

$$f_Y(y) \begin{cases} \frac{y^{\alpha_1-1}(1-y)^{\alpha_2-1}}{B(\alpha_1,\alpha_2)} & y\in(1,0) \\ 0 & \text{else.} \end{cases}$$

⑤ Normal distribution

$X \sim N(\mu,\sigma^2)$. $\boxed{f(x) = \left(\frac{e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}\right) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)}$

$Z \sim N(0,1)$. Standard Normal. $\xrightarrow{pdf} f(z) = \left(\frac{e^{-1/2 z^2}}{\sqrt{2\pi}}\right)$ $z \in \mathbb{R}$. $\to \phi(z)$

$\mu+Z \sim N(\mu,1)$

$X = \mu+\sigma Z \sim N(\mu,\sigma^2)$.

⊛ If $X_1, X_2, \dots, X_n \sim N(\mu,\sigma^2)$
their linear Comb $\sim N(\mu,\sigma^2)$

$\therefore a_1X_1+a_2X_2 \sim N(a_1\mu_1+a_2\mu_2,\ a_1^2\sigma_1^2+a_2^2\sigma_2^2)$

$\sum a_iX_i \sim N(\sum a_i\mu_i,\ \sum a_i^2 V(X_i))$

⑥ $Z_1, Z_2, Z_3, Z_4 \dots Z_n \sim N(0,1)$.

$\sum_{i=1}^{n} z_i^2 = \chi_n^2$. Chi Square distr. with n deg of freedom.

$Z^2 = \chi_1^2 = G(1/2, 1/2)$.

$\sum z_i^2 = \chi_n^2 = G(n/2, 1/2)$.

$E(\chi_n^2) = n$

$Var(\chi_n^2) = 2n$.

⑦ T distribution

If $Z \sim N(0,1)$ and $Y \sim \chi_k^2$ and they are independent then $\boxed{T = \frac{Z}{\sqrt{Y/k}} \sim t_k}$ T-distribution. with k-degree of freedom.

⑧ F distribution

If $\left.\begin{aligned} Y_1 &\sim \chi_{k_1}^2 \\ Y_2 &\sim \chi_{k_2}^2 \end{aligned}\right\}$ Indep. $\boxed{F = \frac{Y_1/k_1}{Y_2/k_2} \sim F_{k_1,k_2}}$

1. Uniform $U[0,1]$.

$$f(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \qquad F(x) = \begin{cases} 0 & x < 0 \\ x & 0 < x \le 1 \\ 1 & x > 1 \end{cases}$$

2. discrete Uniform

$P(X=j) = 1/k \quad j = 1, 2, 3, 4, \dots, k.$

3. Bernoulli (p)

$$X \sim B(p) \qquad f(x) = \begin{cases} p^x (1-p)^{1-x} & x \in (0,1) \\ 0 & \text{otherwise} \end{cases} \qquad F(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 < x < 1 \\ p. & x > 1 \end{cases}$$

$$\boxed{\begin{matrix} E(X) = p \\ V(x) = pq \end{matrix}}.$$

4. Binomial

$X_1, X_2, X_3, \dots, X_n \overset{iid}{\sim} \text{Bernoulli}(p).$

$Y = \sum_{i=1}^{n} X_i \quad Y \sim Bin(n,p).$

$$\text{pmf } p_Y(y) = \begin{cases} \binom{n}{y} p^y (1-p)^{n-y}. & y = 0, 1, 2, \dots n \\ 0 & \text{otherwise} \end{cases}$$

$E(ax) = aE(x)$

$V(ax) = a^2 V(x)$

$$\boxed{\begin{matrix} E(Y) = np \\ V(Y) = npq. \end{matrix}}$$

5. Geometric

Y: No of failures preceding to 1st success

$$f(y) = \begin{cases} (1-p)^y p. & y = 0, 1, 2, 3 \dots \infty \\ 0 & \text{otherwise.} \end{cases}$$

$$\boxed{E(Y) = \frac{q}{p}.} \ \boxed{V(Y) = q/p^2}$$

6. Negative Binomial.

W: No. of failures preceding to r^{th} success.

$$f(w) = \begin{cases} \binom{w+r-1}{w} q^w p^r. & w = 0, 1, 2, \dots \\ 0 & \text{otherwise} \end{cases}$$

$$\boxed{\begin{matrix} E(W) = rq/p. \\ V(W) = rq/p^2 \end{matrix}}.$$

7. Hypergeometric distribution.

No. of items $n_1 \to$ first Category

$n_2 \to$ Second Category.

r samples drawn.

RV X: No of sample from 1st category.

$$f \quad P(X=r_1) = \begin{cases} \dfrac{\binom{n_1}{r_1}\binom{n_2}{r_2}}{\binom{n_1+n_2}{r_1+r_2}} \\ 0 & \text{else.} \end{cases} \qquad \begin{matrix} n_1+n_2 = n \\ r_1+r_2 = r. \end{matrix}$$

$$\boxed{\begin{matrix} E(r) = r\dfrac{n_1}{n_1+n_2} \\ V(r) = r\left(\dfrac{n_1}{n}\right)\left(\dfrac{n_2}{n}\right)\left(1 - \dfrac{r-1}{n-1}\right) \end{matrix}}.$$

8. Poisson distribution.

$X \sim \text{poisson}(\lambda) \quad \lambda > 0$

$$f(x) = \begin{cases} e^{-\lambda} \dfrac{\lambda^x}{x!} & x \in \{0, 1, 2, 3, \dots\}. \\ 0 & \text{otherwise.} \end{cases}$$

$$\boxed{E(X-\lambda)^3 = Var(X) = E(X) = \lambda}$$

$M_X(t) = E(e^{tx}) = e^{(e^t - 1)\lambda}$

(*) $n \uparrow \infty \quad p_n \downarrow 0 \quad np_n \to \lambda. \qquad Z_n \sim bin(n, p_n)$

$Z_n \sim \text{poisson}(\lambda).$

? || Poisson process :- $P(N(s+t) - N(s) = k) = e^{-\lambda t} \dfrac{(\lambda t)^k}{k!}.$

$$\boxed{\begin{matrix} E(N(t)) = \lambda t \\ V(N(t)) = \lambda t \end{matrix}}.$$

Joint Distributions

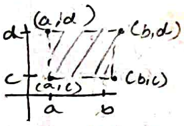
① (X,Y) pair of R.V. in (Ω, A, P)

Joint cdf $F(x,y) = P(X \le x, Y \le y)$

$= P(\{\omega \mid X(\omega) \le x \;\&\; Y(\omega) \le y\})$

② Properties of joint cdf. $F(x,y) = P(X \le x, Y \le y)$

- $\lim_{\substack{x \downarrow -\infty \\ y \downarrow -\infty}} F(x,y) = 0$
- $\lim_{\substack{x \uparrow \infty \\ y \uparrow \infty}} F(x,y) = 1$
- $\lim_{x \uparrow \infty} F(x,y) = F_Y(y)$ Marginal distribution of Y.
- $\lim_{y \uparrow \infty} F(x,y) = F_X(x)$ Marginal distribution of X.



- $P(a \le X \le b, \; c \le Y \le d)$

$= F(b,d) + F(a,c) - F(a,d) - F(b,c)$.

if X & Y are independent

$\Rightarrow F_X(b)F_Y(d) + F_X(a)F_Y(c) - F_X(a)F_Y(d) - F_X(b)F_Y(c)$

$\Rightarrow [F_X(b) - F_X(a)][F_Y(d) - F_Y(c)]$.

\# Marginals $\not\rightarrow$ exact distribution

exact distribution $\longrightarrow$ Marginals

③ joint p.d.f

discrete

$\sum_x \sum_y f(x,y) = 1$

Marginal pmf of X. $\sum_y f(x,y) = f_X(x)$.

Marginal pmf of Y $\sum_x f(x,y) = f_Y(y)$.

X, Y independent

$f(x,y) = P(X=x, Y=y) \Rightarrow P(X=x)P(Y=y)$

$\Rightarrow f_X(x) f_Y(y)$

Continuous

$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial^2 F(x,y)}{\partial y \partial x}$.

$\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} f(x,y) = 1$.

$\int_y f(x,y) = f_X(x)$.

$\int_x f(x,y) = f_Y(y)$

④ Conditional pdf / pmf.

$$\boxed{f_{Y|X}(y \mid X=x) = \frac{f(x,y)}{f_X(x)}}$$

pmf → actual probability

pdf → only density

★ $E(X) = \int_x \int_y x \, f(x,y) \, dy \, dx$

$= \int_x x f_X(x) dx$

⑤ Conditional Expectation (Regression of Y on X)

$E(Y \mid X=x) = \int_{y=-\infty}^{\infty} y \, f_{Y|X}(y \mid X=x) \, dy$

$E(Y \mid X=x) =$
- $\int_{-\infty}^{\infty} y \frac{f(x,y)}{f_X(x)} dy$ Continuous
- $\sum_y y \frac{f(x,y)}{f_X(x)}$. Discrete

Laws

$E(X+Y) = E(X) + E(Y)$ → No need to be indepen ~~X & Y are independent~~.

X, Y are independent ← $E(XY) = E(X)E(Y)$

$E(\alpha X) = \alpha E(X)$

Covariance :- $Cov(X,Y) = E(XY) - E(X)E(Y)$.

$= E(X - E(Y))(Y - E(X))$.

if X, Y independent $Cov(X,Y) = 0$.

$Cov(X,X) = E(X - E(X))^2 = Var(X)$.

But $Cov(X,Y) = 0 \not\Rightarrow$ X, Y are independent.

$$\boxed{\begin{array}{l} Y = a + bX \\ Cov(X,Y) = b^2 V(X) \\ Corr(X,Y) = \frac{b}{|b|} \\ M_Y(t) = e^{ta} M_X(bt) \end{array}}$$

$Var(a + X) = Var(X)$

$Var(a + bX) = b^2 Var(X)$

$V(aX + bY)$

$= a^2 V(X) + b^2 V(Y)$

$+ 2ab \, Cov(X,Y)$

$$\boxed{E(Y) = E_X E_{Y|X}(Y \mid X=x).}$$

$$\boxed{\begin{array}{c} Var(Y) = E_X Var_{Y|X}(Y \mid x) + Var_X E_{Y|X}(Y \mid x) \\ EV + VE. \text{ Rule} \end{array}}$$

Co-relation :- $Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X) Var(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X) Var(Y)}}$.

measures linear dependence.

(X,Y) independent $\Rightarrow Corr(X,Y) = 0$

$Corr(X,Y) = 0 \not\Rightarrow (X,Y)$ independent.

$|Corr(X,Y)| \le 1$

$Y = a + bX$ $Corr(X,Y) = \frac{b}{|b|} = +1 / -1$.

Scanned by TapScanner

Mgf properties :-
- $X \rightarrow mgf = M_X(t)$, $Y = a + bX$, $Y \rightarrow mgf = e^{at} M_X(bt)$
- $Z = X + Y$, X & Y are independent, $Z \rightarrow mgf = M_X(t).M_Y(t)$

⑥ **Bivariate Normal**

$$(X,Y) \sim N(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$$

corr(x, y)

$\rho = 0$ independence

ρ = correlation

joint pdf $$f(x,y) = \frac{e^{-\frac{1}{2}\left(\frac{1}{1-\rho^2}\right)\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}.$$

if x, y independent

$\rho = 0$

$$f(x,y) = f_X(x)\, f_Y(y)$$

$$f(x,y) = \frac{e^{-\frac{1}{2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}}{\left(\sqrt{2\pi}\,\sigma_X\right)\left(\sqrt{2\pi}\,\sigma_Y\right)}.$$

Condition for independence of Bivariate Normal.

When $(X,Y) \sim N(\mu_X\ \mu_Y\ \sigma_X^2\ \sigma_Y^2\ \rho)$

$\hookrightarrow$ $$(Y|X) \sim N\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X),\ (1-\rho^2)\sigma_Y^2\right)$$

linear regression $E(Y|X) = \mu_{Y|x}$; $Var(Y|X)$

pdf $$f_{Y|x}(y|x) = \frac{e^{-1/2\,\frac{1}{(1-\rho^2)\sigma_Y^2}\left(y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)^2}}{\sqrt{2\pi}\,\sigma_Y\sqrt{1-\rho^2}}$$

☆ even though
X, Y individually follow Normal dist.
(X,Y) jointly may not follow Bivariate Normal.

$X_1, X_2, \ldots X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$\rightarrow \bar{X} \sim N(\mu, \sigma^2/n)$
$\rightarrow S^2 = \sum_{i=1}^{n}(X_i - \bar{x})^2 \sim \sigma^2 \chi^2_{n-1}$
$\Big] \rightarrow$ independent

Weak law of Large Numbers :- Sample Mean $\rightarrow$ population mean as sample size increases, given variance is finite

Law of Large Numbers :- $\{X_n\}$ is said to converge to X if for all $w \in \Omega$, $X_n(w) - X(w) \rightarrow 0$ as $n \rightarrow \infty$.

Convergence :- If $\lim_{n\to\infty} F_{Y_n}(a) = F_Y(a)$ or $\lim_{n\to\infty} M_{X_n}(t) = M_Y(t)$
then $\{Y_n\} \rightarrow Y$

TRANSFORMATION

$$(X,Y) \longrightarrow (U,V) \equiv (U(x,y), V(x,y))$$

$$g(u,v) = f(x(u,v), y(u,v)) \left|\left|\frac{\partial(x,y)}{\partial(u,v)}\right|\right|$$

Jacobian $$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}.$$

Cauchy Distribution

$X, Y \overset{iid}{\sim} N(0,1) \Rightarrow \frac{X}{Y} \sim$ Cauchy (0,1)
$\frac{X}{|Y|} \sim$ Cauchy (0,1).

$X, Y \overset{iid}{\sim} N(0,\sigma^2) \rightarrow \frac{X}{Y} \sim$ Cauchy (0,1)
$\frac{X}{|Y|} \sim$ Cauchy (0,1).

$\rightarrow$ No moments for Cauchy

NOTE

t-distribution

$$\frac{X}{|Y|} = \frac{X}{\sqrt{Y^2/1}} = t_1$$

$t_n \rightarrow N(0,1)$ as $n \rightarrow \infty$

for a sequence of RV $\{X_n\}$.

L.L.N $$\lim_{n\uparrow\infty} P(\{w \mid |X_n(w) - X(w)| < \epsilon\}) = 1.$$ ($\epsilon > 0$).

Weak L.L.N $X_1, X_2, X_3, \ldots, X_n \overset{iid}{\sim}$ $E(X_i) = \mu$ $V(X_i) = \sigma^2$.
$\bar{x} = \left(\frac{\sum x_i}{n}\right)$.

$$\lim_{n\uparrow\infty} P(|\bar{x}-\mu| < \epsilon) = 1. \quad \forall\ \epsilon > 0$$

CLT. (Central Limit Theorem)

$X_1, X_2 \ldots X_n \xrightarrow{iid}$ RV , $E(X_i) = \mu$, $V(X_i) = \sigma^2 < \infty$

$$\boxed{T_n} = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \sqrt{n}\,\frac{(\bar{x}-\mu)}{\sigma} = \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{\bar{x} - E(\bar{x})}{\sqrt{V\bar{x}}}$$

$$S_n = \sum_{i=1}^{n} X_i = n\bar{X}$$

$$\lim_{n\to\infty} P(T_n \leq t) = \Phi(t) = \int_{-\infty}^{t} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx$$

$$\boxed{T_n \longrightarrow N(0,1)}$$ as $n \uparrow \infty$

$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 \ldots$

$= 1 + {}^nC_1 x + {}^nC_2 x^2 + {}^nC_3 x^3 + \ldots$

Probability & Statistics

- Permutation :- ${}^nP_r = \frac{n!}{(n-r)!}$
- combination :- ${}^nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$

Sum of geometric progression $= \frac{a(r^n - 1)}{r-1}$

∞ series $= \frac{a}{r-1}$

$\Rightarrow$ No. of ways

1) Maxwell-Boltzmann statistics $\rightarrow r^n$
   n distinguishable balls in r distinguishable boxes
2) Fermi-Dirac statistics $\rightarrow {}^nC_r$ or $\binom{n}{r}$
   choosing r balls from n distinguishable balls
3) Bose-Einstien statistics $\rightarrow \binom{n+r-1}{r-1}$
   n identical balls in r distinguishable boxes
4) $\binom{n-1}{r-1}$
   same as 3, but no empty boxes
5) $a_1 + a_2 + \ldots a_r = n$, $n, r \in \mathbb{N}$, $a_i \in \mathbb{N} \cup \{0\}$
   Probability of solution being only +ve integers
   $= \frac{\binom{n-1}{r-1}}{\binom{n+r-1}{r-1}}$

Probability :- $P(A) = \frac{|A|}{|\Omega|}$ $\rightarrow$ no. of times event occurs / $\rightarrow$ no. of elements in sample space

(A: event)

Acc. to frequency def^n :- $P(A) = \lim_{n \to \infty} \frac{|A_n|}{|\Omega_n|}$ (if lim exists)

$\lim_{n \to \infty} A_n = A$

$\lim_{n \to \infty} \Omega_n = \Omega$

$\rightarrow A_i \subseteq \Omega_i$ always

Algebra ($\mathcal{A}$) :- collection of subsets of Ω

- $\Omega \in \mathcal{A}$, • $A \subseteq \Omega$ & $A \in \mathcal{A} \rightarrow A^c \in \mathcal{A}$ • $A, B \subseteq \Omega$ & $A, B \in \mathcal{A} \rightarrow A \cup B \in \mathcal{A}$

σ-Algebra :- $\mathcal{A}$ is σ-Algebra if $\{A_i\} \subseteq \Omega$, $\{A_i\} \in \mathcal{A} \rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

Total probability :- $P(A) = \sum_n P(A \cap B_n)$

$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots$ | $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \cdots$

Axiomatic Defⁿ (Kolmogorov)
- $\mathcal{A}$ has to be σ-algebra
- probability P, $P: \mathcal{A} \to [0, 1]$ satisfies
  - $P(\Omega) = 1$
  - $P(A) \geq 0$ for any $A \in \mathcal{A}$
  - $\{A_i\} \in \mathcal{A}$ implies $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ if $A_i \cap A_j = \phi \ \forall \ i \neq j$

#Note :- If $B \subseteq \Omega$ but $B \notin \mathcal{A}$, $P(B)$ is not defined

- $P(A) = 0$ doesn't mean A never occurs
- $P(A) = 1$ doesn't mean A always occurs

- Probability space :- $(\Omega, \mathcal{A}, P)$
- Event :- $A \subseteq \Omega$, $A \in \mathcal{A}$ → A is event
- Conditional probability :- $P(A|B) = P(A \cap B) / P(B)$ (Probability of A given B)
- Independent events :- If $P(A|B) = P(A)$ or $P(A \cap B) = P(A).P(B)$ then A, B are independent of each other
- Pairwise independence :- If $P(A_i \cap A_j) = P(A_i).P(A_j) \ \forall \ i \neq j$
- Mutual " :- If $P(\cap_{i=1}^{k} A_i) = \pi_{i=1}^{k} P(A_i) \ \forall$ i distinct.
- Mutual exclusive :- $A_i \cap A_j = \phi \quad \forall \ i \neq j$
- " exhaustive :- $\cup_{i=1}^{\infty} A_i = \Omega$

Note :- Pairwise independence $\neq$ mutual independence

- Partition :- Both mutually Exclusive & exhaustive

Results
- → $P(A^c) = 1 - P(A)$ → $P(\phi) = 0$ → If $A \subseteq B$ then $P(A) \leq P(B)$
- → $1 - P(\cup_{i=1}^{\infty} A_i) = P(\cap_{i=1}^{\infty} A^c_i)$ → $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- → $P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$ → $P(\cup_{i=1}^{\infty} A_i) = \sum_{k=1}^{\infty} (-1)^{k-1} S_k$

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

  $\binom{n}{k}$ terms

Bayes Theorem → $A_1, A_2 \ldots A_k$ are partitions → $P(A_i) > 0 \ \forall \ i$
→ $B \subseteq \Omega$ s.t $P(B) > 0$

$$P(A_i|B) = \frac{P(B|A_i) P(A_i)}{\sum_{j=1}^{k} P(B|A_j) P(A_j)}$$

Total probability theorem :- $P(D) = P(D|A) P(A) + P(D|A^c) P(A^c)$

Table of Common Distributions

taken from *Statistical Inference* by Casella and Berger

Discrete Distrbutions

distribution	pmf	mean	variance	mgf/moment
Bernoulli(p)	$p^x(1-p)^{1-x};\ x=0,1;\ p\in(0,1)$	p	$p(1-p)$	$(1-p)+pe^t$
Beta-binomial(n,α,β)	$\binom{n}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(\alpha+\beta+n)}$	$\frac{n\alpha}{\alpha+\beta}$	$\frac{n\alpha\beta}{(\alpha+\beta)^2}$	

Notes: If $X|P$ is binomial (n,P) and P is beta(α,β), then X is beta-binomial(n,α,β).

distribution	pmf	mean	variance	mgf/moment
Binomial(n,p)	$\binom{n}{x}p^x(1-p)^{n-x};\ x=1,\ldots,n$	np	$np(1-p)$	$[(1-p)+pe^t]^n$
Discrete Uniform(N)	$\frac{1}{N};\ x=1,\ldots,N$	$\frac{N+1}{2}$	$\frac{(N+1)(N-1)}{12}$	$\frac{1}{N}\sum_{i=1}^N e^{it}$
Geometric(p)	$p(1-p)^{x-1};\ p\in(0,1)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$

Note: $Y=X-1$ is negative binomial$(1,p)$. The distribution is memoryless: $P(X>s|X>t)=P(X>s-t)$.

distribution	pmf	mean	variance	mgf/moment
Hypergeometric(N,M,K)	$\frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}};\ x=1,\ldots,K$ $M-(N-K)\le x\le M;\ N,M,K>0$	$\frac{KM}{N}$	$\frac{KM}{N}\frac{(N-M)(N-k)}{N(N-1)}$	?
Negative Binomial(r,p)	$\binom{r+x-1}{x}p^r(1-p)^x;\ p\in(0,1)$ $\binom{y-1}{r-1}p^r(1-p)^{y-r};\ Y=X+r$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
Poisson(λ)	$\frac{e^{-\lambda}\lambda^x}{x!};\ \lambda\ge 0$	λ	λ	$e^{\lambda(e^t-1)}$

Notes: If Y is gamma(α,β), X is Poisson$(\frac{x}{\beta})$, and α is an integer, then $P(X\ge\alpha)=P(Y\le y)$. → additive

↳ $\sum_{i=1}^{n}$ pois(1) = pois(n)

?|| Poisson process

Continuous Distributions

distribution	pdf	mean	variance	mgf/moment
Beta(α,β)	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1};\ x\in(0,1),\ \alpha,\beta>0$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$
Cauchy(θ,σ)	$\frac{1}{\pi\sigma}\frac{1}{1+(\frac{x-\theta}{\sigma})^2};\ \sigma>0$	does not exist	does not exist	does not exist
Notes: Special case of Students's t with 1 degree of freedom. Also, if X,Y are iid $N(0,\sigma^2)$, $\frac{X}{Y}$ is Cauchy $\frac{Y}{\lvert Y\rvert}\sim C(0,1)$, $\frac{X}{Y}\sim C(0,1)$, $t_1\sim C(0,1)$				
χ^2_p	$\frac{1}{\Gamma(\frac{p}{2})2^{\frac{p}{2}}}x^{\frac{p}{2}-1}e^{-\frac{x}{2}};\ x>0,\ p\in N$	p	$2p$	$\left(\frac{1}{1-2t}\right)^{\frac{p}{2}},\ t<\frac{1}{2}$
Notes: Gamma$(\frac{p}{2},2)$. $\theta=1/\lambda$				
Double Exponential(μ,σ)	$\frac{1}{2\sigma}e^{-\frac{\lvert x-\mu\rvert}{\sigma}};\ \sigma>0$	μ	$2\sigma^2$	$\frac{e^{\mu t}}{1-(\sigma t)^2}$
Exponential(θ) $\theta=1/\lambda$	$\frac{1}{\theta}e^{-\frac{x}{\theta}};\ x\ge 0,\ \theta>0$	θ	θ^2	$\frac{1}{1-\theta t},\ t<\frac{1}{\theta}$
Notes: Gamma$(1,\theta)$. exponential → Memoryless. $Y=X^{\frac{1}{\gamma}}$ is *Weibull*. $Y=\sqrt{\frac{2X}{\beta}}$ is *Rayleigh*. $Y=\alpha-\gamma\log\frac{X}{\beta}$ is *Gumbel*.				
F_{ν_1,ν_2}	$\frac{\Gamma(\frac{\nu_1+\nu_2}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}\frac{x^{\frac{\nu_1-2}{2}}}{\left(1+(\frac{\nu_1}{\nu_2})x\right)^{\frac{\nu_1+\nu_2}{2}}};\ x>0$	$\frac{\nu_2}{\nu_2-2},\ \nu_2>2$	$2(\frac{\nu_2}{\nu_2-2})^2\frac{\nu_1+\nu_2-2}{\nu_1(\nu_2-4)},\ \nu_2>4$	$EX^n=\frac{\Gamma(\frac{\nu_1+2n}{2})\Gamma(\frac{\nu_2-2n}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}\left(\frac{\nu_2}{\nu_1}\right)^n,\ n<\frac{\nu_2}{2}$
Notes: $F_{\nu_1,\nu_2}=\frac{\chi^2_{\nu_1}/\nu_1}{\chi^2_{\nu_2}/\nu_2}$, where the χ^2s are independent. $F_{1,\nu}=t^2_\nu$. $Y_1\sim\chi^2_{k_1}$, $Y_2\sim\chi^2_{k_2}$ $F_{k_1,k_2}\sim\frac{Y_1/k_1}{Y_2/k_2}$				
Gamma(α,β) $\beta=1/\lambda$	$\frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-\frac{x}{\beta}};\ x>0,\ \alpha,\beta>0$	$\alpha\beta$	$\alpha\beta^2$	$\left(\frac{1}{1-\beta t}\right)^\alpha,\ t<\frac{1}{\beta}$ | Additive Gamma(α_1,λ) + Gamma(α_2,λ) = Gamma$(\alpha_1+\alpha_2,\lambda)$
Notes: Some special cases are exponential ($\alpha=1$) and χ^2 ($\alpha=\frac{p}{2},\beta=2$). If $\alpha=\frac{3}{2}$, $Y=\sqrt{\frac{X}{\beta}}$ is *Maxwell*. $Y=\frac{1}{X}$ is *inverted gamma*.				
Logistic(μ,β)	$\frac{1}{\beta}\frac{e^{-\frac{x-\mu}{\beta}}}{\left[1+e^{-\frac{x-\mu}{\beta}}\right]^2};\ \beta>0$	μ	$\frac{\pi^2\beta^2}{3}$	$e^{\mu t}\Gamma(1+\beta t),\ \lvert t\rvert<\frac{1}{\beta}$
Notes: The cdf is $F(x\mid\mu,\beta)=\frac{1}{1+e^{-\frac{x-\mu}{\beta}}}$.				
Lognormal(μ,σ^2) $=e^{N(\mu,\sigma^2)}$	$\frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x}e^{-\frac{(\log x-\mu)^2}{2\sigma^2}};\ x>0,\sigma>0$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2(\mu+\sigma^2)}-e^{2\mu+\sigma^2}$	$EX^n=e^{n\mu+\frac{n^2\sigma^2}{2}}$
Normal(μ,σ^2)	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}};\ \sigma>0$	μ	σ^2	$e^{\mu t+\frac{\sigma^2t^2}{2}}$
Pareto(α,β) ↳ Income dist.	$\frac{\beta\alpha^\beta}{x^{\beta+1}};\ x>\alpha,\ \alpha,\beta>0$	$\frac{\beta\alpha}{\beta-1},\ \beta>1$	$\frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)},\ \beta>2$	does not exist
t_ν	$\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\frac{1}{\sqrt{\nu\pi}}\frac{1}{(1+\frac{x^2}{\nu})^{\frac{\nu+1}{2}}}$	$0,\ \nu>1$	$\frac{\nu}{\nu-2},\nu>2$	$EX^n=\frac{\Gamma(\frac{\nu+1}{2})\Gamma(\nu-\frac{n}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})}\nu^{\frac{n}{2}}$, n even
Notes: $t^2_\nu=F_{1,\nu}$. $Z\sim N(0,1)$, $Y\sim\chi^2_k$, $t_k\sim Z/\sqrt{Y/k}$ | $X,Y\overset{iid}{\sim}N(0,1)\to X/\lvert Y\rvert=X/\sqrt{Y^2}=t_1$ | $t_n\to N(0,1)$ as $n\to\infty$				
Uniform(a,b)	$\frac{1}{b-a},\ a\le x\le b$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$
Notes: If $a=0,\ b=1$, this is special case of beta ($\alpha=\beta=1$).				
Weibull(γ,β)	$\frac{\gamma}{\beta}x^{\gamma-1}e^{-\frac{x^\gamma}{\beta}};\ x>0,\ \gamma,\beta>0$	$\beta^{\frac{1}{\gamma}}\Gamma(1+\frac{1}{\gamma})$	$\beta^{\frac{2}{\gamma}}\left[\Gamma(1+\frac{2}{\gamma})-\Gamma^2(1+\frac{1}{\gamma})\right]$	$EX^n=\beta^{\frac{n}{\gamma}}\Gamma(1+\frac{n}{\gamma})$
Notes: The mgf only exists for $\gamma\ge 1$. • waiting time distribution $X\sim$ Weibull if $\left(\frac{X}{\lambda}\right)^k\sim exp(1)$				

Theorem → $h(y)=f(g^{-1}(y))\left\lvert\frac{d}{dy}g^{-1}(y)\right\rvert$

old pdf

old variable in terms of new