Advanced Statistics

→ Experiment

→ Deterministic Experiment
$y = f(x)$

→ Random Experiment

→ Sample space $S = \{ \ldots \}$
↳ discrete (countables (finite or infinite))
↳ continous (ranges, interval)

Discrete (cointoss, dice rooll)
Continous (temperature, height measu

(5) Event
(subset of a sample space given a constraint / desired output)
"An event where all outputs are even in dice rooll"

(6) Probability
(quantification the outcomes of a random experiment)
"Chance" is qualitative "

(7) Axioms of Probability (general rules)

$P(A)$ probability of occurance of A

"Relative frequency of experimental event"

1 * $0 \leq P(A) \leq 1$

2 * $P(S) = 1$ (sum of all prob is 1)

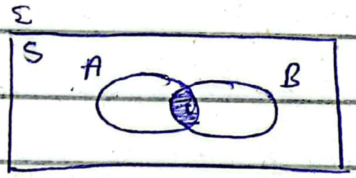
3 * $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A_1 \cup A_2 \cdots \cup A_{n+1}) = P(A_1) + P(A_2) \cdots + P(A_n)$

$P(A) = \dfrac{n(A)}{n(S)}$ ∵ if events are equally likely and sample space discret

$n(A)$ = number of elements in A (event)

$n(S)$ number of elements in sample space

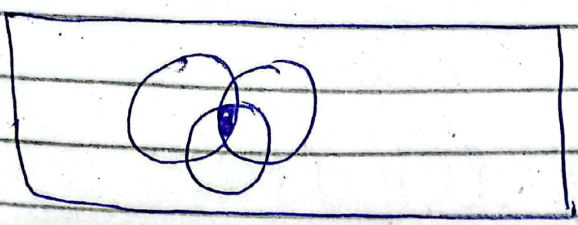
* 

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

→ We have added the shaded region twice so subtract it once.

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

$P(A \cup B \cup C \cup D) =$ something as above.



Complementary

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

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

Additive Rules:-

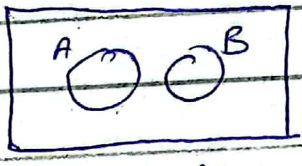
$P(A)$; $P(B)$ known

Conditional Probability

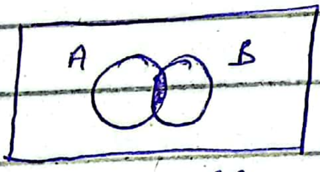
$P(B|A) = ?$

→ conditional prob of B over A

→ prob of B given A has occured already



mutually exclusive



mutually inclusive

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

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

Joint probability of events:

Baye's Theorum :-

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

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

$P(A) =$

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

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

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

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

if P(

Dependent :-

Occurance of A effects occurance of B.

Product Rule

if A and B are independent
then

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

and

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

$$P\left(\bigcap_{i=1}^{N} E_i\right) = P(E_1) \cdots P(E_n)$$

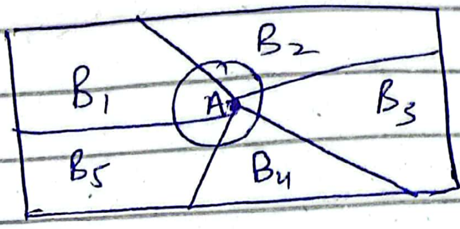
$$= \prod_{i=1}^{N} P(E_i)$$

if $E_1\ E_2 \ldots E_n$ are independent
but mutually inclusive.

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

$$= \sum_{i=1}^{N} P(A_i)$$

if they are mutually exclusive)



→ dividind a sample space in N ~~non~~ mutually exclusive (disjoint partition)

$$P(A) = P(A \cap B_1) \cup P(A \cap B_2) \cup P(A \cap B_3) \cup --- P(A \cap B_n)$$

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + --- + P(A \cap B_n)$$

Because A & B are exclusive within themselves.

$$= \sum_{i=1}^{N} (P(A \cap B_i))$$

$$\boxed{P(A) = \sum_{i=1}^{N} P(A|B_i) P(B_i)}$$

Total probability

$$\boxed{P(B_i|A) = P(A \cap B_i) \div P(A)}$$

$$\boxed{P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{i=1}^{N} P(A|B_i) P(B_i)}}$$

Baye's Theorum

Experiment :-

→ a process that generates some data
→ experiment is performed under certain ~~closed~~ conditions that are constant.
→ environment does not change for an experiment
→ If the outcome does not change, we can model the outcome mathematically

$y = f(x)$

output function input

Deterministic Experiment :-

→

→ If condition are fixed, and then the outcomes are variable then we call such an experiment

Random Experiment :-

→ all natural processes / experiments are random

→ if the experiment generates discrete no. then the experiment is discrete such as tossing a coin or rolling a dices.
→ If the experiment is continous then it will generate ranges and interval as output. Such as measuring height / temp.

Sample Space :-

→ Contains all possible out comes of an experiment
→ It is a set $S = \{ \ldots \}$
→ It can be an interval or a discrete value
→ we can have discrete sample space infinite or finite which is countable
→ we can have continous outputs in samplespace that are finite / infinite.

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

~~Experiment :-~~ Event :-

→ A subset of set / sample space is called an event based on certain checks to derive success.

Probability :-

quantification of the outcomes of a random experiment.
"Chance" is qualitative so we quantify it.

$P(A) \Rightarrow$ probability of A

Axioms ?

Axioms of Probability

① $0 \leq P(A) \leq 1$

② $\sum_{i=1}^{N} P(A_i) = 1$, $P(S) = 1$

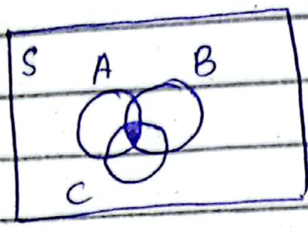
③ If A and B are mutually exclusive then $P(A \cup B) = P(A) + P(B)$

If A and B are mutually inclusive then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

∴ If $A_1, A_2, A_3 \text{---} A_n$ are mutually exclusive then

$$P(A_1 \cup A_2 \text{---} \cup A_N) = P(A_1) + P(A_2) \text{---} + P(A_n)$$

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



$$(A \cup B \cup C) = (A \cup B) \cup (A \cup C) \cup (B \cup C)$$

$$P(A \cup B \cup C) = P(A \cup B) + P(A \cup C) + P(B \cup C) +$$
$$+ P(A \cap B \cap C) - P(A \cap B) - P(A \cap C) - P(B \cap C)$$
$$= P\{\underbrace{A}_{E_1} \cup \underbrace{(B \cup C)}_{E_2}$$
$$= P(E_1 \cup E_2)$$
$$= P(E_1) + P(E_2) - P(E_1 \cap E_2)$$
$$= P(A) + P(B \cup C) - P\{A \cap (B \cup C)\}$$
$$= P(A) + P(B \cup C) - \cancel{P(A \cap B \cup}$$
$$- P\{(A \cap B) \cup (A \cap C)\}$$
$$= P(A) + P(B \cup C) - P\{A \cap B \cup A \cap C$$
$$= P(A) + P(B \cup C) - P(A \cap B) - P(A \cap C$$
$$= P(A) + P(B) + P(C) - P(A \cap B) + P(A \cap C)$$
$$= P(A) + P(B \cup C) - P(A \cap B) + P(C) -$$
$$P[(A \cup B) \cap C]$$

$$= P(A) + P$$

Conditional Probability

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

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

$\Rightarrow$ Independant events :-

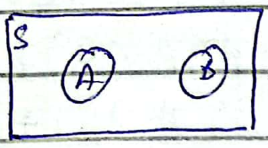
If occurance of A does not impact B, then ~~A~~ and B are not dependant i.e independent

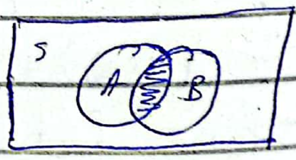
* Two mutually exclusive events are not always independent.

Independent is B of A then

* $P(B|A) = P(B)$

if $P(A \cap B) = P(A)\,P(B)$
then A and B are independent





dependent + m. exclusive

? + m. inclusive

↓

→ i don't know if dependent

→ dependent if $P(A \cap B) \neq P(A)\,P(B)$

Q = If two events occur at the same time, are they always independent ?

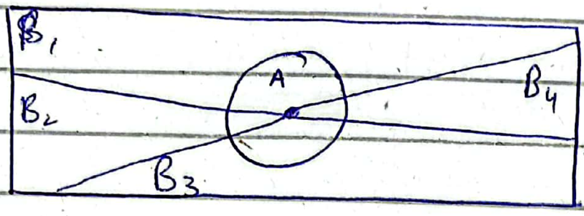
Generalizing Product Rule :-

$P(E_1 \cap E_2 \cap E_3 \cdots \cap E_n) =$ iff E_s are independent

$P(E_1)P(E_2)P(E_3) \cdots P(E_n)$

$$P\left(\bigcap_{i=1}^{N} E_i\right) = \prod_{i=1}^{N} P(E_i)$$

Baye's Theorum :-



$P(B_i)$ $\quad i = 1, 2, 3 \ldots N$

$P(A) = ?$

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

$(A \cap B_3) \cup (A \cap B_4) \cdots$

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

$P(A \cap B_3) \ldots + P(A \cap B_n)$

given that $B_1 B_2 B_3$ are exclusive

$P[(A \cap B_1) \cup (A \cap B_2) \cdots (A \cap B_n)]$

$$\Rightarrow \sum_{i=1}^{N} P(A \cap B_i) = P(A)$$

$$= P(A) = \sum_{i=1}^{N} P(A|B_i)\, P(B_i)$$

$\boxed{\text{Total Probability Law}}$

$$P(B_i|A) = ?$$

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

$$= \frac{P(A|B_i)\, P(B_i)}{P(A)}$$

$$= \frac{P(A|B_i)\, P(B_i)}{\sum_{i=1}^{N} P(A|B_i)\, P(B_i)} \quad \begin{matrix} \} \text{ out of} \\ \} \text{ sum} \end{matrix}$$

Advanced Statistics

Random Variable :-

define by :- X

defined over a whole sample space on a $\mathbb{R}^{+}$ number line.

(i) R.V → X
(ii) Range of X → $S_x = \{ \quad \}$

$S = \{HH, HT, TH, TT\}$
$S_X = \{0, 1, 2\}$
X = shows no. of heads when a pair of coins is tossed.

Y = shows no. of ..

(iii) Probability assignments to outcomes of Random Variable.

"Probability Mass Function"

PMF → $P_X(x)$

$$p_x(x) = P_r[X = x]$$

Probability that X takes value of x

PMF is used to assign masses of probability to random variable values

(ii) $0 \le p_x(x) \le 1$

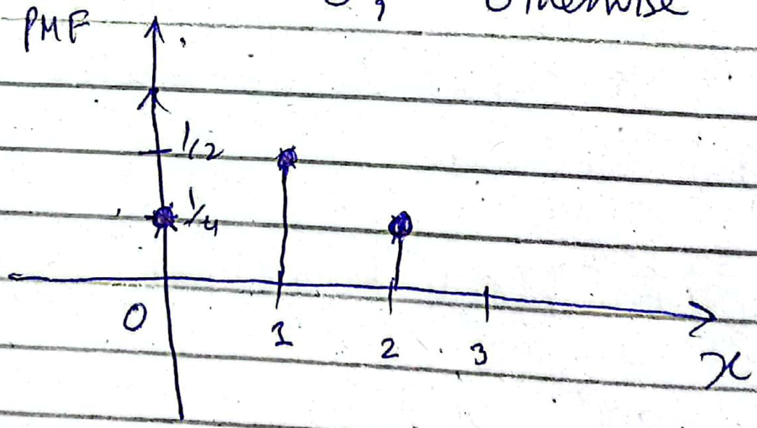
(III) $\sum_{x} p_x(x) = 1$

PMF of $\{0, 1, 2\}$

X = no. of heads when two coins are tossed

$S_x = \{0, 1, 2\}$ $\quad S = \{HH, TT, HT, TH\}$

$$PMF = \begin{cases} 1/4, & x = 0 \\ 1/2, & x = 1 \\ 0, & \text{otherwise} \end{cases}$$



Expected Value

"a prediction of the trend"

"also called the average of or weighted average"

→ Not always an average.

$$E[X] = \sum_{x} x\, p_x(x)$$

∴ statistical inference ↑

$$\mu = \frac{1}{2} \times 1 + \frac{1}{4} \times 2 = \frac{1}{2} + \frac{1}{2} = 1$$

$S_x = \{0, 1, 2\}$ $\quad S \subseteq \{HH, TT, HT, TH\}$

getting 1 head has highest probabili
if two coins are tossed.

$$\frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6_2} \times \cancel{3} + \frac{1}{6} \times 4 + \frac{1}{6} \times 5 + \frac{1}{6} \times 6$$

$$= \frac{1}{6} + \frac{1}{3} + \frac{\cancel{1}}{\cancel{2}} + \frac{\cancel{4}}{6} + \frac{5}{6} + 1$$

$$= \frac{1}{6} + \frac{1}{3} + 0.5 + \frac{1}{2} + 6$$

$= 3.5$ $\quad S_x = \{1, 2, 3, 4, 5, 6\}$

* When the probability of each values is same then $E[x]$ is not useful. (uniform Random variable)

* If PMF is constant i.e in case of tossing a fair die.

~~Random~~ VARIANCE :-
Expected variation around its average in tossing two coins.

$$Var[x] = E\left[\left(X - E[x]\right)^2\right]$$

$$= E[x^2] - (E[x]^2)$$

$$= \sum_x x^2 P[x] - (E[x])^2$$

* variance can never be negative
*

$$STD(x) = \sqrt{var(x)}$$

$$\sigma = \sqrt{\sigma^2}$$

Bernoulli's Trials.

→ We divide a sample space in 2
→ a success is 1, and failure is 0
→ This is a bernoulli trial.
→ first trial is not impacting the other trial.

→ Binomial Random Variable

$S_x = \{0, 1, 2, 3 \ldots n\}$

$P_x(k) = {}^{n}C_{k} \, p^{k} \, (1-p)^{n-k}$

$E(x) = np$ $\quad$ $var(x) = np(1-p)$

Random Variable X :-

→ keep performing independent trials until a success comes.

Geometric Random Variable

$S_x = \{1, 2, 3 \ldots\}$

$P_x(n) = p(1-p)^{n-1}$

$E(X) = \frac{1}{p}$; $Var(x) = \frac{1-p}{p^2}$

Negative Binomial Random Variable

(getting r successes in bernoulli trials)

$S_X = \{r, r+1, r+2 \ldots\}$

$P_X(n) = {}^{n-1}C_{r-1} \, p^r (1-p)^{n-r}$

$E(X) = \frac{r}{p}$; $var(X) = r\left(\frac{1-p}{p^2}\right)$

Hyper Geometric Random variable

$S_x = \{0, 1, 2, 3 \ldots n\}$

if $n \leq K$ and $n \leq N-K$

$$P_X(x) = \frac{({}^{K}C_x)({}^{N-K}C_{n-x})}{{}^{N}C_n}$$

$$E(X) = \frac{nK}{N}$$

$$var(X) = \frac{N-n}{N-1} \cdot \frac{nK}{N} \cdot \left(1 - \frac{K}{N}\right)$$

Poisson Random Variable

$S_x = \{0, 1, 2, 3 \ldots\}$

$$P_x(k) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t}$$

$E(x) = \lambda t$ $\quad$ $Var(x) = \lambda t$

29 January 2023

Advance Statistics

RV $\rightarrow X$

($S_x = \{$ continuous interval $\}$

$\rho = \frac{m}{V}$

$\rho(x, y, z)$ density of definite solid body.

$\rho(v)$

$$m = \int_x \int_y \int_z \rho(x, y, z)\, dx\, dy\, dz$$

$$= \int_V \rho_v\, dv$$

$0 \leq X < 1$

Probability (PDF)

Continous Density Function

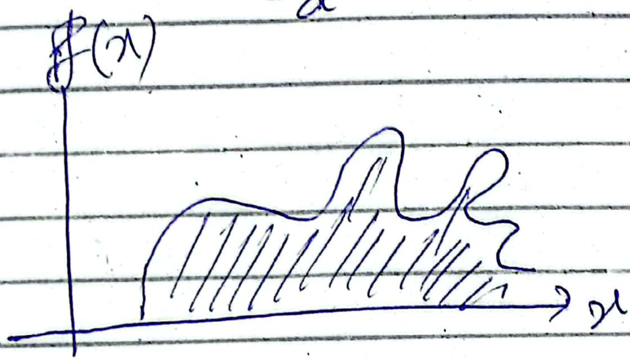
(i) $f_x(x) \geq 0$. $P[X \leq a] = \int_{-\infty}^{a} f_x(x)\,dx$

(II)

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

$P[X \leq a \;\&\; X \geq b]$

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

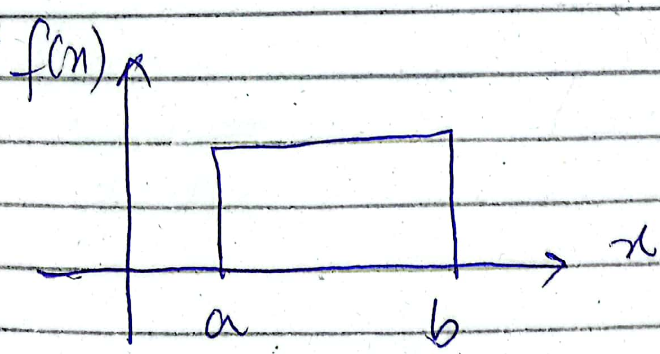


a

$h(b-a) = 1$

$h = \frac{1}{b-a}$

*Uniform RV:-



$\int_{a}^{b} f(x) = 1$

$\text{height} = \frac{1}{b-a}$

(I) Uniform RV :-

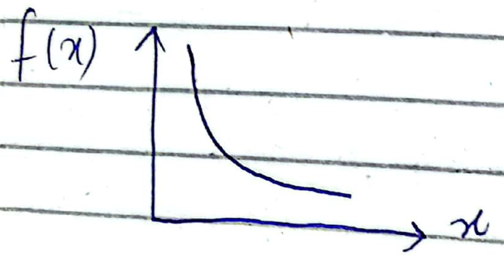
$$f_x(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

$$E[x] = \int_{-\infty}^{\infty} f(x) \cdot x \, dx$$

$$E[x] = \frac{a+b}{2} ;$$

$$V[x] = \frac{(b-a)^2}{12}$$

(II) Exponential RV
(continous)



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

Iter arrival time in poisson dist. is waiting time and is modeled by exponential random variable because it is used to measure inter arrival time.

$E[x] = \frac{1}{\lambda}$ $\quad$ $var[x] = \frac{1}{\lambda^2}$

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

Cummulative distribution Function

CDF $\quad$ F

$Pr(X \le a) = ?$

if discrete, $P(x \le a) = \sum_{x \le a} P_x(x)$

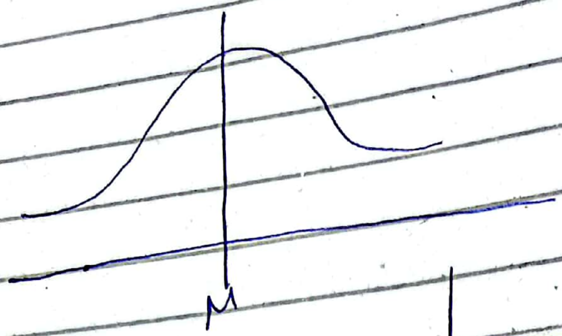
if continuous, $P(x \le a) = \int_{-\infty}^{a} f_x(x)\,dx$

$= \boxed{F_X(a)}$

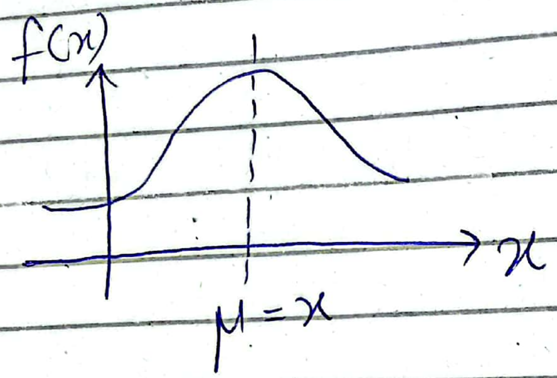
*axiom 3:- $Px(a \le X \le b) = \int_a^b f(x)\,dx$

* $\rightarrow$ $Px(a \le X \le b) = F_X(b) - F_x(a)$

3) Normal Distribution / Gaussian



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



* when $\sigma \uparrow$ width is high
when $\sigma \downarrow$ width is low.

$E[x] = \mu$

$Var[x] = \sigma^2$

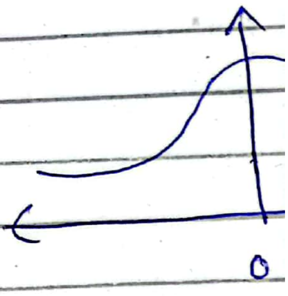
$$P[a \le X \le b] = \int_a^b f_x(x)\, dx$$

Standard

- special

$f_z(z) =$

$\mu = 0$
$var[x]$



$Z \rightarrow$

$P_x [$

$P_x [$

$P [$

$P_x [$

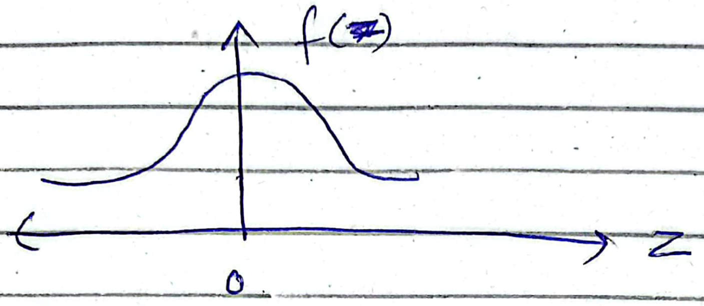
Standard Normal Random Variable

- special

$$f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

$\mu = 0$ $\qquad\qquad$ $E[x] = 0$

$var[x] = 1$



Z → standard normal variable

$$P_x[a \le X \le b] = \cancel{P_x}[\;?\;]$$

$$P_x[X \le a] = \cancel{P\left[\frac{X-\mu}{V}\right]}$$

$$P\left[\frac{X-\mu}{\sigma} \le \frac{a-\mu}{\sigma}\right]$$

$$P_r\left[\cancel{x}\; Z \le \frac{a-\mu}{\sigma}\right]$$

$$P_r[X \leq a] = P\left[\frac{x-\mu}{\sigma} < \frac{a-\mu}{\sigma}\right]$$

$$P_r[X \leq a] = P_r\left[Z \leq \frac{a-\mu}{\sigma}\right]$$

$$P_r(x \leq a) = F_x[a] = F_z\left(\frac{a-\mu}{\sigma}\right)$$

$$f_x(x) = \frac{d}{dx} F_x(x)$$

$$F_x(x) = \int_{-\infty}^{a} f(x)\, dx$$

R.V

(i) $f_{x,}$

$f_{x,}$

$z = f_{x,y}$

$P($

$\int_{-\infty}^{a}$

$cov(X, Y)$

$Cov(x, y) = E[XY] - E[X]E[Y]$

$\star\ \rho = \dfrac{cov(x, y)}{\sqrt{var[X]\,var[y]}}$

$\star\ \rho = -1 \le \rho \le 1$

$\star$ Coefficient of Correlation

31 January 2024 — Wednesday

- fundamentals of prob theory
- RV - discrete & continuous
- mean and variance of RV
- Joint discrete & continuous RV

Random variable X & Y

→ If their covariance or correlation coefficient is zero then they are not related to each other.

Independant Random Variables:-

$$P_{xy}(x,y) = P_X(X) \cdot P_Y(Y)$$

$P_{xy}(x,y)$ ↑ Joint density of x and y ↓ Joint PMF

$P_X(X)$ ↓ probability of X

$P_Y(Y)$ ↓ probability of Y

The above equation holds true if the two variable (Random) are Independant.

$f(x,y) = f(x) \cdot f(y)$ if X and Y are continous random variables.

* if two RV are independant then their covariance is zero but vice versa is not true.
* Independant variables are un-corelated but not vice versa.

Linear Combination of Random Varia

→ let $X_1\ X_2\ X_3 \cdots X_n$ be random variable and $C_1\ C_2\ C_3 \cdots C_n$ are constants

' A new RV "Y" can be a linear combination of X_s and C_s

$$Z = X + 1$$
$$Z = X + Y$$

- We can generate new RV using present RV with help of some mathematical operations

$$Y = X_1 C_1 + X_2 C_2 + X_3 C_3$$

let $\mu_1 = E[X_1]$ $\quad i = 1, 2, 3 \cdots n$

$$\mu_{2i} = E[X_{2i}]$$

$$\sigma^2 = Var[X_i]$$

So, Y is called linear combination of X

$E[Y] = ?$

$$= E[C_1 X_1 + C_2 X_2 \cdots C_n X_n]$$

Quest

$\frac{JPDF}{JPMF}$ = variance / μ = ?

- Linear operators :-
  if $f \cdot x_1 = y_1$ and $f\, x_2 = y_2$ and $f \cdot (x_1 + x_2) = (y_1 + y_2)$ then f is a linear operator.

- Expecatation is a Linear operator

$$E[Y] = E[C_1 X_1 + C_2 X_2 --- C_n X_n]$$
$$E[Y] = C_1\, E[X_1] + C_2\, E[X_2] + --- + C_n\, E[X_n]$$
$$\mu_x = \sum_{i=1}^{N} C_i\, \mu_i$$
$$\mu_Y = E[Y] = [C_1 + C_2 + C_3 --- C_n][E[x_1 + x_2] ---$$

- Variance operator is highly non-linear

IE   <u>Variance</u>

If X_i's are all independent then variance of Y is :-

$$\sigma^2 = Var[Y] = Var[C_1 X_1 + C_2 X_2 --- C_n X_n]$$

$$Var[Y] = C_1^2 Var[X_1] + C_2^2 [Var[X_2]] + C_3^2 Var[X_3] ---- C_n^2 [Var[X_n]]$$

$$\star\ Var[Y] = \sum_{i=1}^{N} C_i^2\, \sigma_i^2 = \sigma_y^2$$

If X_i's are independent i.e not related to each other
And are identically distributed i.e their PMF or PDF is same

$\Rightarrow E[X_i] = \mu \quad \forall\ i = 1, 2, 3 \ldots n$

$\Rightarrow Var[X_i] = \sigma^2 \quad \forall\ i = 1, 2, 3, \ldots N$

because their identically distributed
" $X_i \sim f(x_i)$ "

*

If X_i are IID then

$$Y = \sum_{i=1}^{N} C_i X_i$$

$$\mu_y = \sum_{i=1}^{N} E[X_i] = N\sum_{i=1}^{N}\mu_i$$

$$\mu_y = E\left[\sum_{i=1}^{N} X_i\right] = \sum_{i=1}^{N} E[X_i]$$

$$\boxed{\mu_y = \sum_{i=1}^{N} \mu = N\mu}$$ given that C is 1.

$$\boxed{\sigma_y^2 = N\sigma_x^2}$$

If constants are different :-

$$Y = \sum_{i=1}^{N} C_i X_i$$

$$\mu_y = \mu \cdot \sum_{i=1}^{N} C_i$$

$$\sigma^2 = \sigma^2 \cdot \sum_{i=1}^{N} C_i^2$$

★

If X_i's are IID with Bernoulli distributin then

$$Y = \sum_{i=1}^{N} X_i$$

Y becomes binomial

X can be 0 or 1

$$\mu_y = np \qquad \sigma^2 y = np(1-p)$$

$$PMF(Y) = {}^{N}C_y \cdot p^y \cdot (1-p)^{N-y}$$

★

geometric IID X_i's

$$Y = \sum_{i=1}^{N} X_i \, , \; PMF(Y) = {}^{n-1}C_{r-1} \, p^r (1-p)^{n}$$

Y becomes negative binomial

$$\Rightarrow var(Y) = r\left(\frac{1-p}{p^2}\right)$$

$$\Rightarrow \mu_y = \frac{r}{p}$$

CamScanner

If X_i's are Gaussian / ~~Normal~~ Normal
~~IID~~

$$Y = \sum_{i=1}^{N} C_i X_i$$

Y becomes Normal.

$$\mu_y = \cancel{\mu} \cancel{\quad}$$

$$\mu_y = \sum_{i=1}^{N} C_i \mu_i$$

$$\sigma_y^2 = \sum_{i=1}^{N} C_i^2 \sigma_i^2$$

$$PDF(Y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{y-\mu}{\sigma_x}\right)^2\right\}$$

★

X's are IID Normal then

$$\mu_y = N\mu.$$

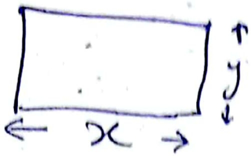
$$\sigma_y^2 = N\sigma^2$$

$$f_y(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left\{-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}$$

pmf

$\frac{1}{\sigma\sqrt{2\pi}} \frac{(x-\mu)^2}{6}$

Question



RV X_1 shows length of component
RV X_2 shows width of component.
Bothe are independent of each other

$E[X_1] = 2\text{ cm}$ $\quad$ $E[X_2] = 5\text{cm}$

$var[X_2] = 0.1\text{ cm}^2$ $\quad$ $Var[X_2] = 0.2\text{ cm}^2$

X_1 and X_2 follow Normal Dist
and are IID.

$peri = 2x + 2y$

$= 2X_1 + 2X_2$

R.V shows perimeter.

R.V. $Y = 2X_1 + 2X_2$

Y is perimeter of component.

What is probability that a selected
Compo. exceeds 14.5 cm in perimeter

$\mu_y = \sum_{i=1}^{N} c_i \mu_i$ $\quad$ because 2 R.V

$= 2 + 5 = 7 \times 2 = 14\text{ cm}$

$\sigma^2_y = N^2 \sigma^2 = 2^2 \cdot (0.1 + 0.2) =$

$\mu_y = 14$ cm

$= 2 \times (0.5 + 0.2) =$

$= 2 \times (7) = 14$ cm

$var(y) = 2^2 \times (0.1)^2 + 2^2 \times (0.2)^2$

$= 4 \times 0.1 + 4 \times 0.04$

$= \cancel{0.4}\ 0.04 + 0.16$

$= 0.2$ cm

$P[x > 14.5 \text{ cm}] = ?$

$P(x < 14.5) = 1 - P(x > 14.5)$

$P(x \leq 14.5) = 1 - F_y(14.5)$

$= 1 - F_z\left(\frac{14.5 - 14}{\sqrt{0.2}}\right)$

$= 1 - F_z(1.136)$

$= 1 - 0.8708$

$= 0.1292.$

Profit in year probability

loss probability

expected gain

$E[x$

$f(x) =$

$f(x) =$

$\int_0^1$

$= \int_0^1 (2$

Profit in year = 4000
Probability = 0.3

loss = 1000
probability = 0.7

expected gain = $0.3 \times 4000 + (-1000 \times 0.7)$

$= \$500$

$E[X] = x\,P(x)$

$$f(x) = \begin{cases} \dfrac{2(x+2)}{5} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$f(x) = \frac{2(x+2)}{5} = \frac{2x+4}{5} \quad 0 < x < 1$$

$$\int_0^1 \frac{2(x)+4}{5}\,dx$$

$$= \int_0^1 \frac{\left(\frac{2x^2}{2}\right)}{5} + \frac{4x}{5}\,dx = \left[\frac{x^2}{5} + \frac{4}{5}x\right]_0^1 \cancel{dx}$$

$$= \frac{1}{5} + \cancel{\frac{4}{5}} - \cancel{\frac{4}{5}} - \frac{1}{5}$$

$2x+4$

$g(x) = (2x+1)^2$

$E[x]_g = g(x) P(x)$

$= (2(-3)+1)^2 \cdot 1/6 +$

$(2(6)+1)^2 \cdot 1/2 +$

$(2(9)+1)^2 \cdot 1/3$

$= 25 \cdot 1/6 +$
$13^2 \cdot 1/2 +$
$19^2 \cdot 1/3$

$= 209$

$f(x) = 3x^{-4} \quad x > 1$

$f(x) = 3x^{-4}$

$F(x) = \int_1^{\infty} 3x^{-4} \, dx$

$= 3 \int_1^{\infty} x^{-4} \, dx$

$= \cancel{3} \left[\frac{x^{-3}}{-\cancel{3}} \right]_1^{\infty}$

$= \cancel{3} \left[-x^{-3} \right]_1^{\infty}$

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

$F(x) = \int_1^t$

$= 3\int$

$= [-x$

$= -t^{-3}$

$P(X > 3)$

$E\{X\} =$

$=$

$=$

$E[x]$

$F(x) = \int_1^t f(x)\,dx = \int_1^t 3x^{-4}\,dx$

$= 3\int_1^t x^{-4}\,dx = \cancel{3}\cdot\left[\frac{x^{-3}}{-\cancel{3}}\right]_1^t$

$= \left[-x^{-3}\right]_1^t = -t^{-3} - (-1)$

$= -t^{-3} + 1 \qquad \boxed{1 - x^{-3}}$ ✓

$1 -$

$P(X > 3) = 1 - P(X < 3)$

$= 1 - (-t^{-3} + 1)$

$= 1 - (-(4)^{-3} + 1)$

$= 1 - \left(-\frac{1}{64}\right) + 1)$

$= 1 - \left(\frac{63}{64}\right) = \frac{1}{64}$

$E[X] = \frac{1}{64} \cdot (3x^{-4})$

$= \frac{1}{64} \cdot (3x^{-4})$

$=$

$E[x] = \int_{-\infty}^{\infty} x f(x)\,dx$

$= \int_1^{\infty} (3x^{-4})(x)$

$= \int_1^{\infty} 3x^{-3}\,dx$

$= \left[\frac{3x^{-2}}{-2}\right]_1^{\infty} = \frac{3}{-2}\left[\frac{1}{x^2} - \frac{1}{1^2}\right]$

$= \frac{3}{-2}\left(0 - \frac{1}{1^2}\right) = \frac{3}{2}$

$3/2$

$$f(y) = \begin{cases} 5(1-y)^4, & 0 \le x \le 1 \\ 0, & \text{else} \end{cases}$$

(a) $c\int_0^1 \cancel{(5-5y)}\,(1-y)^4\,dy$

$= 5\int_0^1 (1-y)^4\,dy$

$= 5\int_0^1 (1-y)^2 \cdot (1-y)^2\,dy$

$= 5\int_0^1 (1+y^2-2y)(1+y^2-2y)\,dy$

$= 5\int_0^1 (1+y^2-2y)+(y^2+y^4-2y^3) -$
$(2y-2y^3+4y^2)\,dy$

$= 5\int_0^1 1+6y^2-4y+y^4\,dy$

$= 5\int_0^1 \cancel{6y^2+y}$

$= 5\int_0^1 y^4+6y^2-4y+1\,dy$

$= 5\left[\left[\frac{y^5}{5}\right]_0^1 + 6\left[\frac{y^3}{3}\right]_0^1 - 4\left[\frac{y^2}{2}\right]_0^1 + y\Big|_0^1\right]$

$= 5\left[\left[\frac{1}{5}\right] + \left[6\cdot\frac{1}{3}\right] - 4\left[\frac{1}{2}\right] + 1\right]$

$= 5\left[\frac{1}{5} + \cancel{2} - \cancel{2} + 1\right]$

$= \frac{1}{5}\times 5 + 1\times 5 = 6$

$= (1-y)^2(1$

$= (1+y^2-2($

$= (1+y^2-$
$(-2y$

$= 1+y^2+$
$-2y^3$

$= 1+6y$

$\int_0^1 (1+6$

$5[6y$

$5[1-$

$\cancel{5}+$

$1($

$= 1$

$= (1-y)^2 (1-y)^2$

$= (1+y^2-2(y))(1+y^2-2y)$

$= (\cancel{1}+\cancel{y^2}-\cancel{2y}) + (\cancel{y^2}+\cancel{y^4}-2y^3) + (-\cancel{2y}-\cancel{2y^3}+\cancel{4y^2})$

$= 1+y^2+y^2+4y^2-2y-2y+y^4-2y^3 -2y^3$

$= 1+6y^2-4y-4y^3+y^4$

$= 5\int_0^1 (1+6y^2-4y-4y^3+y^4)\, dy$

$= 5\left[y+\frac{6y^3}{3}-\frac{4 \cdot y^2}{2}-\frac{4y^4}{4}+\frac{y^5}{5}\right]_0^1$

$= 5\left[1+\frac{6}{3}-\frac{4}{2}-1+\frac{1}{5}\right]$

$= \cancel{5}+\frac{30}{3}-\frac{20}{2}-\cancel{5}+1$

$= \cancel{10}-\cancel{10}+1$

$= 1$

$P(x < 0.1) = ?$

$F(x) = 5\int_0^t 1+6y^2-4y-4y^3+y^4 \, dy$

$= 5\left[y + \frac{6y^3}{3} - \frac{4y^2}{2} - \frac{4y^4}{4} + \frac{y^5}{5} \right]_0^t$

$= 5\left[t + 2t^3 - 2t^2 - t^4 + \frac{t^5}{5} \right]$

$P(x < 0.1) =$

$\cancel{\frac{0.1}{5} = \left[t + 2t^3 - 2t^2 - t^4 + \frac{t^5}{5} \right]}$

$P(x<0.1) = 5\left(0.1 + (0.1)^3 \cdot 2 - 2(0.1)^3 - (0.1)^4 + \frac{(0.1)^5}{5}\right]$

$= \cancel{0.0049}$

$= 0.4095$

$1 - P(x < 0.5)$

$= 1 - (1-y)^5$

$= 1 - (1-0.5)^5 =$

0.3125

$E[x] =$

$$E[X] = \int_0^1 x\, f(x)$$

$$= \int_0^1 y \cdot (1-y)^4\, dy$$

$$= \int_0^1 \cdot$$

$$=$$

$$= y \cdot [5y + 10y^3 - 10y^2 - 5y^4 + y^5]$$

$$= 5y^2 + 10y^4 - 10y^3 - 5y^5 + y^6$$

$$= \cancel{5 + 10 - 10 - 5 + 1}$$

$$= 1$$

$$= 5\int_0^1 (y + 6y^3 - 4y^2 - 4y^4 + y^5)\, dy$$

$$= \left[\frac{y^2}{2} + \frac{6y^4}{4} - \frac{4y^3}{3} - \frac{4y^5}{5} + \frac{y^6}{6} \right]_0^1 5$$

$$= \left[\frac{1}{2} + \frac{3}{2} - \frac{4}{3} - \frac{4}{5} + \frac{1}{6} \right] 5$$

$$= \frac{5}{2} + \frac{15}{2} - \frac{20}{3} - \frac{20}{5} + \frac{5}{6}$$

$$= \frac{5}{2} + \frac{15}{2} - \frac{20}{3} - 4 + \frac{5}{6} =$$

$$= \frac{1}{6}$$

$$\left(P(X > 1/6) = 1 - P(X < 1/6) \right.$$

$$= 1 - (1 - 1/6)^5$$

$$= 0.40187$$

$\mu = 3$

exponential · dist.

$x/1.5$

$f(x) = \lambda e^{-\lambda x}$ $\quad 0 \leq x$

$= 3$

$\left(\frac{1}{3}\right) e$

$E[x] = 1/\lambda$

$var[x] = 1/\lambda^2$

$3 = \frac{1}{\lambda}$

$\boxed{\lambda = \frac{1}{3}}$

$\frac{(\lambda t)^k e^{-\lambda t}}{k!}$

λt $\quad 3.$

$\mu = 3/$

$\frac{(\lambda t)^k}{k}$

$\mu = 3/\text{hour}$

$$\frac{(\lambda t)^k e^{-\lambda t}}{k!} = \frac{(3 \cdot 60)^{10} \cdot e^{-3 \cdot 60}}{10!}$$

$$=$$

$\mu = \frac{1}{x}$ $x = 3$

$$\frac{\cancel{60}^{2}}{\cancel{3}_{1}} = 20$$

$$= 1 - e^{-x/\mu}$$
$$= 1 - e^{-4/3}$$
$$=$$

$\mu = 3$ minutes

$$f(x) = \lambda e^{-\lambda x}$$
$$= 3 \cdot e^{-\cancel{x}(3)(4)}$$
$$=$$

$$F(x) = \int_{1}^{t} \lambda e^{-\lambda x} \cdot x$$
$$= \int \lambda e^{-\lambda}$$
$$= \int 3 e^{-3x} \cdot x$$
$$= 3 \int e^{-3x} \cdot x$$
$$= 3 \int e^{-3x} x (-3x)$$
$$= \boxed{\frac{e^{-3x} \cdot 3x + 1}{3}}$$

CamScanner

PDF for exponential dist.

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

CDF for exponential function

$$= 1 - e^{(-x/\mu)}$$

$$= \int_{0} 1 - e^{-(x/0.3)}$$

66 One man completes reservation with mean time of 3 minutes. It is an exponential distribution. Out of 5 people; what is probability that at-least 4 will complete their reservation in 3 minutes "

→ Known events i.e constant time b/w events
→ Independent of each other (memoryless)

6.45 ask

Exponential	Poisson (PWASAN)
→ how much time passes between one event and another (continuous)	→ how many events occur in a given period of time (discrete)
→ λ (rate of occurrence)	→ μ (expected no. of occurrence)
→ units = 1/time or 1/area	→ units ⇒ counts

∴ Exponential is a special case of Gamma distribution when $\alpha = 1$ and $\beta = 1$

∴ It is continuous analogy of geometric distribution.

∴ The probability of occurrence of an event at a particular exact time will be a bernoulli trial (0 or 1)

||

$W = 2X + 3Y$
$E(X) = 5 \quad Var(X) = 4$
$E(Y) = 10 \quad Var(Y) = 9$

$E(W) = 2 \times 5 + 3 \times 10$
$= 10 + 30 = 40$
$Var(W) = 4 \times 2^2 + 9 \times 3^2$
$= 16 \times 4 + 81 \times 9$
$= \cancel{793}\ 97$

$P(W < 30) = P\left(Z < \frac{30 - 40}{\sqrt{97}}\right)$

$P(Z < -0.335)$
$= \cancel{0.3707}\ 0.1562$ *

(Q2)
$\mu(x) = 2$
$\sigma(x) = 0.1$

$\mu(y) = 0.2$
$\sigma(y) = 0.1$

$P(X > 4.3)$
$= 1 - P(X < 4.3)$
$= 1 - P\left(X < \frac{4.3 - 4}{0.02}\right)$

$y = x + y$
$\mu(y) = 2 + 2 =$
$= 4\text{ mm}$
$\varepsilon(y) = 0.1^2 + 0.1^2$
$= 0.01 + 0.01$
$= 0.02$

$Var(Y) = ($
$= 2$
$\sigma = 0$
$P(Z <$

$X($
$\bar{X} =$
$\sigma(x) =$

$\bar{Y} =$
$\sigma =$

$P($

$Var(Y) = (0.12)^2 + ((0.1)^2)^2$

$= 2 \times 10^{-4}$

$\sigma = 0.0141$

$P(z < 4.3) = \frac{0.3}{0.0141}$

$=$

$X($	$W = X + Y$
$\bar{X} = 0.2$	$\bar{W} = 2 + 2$
$\sigma(X) = 0.1$	$= 4 mm$
$\bar{Y} = 0.2$	$var(W) = (0.1^2) + (0.1)^2$
$\sigma = 0.1$	$= 0.01 + 0.01$
	$= 0.02$

$P(W > 4.3) = 1 - P(W < 4.3)$

$= 1 - P(\bar{w} < 2.121)$

$= 1 - 0.9830$

$= 0.017$

$\bar{X} = 0.1$

$\sigma_x = 0.00031$

$\bar{Y} = 0.23$

$\sigma_y = 0.00017$

$W = X + Y$

$\bar{W} = \bar{X} + \bar{Y}$

$= 0.1 + 0.23$

$= 0.33$

$\sigma_W = \sqrt{(0.00031)^2 + (0.00017)^2}$

$=$

$\sigma = 3.535 \times 10^{-4}$

$P(W < 0.2337)$

$= P(Z <$

$= 0.9967$

7 February 2

$Y = \sum_{i=}$

$E[Y] = ?$

$Var[Y]$

X

1. gau
ind
Non-i
2. Bern
indep
iden
3. Gau
Indep
Iden

Let

inde
and
the

E

Y

17 February 2023 Monday

$$Y = \sum_{i=1}^{N} C_i Y_i$$

$E[Y] = ?$

$Var[Y] = ?$

	X	Y
1.	gaussian, independent, Non-identical	gaussian, independent
2.	Bernaulli, independent, identical	binomial
3.	Gaussian, Independant, Identical.	Normal, IID

Let $X_1, X_2, X_3 \ldots, X_n$ are

independent ~~Normal Random Variables~~ and their PMF & PDF is same i.e they are identical.

$E[X_i] = \mu$ $\quad$ $Var[X_i] = \sigma^2$

$$Y = \sum_{i=1}^{N} X_i$$

- Let $X_1, X_2, X_3 --- X_n$ are IIDs
- $E[X_i] = \mu$ $\quad Var[X_i] = \sigma^2$
- $E[X] = N\mu$ $\quad$ • $Var[X] = n\sigma^2$
- $Y = \dfrac{X - n\mu}{\sqrt{N\sigma^2}}$
- $F_Y(y) = Pr(Y \le y)$

$$= Pr\left(\left(\frac{\sum^{N} X - N\mu}{\sqrt{N\sigma^2}}\right) \le y\right)$$

$=$

CENTRAL LIMIT THEOREM

$\lim\limits_{N\to\infty} F_Y(y) = F_Z(z)$ where z is Standard Normal variable

$$\lim_{N\to\infty} F_Y(y) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-z^2} \, dz$$

C

Thus, ~~probability~~ distribution of X (capital) approaches Normal Random variable when N is large.

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

Example :-

" We load 100 packs they are Independent and they are distributed ~~normally~~ 5 to 50 kg is the variation for each pack (continous range) It is uniformally distributed. What is prob that total weight exceeds 3000 kg-

let Y_i show the weight of i^{th} pack.

$N = 100$

$i = 1$ to 100

X shows total weight

$Pr(X \geq 3000) = ?$

$$E[X_i] = \frac{a+b}{2} = \frac{50+5}{2} = \frac{55}{2} = 27.5$$

$$var[X_i] = \frac{(b-a)^2}{12} = \frac{(45)^2}{12} = \cancel{22.5}$$

$$= 168.75$$

$$X = \sum_{i=1}^{N} X_i$$

$$E[X] = n\mu = 100 \times 27.5$$
$$= 2750$$

Using CLD, $\quad Var[X] = 168.75 \times 100$
$$= 16875$$

$$X \sim N(100 \cdot \mu,\ 100\,\sigma^2)$$

$$P(X > 3000) = 1 - P(X \leq 3000)$$
$$= 1 - F_X(3000)$$
$$= 1 - F_Z\left(\frac{3000 - 100\mu}{\sqrt{1000\,\sigma^2}}\right)$$
$$= 1 - F_Z(1.92)$$
$$= 1 - 0.9726$$
$$= 0.0274$$

Statistical Analysis

Population ⇒ totality of all observatio

Sample ⇒ Subset of population

We always take a random sample (not sample space)

⇒ independently & randomly choosing points / observances without any bias.

⇒ each observation is an outcome of random variable that are independent.

⇒ Point estimation.

→ to guess a single point for parameters that build the distribution of population.

- Sample mean :-

  Let X_i, X_2, X_3 --- are Random Va

$$\bar{X} = \frac{1}{N} \cdot \sum_{i=1}^{N} X_i$$

- Sample variance :-

$$S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2$$

→ mean of sample is also a R.V like variance of sample mean

→ This is because mean is not constant and is based on random observations.

Statistics :- Outcome of mathematical operations applied on the Random Samples

Example :-

X_1 X_2 X_3 --- X_n are IID

$$\bar{X} = \frac{1}{N} \cdot \sum_{i=1}^{N} X_i \; ; \; E[X_i] = \mu$$
$$Var[X_i] = \sigma^2$$

$$E[\bar{X}] = \frac{1}{N} \cdot N \cdot \sum_{i=1} X_i$$
$$= \mu$$

$$Var[\bar{X}] = N \cdot \sigma^2$$
$$= N \cdot \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2$$
$$= N \cdot \frac{1}{N^2} \sigma^2 = \frac{\sigma^2}{N}$$

let $Y = \frac{\bar{X} -}{\sqrt{\sigma^2}}$

according to

when N is (

normal as

$\bar{X} \sim$

CDF$(\bar{X}$

Let $Y = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$

according to CLT, Y becomes normal (standard)

when N is large, Therefore, $\bar{X}$ becomes normal as well when N is large.

$\bar{X} \sim N(\mu, \frac{\sigma^2}{N})$

$CDF(\bar{X}) = Pr(\bar{X} \leq x)$

$= F_Z\left(Z < \frac{\bar{X} - \mu}{\sqrt{\sigma^2/N}}\right.$

$$Var[\hat{\mu}] = Var[\bar{X}] = \frac{\sigma^2}{N}$$

when N is large, variance decreases

Strong Law of large Numbers :-

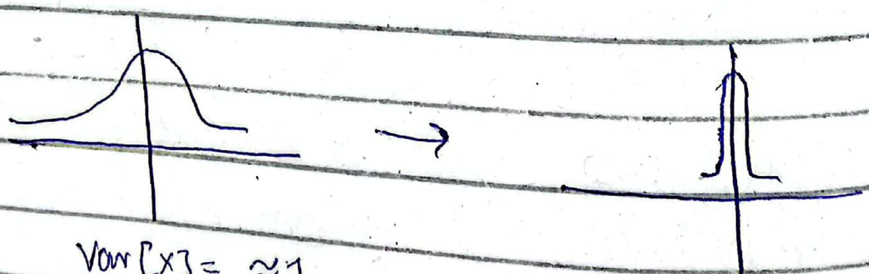
$$\lim_{N\to\infty} \left(\frac{\sum_{i=1}^{N} X_i}{N} \right) = \mu$$

$$Pr\left[\lim_{n\to\infty} \left[\frac{\sum_{i=1}^{N} X_i}{N} \right] = \mu \right] = 1$$

When using CLT, when N is large,

$\cancel{X \sim \hat{\mu}(}$

$$\bar{X} \sim N(\mu, \sigma^2/N)$$



$Var[X] = \approx 1$

$var[X] \leq 1$

Note: we can ...
to make ...

True variance ...
which is unkno...

$\hat{\sigma}^2$ shows t...

$\hat{\sigma}^2 = S^2 =$

$E[\hat{\sigma}^2] =$

$E[S^2] =$

$=$

N

$E[(X_i - X$

$E[$

$= E[$

$= E[(X_i$

$E[(\bar{X}$

$= \sigma^2 +$

$= \sigma^2 + \frac{\sigma^2}{N}$

Note: we can subtract bias from estimator to make it fair.

(X's are IID)

True variance of Population (σ^2) which is unknown

$\hat{\sigma}^2$ shows the estimated value of σ^2

$$\hat{\sigma}^2 = S^2 = \frac{1}{N-1} \sum_{i=1}^{N} [X - \bar{x}]^2$$

$$E[\hat{\sigma}^2] = E[S^2]$$

$$E[S^2] = E\left[\frac{1}{N-1} \sum_{i=1}^{N} (X_i - X)^2\right]$$

$$= \frac{1}{N-1} \sum_{i=1}^{N} E[(X - \bar{x})^2]$$

$$E[(X_i - X)^2] =$$

$$E[\{(X_i - \mu) - (\bar{X} - \mu)\}^2]$$

$$= E[(X_i - \mu)^2 + (\bar{X} - \mu^2) - 2(X_i - \mu)(\bar{X} - \mu)]$$

$$= E[(X_i - \mu)^2] + E[(\bar{X} - \mu^2)] - 2E[(X_i - \mu)(\bar{X} - \mu)]$$

$$= \sigma^2 + \frac{\sigma^2}{N} - 2E\left[(X_i - \mu)\left(\frac{1}{N} \sum X_i - \mu\right)\right]$$

$$= \sigma^2 + \frac{\sigma^2}{N} - \frac{2}{N} E\left[(X_i - \mu)\left[\sum X_i - N\mu\right]\right]$$

$$E[(X_i - \bar{X})^2]$$

$$= \sigma^2 + \frac{\sigma^2}{N} - \frac{2}{N} E\left[(X_i - \mu)\{(X_i - \mu) + (X_2 - \mu) + (X_3 - \mu) \dots\right.$$

$$= \sigma^2 + \frac{\sigma^2}{N} - \frac{2}{N} E[(X_i - \mu)^2]$$

$$= \sigma^2 + \frac{\sigma^2}{N} - \frac{2}{N}\sigma^2$$

$$= \sigma^2 - \frac{1}{N}\sigma^2$$

$$= \sigma^2\left(\frac{N-1}{N}\right)$$

$$E[(X_i - \bar{X})^2] = \sigma^2\left(\frac{N-1}{N}\right)$$

$$E[S^2] = \frac{1}{N-1} \times \sum_{i=1}^{N} \sigma^2 \left(\frac{N-1}{N}\right)$$

$$= \frac{1}{N-1} \cdot N \cdot \sigma^2 \left(\frac{N-1}{N}\right)$$

$$E[S^2] = \sigma^2$$

$$E[\hat{\sigma}^2] = \sigma^2$$

$$S^2 = \frac{1}{N-1} \sum_{i=1}^{N}$$

$$(N-1)S^2 = \frac{1}{1} \sum_{i}^{N} ($$

$$N-1 = S^{-2}$$

$$\left(\frac{N-1}{\sigma^2}\right) S^2 =$$

$$\left(\frac{N-1}{\sigma^2}\right) S^2$$

This distri

Chi -

degree

$$f_x(x) =$$

∴ Chi

of f

subsitute

$$S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2$$

$$(N-1)S^2 = \frac{1}{1} \sum_{1}^{N} (X_i - \bar{X})^2$$

$$N-1 = S^{-2} \sum_{1}^{N} (X_i - \bar{X})^2$$

$$\left(\frac{N-1}{\sigma^2}\right) S^2 = \frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{\sigma^2}$$

$$\left(\frac{N-1}{\sigma^2}\right) S^2 = \frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{\sigma^2} \quad (i)$$

This distribution follows a Chi-square PDF with (N-1) degree of freedom

$$f_x(x) = \frac{1}{2^{v/2} \Gamma(v/2)} x^{(v/2 - 1)} e^{-x/2} \quad (ii)$$

∴ Chi square PDF with v degree of freedom

subsitute v with N-1 to get (i)

$E[X] = \nu \quad var[X] = 2\nu$

Where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, dx$;

Explain

let $= \dfrac{Z}{\sqrt{Y/V}}$ $Z \rightarrow N \sim (0, 1)$, $Y \rightarrow \chi^2 \sim (\nu)$

T follows Student - t distribution

$$T = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}} \longrightarrow N \sim (0, 1)$$

$$\sqrt{\frac{(N-1)S^2}{V^2} \Big/ N-1}$$

Chi-square with N-1 degree of freedom.

$$T = \frac{\bar{X} - \mu}{S/\sqrt{N}}$$

follows t-distribution with (N-1) freedom.

19 February

$N = 16$, $\mu =$

$Pr(\mu_{\bar{X}} - 1.$

$\mu_{\bar{X}} = E[$

$\sigma_x = \sqrt{}$

$Pr(40.5 <$

$Pr(\bar{X} >$

$Pr(\bar{X}$

$\frac{(A,B)}{T_1}$ $T_2(B)$

B	A
4	2
4	3
5	2

where $T_1.A != 1$

1)

2) 3

3) R, A, A

19 February Monday

Advanced Statistics

(Q)

$N = 16, \quad \mu = 50, \quad \sigma^2 = 5^2 = 25$

$Pr(\mu_{\bar{x}} - 1.9\sigma_{\bar{x}} \text{ to } \mu_{\bar{x}} - 0.4\sigma_{\bar{x}})$

$\mu_{\bar{x}} = E[\bar{X}]$

$\sigma_x = \sqrt{Var[\bar{x}]}$

$Pr(40.5 < \bar{X} < 48)$

$Pr(\bar{X} > 40.5) = 1 - Pr(\bar{X} < 40.5)$

$Pr(\bar{X} <$

$N=16$

$\mu = 50$

$\sigma = 5$

$\sigma^2 = 25$

$P_r(\mu_{\bar{x}} - 1.96 < \bar{X} < \bar{\mu}_x - 0.4\sigma\bar{x})$

$P_r(\mu_{\bar{x}} - 1.9\sigma < \bar{X} < \bar{\mu}_x - 0.4\sigma\bar{x})$

$P(\mu_{\bar{x}} - 0.4\sigma_{\bar{x}}) -$ (i)

$= F_X(\mu_{\bar{x}} - 0.4\sigma_{\bar{x}}) -$

$F_X(\mu_{\bar{x}} - 1.9\sigma_{\bar{x}})$ (ii)

$=$ CDF(i) - CDF(ii)

$=$ using central limit theorum

$$F_Z\left(\frac{\mu_{\bar{x}} - 0.4\sigma\bar{x} - \mu\bar{x}}{\sigma\bar{x}}\right) - F_Z\left(\frac{\mu_{\bar{x}} - 1.9\sigma\bar{x} - \mu\bar{x}}{\sigma\bar{x}}\right)$$

∴ we are subtracting CDF of ~~upper~~ lower ~~bound~~ lower bound from upper bound by mapping into "Z" standard normal

$= F_Z(-0.4) - F_Z(-1.9)$

$= 0.3446 - 0.0287$

$= 0.3159$

TRUE ME

~~N~~ $= 36$ $N = 3$

$\sigma = 2$

if std $= 1.2$.

$\lim_{n\to\infty}\left(\frac{\ }{N}\right.$

$N = 36$, $\sigma_{\bar{X}} =$

$Var[\bar{X}] =$

$\sigma^2 =$

$= 3$

$= 14$

TRUE MEAN & TRUE VARIANCE

~~N = 36~~ $N = 36$.

$\sigma = 2$

if std ~~D~~ $= 1.2$.

$$\lim_{n \to \infty} \left(\frac{}{N} \right.$$

$N = 36$, $\sigma_{\bar{x}} = 2$, $N' = ?$

$\sigma_{\bar{x}}' = 1.2$

$$Var[\bar{x}] = \frac{\sigma^2}{N}$$

$$\sigma^2 = N\sigma_{\bar{x}}^2$$

$$= 36 \times 2^2$$

$$= 144$$

$$\sigma^2 = \frac{\sigma^2}{N}$$

$$N' = \frac{\sigma^2}{\sigma_{\bar{x}}^2} = \frac{144}{1.2^2}$$

$$N' = 100$$