Unit - 1

Lec 1 :

- Statistics is defined as collection of quantitative data.
  Main purpose of statistics is making conclusion using limited sample about great population.

- Statistics
  - Descriptive → describes data
  - Inferential → helps you make predictions from data (uses sample statistics)

- Population - entire group that you conclude about
  Species - sample that you collect data from

- parameter is a measure that describes the whole population
  statistic is a parameter for sample.

- sampling error = parameter - statistic

- Data → sampling → Information

- Methods of collecting data : Direct obs. , experiments, surveys

- Conditions for good sample - representativeness, accuracy, size

- sampling : Technique of selecting a group of individual from a population to study them

- Types of sampling : Probability sampling & Non-probability sampling

Probability sampling → involves random selection

Non-probability sampling → involves non-random selection

- Types of probability sampling techniques :

- simple random sampling
- cluster sampling
- Systematic sampling
- Stratified random sampling

---

17 Aug

eg

6R 4W → choosing 3

Solⁿ $X = 0, 1, 2, 3$

$$= \frac{{}^6C_3}{{}^{10}C_6} \;\Big|\; \frac{{}^6C_2 \cdot {}^4C_1}{{}^{10}C_6} \;\Big|\; \ldots\ldots$$

Lec 1-2

- $\sigma = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n}}$ where $\bar{x} = \dfrac{\sum x}{n}$, σ = Root mean square deviation (standard deviation)

- $\sigma = \sqrt{\dfrac{\sum d^2}{n} - \left(\dfrac{\sum d}{n}\right)^2}$, (Assumed mean method)

  $d = x - A$ & A = assumed mean & $\bar{x} = A + \dfrac{\sum d}{n}$

- coeff. of standard deviation $= \dfrac{\sigma}{\bar{x}}$

- variance (σ^2) $= \dfrac{\sum(x-\bar{x})^2}{n}$

- coeff. of variation $= \dfrac{\sigma}{\bar{x}} \times 100$

- Standard deviation & variance for discrete series :

- Assumed mean method

$$\sigma = \sqrt{\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2}, \quad d = x - A, \quad \bar{x} = A + \frac{\sum fd}{\sum f}$$

- Step deviation

$$\sigma = \left(\sqrt{\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2}\right) \times h, \quad d = \frac{x-A}{h}, \quad \bar{x} = A + \left(\frac{\sum fd}{\sum f}\right) \times h$$

fixed gap

- Comparison based q :

coeff. of variance $\uparrow$ $\Rightarrow$ Stability $\downarrow$es

Lec 3 :

eg: 4 cards are drawn

i) 1 King, 1 queen, 1 ace, 1 jack

$$\frac{{}^4C_1 \times 4}{{}^{52}C_4}$$

ii) 2 King, 2 queens — $\dfrac{{}^4C_2 \cdot {}^4C_2}{{}^{52}C_4}$

iii) 2 Black, 2 red — $\dfrac{{}^{26}C_2 \cdot {}^{26}C_2}{{}^{52}C_4}$

eg. 6 white, 4 Red, 3 black $\longrightarrow$ 3

i) 2 W

ii) None are red

Sol^n^ i) $\dfrac{{}^6C_2 \cdot {}^4C_1}{{}^{13}C_3} + \dfrac{{}^6C_2 \cdot {}^3C_1}{{}^{13}C_3}$

ii) 3W, 3B, 2B & 1W, 2W & 1B

or $P(\text{None is Red}) = P(\text{all are W \& B})$

$= {}^9C_3 / {}^{13}C_3$

eg: sum of die is ~~less~~ greater than 3

Sol^n $1 - (\text{sum is } 2, 3)$

$\downarrow$

$(1,1), (2,1), (1,2)$

$$1 - \frac{3}{36}$$

$$= \frac{33}{36} = \frac{11}{12}$$

eg:

3 prod. dept
4 switches dept $\longrightarrow 4$
2 call dept
1 CA

Sol^n i) $\dfrac{{}^3C_1 \cdot {}^4C_1 \cdot 1 \cdot {}^2C_1}{{}^{10}C_4}$

ii) $1 - P(\text{None from switch})$

$$1 - \frac{{}^6C_4}{{}^{10}C_4}$$

iii) $\dfrac{1 \cdot {}^9C_3}{{}^{10}C_4}$

eg: iii)

6W
4R $\longrightarrow 3$
9B

Sol^n i) $\dfrac{{}^6C_2\left(\cancel{{}^6C_1} + {}^4C_1 + {}^9C_1\right)}{{}^{19}C_3}$

ii) $\dfrac{6 \cdot 4 \cdot 9}{{}^{19}C_3}$

iii) $\dfrac{{}^{15}C_3}{{}^{19}C_3}$ or

3W
2W 1B
2B 1W
3B

iv) $P(W \geq 1) = 1 - P(W < 1)$

$$= 1 - \frac{{}^{13}C_3}{{}^{19}C_3}$$

Lec-4:

eg. $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$

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

eg. given: $P(A \cup B)$, $P(A \cap B)$, $P(\bar{A})$

find: $P(B) = ?$

sol^n $P(A \cup B) = 1 - P(\bar{A}) + P(B) - P(A \cap B)$

$P(B) = ?$

eg. $P(\text{plumbing}) = 2/3$, $P(\text{electric}) = 5/9$, $P(A \cup B) = 4/5$, $P(A \cap B) = ?$

$\text{plumbing} \to A$, $\text{electric} \to B$

eg. $P(B) = \frac{3}{2} P(A)$, $P(C) = \frac{1}{2} P(B)$, $P(A) = ?$

sol^n $P(S) = P(A) + P(B) + P(C) = 1$

$P(A) = ?$

$$\boxed{\begin{aligned} P[\bar{A} \cap B] &= P(B) - P(A \cap B) \\ P(A \cap \bar{B}) &= P(A) - P(A \cap B) \end{aligned}}$$

Lec - 5

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

- $P(A \cap B) = P(A) \cdot P(B|A)$ —— Multiplication rule of probability

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

eg: i) $P(\text{colorblind} \mid \text{man}) = \dfrac{0.04}{0.51}$

ii) $P(\text{colorblind} \mid \text{female}) = \dfrac{0.02}{0.49}$

eg: ii

$A \begin{cases} \text{Sum}(6) = (2,4), (4,2), (3,3), (1,5), (5,1) \\ \text{Sum}(8) = (4,4), (5,3), (3,5), (2,6), (6,2) \end{cases}$

$B \begin{cases} \text{one } 4 \longrightarrow (1,4), (2,4), (3,4), (5,4), (6,4) \longrightarrow \leftrightarrow \\ \text{two } 4 \longrightarrow 1 \text{ case } (4,4) \end{cases}$

$P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{3}{11}$

~~eg:~~ 4 sided die

Sol^n

$$S = \{ (1,1), (1,2), (1,3), (1,4)$$
$$(2,1), (2,2), (2,3), (2,4)$$
$$(3,1), (3,2), (3,3), (3,4)$$
$$4 \quad 4 \quad 4 \quad 4$$

$P(B) = 5/16$, $B = \{(2,2), (2,3), (2,4), (3,4), (4,4)\}$

$$A = \{ (x,y) \mid \max(x,y) = m \}$$

$m = 1$ $\quad (1,1) \longrightarrow A \cap B = \emptyset$

$m = 2$ $\quad (2,1), (2,2) \longrightarrow A \cap B = 1$

$m = 3$ $\quad (3,1), (3,2), (3,3), (2,3) \longrightarrow A \cap B = 2$

$m = 4$ $\quad (3,4), (4,1), (4,2), (4,3), (4,4), (2,4), (1,4) \longrightarrow A \cap B = 2$

$$P(A|B) = \begin{cases} 0 & m=1 \\ \dfrac{1/16}{5/16} = 1/5 & m=2 \\ \dfrac{2/16}{5/16} = 2/5 & m=3 \\ \dfrac{2/16}{5/16} = 2/5 & m=4 \end{cases}$$

eg: A : first tube is good
B : second tube is good

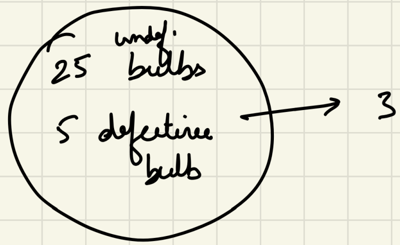
4B
6G $\rightarrow$ 2

sol^n $\quad P(B|A) = \dfrac{P(B \cap A)}{P(A)} = \dfrac{{}^6C_2 / {}^{10}C_2}{6/10} = \dfrac{1/3}{3/5} = \dfrac{5}{9}$

eg. $$P(A|\bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})} = \frac{P(A) - P(A \cap B)}{1 - P(B)}$$

$=$

eg.



Sol^n $$\frac{5}{30} \cdot \frac{4}{29} \cdot \frac{3}{28}$$

$$P(A \cap B \cap C) = P(A) \cdot P(B/A) \cdot P\left(C/B \cap A\right)$$

eg. $P(\text{remaining with company}) = 0.6$ — A

$P(> 10000) = 0.5$ — B

$P(A \cup B) = 0.7$

$P(B|A) = ?$

Sol^n $$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{0.4}{0.6} = 2/3 = 0.67$$

$$\left\{ \begin{aligned} 0.7 &= 0.5 + 0.6 - x \\ x &= 0.4 \end{aligned} \right\}$$

(*) $$P(\bar{A}|\bar{B}) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{A})} = \frac{1 - P(A \cup B)}{1 - P(A)}$$

eg: 1, 2, 3, 4, 5, 6, 7, 8, 9

Sol^n

A = both digits are odd

B = sum is even

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

$P(B) = P(\text{both are odd}) + P(\text{both are even})$

$$= \frac{{}^5C_2}{{}^9C_2} + \frac{{}^4C_2}{{}^9C_2} = \frac{16}{36} =$$

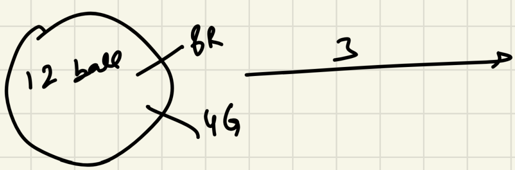
$\frac{5!}{2!\,3!}$ $\quad \frac{4!}{2!\,2!}$

$\frac{9!}{2!\,7!}$

$$P(A \cap B) = \frac{{}^5C_2}{{}^9C_2} = \frac{10}{16}$$

$$\frac{P(A \cap B)}{P(B)} = \frac{{}^5C_2/36}{P(B)} = \frac{{}^5C_2}{\cancel{36}} \times \frac{\cancel{36}}{16} =$$

eg:



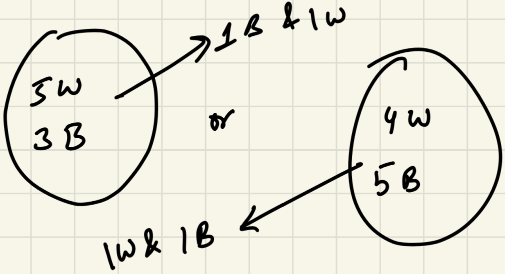
Sol^n

RGR $\longrightarrow$ $\frac{8}{12} \cdot \frac{4}{11} \cdot \frac{7}{10}$

GRG $\longrightarrow$ $\frac{4}{12} \cdot \frac{8}{11} \cdot \frac{3}{10}$

$$P = \frac{8\ \cancel{32} \times \cancel{10}}{3\ \cancel{12} \times 11 \times \cancel{10}} = \frac{8}{33}$$

eg.



Solⁿ

$$\frac{1}{2}\left(\frac{{}^5C_1 \cdot {}^3C_1}{{}^8C_2}\right) + \frac{1}{2}\left(\frac{{}^4C_1 \cdot {}^5C_1}{{}^9C_2}\right)$$

eg. sum = 9

Solⁿ $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{(6,3),\ (3,6)}{(6,3),\ (3,6),\ (4,5),\ (5,4)}$

$$= \frac{2/36}{4/36} = 1/2$$

Lec 6 : Independent events

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

Mutually exclusive events (disjoint) : $P(A \cap B) = 0$ $\because A \cap B = \phi$

whereas Independent " : $P(A \cap B) = P(A) \cdot P(B)$

eg. $P(A) = 0.4$ $P(B) = p$

$P(A \cup B) = 0.7$

Sol^n ① for mutually exclusive & exhaustive events

$$\boxed{P(A) + P(B) = 1}$$

$\Rightarrow \quad 0.4 + p = 1$

$p = 0.6$

② $\quad 0.7 = 0.4 + p - 0$

$p = 0.3$

③ $\quad P(A \cap B) = P(A) \cdot P(B)$

$0.7 = 0.4 + p - P(A \cap B)$

$0.3 = p - 0.4p$

$0.3 = 0.6p$

$p = 1/2$

eg. Event A, Event B

Sol^n To prove A & B are independendent, prove

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

eg. $P(D) \cdot P(A) \neq P(D \cap A) \Rightarrow$ A & D are dependent

Lecture 7 (Total probability theorem):

Lecture - 7

eg. x students

$\frac{2}{3}x$ boys $\longrightarrow P = 0.28$

$\frac{x}{3}$ girls $\longrightarrow P = 0.25$

Partition :

i) $P(A_i) \geq 0$ for each i

ii) Events A_i are pairwise disjoint i.e.
$A_i \cap A_j = \phi$ for $i \neq j$

iii) Union of events A_i equal to sample space S i.e.

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

eg. $S = \{1, 2, 3, 4, 5, 6\}$ has partition ?

sol^n $A_1 = \{1, 2\}$ $\quad$ $A_2 = \{3, 4, 5, 6\}$

$$A_1 \cap A_2 = \phi$$

$$A_1 \cup A_2 = S$$

$\Rightarrow A_1$ & A_2 form a partition of S.

Total probability theorem :

Let $A = \{A_1, A_2 \dots A_n\}$ be partition of sample space S.
If B is any event then

$$\boxed{P(B) = P(A_1)\,P(B|A_1) + \dots = \sum_{i=1}^{n} P(A_i) \cdot P(B|A_i)}$$

Solⁿ

A_1 : student is girl

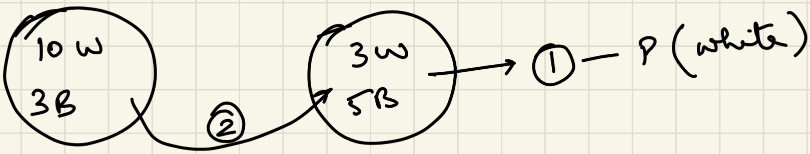
A_2 : " " boy

B : " will get first class

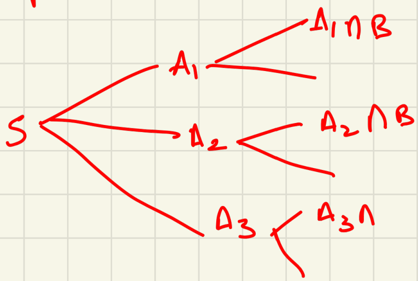
$$P(B/A_1) = \frac{1}{3}(0.25) \qquad P(B/A_2) = \frac{2}{3}(0.28)$$

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

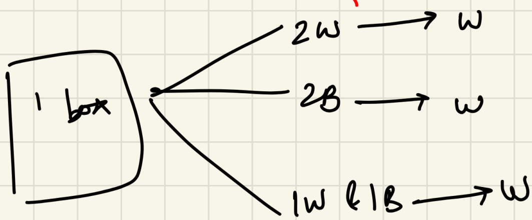
eg.



Solⁿ for applying total probability theorem, our sample space is divided into partitions



$P(B) = \underline{\qquad\qquad}$.



$$P(\text{white}) = \frac{{}^{10}C_2}{{}^{13}C_2}\left(\frac{5}{10}\right) + \frac{{}^{3}C_2}{{}^{13}C_2}\left(\frac{3}{10}\right) + \frac{{}^{10}C_1 \cdot {}^{3}C_1}{{}^{13}C_2}\left(\frac{4}{10}\right)$$

eg. 15 Intelligent Inte fails = 0.005
45 Medium M = 0.05
15 Avg. A = 0.15

P(student passed) = ?

Solⁿ $A_1 : I \quad A_2 : M \quad A_3 : Avg.$

B : passes

$$P(B) = \left(\frac{15}{75}\right)(1-0.005) + \left(\frac{45}{75}\right)(1-0.05) + \frac{15}{75}(1-0.15)$$

Lecture 8 :

- Bayes' theorem : If $A_1, A_2, A_3 \dots A_n$ are mutually disjoint events with $P(A_i) \neq 0$ then for arbitrary event E which is subset of $\bigcup_{i=1}^{n} A_i$ such that $P(E) > 0$ then

$$\boxed{P(A_i \mid E) = \frac{P(A_i) \cdot P(E \mid A_i)}{\sum P(A_i) P(E \mid A_i)}}$$

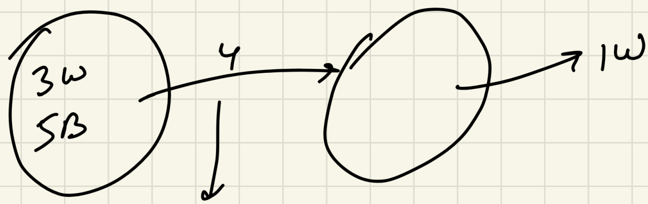
eg. P(knew) = 0.6 P(answered correctly) = 1/5
P(guessed) = 0.4
P(knew | answered correctly) = ?

Sol^n $\quad P = \dfrac{0.6(1)}{0.4\left(\frac{1}{5}\right) + 0.6(1)} = \dfrac{0.6}{0.68}$

eg: $\quad P\left(4/\text{Bonus}\right) = \dfrac{P\left(\text{Bonus}/4\right) \cdot P(4)}{\Sigma \text{ \_\_\_\_\_\_\_\_ }}$

$$= \frac{\frac{2}{9} \cdot \frac{1}{2}}{\frac{4}{9} \cdot \frac{3}{10} + \frac{2}{9} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{4}{5}}$$

eg:



$P(3W, 1B) = ?$

Sol^n

$S \to 3W\ 1B \to 1W$

$S \to 2W\ 2B \to 1W$

$S \to 1W\ 3B \to 1W$

$S \to 4B \to 1W$

$$P\left(3W, 1B \mid 1W\right) = \frac{P\left(1W / 3W\ 1B\right) \cdot P(3W\ 1B)}{\Sigma}$$

$$P = \frac{\frac{1}{3}\cdot\left(\frac{^3C_3\cdot{}^5C_1}{^8C_4}\right)}{\frac{3}{4}\left(\frac{^3C_3\cdot{}^5C_1}{^8C_4}\right)+\frac{9}{4}\left(\frac{^3C_2\cdot{}^5C_2}{^8C_4}\right)+\frac{1}{4}\left(\frac{^3C_1\cdot{}^5C_3}{^8C_4}\right)+0}$$

eg: $P\left(\frac{III}{1W\ 1R}\right) = ?$

sol^n
$$P = \frac{P\left(\frac{1W\ 1R}{III}\right)\cdot P(III)}{\Sigma}$$

$$= \frac{\left(\frac{^4C_1\cdot{}^3C_1}{^{12}C_2}\right)\frac{1}{3}}{\frac{1}{3}\left(\frac{^4C_1\cdot{}^3C_1}{^{12}C_2}\right)+\frac{1}{3}\left(\frac{^2C_1\cdot{}^1C_1}{^4C_2}\right)+\frac{1}{3}\left(\frac{^3C_1\cdot{}^1C_1}{^6C_2}\right)}$$

Lecture 9

- pmf , $p(x) \geq 0$ & $\Sigma p(x) = 1$

Random variable

- DRV
  - $p(x) \geq 0$
  - $\Sigma p(x) = 1$
- CRV
  - $f(x) \geq 0$
  - $\int_{-\infty}^{\infty} f(x)\,dx = 1$

- $E(x)$ = mean (Mathematical expectation)

DRV / CRV

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

eg. 16 G, 4B $\rightarrow$ no. of B in draw of 2

Sol^n

X (no. of bad)	0	1	2	3	4
P(X)	$\frac{^{16}C_2}{^{20}C_2}$	$\frac{^{16}C_1 \, ^4C_1}{^{20}C_2}$	$\frac{^4C_2}{^{20}C_2}$	0	0

This is probability distribution

eg. 5 D, 20 G $\rightarrow$ 4

Sol^n

X	0	1	2	3	4	5
P(X)	$\frac{^{20}C_4}{^{25}C_4}$	$\frac{^5C_1 \cdot {^{20}C_3}}{^{25}C_4}$	$\frac{^5C_2 \cdot {^{20}C_2}}{^{25}C_4}$	$\frac{^5C_3 \cdot {^{20}C_1}}{^{25}C_4}$	$\frac{^5C_4}{^{25}C_4}$	0

CRV

eg. $f(x) = \begin{cases} \alpha(2x - x^2) & 0 < x < 2 \\ 0 \end{cases}$

$\alpha = ?$

$P(X > 1) = ?$

Solⁿ

$$\boxed{\int_{-\infty}^{\infty} f(x)\,dx = 1} \rightarrow \text{pdf by definition}$$

$$\int_{-\infty}^{0} f(x)\,dx + \int_{0}^{2} f(x)\,dx + \int_{2}^{\infty} f(x)\,dx = 1$$

$$\alpha\left(\frac{2x^2}{2} - \frac{x^3}{3}\right)_0^2 = 1$$

$$\left(8/2 - 8/3\right)\alpha = 1$$

$$\frac{4}{3}\alpha = 1 \Rightarrow \alpha = 3/4$$

$$P(X>1) = \int_{1}^{\infty} f(x)\,dx$$

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

$$= \alpha\left[8/2 - 8/3 - 2/2 + 1/3\right]$$

$$= \alpha\left[3 - 7/3\right] = \frac{3}{4} \times \frac{2}{3} = 1/2$$

$$\boxed{\text{cdf}: \quad F_X(x) = P(X<x) = \int_{-\infty}^{x} f(x)\,dx}$$

$\circledast$ pdf : $f(x)$
cdf : $F_x(x)$

$$\boxed{\frac{d\,F_x(x)}{dx} = f(x)}$$

eg. a) $P(X \geq 1.5)$

$$P(X \geq 1.5) = \int_{1.5}^{2} (2-x)\,dx + \int_{2}^{\infty} 0\,dx$$

b) cdf = ? if

$$f(x) = \begin{cases} x & 0 \leq x \leq 1 \\ 2-x & 1 < x < 2 \\ 0 & o/w \end{cases}$$

$$f(x) = \begin{cases} \cancel{x^2}/2 & 0 \leq x \leq 1 \\ 2x - x^2/2 & 1 < x < 2 \\ c & o/w \end{cases}$$

?? slide 20, 33

$\circledast$ If $\phi(x) = x$

Mean $\Rightarrow$ $E(x)$
- DRV: $E(x) = \sum x P(x) = \bar{x}$
- CRV: $E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$

$$\boxed{\text{variance} = E(x^2) - (E(x))^2 = E(x - \bar{x})^2}$$

$$\boxed{E(x^2) = \sum x^2 p(x)}$$

eg: $f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & o/w \end{cases}$, $E(x) = ?$
var = ?

sol:
$$\boxed{E(x) = \int_{-\infty}^{\infty} x \, (2e^{-2x}) \, dx}$$

$$= x\left(\frac{\cancel{2}e^{-2x}}{-\cancel{2}}\right) + \int \frac{\cancel{2}}{\cancel{2}}(e^{-2x}) \, dx$$

$$= \left[-xe^{-2x} - \frac{e^{-2x}}{2}\right]_0^{\infty}$$

$$= 1/2$$

$$\boxed{E(x^2) = \int_{-\infty}^{\infty} x^2 f(x) \, dx}$$

Bernoulli Distribution:

$$p[X=x] = p_X(x) = \begin{cases} p^x (1-p)^{1-x} = p^x q^{1-x} & , \quad x = 0,1 \\ 0 & , \quad o/w \end{cases}$$

$p \rightarrow$ success, $q \rightarrow$ failure

$$\boxed{\text{Mean} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \, , \quad \text{var} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

In terms of frequency,

$$\boxed{\bar{x} = \frac{\sum f_i x_i}{\sum f_i}, \quad var(\bar{x}) = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}, \quad Sd = \sqrt{var}}$$

$$\text{Mean} = p$$

$$\text{Var} = pq$$

$$Sd = \sqrt{pq}$$

$M_x(t) = q + pe^t$ (Moment generating f^n)

To show that a f^n is pmf, prove $\sum p(x) = 1$

$\phi_x(t) = q + pe^{it}$ (characteristic f^n)

$Z_x(t) = q + 3p$ (probability generating f^n)

Bionomial distribution :

$$\boxed{p(x) = {}^nC_x \, p^x \, q^{n-x}}$$

All trials are indepent
No. of trials are finite
probability of success is same for each trial.

eg: $S = \{HHH, HTT, THT, TTH, HHT, THH, HTH, TTT\}$, 2 heads = ?

Solⁿ

X	0	1	2	3
$P(X)$	$1/8$	$3/8$	$3/8$	$1/8$

$n=3 \quad p = 1/2 \quad q = 1/2, \quad x = 2$

$p(x) = {}^3C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{3-2}$

eg: $P = {}^{10}C_5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^{10-5}$

eg: Show $p(x) = {}^nC_x p^x q^{n-x}$ is pmf.

Sol^n $\sum_{x=0}^{n} {}^nC_x p^x q^{n-x} = q^n + pq^{n-1} + \dots + p^n = (p+q)^n = 1$

* Mean for binomial distribution $= np$
  Var " " " $= npq$
  std $= \sqrt{npq}$

represented as $X \sim B(n,p)$

* for binomial distribution : Mean > Variance
  for poisson distribution : Mean = Variance

* Moment generating f^n for binomial distribution $= (q+pe^t)^n$
  charateristic f^n " " " $= (q+pe^{it})^n$
  probability " " " " " $= (q+pz)^n$

eg: Out of 800 families with 5 children each, how many families could be expected to have : ① 3 boys ② 5 girls ③ either 2 or 3 boys ④ Atleast 2 girls

Sol^n $N = 800,\ n = 5$

Boys : $p = \frac{1}{2},\ q = \frac{1}{2}$

i) $P = {}^5C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{5-x} \Rightarrow$ No. of families $= 800 \times P$

ii) 5 girls $\Rightarrow$ 0 boys

$$p = {}^5C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{5-0}$$

iii) $p = {}^5C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 + {}^5C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2$

iv) $p = {}^5C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 + {}^5C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 + {}^5C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 +$ 2, 3, 4, 5 girls
3, 2, 1, 0 boys

$${}^5C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5$$

eg: $p = \frac{1}{2}$, $q = \frac{1}{2}$, $n = 10$, $x = 0, 1, 2, 3$, $N = 100$

Solⁿ $100 \times \left({}^{10}C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x} \right)$

eg: $n = 5$, $8\,P(X=4) = P(X=2)$, find i) p
ii) $P(X \geq 1)$

Solⁿ $8\left({}^5C_4\, p^4 (1-p) \right) = {}^5C_2\, p^2 (1-p)^3 \Rightarrow p, q = ?$

eg: $n = 10$, $p = 0.6$, $q = 0.4$, $P(X \geq 2) = ?$

$$P = {}^nC_x\, p^x q^{n-x}$$

Geometric Distribution:

$$\boxed{P = pq^{x-1}}$$

Lec-11

eg: i) Tossing a coin repeatedly until first head appears.
ii) Shoot the target until it hits
iii) Give test until he'll pass it.
iv) Throwing a die repeatedly until first six appears

pmf of geometric distribution is

$$\boxed{P(X=x) = \begin{cases} pq^{x-1} & , x = 1, 2 \dots \infty \\ 0 & o/w \end{cases}} \quad , \quad p+q=1 \quad , \quad p>0$$

for geometric distribution f^n :

$$\boxed{\text{(Avg) Mean} = 1/p \quad , \quad \text{var} = \frac{q}{p^2}}$$

$$\boxed{M_X(t) = \frac{pe^t}{1-qe^t}} \quad , \quad qe^t < 1$$

$$\boxed{\text{pmf}, \quad P(X > k) = q^k}$$

Lec 10a : Uniform discrete distributions

$$P(X=x) = \begin{cases} 1/N & x \in \{1, 2 \dots N\} \\ 0 \end{cases}$$

$$E(X) = \frac{N+1}{2}, \quad E(X^2) = \frac{(N+1)(2N+1)}{6}, \quad \text{var} = \frac{N^2-1}{12}$$

$$M_X(t) = \begin{cases} \dfrac{e^t(e^{Nt}-1)}{N(e^t-1)} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

Lec-12 : (Negative Binomial) or Pascal distribution

→ x no. of trials to reach fixed no. of successes

$$P = {}^{x-1}C_{r-1}\, p^r q^{x-r}$$

(if $r=1$, $P = pq^{x-1}$)

Mean $= \dfrac{r}{p}$, variance $= \dfrac{rq}{p^2}$

eg: 10th child exposed to the disease will be 3rd to catch

Solⁿ $p = 0.4$, $q = 0.6$, $r = 3$

$$P(x=10) = {}^{9}C_2 (0.4)^3 (0.6)^{10-3}$$

eg: $p = 1/2$, $q = 1/2$, $r = 2$, $x = 6$

Solⁿ ${}^5C_1 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^4$

eg: $p = 0.05$, $q = 0.95$, $r = 3$, $x = 5$

Solⁿ ${}^{x-1}C_{r-1}\, p^r q^{x-r}$

atleast 5 $= (\geq 5) = 1 - (< 5)$

$= 1 - [P(x=3) + P(x=4)]$

Lec-13 (Poisson Distribution) :

$n \to \infty$, $p \to 0$, $np = \lambda$ (infinite)

$$\boxed{pmf = \frac{e^{-\lambda} \lambda^x}{x!}}, \quad x = 0, 1, 2 \ldots \infty$$

eg. poisson distribution is limiting case of binomial distribution

Sol

$$P(x) = {}^nC_x p^x q^{n-x}$$

$$\lim_{n\to\infty} P(x) = \lim_{n\to\infty} \frac{n!}{x!\,(n-x)!} p^x q^{n-x}$$

$$= \lim_{n\to\infty} \frac{n(n-1)(n-2)\ldots(n-(x-1))\cancel{(n-x)!}}{x!\,\cancel{(n-x)!}} p^x q^{n-x}$$

$$= \lim_{n\to\infty} \frac{n^x\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{(x-1)}{n}\right) p^x q^{n-x}}{x!}$$

$$= \lim_{n\to\infty} \frac{n^x}{x!} p^x q^{n-x}$$

$$= \lim_{n\to\infty} \frac{(np)^x}{x!} q^{n-x}$$

$$= \lim_{n\to\infty} \frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x!} \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

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

for poisson distribution :

$$\text{Mean} = \text{Var} = \lambda = E(x)$$

$$E(x^2) = \lambda^2 + \lambda$$

$$M_x(t) = e^{\lambda(e^t - 1)}$$ Moment generating f^n

$$\phi_i(t) = e^{\lambda(e^{it} - 1)}$$ characteristic f^n

$$Z_x(t) = e^{\lambda(z-1)}$$ probability generating f^n

eg. $n = 200$, $p = 0.02$

sol^n (1) atleast one defective fuse $P(\geq 1) = 1 - P(<1)$

$= 1 - P(0)$

(2) $P(\geq 3) = 1 - (P(0) + P(1) + P(2))$

(3) $x = 0$

eg. $n = 10$, $p = 0.02$, $N = 10000$

$$\lambda = \bar{x} = \frac{\Sigma(f \times x)}{\Sigma f}$$

eg. $p(\text{twins}) = 1/80$, $n = 30$, $P(X \geq 2)$

sol^n $P = 1 - (P(0) + P(1))$

eg. mean $= 1.5$,

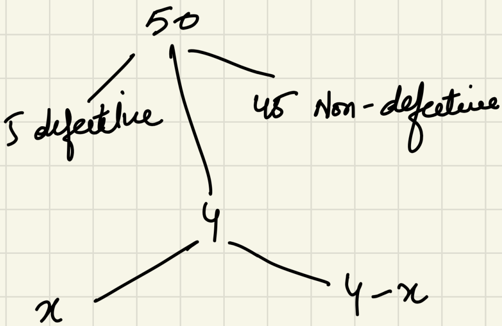
sol^n $\lambda = 1.5$, no car $\longrightarrow x = 0$

some demand is refused $\longrightarrow x > 2$ Not $x = 2$

Lec-14 (Hypergeometric distribution)

sampling without replacement

eg.

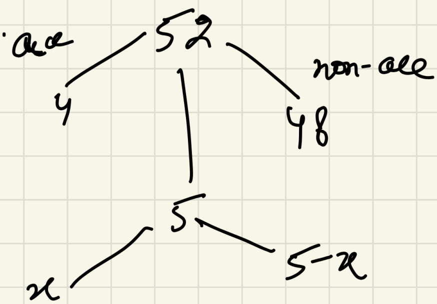


pmf for X = ?

sampling without replacement is done

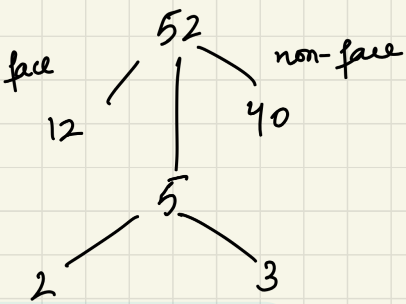
Solⁿ $P(X=x) = \dfrac{{}^{5}C_{x} \cdot {}^{45}C_{4-x}}{{}^{50}C_{4}}$, $x = 0, 1, 2, 3, 4$

eg.



$P(X=x) = \dfrac{{}^{4}C_{x} \cdot {}^{48}C_{5-x}}{{}^{52}C_{5}}$

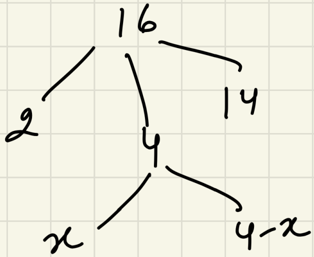
eg.



Made with Goodnotes

eg. $\boxed{E(x) = \frac{nM}{N}}$ $N = 16$ $M = 2$

$n = 4$



Lec-15 (Chebyshev's & Markov's Inequality):

$$\boxed{P(|x-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}}, \quad k > 0$$

eg. mean = 8, var = 9, $P(-4 < x < 20)$ & $P(|x-8| \geq 6)$

Sol^n sd = 3

$$P(-4-8 < x-8 < 20-8) = P(-12 < x-8 < 12)$$

$$= P(|x-8| < 12)$$

$$\Rightarrow k = 4$$

$$P(-4 < x < 20) \geq \frac{15}{16}$$

$$\boxed{P(|x-\mu| \geq k\sigma) \leq \frac{1}{k^2}}$$

eg. $\mu = 12$, $var = 9$, $P(6 < x < 18)$ & $P(3 < x < 21)$

Solⁿ $Sd = 3$ (σ)

$$P(6-12 < x-12 < 18-12) = P(-6 < x-12 < 6)$$

$$= P(|x-12| < 6)$$

$$k = 2$$

$$\Rightarrow P(|x-12| < 6) \geq 3/4$$

$$P(6 < x < 18) \geq 3/4$$

$$P(3 < x < 21) = P(3-12 < x-12 < 21-12)$$

$$= P(-9 < x-12 < 9)$$

$$= P(|x-12| < 9)$$

$$k = 3$$

② **Another form of chebyshev's eqn :**

$$\boxed{\begin{gathered} P(|x-\mu| \geq a) \leq \frac{var(x)}{a^2} \\ P(|x-\mu| < a) \geq 1 - \frac{var(x)}{a^2} \end{gathered}}$$

* Markov's inequality:

$p(x>a)<$ $\boxed{p(|x| \geq a) \leq \frac{E(|x|)}{a}}$

eg.

X	0	1
$p(X)$	$\frac{24}{25}$	$\frac{1}{25}$

$P(X \geq 5) = ?$

solⁿ

$$p(|x| \geq 5) \leq \frac{E(|x|)}{5}$$

$$E(x) = \sum_{x=0}^{1} x P(x) = \frac{1}{25}$$

$$p(x \geq 5) \leq 1/125$$

eg. $p = \frac{1}{5}$, $n = 20$, $x \geq 16$, $P = ?$, find bound for P (coins tossing).

solⁿ

$$p(x \geq 16) \leq \frac{E(x)}{16}$$

$$E(x) = np = 20 \times 1/5$$ (for binomial)

$$\Rightarrow p(x \geq 16) \leq \frac{4}{16}$$

Lec-16 : Joint probability dist., Marginal & Conditional dist.

- for discrete

i) $p(x,y) \geq 0$

ii) $\sum_x \sum_y p(x,y) = 1$

- for continuous :

i) $f(x,y) \geq 0$

ii) $\int_x \int_y f(x,y)\, dy\, dx = 1$

① Marginal of Y, denoted by $p_Y(y)$ or f_Y or $p(Y=y)$ is
x varies

given as :

① Discrete : $p_Y(y) = \sum_x p(X=x, Y=y)$ $\left(\begin{matrix} y = \text{fixed}, \\ x \text{ varies} \end{matrix}\right)$

② Continuous : $\int_x f(x,y)\, dx$ (y fixed, x varies)

- Also joint probability function of X and Y is usually :

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p(X = x_i, Y = y_j) = 1$$

$$\sum_{j=1}^{m} f_Y(y_j) = \sum_{i=1}^{n} f_X(x_i) = 1$$

- Conditional density function :

$$p(X|Y) = \frac{P(X = x \cap Y = y)}{P(Y=y)} = \frac{p(x,y)}{p_Y(y)}$$

↳ marginal density function of y

= Independent Random variable :

Two random variables X and Y are said to be independent if, $\boxed{p(x,y) = p_X(x) \cdot p_Y(y)}$

↳ marginal density functions

eg: $P(X=x, Y=y) = \begin{cases} 1/3 & x=-1, y=0 \\ 1/3 & x=0, y=1 \\ 1/3 & x=1, y=1 \\ 0 & \text{o/w} \end{cases}$

$p_X(x) = ?$

$p_Y(y) = ?$

$P(X \mid Y=1) = ?$

Sol^n

x \ y	0	1	$p_X(x)$
-1	1/3	0	1/3
0	0	1/3	1/3
1	0	1/3	1/3
$p_Y(y)$	1/3	2/3	1

X	-1	0	1
$P_X(x)$	1/3	1/3	1/3

Y	0	1
$P_Y(y)$	1/3	2/3

$$P(X \mid Y=1) = \frac{P(X=x \cap Y=1)}{P(Y=1)} \longrightarrow \text{marginal}$$

for $x=-1$, $P(X \mid Y=1) = \dfrac{0}{2/3}$

$x=0$, $P(X=0 \mid Y=1) = \dfrac{1/3}{2/3} = \dfrac{1}{2}$

$x=1$, $P(X=1 \mid Y=1) = \dfrac{1/3}{2/3} = \dfrac{1}{2}$

$X \mid Y=1$	-1	0	1
$P(X \mid Y=1)$	0	$1/2$	$1/2$

eg:

$x \backslash y$	1	2	$p_X(x)$
1	$2/21$	$3/21$	$5/21$
2	$3/21$	$4/21$	$7/21$
3	$4/21$	$5/21$	$9/21$
$p_Y(y)$	$9/21$	$12/21$	1

Mean of $X, Y, XY, X+Y$

Sol^n

$$\boxed{\text{Mean of } X = E(X) = \sum x\, p_X(x)}$$

$$E(Y) = \sum_y y\, p_Y(y)$$

$$E(XY) = \sum_x \sum_y xy\,(p(x,y))$$

$$E(X+Y) = \sum_x \sum_y (x+y)\,\underline{p(x,y)} \rightarrow \text{joint density } f^n$$

In joint distribution, $P(X=0) = p_X(0)$

eg. 2W, 3R, 4B $\rightarrow$ X : W, 3, Y : R

$P(X, Y)$

Sol^n $X = 0, 1, 2$

$Y = 0, 1, 2, 3$

eg: X : no. of kings $= 4$

Y : no. of aces $= 4$

52 cards $\rightarrow$ 2, without replacement

Sol^n

$Y \backslash X$	0	1	2	$p_Y(y)$
0	${}^{44}C_2 / {}^{52}C_2$	${}^{4}C_1 \cdot {}^{44}C_1 / {}^{52}C_2$	${}^{4}C_2 / {}^{52}C_2$	
1	${}^{4}C_1 \cdot {}^{44}C_1 / {}^{52}C_2$	${}^{4}C_1 \cdot {}^{4}C_1 / {}^{52}C_2$	0	
2	${}^{4}C_2 / {}^{52}C_2$	0	0	
$p_X(x)$				1

$$P(1 \le X \le 2 \mid Y = 0, 2) = \frac{p(1,0) + p(1,2) + p(2,0) + p(2,2)}{p_Y(0) + p_Y(2)}$$

Lec-17

- for continous joint probability

joint density function $= \int_x \int_y f(x,y)\, dy\, dx = 1$

Marginal density function : $f_X(x) = \int_y f(x,y)\, dy$ (x fixed, y varies)

$f_Y(y) = \int_x f(x,y)\, dx$ (y fixed, x varies)

conditional density function : $f(X \mid Y) = \dfrac{f(x,y)}{\underbrace{f_Y(y)}_{\text{marginal}}}$

Independent random variable : $f(x,y) = f_X(x) \cdot f_Y(y)$ for all $x, y \in \mathbb{R}^2$

eg. $f(x,y) = \begin{cases} kxy, & 0 < x < 1,\ 0 < y < 1 \\ 0 & o/w \end{cases}$

value of k = ?

X, Y are independent ?

Marginal density fn of X and Y = ?

$P(X + Y \le 1)$

Solⁿ

① $$\int_0^1 \int_0^1 kxy \, dx \, dy = 1$$

$$\int_0^1 ky \left(\frac{x^2}{2}\right)_0^1 dy = 1$$

$$\int_0^1 ky \left(\frac{1}{2}\right) dy = 1$$

$$k\left(\frac{y^2}{2}\right)_0^1 \frac{1}{2} = 1$$

$$\boxed{k = 4}$$

② $$\int_0^1 4xy \, dy = f_X(x)$$

$$4x\left(\frac{y^2}{2}\right)$$

$$\boxed{2x = f_X(x)}$$

$$\int_0^1 4xy \, dx = f_Y(y)$$

$$4y\left(\frac{x^2}{2}\right)_0^1 = f_Y(y)$$

$$\boxed{2y = f_Y(y)}$$

③ $$f(x,y) = \begin{cases} 4xy & ; \quad 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{o/w} \end{cases}$$

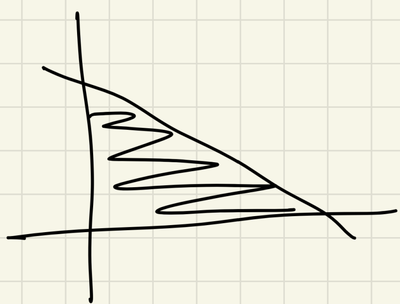
$$f_X(x) = 2x$$
$$f_Y(y) = 2y$$

$$f(x,y) = f_X(x) \cdot f_Y(y) = 4xy$$

$\Rightarrow$ X & Y are independent.

④ $P(X + Y \leq 1) = ?$

$$\int_x \int_y f(x,y) \, dx \, dy$$



$$\int_0^1 \int_{0=y}^{1-x=y} 4xy\, dy\, dx$$

$$\int_0^1 \int_0^{1-x} 4x \cdot \frac{y^2}{2}\, dx$$

$$\int_0^1 \cancel{4}^{2} x \, \frac{(1-x)^2}{\cancel{2}}\, dx$$

$$\int_0^1 2x(1+x^2-2x)\, dx$$

$$\left[\frac{2x^2}{2} + \frac{2x^4}{4} - \frac{4x^3}{3}\right]_0^1$$

eg: $f(x,y) = \begin{cases} e^{-x-y} & x, y \geq 0 \\ 0 & o/w \end{cases}$ $\quad P(X \geq 1) = ?$

Sol^n $f_X(x) = \int_0^\infty e^{-x-y}\, dy = e^{-x}(-e^{-y})$

$$= \left[-e^{-x-y}\right]_{y=0}^{\infty}$$

$$= e^{-x}(-e^{-y})_0^\infty$$

$$= e^{-x}(0 + e^0)$$

$$= e^{-x}$$

$$f_Y(y) = e^{-y}$$

$$P(X>1) = \int_1^{\infty} f_X(x)\,dx$$

$$= \int_1^{\infty} e^{-x}\,dx = 1/e$$

Lec-18 (Conditional mean, & variance, Corr. & regression)

- Independent Random variable : $\boxed{f_{XY}(x,y) = f_X(x) \cdot f_Y(y)}$

- Conditional Mean and Variance

$$E(X) = \sum_{x_i} x_i \cdot P(X = x_i)$$

$$E(Y) = \sum_{y_j} y_j \cdot P(Y = y_j)$$

$$E(X|Y) = \sum_{x_i} x_i \cdot P(X = x_i \mid Y = y_j)$$

$$E(Y|X) = \sum_{y_j} y_j \cdot P(Y = y_j \mid X = x_i)$$

$$E(XY) = \sum_{x_i}\sum_{y_j} x_i y_j \cdot P(X = x_i, Y = y_j)$$

- when X, Y are continuous random variables then

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

(formulas on slide 8)

$$E(Y) = \int_{-\infty}^{\infty} y\, f(y)\, dy$$

conditional expected values

- Correlation and Regression:

① $$\boxed{\text{cov.}(X,Y) = E[XY] - E[X]\cdot E[Y]}$$

② correlation coefficient $= \rho(X,Y)$

$$\boxed{\rho(X,Y) = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}[X]\cdot\text{var}[Y]}} = \frac{\text{cov}(X,Y)}{\sqrt{\sigma_X\,\sigma_Y}}}$$

- Two random variables X and Y are uncorrelated if

$$\text{cov}(X,Y) = 0 \quad \text{or} \quad \rho(X,Y) = 0$$

$\Rightarrow$ X and Y are independent.

-

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\cdot f_{x,y}(x,y)\,dx\,dy$$

$$E[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\cdot f_{x,y}(x,y)\,dx\,dy$$

$$E[Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\cdot f_{x,y}(x,y)\,dx\,dy$$

eg: $f(x,y) = \begin{cases} x+y & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & o/w \end{cases}$

sol^n i) $P\left[0 < X < \frac{1}{2},\ 0 < Y < \frac{1}{4}\right]$

$$= \int_0^{1/2} \int_0^{1/4} (x+y)\, dy\, dx$$

$$= \int_0^{1/2} \left[xy + \frac{y^2}{2}\right]_0^{1/4} dx$$

$$= \int_0^{1/2} \frac{x}{4} + \frac{1}{32}$$

$$= \left[\frac{x^2}{8} + \frac{x}{32}\right]_0^{1/2} = \frac{1}{32} + \frac{1}{64} = \frac{3}{64}$$

Lec 19: (Transformation method)

– CRV
$$\boxed{g(y) = f(x(y)) \left|\frac{dx}{dy}\right|}$$

when eqn $y = v(x)$ has finite no. of sol^n say $x_1, x_2 \ldots x_n$ we can find pdf of w by

$$\boxed{g(y) = f(x_1)\left|\frac{dx_1}{dy}\right| + f(x_2)\left|\frac{dx_2}{dy}\right| + \cdots + f(x_n)\left|\frac{dx_n}{dy}\right|}$$

eg $f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & o/w \end{cases}$

$y = e^x$