17/12/12

Probability & Statistics.

$$\text{Probability} = \frac{\text{no. of favourable ways}}{\text{no of exhaustive ways}}$$

$$\left.\begin{aligned} E_1 &= A \cap B \\ E_2 &= B \cap A \end{aligned}\right\} \text{possible ways} \qquad \begin{aligned} A &= \text{Green} \\ B &= \text{Red} \end{aligned}$$

Probability: It occurs when there is more than one chance get a relation for an action. It is chance of occurence of an event.

Avg represents of a data by a single figure.

deviation - difference between values and average.

Median - To divide the data into two equal halves

Mode - To know the most repeated item.

Terms of probability:

trial - doing the experiment once.

event - either of the outputs

Types of events:

1. Equally likely: All the outputs have equal chances in the event. Ex: Die.
   Drawing ball from the bag.
2. Mutually exclusive: If one event occurs it stops all other events. Ex: Tossing a coin. ($A \cap B = \phi$)
3. Independent events: In one trial, occurence of one event do not depend on another event

$A \cup B \rightarrow$ atleast one of the two events should occur

4) Collection of all possible outcomes of an event is exhaustive events in sample.

5) Favourable events - possible required outcomes.

Rede 1) 3R 5G $\Rightarrow$ $P = {}^3C_2 / {}^8C_2$ [P of selecting 2 red balls].

P of selecting 2 balls of diff. colors $= \frac{{}^3C_1 \times {}^5C_1}{{}^8C_2}$

2) diff $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

$= P(E_1) + P(E_2)$ mutually exclusive

with replacement $= \frac{{}^3C_1}{{}^8C_1} \times \frac{{}^3C_1}{{}^8C_1} + \frac{{}^5C_1}{{}^8C_1} \times \frac{{}^5C_1}{{}^8C_1}$

without replacement $= \frac{{}^3C_1}{{}^8C_1} \times \frac{{}^2C_1}{{}^7C_1} + \frac{{}^5C_1}{{}^8C_1} \times \frac{{}^4C_1}{{}^7C_1}$

2 balls at a time

Same $E_1 - 2R$ $E_2 - 2G$

$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

$= P(E_1) + P(E_2)$

$= \frac{{}^3C_2}{{}^8C_2} + \frac{{}^5C_2}{{}^8C_2}$

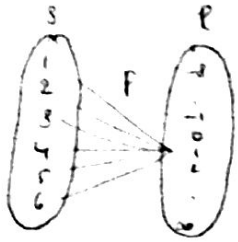
2) diff 1) simultaneous
2) with replacement
3) without replacement

1) $\frac{{}^3C_1 \times {}^5C_1}{{}^8C_2}$ 2) $\frac{{}^3C_1}{{}^8C_1} \times \frac{{}^5C_1}{{}^8C_1}$ or (+) $\frac{{}^5C_1}{{}^8C_1} \times \frac{{}^3C_1}{{}^8C_1}$ 3) $\frac{{}^3C_1}{{}^8C_1} \times \frac{{}^5C_1}{{}^7C_1}$ (+) $\frac{{}^5C_1}{{}^8C_1} \times \frac{{}^3C_1}{{}^7C_1}$

Axiomatic approach in probability.

1) $P(E) \geq 0$ $E \in S$ (non-negativity) $P : S \rightarrow R$.

P is called probability function.



Eg :- $S = \{1, 2, 3, 4, 5, 6\}$.
$P = 1/6, 1/6, 1/6, 1/6, 1/6, 1/6$

2) $P(S) = 1$ (certainity)

3) $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
Mutually exclusive

Addition theorem:

$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

$P(E_1 \cap E_2) = P(E_1) \times P(E_2)$ (E_1 and E_2 are Independent events)

$P(E_1 \cap E_2) = 0$ (E_1 and E_2 are mutually exclusive events)

$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) - [P(E_1 \cap E_2) + P(E_1 \cap E_3) + P(E_2 \cap E_3)] + P(E_1 \cap E_2 \cap E_3)$

Multiplication theorem

Conditional probability:

$P(A/B)$ The probability of occuring event A when subject to the condition that the event B had already occured.

It is given as $\Rightarrow \dfrac{P(A \cap B)}{P(B)}$

If A and B are independent events; then

$P(A/B) = P(A)$ (Since event A is not dependent of event B)

ie probability of occuring event A does not care about whether event B occured or not

$P(B/A) = \dfrac{P(A \cap B)}{P(A)}$. If independent events
$P(B/A) = P(B)$.

$P(\emptyset) = 0$ $\quad$ if $A \subseteq B$ $\quad$ $0 \le P \le 1$

$P(S) = 1$ $\quad$ $P(A) \le P(B)$

$P(\bar{A}) = 1 - P(A)$

$P(A) = 0.3$

$P(B) = 0.5$

$$* \begin{cases} P(\bar{A} \cap B) = P(B) - P(A \cap B) \\ P(A \cap \bar{B}) = P(A) - P(A \cap B). \end{cases}$$

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

(∵ greater value)

Boole's Inequality

1) $P\left(\bigcup_{i=1}^{n} E_i\right) \le \sum_{i=1}^{n} P(E_i) \rightarrow$ upper limit

$P(A) \le P(A \cap B) \le P(A) + P(B)$

(less)

2) $P(\cap E_i) \ge \sum_{i=1}^{n} P(E_i) - (n-1) \rightarrow$ lower limit

It is useful to determine the limits $P(A \cap B)$

Addition Theorem for three events

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

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

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

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

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

1) What is the probability of getting 53rd monday in a non-leap year.

Sol: 365 - 364 (all days equally) 52 weeks 52 mondays

e.e - 7

e.c - $^7C_1$ $\quad$ $\frac{^1C_1}{^7C_1} = \frac{1}{7}$.

f.c - $^1C_1$

2) What is the probability of getting 53rd monday in a leap year.

$0.5 \times 0.4 \times 0.5 = ---$

Multiplication theorem

- $P(A \cap B) = P(A) \times P(B)$

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

$P(A \cap B) = P(A) \times P(B/A)$

7. A box contains 10 chits bearing the numbers 1 to 10 2 chit are drawn at random. What is the probability that their sum is even (ii) odd
a) when drawn together b) out with replacement. c) without replacement

Sol $P(S \text{ is even}) = P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$ Mutually exclusive

$= P(E_1) + P(E_2)$.

a) R.P $= P(E_1) + P(E_2)$

$\frac{5C_2}{10C_2} + \frac{5C_2}{10C_2} = 0.444$

b) R P $= P(E_1) + P(E_2)$

$\frac{5C_1}{10C_1} \times \frac{5C_1}{10C_1} + \frac{5C_1}{10C_1} \times \frac{5C_1}{10C_1} = 0.5$

(c) R.P $= P(E_1) + P(E_2)$

$\frac{5C_1}{10C_1} \times \frac{4C_1}{9C_1} + \frac{5C_1}{10C_1} \times \frac{4C_1}{10C_1}$

(ii) The sum is odd → $P(S \text{ is odd}) = P(\text{even} + \text{odd})$

a) $P(S \text{ is odd}) = \frac{5C_1 \times 5C_1}{10C_2}$

b) E_1 e, o ; E_2 o, e

R.P $= P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

$= P(E_1) + P(E_2)$.

$= \frac{5C_1}{10C_1} \times \frac{5C_1}{10C_1} + \frac{5C_1}{10C_1} \times \frac{5C_1}{10C_1}$

c) $\frac{5C_1}{10C_1} \times \frac{5C_1}{9C_1} + \frac{5C_1}{10C_1} \times \frac{5C_1}{10C_1}$

8 A and B are alternatively throwing a pair of dice the one who first throws a sum 9 wins the game. If A starts the game, what are their respective chances of winning.

A : A get '9'.

B : B get '9'.

$P(A) = \frac{4}{36} = \frac{1}{9}$; $P(B) = \frac{1}{9}$.

Baye's Theorem: If $E_1, E_2, \ldots E_n$ are n mutually disjoint events in the sample space S ($P(E_i) \neq 0 \; \forall \; i = 1, 2, \cdots n$), and A is the subset of union of E_i $\left[A \subseteq (\bigcup_{i=1}^{n} E_i) \; ; \; P(A) > 0\right]$, then

$$P(E_i/A) = \frac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i) P(A/E_i)}$$

$$A \subseteq (\cup E_i)$$

$$A = A \cap (\cup E_i) = \bigcup_{i=1}^{n} (A \cap E_i)$$

$$P(A) = P\left(\cup (A \cap E_i)\right)$$

$P(A \cap E_i) = P(A) P(E_i/A)$ or $P(E_i) P(A/E_i)$

$$= \sum P(A \cap E_i)$$

$$= \sum P(E_i) P(A/E_i) \quad \text{——} \; (1)$$

We should find $\Rightarrow P(E_i/A) = \frac{P(A \cap E_i)}{P(A)}$

$$= \frac{P(E_i) P(A/E_i)}{\sum P(E_i) P(A/E_i)} \quad [\because \text{from } (1)]$$

Here $P(E_i) \rightarrow$ priory probabilities

$P(A/E_i) \rightarrow$ likelyhood probabilities

$P(E_i/A) \rightarrow$ posterior probabilities.

1. An urn A contains 2R, 5G balls. B contains 4R, 3G balls. An urn is chosen at random and 2 balls are drawn. They are foundout to be red. What is the probability that they are from (i) Urn A, (ii) Urn B.

E_1 : Selection of Urn A

E_2 : Selection of Urn B.

A : Selection of 2R balls

$P(E_1) = 1/2 \quad , \quad P(E_2) = 1/2$

$P(A/E1) = \frac{2C_2}{7C_2} \quad , \quad P(A/E_2) = \frac{4C_2}{7C_2}$.

(i) $$P(Ei/A) = \frac{P(E_1) \cdot P(A/E_1)}{\sum_{i=1}^{2} P(E_i) P(A/E_i)}$$

$$= \frac{(1/2)\left(\frac{1}{7C_2}\right)}{\frac{1}{2}\left(\frac{1}{7C_2}\right) + \frac{1}{2}\left(\frac{4C_2}{7C_2}\right)} = 1/7$$

(ii) $$P(E_2/A) = \frac{(1/2)(4C_2)/(7C_2)}{\frac{1}{2}\left(\frac{1}{7C_2}\right) + \frac{1}{2}\left(\frac{4C_2}{7C_2}\right)} = 6/7.$$

UNIT - I - II

RANDOM VARIABLES.

Definition of a random variable:

It is a variable associated with the out comes of a random experiment & it is a function of sample space into the real line.

Discrete and Continuous random variables (r.v's)

A random variable (r.v) X is said to be discrete if it assumes countable set of values, otherwise it is said to be continuous.

Eg. for discrete r.vs: $X = 1, 2, 3, 4, 5, 6$

Eg. for continuous r.vs: $10 < X < 12$ (or)

$-2 < X < 5$

Probability Mass Function: (PMF)

Consider a r.v. 'X' as follows.

X	x_1	x_2	x_3	.	x_n
$P_X(x)$	$P(x_1)$	$P(x_2)$	$P(x_3)$	.	$P(x_n)$.

$P_X(x) = P(X = x)$ is said to be a PMF if it satisfies following 2 conditions.

(i) $P(x_i) \geq 0$

(ii) $\sum_{i=1}^{n} P(x_i) = 1$.

Probability distribution function (PDF)

It is the probability distribution of a continous r.v. represented by f(x). When $a<x<b$ the corresponding f(x) is said to be a PDF if

(i) $f(x) \geq 0 \quad \forall x$.

(ii) $\int_a^b f(x)\,dx = 1$ (or) $\int_{-\infty}^{\infty} f(x)\,dx = 1$ (general).

Cumulative Distribution function (CDF):

It is represented by F(x) and is given as

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

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

Expectation: (E(X))

It is used to know the nature of a r.v. by calculating its average. It is given as

$$E(X) = \begin{cases} \sum x\,P(x) & \text{for discrete r.v.} \\ = \int_{-\infty}^{\infty} f(x)\,dx & \text{for continuous r.v.} \end{cases}$$

Note: i) $E(C) = C$, C is a constant

ii) $E(CX) = CE(X)$

iii) $E(X \pm Y) = E(X) \pm E(Y)$

(iv) $E(XY) = E(X).E(Y)$

Here X and Y are independent R.Vs.

Variance $(V(X))$:

It is the average of squares of deviations of observations taken from their mean.

It is written as

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

$$= E(X^2) - (E(X))^2$$

Note: $V(CX) = C^2 V(X)$

When X and Y are independent r.vs

$$V(X \pm Y) = V(X) \pm V(Y)$$

Example to illustrate discrete r.v:-

1. A r.v, 'x' assumes the following PMF.

X	1	2	3	4	5	6
$P(X)$	a	a^2	$2a$	$2a^2$	$3a$	a^2+2a

Find (i) a

(ii) $P(2 < X \leq 4)$

(iii) $P(X \geq 5)$

(iv) $E(X)$

(v) $V(X)$

(vi) Find minimum value of a such that $P(x \leq a) \geq 0.6$.

(vii) If $y = x^2$ find $E(y \pm 3)$

Solution:

(i) We know that $\sum_{i=1}^{n} P(x_i) = 1$

$\therefore \quad a + a^2 + 2a + 2a^2 + 3a + a^2 + 2a$

$= \quad 8a + 4a^2 = 1$

$\Rightarrow \quad a = 0.12.$

Note: Probability can never take a negative value.

X	1	2	3	4	5	6
P(x)	0·12	0·01	0·24	0·03	0·36	0·25

ii) $P(2 < x \leq 4)$

$= P(3) + P(4)$

$= 0 \cdot 24 + 0 \cdot 03$

$= 0 \cdot 27$

iii) $P(x \geq 5)$

$= P(5) + P(6)$

$= 0 \cdot 36 + 0 \cdot 25 = 0 \cdot 61$

iv) $E(x) = \sum_{i=1}^{n} x P(x_i)$

$= 1(0 \cdot 12) + 2(0 \cdot 01) + 3(0 \cdot 24) + 4(0 \cdot 02) + 5(0 \cdot 36)$
$+ 6(0 \cdot 25)$

$= 4 \cdot 24$

(v) $V(x) = E(x^2) - (E(x))^2$

$= x^2 P(x) - (E(x))^2$

$= 1(0 \cdot 12) + 4(0 \cdot 01) + 9(0 \cdot 24) + 16(0 \cdot 03)$
$+ 25(0 \cdot 36) + 36(0 \cdot 25) - 17 \cdot 98$

$= 20 \cdot 64 - 17 \cdot 98$

$= 2 \cdot 66$

(vi)

X	P(X)	F(x) [Cumulative distribution function].
1	0·12	0·12
2	0·01	0·13
3	0·24	0·37
4	0·02	0·39
5	0·36	0·75
6	0·25	1

$\longrightarrow 0.6$

given $P(x \leq a) \geq 0.6$

$\therefore \quad a = 5$

(vii) $E(y \pm 3) = E(y) \pm 3$

Given $y = x^2$

$\therefore \quad E(x^2) \pm 3$

$E(x^2) + 3 = 20.64 + 3 = 23.64$

$E(x^2) - 3 = 20.64 - 3 = 17.64$

Example to illustrate Continuous r.v.

1. $$f(x) = \begin{cases} ke^{-x} & \text{where } 0 < x < \infty \\ 0 & \text{otherwise} \end{cases}$$

(i) Find k.

(ii) $P(2 < x < 4)$

(iii) $P(x \geq 3)$

(iv) $E(x)$

(v) $V(x)$.

(vi) Cumulative distribution function $(F(x))$

Solution:

We know that

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

$$\therefore \int_{-\infty}^{\infty} ke^{-x}\,dx = 1 \cdot \cancel{\int_{-\infty}^{0} ke^{-x}\,dx} + \int_{0}^{\infty} ke^{-x}\,dx$$

$$\therefore \int_{0}^{\infty} ke^{-x}\,dx = 1 \Rightarrow k\int_{0}^{\infty} e^{-x}\,dx = 1$$

$$\therefore k\left[e^{-x}\right]_{0}^{\infty} = k[0+1] = 1$$

$$\Rightarrow k = 1$$

(ii) $P(2 < x < 4)$

$$\int_{2}^{4} e^{-x}\,dx = -e^{-4} + e^{-2}.$$

(iii) $P(x \geq 3)$

$$= \int_3^\infty e^{-x} dx = -\left[e^{-\infty} - e^{-3}\right] = e^{-3} = \frac{1}{e^3}.$$

(iv) $E(x) = \int_0^\infty x f(x)\, dx$

$$= \int_0^\infty x e^{-x} dx = \Gamma(2) = 1! = 1$$

Note: $\int_0^\infty e^{-x} x^{m-1} dx = \Gamma(m) = (m-1)!$

(v) $V(x) = E(x^2) - (E(x))^2$

$$= \left[\int_0^\infty x^2 f(x)\, dx\right] - (1)^2$$

$$= 2! - 1 \; (2-1) = \underline{\underline{1}}$$

(vi) Cumulative distribution function.

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

Here, $\int_0^x e^{-x} dx = -\left|e^{-x}\right|^x$

$$= \left|-e^{-x} + e^{-0}\right| = \underline{\underline{1 - e^{-x}}}$$

Joint Random Variables.

When 2 r.v. are defined in the same sample space; then x and y are known as bivariate r.vs. Consider the following tables.

Discrete R.Vs

(i)

$X \setminus Y$	1	2	3	4	$P_X(x)$
0	C(2)	C(4)	C(6)	C(8)	20C
1	C(3)	C(5)	C(7)	C(9)	24C
2	C(6)	C(8)	C(10)	C(12)	36C
$P_Y(y)$	11C	17C	23C	29C	$80C = \Sigma\Sigma P_X(x) = \Sigma\Sigma P_Y(y)$

(i) The value of C:

$$(20+24+36)C = 1$$

$$\Rightarrow C = 1/80$$

(ii) $P(X=2, Y=3)$

$$C(10) = \frac{1}{80}(10) = 1/8$$

(iii) $P(X \le 1, Y > 2) =$

$$= C(6) + C(8) + C(7) + C(9)$$

$$= C(30) = \frac{1}{80}(30) = 3/8$$

(iv) Marginal probability of $x = \sum_y P(x,y)$

Marginal probability of $y = \sum_x P(x,y)$

$P(X=0) = 20\left(\frac{1}{80}\right) = 1/4$.

$P(X=1) = 3/10$

$P(X=2) = 9/20$.

$P(Y=1) = 11\left(\frac{1}{80}\right) = \frac{11}{80}$

$P(Y=2) = 17/80$

$P(Y=3) = 23/80$

$P(Y=4) = 29/80$

(v) $E(X) = \sum x P(x)$

$= 0\left(\frac{1}{4}\right) + 1\left(\frac{3}{10}\right) + 2\left(\frac{9}{20}\right) = \underline{\underline{6/5}}$

(vi) $E(Y) = \sum y P(y)$

$= 1\left(\frac{11}{80}\right) + 2\left(\frac{17}{80}\right) + 3\left(\frac{23}{80}\right) + 4\left(\frac{29}{80}\right)$

$= \underline{\underline{\frac{81}{40}}}$.

Continuous r.vs
(Joint density function)

- Determine the constant b such that

$$f(x,y) = 3xy \begin{cases} 0 < x < 1 \\ 0 < y < b \end{cases}$$

is a valid joint density function.
Also calculate mean values of x and y, the joint cumulative distribution function.

$$\int_0^1 \int_0^b 3xy \, dy \, dx = 1$$

$$= \int_0^1 3x \left(\frac{y^2}{2}\right)_0^b dx = 1$$

$$= \frac{b^2}{2}(3) \int_0^1 x \, dx = 1$$

$$= \frac{3b^2}{2}\left(\frac{x^2}{2}\right)_0^1 = 1 \Rightarrow b = +\frac{2}{\sqrt{3}}$$

Mean values of x = $E(x) = \sum x f(x)$

$$f(x) = \int_0^{2/\sqrt{3}} 3xy \, dy = \frac{\cancel{3}x}{\cancel{2}}\left(\frac{4}{\cancel{3}}\right) = 2x$$

$$E(x) = \int_0^1 x f(x) \, dx = \int_0^1 2x^2 = 2\left(\frac{x^3}{3}\right)_0^1$$

$$= 2/3$$

$$E(Y) = \int_0^{2/\sqrt{3}} y f(y)\, dy$$

$$f(y) = \int_0^1 3xy\, dx = 3y\left(\frac{x^2}{2}\right) = \frac{3y}{2}$$

$$E(Y) = \int_0^{2/\sqrt{3}} \frac{3}{2} y^2\, dy = \frac{\cancel{3}}{2}\left(\frac{1}{\cancel{3}}\right)\left[y^3\right]_0^{2/\sqrt{3}}$$

$$= 4/3\sqrt{3}.$$

Joint cumulative distribution function.

$$\int_0^1 \int_0^{2/\sqrt{3}} 3xy\, dy\, dx = 1$$

Distributive function properties:

The distribution func^n gives cumulative probability of the random variable. If X is a random variable, its cumulative distribution is denoted by $F_X(x)$ and is given as $F_X(x) = P(X \le x)$

$= \sum^{x} p(x)$ (if x is discrete)

$= \int_{-\infty}^{x} f(x)\, dx$ (x is cont)

$= \sum^{x} \sum^{y} p(x, y)$

$= \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dy\, dx$

Properties of DF: (i) $f(-\infty) = 0$

(ii) $f(\infty) = 1$

(iii) $P(a \le X \le b) = f(b) - f(a)$.

Note $F(X) = 8x^2$ $\quad 0 < x < 2$

$E(X) = ?$

$f(x) = \frac{d}{dx} F(x)$.

$F(x, y) = 8xy$.

$\frac{\partial^2}{\partial x \partial y} (8xy) = f(x)$.