$$\Rightarrow \tilde{\mu} = \frac{-\ln 0.5}{\lambda}$$

$$\Rightarrow \tilde{\mu} = \frac{0.693}{\lambda}$$

Chapter-5: Joint probability distributions and random samples

5.1

Jointly distributed random variables

(More than 1 random variable at a time)

⇓

In our syllabus upto 2 random variables only.

Two discrete random variables

The probability mass function of a single random variable 'X' specifies how much probability mass is placed on each possible value of X.

The pmf of two random variables X and Y describes how much probability mass can be replaced on each possible pair of values (x, y)

Let, X and Y be two discrete random variables defined on a sample space X on an experiment. Then the joint pmf of X defined for each possible pair of values (x, y) is given by $p(x, y) = P(X = x \text{ and } Y = y)$

Here, $p(x,y) \geq 0$, $\sum_x \sum_y p(x,y) = 1$

For any set 'A' consisting of the pair of values (x,y)
$$P\{(x,y)\in A\} = \sum_{(x,y)\in A}\sum p(x,y)$$

The marginal pmf of X is denoted as $p_X(x)$ and is defined as $p_X(x) = \sum_{y:p(x,y)\geq 0} p(x,y)$; for each possible values of x.

Similarly, $p_Y(y) = \sum_{x:p(x,y)\geq 0} p(x,y)$; for each possible value of y.

The two discrete random variables X and Y are independent if $p(x,y) = p_X(x)\cdot p_Y(y)$
If the above condition does not hold good, then we say the two variables are dependent.

If X and Y are two drv with joint pmf $p(x,y)$ and marginal X pmf $p_X(x) (>0)$, the conditional probability of Y when $X=x$ is given by $\underbrace{p_{Y|X}(y|x)}_{y \text{ given } x} = \frac{p(x,y)}{p_X(x)}$

Similarly $p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}$; $p_Y(y) > 0$

Ex: 5.1

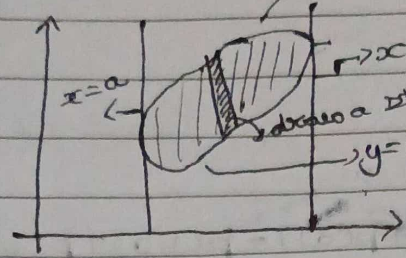
Two continuous random variables:- Let, X and Y be two crv's defined on a sample space 'S'. Then the joint probability density for $f(x,y)$ for each possible pair (x,y) in a two dimensional set A is given by:-

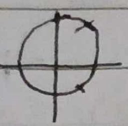
$$P((x,y) \in A) = \iint_A f(x,y)\, dy\, dx$$

It may be noted that

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

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

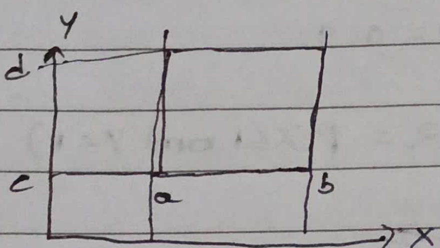




$$\int_a^b \left\{ \int_{\phi(x)}^{\psi(x)} f(x,y)\,dy \right\} dx$$

This will be a function of x which can be integrated

Here variable limit because region is not uniform



$A = \{ a \leq x \leq b,\ c \leq y \leq d \}$

$$\int_a^b \int_c^d f(x,y)\,dy\,dx$$

If A is the set of two dimensional rectangles,

$A = \{ (x,y) : a \leq x \leq b \ \&\ c \leq y \leq d \}$

then

$$\iint_A f(x,y)\,dy\,dx = \int_a^b \int_c^d f(x,y)\,dy\,dx$$

$$= \int_a^b \left\{ \int_c^d f(x,y)\,dy \right\} dx$$

It can contain both x and y.

x: Treat as a constant.

The marginal probability density function for X is given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \quad \text{for } -\infty < x < \infty$$

Similarly,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \quad \text{for } -\infty < y < \infty$$

If X and Y are independent, then

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

If X and Y are two crv with joint pdf $f(x,y)$ and marginal X pdf $f_X(x) > 0$, then the conditional probability for Y when $X = x$ is given by:-

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

Similarly,

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}; \quad f_Y(y) > 0.$$

Note

$E(X) = \sum_x x \cdot p_X(x)$, $E(Y) = \sum_y y \; p_Y(y)$ if X and Y are discrete random variable.

If X and Y are continuous,

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx \; , \; E(Y) = \int_{-\infty}^{\infty} y \cdot f_Y(y)\, dy$$

5.2

The expected values

1. If X and Y are two drv's with joint pmf $p(x,y)$, then the expected value of $h(x,y)$ is given by:-

$$E\{h(x,y)\} = \sum_x \sum_y h(x,y)\, p(x,y)$$

2. If X and Y are two crv's with joint pmf $f(x,y)$, then the expected value of $h(x,y)$ is given by:-

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

PAGE No.
DATE / /

Co-variance: If X and Y are two random variables, then the co-variance of x and Y is denoted as cov(x,Y) and is defined as::

$$cov(X,Y) = E\big(\underbrace{(x-\mu_x)}_{\text{Deviation of mean of } x}\cdot\underbrace{(Y-\mu_Y)}_{\text{Deviation of mean of } y}\big)$$

$$= \begin{cases} \sum_x \sum_y (x-\mu_x)\cdot(y-\mu_y)\, p(x,y) & \text{if X and Y are discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)\cdot(y-\mu_y)\cdot f(x,y)\, dy\, dx & \text{if X and Y are continuous} \end{cases}$$

Note-1)

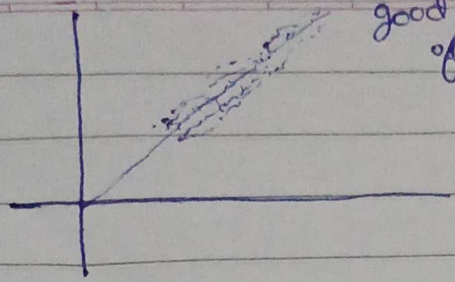
1. $cov(X,X) = E\left((x-\mu_x)^2\right) = V(x)$
2. $cov(X,Y) = E(X,Y) - E(X)\cdot E(Y)$ (short cut formula)
   $= E(XY) - \mu_X \mu_Y$ → Prove this

$$cov(X,Y) = E\{(x-\mu_x)\cdot(y-\mu_y)\}$$
$$= E\{xY - x\mu_y - Y\mu_x + \mu_x\mu_Y\}$$
$$= E(XY) - \mu_y E(X) - \mu_x E(Y) + \mu_x\mu_y E(1)$$
$$= E(XY) - \mu_x\mu_y - \mu_x\mu_Y + \mu_x\mu_y \quad [\because E(1)=1]$$
$$= E(XY) - \mu_x\mu_y$$

Correlation

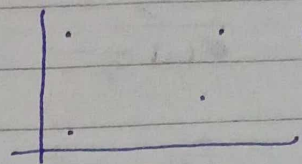
If X and Y are two random variables, then the correlation between X and Y is denoted as ($corr(X,Y)$ or ρ_{XY} or ρ) and is defined as $corr(x,Y) = \dfrac{Cov(X,Y)}{\sigma_x\sigma_y}$

σ_x and σ_y are the standard deviation of X and Y respectively.



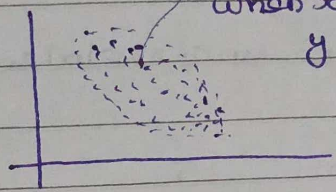
good correlation of x and y

Positive correlation (when x is increasing y increasing)



Poor correlation of x and y

Good correlation



when x is increasing, y is decreasing

negative correlation

Note.

1) For positive values of a & c / negative values of a & c

$$\operatorname{corr}(aX+b,\ cY+d) = \operatorname{corr}(X,Y)$$

2) For any two random variables X and Y,

$$-1 \leq \rho \leq 1$$

3) If X and Y are independent, then $\rho=0$ but $\rho=0$ does not imply X and Y are independent.

4) If $a \neq 0$, $\operatorname{corr}(X, aX+b) = \pm 1$, Here $Y = aX+b$

"prove this"

$$\operatorname{corr}(X, aX+b) = \frac{\operatorname{cov}(X, aX+b)}{\sigma_X\, \sigma_{aX+b}}$$

$$= \frac{E((x-\mu_x)\cdot(ax+b-\mu_{ax+b}))}{\sigma_x \cdot \sigma_{ax+b}}$$

$\mu_{ax+b} = aE(x) + b = a\mu_x + b$

$$= \frac{E((x-\mu_x)\cdot(ax+b-a\mu_x-b))}{\sigma_x \cdot \sigma_{ax+b}}$$

$\sigma_{ax+b} = |a| \sigma$

$$= \frac{a \cdot E((X-\mu_X)^2)}{\sigma_X \cdot |a| \sigma_X}$$

$$= \frac{a \cdot \sigma_X^2}{|a| \sigma_X^2}$$

$$= \frac{a}{|a|} = \frac{a}{\pm a}$$

$$= \pm 1$$

PAGE NO.
DATE / /

5.3 Statistics and their distribution
5.4 Distribution of the sample mean. } we'll discuss simultaneously.

→ Different from statistics

What is Statistic ?

Ans:- It is a quantity whose value can be calculated from the sample data. To evaluate it, we check the mathematical models (affected by chance) against the observable reality. It is usually done by sampling i.e, by drawing samples. Sample values are the set of values taken from a larger set of values known as population.

eg. ① 20 screws from a lot of 1000 screws
② 50 voters from 50,000 voters.

Defination:- The rv's $X_1, X_2, \dots, X_n$ forms a random sample if :-

i) All X_i's $(i=1,2,\dots,n)$ are independent
ii) Each X_i (for $i=1,2,\dots,n$) has same probability distribution.

The random variables $X_1, X_2, \dots, X_n$ are independent and identically distributed (iid) if the above two conditions are satisfied.

Eg: Let, 1,3,4 & 7 are the values of ~~the random varia~~ X. Consider the samples of two from the given population and verify whether the sample mean is equal to the population mean or not.

Ans:- (1,3), (1,4), (1,7), (3,4), (3,7), (4,7)

PAGE No.		
DATE	/ /	

Ans-

S.L. No	Samples	Sample mean
1	1, 3	2
2	1, 4	2.5
3	1, 7	4
4	3, 4	3.5
5	3, 7	5
6	4, 7	5.5

~~For each of these samples, we're select~~

$\therefore$ Sample mean $(\bar{x}) = \frac{1}{6}(2+2.5+4+3.5+5+5.5)$

$= 3.75$

Population mean $(\mu) = \frac{1+3+4+7}{4} = 3.75$

Note:

1. Sample mean $(\bar{x}) = \frac{1}{n}\sum_{j=1}^{n} x_j$ n is the sample size.

2. Sample variance $(\sigma_{\bar{x}}^2) = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2$

Defination:- Let, $x_1, x_2, \ldots\ldots, x_n$ be a random sample from a distribution with mean 'μ' and standard deviation 'σ', then

i) $\mu_{\bar{x}} = E(\bar{x}) = \mu$ → (Expectation of sample mean)

ii) $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ → Population variance

iii) $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

iv) If the sample total $(T_0) = x_1 + x_2 + \ldots\ldots + x_n$, then $E(T_0)$ or $\mu_{T_0} = n\mu$, $V(T_0) = \sigma_{T_0}^2 = n\sigma^2$ → Population variance

$n\mu$ ↓ Population mean

PAGE NO
DATE / /

$\sigma_{T_0} = \sqrt{n}\,\sigma$

The central limit theorem (CLT) - (when n is strictly greater than 30)

Let, $X_1, X_2, \ldots, X_n$ be a random sample with mean 'μ' and standard deviation 'σ'. Then for sufficiently large 'n', the sample mean $\bar{X}$ has approximately a normal distribution with $\mu_{\bar{X}} = \mu$, $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ and the sample total T_0 also has approximately a normal distribution with mean $\mu_{T_0} = n\mu$ & $\sigma_{T_0}^2 = n\sigma^2$

(n>30)

Note

1) CLT is applicable when n is strictly greater than 30.

2) Since for sufficiently large n, $\bar{X}$ and T_0 has approximately a normal distribution, to calculate $P(\bar{X} \le a) / P(T_0 \le a)$, we standardize the sample total and the sample mean and apply/use table A_3 to calculate the probabilities.

$$P(X \le x)$$
$$= P\left(\frac{x-\mu}{\sigma} \le \frac{a-\mu}{\sigma}\right)$$
$$= P\left(z \le \frac{a-\mu}{\sigma}\right)$$

$\therefore P(\bar{X} \le a) =$ we standardize by

$$= P\left(\frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} \le \frac{a - \mu_{\bar{X}}}{\sigma_{\bar{X}}}\right)$$
$$= P\left(Z \le \frac{a - \mu_{\bar{X}}}{\sigma_{\bar{X}}}\right)$$
$$= \Phi\left(\frac{a - \mu_{\bar{X}}}{\sigma_{\bar{X}}}\right)$$

$$P(T_0 \le a) = P\left(\frac{T_0 - \mu_{T_0}}{\sigma_{T_0}} \le \frac{a - \mu_{T_0}}{\sigma_{T_0}}\right)$$

$$= P\left(Z \le \frac{a - \mu_{T_0}}{\sigma_{T_0}}\right)$$

$$= \phi\left(\frac{a-\mu_{T_0}}{\sigma_{T_0}}\right)$$

Ex: 5.5

The distribution of a linear combination

Given a collection of random variables $X_1, X_2, \ldots\ldots, X_n$

and the numerical constants $a_1, a_2, \ldots\ldots, a_n$

Then, $Y = a_1x_1 + a_2x_2 + \ldots\ldots + a_nx_n$

$$= \sum_{i=1}^{n} a_ix_i$$

is called the linear combination of $x_1, x_2, \ldots\ldots, x_n$

Definition :- Let, $x_1, x_2, \ldots, x_n$ be the random variables with means $\mu_1, \mu_2, \ldots\ldots, \mu_n$ and variance $\sigma_1^2, \sigma_2^2, \ldots\ldots, \sigma_n^2$ respectively
then

1. whether X_i's are independent or not

$$E(Y) = E(a_1x_1 + a_2x_2 + \ldots\ldots + a_nx_n)$$
$$= a_1E(x_1) + a_2E(x_2) + \ldots\ldots + a_nE(x_n)$$
$$= a_1\mu_1 + a_2\mu_2 + \ldots\ldots + a_n\mu_n$$

In particular, $E(x_1 - x_2) = E(x_1) - E(x_2) = \mu_1 - \mu_2$

2. If X_i's are independent, then

$$V(a_1x_1 + a_2x_2 + \ldots\ldots + a_nx_n) = a_1^2V(x_1) + a_2^2V(x_2) + \ldots\ldots + a_n^2V(x_n)$$
$$= a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots\ldots + a_n^2\sigma_n^2$$

$$\sigma_{a_1x_1 + a_2x_2 + \ldots\ldots + a_nx_n} = \sigma_y = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots\ldots + a_n^2\sigma_n^2}$$

In particular,

$$V(x_1 + x_2) = V(x_1) + V(x_2) = \sigma_1^2 + \sigma_2^2$$

3) For any x_i $(i = 1, 2, \ldots\ldots, n)$

$$V(a_1x_1 + a_2x_2 + \ldots\ldots + a_nx_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \operatorname{cov}(x_i, x_j)$$

4) If $x_1, x_2, \ldots, x_n$ are linearly independent and normally distributed then the linear combination of $x_1, x_2, \ldots, x_n$ is also normally distributed