MATH4441: Probability and Statistics

1. Probability law: Sets, Probabilistic Models, Conditional Probability**, Independence*, Total probability* theorem, Baye's theorem*.

2. Discrete Random Variables: Probability Mass Function (PMF), Cumulative Distribution Function (CDF), Expectation, Variance,
Well known distribution: Uniform Distribution, Bernoulli, Binomial, Poisson

3. Continuous Random Variable: Probability Density Function (PDF), Cumulative Distribution Function (CDF), Expectation, Variance

4. Joint Random Variable: Joint PMFs, Joint PDFs, Conditional Expectation, Covariance, Correlation, Independence of Random Variables, Inferential Statistics and Probability models, populations and samples.

Ref books: Syllabus

[Quiz may contain definition but Finals will have app. based]

Probability:

The term probability is an estimate of the proportion of one or more uncertain experimental outcomes when the experiment is performed randomly.

i. What is the probability that the stock market will show an abrupt rise soon after the announcement of forthcoming budget??

A review of set & set notations:

Set: → should be a collection of well defined object
→ must have distinct values/elements.

. Universal set: The set which contains all possible outcomes.

$U = \{x : x$ is the sum of points uppermost side ~~of~~ on the two dice$\} = \{2, 3, 4, 5, 6 \dots 12\}$

Null set: Φ [subset of any set]

For any set A, $\Phi \subset A$; $A \neq \Phi$

Venn diagram:

Set algebra: Through which properties of set are defined.

Set operations: i) Union ii) Intersection iii) Complement

↳ used to prepare new sets from given states.

Union:

$\boxed{x \in A \cup B}$ iff $x \in A$ or $x \in B$ or $x \in$ both A & B

Intersection:

$\boxed{x \in A \cap B}$ iff $x \in A$ and $x \in B$

Complement:

$\boxed{x \in A^c}$ iff $x \notin A$

Two important properties of collection of sets while working with probability in real life:

1. Mutually Exclusive/Disjoint

A collection of sets $A_1, A_2, \ldots, A_n$ is mutually exclusive iff $\boxed{A_i \cap A_j = \phi}$

where $i \neq j$

2. Collectively exhaustive:

A collection of sets $A_1, A_2, \ldots A_n$ is collectively exhaustive iff $A_1 \cup A_2 \cup \ldots \cup A_n$ i.e. $\bigcup_{i=1}^{n} A_i = S$

Partition: A collection of sets $A_1, A_2, \ldots A_n$ is a partition if it is both mutually exclusive and collectively exhaustive.

De Morgan's Law: i) $(A \cap B)^c = A^c \cup B^c$

ii) $(A \cup B)^c = A^c \cap B^c$

Application of set theory in probability:

- An outcome is an observation
- An event is a set of outcomes

Outcome

An outcome of an experiment is any possible observation of that experiment.

Random experiment/ Probabilistic experiment:

An experiment that can result in different outcomes even though it is repeated in the same manner. Things to consider for random experiment:

Example

- **Procedure:** Monitor activity at a mobile store.
- **Observation:** Observe which type of mobile sets the next customer purchase
- **Model:** Android and IOS are equally likely. The result of each purchase is unrelated to the results of previous purchase.

Set theory	Probability theory
Set →	Event
Universal set →	Sample space
Element →	Outcome

Sample space:

→ Discrete Sample Space

→ Continuous Sample Space

Event has 3 operations which are Union, Intersection, Complement

if E_1 and E_2 are two events and are mutually exclusive then,

$$\Rightarrow P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

Exioms of probability:

$S \to$ Sample Space $E \to$ Event

1. $P(S) = 1$

II. $0 \leq P(E) \leq 1$

III. If mutually exclusive, $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

Laws of sets

(I) For any sets A and B, $A \cap B = B \cap A$ and $A \cup B = B \cup A$ [Commutative law]

(II) For any set A, B and C, $(A \cup B) \cup C = A \cup (B \cup C)$
$(A \cap B) \cap C = A \cap (B \cap C)$ [Associative Law]

(III) For any set A, B and C ~~A~~ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ [Distributive law]

(IV) For any given set A and any null set ϕ, universal set U

$$\left.\begin{aligned} A \cup \phi &= A \\ A \cap \phi &= \phi \\ A \cap U &= A \\ A \cup U &= S \end{aligned}\right\} \text{identity law}$$

(V) For any given set A. $A \cup A = A$ and $A \cap A = A$ [idempotent law]

(VI) For any subset ~~of~~ A of a universal set U, there is one and only complement of A, A^c that follows.

$A \cup A^c = U$

$A \cap A^c = \phi$

$(A^c)^c = A$

$P(A) + P(A^c) = 1$ $\quad$ $P(S) = 1 \Rightarrow P(A \cup A^c) = 1 \Rightarrow P(A) + P(A^c) = 1$

Joint Probability:

Two or more events form a joint event if all of them occur simultaneously and the probability of these joint events are called joint probability.

A: represents the event smoker
B: " " " heart disease

$A \cap B$ = represents a joint event describing that a randomly chosen person is a smoker who is suffering from heart disease.

Problem:

	Employment Status		
Sex	Employed(E)	Unemployed (U)	Total
Male (M)	255	20	275
Female (F)	80	145	225
Total	335	165	500

$n(S) = 500$ $\quad$ $n(M) = 275$

let us define the events as follows —

M: Randomly selected person is a male.
F: " " " " " female.
E: " " " " employed.
U: " " " " unemployed.

$n(F) = 225$ $\quad$ $n(E) = 335$ $\quad$ $n(U) = 165$

$M \cap E$: Randomly selected person is male and employed.
$M \cap U$:
$F \cap E$:
$F \cap U$:

$$P(M) = \frac{n(M)}{n(S)} = \frac{275}{500} = 0.55$$

$$P(E) = \frac{n(E)}{n(S)} = \frac{335}{500} = 0.67$$

$$P(M \cap E) = \frac{n(M \cap E)}{n(S)} = \frac{255}{500} = 0.51$$

$$P(M \cap U) = \frac{n(M \cap U)}{n(S)} = 0.04$$

$$P(F \cap E) = \frac{n(F \cap E)}{n(S)} = \frac{80}{500} = 0.16$$

$$P(F \cap U) = \frac{n \cancel{E}(F \cap U)}{n(S)} = \frac{145}{500} = 0.29$$

$P(M) = P(M \cap E) + P(M \cap U)$ [sum of disjoint event]

$$= 0.51 + 0.44 = 0.55$$

Problem:

$n(S) = 100$; $n(E) = 75$; $n(B) = 50$ $n(E \cap B) = 40$

E : reads English newspaper

B : " Bengali "

$E \cap B$: " both

a) Read English newspaper probability.

b) Reads none of the newspaper.

c) Reads English but not Bengali.

(a) $P(E) = \frac{n(E)}{n(S)} = \frac{75}{100} = 0.75$

(b) $P(E^c \cap B^c) = \cancel{\frac{85}{100}}\ 1 - \frac{85}{100}$
$= 0.15$

(c) $P(E \cap B^c) = \frac{35}{100} = 0.35$

	E	E^c	total
B	$B \cap E = 40$	$B \cap E^c$ 10	50
B^c	$B^c \cap E$ 35	$B^c \cap E^c$ 15	50
Total	75	25	100

Conditional Probability:

Conditional Probability correspond to a modified probability model that reflects partial information about the outcome of an experiment. ~~The~~ This modified model has a reduced sample space than the original model.

Definition: The probability of an event E when it is known that some other event F has occured is called a conditional probability of E and is denoted with P(E/F) and defined as ~~probability of~~

$$\boxed{P(E/F) = \frac{P(E \cap F)}{P(F)} \; ; \; P(F) > 0}$$

$P(EF) = P(E \cap F)$

Similarly, $P(F/E) = \frac{P(F \cap E)}{P(E)} = \frac{P(E \cap F)}{P(E)} \; ; \; P(E) > 0$

$\boxed{P(E/F) \cdot P(F) = P(F/E) \cdot P(E)}$] multiplication theorem/multiplication law/ law of compound probability.

i.e. $P(E \cap F) = \underbrace{P(E)}_{\text{undonditional}} \cdot \underbrace{P(F/E)}_{\text{conditional}}$ = The probability of two events taking place happening at the same time is equal to the product of the probability of unconditional and conditional probability.

$= P(F) \cdot P(E/F)$

For 3 events A, B, C: $P(A \cap B \cap C) = P(A) \cdot P(B/A) \cdot P(C/A \cap B)$

For n num of events, $A_1, A_2 \ldots A_n$, $P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \cdots$

$P(A_n / A_1 \cap A_2 \ldots \cap A_{n-1})$

26.11.2021

Problem:

A family has two children. What is their conditional probability that both are boys. Given that at least both are boys given that one of them at least is a boy?

Sol^n: Let, the sample space is given by:

$S = \{(b,b), (b,g), (g,b), (g,g)\}$ where all outcomes are equally likely.

E: The event that both children are boys.

F: The event that at least one of the children is a boy.

Here, $E = \{(b,b)\}$ $F = \{(b,b), (b,g), (g,b)\}$

$E \cap F = \{(b,b)\}$

Hence the required probability is, $P(E/F) = \dfrac{P(E \cap F)}{P(F)}$

$$= \frac{\frac{n(E \cap F)}{n(S)}}{\frac{n(F)}{n(S)}} = \frac{1/4}{3/4}$$

$$= 1/3$$

Scanned with CamScan

b) The probability that a married woman watches a TV show is 0.5, and that of her husband is 0.4.

The probability that a man watches the show given that his wife does is 0.7. Then find the following probabilities:

(a) The probability that the married couple watches the show.

(b) a wife watches the show given that her husband does.

(c) at least one of the partner watch the show.

Sol^n: Let us define the event as follows;

H: Husband watches the show.

W: Wife " " " .

According to Question,

$$P(H) = 0.4 \quad P(W) = 0.5 \quad P(H/W) = 0.7$$

(a) $P(H \cap W) = P(H/W) \cdot P(W) = 0.35$

(b) $P(W/H) = 0.35/0.4 = 0.875$

(c) $P(H \cup W) = P(H) + P(W) - P(W \cap F) = 0.9 - 0.35$

$= 0.55$

Problem:

A box contains 7 red balls and 3 black balls. 3 balls are drawn from the box one after another. Find the probability that the 1st 2 balls are red and 3rd one is black? if

Red = 7
Black = 3

(i) The balls are replaced before the next draw.

(ii) The balls are not replaced.

Solution:

Let us define the event as follows:

R: The event of drawing red balls.

B: The " " " black balls.

i) If a ball is replaced before the next draw then the subsequent draws are not affected as the number of balls in the box remains the same. So this is a case of unconditional probability.

R_1: the first ball is red.

R_2: " second " " red.

B_3: " third " " black.

$$P(R_1 \cap R_2 \cap B_3) = P(R_1) \cdot P(R_2) \cdot P(B_3)$$

$$= \cancel{7 \times 7} \; \frac{7}{10} \cdot \frac{7}{10} \cdot \frac{3}{10}$$

$$= \frac{147}{1000}$$

ii) Since the balls will not be replaced after each draw so the number of balls will decrease after each subsequent drawing. So this is a conditional probability case.

$$P(R_1 \cap R_2 \cap B_3) = P(R_1) \cdot P(R_2 | R_1) \cdot P(B_3 | R_1 \cap R_2)$$

$$= \cancel{\frac{7}{10} \cdot \frac{6}{9} \cdot \frac{3}{8}} = \cancel{\frac{7}{40}}$$

$$= \frac{7}{10} \cdot \frac{6}{9} \cdot \frac{3}{8} = \frac{7}{40}$$

4) A coin is tossed until a head appers or it has been tossed 3 times given that the head doesn't appear on the 1st toss then what is the probability that the coin is tossed 3 times?

Solution:

$S = \{H, TH, TTH, TTT\}$

$P(H) = P(T) = 1/2$

$P(TH) = 1/4$

$P(TTH) = 1/8 \quad P(TTT) = 1/8$

H: The event that head comes

T: The event that tail comes

A: coin is tossed 3 times

B: no head appears on the 1st toss.

$A = \{TTH, TTT\}$

$B = \{TH, TTH, TTT\}$

$\therefore A \cap B = \{TTH, TTT\}$

$P(A) = \frac{1}{8} + \frac{1}{8} = 1/4$

$P(B) = \frac{1}{4} + \frac{1}{8} + \frac{1}{8} = 1/2$

$P(A \cap B) = \frac{1}{8} + \frac{1}{8} = 1/4$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = 1/4 \times 1/2 = 1/2$$

Problem Three cards are drawn in succession without replacement from an ordinary deck of playing card. Find the probability that the first card is a red ace, the second card is a 10 or a Jack and the third card is greater than 3 but less than 8.

Soln:

A : the first card is a red ace -

B : " second " " " Jack or 10.

C : " third " " greater than 3 but less than 8.

$P(A) = \frac{2}{52}$ $P(B/A) = 8/51$ $P(C/A \cap B) = \boxed{16/50}$ A,

$P(A \cap B \cap C) = \frac{2}{52} \cdot \frac{8}{51} \cdot 16/50 =$

Independent Events: If E and F are two events and if the occurence of E doesn't affect and isn't affected by the occurence of F then E and F are said to be independent events.

$$\boxed{P(P \cap F) = P(E) \cdot P(F)}$$

From conditional property we know,

$$P(E/F) = \frac{P(E \cap F)}{P(F)} = \frac{P(E) \cdot P(F)}{P(F)} = P(E)$$

$$\boxed{P(E/F) = P(E)}$$

$$\boxed{P(F/E) = P(F)}$$

Two ideal coins are tossed :

A: head on the 1st coin

B: head on the 2nd coin.

Sol^n: Sample space, $S = \{HH, HT, TH, TT\}$

$A = \{HH, HT\}$

$B = \{HH, TH\}$

$P(A) = \frac{2}{4} = 1/2$

$P(B) = \frac{2}{4} = 1/2$

$\cancel{P(A) =}$ $(A \cap B) = \{HH\}$

$P(A \cap B) = 1/4$

$P(A) \cdot P(B) = 1/2 \times 1/2 = 1/4$

$= P(A \cap B)$

$\therefore$ The events are independent.

$\boxed{\text{Problem:}}$

Three coins are tossed, show that the events "head on the first coin" and the event "tails on the last two coins" are independent.

	HH	HT	TH	TT
H	HHH	HHT	HTH	$\boxed{\text{HTT}}$
T	THH	THT	TTH	TTT

A: heads on the first coin ✓

B: tails on the last two coins ✓

→ Ans

Problem:

Suppose we toss two fair dice. Let E_1 denotes the event that the sum of the dice is six and F denotes the event that the first dice equals four. Check whether the events E_1 & F are independent or not.

$E_1 = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$

$P(E_1) = 5/36$

$F = \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$ $P(F) = 1/6$

$E_1 \cap F = \{(4,2)\}$ $P(E_1 \cap F) = 1/36$

$P_1(E_1) \cdot P(F) = \frac{5}{36} \times \frac{1}{6}$

$= \frac{5}{216} \neq P(E_1 \cap F)$

So the events are not independent.

E_2 = sum of two dices are seven

$E_2 = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

$P(E_2) = 1/6$

$P(E_2) \cdot P(F) = 1/6 \times 1/6 = 1/36$

$E_2 \cap F = \{(4,3)\}$ ~~$1/36$~~

$P(E_2 \cap F) = 1/36$

$\boxed{\text{Problem:}}$

A fire brigade has 2 fire engines operating independently. The probability that a specific fire engine is available when needed is 0.99 then

ⓐ What is the probability an engine is available when needed?

ⓑ What is the probability neither engine 〃 〃 〃 〃

Soln: Let us define the events as follows:

A: The event that first engine is available when needed.

B: 〃 〃 〃 second 〃 〃 〃 〃 〃

Given, $P(A) = P(B) = 0.99$ //

$P(A \cap B) = P(A) * P(B) = 0.99 \times 0.99 = 0.9801$

(a)

$P(A \cup B) = P(A) + P(B) - P(A \cancel{\cup} \cap B) = 0.99 + 0.99 - 0.9801$

$= 0.9999$

(b)

$P(A^c \cap B^c) = (P(A \cup B)^c) = 1 - P(A \cup B)$

$= 1 - 0.9999$

$= 0.0001$

$P(A_1 \cap A_2 \cap A_3) \neq P(A_1) \cdot P(A_2) \cdot P(A_3)$

$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$

$P(A_2 \cap A_3) = P(A_2) \cdot P(A_3)$

$P(A_1 \cap A_3) = P(A_1) \cdot P(A_3)$

pairwise independent

Q) Two coins are tossed :-

A: "heads on the first coin"

B: "heads on the second coin"

C: "coins fall alike"

Check whether they are pair-wise independent or completely independent.

$A = \{(H,H), (H,T)\}$ $B = \{(H,H), (T,H)\}$

$C = \{(H,H), (T,T)\}$

$P(A) \cdot P(B|A) \cdot P(C|A \cap B)$
$= \frac{1}{2} \cdot \frac{1}{2} = 1/4$

$P(A) = 2/4 = 1/2$ $P(B) = 1/2$ $P(C) = 1/2$

$P(A \cap B) = 1/4$ $P(B \cap C) = 1/4$ $P(A \cap C) = 1/4$

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

but, $P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C)$

$\therefore$ It is pairwise independent

how?

Conditional Probability and Partition:

Let S denotes the sample space of some random experiments and consider n events $A_1, A_2, \dots A_n$ such that they are mutually exclusive and collectively exhaustive i.e. $A_i \cap A_j = \phi \ \forall i \neq j$ and $S = A_1 \cup A_2 \cup A_3 \dots \cup A_n$

It is said that these events form a partition of S.

Total Probability Theorem/Rule:

Suppose that the events $A_1, A_2, \dots, A_n$ make partition of the sample space S, and $P(A_j) > 0$ for $j = 1, 2, \dots n$ Then for any event $B \in S$;

$$\boxed{P(B) = \sum_{j=1}^{n} P(A_j) \cdot P(B|A_j)}$$

Bayes' theorem:

Application: Let us consider n mutually exclusive and collectively exhaustive events $A_1, A_2, \dots A_n$ and B be any event,

If $P(B|A_1), P(B|A_2), P(B|A_3) \dots P(B|A_n)$ are known then Bayes' theorem is useful to complete the conditional probabilities of A_j events given B,

$P(A_1|B), P(A_2|B) \dots P(A_n|B)$

Prior probability: আগে থেকেই জানা সেই probability.

Posterior " : prior probability use করে সেই new probability বের করা হয়,

---

Bayes' theorem:

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

Problem: In a factory 3 different machines M_1, M_2, M_3 were used for producing large batch of similar manufactured items. Suppose that M_1, M_2 and M_3 produce 25%, 35% and 40% of the total items respectively. Suppose further that their items, respectively 5%, 4% and 2% are defective. Finally suppose that one item is selected at random from the entire batch and it is found to be defective. Determine the probability that this item is produced by machine M_1 (M_2 or M_3)

Solution:

Let us define the events as follows:

A_1: the selected item is produced by M_1.

A_2: " " " " " " M_2.

A_3: " " " " " " M_3.

B : " " " " defective.

Given that,

$P(A_1) = 25\% = 0.25\%$ $P(A_2) = 35\% = 0.35$ $P(A_3) = 40\% = 0.4$

Also,

$P(B|A_1) = 5\% = 0.05$ $P(B|A_2) = 4\% = 0.04$ $P(B|A_3) = 2\% = 0.02$

Using Bayes' theorem, the required probability that the selected defected item is produced by M_1 is:

$$P(A_1|B) = \frac{P(A_1)\,P(B|A_1)}{\sum_{j=1}^{3} P(A_j)\cdot P(B/A_j)} = \frac{0.25 \times 0.05}{(0.25\times0.05)+(0.35\times0.04)+(0.4\times0.02)}$$

$$= \frac{0.0125}{0.0125+0.014+0.008}$$

$$= \frac{0.0125}{0.0345}$$

$$\doteq 0.36$$

(b) $P(A_2 \cup A_3 | B) = P(A_2|B) + P(A_3|B)$

$$= \frac{P(A_2)\cdot P(B|A_2)}{\text{Total Probability}} + \frac{P(A_3)\cdot P(B|A_3)}{\text{Total probability}}$$

$$= \frac{P(A_2)\cdot P(B|A_2) + P(A_3)\cdot P(B|A_3)}{\text{Total Probability}}$$

$$= \frac{0.35\times0.04 + 0.4\times0.02}{0.0345} = \frac{0.014+0.008}{0.0345} = 0.64$$

Problem:

P_1 থেকে produce করে 2000

~~B.~~ P_2 " " " 3000

P_1 : 10% defective P_2 : 15% defective

Problem: Three urns contain

U_1 has 2 red balls, 4 black balls

U_2 has 3 " " 1 " ball

U_3 has 3 " " 4 black balls

One urn is chosen at random and then two balls are randomly drawn from the selected urn. If the selected balls are different in colour, what is the probability that the balls come from urn 2?

U_1 : total 6 $\quad A_1 : U_1$ is chosen $\quad$ B: the selected balls are different in colour

U_2 total 4 $\quad A_2 : U_2$ is chosen

U_3 : total 7 $\quad A_3 : U_3$ is chosen

$A_2 | B$

$$P(A_1) = P(A_2) = P(A_3) = 1/3$$

$$P(B|A_1) = \frac{\binom{2}{1}\binom{4}{1}}{\binom{6}{2}} = \frac{2\times 4}{15} = 8/15 \qquad P(B|A_2) = 1/2$$

$$P(B|A_3) = 4/7$$

$$P(A_2|B) = \frac{P(A_2)\,P(B|A_2)}{\sum_{n=1}^{3} P(A_j)\cdot P(B|A_j)} = \frac{1/3 \times 1/2}{\frac{1}{3}\times\frac{8}{15} + 1/3\times\frac{1}{2} + 1/3 \times 4/7}$$

$$= \frac{\frac{1}{6}}{\frac{8}{45} + \frac{1}{6} + \frac{4}{21}}$$

$$= \frac{105}{337}$$

$$= 0.31$$

Problem : Principal Selection

This year suppose that there will be three candidates for the post of principal in NDC. They are father Costa, Father Hemonto and Father Adam. The chances that they get the post are 4:2:3. The probability that Father Costa if gets selected will introduce co-education in college is 0.3. The probability that Fr. Hemonto and Fr. Adam doing the same are 0.5 and 0.8, respectively. What is the probability that there will be co-education in the college this year. If co-education will be introduced, what will be the chance that it will be introduced, what will be the chance that it will be introduced by Fr. Costa?

Solution:

A_1 : Fr. Costa will be selected as Principal

A_2 : Fr. Hemata " " " " "

A_3 : Fr. Adam " " " " "

B : Co-education will be introduced.

Given,

$$P(A_1) = 4/9 \quad P(A_2) = 2/9 \quad P(A_3) = 3/9$$

Also $P(B|A_1) = 0.3 \quad P(B|A_2) = 0.5 \quad P(B|A_3) = 0.8$

Now,

$$P(B) = \sum_{j=1}^{3} P(A_j) \cdot P(B|A_j)$$

$$= P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + P(A_3) \cdot P(B|A_3)$$

$$= \left(\frac{4}{9} \times 0.3\right) + \left(\frac{2}{9} \times 0.5\right) + \left(\frac{3}{9} \times 0.8\right)$$

$$= \frac{23}{45} \text{ (Ans)}$$

Using Baye's theorem,

$$P(A_1|B) = \frac{P(A_1) \cdot P(B|A_1)}{P(B)} = \frac{4/9 \times 0.3}{23/45} = 6/23 \text{ (Ans)}$$

Chap-2 and 3

Random variable (R.v)/(r.v)

A ~~re~~ variable whose values are any definite numbers or quantities that arise as a result of chance factors such that they cannot be predicted in advance is called a random variable (r.v).

A ~~real~~ r.v is a real-valued function defined over a sample space.

Random variable
- discrete random variable
- Continuous random variable

Examples:

discrete random variable	Continuous random variable
1. no. of telephone cells recieved in a telephone booth	~~in a cte~~ ① Height, Weight, Time of exam, Exam score, Temp.
2. no. of correct ans. in MCQ	② Rate of bank interest
3. no. of defective bulbs/fans	
4. no. of children in a family	

Probability Distribution:

The idea of probability distribution exactly parallels to that of a frequency distribution. Each type of distributions is based on a set of mutually exclusive and collectively exhaustive measurement classes or class intervals.

Frequency Distribution: Class-interval → frequency } N →

Probability Distribution: Class-interval → probability values } 1

↳ hypothetical distribution

theoritical distribution

Probability Distribution:

Any statement of a function associating each of a set of mutually exclusive and collectively exhaustive classes or class intervals with its probability is a probability distribution. (P.d).

2 types:-

Discrete Probability Distribution:

A discrete random variable assumes each of its values on numbers with a certain probability. Sometimes these numbers are equally likely but sometimes some of the possible values of the rv are more likely to occur than others.

In either case, the func. that gives the probability associated with each possible values of the rv will result in a probability distribution of the r.v.

Example:

A coin is tossed 3 times

$S = \{HHH, HHT, HTH, HTT, THT, THH, TTH, TTT\}$

x: A discrete r.v. defining the no. of heads.

$X : x$	0	1	2	3
$P(X=x) = f(x)$	$1/8$	$3/8$	$3/8$	$1/8$

When $P(x=2) = 3/8$

{No. of heads, probability of x}

$\rightarrow (x, f(x)) = \{(0, 1/8), (1, 3/8), (2, 3/8), (3, 1/8)\}$

$\Rightarrow$ Probability Mass Function (PMF) of a discrete random variable.

DRV · Def^n:

If a r.v X has a discrete distribution, the probability distribution of x is defined as the function "f" such that for any real number x,

$$f(x) = P(X = x)$$

It is important to note that all functions are not PMFs ∴

Conditions for being a $\boxed{\text{PMF}}$:

1. $f(x) \geq 0$

2. $\sum_{x} f(x) = 1$

3. $P(X = x) = f(x)$

Problem:

Verify whether the following functions are PMFs or not:

a. $f(x) = \frac{2x-1}{8}$; $x = 0, 1, 2, 3$

b. $f(x) = \frac{x+1}{16}$; $x = 0, 1, 2, 3$

c. $f(x) = \frac{3x+6}{21}$; $x = 1, 2$

Solution:

a. Given, $f(x) = \frac{2x-1}{8}$; $x = 0, 1, 2, 3$

$X : x$	0	1	2	3
$P(X=x) = f(x)$	$-\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{5}{8}$

i. $f(x) \geq 0$ $\quad f(0) = -1/8 < 0$

but, $\sum_{x} f(x) = f(0) + f(1) + f(2) + f(3)$

$$= -1/8 + 1/8 + 3/8 + 5/8 = 1$$

Ex :

Total $T = 7$; Male $= 4$; Female $= 3$

Committee : 2

Solution:

x : A discrete random variable defying the number of males

Events	Sequence of events	$X = x$
E_1	Male, Male	2
E_2	Male, Female	1
E_3	Female, Female	0

$\therefore f(0) = P(X = 0)$

$$= \frac{{}^4C_0 \times {}^3C_2}{{}^7C_2} = 3/21$$

$f(1) = P(X = 1)$

$$= \frac{{}^4C_1 \times {}^3C_1}{{}^7C_2} = \cancel{9/9} \; \cancel{4/21} \; \frac{4}{7}$$

$$f(2) = \frac{{}^4C_2 \times {}^3C_0}{{}^7C_2} = 6/21$$

$\boxed{\text{Example}}$ A bag contains 10 balls of which 4 are black. If 3 balls are drawn at random without replacement, obtain the probability distribution for the number of black balls drawn.

Solution:

Let, the random variable X denotes the number of black balls drawn.

Clearly, X can assume the values 0, 1, 2, 3

To obtain the probability distribution of X we need to compute the probabilities associated with 0, 1, 2, 3.

Since 3 balls are to be chosen the number of ways in which this choice can be made is $^{10}C_3$.

Thus, $f(0) = \dfrac{^4C_0 \times {^6C_3}}{^{10}C_3} = \dfrac{20}{120} = 1/6$

$f(1) = \dfrac{^4C_1 \times {^6C_2}}{^{10}C_3} = \dfrac{60}{120} = 1/2$

$f(2) = P(X=2) = \dfrac{^4C_2 \times {^6C_1}}{^{10}C_3} = 36/120 = 9/30 = 3/10$

$P(3) = P(X=2) = \dfrac{^4C_3 \times {^6C_0}}{^{10}C_3} = \dfrac{4}{120} = 1/30$

$X : x$	0	1	2	3
$P(X=x) = f(x)$	$\frac{20}{120}$	$\frac{60}{120}$	$\frac{36}{120}$	$\frac{4}{120}$

Now, if we wish to check whether the distribution is PMF or not, we have to check $\sum f(x) = 1$; $f(x) \geq 0$ for all x

$$\sum_{x=0}^{3} f(x) = f(0) + f(1) + f(2) + f(3) = 1 \quad \therefore \text{PMF}$$

Example:

The probability function of a d.r.v X is given by

$$f(x) = \alpha \cdot \left(\tfrac{3}{4}\right)^{x}; \quad x = 0, 1, 2 \ldots \infty$$

$$= 0 \quad ; \text{elsewhere}$$

Find α and evaluate $P(X \leq 3)$

Solution:

Since,

$f(x)$ is a probability function, $\sum_{x=0}^{\infty} f(x) = 1$

$f(0) = \alpha \cdot \left(\frac{3}{4}\right)^0 = \alpha$

$f(1) = \alpha \cdot \frac{3}{4}$

$f(2) = \alpha \cdot \left(\frac{3}{4}\right)^2$

$f(3) = \alpha \cdot \left(\frac{3}{4}\right)^3$

$\vdots$

$1 + x + x^2 + \ldots$

$S_\infty = \frac{a}{1-r}$

$|r| < 1$

$$\Rightarrow \alpha + \alpha \cdot \tfrac{3}{4} + \alpha \cdot \left(\tfrac{3}{4}\right)^2 + \ldots = 1$$

$$\Rightarrow \alpha \left[1 + \tfrac{3}{4} + \left(\tfrac{3}{4}\right)^2 + \ldots\right] = 1$$

$$= \alpha \cdot \left[\frac{1}{1 - 3/4}\right] = 1 \Rightarrow \alpha = 1/4$$

Hence,

the complete PMF of x is

$$f(x) = \frac{1}{4}\left(\frac{3}{4}\right)^x \quad ; x = 0, 1, 2, \ldots \infty$$

$$= 0 \text{ ; elsewhere}$$

Also,

$$P(X \leq 3) = f(0) + f(1) + f(2) + f(3)$$

$$= \alpha\left[1 + \frac{3}{4} + \left(\frac{3}{4}\right)^2 + \left(\frac{3}{4}\right)^3\right]$$

$$= \frac{1}{4}\left[1 + \frac{3}{4} + \frac{9}{16} + \frac{27}{64}\right]$$

$$= \frac{175}{256}$$

Example

The drv Y has a PMF as shown below:

Y : y	-3	-2	-1	0	1
P(Y=y) = f(y)	0.10	0.25	0.30	0.15	K

1. Find the value of k
2. $P(-3 < Y < 0)$
3. $P(Y \geq -1)$

1. $f(-3) + f(-2) + f(-1) + f(0) + f(1) = 1$

   $k = 0.2$ (Ans)

2. $f(-2) + f(-1) = 0.55$ (Ans)

3. ~~f(1) = 0.2~~ $f(-1) + f(0) + f(1)$

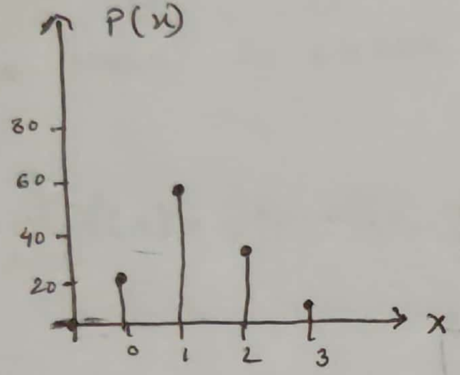
   $= 0.3 + 0.15 + 0.2$

   $= 0.65$ (Ans)

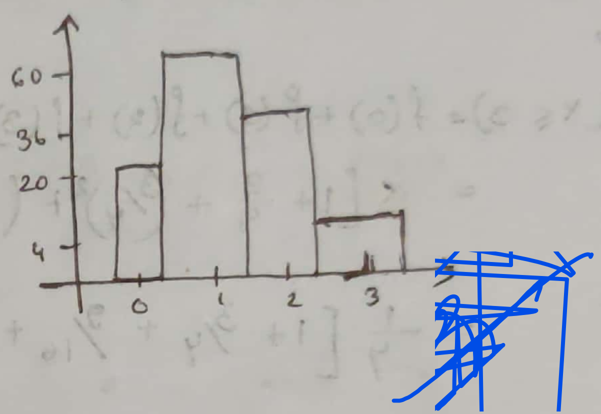
Example:

$X : x$	0	1	2	3
$P(X=x)=f(x)$	$\frac{20}{120}$	$\frac{60}{120}$	$\frac{36}{120}$	$\frac{4}{120}$

Graphical representation:



using histogram:



Cumulative Distribution Function (CDF):

The cumulative distribution function CDF or simply the distribution function $F(x)$ of a discrete random variable X with probability function $f(x)$ defined over all real numbers x is the cumulative probability upto and including the point x.

Symbolically, it is defined as $\boxed{F(x) = P(X \leq x) = \sum_{t \leq x} f(t)}$

$F(x)$ is always a monotonic increasing function. i.e.

$F(a) \leq F(b)$ for $a \leq b$

$$\left.\begin{aligned} \lim_{x \to -\infty} F(x) &= 0 \\ \lim_{x \to +\infty} F(x) &= 1 \end{aligned}\right\}$$

Example:

A coin is tossed 3 times. If X is the random variable representing the no. of heads obtained. Find the probability distribution of X and hence obtain $F(x)$.

Solution:

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

$X: x$	0	1	2	3
$P(X=x) = f(x)$	$1/8$	$3/8$	$3/8$	$1/8$

X : no. of head obtained

$x = 0, 1, 2, 3$

$$F(x) = \sum_{t=0}^{3} f(x)$$

$$= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8}$$

$$= 1$$

$F(0) = f(0) = 1/8$

$F(1) = f(0) + f(1) = F(0) + f(1) = 4/8$

$F(2) = F(1) + f(2) = 7/8$

$F(3) = 1$

Therefore,

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/8, & 0 \leq x \leq 1 \\ 4/8, & 1 \leq x < 2 \\ 7/8, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

?

Continuous Probability Distribution

The probability distribution of a continuous random variable
cannot be presented in a tabular form, because ~~of~~ a continuous
variable can take on a non-innumerable, infinite set of values.
More precisely, a continuous variable can take on any value in the
given interval $a \leq X \leq b$. As a result a continuous random
variable has a probability zero of assuming exactly one of its
values,

This implies $\boxed{P(a \leq X \leq b)} = P(\cancel{X = a}) + P(a < X < b) + P(\cancel{X = b})$

$$\boxed{= P(a < X < b)}$$

Probability Density Function (PDF): A PDF is a non-negative ~~and is~~ function and is constructed so that the area under its curve bounded by x-axis is equal to unity when computed over the range x, for which $f(x)$ is defined.

Conditions for being a PDF:

i. $f(x) \geq 0$

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

iii. $P(a < x < b) = \int_a^b f(x)dx$.

Example:

A random variable X has the following functional form:

$$f(x) = kx \quad ; \quad 0 < x < 4$$
$$= 0 \quad , \text{ elsewhere}$$

i. Find the value of k.

ii. Find $P(1 < x < 2)$ and $P(X > 2)$

Soln:

i) For $f(x)$ to be a PDF, we must have $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\Rightarrow \int_0^4 kx\,dx = 1$$

$$\Rightarrow k\left[\frac{x^2}{2}\right]_0^4 = 1$$

$$\Rightarrow 8k = 1 \quad \therefore k = 1/8$$

The complete PDF is $f(x) = 1/8\, x \quad ; \quad 0 < x < 4$

$= 0, \text{ elsewhere}$

ii. $P(1 < x < 2) = \int_1^2 1/8\, x\,dx = \frac{1}{16}[x^2]_1^2 = 3/16$

$P(x > 2) = \int_2^4 f(x)dx = 1/8 \int_2^4 x\,dx = 1/16\,[x^2]_2^4 = 3/4$

Example: A continuous random variable x has the following density function

$$f(x) = \left.\begin{array}{ll} \frac{2}{27}(1+x) \; ; & 2<x<5 \\ 0 \; ; & \text{elsewhere} \end{array}\right\} \text{piecewise function}$$

a) Verify that ~~if~~ it satisfies the condition $\int_{-\infty}^{\infty} f(x)\,dx = 1$

b) Find $P(x<4)$ and c) Find $P(3<x<4)$

~~CDF~~ a) $\int_2^5 \frac{2}{27}(1+x)\,dx = \frac{2}{27}\int_2^5 (1+x)\,dx$

$$= \frac{2}{27}\left[x+\frac{x^2}{2}\right]_2^5$$

$$= \frac{1}{27}\left(2x+x^2\right]_2^5$$

$$= \frac{1}{27}\left(35-8\right)$$

$$= \frac{1}{27}\times 27 = 1$$

b) $\int_2^4 \frac{2}{27}(1+x)\,dx = \frac{2}{27}\left[x+\frac{x^2}{2}\right]_2^4 = \frac{1}{27}\left[2x+x^2\right]_2^4$

$$= \frac{1}{27}\left[24-8\right)$$

$$= 16/27$$

c) $\int_3^4 \frac{2}{27}\,\cancel{t}\,(1+x)\,dx = \frac{1}{27}\left[2x+x^2\right]_3^4 = \frac{1}{27}\left(24-15\right) = 9/27$

$$= 1/3$$

Scanned with CamScan

CDF for continuous variable:

The cumulative distribution or distribution function, $F(x)$ of a continuous random variable X with density function $f(x)$ is defined as

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

, If the derivative of $F(x)$ exists, then

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

$$f(x)\,dx = d(F(x))$$

$f(x)\,dx$ → PDF, $d(F(x))$ → CDF

The CDF, $F(x)$ possesses the following properties,

i. $F'(x) = f(x) > 0$

ii. $F(-\infty) = 0$

iii. $F(+\infty) = 1$

iv. $F(x)$ is is defined at every point in a continuous range and is continuous.

$$P(a < X < b) = \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx$$

$$= F(b) - F(a)$$

Cumulative distribution curves often have more or less the S shape pattern.

Example:

If X has the density function

$$f(x) = \frac{2}{27} \; ; \; 2 < x < 5$$

$$= 0 \; ; \; \text{elsewhere}$$

Obtain the distribution function and hence find $F(3)$ and $F(4)$.

Also verify that $P(3 < x < 4) = F(4) - F(3)$

Solution:

The CDF of the given random variable is

$$F(x) = \frac{2}{27} \int_{t=2}^{x} (1+t)\,dt \quad = \frac{2}{27} \left[t + \frac{t^2}{2} \right]_2^x$$

$$F(x) = \frac{1}{27} \left[x^2 + 2x - 8 \right]$$

for $x = 3$ and $x = 4$

$F(3) = 7/27$ $\quad F(4) = \frac{16}{27}$ $\quad$ Hence, $F(4) - F(3) = 1/3$

Also, $P(3 < x < 4) = \int_3^4 f(x)\,dx = \frac{2}{27} \int_3^4 (1+x)\,dx = 1/3$

Example: Suppose X is a random variable with density function.

$$f(x) = \frac{k}{(1+x)^2}, \quad x > 0$$
$$= 0; \text{ elsewhere}$$

a. Find the value of k.

b. Find $F(x)$. c. Find $P(X < 1)$ d. Find $P(X > 1)$

Solution:

(a) Since, $f(x)$ is a density function, it must satisfy

$$\int_0^\infty f(x)dx = 1$$

$$\Rightarrow \int_0^\infty \frac{k}{(1+x)^2}dx = 1$$

$$\Rightarrow k\left[-\frac{1}{1+x}\right]_0^\infty = 1$$

$$\Rightarrow k\left[\frac{1}{1+x}\right]_\infty^0 = 1$$

$$\Rightarrow k(1-0) = 1$$

$$\Rightarrow k = 1$$

Thus the complete density function is $f(x) = \frac{1}{(1+x)^2}; \ x > 0$

(b) $$F(x) = P(X \leq x) = \int_{t=0}^{x} \frac{1}{(1+t)^2}dt = \left[-\frac{1}{1+t}\right]_0^x = \left[\frac{1}{1+t}\right]_x^0$$

$$= 1 - \frac{1}{1+x} = \frac{x}{1+x}$$

Scanned with CamScan

(c) $P(x<1) = F(1) = \frac{1}{1+1} = 1/2$

(d) $P(x>1) = \int_1^\infty f(x)dx = \int_0^\infty f(x)dx - \int_0^1 f(x)dx$

$$= \int_0^\infty \frac{1}{(1+x)^2}dx - \int_0^1 \frac{1}{(1+x)^2}dx$$

$$= \left[-\frac{1}{1+x}\right]_0^\infty - \left[\frac{1}{1+x}\right]$$

$$= \left[\frac{1}{1+x}\right]_\infty^0 - 1/2$$

$$= 1 - 0 - 1/2$$

$$= 1/2 \text{ (Answer)}$$

Example (HW):

find the constant k so that the following function may be a density function, $f(x) = 1/k;\ a \leq x \leq b$

$= 0;$ elsewhere

$$\int_0^\infty \frac{1}{k}dx = 1$$

$$\left[\frac{x}{k}\right]_a^b = 1$$

$$\Rightarrow k = b-a$$

Find the CDF of the random variable X.

$$f(x) = \frac{1}{b-a}$$

$$\int_{t=0}^{x} \frac{1}{b-a}dt = \frac{1}{b-a}[t]_0^x = \frac{x}{b-a}$$

Start

Joint Probability Distribution:

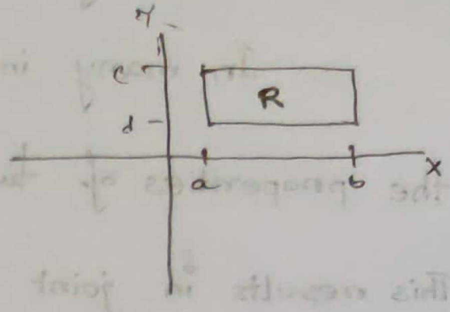
In many instances it is necessary to find or consider the properties of two or more random variables simultaneously. This results in joint probability distributions and when these distributions involve two variables they are referred to as bivariable bivariate probability distributions. Clearly, bivariate probability distribution is a special case of joint probability distributions for two random variables.

Joint probability distribution for discreet random variable

Suppose that, a given experiment involves 2 random variables X and Y each of which has a discreet probability distribution. It is customary to refer to $f(x,y)$ as the joint probability distribution of X and Y.

Since X and Y are discrete, $f(x,y) = P(X=x, Y=y)$. That is $f(x,y)$ gives the probability that the outcomes x and y at the same time.

The function $f(x,y)$ will be called a joint probability distribution of the random variables X and Y if it possesses the following properties:



i. $f(x,y) \geq 0$; for all (x,y)

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

iii. $P[(X,Y) \in R] = \sum\sum_{(x,y)\in R} f(x,y)$ for any region R in the xy plane.

$R = (a,b) \times (c,d)$

Example: A coin is tossed three times. If X denotes the no. of heads and Y denotes the no. of tails in the last two tosses. Find JPD of X and Y.

Solution:

The outcomes of the experiment and its associated probabilities are shown in the following table:

Outcome	X	Y	$P(X,Y)$
HHH	3	0	1/8
HHT	2	1	1/8
HTH	2	1	1/8
HTT	1	2	1/8
THH	2	0	1/8
THT	1	1	1/8
TTH	1	1	1/8
TTT	0	2	1/8

Scanned with CamScan

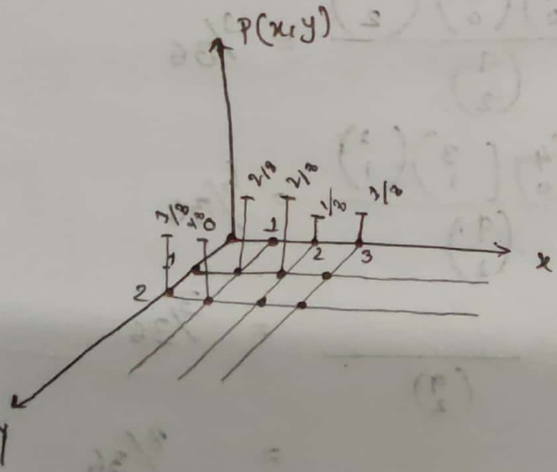
It is easy to observe that X assumes ~~variable~~ values $(x = 0, 1, 2, 3)$ and Y assumes values $(y = 0, 1, 2)$

So, JPD can now be put in the following table:

Solution:

	X values				
Y-values	0	1	2	3	Row sum
0	0	0	1/8	1/8	2/8
1	0	2/8	2/8	0	4/8
2	1/8	1/8	0	0	2/8
Column sum	1/8	3/8	3/8	1/8	1

Graphical representation:



Example:

A box contains 4 white balls, 3 black balls and 2 red balls. Two balls are to be drawn without replacement. Let, X denotes the no. of white balls and Y denotes the no. of black balls drawn. Find the JPD of (X, Y). Find also the probability that $X+Y \geq 3$. Find the probability that $X = 1$.

Solution:

X : no. of white balls drawn $x = 0, 1, 2$

Y : no. of black balls drawn $y = 0, 1, 2$

The possible pair of points are: (0, 0), (0, 1), (1, 0), (1, 1), (2, 0), (0, 2)

$$f(0,0) = P(X=0, Y=0) = \frac{\binom{4}{0}\binom{3}{0}\binom{2}{2}}{\binom{9}{2}} = 1/36$$

$$f(0,1) = P(X=0, Y=1) = \frac{\binom{4}{0}\binom{3}{1}\binom{2}{1}}{\binom{9}{2}} = 6/36$$

$$f(0,2) = P(X=0, Y=2) = \frac{\qquad}{\binom{9}{2}} = 3/36$$

$$f(1,0) = P(X=1, Y=1) = \frac{\qquad}{\binom{9}{2}} = 8/36$$

$$f(1,1) = P(X=1, Y=1) = \frac{\qquad}{\binom{9}{2}} = 12/36$$

$$f(2,0) = P(X=2, Y=0) = \frac{\qquad}{\binom{9}{2}} = 6/36$$

The JPD can now be placed in the following table:

	X values			
Y values	0	1	2	Row sum
0	1/36	8/36	6/36	15/36
1	6/36	12/36	0	18/36
2	3/36	0	0	3/36
Column sum	10/36	20/36	6/36	1

Now,

$$P(X+Y \geq 3) = f(1,2) + f(2,1) + f(2,2) = 0$$

Also,

$$P(X=1) = \sum_{Y=0}^{2} f(1,Y) = f(1,0) + f(1,1) + f(1,2)$$

$$= \frac{8}{36} + \frac{12}{36} + 0$$

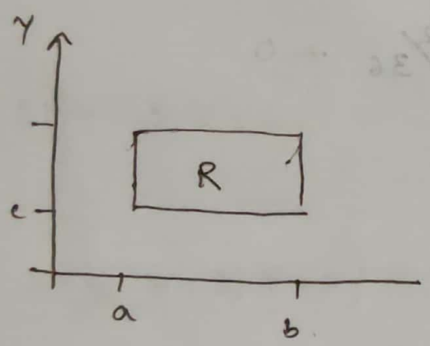
$$= \frac{20}{36}$$

Joint distribution for continuous variables:

Let X and Y be two continuous random variables. If there exists a non negative $f(x,y)$ defined on the entire x,y plane such that for every subset R of pairs of numbers.

$$P[(X,Y) \in R] = \iint_R f(x,y)\,dx\,dy$$

then $f(x,y)$ is called the joint density function (joint PDF) of (X,Y)



$$R = (a,b) \times (c,d)$$

$$\iint_R f(x,y)\,dx\,dy = \int_{x=a}^{b}\int_{y=c}^{d} f(x,y)\,dy\,dx$$

Like all other probability functions, the requirements for a continuous function $f(x,y)$ to be a joint PDF are:

1. $f(x,y) \geq 0,\ \forall (x,y) \in \mathbb{R}$ on $-\infty < x < \infty,\ -\infty < y < \infty$

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

3. $P[(X,Y) \in R] = \iint_R f(x,y)\, dx\, dy = P(a \leq X \leq b,\ c \leq Y \leq d)$

$$= \int_{x=a}^{b} \int_{y=c}^{d} f(x,y)\, dy\, dx$$

for any region R in the xy plane.

The cumulative distribution function (CDF) for $f(x,y)$ is defined by -

$$F(x,y) = P[X \leq x,\ Y \leq y] = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y)\, dy\, dx$$

the joint cumulative distribution function has the following properties -

1. $0 \leq F(x,y) \leq 1$
2. $\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$, where $F(x,y)$ is differential
3. $F(x,-\infty) = F(-\infty,y) = 0$
4. $F(-\infty,\infty) = 1$

Ex:

Let X and Y have the following distributions ▷

$$f(x,y) = x^2 + \frac{xy}{3} \; ; \; 0 \leq x \leq 1$$
$$= 0 \; ; \text{ elsewhere}$$

check whether $f(x,y)$ is a density function or not

Solution:

The function $f(x,y)$ will be a joint PDF is

$$f(x,y) \geq 0 \; ; \; \forall (x,y) \in \mathbb{R}$$

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

clearly $f(x,y) \geq 0$ for values of x and y in the given range.

($0 \leq x \leq 1$ and $0 \leq y \leq 2$)

Now, $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx = \int_{x=0}^{1}\int_{y=0}^{2} \left[x^2 + \frac{2y}{3}\right] dy\,dx$

$$= \int_{x=0}^{1} \left[x^2 y + \frac{xy^2}{6}\right]_0^2 dx$$

$$= \int_{x=0}^{1} \left(2x^2 + \frac{2x}{3}\right) dx = 2\int_{x=0}^{1} \left(x^2 + \tfrac{1}{3}x\right) dx$$

$$= 2\left[\frac{x^3}{3} + \frac{x^2}{6}\right]_0^1 = 2\left(\frac{1}{3} + \frac{1}{6}\right) = 2 \times \frac{3}{6} = 1$$

Hence $f(x,y)$ is a joint PDF.

Ex: Suppose that X and Y has the following density function $f(x,y)$

$= ky^2$; $0 \leq x \leq 2$; $0 \leq y \leq 1$

$= 0$; elsewhere

determine the value of the constant k and find $P(X \leq 1)$

[$k = 3/2$ Ans]

Solution:

$$\int_{x=0}^{2}\int_{y=0}^{1} f(x,y)\,dy\,dx = 1$$

Ex: Let the joint PDF of the continuous random variable (x, y) be

$$f(x,y) = \begin{cases} k(x^2 + 2xy) \,; & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 \,; & \text{elsewhere} \end{cases}$$

a. Find the value of k $\qquad k = 6/5$ (Ans)

b. Find $P(X \leq 1/2,\ Y \geq 1/2)$

c. What is the probability of the event $(X \leq Y)$?

Solution:

Hints:

$$\int_0^{1/2} \int_0^{1} f(x,y)\, dy\, dx = 11/30 \text{ (Ans)}$$

c. $$P(x \leq y) = 6/5 \int_{y=0}^{1} \int_{x=0}^{y} (x^2 + 2xy)\, dx\, dy$$

$$= 2/5$$

Marginal Distribution:

We have noted earlier that if the joint probability function $f(x,y)$ of two random variables X and Y is known, the probability function of X and Y can be derived separately.

In that case when the distribution of X is derived from $f(x,y)$ the resulting distribution is called marginal distribution of X.

Similarly in case of Y when the distribution of Y is derived from $f(x,y)$ the resulting distribution is called marginal distribution of Y.

Here we denote these two distributions by $g(x)$ and $h(y)$ respectively when they are derived from $f(x,y)$. Thus given the joint probability distribution $f(x,y)$ of two discrete random variable X and Y, the probability distribution of X alone is

$$g(x) = \sum_{y} f(x,y)$$

And that of Y is,

$$h(y) = \sum_{x} f(x,y)$$

For two continuous random variables X and Y

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

** Marginal Distributions are indeed probability distributions; since they satisfy, all the properties of a probability distribution

For continuous case:

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

and $P[a < X < b]$

$$= P[a < X < b, -\infty < Y < \infty]$$

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

Definition: Let $F(x,y)$ be the joint CDF of (X,Y). Then the marginal CDF of X is

$$F_1(x) = \lim_{y \to \infty} F(x,y)$$

Similarly, the marginal CDF of Y is

$$F_2(Y) = \lim_{x \to \infty} F(x,y)$$

The probability function or PDF of X associated with the marginal CDF of X is called the marginal PDF of X.

Ex:

Suppose X and Y have the following joint distribution

$Y \backslash X$	0	1	2	3	Rowsum
0	0	$1/8$	$2/8$	$1/8$	$4/8$
1	$1/8$	$2/8$	$1/8$	0	$4/8$
Columnsum	$1/8$	$3/8$	$3/8$	$1/8$	1

Find the marginal distribution of X and Y $(g(x), h(y))$

Hints: [rowsum and columnsum are results.]

Solution:

For the random variable X

$g(0) = P(X=0) = \sum_{Y=0}^{1} f(0,y) = f(0,0) + f(0,1) = 0 + 1/8 = 1/8$

$g(1) = 3/8$

$g(2) = 3/8$

$g(3) = 1/8$

Similarly for the r.v. Y

$h(0) = P(Y=0) = \sum_{x=0}^{3} f(x,0) = f(0,0) + f(1,0) + f(2,0) + f(3,0)$

$= 0 + 1/8 + 2/8 + 1/8$

$= 4/8$

$h(1) = P(Y=1) = \sum_{x=0}^{3} f(x,1)$

$= f(0,1) + f(1,1) + f(2,1) + f(3,1)$

$= 4/8$

The marginal distributions derived above may be written as

MD of x:

$X : x$	0	1	2	3	Sum
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

MD of Y:

$Y : Y$	0	1	Sum
$f(y)$	$\frac{4}{8}$	$\frac{4}{8}$	1

Example:

Find the marginal densities of X and Y from the following joint density function and verify that marginal distributions are also probability distributions.

$$f(x,y) = \frac{1}{8}(6-x-y) \quad \text{for} \quad 0<x<2, \; 2<y<4$$

$$= 0; \text{ elsewhere}$$

Also compute $P(X+Y<3)$ and $P(X<1.5, Y<2.5)$

Solution:

By definition, the marginal dsty of X is

$$g(x) = \frac{1}{8}\int_{y=2}^{4}(6-x-y)\,dy = \frac{1}{4}(3-x); \text{ for } 0<x<2$$

Scanned with CamScan

Similarly the marginal density of Y is

$$h(y) = \int_{x=0}^{2} f(x,y)\,dx = \frac{1}{8}\int_{x=0}^{2} (6-x-y)\,dx = \frac{1}{4}(5-y) \; ; \text{ for } 2<y<4$$

We need to verify that $g(x)$ and $h(y)$ are probably distributions

It is clear that in the given range of the variables X and Y,

$g(x) \geq 0$ and $h(y) \geq 0$

Also,

$$\int_{x=0}^{2} g(x)\,dx = \frac{1}{4}\int_{0}^{2} (3-x)\,dx = 1$$

$$\int_{y=2}^{4} h(y)\,dy = \frac{1}{4}\int_{2}^{4} (5-y)\,dy = 1$$

Thus $g(x)$ and $h(y)$ satisfy all the conditions of being a probability density function.

$$\frac{1}{8}\int_{y=2}^{4}\int_{x=0}^{3-y}(6-x-y)\,dx\,dy$$

$$= \frac{1}{8}\int_{y=2}^{4}\left[6x-\frac{x^2}{2}-xy\right]_{0}^{3-y} dy$$

$$= \frac{1}{8}\int_{y=2}^{4}\left(18-6y-\frac{(3-y)^2}{2}-(3-y)\times y\right) dy$$

$$= \frac{1}{8}\int_{y=2}^{4}\left(18-6y-\frac{9-6y+y^2}{2}-3y+y^2\right) dy$$

$$= \frac{1}{8}\int_{y=2}^{4}\left(18-\frac{9}{2}-9y+3y-\frac{y^2}{2}+y^2\right) dy$$

$$= \frac{1}{8}\int_{y=2}^{4}\left(13.5-6y+\frac{y^2}{2}\right) dy$$

$$= \frac{1}{8}\left[13.5y-3y^2+\frac{y^3}{3}\right]_{2}^{4}$$

$$= \frac{1}{8}\left(\cancel{13}\ 54-48+\frac{64}{3}-27+12-\frac{8}{3}\right)$$

$$= \frac{1}{8}\left(-9+\frac{56}{3}\right)$$

$$= 1/24 \ \ldots$$

and $P\left(X < \frac{3}{2}, Y < \frac{5}{2}\right)$

$$= \int_{x=0}^{3/2} \int_{y=2}^{5/2} f(x,y)\, dy\, dx = \frac{1}{8} \int_{x=0}^{3/2} \int_{y=2}^{5/2} (6-x-y)\, dy\, dx = \frac{9}{32}$$

Example:

The random probability function of a discrete bivariate r.v. is as follows:

$$f(x,y) = \begin{cases} \frac{1}{36} & \text{for } 1 \leq x = y \leq 6 \\ \frac{2}{36} & \text{for } 1 \leq y < x \leq 6 \\ 0 ; & \cancel{\text{oth}}\ \cancel{\text{elsewhere}}\ \text{otherwise} \end{cases}$$

Find $g(x)$ and $h(y)$

Bivariate cumulative distribution function:

The joint distribution function or joint cumulative distribution function (joint CDF) $F(x,y)$ of two dimensional random variable (X, Y) is defined as a function.

$$F(x,y) = P\left[X \leq x \text{ and } Y \leq y\right] \; ; -\infty < x < \infty \text{ and } -\infty < y < \infty$$

$F(x,y)$ is a monotonic increasing function in x for each fixed y and is a monotonic increasing function in y for each fixed x.

Here, the CDF of r.v. X is $F_1(x) = \lim_{y \to \infty} F(x,y)$

the CDF of r.v. Y is $F_2(y) = \lim_{x \to \infty} F(x,y)$

If the bivariate random variable (X, Y) is discrete

$$F(x,y) = P\left(X \leq x \text{ and } Y \leq y\right) = \sum_{r \leq x} \sum_{s \leq y} F(r,s) \text{ for } -\infty < x < \infty$$

and $-\infty < y < \infty$

If the bivariate random variable (X, Y) is continuous

$$F(x,y) = P\left(X \leq x \text{ and } Y \leq y\right) = \int_{y=-\infty}^{y} \int_{x=-\infty}^{x} f(r,s)\, dr\, ds \text{ for, }$$

$-\infty < x < \infty$

$-\infty < y < \infty$

If the bivariate random variable (X, Y) is continuous then,

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} \qquad \rightarrow CDF \rightarrow JPD$$

Note:

For given numbers $a < b$ and $c < d$

$P(a \leq X \leq b$ and $c \leq Y \leq d)$

$= P(a \leq X \leq b$ and $y \leq d) - P(a \leq X \leq b$ and $y \leq c)$

$= P(X \leq b$ and $y \leq d) - P(X \leq a$ and $y \leq d) - \{P(X \leq b$ and $y \leq c) - P(X \leq a$ & $Y \leq c)\}$

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

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

Scanned with CamScan

Problem:

The joint probability density function of (X, Y) is given as

$$f(x,y) = \begin{cases} \frac{1}{8}(x+y) \text{ for } 0<x<2 \text{ and } 0<y<2 \\ 0; \text{ otherwise} \end{cases}$$

Find it's joint CDF.

Solution: The joint CDF of (X, Y) is given by

For $x<0$ or $y<0$,

$F(x,y) = 0$

For, $0 \le x < 2$ or $0 \le y < 2$

$$F(x,y) = \int_0^x \int_0^y \frac{1}{8}(x+y)\,dy\,dx$$

$$= \frac{1}{8}\int_0^x \int_0^y (x+y)\,dy\,dx = \frac{1}{8}\int_0^x \left[xy + \frac{y^2}{2}\right]_0^y dx$$

$$= \frac{1}{8}\left(\frac{x^2 y}{2} + \frac{xy^2}{2}\right)$$

$$= \frac{1}{16} xy(x+y)$$

For $0 \le x < 2$ and $y \ge 2$

$$F(x,y) = \int_{x=0}^{x}\int_{y=0}^{2} \frac{1}{8}(x+y)\,dy\,dx = \frac{1}{8}x(x+2)$$

For $x \geq 2$ and $0 \leq y < 2$

$$F(x,y) = \int_{x=0}^{2}\int_{y=0}^{y} \frac{1}{8}(x+y)\,dy\,dx = \frac{1}{8}y(y+2)$$

For $x \geq 2$ and $y \geq 2$

$$F(x,y) = 1$$

Thus the complete joint CDF of (x,y) is given by

$$F(x,y) = \begin{cases} 0 \,;\ \text{for } x<0 \text{ or } y<0 \\ \frac{1}{16}xy(x+y)\,;\ \text{for } 0 \leq x < 2 \text{ and } 0 \leq y < 2 \\ \frac{1}{8}x(x+2)\,;\ \text{for } 0 \leq x < 2 \text{ and } y \geq 2 \\ \frac{1}{8}y(y+2)\,;\ \text{for } x \geq 2 \text{ and } 0 \leq y < 2 \\ 1\,;\ \text{for } x \geq 2 \text{ and } y \geq 2 \end{cases}$$

Problem: The joint CDF for water and electric demand (X, Y) is

$$F(x,y) = \begin{cases} 0 \; ; \text{ for } x < 4 \text{ or } y < 2 \\ \dfrac{(x-4)(y-2)}{19208} \; ; \text{ for } 4 \leq x < 200 \text{ and } 2 \leq y < 100 \\ \dfrac{x-4}{196} \; ; \text{ for } 4 \leq x < 200 \text{ and } y \geq 100 \\ \dfrac{y-2}{98} \; ; \text{ for } x \geq 200 \text{ and } 2 \leq y < 100 \\ 1 \; ; \text{ for } x \geq 200 \text{ and } y \geq 100 \end{cases}$$

a) Find the joint PDF of (X, Y);

b) $P(50 \leq x \leq 100, \; 50 \leq Y \leq 75)$

c) The CDF of X.

d) The CDF of Y.

Solution:

a) For $4 \leq x \leq 200$ and $2 \leq Y \leq 100$

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial^2}{\partial x \partial y}\left\{\frac{(x-4)(y-2)}{19208}\right\} = \frac{\partial^2}{\partial x}\left\{\frac{(x-4)}{19208}\right\}$$

$$= \frac{1}{19208}$$

Thus, $f(x,y) = \begin{cases} \dfrac{1}{19208} \text{ for } 4 \leq x \leq 200 \text{ and } 2 \leq y < 100 \\ 0 \; ; \text{ otherwise} \end{cases}$

b) $P(50 \leq x \leq 100,\ 50 \leq y \leq 75) = F(100,75) - F(50,75) - F(100,50) + F(50,50)$

$$= \left\{\frac{(100-4)(75-2)}{19208}\right\} - \left\{\frac{(50-4)(75-2)}{19208}\right\} - \left\{\frac{(100-4)(50-2)}{19208}\right\} \cancel{+ \left\{\frac{(100-50)(75-50)}{19208}\right\}} + \left\{\frac{(50-4)(50-2)}{19208}\right\}$$

$$= \frac{625}{9604}$$

c) If $4 \leq x < 200$ and $y \geq 100$

$F(x,y) = F(x,100)$ and it follows that

$F(x,y) = \dfrac{x-4}{196}$; Thus by letting $y \to \infty$ then the CDF of X is:

$$F_1(x) = \begin{cases} 0 & ; x < 4 \\ \dfrac{x-4}{196} & ; 4 \leq x < 200 \\ 1 & x \geq 200 \end{cases}$$

d) If $2 \leq y \leq 100$ and $x \geq 200$ then

$F(x,y) = F(200,y)$ and it follows that

$F(x,y) = \dfrac{y-2}{198}$

Thus by letting $x \to \infty$, the CDF of Y is

$$F_2(y) = \begin{cases} 0 ; & y < 2 \\ \frac{y-2}{198} ; & 2 \leq y < 100 \\ 1 ; & y \geq 100 \end{cases}$$

Conditional Distribution:

The conditional distribution are exactly analogous to the conditional probabilities of the type $P(A|B)$ or $P(B|A)$ where A and B are two events in a sample space.

$$P(B|A) = \frac{P(A \cap B)}{P(A)} ; P(A) > 0 \quad \Bigg| \quad P(A|B) = \frac{P(A \cap B)}{P(B)} ; P(B) > 0$$

Replacing the events A and B by the random variables X and Y respectively we can define the conditional probability of Y for given X as follows:

$$P(Y=y \mid X=x) = \frac{\{P(X=x), P(Y=y)\}}{P(X=x)} = \frac{f(x,y)}{g(x)}, \text{ where}$$

X and Y are discreet random variables.

The function $f(x,y)/g(x)$ is a function of y with x held fixed. We call this function conditional probability distribution of the discrete random ~~var~~ variable Y given $X = x$.

Scanned with CamScan

It satisfies all the properties of a probability distribution.

$$f(Y|x) = \frac{f(x,y)}{g(x)}, \quad g(x) > 0 \qquad g(x) = \sum_{Y} f(x,y)$$

$$f(x|Y) = \frac{f(x,y)}{h(Y)}, \quad h(Y) > 0 \qquad h(Y) = \sum_{x} f(x,y)$$

If X and Y are continuous random variable,

$$f(Y|x) = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)\,dy} = \frac{f(x,y)}{g(x)} \; ; \; f(x|Y) = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)\,dx} = \frac{f(x,y)}{h(Y)}$$

$$P(a < X < b \mid Y = Y) = \begin{cases} \sum_{x} f(x|Y), & \text{if } X \text{ is discrete} \\ \int_{a}^{b} f(x|Y) \; ; & \text{if } X \text{ is continuous.} \end{cases}$$

Example: Given the following joint distribution of the random variables X and Y.

	X		
Y	0	1	2
0	$3/28$	$9/28$	$3/28$
1	$6/28$	$6/28$	0
2	$1/28$	0	0

Find $f(x|1)$, $f(y|1)$ and $P(X=0|Y=1)$

Solution:

By definition $f(x|1) = \frac{f(x,1)}{h(1)}$ — ①

Now, $h(1) = \sum_{x=0}^{2} f(x,1) = f(0,1) + f(1,1) + f(2,1) = 6/28 + 6/28 + 0$
$= 3/7$

Hence, the conditional distribution of X given Y is

$$f(x|1) = \frac{f(x,1)}{h(1)} = 7/3\, f(x,1) \quad \text{for } x = 0, 1, 2$$

Therefore, $f(0|1) = 7/3\, f(0,1) = 7/3 \times 6/28 = 1/2$

$f(1,1) = 7/3\, f(1,1) = 7/3 \times 6/28 = 1/2$

$f(2,1) = 7/3\, f(2,1) = 7/3 \times 0$

x	0	1	2
$f(x\|1)$	$1/2$	$1/2$	0

Again, $f(y|1) = \frac{f(1,y)}{g(1)}$ —— ①

$g(1) = \sum_{y=0}^{2} f(1,y) = f(1,0) + f(1,1) + f(1,2)$

$= 9/28 + 6/28 + 0 = 15/28$

From 2 ⇒

$f(y|1) = \frac{f(1,y)}{g(1)} = \frac{28}{15} f(1,y) \; ; \; y = 0, 1, 2$

Therefore,

$f(0|1) = \frac{28}{15} f(1,0) = \frac{28}{15} \times \frac{9}{28} = 3/5$

$f(1,1) = \frac{28}{15} f(1,1) = 28/15 \times \frac{6}{28} = 2/5$

$f(2|1) = \frac{28}{15} f(1,2) = \frac{28}{15} \times 0 = 0$

y	0	1	2
f(y|1	3/5	3/5	0

Also $P(X=0|\; Y=1) = \frac{f(0,1)}{h(1)} = f(0|1) = 1/2$

$P(X=0|Y=1) = \frac{6/28}{3/7}$

$= 1/2$

Example:

$$f(x,y) = \frac{x+y}{21} \; ; \; x = 1,2,3 \; ; \; y = 1,2$$

i. Obtain the marginal distributions and conditional distributions of X and Y

ii. Find $f(x|1)$ iii. Find $P(X=2|Y=1)$

Example:

Find the conditional density of Y given X for the following distributions

$$f(x,y) = \begin{cases} \frac{6-x-y}{8} & ; 0<x<2, 2<y<4 \\ 0 & ; \text{elsewhere} \end{cases}$$

Also compute $P(2<Y<3 \mid X=2)$

Solution:

The conditional density of Y given X is

$$f(y|x) = \frac{f(x,y)}{g(x)} \quad \text{where } g(x) = \int_{y=2}^{4} f(x,y)\,dy$$

$$= \frac{1}{8}\int_{2}^{4} 6-x-y \; dy = \frac{1}{4}(3-x)$$

$$\therefore g(x) = \frac{1}{4}(3-x) \; ; 0<x<2$$

Thus $f(y|x) = \frac{6-x-y}{8} \times \frac{4}{(3-x)} = \frac{6-x-y}{2(3-x)}$; $0<x<2$; $2<y<4$

$$P(2<Y<3 \mid X=2) = \int_{y=2}^{3} f(y|2)\,dy = \int_{y=2}^{3} \frac{6-2-y}{2(3-2)}\,dy = \frac{3}{4}$$

; $2<y<4$

Example,

Two random variables X and Y have the following pdf

$$f(x,y) = \begin{cases} \frac{1}{2} & ; 0 \le x \le Y \le 2 \\ 0 & ; \text{elsewhere} \end{cases}$$

a) Find the marginal density of Y and hence the conditional density of X.

b) Find also $P(X \le 0.5 \mid Y = 1.5)$

Independence of random variable:

Two random variables X and Y with marginal densities $g(x)$ and $h(Y)$ respectively are said to be independent if and only if:

$$f(x|Y) = g(x)$$

or $f(Y|x) = h(Y)$

where $f(x|Y)$ is the conditional density of X for a given Y and $f(Y|x)$ is the conditional density of Y for a given X.

If X and Y are idependent then for any real numbers x and y it must be true that

$$f(x,y) = g(x) * h(Y)$$

Example: Suppose x and Y have the following jpd.

	X		
Y	2	4	Row sum
1	0.10	0.15	0.25
3	0.20	0.30	0.50
5	0.10	0.15	0.25
Column sum	0.40	0.60	1

Check whether X and Y are independent or not.

Solution:

To check the independence of X and Y we must compute $g(x)$ and $h(y)$ for all values of the random variables X and Y and equate to $f(x,y)$

i. $f(2,1) = 0.10$; $g(2) = 0.40$; $h(1) = 0.25$

$g(2) \times h(1) = 0.40 \times 0.25 = 0.1 = f(2,1)$

ii. $f(4,1) =$ $g(4) =$ $h(1) =$

iii. $f(2,3) =$ $g(2) =$ $h(3) =$

iv. $f(4,3) =$ $g(4) =$ $h(3) =$

v. $f(2,5) =$ $g(2) =$ $h(5) =$

vi. $f(4,5) =$ $g(4) =$ $h(5) =$

$P(Y=i \mid X=j)$

i. $P(Y=1 \mid X=2) = f\left(\overset{y}{1} \mid \overset{x}{2}\right) = \dfrac{f(2,1)}{g(2)}$

$$= \frac{0.10}{0.40}$$

$$= 0.25$$

$$= h(1)$$

ii.