STATISTICS

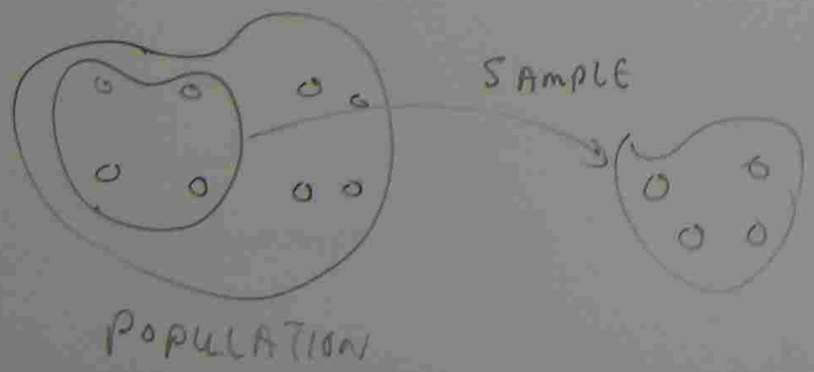
STATISTICS INVOLVES COLLECTING, SUMMARISING ANALYZING AND INTERPRETING DATA.

TYPES OF DATA

CATEGORICAL DATA

NUMERICAL DATA

- DISCRETE
- CONTINUOUS



DISCRETE — COUNTABLE SUCH AS NUMBER OF DAYS.

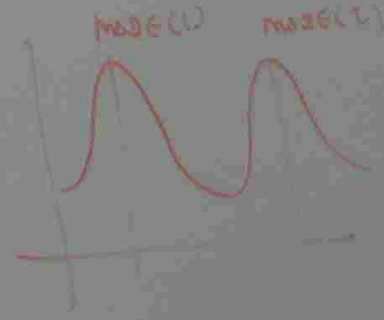
CONTINUOUS — INFINITE SUCH AS DAILY TEMPERATURE

PARAMETER

A NUMERICAL MEASURE THAT DESCRIBES SOME CHARACTERISTICS OF POPULATION

STATISTICS - NUMERICAL MEASURE THAT DESCRIBES SOME CHARACTERISTICS OF SAMPLE

MODE THE DATA SET THAT OCCURS MOST OFTEN



OF DAYS.

PERATURE

SOME

SOME

OFTEN

MODE(2)

NO MODE

Pb (1) FIND THE MODE OF THE FOLLOWING SAMPLES OF DATA

(a) 18, 19, 18, 20, 18, 18, 20, 21, 37, 18

(b) 27, 28, 29, 36, 14, 24

(c) 19, 18, 20, 35, 18, 18, 19, 46, 19

(a) MODE = 18

(b) NO MODE

(c) 19, 18 = MODES

18 = MODE 1

19 = MODE 2

MEDIAN

THE MIDDLE NUMBER IN AN ORDERED SET.

Pb (2) FIND THE MEDIAN OF EACH OF THE FOLLOWING SAMPLES OF DATA

(a) X: 7, 3, 2, 8, 11

(b) X: 18, 19, 18, 20, 18, 18, 20, 21, 37, 18

(a) 2, 3, 7, 8, 11

MEDIAN = 7

(b) X = 18, 18, 18, 18, 18, 19, 20, 20, 21, 37

MEDIAN =

= 18.

MEAN

THE AR

MEASUREMENTS

BY ADDING TH

DIVIDING THI

THE DATA S

$\bar{X}$ =

$$\text{MEDIAN} = \frac{18+19}{2}$$

$$= 18.5$$

MEAN THE ARITHMETIC AVERAGE OF ALL MEASUREMENTS IN THE DATA SET OBTAINED BY ADDING THEM ALL UP AND DIVIDING THIS SUM BY THE SIZE OF THE DATA SET.

$$\bar{X} = \frac{\Sigma X}{n}$$

21, 37

pb (3) FIND THE MEAN OF THE FOLLOWING SAMPLES OF DATA

(i) X: 11, 14, 10, 5

(ii) X: 2, 2, 4, 5, 5, 5, 6, 7, 8, 9, 11

(i) $\text{MEAN} = \bar{x} = \frac{11+14+10+5}{4} = \frac{40}{4} = 10$

(ii) $\text{MEAN} = \bar{X} = \frac{2+2+4+5+5+5+6+7+8+9+11}{11}$

$= \frac{64}{11} = 5.818$

pb (4) FIND THE MEAN OF THE FOLLOWING GROUP FREQUENCY DISTRIBUTION (GFD)

INTERVAL	FREQUENCY
0 → 9	1
10 → 19	2
20 → 29	4
30 → 39	8
40 → 49	4
50 → 59	1

INTERVAL	MEAN	f	fx
0 ⟶ 9	$\frac{0+9}{2} = 4.5$	1	$4.5 \times 1 = 4.5$
10 ⟶ 19	$\frac{10+19}{2} = 14.5$	2	$14.5 \times 2 = 29$
20 ⟶ 29	$\frac{20+29}{2} = 24.5$	4	$24.5 \times 4 = 98$
30 ⟶ 39	$\frac{30+39}{2} = 34.5$	8	$34.5 \times 8 = 276$
40 ⟶ 49	$\frac{40+49}{2} = 44.5$	4	$44.5 \times 4 = 178$
50 ⟶ 59	$\frac{50+59}{2} = 54.5$	1	$54.5 \times 1 = 54.5$
		$\Sigma f = 20$	$\Sigma fx = 640$

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{640}{20} = 32$$

+11

P FREQUENCY

RANGE

THE DIFFERENCE BETWEEN THE LARGEST AND SMALLEST VALUES IN THE SAMPLE.

Pb (5) FIND THE RANGE OF THE FOLLOWING SAMPLE OF DATA.

(i) 18, 19, 18, 20, 18, 37, 21, 20, 18, 18

(ii) 18, 19, 18, 20, 18, 21, 20, 18, 18

(i) RANGE = LARGEST − SMALLEST

= 37 − 18 = 19

(ii) RANGE = LARGEST − SMALLEST

= 21 − 18 = 3

RESIDUAL

THE DIFFERENCE BETWEEN EACH SAMPLE x AND THE MEAN $\bar{x}$ IS CALLED THE DEVIATION FROM THE MEAN OR RESIDUAL.

$$\text{RESIDUAL} = x - \bar{x}$$

Pb (6) 10, 12, 15, 17, 21

CALCULATE (i) MEAN (ii) RESIDUAL (DEVIATION)

(i) $\bar{x} = \frac{10+12+15+17+21}{5} = \frac{75}{5} = 15$

(ii) RESIDUAL $= (x - \bar{x}) = (10-15), (12-15), (15-15), (17-15)$

$= -5, -3, 0, 2, 6$

VARIANCE

$$\text{VARIANCE} = \frac{\sum (x - \bar{x})^2}{n-1}$$

CE BETWEEN EACH SAMPLE x AND THE MEAN $\bar{x}$

THE DEVIATION FROM THE MEAN OR RESIDUAL.

IDUAL $= x - \bar{x}$

, 12, 15, 17, 21

(i) MEAN (ii) RESIDUAL (DEVIATION)

$\frac{10+12+15+17+21}{5} = \frac{75}{5} = 15$

$= (x - \bar{x}) = (10-15), (12-15), (16-15), (17-15), (21-15)$

$= -5, -3, 0, 2, 6$

$\text{VARIANCE} = \frac{\sum (x-\bar{x})^2}{n-1}$

pb (7) CALCULATE VARIANCE OF 10, 12, 15, 17, 21

$\bar{x} = \text{MEAN} = \frac{\sum x}{n} = \frac{10+12+15+17+21}{5} = 15$

$\text{VARIANCE} = s^2 = \frac{\sum (x-\bar{x})^2}{n-1} = \frac{((0-15)^2 + (12-15)^2 + (15-15)^2 + (17-15)^2 + (21-$

$= \frac{(-5)^2 + (-3)^2 + (0)^2 + (2)^2 + (6)^2}{4}$

$= \frac{25+9+4+36}{4} = \frac{74}{4} = 18.5$

STANDARD DEVIATION

THE STANDARD DEVIATION OF A SAMPLE OF DATA IS DEFINED TO THE POSITIVE SQUARE ROOT OF THE VARIANCE

$s = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}}$

$(21-15)^2$

Pb (8) CALCULATE STANDARD DEVIATION of 10, 12, 15, 17, 21

$$S = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}} = \sqrt{\frac{(10-15)^2+(12-15)^2+(15-15)^2+(17-15)^2+(21-15)^2}{5-1}}$$

$$\bar{x} = \frac{10+12+15+17+21}{5} = 15$$

$$= \sqrt{18.5} = 4.3$$

CALCULATION of S^2 (VARIANCE) AND S (STANDARD DEVIATION) FOR A SAMPLE OF RAW DATA

$$S^2 = \frac{\sum (x-\bar{x})^2}{n-1} = \frac{\sum x^2 - n(\bar{x})^2}{n-1}$$

Pb (9) CALCULATE THE VARIANCE AND STANDARD DEVIATION FOR THE FOL SAMPLE OF DATA.

$$X: 2, 3, 5, 5, 9$$

$$\bar{x} = \frac{\sum x}{n} = \frac{2+3+5+5+9}{5} = \frac{24}{5} = 4.8$$

$$\sum x^2 = 2^2+3^2+5^2+5^2+9^2 = 144$$

$$n(\bar{x})^2 = 5 \times (4.8)^2 = 115.2$$

$$S^2 = \frac{\sum (x-\bar{x})^2}{n-1} = \frac{\sum x^2 - n(\bar{x})^2}{n-1}$$

$$S^2 = \frac{144-115.2}{5-1} = 7.2$$

$$S = \sqrt{7.2} = 2.68$$

VARIAN

STAND DEVIA

Pb (1

$\bar{x}$=

$S^2 = \frac{\sum fx^2 - n($

$n-1$

$S = \sqrt{32.8}$ = 5.7

ANCE AND STANDARD DEVIATION FOR THE FOLLOWING

$s_1, s_2, 9$

$\frac{s+s+9}{} = \frac{24}{5} = 4.8$

$^2 + s^2 + 9^2 = 144$

$)^2 = 115.2$

$= \frac{\sum x^2 - n(\bar{x})^2}{n-1}$

$^2 = 7.2$

68

CALCULATION OF s^2 AND s FOR SAMPLE DATA PRESENTED AS GFD (GROUP FREQUENCY DISTRIBUTION).

VARIANCE: $s^2 = \frac{\sum f(x-\bar{x})^2}{n-1} = \frac{\sum f(x-\bar{x})^2}{\sum f - 1} = \frac{\sum fx^2 - n(\bar{x})^2}{n-1}$

STANDARD DEVIATION $= s = \sqrt{\frac{\sum fx^2 - n(\bar{x})^2}{n-1}}$

Pb (10) FIND THE VARIANCE AND STANDARD DEVIATION OF THE FOLLOWING SAMPLE OF GROUPED DATA.

INTERVAL	FREQUENCY
10 – 14	4
15 – 19	7
20 – 24	5
25 – 29	3
30 – 34	1

INTERVAL	x (MID POINT)	f	fx	fx^2
10 – 14	$\frac{10+14}{2} = \frac{24}{2} = 12$	4	$4 \times 12 = 48$	$4 \times 12^2 = 576$
15 – 19	$\frac{15+19}{2} = 17$	7	$17 \times 7 = 119$	$7 \times 17^2 = 2023$
20 – 24	$\frac{20+24}{2} = 22$	5	$22 \times 5 = 110$	$5 \times 22^2 = 2420$
25 – 29	$\frac{25+29}{2} = 27$	3	$27 \times 3 = 81$	$3 \times 27^2 = 2187$
30 – 34	$\frac{30+34}{2} = 32$	1	$32 \times 1 = 32$	$1 \times 32^2 = 1024$
		$\sum f = 20$	$\sum fx = 390$	$\sum fx^2 = 8230$

$\bar{x} = \frac{\sum fx}{\sum f} = \frac{390}{20} = 19.5$

$s^2 = \frac{\sum fx^2 - n(\bar{x})^2}{n-1} = \frac{8230 - 20 \times (19.5)^2}{20-1}$

$= \frac{8230 - 7605}{19} = 32.89$

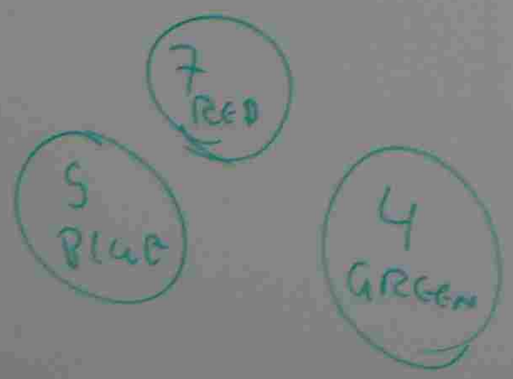
$s = \sqrt{32.89} = 5.739$

SET THEORY AND ELEMENTARY PROBABILITY

DEFINITION of PROBABILITY

$$P(A) = \frac{\text{NUMBER of TIMES A OCCUR}}{\text{TOTAL NUMBER of REPETITIONS OF EXPERIMENT}} = \frac{m}{N}$$

pb (11) A BAG CONTAINS 7 RED BALLS, 5 BLUE AND 4 GREEN BALLS. FIND THE PROBABILITY THAT IF ONE BALL ONLY IS DRAWN RANDOMLY, IT WILL BE GREEN.



N = TOTAL NUMBER OF BALLS = 7 + 5 + 4 = 16

m = NO. OF GREEN BALL = 4

$$P(\text{GREEN}) = \frac{m}{N} = \frac{4}{16} = \frac{1}{4}$$

pb (12)

Pb (12) IF A SINGLE DRAW IS MADE FROM A STANDARD PACK OF CARDS. FIND THE PROBABILITY OF DRAWING A JACK (OR) A QUEEN (OR) A KING.

1 PACK = 52 CARDS.

JACK = 4 CARDS

QUEEN = 4 CARDS

KING = 4 CARDS.

$P(A) = P(\text{JACK}) = \frac{4}{52} = \frac{1}{13}$

$P(B) = P(\text{QUEEN}) = \frac{4}{52} = \frac{1}{13}$

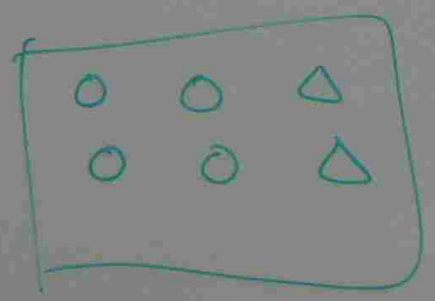
$P(C) = P(\text{KING}) = \frac{4}{52} = \frac{1}{13}$

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

$= \frac{1}{13} + \frac{1}{13} + \frac{1}{13}$

$= \frac{3}{13}$

$-4 = 16$



$\frac{2}{6}$ $P(\triangle)$

$1 - \frac{2}{6} = \frac{4}{6}$ $1 - P(\triangle)$

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

Pb (14) THERE ARE 52 CARDS IN A STANDARD PACK IN WHICH 39 CARDS CONTAIN HEART. CALCULATE THE PROBABILITY THAT THE CARD CONTAINING HEART IS DRAWN AND PROBABILITY THAT THE CARD NOT CONTAINING HEART IS DRAWN.

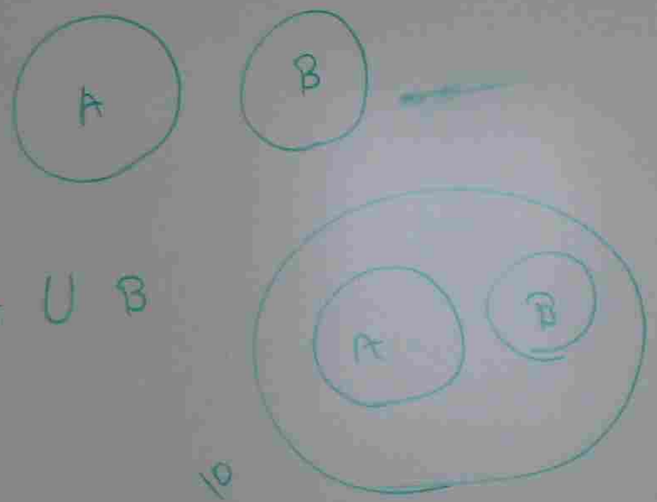
$P(A) = P(\text{HEART}) = \frac{39}{52} = \frac{3}{4}$

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

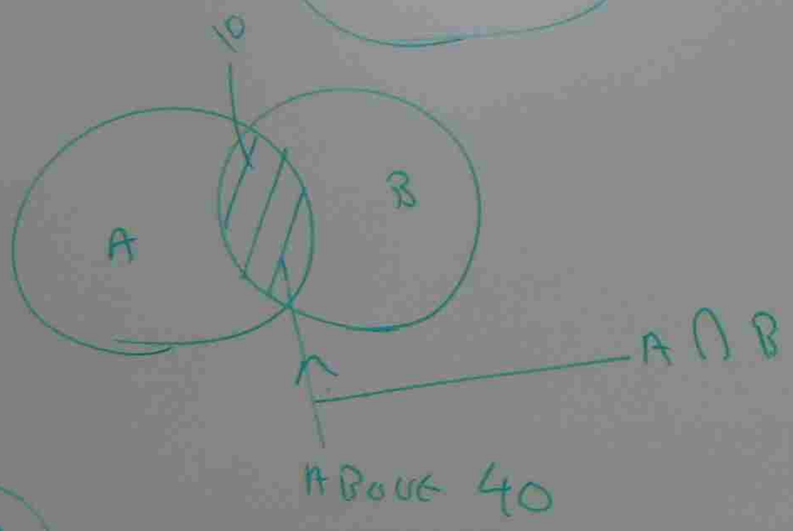
$= 1 - \frac{3}{4}$

$= \frac{1}{4}$

UNION, INTERSECT, SUBSET



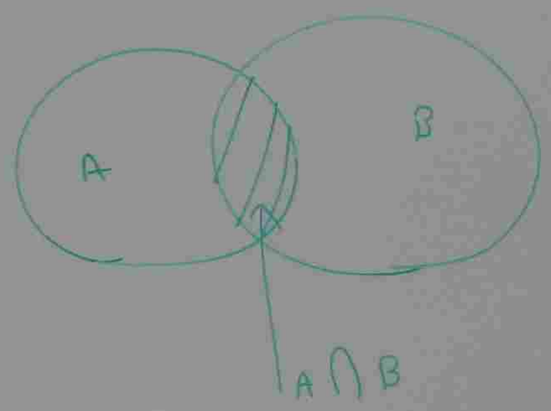
A U B



(A) 60 STUDENTS, 10 ABOVE AGE 40

(B) 50 STUDENTS, 10 ABOVE AGE 40

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



$$P(A) = \frac{60}{110}$$

$$P(B) = \frac{50}{110}$$

$$P(A \cap B) = \frac{10}{110}$$

$$P(A \cup B) = \frac{60}{110} + \frac{50}{110} - \frac{10}{110}$$

$$= \frac{100}{110}$$

Pb (15) $P(S) = 0.4$

FIND $P($

$P(S \cup T) = P(S)$

$= 0.4$

$=$

Pb (16) A SI

IS THE P

ON THE U

NUMBER

Pb (15) $P(S) = 0.4$, $P(T) = 0.5$, $P(S \cap T) = 0.15$

FIND $P(S \cup T)$

$$P(S \cup T) = P(S) + P(T) - P(S \cap T)$$

$$= 0.4 + 0.5 - 0.15$$

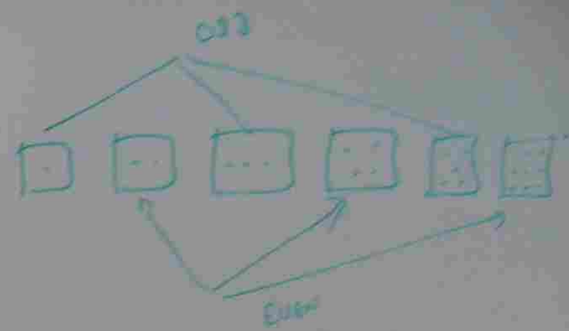
$$= 0.75$$

PROBABILITY OF SUBSET

$A \subset B$, A IS SUB SET of B

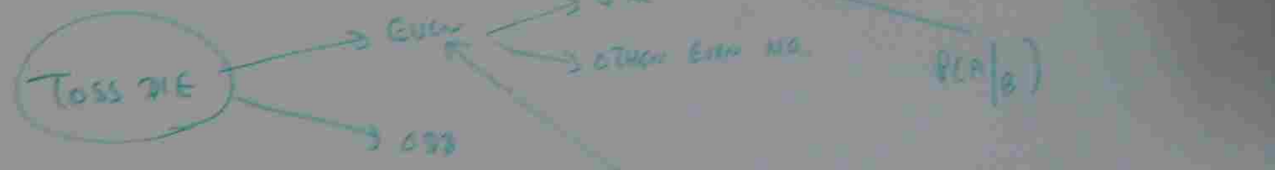
$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Pb (16) A SINGLE DIE IS TOSSED. WHAT IS THE PROBABILITY OF A 6 APPEARING ON THE UPPER MOST FACE GIVEN THAT NUMBER SHOWING IS EVEN.



A = EVENT THAT "6" IS SHOWING

B = EVENT THAT EVEN NUMBER IS SHOWING



$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

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

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

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

pb (17) SUPPOSE WE HAVE 3 EVENTS X, Y, Z

$P(X) = \frac{1}{2}$, $P(Y) = \frac{1}{4}$, $P(Z) = \frac{1}{8}$

FIND (a) $P(Z \cap X)$, (b) $P(Z \cup X)$

(c) $P(X \cup Y)$, (d) $P(X \cup \bar{Y})$

(a) $P(Z \cap X) = P(Z) \times P(X)$

$= \frac{1}{8} \times \frac{1}{2} = \frac{1}{16}$

(b) $P(Z \cup X) = P(Z) + P(X) - P(Z \cap X)$

$= \frac{1}{8} + \frac{1}{2} - \frac{1}{16}$

$= \frac{2+8}{16} - \frac{1}{16}$

$= \frac{10}{16} - \frac{1}{16}$

$= \frac{10-1}{16} = \frac{9}{16}$

(c) $P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$

$= P(X) + P(Y) - P(X) \times P(Y)$

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

$= \frac{4+2}{8} - \frac{1}{8} = \frac{6}{8} - \frac{1}{8} = \frac{5}{8}$

(d)

(d) $P(X \cup \bar{Y}) = P(X) + P(\bar{Y}) - P(X \cap \bar{Y})$

$= P(X) + [1 - P(Y)] - P(X) \times P(\bar{Y})$

$= P(X) + [1 - P(Y)] - P(X) \times [1 - P(Y)]$

$= \frac{1}{2} + \left[1 - \frac{1}{4}\right] - \frac{1}{2} \times \left[1 - \frac{1}{4}\right]$

$= \frac{1}{2} + \frac{3}{4} - \frac{1}{2} \times \frac{3}{4}$

$= \frac{4+6}{8} - \frac{3}{8}$

$= \frac{10}{8} - \frac{3}{8} = \frac{7}{8}$ //

Pb(18) SUPPOSE THAT WE KNOW FROM EXPERIMENT THAT 28% OF PEOPLE READ THE WOMAN'S DAY ARTICLE AND 45% OF PEOPLE READ THE NEW IDEA.

READING WOMEN'S DAY AND NEW IDEA ARE INDEPENDENT OF EACH OTHER.

TWO PEOPLE ARE SELECTED AT RANDOM

FIND THE PROBABILITY THAT

(a) NEITHER READ NEW IDEA

(b) BOTH READ WOMEN'S DAY

(c) BOTH DO NOT READ WOMEN'S DAY

$P(W) = P(\text{WOMEN'S DAY}) = 0.28$

$P(\bar{W}) = P(\text{DO NOT READ WOMEN DAY}) = 1 - P(W)$

$= 1 - 0.28$

$= 0.72$

$P(N) = 0.45$

$P(\bar{N}) = 1 - P(N) = 1 - 0.45 = 0.55$

(a) NEITHER READS NEW IDEA (FROM TWO PEOPLE)

$P(\bar{N} \cap \bar{N}) = P(\bar{N}) \times P(\bar{N})$

$= 0.55 \times 0.55 = 0.3025$

(b) P

(c) P

Pb(19

6

OF

UF

FIND

(i) F

AD

(

( )

$\frac{1}{8} = \frac{5}{8}$

(b) $P(W \cap W) = P(W) \times P(W)$

$= 0.28 \times 0.28 = 0.0784$

(c) $P(\bar{W} \cap \bar{W}) = P(\bar{W}) \times P(\bar{W})$

$= 0.72 \times 0.72 = 0.5184$

Pb (19) THERE ARE 10 PEOPLE TO BE CLASSIFIED ACCORDING TO SEX & AGE

6 ARE FEMALES AND 4 ARE MALES.

OF 6 FEMALES, 4 ARE UNDER 40, 2 ARE 40 (OR) MORE

OF 4 MALES. 3 ARE UNDER 40, 1 IS 40 (OR) MORE

FIND THE PROBABILITY THAT PERSON IS

(i) FEMALE (ii) UNDER 40, (iii) FEMALE UNDER 40 (iv) UNDER 40 GIVEN THAT SHE IS FEMALE

(v) MALE AND 40 (OR) MORE, (vi) MALE

GIVEN THAT HE IS UNDER 40.

(i) $P(F) = \frac{6}{10}$

$P(A)$ UNDER 40

(ii) $P(A) = \frac{4+3}{10} = \frac{7}{10}$

$P(F)$, $P(A)$
FEMALE , UNDER 40

$P(F \cap A) =$ FEMALE UNDER 40 $= \frac{4}{10}$

$P(A \cap F) = P(F \cap A)$

(iii)

(iv) $P(A/F) = \frac{P(A \cap F)}{P(F)} = \frac{P(F \cap A)}{P(F)} = \frac{\frac{4}{10}}{\frac{6}{10}}$

$= \frac{4}{10} \times \frac{10}{6}$

$= \frac{4}{6}$

→ AMONG THE FEMALES, FIND UNDER 40

(vi)

(v) $P(A)$ = UNDER 40

40 OR MORE = $P(\bar{A})$

$P(m)$ = MALE

MALE AND 40 (OR) MORE = $P(m \cap \bar{A}) = P(m) \times P(\bar{A})$

$= P(m) \times (1 - P(A))$

$= \frac{4}{10} \times \left(1 - \frac{7}{10}\right)$

$= \frac{4}{10} \times \frac{3}{10} = \frac{12}{100} \simeq \frac{}{10}$

(vi) $P(\bar{A}/m) = \frac{P(m \cap A)}{P(m)}$

$= \frac{3/10}{4/10} = \frac{3}{4}$

$)=P(E \cap A)$

$A) = \frac{4}{10}$

$\frac{}{} = \frac{6}{10}$

$= \frac{4}{10} \times \frac{10}{6}$

$= \frac{4}{6}$