Knowledge Discovery and Data Mining

HOMEWORK-1

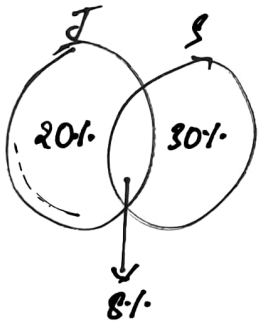
Homework 1.1

Jerry (J) and Susan (S) have a joint bank account

J goes to the bank 20% of the days

S goes to the bank 30% of the days

Both at the bank together 8% of the days



	(S) at bank	(S) not at bank
(J) at bank	8%	12%
(J) not at bank	22%	58%

a) probability $P(J \text{ at bank} / S \text{ at bank}) = \frac{8}{30} = 0.266 = 26.6\%$

b) probability $P(J \text{ at bank} / S \text{ not at bank}) = \frac{12}{70} = 0.1714 = 17.1\%$

c) probability $P(J \cap S / J \cup S) = \frac{8}{42} = 19.04\%$

Homework 1.2

Harold (H) and Sharon (S) are studying for a test

Harold (H) chances of getting "B" are 80%. P(H) = 80%.

Sharon (S) chances of getting "B" are 90%. P(S) = 90%.

The probability of atleast one of them getting a "B" is 91%. P(HUS) = 91%.

$$P(H \cap S) = P(H) + P(S) - P(H \cup S) = 80 + 90 - 91 = 79\%$$

a) Probability of only H = $P(H) - P(H \cap S) = 80 - 79 = 1\%$.

b) probability of only S = $P(S) - P(H \cap S) = 90 - 79 = 11\%$.

c) probability of no one = $100 - P(H \cup S) = 100 - 91 = 9\%$.

Homework 1.3

Jerry (J) and Susan (S) have a joint account

P(J goes to bank) = 20%.

P(S goes to bank) = 30%.

$P(J \cap S) = 8\%$. → given

Symbolically,
If J and S are independent then,

$P(J \cap S) = P(J)\ P(S)$

$= 20 \times 30$

$= 60\%$.

Both are not same, so they are not independent events.

Homework 1.4 :

2 Dice are rolled

The probabilities are as follows in the table

	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

By the table,

a) probability of sum is 6 = $\frac{5}{36}$

probability of second die is 5 = $\frac{6}{36} = \frac{1}{6}$

Both events intersection probability = $\frac{1}{36}$

According to independent events,

probability of both events i.e, sum is 6 and 2nd die shows 5 = $\frac{5}{36} \cdot \frac{1}{6}$

$= \frac{5}{216}$

Both are not equal so, they are not independent

b) probability of sum is 7 = $\frac{1}{6}$ — (A)

Probability of 1st die shows 5 = $\frac{1}{6}$ — (B)

Probability of both at once = $\frac{1}{36}$

Symbolically, if two events (A,B) are independent then,

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

$$= \frac{1}{6} \cdot \frac{1}{6}$$

$$= \frac{1}{36}$$

It satisfies the condition so both are independent events.

Homework 1.5 :

- Oil company drilling in either TX, AK and NJ
- only opeartes in one state
- Chances of finding oil are : Tx 30%.
  AK 20%.
  NJ 10%.
- chance of company choosing the state are : Tx 60%.
  100 - (60+10) = 100 - 70 → AK 30%.
  NJ 10%.

Probability of finding oil in Tx = $60 \times \frac{30}{100} = 18\%$.

AK = $30 \times \frac{20}{100} = 6\%$.

NJ = $10 \times \frac{10}{100} = 1\%$.

a) probability of finding oil is

$= 18 + 6 + 1$

$= 25\%$.

b) Probability of drilled and found oil in Tx

$= \frac{18}{25}$

$= 0.72 = 72\%$.

Homework 1.6 :

Not survived passengers total $= 528 + 167 + 122$

$= 817$

Total passengers $= 706 + 285 + 325 = 1316$

a) Probability of passengers didnot survive is $817/1316 = 62.08\%$

b) first class passengers total $= 325$

probability of passengers staying first class is $325/1316 = 24.69\%$.

c) probability of first class passengers survived $= 203/499 = 40.68\%$.

d) $P(\text{survival}) = 711/2201 = 32.5\%$.

$P(\text{survival not crew member}) = 499/1316 = 0.5791$

$P(1^{st}\text{ class}) = 24.69\%$. → from (b) question

$P(1^{st}\text{ class} \cap \text{survival}) = 203/1316 = 15.42\%$.

independant events formula,

$$P(S \cap FC) = P(S) \times P(FC)$$

$$= 499/1316 \cdot 325/1316$$

$$= 0.09865$$

It is not same so events are not independant

e) Probability of passenger survived is child and 1st class is 6/499
survived childs = 6 / 1st class survived = 499
= 1.201

f) Total passengers survived are 499
adult passengers = 442

Probability of adult passenger is survived = 442/499 = 88.57%

g) lets imagine age as (A) and staying in 1st class as (B)

$$P(\text{adult } A/B) = \frac{197}{711} = 0.277$$

$$P(A \text{ adult}) = \frac{654}{711} = 0.919$$

$$P(A \text{ child}/B) = \frac{6}{711} = 0.0084$$

$$P(A \text{ child}) = \frac{57}{711} = 0.0801$$

$P(A \text{ Adult}/B) \neq P(A \text{ adult})$

&

$P(A \text{ child}/B) \neq P(A \text{ child})$

Both are not same, so age and 1st class are not independent