

① Air traffic Problem

Days	Season	Fog	Rain	Class
Weekday	Spring	None	None	On time
Weekday	Winter	None	Slight	On time
Weekday	Winter	None	None	On time
Holiday	Winter	High	Slight	Late
Saturday	Summer	Normal	None	On time
Weekday	Autumn	Normal	None	Very late
Holiday	Summer	High	Slight	On time
Sunday	Summer	Normal	None	On time
Weekday	Winter	High	Heavy	Very late
Weekday	Summer	None	Slight	On time
Saturday	Spring	High	Heavy	Cancelled
Weekday	Summer	High	Slight	On time
Weekday	Winter	Normal	None	Late
Weekday	Summer	High	None	On time
Weekday	Winter	Normal	Heavy	Very late
Saturday	Autumn	High	Slight	On time
Weekday	Autumn	None	Heavy	On time
Holiday	Spring	Normal	Slight	On time
Weekday	Spring	Normal	None	On time
Weekday	Spring	Normal	Heavy	On time

Let's calculate all the posterior and prior probabilities.

Prior probabilities -

Day :-

	ontime	late	very late	cancelled
Weekday	$9/14 = 0.64$	$1/2 = 0.5$	$3/3 = 1$	$0/1 = 0$
Saturday	$2/14 = 0.14$	$1/2 = 0.5$	$0/3 = 0$	$1/1 = 1$
Sunday	$1/14 = 0.07$	$0/2 = 0$	$0/3 = 0$	$0/1 = 0$
Holiday	$2/14 = 0.14$	$0/2 = 0$	$0/3 = 0$	$0/1 = 0$

Season -

	ontime	late	very late	Cancelled
Spring	$9/14 = 0.29$	$0/2 = 0$	$0/3 = 0$	$0/1 = 0$
Summer	$6/14 = 0.43$	$0/2 = 0$	$0/3 = 0$	$0/1 = 0$
Autumn	$2/14 = 0.14$	$0/2 = 0$	$1/3 = 0.33$	$0/1 = 0$
Winter	$2/14 = 0.14$	$2/2 = 1$	$2/3 = 0.67$	$0/1 = 0$

Fog -

	ontime	late	very late	Cancelled
None	$5/14 = 0.36$	$0/2 = 0$	$0/3 = 0$	$0/0 = 0$
High	$4/14 = 0.29$	$1/2 = 0.5$	$1/3 = 0.33$	$1/1 = 1$
Normal	$5/14 = 0.36$	$1/2 = 0.5$	$4/3 = 0.67$	$0/0 = 0$

Rain -

	ontime	late	very late	Cancelled
None	$5/14 = 0.36$	$1/2 = 0.5$	$1/3 = 0.33$	$0/1 = 0$
Slight	$8/14 = 0.07$	$0/2 = 0$	$0/3 = 0$	$0/1 = 0$
Heavy	$1/14 = 0.07$	$1/2 = 0.5$	$4/3 = 0.67$	$1/1 = 1$

Prior probabilities \rightarrow $14/20 = 0.70$ | $2/20 = 0.10$ | $3/20 = 0.15$ | $1/20 = 0.05$

Instance

Weekday, Winter, High, ~~late~~ Heavy, ???

Case 1 -

Class = ontime

$$= 0.70 \times 0.64 \times 0.14 \times 0.29 \times 0.36 \times 0.02$$

$$= 0.0013$$

Case 2 - Class = late.

$$= 0.10 \times 0.50 \times 1.0 \times 0.50$$

$$\times 0.50 = 0.0125$$

Case 3 - Class = Very late.

$$= 0.15 \times 1.0 \times 0.67 \times 0.33 \times 0.67$$

$$=$$

Case 4 \Rightarrow Class = Cancelled:

$$= 0.05 \times 0.0 \times 0.0 \times 1.0 \times 1.0$$

$$= 0.0$$

Hence,

For the instance - Weekday, Winter, High, Heavy, ???.

The correct classification is Very late.

(2)

	male	female	total
fiction	250 (90)	200 (360)	450
non fiction	50 (210)	1000 (840)	1050
Total	300	1200	1500

H₀ :- The two attributes preferred reading and gender are correlated.

H₁ :- The preferred reading and age are not correlated.

→ Let's perform correlation analysis on the above data using Chi square test.

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

$$\chi^2 = \frac{(250 - 90)^2}{90} + \frac{(50 - 210)^2}{210} + \frac{(200 - 360)^2}{360} + \frac{(1000 - 840)^2}{840}$$

$$= 284.44 + 121.9 + 71.11 + 30.48$$

$$\boxed{\chi^2 = 507.93}$$

∴ The above correlation analysis shows that preferred reading & gender are correlated.

∴ We accept the ^{null} hypothesis that preferred reading and gender are correlated for the given group.