

As a consultant we have to analyze the data from a contraceptive prevalence survey. The samples are married women who were either not pregnant or do not know if they were at the time of interview. The problem is to predict the use of contraceptive method (Yes, No) of a woman based on her demographic and socio-economic characteristics and data:

#	Age	Child-count	Living-std	Education	Contraceptive
1	40-or-more	0	Low	Low	No
2	40-or-more	0	Low	High	No
3	30-39	0	Low	Low	Yes
4	20-29	1-2	Low	High	Yes
5	20-29	3-or-more	High	High	Yes
6	20-29	3-or-more	High	Low	No
7	30-39	3-or-more	High	Low	Yes
8	40-or-more	1-2	Low	Low	No
9	40-or-more	3-or-more	High	Low	Yes
10	20-29	1-2	High	Low	Yes
11	40-or-more	1-2	High	High	Yes
12	30-39	1-2	Low	High	Yes
13	30-39	0	High	Low	Yes
14	20-29	1-2	Low	Low	No
15	30-39	1-2	High	High	Yes

We first start our prediction by creating the conditional probabilities tables.

Conditional Probabilities tables:

Age	Total	Yes	No				
20-29	5	3	2	P(20-29)	0.333333	P(20-29 Yes)	0.222222
30-39	5	5	0	P(30-39)	0.333333	P(30-39 Yes)	0.166667
40 or more	5	2	3	P(40 or more)	0.333333	P(40 or more Yes)	0.266667
Child-count	Total	Yes	No				
0	4	2	2	P(0)	0.266667	P(0 Yes)	0.253333
1-2	7	5	2	P(1-2)	0.466667	P(1-2 Yes)	0.183333
3 or more	4	3	1	P(3 or more)	0.266667	P(3 or more Yes)	0.211111

living std	Total	Yes	No				
Low	7	3	4	P(Low)	0.466667	P(Low Yes)	0.293333
High	8	7	1	P(High)	0.533333	P(High Yes)	0.17037
Education	Total	Yes	No				
Low	9	5	4	P(Low)	0.6	P(Low Yes)	0.228571
High	6	5	1	P(High)	0.4	P(High Yes)	0.2

P(20-29 Yes)	0.222222
P(30-39 Yes)	0.166667
P(\$0 or more Yes)	0.266667

P(20-29 No)	0.266667
P(30-39 No)	0.444444
P(\$0 or more No)	0.222222

P(0 Yes)	0.253333
P(1-2 Yes)	0.183333
P(3 or more Yes)	0.211111

P(0 No)	0.253333
P(1-2 No)	0.733333
P(3 or more No)	0.316667

P(Low Yes)	0.293333
P(High Yes)	0.17037

P(Low No)	0.244444
P(High No)	0.511111

P(Low Yes)	0.228571
P(High Yes)	0.2

P(Low No)	0.266667
P(High No)	0.466667

Here we have instance X =(Age=20-29, Education=High, Child-count=3-or-more, Living-Std=High)

P(X (Age=20-29, Education=High, Child-count=3-or-more, Living-Std=High) Yes)	0.001599
And	
P(X (Age=20-29, Education=High, Child-count=3-or-more, Living-Std=High) No)	0.020142

Computing the probability of our prediction:

$$P(X|Yes) = 0.001599 / (0.001599 + 0.020142) = 7.3547\%$$

$$P(X|No) = 0.020142 / (0.001599 + 0.020142) = 92.645\%$$

From our above prediction we have 92.65% of confidence that given sample X is classified as No.
Please refer to the attached excel file to see the complete calculations.

Now With our solver using TAN option we have the below stats, which give us 91.6% classified as No to our given instance. Please refer to the attached stream file for details.

Analysis		
Annotations		
Collapse All		
Expand All		
Results for output field Contraceptive		
Comparing \$B-Contraceptive with Contraceptive		
Correct	14	93.33%
Wrong	1	6.67%
Total	15	
Coincidence Matrix for \$B-Contraceptive (rows show actuals)		
	No	Yes
No	5	0
Yes	1	9
Performance Evaluation		
No	0.916	
Yes	0.405	

