Week 3 Probability Assignment

Benjamin Siebold

Sep, 08, 2020

Introduction

Question 1

1. In a study of pleas and prison sentences, it is reported that 42% of the subjects were sent to prison. Among those sent to prison, 38% plead guilty. Among those not sent to prison, 50% plead guilty.

With the above information, the following probability assignments are made:
P(P) = sent to prison
P(!P) = not sent to prison
P(G|P) = plead guilty prison
P(G|!P) = plead guilty no prison
P(!G|P) = plead not guilty prison
P(!G|!P) = plead not guilty no prison

a. If a subject is randomly selected, what is the probability of getting a person who was not sent to prison.
P(!P) = 1 - P(P)

```
1-.42
```

```
## [1] 0.58
```

P(!P) = .58

b. If a subject is randomly selected, and it is known that the subject entered a guilty plea, what is the probability that this subject was not sent to prison.

unknown: P(!P|G)
known:
P(P) = .42
P(!P) = .58
P(G|P) = .38
P(!G|P) = .62
P(G|!P) = .50
P(!G|!P) = .50

equation: P(!P|G) = P(G|!P)P(!P)/(P(G|!P)P(!P)+P(G|P)P(P))

```
(.58)*(.50)/((.58)*(.50)+(.42)*(.38))
```

```
## [1] 0.6450178
```

P(P|G) = .645

c. If a subject is randomly selected, what is the probability of getting someone who was sent to prison.

The answer is given in the description of the problem: .42

d. If a subject is randomly selected, and it is known that the subject entered a guilty plea, what is the probability that this person was sent to prison.

unknown: P(P|G)
known:
P(P) = .42
P(!P) = .58
P(G|P) = .38
P(!G|P) = .62
P(G|!P) = .50
P(!G|!P) = .50

equation: P(P|G) = P(G|P)P(P)/(P(G|P)P(P)+P(G|!P)P(!P))

```
(.38)*(.42)/((.38)*(.42)+(.50)*(.58))
```

```
## [1] 0.3549822
```

P(P|G) = .355

Question 2

Given the following table:

Customer ID	Gender	Car Type	Shirt Size	Class
1	M	Family	Small	C0
2	M	Sports	Medium	C0
3	M	Sports	Medium	C0
4	M	Sports	Large	C0
5	M	Sports	Extra Large	C0
6	M	Sports	Extra Large	C0
7	F	Sports	Small	C0
8	F	Sports	Small	C0
9	F	Sports	Medium	C0
10	F	Luxury	Large	C0
11	M	Family	Large	C1
12	M	Family	Extra Large	C1
13	M	Family	Medium	C1
14	M	Luxury	Extra Large	C1
15	F	Luxury	Small	C1
16	F	Luxury	Small	C1
17	F	Luxury	Medium	C1
18	F	Luxury	Medium	C1
19	F	Luxury	Medium	C1
20	F	Luxury	Large	C1

a. What is the value of each of the following probability?

The below can be determined by looking at the number of rows in which both conditions are true. For example in p1 there is a probability that Gender = M of 10/20, but only 6 of those rows have class = C0, thus, 6/20.

P(Gender = M | Class = C0) = 6/20

P(Gender = F | Class = C1) = 6/20

P(Car Type = Family | Class = C0) = 1/20

P(Car Type = Family | Class = C1) = 3/20

P(Shirt Size = Medium | Class = C0) = 3/20

P(Shirt Size = Medium | Class = C1) = 4/20

b. Use Naive Bayes' Classifier to find the class of:

P(Gender=F, Car Type = Family, Shirt Size = Medium)

Class=C0

P(X|Class=C0) = P(Gender=F|Class=C0) *
P(Car Type=Family|Class=C0) *
P(Shirt Size = Medium|Class=C0)

```
(4/20)*(1/20)*(3/20)
```

```
## [1] 0.0015
```

Class=C1

P(X|Class=C1) = P(Gender=F|Class=C1) *
P(Car Type=Family|Class=C1) *
P(Shirt Size = Medium|Class=C1)

```
(6/20)*(3/20)*(4/20)
```

```
## [1] 0.009
```

Because P(X|Class=C1) > P(X|Class=C0) (0.009 > 0.0015)

the class of P(Gender=F, Car Type = Family, Shirt Size = Medium) = C1