1 Marginalization

1.1 Review

Assume you have a problem statement involving two variables, X and Y. But we only want to know the probability of X. Then we can marginalize, or sum out all possibilities of Y. For example. $P(X) = P(X, y) + P(X, \neg y)$

Written more generally:

$$P(A) = \sum_{B} P(A, B)$$

1.2 Examples

Vars	P(Vars)
(a, b, c)	x_1
$(a, b, \neg c)$	x_2
$(a, \neg b, c)$	x_3
$(a, \neg b, \neg c)$	x_4
$(\neg a, b, c)$	x_5
$(\neg a, b, \neg c)$	x_6
$(\neg a, \neg b, c)$	x_7
$(\neg a, \neg b, \neg c)$	x_8

What is $P(a, \neg b)$? $x_3 + x_4$. What about $P(\neg a)$?

2 Independence

2.1 Review

Independence is met if the following rule applies:

$P(A, B) = P(A)P(B)$

2.2 Problem Statement

Consider we wanted to know the probability you had a good break (G). Additionally, we know whether your flight got delayed (D) and whether you got sick over break (S).

2.3 Delayed Flight and Good Break

Assume you have marginalized out sickness, and have a distribution over delayed flight and good break.

Vars	P(Vars)
(g, d)	0.08
(g, d)	0.02
$(\neg g, d)$	0.72
$(\neg g, \neg d)$	0.18

Are these two variables independent? Calculate $P(G)$ and $P(D)$ and see if the conditional relationship holds.

$P(g, d) = P(g)P(d)$
$P(g, \neg d) = P(g)(1 - P(d))$
$P(\neg g, d) = (1 - P(g))P(d)$
$P(\neg g, \neg d) = (1 - P(g))(1 - P(d))$

Answer: $P(g) = 0.08 + 0.02 = 0.1$, $P(d) = 0.72 + 0.8 = 0.8$. Yes, they are independent. Plugging in values into above rules all are equivalent.

2.4 Sickness and Good Break

Assume you have marginalized out delayed flight, and have a distribution over sickness and good break.

Vars	P(Vars)
(s, g)	0.1
$(s, \neg g)$	0.3
$(\neg s, g)$	0.4
$(\neg s, \neg g)$	0.2

Are these two variables independent? Calculate $P(S)$ and $P(G)$ and see if the conditional relationship holds.

$P(s) = 0.1 + 0.3 = 0.4$, $P(g) = 0.1 + 0.4 = 0.5$
$P(s, g)0.1 = P(s)P(g) = 0.4 * 0.5 = 0.2$...not equal!

Answer: No, they are not independent.

3 Conditional Independence

3.1 Review

Conditional independence is met if the following rules apply:

$P(A|B, C) = P(A|C)$ or
$P(A, B|C) = P(A|C)P(B|C)$

In both examples you would say A is conditionally independent of B given C. In other words, knowing B gives no additional information if we already know C.

3.2 Deriving Second Cond. Indep Equation Using the First

$$\begin{aligned}
P(A, B|C) &= \frac{P(A, B, C)}{P(C)} \\
&= \frac{P(A|B, C)P(B, C)}{P(C)} \\
&= \frac{P(A|B, C)P(B|C)P(C)}{P(C)} \\
&= P(A|B, C)P(B|C) \\
&= P(A|C)P(B|C) \quad \text{Using first rule}
\end{aligned}$$

You can also derive it more succinctly by just stating:

$$\begin{aligned}
P(A, B|C) &= P(A|B, C)P(B|C) \\
&= P(A|C)P(B|C) \quad \text{Using first rule}
\end{aligned}$$

3.3 Example Problem

Consider the following problem:

We are trying to determine the probability that a student will do well on the second midterm (W). The two other variables we know are whether the student understood everything in class (C), and whether the student understood the practice midterm (M).

Assume we know the following probabilities:

$P(w|c) = 0.6$
$P(m|c) = 0.8$
$P(w|\neg c) = 0.1$
$P(m|\neg c) = 0.2$

3.4 If W is Conditionally Independent of M Given C...

If W is conditionally independent of M given C, what can we say about the values $P(W, M|C)$ must take on? Write out the truth table for $P(W, M|C)$:

3.5 For Example

$P(w, m|c)$ must equal $P(m|c) * P(w|c)$ or 0.48

$P(w, \neg m|\neg c)$ must equal $P(w|\neg c) * P(\neg m|\neg c)$ which is equivalent to $P(w|\neg c) * (1 - P(m|\neg c))$ or 0.08

And so on