CS4LL5 Advanced Computational Linguistics

Séamus Woods
15317173

23/09/2019

0.1 Question 1

Consider the following equations:

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

ii $P(A|B) = P(A)$

Show that (i) implies (ii), and also that (ii) implies (i).

Answer:
If $P(A,B)$ happening is the same as $P(A) * P(B)$, then A and B are independent. Similarly if $P(A|B)$ is the same as $P(A)$, then A and B are independent.

(i) implies (ii):

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

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

$$P(A) * P(B) = P(A|B) * P(B)$$

$$P(A) = P(A|B)$$

$$P(A|B) = P(A)$$

(ii) implies (i):

$$P(A|B) = P(A)$$

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

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

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

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

0.2 Question 2

(a) Calculate $P(gw|ps)$:
Sample space of ps is 30, and out of those 30, gw happened 28 times, therefore $P(gw|ps) = 28/30$.
The sample space of not-ps is irrelevant since we're working with the probability gw happened <u>given</u> ps already happened.

(b) Calculate $P(ps|gw)$
Sample space of gw is 168, and out of that ps occurred 28 times. Therefore $P(ps|gw) = 28/168$.
The sample space of not-gw is irrelevant here because we are only concerned about the chance ps occurred <u>given</u> gw has already occurred.

0.3 Question 3

Let vmel stand for Speaker = 'Victor Meldrew'.
Let dbi stand for DBI = true.

Work out which of vmel or not-vmel is likelier, given dbi, supposing the following probabilities:
(a)
$P(vmel) = 0.01$
$P(dbi|vmel) = 0.95$
$P(dbi|notvmel) = 0.01$

$P(dbi|vmel)P(vmel) = 0.95 * 0.01 = 0.0095$
$P(dbi|notvmel)P(notvmel) = 0.01 * 1 - 0.01 = 0.0099$
Therefore not-vmel is likelier given dbi.

(b)
$P(vmel) = 0.15$
$P(dbi|vmel) = 0.95$
$P(dbi|notvmel) = 0.01$

$P(dbi|vmel)P(vmel) = 0.95 * 0.15 = 0.1425$
$P(dbi|notvmel)P(notvmel) = 0.01 * 1 - 0.15 = 0.0.0085$

Therefore vmel is likelier given dbi.

(c)
$P(vmel) = 0.01$
$P(dbi|vmel) = 0.95$
$P(dbi|notvmel) = 0.001$

$P(dbi|vmel)P(vmel) = 0.95 * 0.01 = 0.0095$
$P(dbi|notvmel)P(notvmel) = 0.001 * 1 - 0.01 = 0.00099$
Therefore vmel is likelier given dbi.

0.4 Question 4

Find $P(cool+)$ and $P(cool+|noisy+)$ and conclude from this whether or not $cool+$ is independent of $noisy+$

$P(cool+) = 170/500 = 17/50$
$P(cool + |noisy+) = 62/100$
A condition of independence is $P(A|B) = P(A)$, so if $cool+$ and $noisy+$ are independent they must follow this rule.
$P(cool+) = P(cool + |noisy+)$
$17/50 = 62/100$ is not true, and therefore $cool+$ is not independent of $noisy+$.

0.5 Question 5

With reference to table 2, find $P(cool + |open+)$, $P(cool + |open+, noisy+)$.
$P(cool + |open+) = 90/100$
$P(cool + |open+, noisy+) = 54/60$
Is $cool+$ conditionally independent of $noisy+$ given $open+$?
Conditional independence $= P(X|Y, Z) = P(X|Z)$.
Let $X = cool+$, $Y = noisy+$ and $Z = open+$.
$P(cool + |open+, noisy+) = P(cool+, open+)$
$90/100 = 54/60$
$0.9 = 0.9$

Therefore $cool+$ is conditionally independent of $noisy+$ given $open+$.

With reference to table 3, find $P(cool+|open-)$ and $P(cool+|open-, noisy+)$.
$P(cool+|open-) = 80/400$
$P(cool+|open-, noisy+) = 8/40$
Is $cool+$ conditionally independent of $noisy+$ given $open-$?
$P(cool+|open-, noisy+) = P(cool+|open-)$
$8/40 = 80/400$
$0.2 = 0.2$
Therefore, $cool+$ is conditionally independent of $noisy+$ given $open-$.