Sanjana Agarwal
Username: sa14593

ADVANCED NLP ASSIGNMENT 1

1) Let P(A) = P(three-letter-word) = 0.0003
Let P(B) = P(is-abbreviation)
Let P(B|A) = P(is-abbreviation|three-letter-word) = 0.8

Therefore, P(A∩B) = ?
According to Bayes Theorem,

P(B|A) = $\left(\frac{P(A \bigcap B)}{P(A)}\right)$

P(A∩B) = P(A) * P(B|A)
= 0.0003 * 0.8
= 0.00024
Thus the probability is 0.00024.

2) For X and Y to be independently distributed, we need,
P(X∩Y) = P(X) * P(Y)
The marginal distributions of X and Y are,

$$P_X(x) = \sum_y p(x, y)$$

$$P_Y(y) = \sum_x p(x, y)$$

Thus, we derive,

x	P(X=x)
0	0.4
1	0.6

y	P(Y=y)
0	0.8
1	0.2

P(X) = 0.6, P(Y) = 0.2,
Thus, P(X) * P(Y) = 0.12

And from the given table in the question, P(X∩Y) = 0.12 (where P(X=1,Y=1))

Thus, X and Y are independently distributed.

3 a) Let P(A) = At least one dice lands on 1.
Let P(B) = Both dices land on different numbers, that is, excluding all the doubles = $1-\frac{1}{6}=\frac{5}{6}$

Thus, P(A∩B) = P(At-lease-one-on-a-dice ∩different-numbers-on-dice)
The event space is,
(1,2), (1,3), (1,4), (1,5), (1,6)
(2,1)
(3,1)
(4,1)
(5,1)
(6,1)
= $\frac{10}{36}=\frac{5}{18}$
Thus, $P(A|B)=\frac{P(A\bigcap B)}{P(B)}$
= $\frac{\frac{5}{18}}{\frac{5}{6}}=\frac{1}{3}$

Thus, the probability is $\frac{1}{3}$.

3 b) Let P(A) be the probability that bulb works for a week.
Let P(B) be the probability that the bulb lights up initially
= no. of bulbs that lit up initially / total no. of bulbs
$=\frac{10+5}{25}=\frac{15}{25}=\frac{3}{5}$

Now, $P(A\bigcap B)=\frac{5}{25}=\frac{1}{5}$

Thus, using Bayes theorem,
$P(A|B)=\frac{P(A\bigcap B)}{P(B)}=\frac{\frac{1}{5}}{\frac{3}{5}}=\frac{1}{3}$

Thus, the probability is $\frac{1}{3}$.

4) Let P(A) be the probability that the writer is American = 0.6
Let P(B) be the probability that the writer is British = 0.4

Let P(C) be the probability that a letter taken at random from the spelling is a vowel

= letter taken from Britisher's spelling and turns out to be a vowel **OR** letter taken from the American's spelling and turns out to be a vowel.
= P(B) * P(C|B) + P(A) * P(C|A) ---- Eqn. 1

Now, P(C|B) = no. of vowels in RIGOR / total no. of alphabets in RIGOR

$= \frac{2}{5}$

Similarly, P(C|A) = no. of vowels in RIGOUR / total no. of alphabets in RIGOUR
$= \frac{1}{2}$

Thus, getting back to eqn. 1, we get,
$$P(C) = 0.4 * \frac{1}{2} + 0.6 * \frac{2}{5}$$
P(C) = 0.44

Now, probability that the writer is British,
$$P(B|C) = \frac{P(B \bigcap C)}{P(C)} = \frac{P(B) * P(C|B)}{P(C)} = \frac{0.4 * 0.5}{0.44} = \frac{5}{11}$$

Thus, probability that the writer is British is $\frac{5}{11}$.