Probability Theory

Contents

1 Bayes' Theorem and Independence

From StackExchange.

1.1 Question

I'm deriving Probabilistic latent semantic analysis model. In the model, documents d and words w are observed.

$$\begin{aligned}
\Pr(d, w) &= \sum_c \Pr(d, w, c) \\
&= \sum_c \Pr(d) \Pr(w, c|d) \\
&= \Pr(d) \sum_c \frac{\Pr(w, c, d)}{\Pr(d)} \\
&= \Pr(d) \sum_c \Pr(w|c, d) \Pr(c|d)
\end{aligned}$$

I used Bayes' theorem from the second line to the third. However, Wikipedia says $\Pr(d) \sum_c \Pr(w|c) \Pr(c|d)$. Is this because d and w are independent? If so, how can I know the independence from the grahical model?

1.2 Answer 1

From what Wikipedia says, d and w are assumed to be conditionally independent given c. This means that

$$\Pr(d, w, c) = \Pr(c) \Pr(d|c) \Pr(w|c).$$

Now using Bayes's theorem we obtain that this quantity equals $\Pr(d) \Pr(c|d) \Pr(w|c)$.

You could also continue your line of thought, because conditional independence is equivalent to saying that $\Pr(w|c, d) = \Pr(w|c)$.

1.3 Answer 2

Immediately skip to the last line by just applying the definition of conditional probability.

$$
\begin{aligned}
\Pr(d, w) &= \sum_c \Pr(d, w, c) \\
&= \sum_c \Pr(d) \Pr(w, c \mid d) \\
&= \Pr(d) \sum_c \Pr(w \mid c, d) \Pr(c \mid d)
\end{aligned}
$$

Anyhow, to say that $\Pr(w \mid c, d) = \Pr(w \mid c)$ is to assert that w, d are conditionally independent for any given c. Conditional Independence of w, d given c is defined as when: $\Pr(w, d \mid c) = \Pr(w \mid c) \Pr(d \mid c)$

So we have:

$$
\begin{aligned}
\Pr(d, w) &= \sum_c \Pr(d, w, c) && \text{Law of Total Probability} \\
&= \sum_c \Pr(c) \Pr(d, w \mid c) && \text{defn. Conditional Probability} \\
&= \sum_c \Pr(c) \Pr(d \mid c) \Pr(w \mid c) && \text{because conditional independence} \\
&= \sum_c \Pr(c \mid d) \Pr(d) \Pr(w \mid c) && \text{Bayes' Rule} \\
&= \Pr(d) \sum_c \Pr(w \mid c) \Pr(c \mid d)
\end{aligned}
$$