Independence and Conditional Probability

CS 2800: Discrete Structures, Spring 2015

Sid Chaudhuri
Independence of Events

Two events $A$ and $B$ in a probability space are independent if and only if

$$P(A \cap B) = P(A) \cdot P(B)$$

Mathematical definition of independence
WTF?

Why does this even make sense?
Independence of Events

\[ P(A \cap B) = P(A)P(B) \]
Independence of Events

\[ P(A \cap B) = P(A) P(B) \]

if and only if

\[ P(B) = \frac{P(A \cap B)}{P(A)} \]

(assuming \( P(A) \neq 0 \))
Independence of Events

\[ P(A \cap B) = P(A)P(B) \]

if and only if

\[ P(B) = \frac{P(A \cap B)}{P(A)} \]

(assuming \( P(A) \neq 0 \))
Conditional Probability

The conditional probability of $B$, given $A$, is written

$$P(B|A)$$
Conditional Probability

The **conditional probability** of $B$, given $A$, is written

$$P(B|A)$$

The probability of
The **conditional probability** of $B$, given $A$, is written

$$P(B|A)$$
Conditional Probability

The **conditional probability** of $B$, given $A$, is written

$$P(B|A)$$

The probability of $B$ given

The probability of $B$, given
Conditional Probability

The conditional probability of \( B \), given \( A \), is written

\[
P(B|A)
\]
Conditional Probability

The conditional probability of B, given A, is written

\[ P(B|A) \]

and defined as

\[ \frac{P(A \cap B)}{P(A)} \]
Conditional Probability

The **conditional probability** of B, given A, is written

\[ P(B|A) \]

and defined as

\[ \frac{P(A \cap B)}{P(A)} \]
WTF #2?

Why does this make sense?
Intuitively, \( P(B \mid A) \) is the probability that event \( B \) occurs, given that event \( A \) has already occurred

(This is NOT the formal math definition)

\( \text{(A and } B \text{ need not actually occur in temporal order) } \)
Cases where, given that $A$ happens, $B$ also happens.
$A \cap B$
$S$ acts as new sample space ("universe of outcomes where $A$ happens")
acts as new sample space
(“universe of outcomes
where A happens”)

all outcomes where B
happens, in this restricted
space, i.e. given that A is
known to have happened
Thought for the Day #1

If the conditional probability $P(B \mid A)$ is defined as $P(A \cap B) / P(A)$, and $P(A) \neq 0$, then show that $(A, Q)$, where $Q(B) = P(B \mid A)$, is a valid probability space satisfying Kolmogorov's axioms.
Independence of Events

\[ P(A \cap B) = P(B \mid A) \ P(A) \]

(by definition)

\[ P(A \cap B) = P(B) \ P(A) \]

(if independent)
Independence of Events

In other words, assuming $P(A) \neq 0$, $A$ and $B$ are independent if and only if

$$P(B \mid A) = P(B)$$
Independence of Events

In other words, assuming $P(A) \neq 0$, $A$ and $B$ are independent if and only if

$$P(B \mid A) = P(B)$$

(Intuitively: the probability of $B$ happening is unaffected by whether $A$ is known to have happened)
Independence of Events

In other words, assuming \( P(A) \neq 0 \), \( A \) and \( B \) are independent if and only if

\[
P(B \mid A) = P(B)
\]

(Intuitively: the probability of \( B \) happening is unaffected by whether \( A \) is known to have happened)

(Note: \( A \) and \( B \) can be swapped, if \( P(B) \neq 0 \))
Bayes' Theorem

Assuming $P(A), P(B) \neq 0,$

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)}$$
Bayes' Theorem

Assuming $P(A), P(B) \neq 0$,

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

since $P(A | B) P(B) = P(A \cap B) = P(B | A) P(A)$

(by definition of conditional probability)
Bayes' Theorem

Assuming $P(A), P(B) \neq 0$,

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

since $P(A | B) P(B) = P(A \cap B) = P(B | A) P(A)$

(by definition of conditional probability)
Bayes' Theorem

Assuming $P(A), P(B) \neq 0$, since

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

(by definition of conditional probability)

\begin{align*}
\text{Prior probability of } A \\
\text{Posterior probability of } A, \text{ given evidence } B
\end{align*}
How do we estimate $P(B)$?

- Theorem of Total Probability (special case):

  If $P(A) \neq 0$ or $1$,

  $$P(B) = P((B \cap A) \cup (B \cap A'))$$
  $$= P(B \cap A) + P(B \cap A')$$  \hspace{1cm} \text{(Axiom 3)}
  $$= P(B \mid A) P(A) + P(B \mid A') P(A')$$ \hspace{1cm} \text{(Definition of conditional probability)}
Example: Medical Diagnosis

• Suppose:
  – 1 in 1000 people carry a disease, for which there is a pretty reliable test
Example: Medical Diagnosis

• Suppose:
  – 1 in 1000 people carry a disease, for which there is a pretty reliable test
  – Probability of a false negative (carrier tests negative) is 1% (so probability of carrier testing positive is 99%)
Example: Medical Diagnosis

• Suppose:
  
  – 1 in 1000 people carry a disease, for which there is a pretty reliable test
  
  – Probability of a false negative (carrier tests negative) is 1% (so probability of carrier testing positive is 99%)
  
  – Probability of a false positive (non-carrier tests positive) is 5%
Example: Medical Diagnosis

• Suppose:
  – 1 in 1000 people carry a disease, for which there is a pretty reliable test
  – Probability of a false negative (carrier tests negative) is 1% (so probability of carrier testing positive is 99%)
  – Probability of a false positive (non-carrier tests positive) is 5%

• A person just tested positive. What are the chances (s)he is a carrier of the disease?
Example: Medical Diagnosis

- P(Carrier) = 0.001
- P(NotCarrier) = 1 - 0.001 = 0.999
Example: Medical Diagnosis

- **Priors:**
  - \( P(Carrier) = 0.001 \)
  - \( P(NotCarrier) = 1 - 0.001 = 0.999 \)

- **Conditional probabilities:**
  - \( P(Positive \mid Carrier) = 0.99 \)
  - \( P(Positive \mid NotCarrier) = 0.05 \)
Example: Medical Diagnosis

\[ P(Carrier \mid Positive) = \frac{P(Positive \mid Carrier) \cdot P(Carrier)}{P(Positive)} \]  
(by Bayes' Theorem)
Example: Medical Diagnosis

\[ P(\text{Carrier} \mid \text{Positive}) \]

\[ = \frac{P(\text{Positive} \mid \text{Carrier}) \ P(\text{Carrier})}{P(\text{Positive})} \]

(by Bayes' Theorem)

\[ = \frac{P(\text{Positive} \mid \text{Carrier}) \ P(\text{Carrier})}{P(\text{Positive} \mid \text{Carrier}) \ P(\text{Carrier}) + P(\text{Positive} \mid \text{NotCarrier}) \ P(\text{NotCarrier})} \]

(by Theorem of Total Probability)
Example: Medical Diagnosis

\[
\frac{P(\text{Positive} \mid \text{Carrier}) \cdot P(\text{Carrier})}{P(\text{Positive} \mid \text{Carrier}) \cdot P(\text{Carrier}) + P(\text{Positive} \mid \text{NotCarrier}) \cdot P(\text{NotCarrier})} = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.05 \times 0.999} = 0.0194
\]
I used to think correlation implied causation.

Then I took a statistics class. Now I don’t.

Sounds like the class helped. Well, maybe.