Probability Axioms, Conditional Probability, Statistical (In)dependence, Bayes Theorem
The **conditional probability** of an event $B$ given an event $A$, denoted as $P(B|A)$, is

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

for $P(A) > 0$.

This definition can be understood in a special case in which all outcomes of a random experiment are equally likely. If there are $n$ total outcomes,

$$P(A) = \frac{\text{number of outcomes in } A}{n}$$

Also,

$$P(A \cap B) = \frac{\text{number of outcomes in } A \cap B}{n}$$

Consequently,

$$\frac{P(A \cap B)}{P(A)} = \frac{\text{number of outcomes in } A \cap B}{\text{number of outcomes in } A}$$

Therefore, $P(B|A)$ can be interpreted as the relative frequency of event $B$ among the trials that produce an outcome in event $A$. 
Statistically independent events

Always true: \( P(A \cap B) = P(A | B) \cdot P(B) = P(B | A) \cdot P(A) \)

- **Two events**

Two events are **independent** if any one of the following equivalent statements is true:

1. \( P(A | B) = P(A) \)
2. \( P(B | A) = P(B) \)
3. \( P(A \cap B) = P(A)P(B) \)

- **Multiple events**

The events \( E_1, E_2, \ldots, E_n \) are independent if and only if for any subset of these events \( E_{i_1}, E_{i_2}, \ldots, E_{i_k} \):

\[
P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) = P(E_{i_1}) \times P(E_{i_2}) \times \cdots \times P(E_{i_k})
\]
Bayes’ theorem (1812)

Thomas Bayes
(1701-1761)
English statistician, philosopher, and Presbyterian minister
Bayes’ theorem (simple)

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

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

- In Science we often want to know: “How much faith should I put into hypothesis, given the data?” or \( P(H|D) \)
- What we usually can calculate if the hypothesis/model is OK: “Assuming that this hypothesis is true, what is the probability of the observed data?” or \( P(D|H) \)
- Bayes’ theorem can help: \( P(H|D) = P(D|H) \cdot P(H)/P(D) \)
- The problem is \( P(H) \) (so-called prior) is often not known
Bayes theorem (continued)

Works best with exhaustive and mutually-exclusive hypotheses: $E_1, E_2, \ldots, E_n$ such that $E_1 \cup E_2 \cup E_3 \ldots \cup E_n = S$ and $E_i \cap E_j = \emptyset$ for $i \neq j$

$$P(E_1 | B) = P(B | E_1) \cdot P(E_1) / P(B)$$

where:

$$P(B) = P(B | E_1) \cdot P(E_1) + P(B | E_2) \cdot P(E_2) + \ldots + P(B | E_n) \cdot P(E_n)$$

$$B = (B \cap E_1) \cup (B \cap E_2) \cup (B \cap E_3) \cup (B \cap E_4)$$
An **awesome new test** has been invented for an early detection of cancer. The probability that it **correctly identifies someone with cancer as positive** is 95%, and the probability that it **correctly identifies someone without cancer as negative** is 99%. The **incidence** of this type of cancer in the general population is $10^{-4}$. A random person in the population takes the test, and the result is positive.

**What is the probability that he/she has cancer?**

A. 99%
B. 95%
C. 30%
D. 1%

Get your i-clickers
Events: \( C \) - cancer, \( C' \) - no cancer
Test events: \( Y \) - positive, \( N \) - negative

We know:

\[
P(C) = 10^{-4}, \quad P(Y|C) = 0.95, \quad P(N|C') = 0.99
\]

We need

\[
P(C|Y)
\]

Bayes:

\[
P(C|Y) = \frac{P(Y|C) \cdot P(C)}{P(Y)}
\]
$P(Y)$—probability that a random person will test positive

$P(Y) = P(Y \cap C) + P(Y \cap C') = P(Y | C) P(C) + P(Y | C') P(C')$

$= 0.95 \times 10^{-4} + (1 - 0.99) \times (1 - 10^{-4}) \approx 10^{-4} + 10^{-2} \approx 10^{-2} = 1\%$

$P(C | Y) = P(Y | C) \cdot \frac{P(C)}{P(Y)} = 0.95 \times \frac{10^{-4}}{10^{-2}} \approx 1\%$
What if a doctor is already 50% sure of cancer based on other tests?

That changes things!

Now \( P(C) = P(C') = 0.5 \)

\[
P(C|y) = \frac{P(y|c) \cdot P(c)}{P(y|c) \cdot P(c) + P(y|c') \cdot P(c')}
\]

\[
= \frac{0.95 \times 0.5}{0.95 \times 0.5 + (1 - 0.99) \times 0.5} \approx 0.99
\]
How come?
We thought it was a great test..

- Let $C$ – be the event that the patient has cancer; $C'$ – patient is cancer free
- $Y/N$ – events that test is Positive/Negative ($N=Y'$)
- We know: $P(C)=10^{-4}$, $P(Y|C)=0.95$, $P(N|C')=0.99$
- We need to find $P(C|Y)$
- Bayes to the rescue: $P(C|Y)=P(Y|C)*P(C)/P(Y)$
- What on earth is $P(Y)$ ???
What on Earth is $P(Y)$ ???

• Likelihood that a random patient would test $Y$:
  
  $P(Y) = P(Y \cap C) + P(Y \cap C') = P(Y|C)P(C) + P(Y|C')P(C') = 0.95 \times 10^{-4} + (1 - 0.99) \times (1 - 10^{-4}) \approx 0.01$

• Hence $P(C|Y) = P(Y|C) \times P(C)/P(Y)$
  
  $\approx 10^{-4}/0.01 = 0.01 = 1\%$

• But we would like it to be 100%, please!!!

• Until the false positive discovery rate $1 - P(N|C')$ does not fall below the general population prevalence the result will never be close 100%
What if I am already 50% sure (based on other tests) that a patient has cancer?

• That changes everything!
• Now $P(C)=P(C')=0.5$
• $P(C|Y)=P(Y|C)*P(C)/[P(Y|C)*P(C)+P(Y|C')*P(C')]=0.95*0.5/[0.95*0.5+(1-0.99)*0.5]=0.99$
• Now the doctor can be almost 100% sure.
• The importance of prior:
  – If prior belief that one has cancer is $10^{-4}$ – test is useless
  – If prior belief is at least 1% - the test is useful
Prostate cancer is the most common type of cancer found in males. It is checked by PSA test that is notoriously unreliable. The probability that a noncancerous man will have an elevated PSA level is approximately 0.135, with this probability increasing to approximately 0.268 if the man does have cancer. If, based on other factors, a physician is 70 percent certain that a male has prostate cancer, what is the conditional probability that he has the cancer given that the test indicates an elevated PSA level?

A. 99.99%
B. 99%
C. 82%
D. 71%

Get your i-clickers
Use Bayes theorem again

Events: $C$ - cancer, $E$ - PSA elevated

$p(C) = 0.7$: doctor's prior belief

$p(C | E) = p(E | C) \cdot \frac{p(C)}{p(E)}$

$p(E) = p(E | C) \cdot p(C) + p(E | C') \cdot p(C')$

$= 0.268 \cdot 0.7 + 0.135 \cdot 0.3 = 0.23$

$p(C | E) = \frac{0.268 \cdot 0.7}{0.23} = 0.82 = 82\%$

VS: 70\%
All this trouble for a lousy 12% gain in confidence? I don’t believe you!

- Let $C$ – be the event that the patient has cancer; $C'$ – patient is cancer free, $E$ – events that the PSA test was elevated.
- We know **doctor’s prior belief**: $P(C)=0.7$
- We know test stats: $P(E|C)=0.268$, $P(E|C')=0.135$
- We need to find $P(C|E)=P(E|C)*P(C)/P(E)$
- $P(E)=P(E|C)*P(C)+P(E|C')*P(C')=0.268*0.7+0.135*0.3=0.23$
- $P(C|E)=0.7*0.268/0.23=0.82=82%$