Probability Axioms, Conditional Probability, Statistical (In)dependence, Circuit Problems

Axioms of probability

Probability is a number that is assigned to each member of a collection of events from a random experiment that satisfies the following properties:

If S is the sample space and E is any event in a random experiment,

(1) $P(S) = 1$

(2) $0 \leq P(E) \leq 1$

(3) For two events E_1 and E_2 with $E_1 \cap E_2 = \emptyset$

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

These axioms imply that:

$$P(\emptyset) = 0$$

$$P(E') = 1 - P(E)$$

if the event E_1 is contained in the event E_2

$$P(E_1) \leq P(E_2)$$

Conditional probability

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,

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

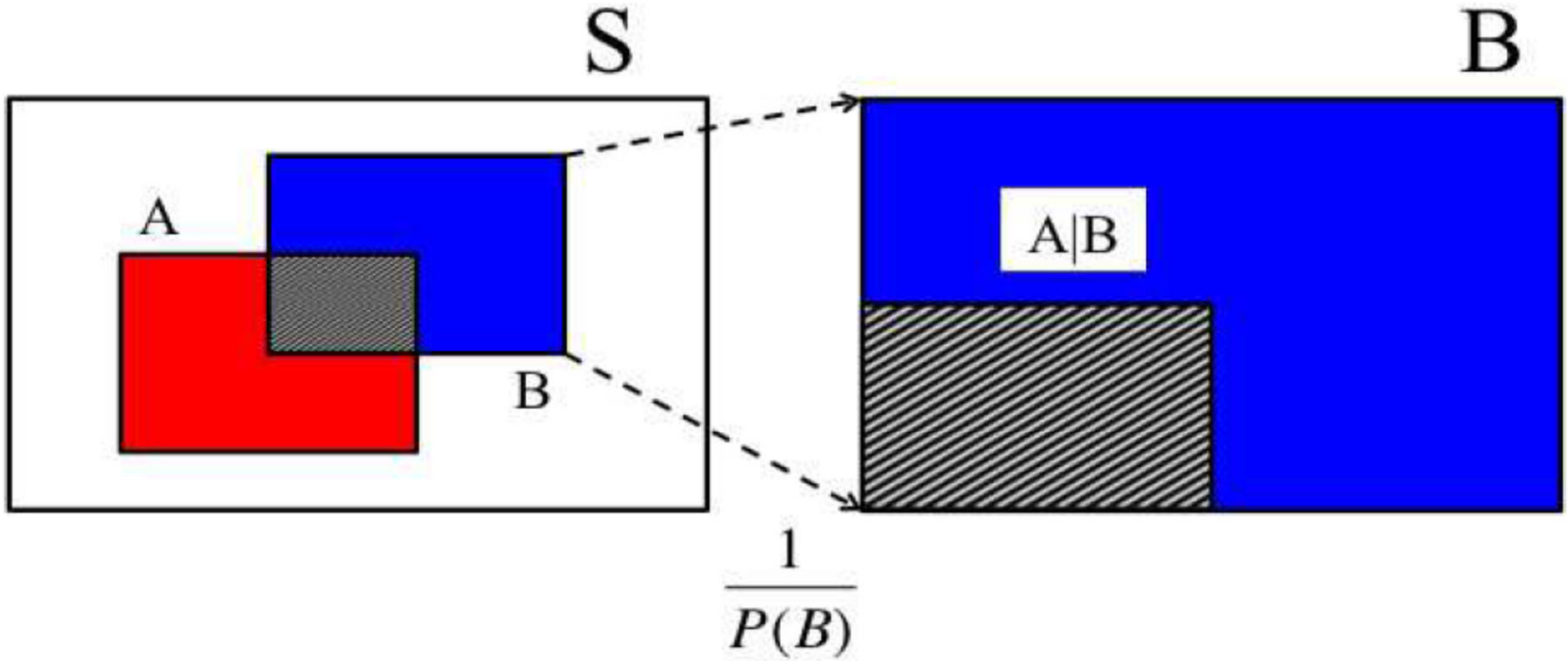
Also,

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

Consequently,

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

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


S
B
A
B
A|B
$\frac{1}{P(B)}$

Multiplication rule

is just definition of conditional probability

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

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

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

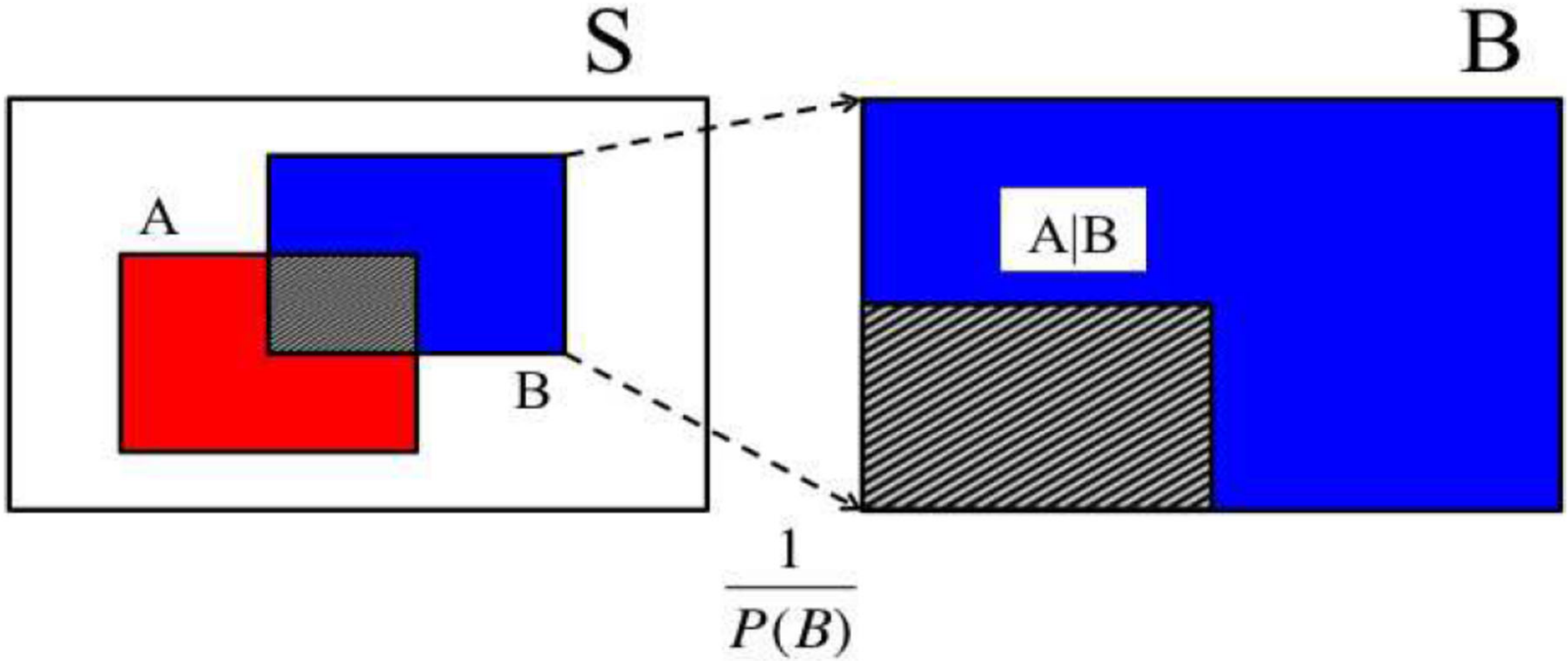
(2) $P(B|A) = P(B)$

(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 $E_{i_1}, E_{i_2}, \ldots, E_{i_k}$,

$$P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) = P(E_{i_1}) \times P(E_{i_2}) \times \cdots \times P(E_{i_k})$$


S
A
B
B
A|B
$\frac{1}{P(B)}$

Example 3.10. Let an experiment consist of drawing a card at random from a standard deck of 52 playing cards. Define events A and B as "the card is a $\spadesuit$" and "the card is a queen." Are the events A and B independent? By definition, $P(A \cdot B) = P(Q\spadesuit) = \frac{1}{52}$. This is the product of $P(\spadesuit) = \frac{13}{52}$ and $P(Q) = \frac{4}{52}$, and events A and B in question are independent. In this situation, intuition provides no help. Now, pretend that the $2\heartsuit$ is drawn and excluded from the deck prior to the experiment. Events A and B become dependent since

$$\mathbb{P}(A) \cdot \mathbb{P}(B) = \frac{13}{51} \cdot \frac{4}{51} \neq \frac{1}{51} = \mathbb{P}(A \cdot B).$$

Bayes' theorem (1812)

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

Bayes' theorem (simple)

Definitions: $P(B|A) = P(B \cap A)/P(A)$; $P(A|B) = P(A \cap B)/P(B)$

$$P(A \cap B) = \underline{P(A|B)P(B)} = P(B \cap A) = \underline{P(B|A)P(A)}$$

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

- 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 is:
  "Assuming that this hypothesis is true, what is the probability of observing this data?" or $P(D|H)$
- Bayes' theorem can help: $P(H|D)=P(D|H)\bullet P(H)/P(D)$
- The problem is $P(H)$ (so-called <u>prior</u>) is often not known

Bayes theorem (continued)

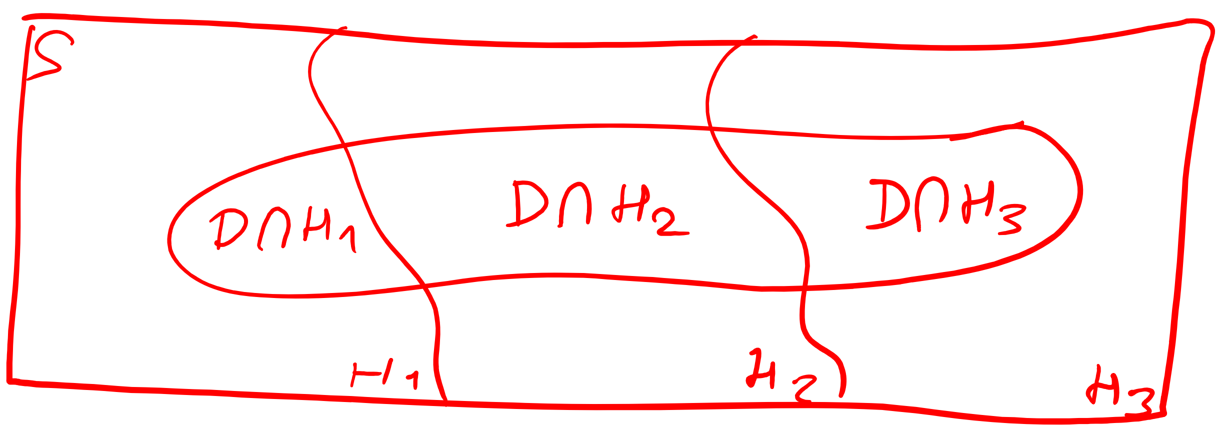
Works best with exhaustive and mutually-exclusive hypotheses:
$H_1, H_2, \dots H_n$ such that $H_1 \cup H_2 \cup H_3 \dots \cup H_n = S$ and $H_i \cap H_j = \emptyset$ for $i \neq j$

$P(H_k|D) = P(D|H_k) \bullet P(H_k)/P(D)$

- $P(H_k)$ is a prior of hypothesis k. But what is P(D)?

- $P(D) = P(D \cap H_1) + P(D \cap H_2) + \dots P(D \cap H_n) =$

$= P(D|H_1) \bullet P(H_1) + P(D|H_2) \bullet P(H_2) + \dots P(D|H_n) \bullet P(H_n)$



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

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

participants

$10^6 <$ 100 - cancer → 95 positive tests

$10^6 - 100 \approx 10^6$ no cancer

10^6 participants with no cancer → 10,000 positive tests

$$P(C|P) = \frac{95}{10{,}000 + 95} \approx 1\%$$

Let’s try to check if it makes sense

- Consider 1,000,000 people from the street
- Only 1e6*1e-4=100 will have this rare cancer
- 95 of them (95%) will be identified by the test
- Out of 1,000,000-100 ≈1,000,000 people without cancer, 1% will get a false positive diagnosis
- That is 10,000 people!!!
- The probability that a random person has cancer if diagnosed is 95/(10,000+95) ≈ 1%

Events: C - cancer, C' - no cancer

Test events P - positive, N - negative

We know:

$P(C) = 10^{-4}$, $P(P|C) = 0.95$

$P(N|C') = 0.99$

We need

$P(C|P)$

Bayes:

$$P(C|P) = P(P|C) \cdot \frac{P(C)}{P(P)} \quad ?$$

$P(P)$ — probability that a random person will test positive

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

$$= P(P|C)P(C) + P(P|C')P(C') =$$

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

$$\approx 10^{-4} + 10^{-2} \approx 10^{-2} = 1\%$$

$$P(C|P) = P(P|C) \cdot \frac{P(C)}{P(P)} = 0.95 \times \frac{10^{-4}}{10^{-2}} \approx$$

$$\approx 1\%$$

An **<u>awesome new test</u>** 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 <u>suspected cancer patient with likelihood of cancer 50%</u> takes the test, and the result is positive.

<u>What is the probability that he/she has cancer?</u>

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

Get your i-clickers

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 suspected cancer patient with likelihood of cancer 50% takes the test, and the result is positive.

What is the probability that he/she has cancer?

A. $\boxed{99\%}$
B. 95%
C. 30%
D. 1%

Get your i-clickers

What if a doctor is already 50% sure of cancer based on other tests?

That changes things!

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

$$P(C \mid P) = \frac{P(P \mid C) \cdot P(C)}{P(P \mid C) \cdot P(C) + P(P \mid C')\, P(C')}$$

$$= \frac{0.95 \times 0.5}{0.95 \times 0.5 + (1 - 0.99) \times 0.5} \approx 0.99$$

How come?

I thought it was a ***great*** test..

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

What on Earth is P(P) ???

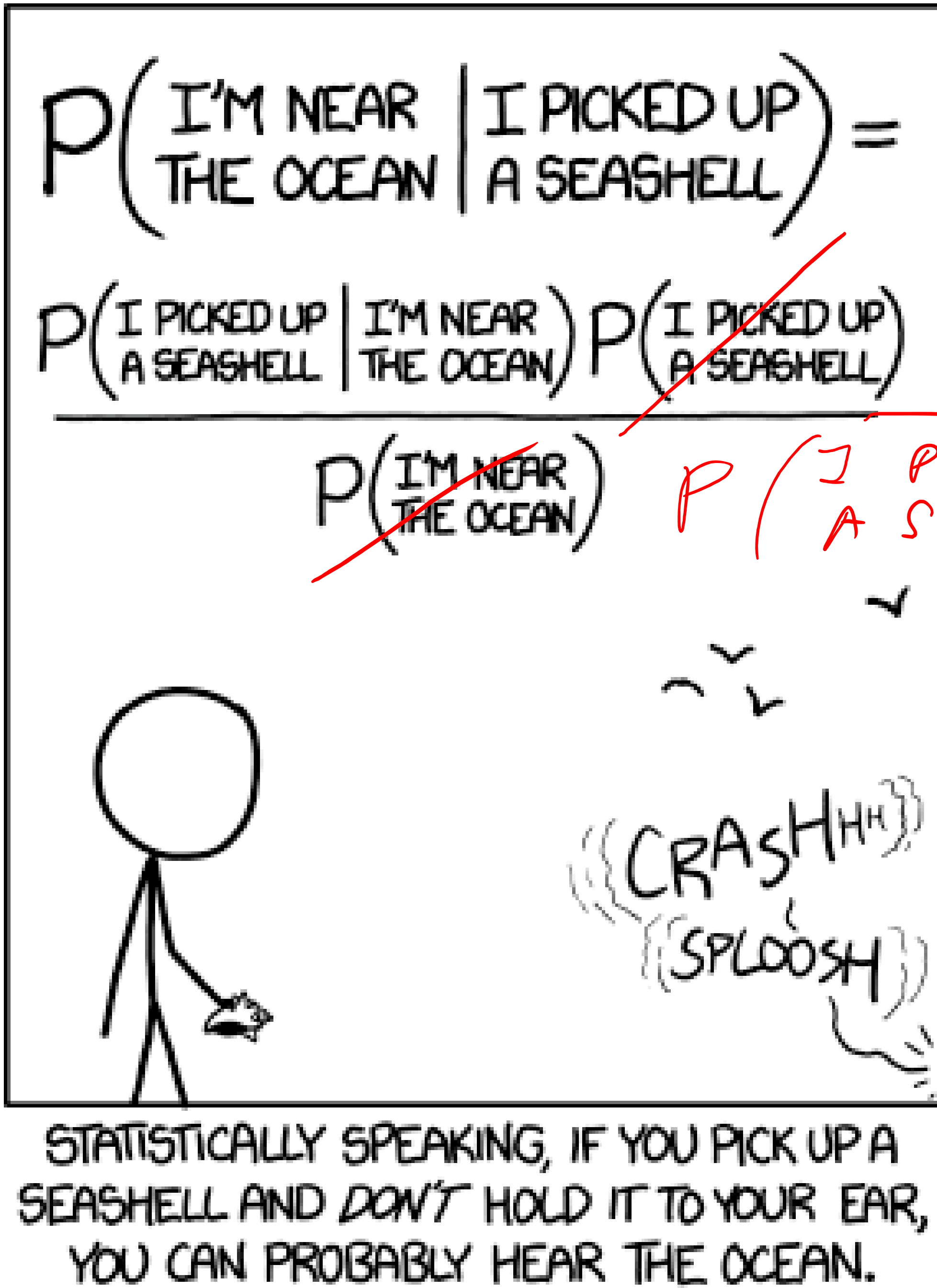
- Likelihood that a random patient would test Y: $P(P)=P(P \cap C)+P(P \cap C')=P(P|C)P(C)+P(P|C')P(C')= 0.95*10^{-4}+(1-0.99)*(1-10^{-4})\approx 0.01$
- Hence $P(C|P)=P(P|C)*P(C)/P(P) \approx 10^{-4}/0.01=0.01=1\%$
- But we would like it to be 100%, please!!!
- Until the false positive discovery rate $1-P(N|C')$ does not fall below the general population prevalence the <u>result will never be close 100%</u>

What if I am already 50% sure (based on other tests) that a patient has cancer?

- That changes everything!
- Now $P(C)=P(C')=0.5$
- $P(C|P)=P(P|C)*P(C)/[P(P|C)*P(C)+ P(P|C')*P(C')]=$

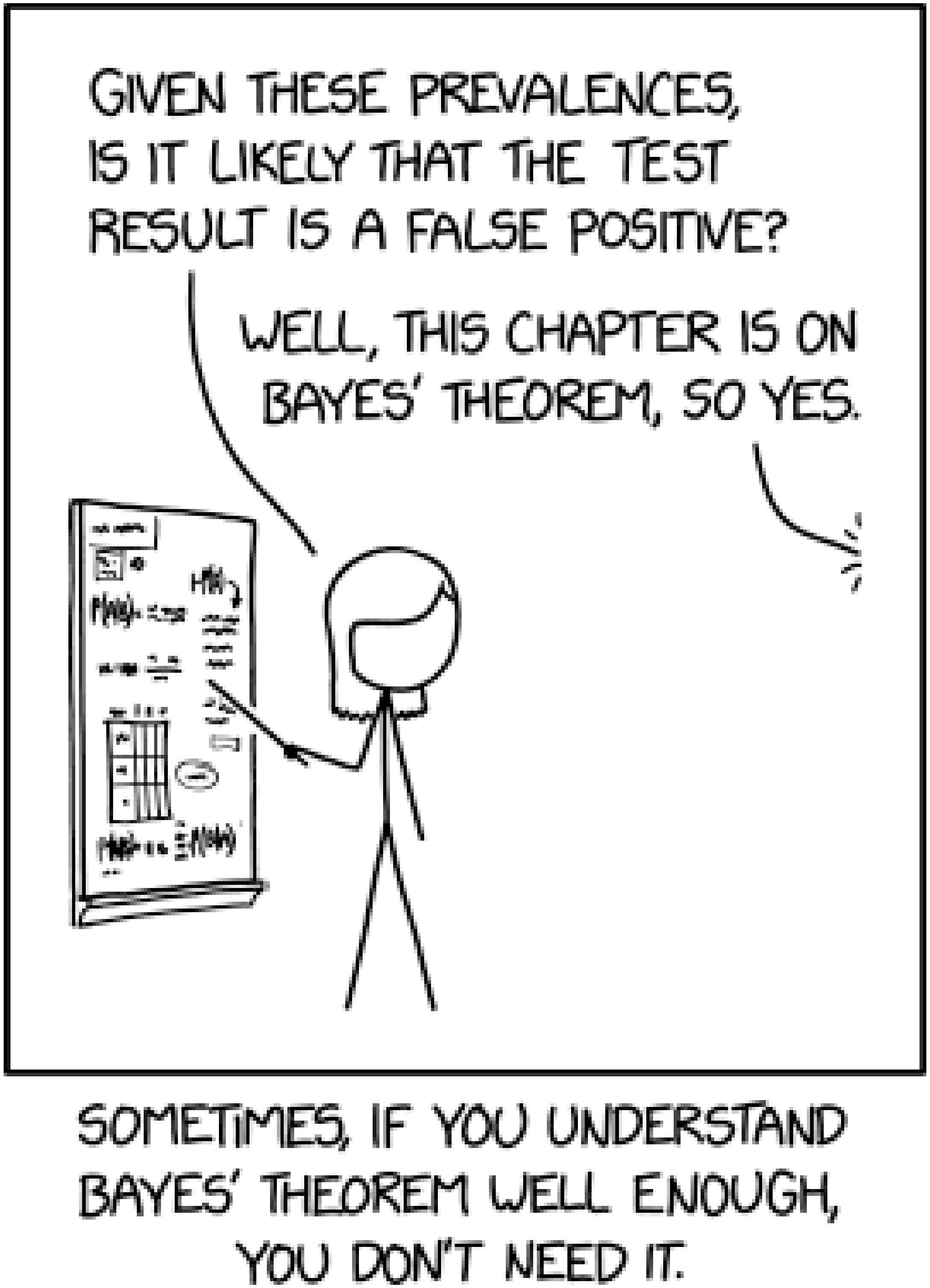
$0.95*0.5/[0.95*0.5+(1-0.99)*0.5]=0.99$

- 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



What is wrong in this comics?

If you are not yet reading XKCD comics https://xkcd.com/ you should start



If you are not yet reading XKCD comics https://xkcd.com/ you should start

Sensitivity/specificity of the standard test for prostate cancer: <u>PSA level > 4.0ng/mL</u>

- Sensitivity of the test is 71.9%
  - fraction of cancer patients who will test positive
  - False negative rate is 28.1%
- Specificity of the test is 90%
  - fraction of healthy patients who will test negative
  - False positive rate is 10%

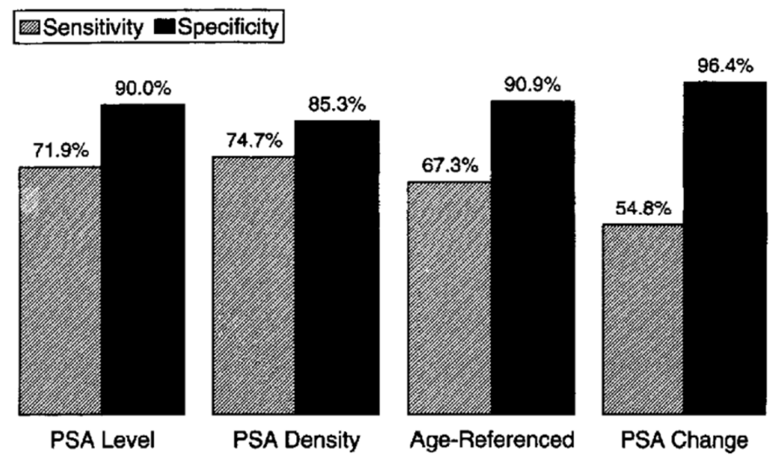
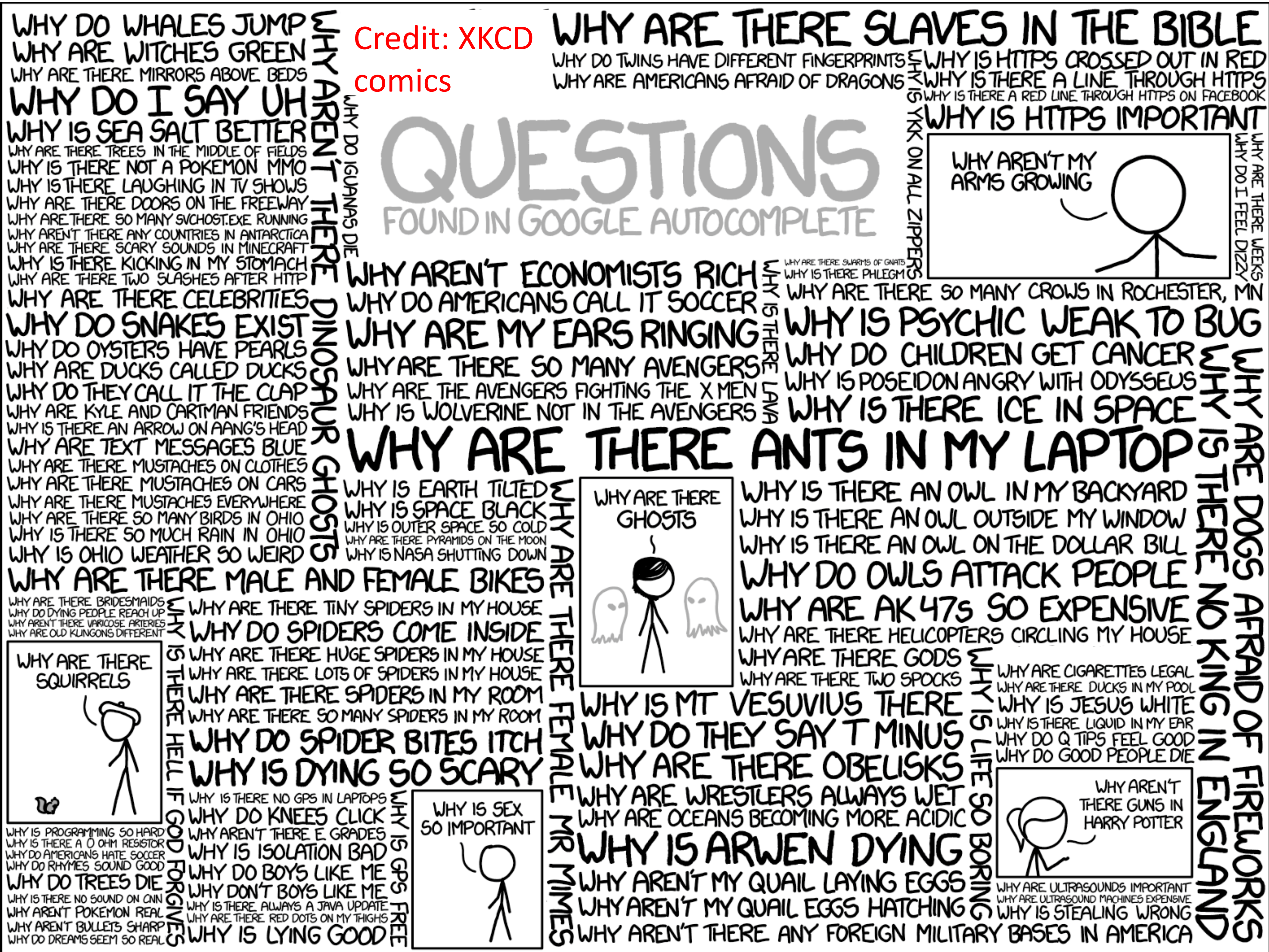


Figure 1. The relative sensitivity and specificity of different indexes of prostate specific antigen (PSA). Except for PSA change, sensitivity is the proportion of 171 known cancers cases for whom the index was positive and specificity is the proportion of 2011 men with normal transrectal ultrasound and digital rectal examinations not known to have prostate cancer who were negative on the index. The sensitivity and specificity of PSA change was evaluated in 84 men with prostate cancer and 1473 men without prostate cancer for on whom multiple PSA measures were available. A PSA level of 4.0 ng/ml or less was considered normal. A PSA density of 0.1 ng/ml per cubic centimeter of ultrasound-measured gland volume was considered normal. Age-referenced PSA was considered normal if it was 3.5 ng/ml or less in men aged 50–59, 4.5 ng/ml in men aged 60–69, and 6.5 ng/ml in men aged 70–79. PSA change was considered normal if the annual rate of PSA change was [illegible] or less per year.

Mettlin C, Littrup PJ, Kane RA, Murphy GP, Lee F, Chesley A, et al. Cancer. 1994;74: 1615–1620.

All this trouble for a lousy 38% gain in confidence? I don't believe you!

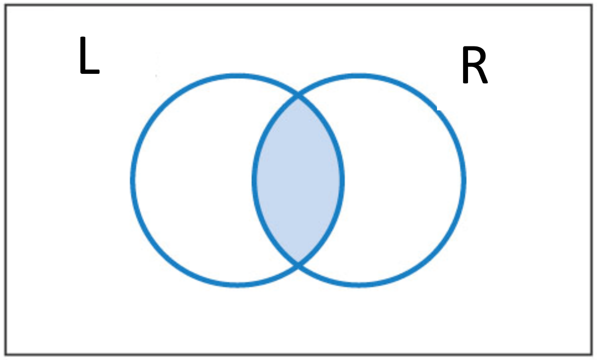
- Let C – be the event that the patient has cancer; C' – patient is cancer free, E – events that the PSA test was elevated
- We know **doctor's prior belief**: $P(C)=0.5$
- We know test stats: $P(E|C)=0.719$, $P(E|C')=0.1$
- We need to find $P(C|E)=P(E|C)*P(C)/P(E)$
- $P(E)=P(E|C)*P(C)+P(E|C')*P(C')=$
  $=0.719*0.5+0.1*0.5=0.41$
- $P(C|E)=0.5*0.719/0.41=0.88$ or 88%



Credit: XKCD comics

Series Circuit

This circuit operates only if there is at least one path of functional devices from left to right. The probability that each device functions is shown on the graph. Assume that the devices fail independently. What is the probability that the circuit operates?

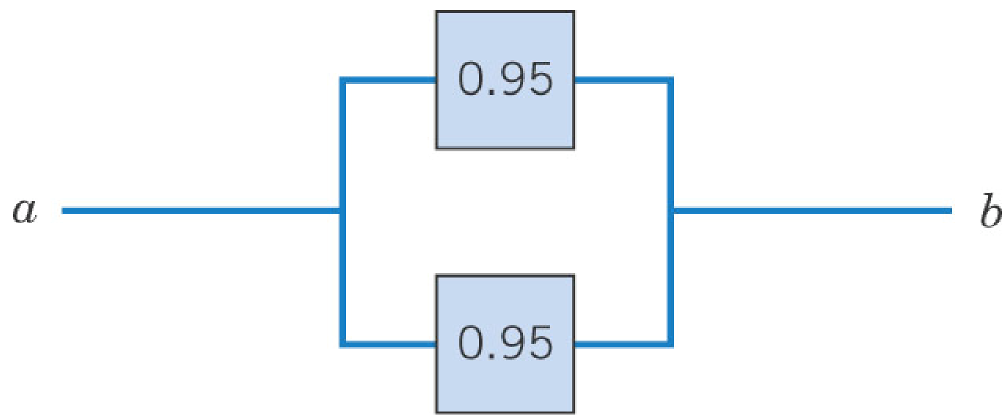


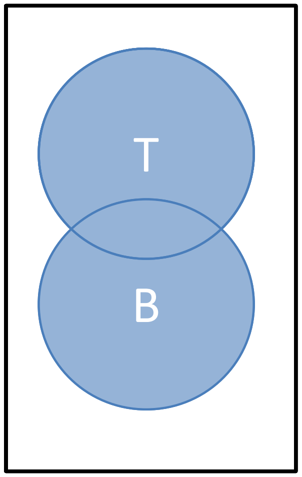
Let L & R denote the events that the left and right devices operate. The probability that the circuit operates is:

$$P(L \text{ and } R) = P(L \cap R) = P(L) * P(R) = 0.8 * 0.9 = 0.72.$$

Parallel Circuit

This circuit operates only if there is a path of functional devices from left to right. The probability that each device functions is shown. Each device fails independently.



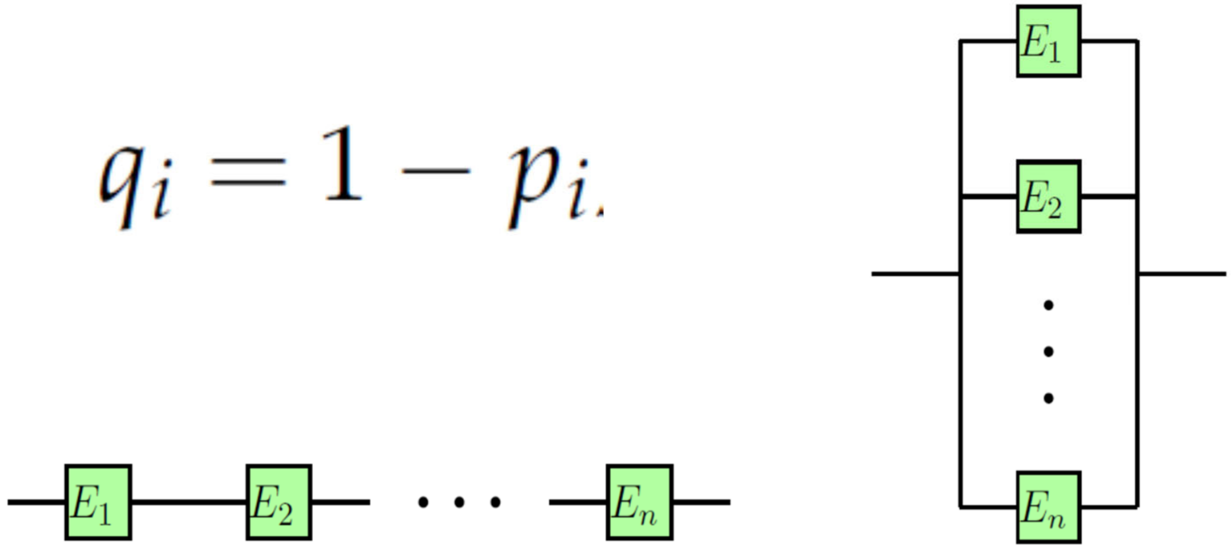


Let T & B denote the events that the top and bottom devices operate. The probability that the circuit operates is:

$P(T \cup B) = 1 - P(T' \cap B') = 1 - P(T')*P(B') = 1 – 0.05^2 = 1 – 0.0025 – 0.9975.$

Duality between parallel and series circuits

$$q_i = 1 - p_i.$$



(a) (b)

Connection	Notation	Works with prob	Fails with prob
Serial	$E_1 \cap E_2 \cap \cdots \cap E_n$	$p_1 p_2 \ldots p_n$	$1 - p_1 p_2 \ldots p_n$
Parallel	$E_1 \cup E_2 \cup \cdots \cup E_n$	$1 - q_1 q_2 \ldots q_n$	$q_1 q_2 \ldots q_n$