Bayes

One of the very best simple visual explanations of Bayes' theorem is found here (it uses LEGO):

https://www.countbayesie.com/blog/2015/2/18/bayes-theorem-with-lego

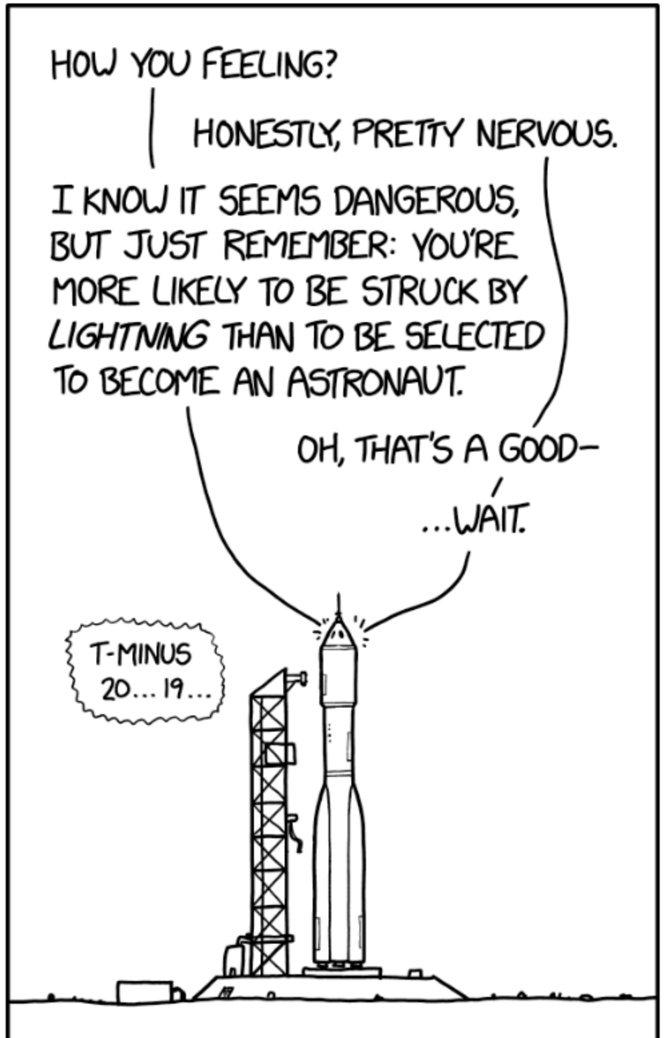
Additionally, Count Bayesie has put together a guide to Bayesian statistics which includes some useful resources if you are struggling with some of the concepts that we have covered so far in lecture:
https://www.countbayesie.com/blog/2016/5/1/a-guide-to-bayesian-statistics

There is an interesting video about using Bayesian inference to search for sunken treasure:
http://fivethirtyeight.com/features/how-data-nerds-found-a-131-year-old-sunken-treasure/

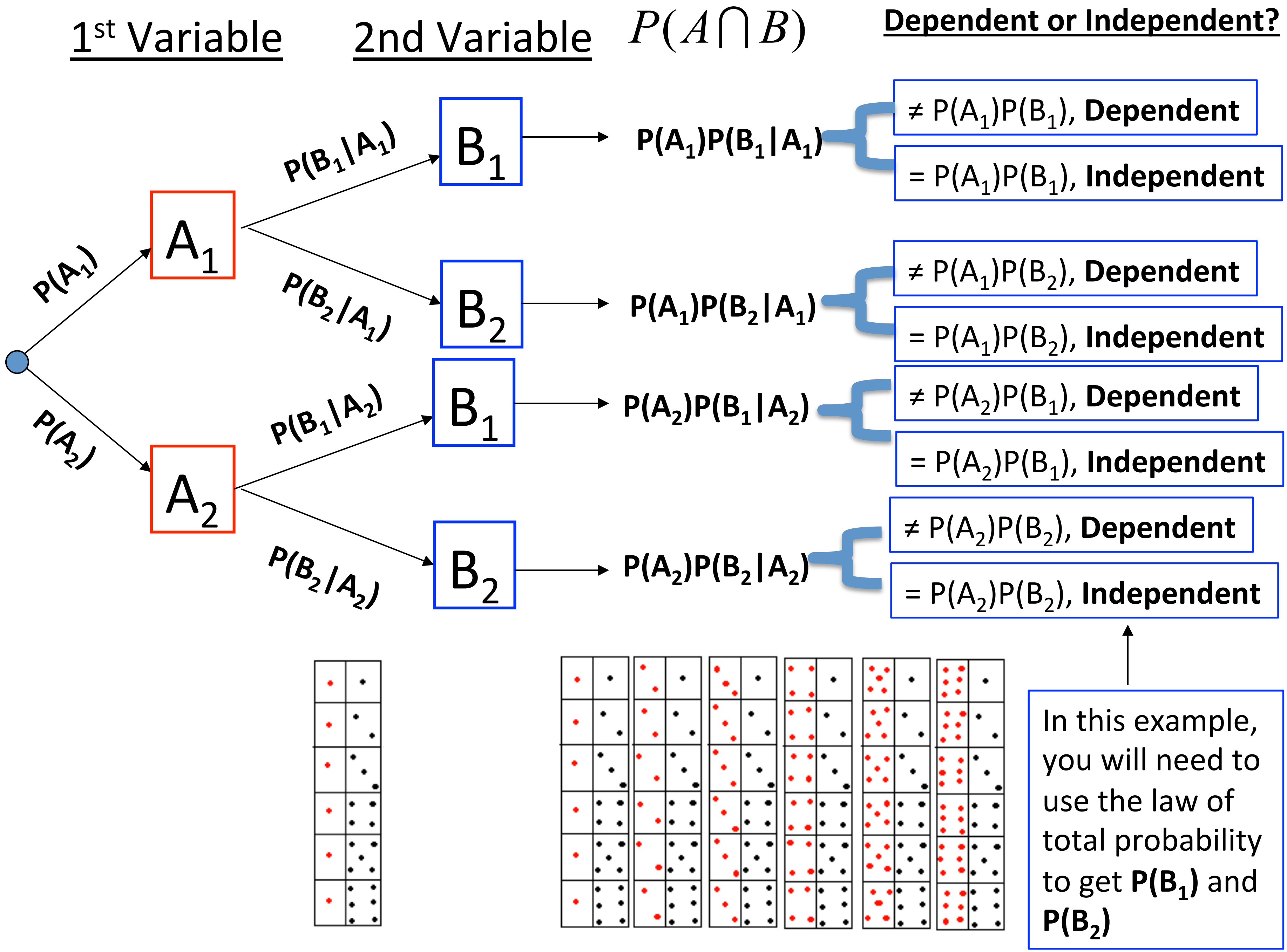
(Has also been used to find "black boxes" of airplanes that have crashed)

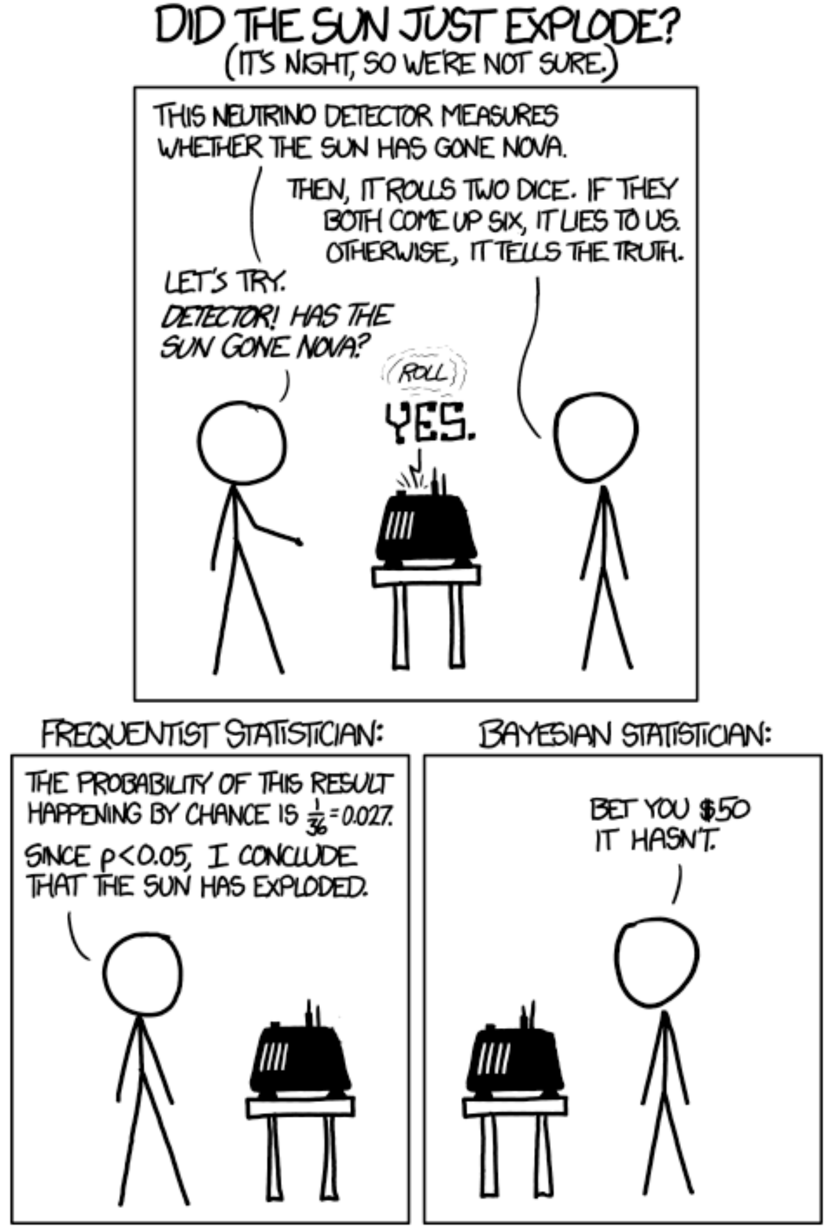
Launch Risk

|< | < Prev | Random | Next > | >|



http://xkcd.com/2107/


1st Variable
2nd Variable
$P(A \cap B)$
Dependent or Independent?
$P(A_1)$
$P(A_2)$
A_1
A_2
$P(B_1|A_1)$
$P(B_2/A_1)$
$P(B_1|A_2)$
$P(B_2/A_2)$
B_1
B_2
B_1
B_2
$P(A_1)P(B_1|A_1)$
$P(A_1)P(B_2|A_1)$
$P(A_2)P(B_1|A_2)$
$P(A_2)P(B_2|A_2)$
$\neq P(A_1)P(B_1)$, **Dependent**
$= P(A_1)P(B_1)$, **Independent**
$\neq P(A_1)P(B_2)$, **Dependent**
$= P(A_1)P(B_2)$, **Independent**
$\neq P(A_2)P(B_1)$, **Dependent**
$= P(A_2)P(B_1)$, **Independent**
$\neq P(A_2)P(B_2)$, **Dependent**
$= P(A_2)P(B_2)$, **Independent**
In this example, you will need to use the law of total probability to get **$P(B_1)$** and **$P(B_2)$**



http://xkcd.com/1132/

For personal enrichment/deepening your understanding, you may find the following websites interesting:

This website https://arbital.com/p/bayes_rule/?l=1zq has one of the best introductions to Bayes theorem − includes different levels of depth for you to choose from and explains why it is so important!

Also see this article from 2014 on the particular importance of understanding statistics/test results as a physician: http://www.bbc.com/news/magazine-28166019

Here's an alternately phrased version of the problem on which doctors fare somewhat better:

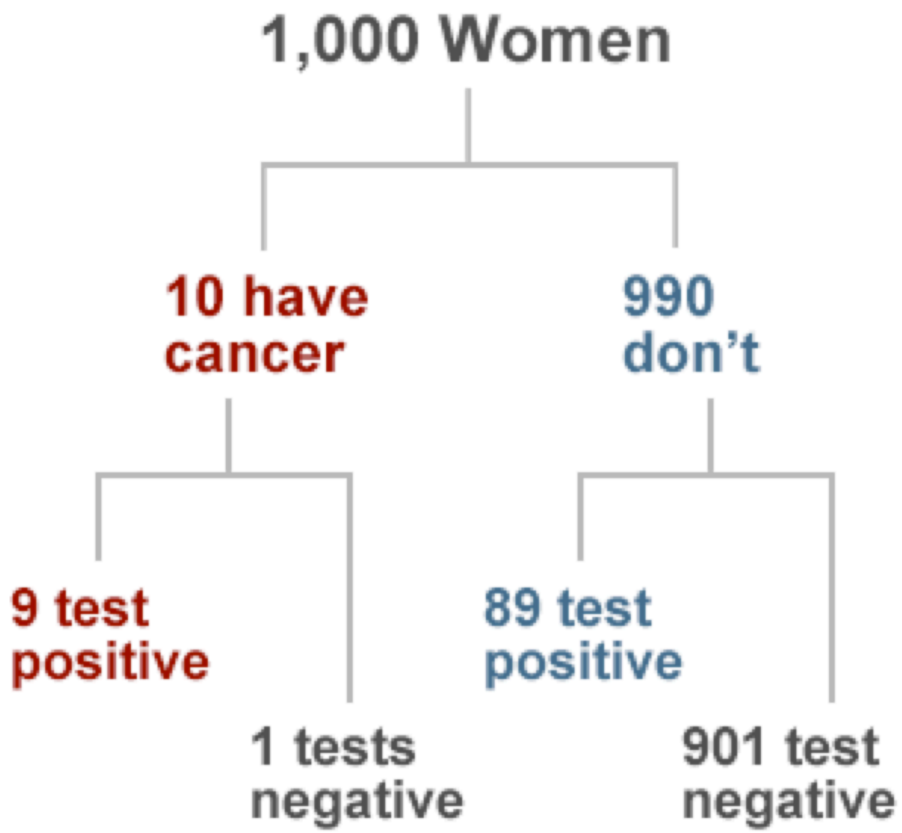
----

10 out of 1000 women at age forty who participate in routine screening **have breast cancer**. 900 out of 1000 women with breast cancer will get positive mammographies. Approx. 89 out of 1000 women without breast cancer will also get positive mammographies. If 1000 women in this age group undergo a routine screening, about what fraction of women with positive mammographies will actually have breast cancer?

And finally, here's the phrasing of the problem on which doctors fare best of all, with 46% - nearly half - arriving at the correct answer:

----

10 out of 1000 women at age forty who participate in routine screening have breast cancer. 90 of every 100 women with breast cancer will get a positive mammography. 89 out of 990 women without breast cancer will also get a positive mammography. If 1000 women in this age group undergo a routine screening, about what fraction of women with positive mammographies will actually have breast cancer?

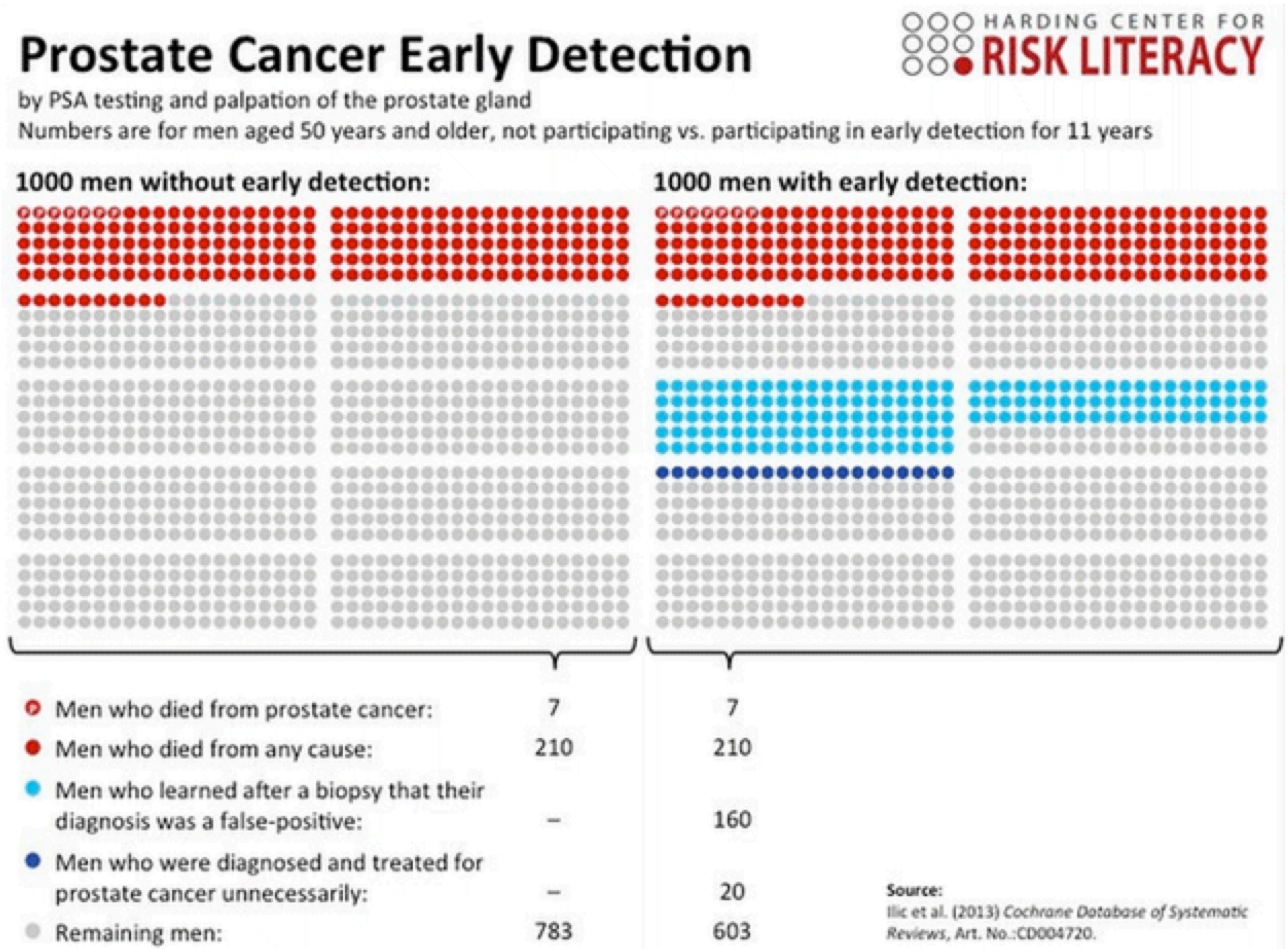


http://www.bbc.com/news/magazine-28166019

But…. doesn’t screening lead to saving lives?

After all, In the UK Men are not usually screened for PSA until **after** they start to demonstrate problems and men in the UK have a five year survival rate (prostate cancer) of 81%. In the USA, men are screened routinely for PSA and have a five year survival rate of 99%.

Inappropriate screening doesn't necessarily lead to more lives saved – it just results in a lot of **false positives** which are then included in the five year 'survival' count so it is inflated.



http://www.bbc.com/news/magazine-28166019

Inappropriate screening doesn't lead to more lives saved – just a lot of false positives
That are then included in the five year 'survival' count.

Breast Cancer Early Detection

HARDING CENTER FOR RISK LITERACY

by mammography screening

Numbers for women aged 50 years or older who participated in screening for 10 years

Benefits	1,000 women without screening	1,000 women with screening
How many women died from breast cancer?	5	4*
How many women died from all types of cancer?	21	21
Harms		
How frequent were false diagnoses, often associated with months of waiting for all-clear?	–	100
How many women were additionally diagnosed and operated** for breast cancer?	–	5

* This means that about 4 out of 1,000 women (50+ years of age) with screening died from breast cancer within 10 years – one less than without screening.

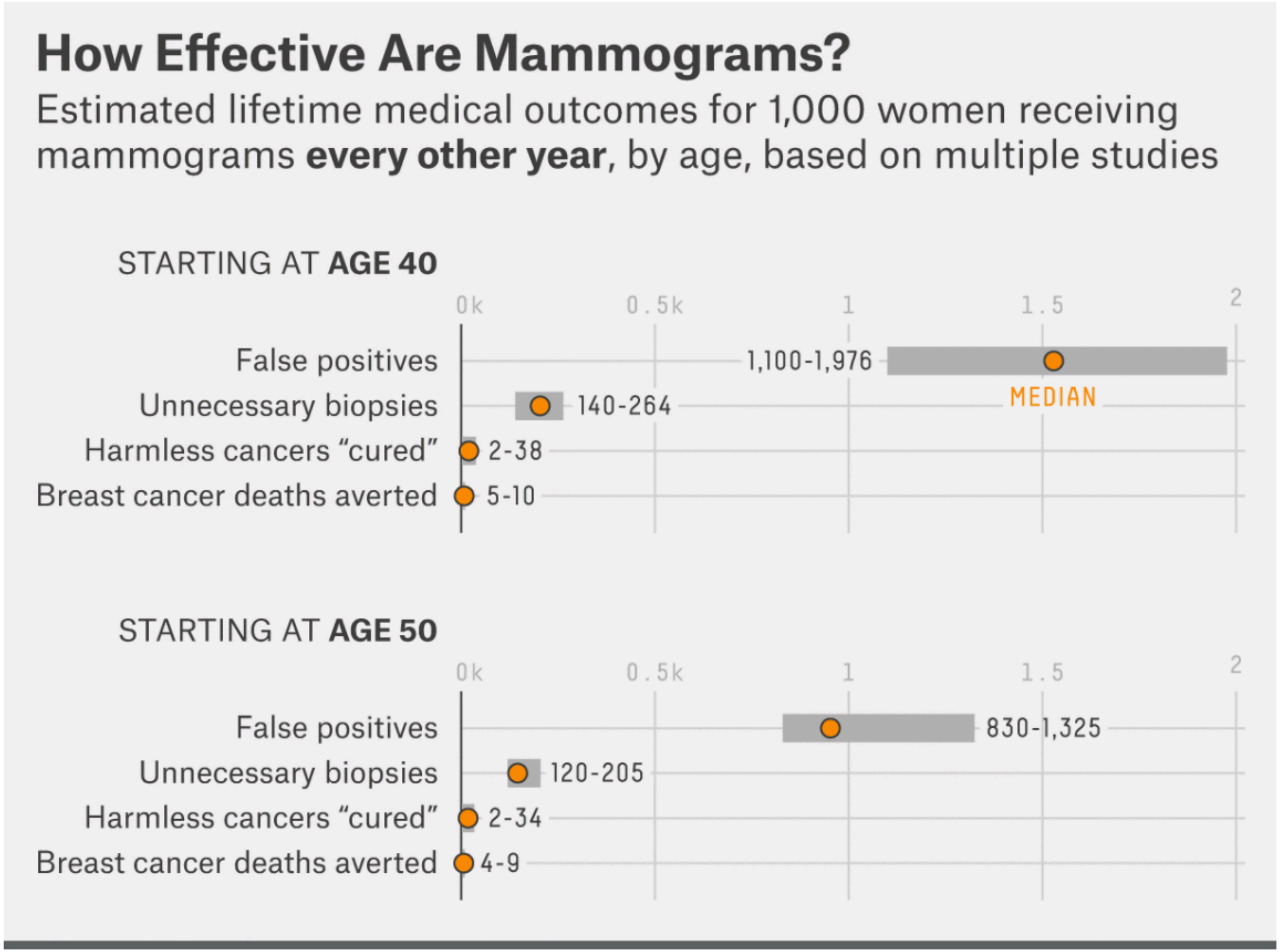
** Complete or partial breast removal

Source: Gøtzsche, PC, Nielsen, M (2011). *Cochrane database of systematic reviews* (1): CD001877.
Where no data for women above 50 years of age are available, numbers refer to women above 40 years of age.

http://www.bbc.com/news/magazine-28166019

Updated discussion:

http://fivethirtyeight.com/features/science-wont-settle-the-mammogram-debate/



Epistemology

Using a spade for some jobs and shovel for others does *not* require you to sign up to a lifetime of using only Spadian or Shovelist philosophy, or to believing that *only* spades or *only* shovels represent the One True Path to garden neatness.

There are different ways of tackling statistical problems, too.

http://faculty.washington.edu/kenrice/BayesIntroClassEpi515.pdf

- **Bayes inference is different from what we normally do**
- Probability is a way of quantifying uncertainty and is based on the assumption that with random sampling we can infer meaningfully from repeated experiments (**Frequentist)**
- Some phenomenon make sense under this definition….
  - If we toss a fair coin, what is the probability of 10 heads in a row?
  - If we assign treatments randomly to subjects, what is the probability that a sample mean difference between treatments will be greater than 20%?
  - What is the probability that, given the null hypothesis is true, of obtaining data that is at least as extreme as that observed?
- Some phenomenon don't….
  - What is the probability that polar bears will be extinct in 30 years?
  - What is the probability that hippos are sister group to whales?
- There is no random trials inherent in certain meaningful questions (Hippos either ARE or ARE NOT sister groups; polar bears will either be EXTINCT or NOT in 30 years)

<u>Bayes probability definition:</u>

"Probability is a measure of belief associated with the occurrence of an event "

Probability is subjective and can be **updated** when new information is available. Start with a **PRIOR** probability (or belief) and update it with a **LIKELIHOOD** and end with a **POSTERIOR** probability.

Ie. Will it rain tomorrow?

Frequentist Answer: probability of rain in Rochester over time.
Bayesian Answer: it will depend on whether it is Summer or Winter.

$$\boxed{P[A \mid B] = \frac{P[A]P[B \mid A]}{P[B]}}$$

You can interpret Bayes as reducing the state space, like the following equations which use the proportion of A intersecting with B over the WHOLE universe (ie. black die and red die both equal 1 is 1/36) and then reduce the proportion by dividing by the probability of the first event

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} \quad \text{OR} \quad P[B \mid A] = \frac{P[A \cap B]}{P[A]}$$

You might notice that both of these equations involve the SAME **numerator** whereas the **denominator** changes based on what event has happened first ie. what we already know and what we still want to know

The **PRIOR** hypothesis:
The original probability of the hypothesis without any additional information

The **LIKELIHOOD** interpreted as:
P(observation GIVEN the hypothesis)

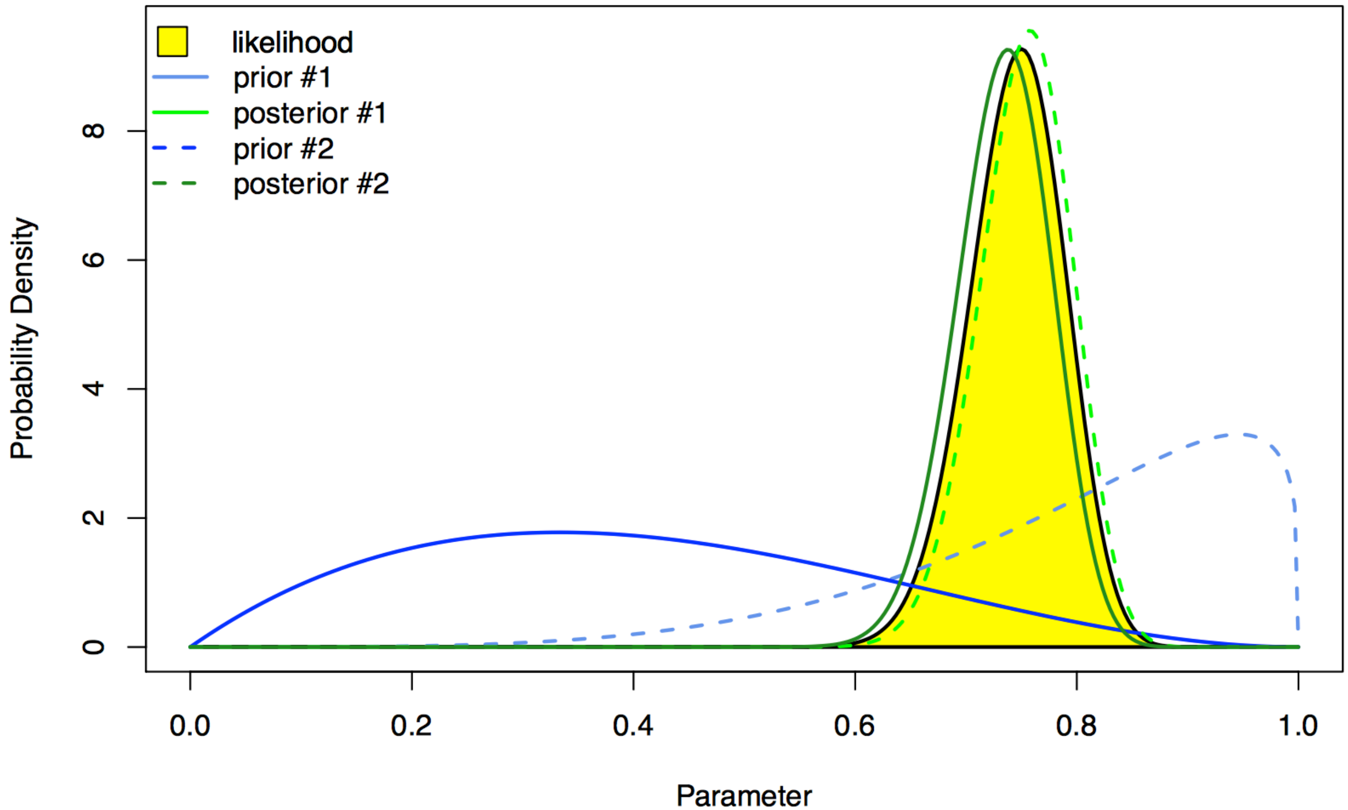
$$P[A \mid B] = \frac{P[A]P[B \mid A]}{P[B]}$$

the **POSTERIOR probability**
sometimes interpreted as
the P(hypothesis GIVEN the observation)

The **observation/data/ Evidence** that has been observed

A tangent:

When the data provide a lot more information than the prior, this happens; (recall the stained glass color-scheme)



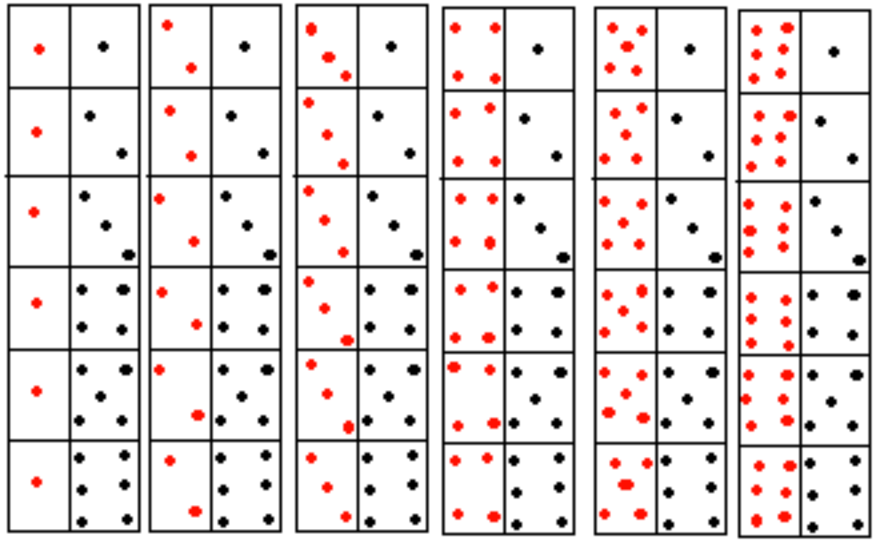
These priors (& many more) are *dominated* by the likelihood, and they give very similar posteriors – i.e. everyone agrees. (Phew!)

http://faculty.washington.edu/kenrice/BayesIntroClassEpi515.pdf

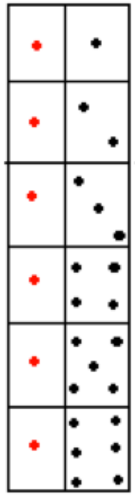
Venn diagram of Bayes'

<u>probability that two dice will sum to three?</u>

- *You have no new information*
- *You have to consider the entire universe*
- *P[Sum to three] =2/36*



- You have updated information
- Red die=1
- P[Sum to three|red=1]=1/6



Remember not to confuse P(A and B) with P(A|B)….

Prosecutor's Fallacy; a small probability of evidence given innocence need NOT mean a small probability of innocence given evidence.

https://www.geeksforgeeks.org/prosecutors-fallacy/

<u>Example:</u> Suppose we want to calculate the probability that someone will die of lung cancer given that they smoke. We study a cohort of individuals, determining which ones smoke and which ones do not and track them until they died. Then we could calculate the number of smokers who died of lung cancer.

There is an easier way, however….

USE BAYES

Example: Suppose we want to calculate the probability that someone will die of lung cancer given that they smoke. We study a cohort of individuals, determining which ones smoke and which ones do not and track them until they died. Then we could calculate the number of smokers who died of lung cancer.

There is an easier way, however….

USE BAYES!

$$\boxed{P[A \mid B] = \frac{P[A]P[B \mid A]}{P[B]}}$$

<u>Example:</u> Suppose we want to calculate the probability that someone will die of lung cancer given that they smoke.

<u>Answer:</u>

Specify the question: What is event 'A' and what is event 'B'?

$$\boxed{P[A \mid B] = \frac{P[A]P[B \mid A]}{P[B]}}$$

<u>Example:</u> Suppose we want to calculate the probability that someone will die of lung cancer given that they smoke.

<u>Answer:</u>

Specify the question: What is event ‘A’ and what is event ‘B’?

P[Death due to lung cancer|smoker]

This means that Bayes formula will be:

$$\boxed{P[Lung\_Cancer\_Death\,|\,Smoker] = \frac{P[Lung\_Cancer\_Death]P[Smoker\,|\,Lung\_Cancer\_Death]}{P[Smoker]}}$$

<u>Example:</u> Suppose we want to calculate the probability that someone will die of lung cancer given that they smoke.

<u>Answer:</u>

Specify the question: What is event 'A' and what is event 'B'?

P[Death due to lung cancer|smoker]

This means that Bayes formula will be:

$$\boxed{P[Lung\_Cancer\_Death|Smoker] = \frac{P[Lung\_Cancer\_Death]P[Smoker|Lung\_Cancer\_Death]}{P[Smoker]}}$$

The probabilities on the right are already present in public databases. If we use Bayes, there is no need for costly study.

Example: Suppose we want to calculate the probability that someone will die of lung cancer given that they smoke.

Answer:

Specify the question: What is event 'A' and what is event 'B'?

P[Death due to lung cancer|smoker]

Estimated from death records

$$P[Lung\_Cancer\_Death \mid Smoker] = \frac{P[Lung\_Cancer\_Death]\,P[Smoker \mid Lung\_Cancer\_Death]}{P[Smoker]}$$

Polling appropriate control population

Example: Suppose we want to calculate the probability that someone will die of lung cancer given that they smoke.

Answer:

Specify the question: What is event 'A' and what is event 'B'?

P[Death due to lung cancer|smoker]

P[smoker] =0.5
P[smoker|Death due to lung cancer] = 0.9
P[Death due to lung cancer] = 0.3

$$P[\text{Death due to lung cancer}|\text{smoker}] = \frac{0.9 \times 0.3}{0.5} = 0.54$$

This study also gives the
P[Death due to lung cancer|Non-smoker] = 0.06

Bayes

- Using data collected in 1975, the probability that a woman had cervical cancer was 0.0001
- The probability that a biopsy would correctly identify these women as having cancer was 0.90
- The probability of a "false positive" (the test claims that cancer is present when, in fact, there is no cancer) was 0.001

What is the probability that a woman with a positive result actually has cervical cancer?

- Using data collected in 1975, the probability that a woman had cervical cancer was 0.0001

  P[cancer] = 0.0001

- The probability that a biopsy would correctly identify these women as having cancer was 0.90

  P[positive test| cancer] = 0.90

- The probability of a "false positive" (the test claims that cancer is present when, in fact, there is no cancer) was 0.001

  P[positive test| no cancer] = 0.001

What is the probability that a woman with a positive result actually has cervical cancer?

P[cancer| positive test] = ?

Bayes

$$P[\text{cancer}|\text{positive result}] = \frac{P[\text{positive result}|\ \text{cancer}] \times P[\text{cancer}]}{P[\text{positive result}]}$$

P[cancer] = 0.0001

P[no cancer] = 1 - 0.0001 = 0.9999

P[positive result| cancer] = 0.9

P[positive result| no cancer] = 0.001

P[positive result] = ??

P[positive result] = P[positive result| cancer]xP[cancer] +
P[positive result| no cancer] x P[no cancer]
= (0.9)(0.0001)+(0.001)(0.9999)
= 0.0010899

Bayes

$$P[\text{cancer}|\text{positive result}] = \frac{P[\text{positive result}|\ \text{cancer}]\text{x}P[\text{cancer}]}{P[\text{positive result}]}$$

P[cancer] = 0.0001

P[no cancer] = 1 - 0.0001 = 0.9999

P[positive result| cancer] = 0.9

P[positive result| no cancer] = 0.001

P[positive result] = 0.0010899

$$P[\text{cancer}|\text{positive result}] = \frac{0.9\text{x}0.0001}{0.0010899} = 0.0826$$

Bayesian Probability:

- Tension between frequentists and bayesians

  → Where does the prior come from?

  * all data that is external to the current study
  * mitigate this issue: use non-informative prior

- Most events are not repeatable so bayesian probability attempts to utilize a personal assessment of an outcomes likelihood
  - Do this by weighting probability with a prior probability
    - Some scientists may use different priors!
    - Increasing data/information means that the prior should have less impact on the posterior
  - This may seem subtle but it is quite different from what science traditionally does:

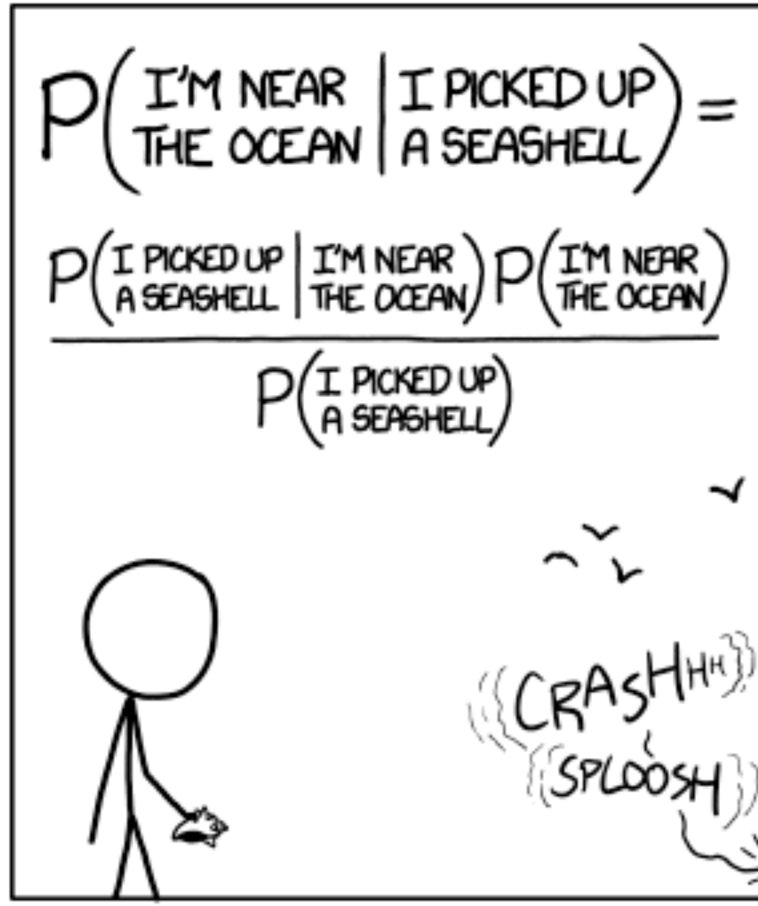
    The parameter is treated as though it were a random variable instead of a constant truth/value

Bayesian Probability:

Bayes Factor:

- *Allows strength of two competing hypotheses (a null and an alternate) to be quantified*
- *Both hypotheses must use same prior*
- *Null hypothesis isn't assumed (it doesn't have primacy over the alternative hypothesis)*
- *Similar to likelihood that we will see towards the end of this course*

http://xkcd.com/1236/



STATISTICALLY SPEAKING, IF YOU PICK UP A SEASHELL AND *DON'T* HOLD IT TO YOUR EAR, YOU CAN PROBABLY HEAR THE OCEAN.