Bayes Theorem

P(A)

P(B)

P(A & B)

**Examples:**

You might be interested in finding out a patient’s probability of having liver disease if they are an alcoholic. “Being an alcoholic” is the **test** (kind of like a litmus test) for liver disease.

* **D** could mean the event “Patient has liver disease.” Past data tells you that 10% of patients entering your clinic have liver disease. P(D) = 0.10.
* **A** could mean the litmus test that “Patient is an alcoholic.” Five percent of the clinic’s patients are alcoholics. P(A) = 0.05.
* You might also know that among those patients diagnosed with liver disease, 7% are alcoholics. This is your **A|D:** the probability that a patient is alcoholic, given that they have liver disease, is 7%.

Find P(D|A)

P(D) = .1, P(A) = .05, P(A|D) = .07

* P(D|A) = .07 \* .1 / .05 = .7/5 = .14

In a particular pain clinic, 10% of patients are prescribed narcotic pain killers. Overall, five percent of the clinic’s patients are addicted to narcotics (including pain killers and illegal substances). Out of all the people prescribed pain pills, 8% are addicts. If a patient is an addict, what is the probability that they will be prescribed pain pills?

A = addicted

Rx = prescribed pills

P(Rx) = .1

P(A) = .05

P(A|Rx) = .08

* P(Rx|A) = .08 \* .1 / .05 = .8/5 = .16

Example #3: the Medical Test

A slightly more complicated example involves a medical test (in this case, a genetic test):

There are **several forms of Bayes’ Theorem**out there, and they are all equivalent (they are just written in slightly different ways). In this next equation, “X” is used in place of “B.” In addition, you’ll see some changes in the denominator. The proof of why we can rearrange the equation like this is beyond the scope of this article (otherwise it would be 5,000 words instead of 2,000!). However, if you come across a question

P(GD) = 1% of people have a certain [genetic defect](https://www.genome.gov/10001204).

P(!GD) = 99% of pop.  
P(+|GD) = 90% of tests for the gene detect the defect (true positives).  
P(+|!GD) = 9.6% of the tests are [false positives](https://www.statisticshowto.datasciencecentral.com/false-positive-definition-and-examples/).  
If a person gets a positive test result, **what are the odds they actually have the genetic defect?**

Find: P(GD|+)

* P(GD|+) = P(+|GD) \* P(GD) / P(+), where P(+) = P(GD) \* P(+|GD) + P(!GD)\*P(+|!GD)

= .9 \* .01 / (.9\*.01 + .99\*.096) = 0.0865

🡪8.65% changes

The General Form:

**Given the following statistics, what is the probability that a woman has cancer if she has a positive mammogram result?**

* One percent of women over 50 have breast cancer.
* Ninety percent of women who have breast cancer test positive on mammograms.
* Eight percent of women will have false positives.

P(BC) = .01

P(!BC) = .99

P(+|BC) = .90

P(+|!BC) = .08

🡪 P(BC|+) = .90 \* .01 / (.90 \* .01 + .08 \* .99) = .102 🡪 10.2% of women

**Naïve Bayes Vs. Linear Discriminate analysis:**

Both are of the form: P(Ci|x) = P(x|Ci)

Both estimate P(Ci) the Same.

Assuming that x is Normally distributed, they differ in how they estimate parameters. Mu\_x is the same, the mean(x|Ci).

* NB assumes data independence: sigma is conditional
  + Models var(x|Ci­) 🡪 variance estimated for variable subsets associated with each class
* LDA uses variance for all of x, so there is dependence there.
  + Models var(x) 🡪 Total variance in the data.

So, these two models are quite similar.