Lecture 3.1 Bayes Theorem

2013/09/16

Previously... Partition

We say a collection of events $B_1, \ldots, B_k$ form a partition of the sample space if:

1. they are mutually exclusive
2. they are exhaustive i.e.

$$S = \bigcup_{i=1}^{k} B_i$$

Previously... Law of Total Probability

If $B_1, \ldots, B_k$ is a partition of the sample space S, then for any event A

$$\mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$$

Just think of the multiplication rule:

$$\mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{k} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$$

Goals for Today

1. Introduce Bayes Theorem
2. Brief Tangent: Bayesian Statistics
3. Example from XKCD

Bayes Theorem (Simplified Form)

Bayes Theorem (Expanded Form)

Bayes Theorem Example

Important Example: Screening for a Rare Disease

Suppose we have a test for a disease which has the following characteristics, where P = 'tests positive', and D = 'has the disease':

- Sensitivity (true positive rate): $\mathbb{P}(P|D) = 0.99$
- Specificity (true negative rate): $\mathbb{P}(P|D^c) = 0.01$
- Prevalence of the disease: $\mathbb{P}(D) = 0.01$

Question: What is the probability that someone who tests positive actually has the disease?

Important Example: Screening for a Rare Disease

Important Example: Screening for a Common Disease

Say now instead the prevalence is 10% (and not 1%) i.e. the disease is more common

- Sensitivity: $\mathbb{P}(P|D) = 0.99$
- Specificity: $\mathbb{P}(P|D^c) = 0.01$
- Prevalence of the disease: $\mathbb{P}(D) = 0.10$ (not 0.01)

Important Example: Screening for a Common Disease

Important Example: Screening for a Disease

For two tests with the exact same sensitivity and specificity:

$\mathbb{P}(D)$	$\mathbb{P}(D\|P)$	$\mathbb{P}(D^c\|P)$
0.01	0.5	0.5
0.1	0.917	0.083

The error rate $\mathbb{P}(D^c|P)$ dropped from 0.5 to 0.083! The change is due to the prevalence of the disease $\mathbb{P}(D)$.

Moral of the Story: In many cases, how well your diagnostic test performs depends on the prevalence of the disease.

Preview of MATH 392: Bayesian Statistics

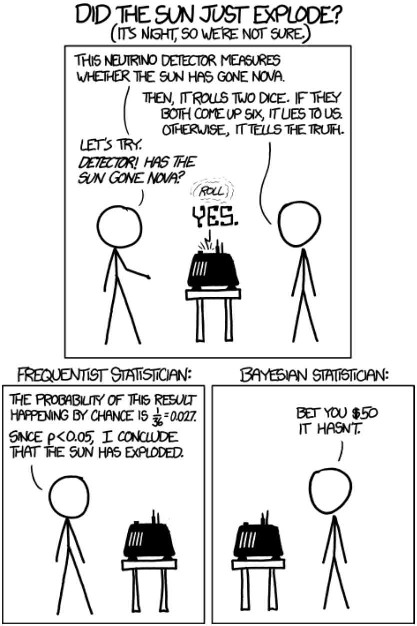
Bayesian statistics differs from frequentist statistics in how we think of the unknown parameter θ:

Preview of MATH 392: Bayesian Statistics

To express our belief about θ from a Bayesian perspective, we have:

Frequentists do have a notion of 1. and 3.

Bayes Theorem XKCD



Bayes Theorem XKCD

Bayes Theorem XKCD

Bayes Theorem XKCD

$\mathbb{P}(N|X) \approx 0$ means our posterior belief that the sun exploded is essentially zero, hence the Bayesian's confidence in making the 50 bet.

i.e. despite what the detector said, there wasn't strong enough evidence the data to overcome our prior belief.

Bayes Theorem XKCD

The comic is arguing the ridiculousness of not incorporating this prior.

However, this goes to the core criticism about Bayesian statistics: that priors are subjective. i.e. where did we come up with the very skeptical prior probability of the sun having exploded of 0.000001?

The Frequentist's argument is based on p-values, which we will see in MATH 392.

Next Time:

We define random variables in the discrete case.