Probability and stats mop-up unit

- Statistical inference and related concepts (last week Monday)
- Hypothesis testing (last week Wednesday)
- Benford's law (Monday)
- Bayesian inference (**today**)
- Begin graph theory (Friday)

Today:

- Examples, part 1
- Bayesian philosophy and methods
- Examples, part 2

Suppose a rare disease affects .1% of the population. A new diagnostic test has been developed that is 98% accurate, with the same rate of false positives as false negatives. If the test is positive, what is the probability that a person has (or doesn't have) the disease?

Let D be the event that a person has the disease and T be the event that the test result is positive. The information given is

$P(D)$	$P(\overline{D})$	$P(T \mid D)$	$P(\overline{T} \mid D)$	$P(T \mid \overline{D})$	$P(\overline{T} \mid \overline{D})$
.001	.999	.98	.02	.02	.98

What we want are $P(D \mid T)$ and $P(\overline{D} \mid T)$.

Adapted from Dworsky, *Probably Not*, 2019 p 200–201

Suppose your friend has a coin that she won't let you examine. You are told that i's either a standard coin and evenly balanced, or it is a two-headed coin or a two-tailed coin. She will flip the coin as many times as you wish and share the results of the flip with you. This will be your only source of information about the coin.
You are asked to predict, in terms of probabilities, which of the above three coin types this coin is.

Adapted from Dworsky, *Probably Not*, 2019 p 202

Suppose one takes a coin and tosses it 10 times and gets 8 heads. Then from a frequentist point of view, the result is that this coin comes down heads 8 times out of 10. This what is called the maximum likelihood estimate.... However, if one has looked the coin over, and there doesn't seem anything wrong with it, one would be very reluctant to accept this estimate. Rather, one would tend to think that the coin would come down equally head and tails over the long run, and getting 8 heads out of 10 is just the kind of thing that happens sometimes given a small sample. In other words, one has a *prior belief* that influences one's beliefs even in the face of apparent evidence against it. Bayesian statistics measure degrees of belief, and are calculated by staring with prior beliefs and updating them in the face of evidence, by use of Bayes's theorem.

Manning and Schütze, *Foundations of Statistical Natural Language Processing*, 1999 , pg ...

Bayesian vs Frequentist Probability

The frequentist point of view is based on the following postulates:

- F1 *Probability refers to limiting relative frequencies. Probabilities are objective properties of the real world.*
- F2 *Parameters are fixed, unknown constants.*
- F3 *Statistical procedures should be designed to have well-defined long run frequency properties.*

The Bayesian approach is based on the following postulates:

- B1 *Probability describes degree of belief, not limiting frequency. "The probability that Albert Einstein drank a cup of tea on August 1, 1948 is .35" does not refer to any limiting frequency but reflects my strength of belief that the proposition is true.*
- B2 *We can make probabilty statements about parameters, even though they are fixed constants.*
- B3 *We make inferences about a parameter by producing a probability distribution for it.*

Wasserman, All of Statistics, 2004 pg 175-176, abridged.

$$\overbrace{P(A|B)}^{\text{posterior}} = \frac{\overbrace{P(B|A)}^{\text{likelihood}}\ \overbrace{P(A)}^{\text{prior}}}{\underbrace{P(B)}_{\text{marginal}}}$$

Suppose that initially your friend had told you, “I think it’s a normal (HT) coin” which you interpreted to mean there’s a 50% probability that it is indeed an HT coin, a 25% probability that it’s an HH coin, and a 25% probability that it’s a TT coin. These numbers are now your original priors What effect does this change of prior have on our beliefs after seeing a sequence of flips?

Adapted from Dworsky, *Probably Not*, 2019 pg 204

Suppose we have a coin that has a head and tail but is weighted. We consider five ways in which the coin's behavior could be affected by this weighting: 100% probability of landing head, 75%, 50%, 25%, and 0%.
Initially we believe that each of these is equally probable. How is that belief updated after a sequence of flips?

Adapted from Dworsky, *Probably Not*, 2019 pg 205

For next time:

Find Benford (and non-Benford) data in the wild.