Must it be Santa? A case study in Bayesian updating

Sam Urmy

December 24, 2014

1 Introduction

For children in many parts of the world, December 24th is a day of unbearable anticipation. The reason? According to popular Christian (and secular) tradition, on the night before Christmas, a man named Saint Nicholas, or Santa Claus, flies around the world in a sleigh drawn by eight or nine magical reindeer. In the course of his trip, he stops at the house of every Christmas-celebrating boy and girl, leaving gifts and consuming any food left for him. He is usually understood to gain entry by sliding down the chimney, but presumably uses alternative methods (e.g. central air-conditioning ducts, windows) in domiciles without fireplaces.

This belief, while popular, is also problematic, since entering a home without the permission of those living there is in other circumstances known as "breaking and entering" or "home invasion," and is a felonious crime in most jurisdictions. On the other hand, if concerned parents call the police when the strange man *is* Santa, Christmas will be ruined not only for their children, but for everyone else lower down on Santa's list while he is booked into the holding cell at the local precinct.[1]

The challenge is to quickly identify Santa based on his known characteristics. Given that the visit in question is most likely happening late at night in the dark, these characteristics may only be revealed to us one at a time. In such a situation, Bayesian reasoning about probabilities provides a logical method for updating our belief that the home invader is actually Santa, as opposed to a criminal trespasser.

2 Bayes' Rule

Bayes' Rule, named for its originator, the Rev. Thomas Bayes (c. 1701-1761), is a statement about the joint probability $P(A, B)$ that two events A and B occur

---

[1]The possibility of parents "standing their ground" and exercising their Second Amendment rights is beyond the scope of this paper.

together. This expression can be decomposed into either of two condidional probability statements,

$$P(A, B) = P(A|B)\,P(B) = P(B|A)\,P(A). \tag{1}$$

In other words, the probability that both A and B are true is equal to the probability that A is true, multiplied by the probability that B is true *conditional on A* being true (or vice versa).

In Bayesian statistics, we are usually interested in the probability of some hypothesis H, given that we have observed some data D. In this context, these quantities are plugged into Equation 1, which is then rearranged as follows:

$$P(H|D) = \frac{P(D|H)\,P(H)}{P(D)} \tag{2}$$

In this equation, $P(D|H)$ is known as the *likelihood*, and is a function telling us how likely we are to observe the data D if the hypothesis H is true. The other two probabilities on the right-hand side, $P(D)$ and $P(H)$, are the probabilities of the data and hypothesis unconditional on anything else. $P(D)$ is sometimes known, but often is not; in that case computational methods are available to solve the equation approximately. $P(H)$, known as the *prior* probability of the hypothesis, is one of the keys to Bayesian analysis. It represents our (subjective) degree of belief in the hypothesis *before* we analyze the data. The left-hand side of Equation 2, known as the *posterior* probability of the hypothesis, is thus a combination of our prior and the likelihood.

In other words, our initial belief is modified by the data to yield a conclusion that involves both. One of the attractive features of Bayesian analysis is that the posterior from one experiment or test can be plugged in as the prior for the next one, allowing our knowledge to be continuously updated as more data arrive. This approach is the one we will use to estimate the likelihood that an intruder is Santa Claus.

3 Must be Santa (?)

The popular Christmas song "Must be Santa," written by Bill Fredericks and Hal Moore, can be read as a folk implementation of a Bayesian updating procedure. Each verse, sung in call-and-response fashion, asks about a known characteristic of Santa Claus, with the previously identified characteristics repeated each verse, before the chorus confirms that the visitor "Must be Santa, Santa Claus." The lyrics are reproduced in full below.

Who's got a beard that's long and white
(Santa's got a beard that's long and white)

Who comes around on a special night
Santa comes around on a special night

Special Night, beard that's white

Chorus: Must be Santa
Must be Santa
Must be Santa, Santa Claus

Who wears boots and a suit of red
(Santa wears boots and a suit of red)

Who wears a long cap on his head
(Santa wears a long cap on his head)

Cap on head, suit that's red
Special night, beard that's white

(Chorus)

Who's got a big red cherry nose
(Santa's got a big red cherry nose)

Who laughs this way HO HO HO
(Santa laughs this way HO HO HO)

HO HO HO, cherry nose
Cap on head, suit that's red
Special night, beard that's white

(Chorus)

Who very soon will come our way
(Santa very soon will come our way)

Eight little reindeer pull his sleigh
(Santa's little reindeer pull his sleigh)

Reindeer sleigh, come our way
HO HO HO, cherry nose
Cap on head, suit that's red
Special night, beard that's white

(Chorus)

Dasher, Dancer, Prancer, Vixen,
Comet, Cupid, Donner and Blitzen

Reindeer sleigh, come our way
HO HO HO, cherry nose
Cap on head, suit that's red
Special night, beard that's white

(Chorus)

As the evidence accumulates, our confidence that the visitor is Santa presumably increases. A more rigorous numerical analysis makes it clear just how much our belief should change.

4 Calculations

In each verse of "Must be Santa," a new piece of data is revealed to us about the nocturnal visitor. At each stage, then, we must update our belief that the visitor is Santa, using Bayes' Rule:

$$P(Santa|Data) = \frac{P(Data|Santa)\,P(Santa)}{P(Data)} \tag{3}$$

The challenge (involving a bit of subjectivity) is to find reasonable numbers for each of these quantitities.

Likelihood

This is a fairly special case, in that all of the characteristics identified here are known to apply to Santa: he has a red suit, comes on Christmas Eve, has reindeer, etc. Consequently, the likelihood at each stage is identically equal to 1.

Priors

The prior probability that the visitor is Santa is a subjective aspect

Probability of data

```
> library(xtable)
> library(ggplot2)
> library(reshape2)
> evidence <- c("None", "Beard", "Date", "Suit", "Hat", "Nose", "Laugh", "Soon",
+               "Reindeer")
> p.evidence <- c(NA, 0.01, 1/365, 0.001, 0.1, 0.3, 0.2, 0.8,
+                 1e-9)
> p.santa <- rep(0, length(p.evidence))
> likelihood <- c(NA, rep(1, length(p.evidence) - 1))
> tab <- data.frame(evidence, p.evidence, likelihood,
+                   p.santa, p.santa, p.santa, p.santa)
> priors <- c(0.9, 0.1, 0.001, 1e-9)
> lty <- 1:length(priors)
> plot.new()
> plot.window(xlim=c(1, 9), ylim=c(0, 1))
> axis(1, at=1:length(p.evidence), labels=evidence)
> mtext("Evidence", 1, 3)
> axis(2)
> mtext("P(Santa)", 2, 3)
> for (i in 1:length(priors)) {
+   tab[1, i + 3] <- priors[i]
+   for (j in 2:length(p.evidence)) {
```

```
+       tab[j, i+3] <- likelihood[j] * tab[j - 1, i+3] /
+         ((1 - tab[j-1, i+3]) * p.evidence[j] + likelihood[j] * tab[j - 1, i+3])
+     }
+   lines(tab[ , i+3], lty=lty[i])
+   points(tab[ , i+3], pch=lty[i])
+ }
>
```
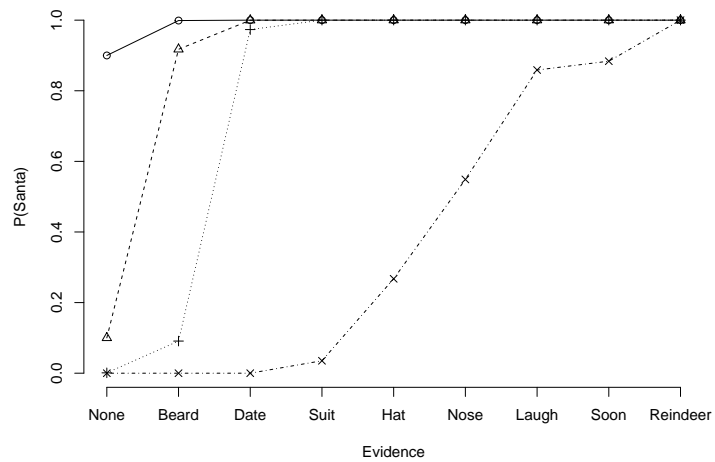

	D	P(D)	P(D|S)	P(S)=0.9	P(S)=0.5	P(S)=0.1	P(S)=1E-9
1	None			0.9	0.1	0.001	1E-09
2	Beard	0.01	1	0.999	0.917	0.091	1E-07
3	Date	0.00274	1	1	1	0.973	3.65E-05
4	Suit	0.001	1	1	1	1	0.0352
5	Hat	0.1	1	1	1	1	0.267
6	Nose	0.3	1	1	1	1	0.549
7	Laugh	0.2	1	1	1	1	0.859
8	Soon	0.8	1	1	1	1	0.884
9	Reindeer	1E-09	1	1	1	1	1