STSCI 4780: Inference in discrete spaces (cont'd)

Tom Loredo, CCAPS & DSS, Cornell University

2018-02-06

Recap: Bayesian inference in one slide

Probability as generalized logic

Probability quantifies the *strength of arguments*

To appraise hypotheses, calculate probabilities for arguments from data and modeling assumptions to each hypothesis

Use *all* of probability theory for this

Bayes's theorem

$$p(\text{Hypothesis} \mid \text{Data}) \quad \propto \quad p(\text{Hypothesis}) \times p(\text{Data} \mid \text{Hypothesis})$$

Data *change* the support for a hypothesis $\propto$ ability of hypothesis to *predict* the data

Law of total probability

$$p(\text{Hypothes}\underline{\textbf{es}} \mid \text{Data}) \quad = \quad \sum p(\text{Hypothes}\underline{\textbf{is}} \mid \text{Data})$$

The support for a *compound/composite* hypothesis must account for all the ways it could be true

Contextual/prior/background information

Bayes's theorem moves the data and hypothesis propositions wrt the solidus:

$$P(H_i|D_{\rm obs},\mathcal{C}) = P(H_i|\mathcal{C})\,\frac{P(D_{\rm obs}|H_i,\mathcal{C})}{P(D_{\rm obs}|\mathcal{C})}$$

It lets us *change the premises*

"Context" or "prior information" or "background information" = information that is **always** a premise (for the current calculation)

Notation: $P(\cdot|\cdot,\mathcal{C})$ or $P(\cdot|\cdot,I)$ or $P(\cdot|\cdot,M)$ or ...

The context can be a notational nuisance! "Skilling conditional":

$$P(H_i|D_{\rm obs}) = P(H_i)\,\frac{P(D_{\rm obs}|H_i)}{P(D_{\rm obs})} \qquad ||\,\mathcal{C}$$

Often just drop $\mathcal{C}$—but be careful!

Essential contextual information

We can only be uncertain about a proposition, A, if there are alternatives (at least $\overline{A}$!); what they are will bear on our uncertainty. *We must explicitly specify relevant alternatives.*

Hypothesis space: The set of alternative hypotheses of interest (and auxiliary hypotheses needed to predict the data, e.g., for LTP)

Data/sample space: The set of possible data we may have predicted before learning of the observed data

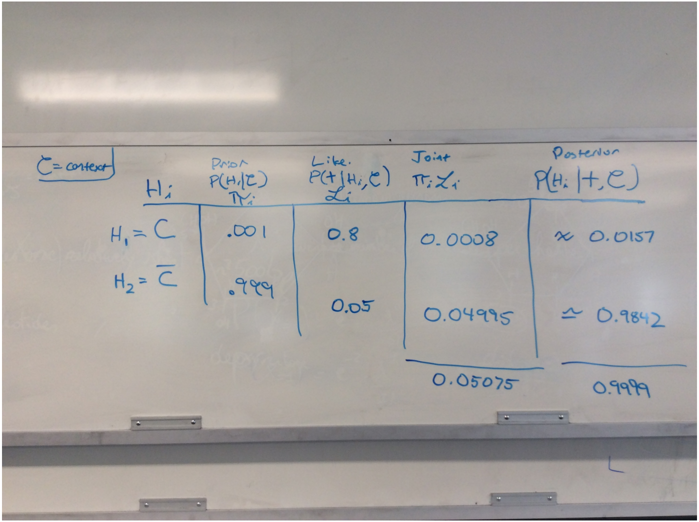
Predictive model: Information specifying the likelihood function (e.g., the conditional predictive dist'n/sampling dist'n)

Other prior information: Any further information available or necessary to assume to make the problem *well posed*

Bayesian literature often uses **model** to refer to *all* of the contextual information used to study a particular dataset and predictive model

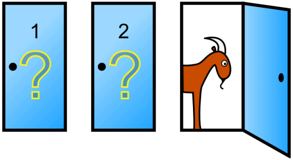
Recap: Binary hypotheses and data

Post-test probabilities after positive test result



Monty Hall problem: 3 hypotheses

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then asks you, "Do you want to switch to door No. 2?" Is it to your advantage to switch?



Wikipedia

Important rules:

- The prize door is set randomly (with equal probability)
- The host will only reveal a goat
- When both unchosen doors hide a goat, the host picks one at random (with equal probability)

Setup:

Hypothesis space: H_i = prize behind door i

Data/sample space: D_i = host opens door i; $D_{\text{obs}} = D_3$

Compute $P(H_i|D_3, \mathcal{C})$

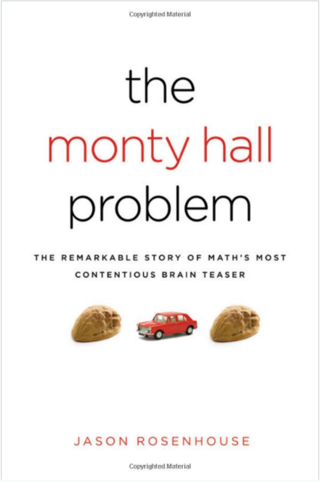
Controversy:

Parade columnist Marilyn vos Savant (IQ $> 200!$) posed the problem in her column and argued that one should switch.

She received $\approx 10,000$ letters, the great majority disagreeing with her. The strongest opposition came from mathematicians and scientists.

A common line of intuitive reasoning: The doors have equal probability for hiding the prize. The host has ruled out one of them. The remaining doors now each have probabilty $1/2$ of having the prize, so there is no advantage to switching.

See statistician Luke Tierney's summary: "Behind Monty Hall's Doors: Puzzle, Debate and Answer?" (NYT 1991)



The Monty Hall Problem: The Remarkable Story of Math's Most Contentious Brain Teaser, by Jason Rosenhouse (2009)

From Tierney's essay:

> *Persi Diaconis, a former professional magician who is now a Harvard* University professor specializing in probability and statistics, said there was no disgrace in getting this one wrong.*
>
> *"I can't remember what my first reaction to it was," he said, "because I've known about it for so many years. I'm one of the many people who have written papers about it. But I do know that my first reaction has been wrong time after time on similar problems. Our brains are just not wired to do probability problems very well, so I'm not surprised there were mistakes."*

*Diaconis subsequently moved to Cornell, and is now at Stanford.