Hierarchical Bayesian Models

Kevin Dorst
kmdorst@mit.edu

24.223 Rationality

I. "Learning to learn"

Laplace's rule of succession: if seen a successes and b failures, probability that the next will be a success is $\frac{a+1}{a+b+2}$

Beta distribution

Humans are better than this!

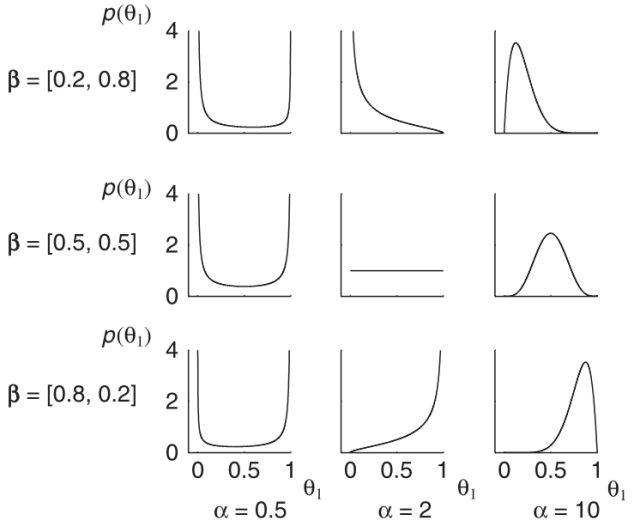
One-shot learning

Levels of uncertainty:

- Level-0: color of marble (blue/green)
- Level-1: proportion of green in bag
- Level-2: distribution of proportions of green in bags — rooms of bags
- Level-3: distribution of distribution of proportions of green in bags — building of rooms of bags

Levels 2, 3,... are **over hypotheses**: abstract knowledge that 'sets up a hypothesis space'—i.e. sets your conditional credences for what to expect at lower levels.

Rooms have different distributions of bags of various proportions:



Buildings have different distributions of rooms.

Relations of *screening-off* between levels.

When you observe marbles, and track whether they're coming from the same bag or not, you get info about (1) the proportion in this bag, (2) the distribution of proportions in the room this bag came from, and (3) the distribution of distributions of proportions in the building that room is in.

Eg: (i) 10 all-green and 10 all-blue bags, vs. (ii) 20 bags of 50-50 each.

HBM treats differently; "conventional" BM treats the same.

HBM exhibits much more human-like learning:

- Do one-shot learning for some categories (university) and not others (birthday month).
- Learn 'shape bias' with new categories[1];
- Learn that some categories unified by shape and others by texture.

[1] Shape > texture, color

Etc.

II. What's happening, mathematically?

1) Hierarchical Bayesian Models are (just) Bayesian models.

- If uncertain about value of variable, it should vary in your state space!
- HBMs give convenient ways to specify rich BMs, and to set overall uncertainty with minimal parameters.

"Conventional" BM is simply dogmatic about the independence of the bags' proportions

How?

2) Nested series of partitions (questions).

- Levels chosen to induce clear posteriors, given answer at that level.
- So you *defer* to each lower level.
- If uncertain, Total Probability tells you how to set credences.

Which building? Which room? Which bag?

Reflection / Principal Principle

We've seen this before.

Gambler's-fallacy paper had {Sticky, Steady, Switchy} be the Level-1 hypotheses.
Could make (more) hierarchical by having distribution over different processes' chance of each.

III. What's happening, conceptually?

Regress?

Do they think people are actually (implicitly) doing this?

If they are, why are they so *bad* at probabilistic reasoning?