Discussion section
week of 20 - 24 Apr 20

Case Study: Monty Hall
(On course webpage)

Structure of all applied math problem-solving
(meta-code)

Real world

Start here
State problem clearly and accurately

(natural language)

Math world

Step 1
Translate

Step 2
Express

Step 3
Solve

Step 4
Obtain math solution in real-world terms
(natural language)
Example of how inaccurate translation from natural language to math language leads to a dead end in probability calculations.

1. **Crucial First Step**

   - Take-home test 1
   - Problem 3

2. Identify how many different sources of information there are in the (natural-language) problem statement, and create symbols to stand for all relevant (true/false) propositions.

---

Here there are 2 different kinds of information:

- 1. who will be pardoned (referred to as \( P \));
- 2. who the warden says will not be pardoned (\( Q \)).

If we let \( P = (A \text{ will be pardoned}) \) and similarly for \( Q \) and \( R \), you might think that the relevant probability, after \( A \) has heard
what the warden said, is but this: $P(A | \text{not } B)$ is not right:

A has not learned that $B$ will not be pardoned; all A has learned is that (warden says $B$ will not be pardoned)

we need new symbols for information of type [2] above. Let $(W = w)$ stand (tells A)

for (warden says $w$ will not be pardoned)

Then what we want is $P(A | W = B)$, and this is not at all the same thing as $P(A | \text{not } B)$. Monty Hall case study

In full generality there are three
different kinds of information here.

1. Which door you initially choose;
2. Which door Monty opens to show you a goat; and
3. Where the car really is.

So let's let \( \mathcal{E}_i = \{ \text{door } i \} \) 
\((i = 1, 2, 3)\)

Monty reveals \( M_j = \{ \text{goat behind door } j \} \) 
\((j = 1, 2, 3)\)

Car really is behind door \( C_k = \{ \text{car really is behind door } k \} \) 
\((k = 1, 2, 3)\)

Without loss of generality, let's look (condition on) \( \mathcal{E}_1 \) and \( M_2 \); we want to compute \( P( C_3 | \mathcal{E}_1, M_2) \) and compare it with \( P( C_1 | \mathcal{E}_1, M_2) \). Naive intuitive argument: since Monty has now shown...
Given a goat behind door 2, it's become a 50/50 proposition between C₁ and C₃, so there's no advantage in switching:

\[ P(C₃ \mid I_i, m₂) = \frac{1}{2} \]  

Intuitive but wrong:

\[ P \text{ intuitive}(C₁ \mid I_i, m₂) \]  

The above argument is guesswork, not math, let's try math.  

Step I: Give I_i and m₂, under the rules of the game the only possibilities left for C_k are C₁ and C₃ (why?), so C₁ and C₃ have become opposite.

\[ P(C₁ \mid I_i, m₂) + P(C₃ \mid I_i, m₂) = 1 \]  

Step II: \[ I_i \text{ adds no} \]
information to the game; why? \( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \) (but much needs to learn \( \frac{1}{3} \))
notice that \( C_k \) acts in
this problem like the \[ \text{unknown} \]
thing (the truth) you're wondering about and
\( m_j \) acts like \[ \text{data} \], so \[ P(C_2 \mid z_1, m_1) \]
is of the form \[ P(\text{unknown} \mid \text{randomization}, \text{data}) \]

Step III Notice further that the rules of the game are giving us information of the form \[ P(\text{data} \mid \text{randomization}, \text{rules}) \]

Translating into math, we know these probabilities:
\[
P(M_2 \mid z_1, C_2) = 0 \]
\[
P(M_2 \mid z_1, C_3) = 0 \]
\[
P(M_2 \mid z_1, C_1) = \frac{1}{2} \]
Step IV

Going from what we know, $p(m_j \mid E_i, C_k)$, to what we want, $p(C_k \mid E_i, m_j)$, is a job for Bayes's Theorem, and since (from Step I) $C_1$ and $C_2$ are opposites given $(E_i, m_2)$, let's use Bayes's Theorem in odds form:

\[
\frac{p(A \mid B)}{p(\neg A \mid B)} = \frac{p(A)}{p(\neg A)} \cdot \frac{p(B \mid A)}{p(B \mid \neg A)}
\]
It's straightforward (if a bit tedious) to show that this leads to situations where there's a third information source (E, say) in the form of a true/false statement that we know to be true: just stick E everywhere to the right of the conditioning bar:

\[
\frac{P(A | E, B)}{P(A | E, \overline{B})} = \frac{P(A | E)}{P(A | E)} \left( \frac{P(B | E, A)}{P(B | E, \overline{A})} \right)
\]

In this case study this becomes

\[
\frac{P(C_3 | E_1, m_0)}{P(C_1 | E_1, m_0)} = \left( \frac{P(C_3 | E_1)}{P(C_1 | E_1)} \right) \left( \frac{P(m_0 | E_1, C_3)}{P(m_0 | E_1, C_1)} \right)
\]
Next, we know that (where the (C_i) (which dav you (initially pick) are independent, so the priors odds ratio is

\[
\frac{P(C_3 | \Xi_1)}{P(C_1 | \Xi_1)} = \frac{P(C_3)}{P(C_1)} = \frac{1}{3} \cdot \frac{1}{3} = 1
\]

But from step III we already have that the Bayes factor in favor of (C_3 over C_1)

\[
\frac{P(m_2 | \Xi, C_3)}{P(m_2 | \Xi, C_1)} = \frac{1}{2} > 2
\]
\[ p_\alpha = \frac{OA}{OA + OB} = \frac{OA}{1 + OA} \]

and since \( p_\alpha \) is the ratio of to probabilities \( p_\alpha \) is

The general way to get from one

You'll recall that.

\[ \frac{p_{C_{31,32}}}{p_{C_{12,3}}} = \left( \frac{\frac{1}{3}}{\frac{1}{2}} \right) \]

So were done. The posterior odds are

in favor of \( C_2 \) over \( C_3 \) is

and

\[ \sum_{i=0}^{3} = \frac{2}{3} + 1 = \frac{5}{3} \]

and

\[ p_{C_{3}} = \frac{2 + 1}{3} = \frac{3}{3} = 1 \]

So can double your chances if sticking
What was wrong with the previous intuitive argument? At first glance it seems that Monty behind door 2 (after you initially chose door 1) doesn't give you new information about where the car is, but as the following correct intuitive argument shows: On those occasions (which happen \(\frac{1}{3}\) of the time) on which you (at random, and not known to you) pick the door where the car...
really is, Monty is forced to randomize between (showing you a foot behind door 2) and (showing you a foot behind door 3), i.e., on those occasions you really do learn nothing about the true location of the car;  

**But** on those occasions on which you (at random, and not known to you) don't pick the door where the car really is (and this happens \( \frac{2}{3} \) of the time), by showing you a foot behind (so you) door 2 he is telling you that the car is behind door 3 (1)
So 1/3 of the time it doesn't matter if you switch or not, and 2/3 of the time you get the car by switching! So you should switch. Here's a math argument that formalizes this second intuitive story above. A direct application of Bayes' Theorem in the extended form

\[ P(A|E, B) = \frac{P(A|E)P(B|E, A)}{P(B|E)} \]

So this problem is

\[ P(C_3|E, m_2) = \frac{P(C_3|E)P(m_2|E, C_3)}{P(m_2|E)} \]
Now, as before, \( Pr(C_3 | E_1) = Pr(C_3) = \frac{1}{3} \)
and \( Pr(M_2 | E_1, C_3) = 1 \), so
\[
Pr(C_3 | E_1, M_2) = \frac{\frac{1}{3} \cdot 1}{Pr(M_2 | E_1)}
\]

and (as usual) we're stuck with evaluating the annoying denominator / normalizing constant.

Crucial: we can't predict what exactly observation will do after we've chosen data.

The unknown location of the car \( C_k \) is

Following Dennis Lindley, let's extend the conversation by bringing \( C_k \) into the calculation by partitioning over it.
\[
\frac{e}{2} = \frac{(\frac{2}{3})}{(\frac{2}{3})} = (\frac{2}{3}) \frac{(\frac{2}{3})}{(\frac{2}{3})} = (\frac{3}{3}) \frac{(\frac{2}{3})}{(\frac{2}{3})} = (\frac{3}{3}) \frac{(\frac{2}{3})}{(\frac{2}{3})}
\]

\[
\frac{\pi(\frac{3}{3})}{\pi(\frac{3}{3})} + \frac{\pi(\frac{3}{3})}{\pi(\frac{3}{3})} + \frac{\pi(\frac{3}{3})}{\pi(\frac{3}{3})}
\]

\[
(\frac{3}{3}) \frac{(\frac{2}{3})}{(\frac{2}{3})} = (\frac{3}{3}) \frac{(\frac{2}{3})}{(\frac{2}{3})}
\]

\[
(3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})} + (3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})} + (3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})}
\]

\[
(3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})} + (3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})} + (3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})}
\]

\[
(3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})} + (3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})} + (3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})}
\]

\[
(3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})} + (3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})} + (3\pi) \frac{(\frac{2}{3})}{(\frac{2}{3})}
\]