This conditional time: probability, discrete distributions.

Next: continuous time: distributions, expectation.

\[ y_i = \{ \text{you initially choose door } i \} \]

\[ m_j = \{ \text{Monte Hall then opens door } j \} \]

\[ c_k = \{ \text{car actually behind door } k \} \]

\[ i, j, k = 1, 2, 3 \]

You pick door 1 (Y1) & Monte Hall opens door 2 to reveal a goat.

Plan ahead!
we want to compare \( P(C_1 | M_2, Y_1) \) to

with \( P(C_3 | M_2, Y_1) \).

This is like ELISA:

| ELISA | unknown; location of car | true HIV status
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>data; route showing you a good behind</td>
<td>what ELISA said</td>
<td></td>
</tr>
</tbody>
</table>

we want \( P(\text{unknown} | \text{data}) \) but

problem setup gave us \( P(\text{data} | \text{unknown}) \)

so let's use Bayes's theorem to reverse order of conditioning:

\[
P(C_2 | M_2, Y_1) = \frac{P(C_1)}{P(C_3)} = \left[ \begin{array}{c} P(M_2, Y_1 | C_1) \\ P(M_2, Y_1 | C_3) \end{array} \right].
\]

Bayes factor
Now by the rules \( P(c_1) = P(c_3) = \frac{3}{10} \), so the prior odds are \( \frac{P(c_1)}{P(c_3)} = \frac{1}{1} = 1 \) to evaluate probabilities like \( P(m_2, y_1 | c_1) \), let's use the general form of product rule for \( \text{and} \):

\[
\frac{P(m_2, y_1 | c_1)}{P(m_2, y_1 | c_3)} = \frac{P(y_1 | c_1) \cdot P(m_2 | y_1, c_1)}{P(y_1 | c_3) \cdot P(m_2 | y_1, c_3)}
\]

but \( y_1 \) and \( c_j \) are independent so

\[
P(y_1 | c_1) = P(y_1) = \frac{1}{3}
\]

and

\[
P(y_1 | c_3) = P(y_1) = \frac{1}{3}
\]

so

\[
P(c_1 | m_2, y_1) = \frac{P(m_2 | y_1, c_1)}{P(m_2 | y_1, c_3)} = 1
\]

\[
P(c_3 | m_2, y_1) = \frac{P(m_2 | y_1, c_3)}{P(m_2 | y_1, c_3)} = 1
\]
So: after \( m_2 \) (given \( y_1, c_j \)), the posterior odds in favor of car behind door 2:1, so \( P(c_3 | m_2, y_1) = \frac{2}{3} \) you should switch.

Case study: Cromwell's rule

For any \( D \) such that \( P(D) > 0 \) \\
and \\
(a) if \( P(A) = 0 \) then \( P(A | D) = 0 \) \\
(b) if \( P(A) = 1 \) then \( P(A | D) = 1 \)

\[ D = \text{data} \quad P(A) = \text{prior information} \quad \sqrt{P(A | D) = \text{posterior info about A}} \]

\[ A = \text{unknown} \]
Anything you put prior probability on has to have posterior probability 1 no matter how the dataset comes out; this destroys the possibility of learning from data.

\[ P(A \mid D) = \frac{P(A \text{ and } D)}{P(D)} \]

But if \( P(A) = 0 \) \( \phi \)

then \( P(A \text{ and } D) = 0 \) \( \checkmark \)

\[ P(A \mid D) = \frac{P(A \text{ and } D)}{P(D)} \]

so \( (A \text{ and } D) = D \)

and \( P(D / D) = 1 \)

\[ (9.50) \]
Case Study: The Rasmussen Report

WASH-1400, "The Reactor Safety Study"

Problem] Estimate \( P(\text{catastrophic accident at nuclear power plant}) \)

at a moment in history when no such events had ever occurred (nuclear power began in about 1955, ...)

3 Mile Island 1979

Solution] Use expert judgment to break down \( \Theta \) into a collection of simpler events connected together with \( \land, \lor, \ldots \); for example

\( \Theta = (\text{hi, y1 & above # faults start off reactor} \land \ldots) \)
Estimate of \( P(\Theta) \) was extremely small: \( 10^{-12} \), yet only 4 years later: 3 miles inland.

What went wrong?

Right calculation:

\[
P(\Theta) = P(\text{tiny} \mid \text{brakes}) \cdot P(\text{brakes}) \]

\[
P(\text{tiny} \mid \text{brakes}) \cdot P(\text{brakes}) \cdot P(\text{brakes} \mid \text{small})
\]

What they did instead: they assumed independence.

\[
P(\Theta) = P(\text{tiny} \mid \text{brakes}) \cdot P(\text{brakes}) \cdot P(\text{brakes} \mid \text{small})
\]

\[
\frac{\text{small}}{\text{small}} \quad \frac{\text{small}}{\text{small}} \quad \frac{\text{small}}{\text{small}}
\]

= tiny just because many monkeys close to a multiplied