The Monte Hall Problem

Richard Lockhart

Simon Fraser University

STAT 870 — Summer 2013

Purposes of Today's Lecture

- Illustrate modelling process.
- Discuss choice of sample or probability space.
- Illustrate ambiguities in real world problems.
- Compare conditional and unconditional modelling assumptions.

The Monte Hall Problem

- Monte Hall hosted a TV game show called "Let's Make a Deal".
- At the end of the show top two winners each got to pick one of three doors, numbered 1, 2 3.
- One door had big prize. One had small prize. One had "goat".
- This problem is not quite the same – modelled on that one, though.
- Imagine only one player, not two.
- Three doors, one prize, 2 goats.

Monte Hall set-up

- Player picks a door.
- Monte Hall opens some other door to show you it had a goat – this is always possible.
- Monte offers you the chance to switch the door you picked for the one he did not open.
- Should you switch?

Monte Hall sample space

- What is the set of possible outcomes — the sample space?
- Ingredients: Monte hides prize (H), player chooses door (C), Monte opens door (O), player chooses whether or not to switch (S).
- So typical outcome is sequence (H, C, O, S). The notation is
  - H is 1, 2 or 3 – the door where Monte hides the prize;
  - C is the door the player initially picks – again 1, 2 or 3.
  - O is door Monte opens; can't be the door where the prize is hidden;
  - S is either 1 for switching or 0 for not switching.
- Total of $3 \times 3 \times 2 \times 2 = 36$ possible outcomes.
- or, for simplicity allowing $O = H$, $3 \times 3 \times 3 \times 2 = 54$ possible outcomes.

Some events, random variables and probabilities

- Have 4 obvious random variables: H, C, O, S.
- Another random variable of interest X which is the number of the door the player ends up with.
- What do we know about probabilities of events or distributions of random variables?
- Assume that player has no knowledge of how the prize is hidden.
- Convert this to H and C are independent.
- Convert this to assumption:

$$Pr(H = C) = \frac{1}{3}.$$

- That is: no matter how player picks the original door s/he cannot improve on picking a door at random.
- See homework for discussion of $P(H = i) = 1/3$ for $i = 1, 2, 3$.

What is a strategy?

- The player gets to pick
  - $p_j = P(C = j)$ – the probability that the player chooses door j.
  - A *strategy*: $q_{ij} = P(\text{Switch}|C = i, O = j)$.
- Want to compute

  $P(H = X|S = 1, C = i, O = j)$ and $P(H = X|S = 0, C = i, O = j)$ and

  for given strategy for switching.
- You control S as a function of D and O (and external randomization).

What else do we know?

- We *don't* know $P(O = k|C = j, H = i)$ which is Monte's strategy.
- We *do* know that if $i \neq j$ (the player has chosen wrongly) then $P(O = k|C = j, H = i) = 1$ for the k which is not i and not j.
- It really remains to specify $P(O = k|C = j, H = j)$. It might be natural for Monte to be making this value $1/2$ for the two possible k values. But this is not really clearly specified by the problem.

First Analysis

- Imagine you adopt the strategy $S = 1$; that is $P(S = 1) = 1$. In fact

$$P(S = 1|C, O) = 1 \text{ for all } C, O.$$

- Then the event $H = X$ is exactly the same event as $H \neq C$ because if the player picked the wrong door then the one you can switch to is the right door.
- So for this strategy $P(H = X) = 1 - P(H = C) = 2/3$.
- Similarly if you never switch you win with probability $1/3$.
- So you should switch.

Criticism of First Analysis

- The analysis did not answer the question about $P(H = X|S = 1)$ for a general strategy.
- What if you switch whenever Monte opens the door with the larger number? Could that be good?
- The strategy is

$$\begin{array}{ll} P(S=1|C=1,O=3)=1 & P(S=0|C=1,O=2)=0 \\ P(S=1|C=2,O=3)=1 & P(S=0|C=2,O=1)=0 \\ P(S=1|C=3,O=2)=1 & P(S=0|C=3,O=1)=0 \end{array} .$$

A computation for this strategy

- We try to compute as an example $P(X = H|S = 1)$.
- Back to basics, keeping track of when Monte has a choice $(H = C)$ and when not:

$$\begin{aligned} P(H = X|S = 1) &= \sum_{i=1}^{3} P(H = i, X = i|S = 1) \\ &= \sum_{i=1}^{3} P(H = i, X = i, S = 1)/P(S = 1) \\ &= \{P(H = 1, C = 2) + P(H = 2, C = 1)\}/P(S = 1) \end{aligned}$$

Computation Continued

Event $S = 1$ has following pieces:

1. I switch if $H = 1$ and $C = 2$ because Monte opens door 3.
2. I switch if $H = 2$ and $C = 1$ because Monte opens door 3.
3. I switch if $H = 3$, $C = 3$ and Monte opens door 2.
4. I switch if $H = 1$, $C = 1$ and Monte opens door 3.
5. I switch if $H = 2$, $C = 2$ and Monte opens door 3.

- Notice several probabilities depend on how I model Monte's behaviour. For instance, what is $P(O = 2|H = 3, C = 3)$?
- In general the story does not specify enough to compute all possible probabilities.

Related Problems

- Many game theory examples.
- Three prisoners; one to be executed.
- Three cards problem.
- Lottery strategy in 6/49.
- *tit for tat*
- Relation is in incomplete specification of the problem.

Step by Step Probabilities

- Game is sequence in time: $H \to C \to O \to S$
- Decompose joint distribution in same way

$$P(H = i, C = J, O = k, S = l) =$$
$$P(H = i)P(C = j|H = i)P(O = k|H = i, C = j)P(S{+}l|H = i, C = j, O$$

- Apply modelling assumptions to pieces.
- $P(C = j|H = i) = P(C = j)$ because player has no information about where the prize is hidden.
- $P(S = l|H = i, C = j, O = k) = P(S = l|C = j, O = k)$ for same reason.
- $P(S = l|C = j, O = k)$ is the player's strategy.
- $P(O = k|H = i, C = j)$ and $P(H = i)$ are summaries of what the player knows about Monte's strategy.