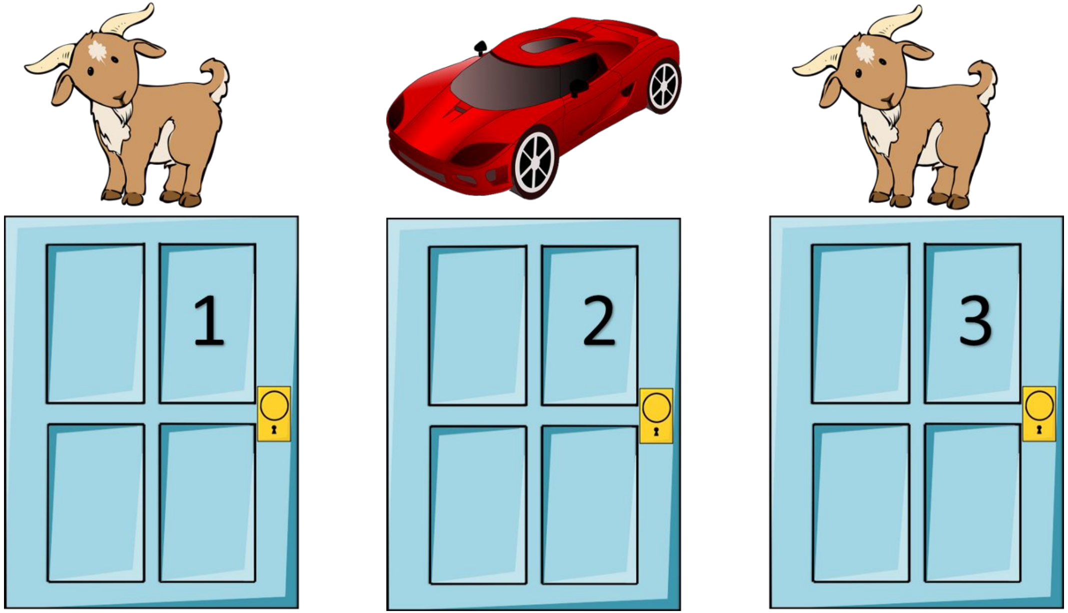
PHIL 313Q: Inductive Logic

ch. 8 of “Probability and Inductive Logic” by Ian Hacking

Alex Rausch
PhD Candidate, The University of Texas at Austin

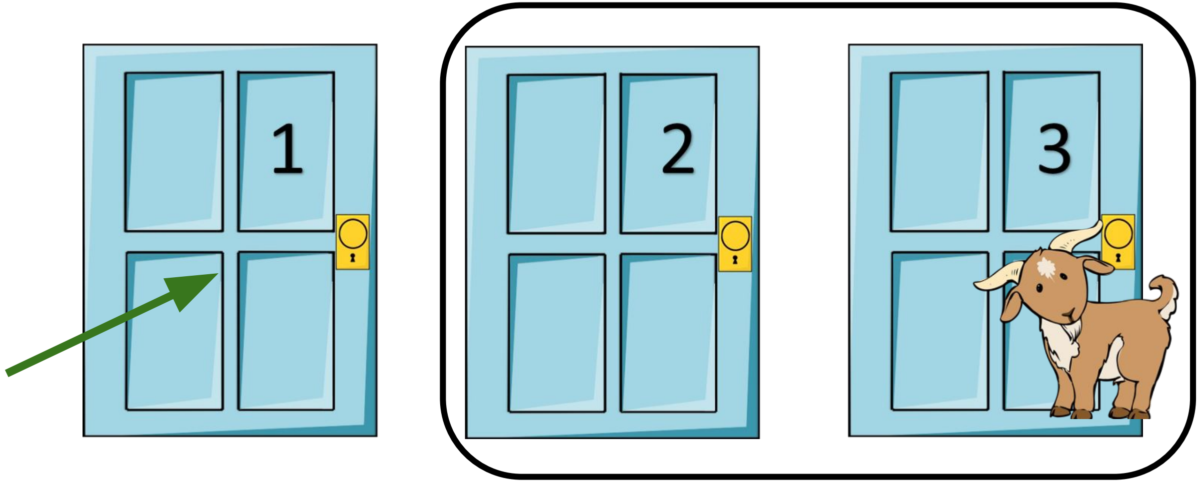
The Monty Hall Problem

- There are three doors. Behind one of them is a brand new car. Behind the other two are worthless goats.



The Monty Hall Problem

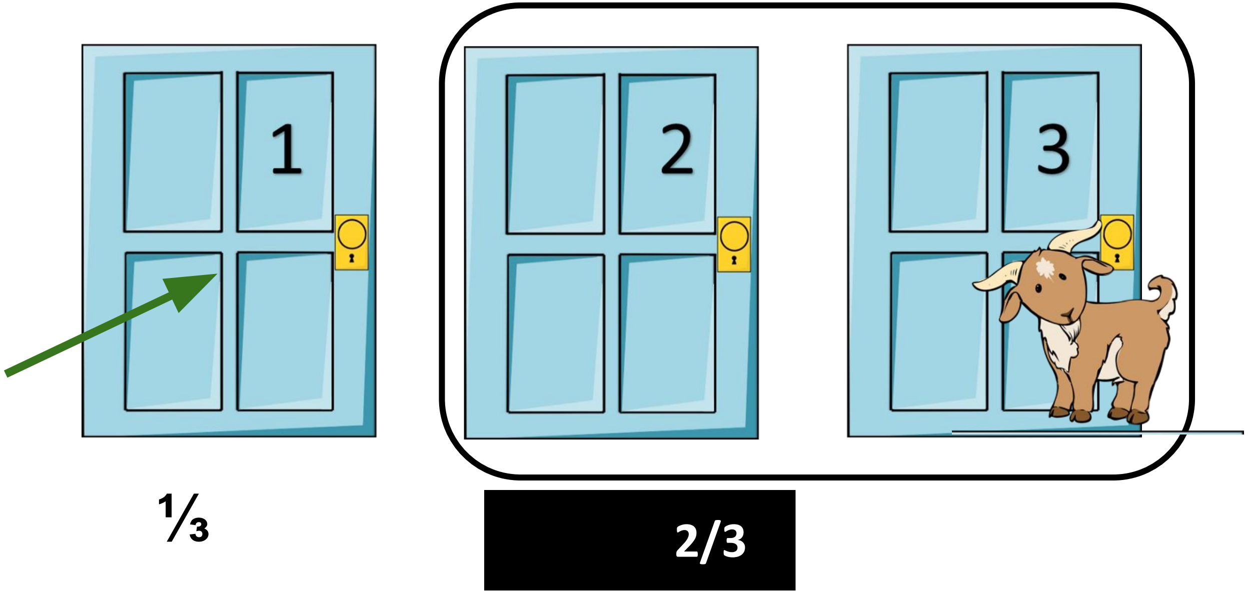
- Here's how the game works.
- First, you pick a door. Let's say you pick #1.



- I'll look behind the other two doors and show you the goat behind one of them.
- You now have the option to stick with your original choice (#1) or switch (#2).
- What do you do?

The Monty Hall Problem

- There is a ⅓ chance that the car is behind each door.
- There is a ⅔ chance that the car is behind #2 or #3.
- You select and I reveal…there is a ⅔ chance that the car is behind #2.



Expected Value

- Suppose there's a lottery with 100 tickets. The prize is $90.
- Actions, like *accepting a free ticket* (**A**), have possible consequences, like *winning the lottery* (**W**).
- A consequence has a chance, or probability, of occurring given an action.
  - The probability of *winning the lottery* given *accepting a free ticket* is 1/100.
  - **Pr(W | A)** = 1/100
- A consequence can be assigned a number representing its worth to us. Call this the **utility** of the consequence.
  - The utility of *winning the lottery* is $90.
  - **U(W)** = $90
- The **weighted utility** of a consequence (given an action) is the consequence's utility times its probability (given the action).
  - The weighted utility of *winning the lottery* given *accepting a free ticket* is **U(W) x Pr(W | A)** = $90 x 1/100 = $0.90

Expected Value

- The **expected value** of an action is the sum of its possible consequences' weighted utilities.
- What are the possible consequences of *accepting a free lottery ticket (**A**)* ?
  - **W** = Winning the lottery, **~W** = Not winning the lottery
- What's the utility and probability of each possible consequence given the action?
  - **U(W)** = $90, **Pr(W | A)** = 1/100
  - **U(~W)** = $0, **Pr(~W | A)** = 99/100
- What are the weighted utility of each possible consequence?
  - **U(W)** x **Pr(W | A)** = $90 x (1/100) = $0.90
  - **U(~W)** x **Pr(~W | A)** = $0 x (99/100) = $0
- What is the **expected value** of *accepting a free lottery ticket* (**A**)?
  - **Exp(A)** = (**U(W)** x **Pr(W | A)**) + (**U(~W)** x **Pr(~W | A)**) **=** $0.90 + $0 = $0.90
- A ticket is "worth" $0.90. *(Image you bought all 100 tickets. You'd be guaranteed to win $90. 100 tickets are worth $90. $90 / 100 tickets = $0.90 per ticket)*

Expected Value

- Suppose there is a lottery with 1250 tickets, each with a fair chance of being drawn.
- The tickets cost $0.75 each. The winner gets $937. Each loser receives a consolation prize of $0.05.
- Assuming you only care about making money on average, should you buy a ticket?
- **Exp(***buying a ticket***)** = (1)x(-0.75) + (1/1,250)x(937) + (1,249/1,250)x(0.05)
  = -0.75 + .7496 + .04996 = .04956
- Yes! You expect to make 4-5 cents on average per ticket.
  - Suppose you bought all 1,250 tickets at $0.75 each, so 937.5.
  - Then you win $937 and get ($0.05 x 1,249) = $62.45 in consolation.
  - $937 + $62.45 - $937.5 = $61.95 / 1,250 = $0.04956

Expected Value

- Let's play a game. Playing will cost you \$5. Here are the rules:
  - I draw a random card from a shuffled deck.
  - If it's red, you win \$8.
  - If it's a club, you win \$4.
  - If it's a spade, you win \$0.
- Want to play?
- **P** = you play the game; **F** = spend \$5; **R** = it's red; **C** = it's clubs, **S** = it's spades
- What's **Exp(P)**?
  - The consequences are **F**, **R**, **C**, **S**.
  - **Pr(F | P)** = 1; **Pr(R | P)** = ½; **Pr(C | P)** = ¼; **Pr(S | P)** = ¼
  - **U(F)** = -5; **U(R)** = 8; **U(C)** = 4; **U(S)** = 0
- **Exp(P) = U(F)Pr(F | P) + U(R)Pr(R | P) + U(C)Pr(C | P) + U(S)Pr(S | P)**

  = (-5)x(1) + (8)x(½) + (4)x(¼) + (0)x(¼)

  = -5 + 4 + 1 + 0 = 0
- The expected value of playing this game is \$0. You can expect to *break even*.

Expected Value

- But how do we *know* what *all* the possible consequences of an action are?
- Aren't there other possible consequences of playing the break-even game?
  - **U(***winning a single draw***)** = ?
  - **U(***losing a single draw***)** = ?
  - **U(***wasting time playing a break-even game)* = - ?
  - **U(***playing a game with Alex***)** = ?
  - …
- **Exp(P)** = **U(F)Pr(F | P)** + **U(R)Pr(R | P)** + **U(C)Pr(C | P)** + **U(S)Pr(S | P)** +
  U(*winning a single draw***)Pr(***winning a single draw* **| P)** +
  U(*losing a single draw***)Pr(***losing a single draw* **| P)** +
  U(*wasting time …* **)Pr(***wasting time…* **| P)** +
  U(*playing a game with Alex***)Pr(***playing a game with Alex* **| P)** +
  …
- There's *nothing* wrong with this. But most of our expected value calculations are **restricted for practical purposes to a set of relevant consequences**.

Expected Value

- Say action **A** has 3 consequences, $\mathbf{C_1}$, $\mathbf{C_2}$, and $\mathbf{C_3}$. Then

  (EQ) $\mathbf{Exp(A)} = \mathbf{U(C_1)Pr(C_1 \mid A)} + \mathbf{U(C_2)Pr(C_2 \mid A)} + \mathbf{U(C_3)Pr(C_3 \mid A)}$
- Why *this* formula?
  - If the utility of a consequence goes up (down), the expected value goes up (down)
  - The utility of more probable consequences counts for more
- But there are infinitely many formulas that satisfy those requirements, such as:
  - $\mathbf{Exp^*(A)} = \mathbf{U(C_1)^2Pr(C_1 \mid A)} + \mathbf{U(C_2)^2Pr(C_2 \mid A)} + \mathbf{U(C_3)^2Pr(C_3 \mid A)}$
  - $\mathbf{Exp^*(A)} = \mathbf{U(C_1)}/(1.01\text{-}\mathbf{Pr(C_1 \mid A)}) + \mathbf{U(C_2)}/(1.01\text{-}\mathbf{Pr(C_2 \mid A)}) + \mathbf{U(C_3)}/(1.01\text{-}\mathbf{Pr(C_3 \mid A)})$
- One reason to prefer (EQ): it calculates the average utility out of 100 trials given the relative frequencies of the consequences. Suppose $\mathbf{C_1 : C_2 : C_3} = 3:2:5$. Then

  $\mathbf{Exp(A)} = \mathbf{U(C_1)}(30/100) + \mathbf{U(C_2)}(20/100) + \mathbf{U(C_3)}(50/100)$

  $= (\mathbf{U(C_1)}30 + \mathbf{U(C_2)}20 + \mathbf{U(C_3)}50) / 100$
- Perhaps EQ is easier to grasp, or more intuitive, given how we think (100 = one whole)
- But other functions *could* play a similar role – the role of *letting us compare the values*

Expected Value

- A bad storm has 20% probability. You have to decide between train and plane.
- The **TRAIN** will take at least 1 hour no matter what
  - Plus 5 additional hours *if no storm*
  - Plus 7 additional hours *if storm*
- The **PLANE** will take at least 3 hours no matter what
  - Plus 1 additional hour *if no storm*
  - Plus 10 additional hours *if storm*
- What's the expected travel time (expected value) of traveling by each?
- **Exp(TRAIN)** = 1 + 5x(.80) + 7x(.20) = 6.4 hours
- **Exp(PLANE)** = 3 + 1x(.80) + 10x(.20) = 5.8 hours
- So, basing our decision *only* on the expected travel time, we should choose the plane.
- Of course, if we allow other consequences into our calculation, and assign these consequences utility in terms of "hours"...that can change our answer.
  - Maybe spending 7 hours in a train station is *really* bad, even worse than spending 10 hours in an airport!

Expected Value

- Imagine a raffle with only 4 tickets, each with ¼ probability of being drawn.
- Two tickets will be drawn. The first ticket wins $90. The second wins $9.
- What's the expected value of accepting a free ticket?

Ticket 1	Ticket 2	Ticket 3	Ticket 4

- A ticket has a ¼ chance of being Ticket 1, a ¼ chance of being Ticket 2, a ¼ chance of being Ticket 3, and a ¼ chance of being Ticket 4.
- **Exp(**accepting a free ticket**)** = $90x(¼) + <u>$9x(¼)</u> + $0x(¼) + $0x(¼) = 24.75.

^note that this is **NOT** $9(⅓),
i.e. **NOT** $9 times the conditional probability of getting your ticket pulled 2nd, assuming it did not get pulled first

Martingale Betting Strategy

- Consider a simplified version of roulette, where there's a 50% chance of landing red.
- If you bet $\$y$ on red and win, you get $\$2y$.
- You deploy this strategy:
  - Bet \$10 on red on spin 0. If you win, claim \$20 and stop. Otherwise:
  - Bet \$20 on red on spin 1. If you win, claim \$40 and stop. Otherwise:
  - Bet \$40 on red on spin 2. If you win, claim \$80 and stop. Otherwise:
  - …
  - Bet $\$(2^n)(10)$ on red on spin *n*. If you win, claim $\$(2^n)(20)$ and stop.
- If you win on Spin 1, net earnings are: +20 - 10 = 10
- If you win on Spin 2, net earnings are: +40 - 20 - 10 = 10
- If you win on Spin 3, net earnings are: +80 - 40 - 20 - 10 = 10
- …
- If you win on Spin *n,* net earnings are: $+(2^n)(20) - 2^n(10) - (2^n)(5) - \ldots - 20 - 10 = 10$

Martingale Betting Strategy

Spin #	Bet required
0	$10
1	$20
2	$40
3	$80
4	$160
5	$320
6	$640
7	$1280
8	$2560

Recall: there is a 52% chance of 7 heads in a row in 100 flips

St. Petersburg Paradox

- Imagine a fair coin that you flip until you get heads, with the following payouts:

Flip #	Payout
1	$2
2	$4
3	$8
4	$16
n	$\$2^n$

- How much would you pay to play this game?

St. Petersburg Paradox

Flip #	Payout
1	\$2
2	\$4
3	\$8
4	\$16
n	$\$2^n$

Exp(playing**)** $= \frac{1}{2}(2) + \frac{1}{4}(4) + (\frac{1}{8})8 + (1/16)16 + \ldots$

$= 1 + 1 + 1 + 1 + \ldots$

$= \infty$

There is no upper bound on what you should be willing to pay to play this game.

St. Petersburg Paradox: What Went Wrong?

1. **“Nothing! Most of the time, the prize is small. But for any amount of money, there is some chance you’ll win it!”**
   a. No one in their right mind would pay much to enter this game.
2. **“The game is impossible. You can’t flip a real coin infinitely many times.”**
   a. Let’s imagine the game ends at 1,000 flips, no matter what. Then the expected value of playing is $1,000. Would you even pay 10% of that ($100) to enter?
3. **“The utility of money diminishes as you make more of it. So, the $1 you can make from the first few flips really is worth 1 unit of utility, but the $1s you can make after 500 flips are worth less than 1 unit of utility.”**
   a. Diminishing returns is a useful concept, but most people won’t even pay $20 to enter this game. The value of money would need to *drastically* drop off after about $32-$64 for this explanation to be plausible.
4. **“I simply don’t care about utilities under a certain threshold of consequences with *extremely low* probability. This game might as well stop after the 20th flip for me – it’s less than 1 in a million odds that the game will get that far.”**

Decision Theory

- According to the *expected value rule*, if you are trying to decide between multiple actions, you should (rationally) do whatever maximizes expected value.
- Suppose you are deciding between parking illegally on the street (10% chance of a $20 fine) or in a lot (100% chance of 3$ fee).
  - **Exp(***illegally parking***)** = .1 x (-$20) = -$2
  - **Exp(***legally parking***)** = 1 x (-$3) = -$3
- So the expected value rule says you should illegally park.
- But surely someone can rationally choose to legally park. These folks are “risk-averse”.
- And others can rationally choose to illegally park. These folks are “risk-seeking”.
- Two Options:
  - 1. The expected value rule is just wrong. There is more to what you should rationally do than maximize expected value.
  - 2. The expected value rule is correct, but *taking a risk* can be assigned a positive or negative utility depending on the person.

Decision Theory

- Expected Value Rule Advocate: “Risk-averse folks assign a -$5 value to taking the risk. Risk-seeing folks assign $2 to taking the risk. Factor in the utility of risk for each kind of person and you’ll find that it is *always* rational to maximize utility.”
- Suppose you are deciding between parking illegally on the street (10% chance of a $20 fine **and 100% chance of taking a risk**) or in a lot (100% chance of 3$ fee **and 100% of not taking a risk**).
- If you are risk-averse:
  - **Exp(***illegally parking***)** = (.1 x -$20) + **(1 x -$5)** = -$7
  - **Exp(***legally parking***)** = (1 x -$3) + **(1 x $0)** = -$3
- If you are risk-seeking:
  - **Exp(***illegally parking***)** = (.1 x -$20) + **(1 x $2)** = $0
  - **Exp(***legally parking***)** = (1 x -$3) + **(1 x $0)** = -$3
- Objection: This assumes that all different kinds of consequences can be weighed against each other on a single scale (money, risk, anxiety, etc.). It’s not plausible to represent risk-aversion as a utility, since the exact number would be arbitrary.

The Allais Paradox

- **Choice A.** 61% chance to win $520,000
- **Choice B.** 63% chance to win $500,000
    - Many people prefer Choice A – as per Expected Value Rule.
    - **Exp(A)** = .61 x $520,000 = **$317,200**
    - **Exp(B)** = .63 x $500,000 = $315,000
- **Choice F:** 98% chance to win $520,000
- **Choice G:** 100% chance to win $500,000
    - Many people prefer Choice G – against Expected Value Rule.
    - **Exp(F)** = .98 x $520,000 = **$509,600**
    - **Exp(G)** = 1 x $500,000 = $500,000
- What does the Allais Paradox show about these people?
    - 1. They are irrational.
    - 2. The simple expected value equation is incorrect because more probable consequences (approaching 1) should have more weight.

Expected Value

- (EQ) $\mathbf{Exp(A) = U(C_1)Pr(C_1 \mid A) + U(C_2)Pr(C_2 \mid A) + \ldots + U(C_n)Pr(C_n \mid A)}$
- St. Petersburg Paradox provides some evidence that EQ needs to be updated because *the utility of consequences with very small probabilities should weigh less*.
- Allais Paradox provides some evidence that EQ needs to be updated because *the utility of consequences with higher probabilities (approaching 1) should weigh more*.
- So let's define a function **α** from [0,1] into [0,1].
    - **α** maps the "lower end" of [0,1] to [0,.1].
    - **α** maps the "higher end" of [0,1] to [1,1.1]
- (EQ+) $\mathbf{Exp(A) = U(C_1)Pr(C_1 \mid A)\alpha Pr(C_1 \mid A) + U(C_2)Pr(C_2 \mid A)\alpha Pr(C_2 \mid A) + \ldots}$
- Are there counter-examples to EQ+, which would require further revision?
- What's the goal here? Are we discovering the *rules of rational decision making* or are we merely *describing the actual decision making behavior of folks*?

Group Exercises

- Ch. 8, p. 96 #2
- Get a head start on problem set, ask any questions about it