Decision Theory

CompSci 370
Department of Computer Science
Duke University
Ron Parr

Decision Theory

What does it mean to make an optimal decision?

- Asked by economists to study consumer behavior
- Asked by MBAs to maximize profit
- Asked by leaders to allocate resources
- Asked in OR to maximize efficiency of operations
- Asked in AI to model intelligence

By Nick Ares - https://www.flickr.com/photos/aresauburnphotos/4956000628/, CC BY-SA 2.0, https://commons.wikimedia.org/w/index.php?curid=16266369

- Asked (sort of) by any intelligent person every day

Utility Functions

- A *utility function* is a mapping from world states to real numbers
- Sometimes called a *value function*
- Rational or optimal behavior is typically viewed as maximizing expected utility:

$$\max_a \sum_s P(s \mid a) U(s)$$

a = actions, s = states

Are Utility Functions Natural?

- Some have argued that people don't really have utility functions
  - What is the utility of the current state?
  - What was your utility at 8:00pm last night?
  - *Utility elicitation* is difficult problem

- It's easy to communicate *preferences*
- Theorem (sorta): Given a plausible set of assumptions about your preferences, there must exist a consistent utility function

Axioms of Utility Theory

- Orderability: $(A \succ B) \vee (A \prec B) \vee (A \sim B)$
- Transitivity: $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
- Continuity: $A \succ B \succ C \Rightarrow \exists p [p, A; 1-p, C] \sim B$
- Substitutability: $A \sim B \Rightarrow [p, A; 1-p, C] \sim [p, B; 1-p, C]$
- Monotonicity: $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; 1-p, B] \geq [q, A; 1-q, B])$
- Decomposability:
  $[p, A; (1-p), [q, B; (1-q), C]] \sim [p, A; (1-p)q, B; (1-p)(1-q), C]$

Bet/gamble between A and C

Consequences of Preference Axioms

- Utility Principle
  - There exists a real-valued function U:

$$U(A) > U(B) \Leftrightarrow A \succ B$$

$$U(A) = U(B) \Leftrightarrow A \sim B$$

- Expected Utility Principle
  - The utility of a lottery can be calculated as:

$$U([p_1, S_1; \ldots; p_n, S_n]) = \sum_i p_i U(S_i)$$

More Consequences

- Scale invariance

- Shift invariance

Maximizing Utility

- Suppose you want to be famous
- You can be either (N,M,C)
  - Nobody
  - Modestly Famous
  - Celebrity
- Your utility function:
  - U(N) = 20
  - U(M) = 50
  - U(C) = 100

- You have to decide between going to grad school and becoming a professor (G) or going to Hollywood and becoming an actor (A)

Outcome Probabilities

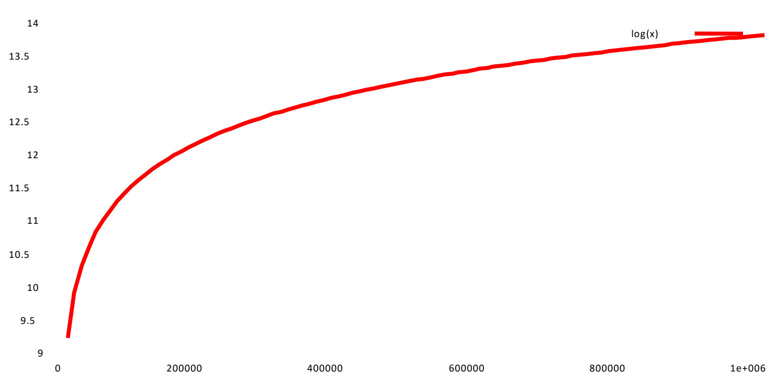
- P(N|G)=0.5, P(M|G)=0.4, P(C|G)=0.1
- P(N|H)=0.6, P(M|H)=0.2, P(C|H)=0.2
- Maximize expected utility:
  - U(N) = 20, U(M) = 50, U(C) = 100

$$EU_G = 0.5(20) + 0.4(50) + 0.1(100) = 40$$
$$EU_H = 0.6(20) + 0.2(50) + 0.2(100) = 42$$

Hollywood wins!

Utility of Money

- How much happier are you with an extra $1M?
- How much happier is Jeff Bezos with an extra $1M?
- Some have proposed:

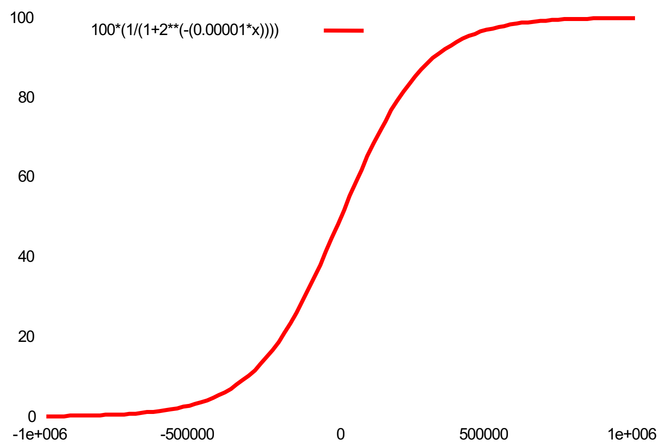


Utility of Money

- U(money) should drop slowly in negative region too
- If you're solvent, losing $1M is pretty bad
- If already $10M in debt, losing another $1M isn't that bad
- Utility of money is probably sigmoidal (S shaped)

A Sigmoidal Utility Function

$$U(\$X) = 100 \frac{1}{1 + 2^{-0.00001X}}$$



Utility & Gambling

- Suppose U($X)=X, would you spend $1 for a 1 in a million chance of winning $1M?

- Suppose you start with c dollars:
  - EU(gamble)=1/1000000(1000000+(c-1))+(1-1/1000000)(c-1)=c
  - EU(do_nothing)=c

- Starting amount doesn't matter
- You have no expected benefit from gambling

Sigmoidal Utility & Gambling

- Suppose: $U(\$X) = 100\frac{1}{1+2^{-0.00001X}}$

- Suppose you start with $1M
  - EU(gamble)-EU(do_nothing)=$-5.7*10^{-7}$
  - Winning is worthless

- Suppose you start with -$1M
  - EU(gamble)-EU(do_nothing)=$+4.9*10^{-5}$
  - Gambling is rational because losing doesn't hurt

Convexity & Gambling

- Convexity: $f(\alpha x+(1-\alpha)y) \le \alpha f(x)+(1-\alpha)f(y)$

  $0 \le \alpha \le 1$
- Suppose x and y are in the convex region of the utility function and are possible outcomes of a bet
- Current cash on hand is x<z<y
- Suppose bet has 0 expected change in monetary value: $z = \alpha x + (1-\alpha)y$
- Will the bet be accepted?
  - Utility of doing nothing: $U(z) = U(\alpha x + (1-\alpha)y)$
  - Expected utility of accepting the bet: $\alpha U(x)+(1-\alpha)U(y)$

Value of Information

- Expected utility of action a with evidence E:

$$\mathrm{EU}_{\mathrm{E}}(A \mid E) = \max_{a\in A} \sum_i P(S_i \mid E, a) U(S_i)$$

- Expected utility given new evidence E'

$$\mathrm{EU}_{E,E'}(A \mid E, E') = \max_{a\in A} \sum_i P(S_i \mid E, E', a) U(S_i)$$

- Value of knowing E' (**V**alue of **P**erfect **I**nformation)

$$\mathrm{VPI}_E(E') = \left(\sum_{E'} P(E' \mid E)\,\mathrm{EU}_{E,E'}(A' \mid E, E') \right) - \mathrm{EU}_{\mathrm{E}}(A \mid E)$$

Expected utility given New information (weighted)

Previous Expected utility

VPI Example

- Should you pay to subscribe for traffic information? Assume:
  - Time = cost = -utility
  - Cost of taking highway to work (w/o traffic_jam) = 15
  - Cost of taking highway to work (w/traffic_jam) = 30
  - Cost of taking local roads to work = 20
  - P(traffic_jam) = 0.15

- Steps:
  - Determine optimal decision w/o information: EU(A|{})
  - Determine optimal decisions given information: $EU_T(A|T)$
  - Compute expected value of optimal decisions given T
  - Estimate value of information (difference in prev. slide)

VPI for Traffic Info

- Cost of local roads = 20
- Cost of highway = 0.15*30 + 0.85*15 = 17.25

EU(A|{})

- Traffic = true case: Take local roads; cost = 20
- Traffic = false case: Take highway; cost = 15

$EU_T(A|T)$

- Expected cost: 0.15*20 + 0.85*15 = 15.75

Expectation

- Value = 1.5

VPI

- **Important:** In this case, the optimal choice given the information was trivial. In general, we may to do more computation to determine the optimal choice given new information – not all decisions are "one shot"

How Information is Doled Out

- VPI = Value of Perfect Information
- In practice, information is:
  - Partial
  - Imperfect
- Partial information:
  - We learn about some state variables, but don't learn the exact state of the world
  - Example: We can see a traffic camera at one intersection, but we don't have coverage of our entire route
- Imperfect information:
  - We learning something that may not be reliable
  - Example: There may be a lag in our traffic data
- Our framework can handle this by introducing an extra variable. (We get perfect information about the observed variable, and this influences the distribution over the others.)

Examples Where Value of Information is (should be) Considered

- Medical tests (x-rays, CT-scans, mammograms, etc.)
- Pregnancy tests
- Pre-purchase house/car inspections
- Subscribing to Consumer Reports
- Hiring consultants
- Hiring a trainer
- Funding research
- Checking one's own credit score
- Checking somebody else's credit score
- Background checks
- Drug tests
- Real time stock prices
- Etc.

Properties of VPI

- VPI is non-negative!
- VPI is order independent
- VPI is not additive

- VPI is easy to compute and is often used to determine how much you should pay for ***one*** extra piece of information. Why is this myopic?

For example, knowing X AND Y together may useful, while knowing just one alone may be useless.

More Properties of VPI

- Acquiring information optimally is very difficult

- Need to construct a conditional plan for every possible outcome before you ask for even the first piece of information
  - Suppose you're a doctor planning to treat a patient
  - Picking the optimal test to do first requires that you consider all subsequent tests and all possible treatments as a result of these tests

- General versions of this problem are intractable!

Conclusions

- Decision theory provides a framework for optimal decision making
- Principle: Maximize Expected Utility
- Easy to describe in principle
- Application to multistep problems can require advanced planning and probabilistic reasoning techniques