

Bayes AI

Unit 2: Utility and Decision Theory

Vadim Sokolov George Mason University Spring 2025

Probability and Psychology

How do people form probabilities or expectations in reality?

Psychologists have categorized many different biases that people have in their beliefs or judgments.

Loss Aversion

The most important finding of Kahneman and Tversky is that people are loss averse.

Utilities are defined over gains and losses rather than over final (or terminal) wealth, an idea first proposed by Markowitz. This is a violation of the EU postulates. Let (x, y) denote a bet with gain x with probability y .

To illustrate this subjects were asked:

In addition to whatever you own, you have been given \$1000, now choose between the gambles $A = (1000, 0.5)$ and $B = (500, 1)$.

B was the more popular choice.

Example

The same subjects were then asked: *In addition to whatever you own, you have been given \$2000, now choose between the gambles $C = (-1000, 0.5)$ and $D = (-500, 1)$.*

- ▶ This time C was more popular.
- ▶ The key here is that their final wealth positions are identical yet people chose differently. The subjects are apparently focusing only on gains and losses.

When they are not given any information about prior winnings, they choose B over A and C over D . Clearly for a risk averse people this is the rational choice.

- ▶ This effect is known as loss aversion.

Representativeness

Representativeness

When people try to determine the probability that evidence A was generated by model B , they often use the representative heuristic. This means that they evaluate the probability by the degree to which A reflects the essential characteristics of B .

- ▶ A common bias is *base rate neglect* or ignoring prior evidence.
- ▶ For example, in tossing a fair coin the sequence *HHTHTHHTHH* with seven heads is likely to appear and yet people draw conclusions from too few data points and think 7 heads is representative of the true process and conclude $p = 0.7$.

Expected Utility (EU) Theory: Normative

Let P, Q be two probability distributions or *risky gambles/lotteries*.

$pP + (1 - p)Q$ is the *compound* or *mixture* lottery.

The rational agent (You) will have preferences between gambles.

- ▶ We write $P \succeq Q$ if and only if You strictly prefer P to Q . If two lotteries are *indifferent* we write $P \sim Q$.
- ▶ EU – a number of plausible axioms – completeness, transitivity, continuity and independence – then preferences are an expectation of a utility function.
- ▶ The theory is a *normative* one and not necessarily *descriptive*. It suggests how a rational agent should formulate beliefs and preferences and not how they actually behave.
- ▶ Expected utility $U(P)$ of a risky gamble is then

$$P \succeq Q \iff U(P) \geq U(Q)$$

Key Facts

The two key facts then are uniqueness of probability and existence of expected utility. Formally,

1. If $P \succeq R \succeq Q$ and $wP + (1 - w)Q \sim R$ then w is unique.
2. There exists an expected utility $U(\cdot)$ such that
 $P \succeq Q \iff U(P) \geq U(Q)$. Furthermore

$$U(wP + (1 - w)Q) = wU(P) + (1 - w)U(Q)$$

for any P, Q and $0 \leq w \leq 1$.

This implies that U is additive and it is also unique up to affine transformation.

St. Petersburg Paradox

What are you willing to pay to enter the following game?

- ▶ I toss a fair game and when the first head appears, on the T th toss, I pay you $\$2^T$ dollars.
- ▶ First, probability of first head on T th toss is 2^{-T}

$$\begin{aligned}E(X) &= \sum_{T=1}^{\infty} 2^T 2^{-T} \\&= 2(1/2) + 4(1/4) + 8(1/8) + \dots \\&= 1 + 1 + 1 + \dots \rightarrow \infty\end{aligned}$$

- Bernoulli (1754) constructed utility to value bets with $E(u(X))$.

Some examples of utility functions are,

- ▶ $U(x) = V_0(1 - x^{-\alpha})$, $\alpha > 0$, which gives an expected utility of $V_0 \left(1 - \frac{1}{2^{\alpha+1}-1}\right)$
- ▶ Log utility, $U(x) = \log(x)$, with expected value $2 \log(2)$.

Notice that after obtaining an expected utility value you'll have to

Attitudes to Risk

Two gambles

1. get P_1 for sure
 2. get $P_2 = P_1 + k$ and $P_3 = P_1 - k$ with probability $1/2$.
- ▶ Risk neutral: indifferent about fair bets, has linear utility.
 - ▶ Risk averse: prefers certainty over fair bets, concave utility.
 - ▶ Risk loving: prefers fair bets over certainty, convex utility.

Then we will compare the utility of those gambles

Attitudes to Risk

The solution depends on your *risk* preferences:

- ▶ *Risk neutral*: a risk neutral person is indifferent about fair bets.

Linear Utility

- ▶ *Risk averse*: a risk averse person prefers certainty over fair bets.

$$\mathbb{E}(U(X)) < U(\mathbb{E}(X)) .$$

Concave utility

- ▶ *Risk loving*: a risk loving person prefer fair bets over certainty.

Depends on your preferences.

Ellsberg Paradox

Probability is counter-intuitive!!!

Two urns

1. 100 balls with 50 red and 50 blue.
 2. A mix of red and blue but you don't know the proportion.
- ▶ Which urn would you like to bet on?
 - ▶ People don't like the "uncertainty" about the distribution of red/blue balls in the second urn.

Allais Paradox

You have to make a choice between the following gambles

First compare the “Gambles”

Experiment 1		Experiment 2			
Gamble G_1		Gamble G_2	Gamble G_3		Gamble G_4
Win	Chance	Win	Win	Chance	Win
\$25m	0	\$25m	\$25m	0	\$25m
\$5m	1	\$5m	\$5m	0.11	\$5m
\$0m	0	\$0m	\$0m	0.89	\$0m

If $G_1 \geq G_2$ then $G_3 \geq G_4$ and vice-versa.

Solution: Expected Utility

Given (subjective) probabilities $P = (p_1, p_2, p_3)$. Write $E(U|P)$ for expected utility. W.l.o.g. set $u(0) = 0$ and for the high prize set $u(\$25 \text{ million}) = 1$. Which leaves one free parameter $u = u(\$5 \text{ million})$.

- ▶ Hence to compare gambles with probabilities P and Q we look at the difference

$$E(u|P) - E(u|Q) = (p_2 - q_2)u + (p_3 - q_3)$$

- ▶ For comparing G_1 and G_2 we get

$$E(u|G_1) - E(u|G_2) = 0.11u - 0.1$$

$$E(u|G_3) - E(u|G_4) = 0.11u - 0.1$$

The order is the same, given *your* u .

- ▶ If your utility satisfies $u < 0.1/0.11 = 0.909$ you take the “riskier” gamble.

Power Utility

Power and log-utilities

- ▶ Constant relative risk aversion (CRRA).
- ▶ Advantage that the optimal rule is unaffected by wealth effects.
The CRRA utility of wealth takes the form

$$U_{\gamma}(W) = \frac{W^{1-\gamma} - 1}{1 - \gamma}$$

- The special case $U(W) = \log(W)$ for $\gamma = 1$.

This leads to a myopic Kelly criterion rule.

Kelly Criterion

Kelly Criterion corresponds to betting under binary uncertainty. -
Consider a sequence of i.i.d. bets where

$$p(X_t = 1) = p \text{ and } p(X_t = -1) = q = 1 - p$$

The optimal allocation is $\omega^* = p - q = 2p - 1$. - Maximising the expected long-run growth rate leads to the solution

$$\begin{aligned} \max_{\omega} \mathbb{E}(\ln(1 + \omega W_T)) &= p \ln(1 + \omega) + (1 - p) \ln(1 - \omega) \\ &\leq p \ln p + q \ln q + \ln 2 \text{ and } \omega^* = p - q \end{aligned}$$

Kelly Criterion

- ▶ Let p denote the probability of a gain and $O = (1 - p)/p$ the odds. We can generalize the rule to the case of asymmetric payouts (a, b) where

$$p(X_t = 1) = p \text{ and } p(X_t = -1) = q = 1 - p$$

- Then the expected utility function is

$$p \ln(1 + b\omega) + (1 - p) \ln(1 - a\omega)$$

- The optimal solution is

$$\omega^* = \frac{bp - aq}{ab} = \frac{p - q}{\sigma}$$

Kelly Criterion

- ▶ If $a = b = 1$ this reduces to the pure Kelly criterion.
- ▶ A common case occurs when $a = 1$ and market odds $b = O$.
The rule becomes

$$\omega^* = \frac{p \cdot O - q}{O}$$

Two possible market opportunities: one where it offers you 4/1 when you have personal odds of 3/1 and a second one when it offers you 12/1 while you think the odds are 9/1.

In expected return these two scenarios are identical both offering a 33% gain.

In terms of maximizing long-run growth, however, they are not identical.

Example

- Table below shows the Kelly criteria advises an allocation that is twice as much capital to the lower odds proposition: 1/16 weight versus 1/40.

Market	You	p	ω^*
4/1	3/1	1/4	1/16
12/1	9/1	1/10	1/40

- The optimal allocation $\omega^* = (pO - q)/O$ is

$$\frac{(1/4) \times 4 - (3/4)}{4} = \frac{1}{16} \text{ and } \frac{(1/10) \times 12 - (9/10)}{12} = \frac{1}{40}$$

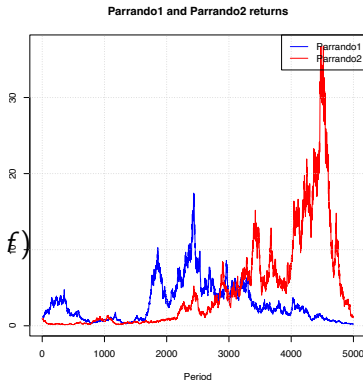
Parrando's Paradoxes

Two losing bets can be combined to a winner

- ▶ Bernoulli market: $1 + f$ or $1 - f$ with $p = 0.51$ and $f = 0.05$
- ▶ Positive Expectation
- ▶ Caveat: Growth governed by the median/entropy

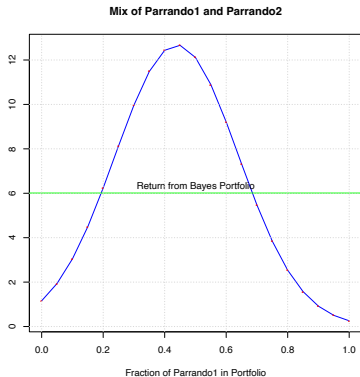
$$p \log(1 + f) + (1 - p) \log(1 - f) \\ = -0.00025 < 0$$

Brownian Ratchets and
cross-entropy of Markov
processes

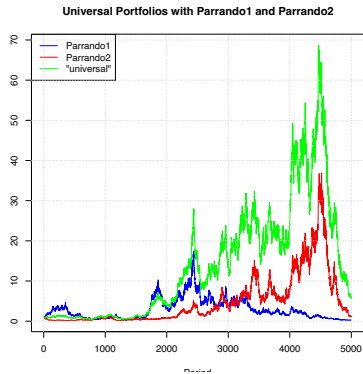


Two Losing Bets+Volatility

Parrando



(a) Optimal allocation ω



(a) Ex Ante vs Ex Post Realisation

Breiman-Kelly-Merton Rule

Kelly Criterion: Optimal wager in binary setting

$$\omega^* = \frac{p \cdot O - q}{O}$$

Merton's Rule: in continuous setting is Kelly

$$\omega^* = \frac{1}{\gamma} \frac{\mu}{\sigma^2}$$

1. μ : (excess) expected return
2. σ : volatility
3. γ : risk aversion
4. ω^* : optimal position size
5. $p = \text{Prob}(Up)$, $q = \text{Prob}(Down)$, $O = \text{Odds}$

Example: Kelly Criterion S&P500:

Consider logarithmic utility (CRRA with $\gamma = 1$). This is a pure Kelly rule.

- ▶ We assume iid log-normal stock returns with an annualized expected excess return of 5.7% and a volatility of 16% which is consistent with long-run equity returns. In our continuous time formulation $\omega^* = 0.057/0.16^2 = 2.22$ and the Kelly criterion which imply that the investor borrows 122% of wealth to invest a total of 220% in stocks. This is the risk-profile of the Kelly criterion.
- ▶ One also sees that the allocation is highly sensitive to estimation error in $\hat{\mu}$. We consider dynamic learning in a later section and show how the long horizon and learning affects the allocation today.

Fractional Kelly

The fractional Kelly rule leads to a more realistic allocation.

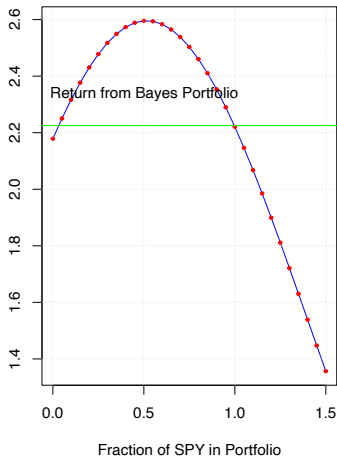
- ▶ Suppose that $\gamma = 3$. Then the informational ratio is

$$\frac{\mu}{\sigma^2} = \frac{0.057}{0.16} = 0.357 \text{ and } \omega^* = \frac{1}{3} \frac{0.057}{0.16^2} = 74.2\%$$

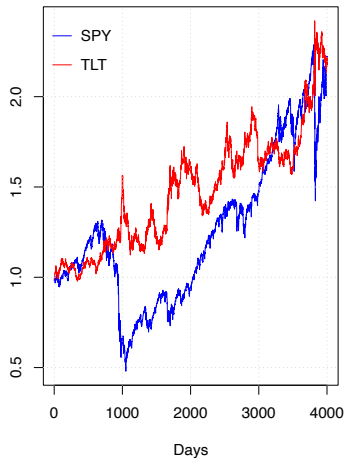
- ▶ An investor with such a level of risk aversion then has a more reasonable 74.2% allocation.
- ▶ This analysis ignores the equilibrium implications. If every investor acted this way, then this would drive up prices and drive down the equity premium of 5.7%.

60-40 Rule

Mix of SPY and TLT

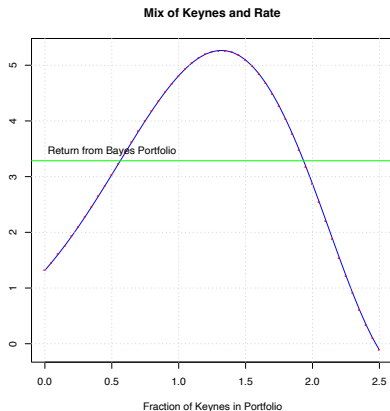


SPY and TLT returns

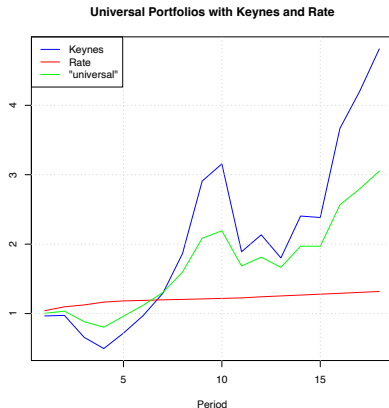


Keynes

Optimal Bayes Rebalancing



(a) Optimal allocation ω



(a) Keynes vs Universal vs Cash

Winner's Curse

Immediately after you have win, you should feel a little regret!

Claiming racehorse whose value is uncertain

Value	Outcome
0	horse never wins
50,000	horse improves

Simple expected value tells you

$$E(X) = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 50,000 = \$25,000.$$

In a \$20,000 claiming race (you can buy the horse for this fixed fee ahead of time from the owner) it looks like a simple decision to claim the horse.

Asymmetric information!

Lemon's Problem

- ▶ Asymmetric information.
- ▶ Proposed by George Akerlof in his 1970 paper "The Market for Lemons: Quality Uncertainty and the Market Mechanism."
- ▶ The lemons principle: low-value cars force high-value cars out of the market because of the asymmetrical information
- ▶ Seller does not know what the true value of a used car is and, not willing to pay a premium o
- ▶ Sellers are not willing to sell below the premium price so this results in only lemons being sold.

Lemon's Problem

- ▶ Suppose that a dealer pays \$20K for a car and wants to sell for \$25K, a lemon is only worth \$5K.
- ▶ Let's first suppose only 10% of cars are lemons, the customer's calculations are

$$E(X) = \frac{9}{10} \cdot 25 + \frac{1}{10} \cdot 5 = \$23K$$

- ▶ Dealer is missing \$2000. Therefore, they should try and persuade the customer its not a lemon by offering a warranty for example.

Lemon's Problem

- ▶ The more interesting case is when $p = 0.5$. The customer now values the car at

$$E(X) = \frac{1}{2} \cdot 25 + \frac{1}{2} \cdot 5 = \$15K$$

- ▶ This is lower than the \$20K – the reservation price that the dealer would have for a good car. Now what type of car and at what price do they sell?
- ▶ At \$15K dealer is only willing to sell a lemon.
- ▶ But then if the customer computes a conditional expectation

$$E(X | L) = 1 \cdot 5 = \$5K$$

Therefore only lemons sell, at \$ 5K, even if the dealer has a perfectly good car the customer is not willing to buy!

Again what should the dealer do?

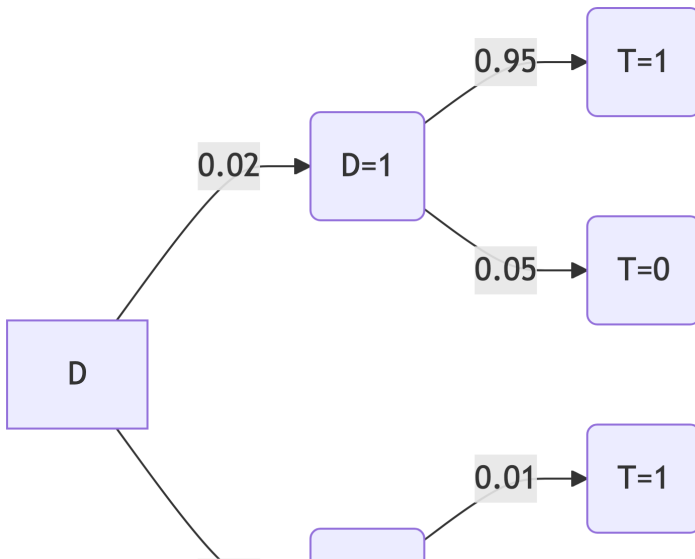
Decision Trees

Medical Testing

- ▶ A patient goes to see a doctor.
- ▶ The doctor performs a test which is 95% sensitive – that is 95 percent of people who are sick test positive and 99% specific – that is 99 percent of the healthy people test negative.
- ▶ The doctor also knows that only 1 percent of the people in the country are sick. Now the question is: if the patient tests positive, what are the chances the patient is sick? The intuitive answer is 99 percent, but the correct answer is 66 percent.

Decision Trees: Medical Testing

- ▶ $D = 1$ that indicates you have a disease
- ▶ $T = 1$ that indicates you tested positive



Medical Testing: Intuition

- ▶ Imagine that the above story takes place in a small town, with 1,000 people.
- ▶ Prior: 20 people, are sick, and 980 are healthy.
- ▶ Administer the test to everyone: 19 of the 20 sick people test positive, 9.8 of the healthy people test positive, we round it to 10.
- ▶ Now if the doctor sends everyone who tests positive to the national hospital, there will be 10 healthy and 19 sick patients. 1 to 2 ratio or 66 percent of the patients are healthy.

Medical Testing: With utility

- The decision problem is to treat a_T or not to treat a_N

Table 5: Utility of the test and the treatment.

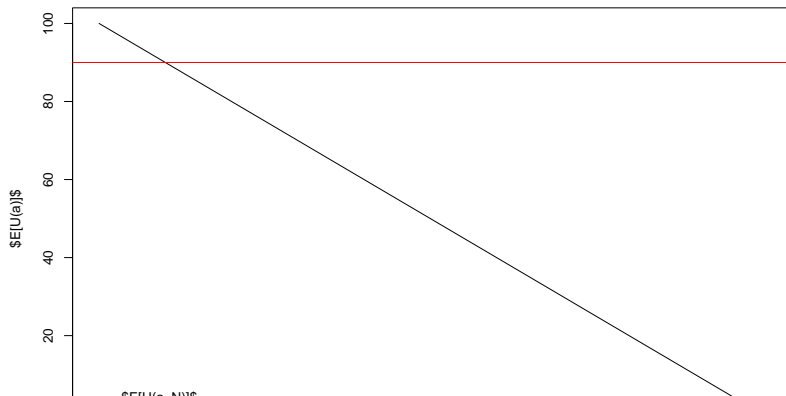
A/S	a_T	a_N
D_0	90	100
D_1	90	0

Then expected (unconditional) utility of the treatment is 90 and no treatment is 98. A huge difference. Given our prior knowledge, we should not treat everyone.

How does the utility will change when our probability of disease changes?

Medical Testing: With utility

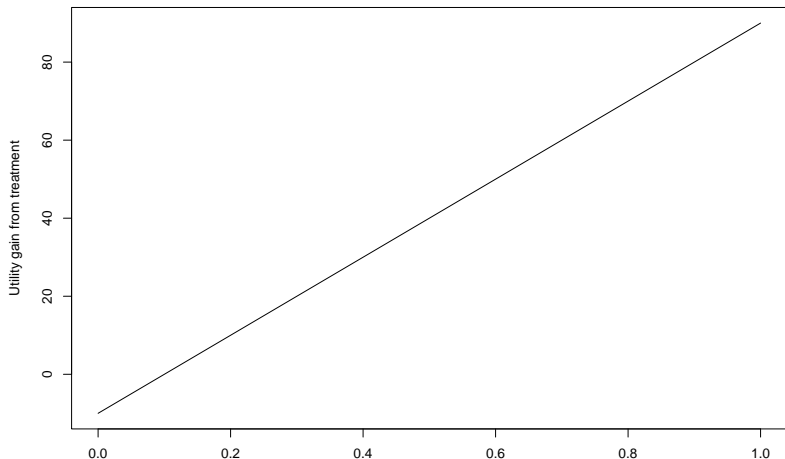
```
p = seq(0,1,0.01)
plot(p, 100*(1-p), type = "l", xlab = "p", ylab = "$E[U(a)]")
abline(h=90, col="red")
legend("bottomleft", legend = c("$E[U(a_N)]$", "$E[U(a_T)]$"))
```



Medical Testing: With utility

The gap of $0.9 - 100(1 - p)$ is the expected gain from treatment.

```
plot(p, 90-100*(1-p), type = "l", xlab = "p", ylab = "Util
```



Medical Testing: The value of test

We will need to calculate the posterior probabilities

```
#  $P(D \mid T = 0) = P(T = 0 \mid D) P(D) / P(T = 0)$   
pdt0 = 0.05*0.02/(0.05*0.02 + 0.99*0.98)  
print(pdt0)
```

```
[1] 0.001029654
```

```
# Expected utility given the test is negative  
#  $E[U(a_N \mid T=0)]$   
UN0 = pdt0*0 + (1-pdt0)*100  
print(UN0)
```

```
[1] 99.89703
```

```
#  $E[U(a_T \mid T=0)]$   
UT0 = pdt0*90 + (1-pdt0)*90  
print(UT0)
```

```
[1] 90
```

Medical Testing: The value of test

```
#  $P(D \mid T = 1) = P(T = 1 \mid D) P(D) / P(T = 1)$   
pdt = 0.95*0.02/(0.95*0.02 + 0.01*0.98)  
print(pdt)
```

```
[1] 0.6597222
```

```
#  $E[U(a_N \mid T=1)]$   
UN1 = pdt*0 + (1-pdt)*100  
print(UN1)
```

```
[1] 34.02778
```

```
#  $E[U(a_T \mid T=1)]$   
UT1 = pdt*90 + (1-pdt)*90  
print(UT1)
```

```
[1] 90
```

The best option is to treat now! Given the test our strategy is to treat if the test is positive and not treat if the test is negative.

Medical Testing: The value of test

Let's calculate the expected utility of this strategy.

```
#  $P(T=1) = P(T=1 \mid D) P(D) + P(T=1 \mid D=0) P(D=0)$   
pt = 0.95*0.02 + 0.01*0.98  
print(pt)
```

```
[1] 0.0288
```

```
#  $P(T=0) = P(T=0 \mid D) P(D) + P(T=0 \mid D=0) P(D=0)$   
pt0 = 0.05*0.02 + 0.99*0.98  
print(pt0)
```

```
[1] 0.9712
```

```
# Expected utility of the strategy  
pt*UT1 + pt0*UN0
```

```
[1] 99.612
```

The utility of our strategy of 100 is above of the strategy prior to testing (98), this difference of 2 is called the *value of information*.

Nash Equilibrium

- ▶ When multiple decision makers interact with each other, meaning the decision of one player changes the state of the “world” and thus affects the decision of another player - Nash equilibrium: a set of strategies where no player can improve their payoff by unilaterally changing their strategy, assuming others keep their strategies constant.
- ▶ No player has an incentive to deviate from their current strategy, given the strategies of the other players.

Nash Equilibrium

- ▶ Prisoner's Dilemma: Two prisoners must decide whether to cooperate with each other or defect. The Nash equilibrium is for both to defect, even though they would be better off if they both cooperated.
- ▶ Pricing Strategies: Firms in a market choose prices to maximize profits, taking into account their competitors' pricing decisions. The equilibrium is the set of prices where no firm can increase profits by changing its price unilaterally.
- ▶ Traffic Flow: Drivers choose routes to minimize travel time, based on their expectations of other drivers' choices. The equilibrium is the pattern of traffic flow where no driver can reduce their travel time by choosing a different route.

Marble Game

Two players A and B have both a red and a blue marble. They present one marble to each other. The payoff table is as follows:

- ▶ If both present red, A wins \$3.
- ▶ If both present blue, A wins \$1.
- ▶ If the colors do not match, B wins \$2

The *tit-for-tat* strategy, where you cooperate until your opponent defects. Then you match his last response.

Marble Game

Nash equilibrium will also allow us to study the concept of a *randomized strategy* (ie. picking a choice with a certain probability) which turns out to be optimal in many game theory problems.

First, assume that the players have a $\frac{1}{2}$ probability of playing Red or Blue. Thus each player has the same expected payoff $E(A) = \$1$

$$E(A) = \frac{1}{4} \cdot 3 + \frac{1}{4} \cdot 1 = 1$$

$$E(B) = \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 2 = 1$$

Marble Game

We might go one step further and look at the risk (and measured by a standard deviation) and calculate the variances of each players payout

$$\text{Var}(A) = (1 - 1)^2 \cdot \frac{1}{4} + (3 - 1)^2 \cdot \frac{1}{4} + (0 - 1)^2 \cdot \frac{1}{2} = 1.5$$

$$\text{Var}(B) = 1^2 \cdot \frac{1}{2} + (2 - 1)^2 \cdot \frac{1}{2} = 1$$

Therefore, under this scenario, if you are risk averse, player B position is favored.

Marble Game

The matrix of probabilities with equally likely choices is given by

A, B	Probability
$P(\text{red}, \text{red})$	$(1/2)(1/2)=1/4$
$P(\text{red}, \text{blue})$	$(1/2)(1/2)=1/4$
$P(\text{blue}, \text{red})$	$(1/2)(1/2)=1/4$
$P(\text{blue}, \text{blue})$	$(1/2)(1/2)=1/4$

Now there is no reason to assume ahead of time that the players will decide to play 50/50. We will show that there's a mixed strategy (randomized) that is a *Nash equilibrium* that is, both players won't deviate from the strategy.

Marble Game

We'll prove that the following equilibrium happens:

- ▶ A plays Red with probability $1/2$ and blue $1/2$
- ▶ B plays Red with probability $1/4$ and blue $3/4$

In this case the expected payoff to playing Red equals that of playing Blue for each player. We can simply calculate: A 's expected payoff is $3/4$ and B 's is \$1

$$E(A) = \frac{1}{8} \cdot 3 + \frac{3}{8} \cdot 1 = \frac{3}{4}$$

Moreover, $E(B) = 1$, thus $E(B) > E(A)$. We see that B is the favored position. It is simple that if I know that you are going to play this strategy and vice-versa, neither of us will deviate from this strategy – hence the Nash equilibrium concept.

Marble Game

Nash equilibrium probabilities are:

$p = P(A \text{ red}) = 1/2, p_1 = P(B \text{ red}) = 1/4$ with payout matrix

A, B	Probability
$P(\text{red}, \text{red})$	$(1/2)(1/4)=1/8$
$P(\text{red}, \text{blue})$	$(1/2)(3/4)=3/8$
$P(\text{blue}, \text{red})$	$(1/2)(1/4)=1/8$
$P(\text{blue}, \text{blue})$	$(1/2)(3/4)=3/8$

We have general payoff probabilities: $p = P(A \text{ red}), p_1 = P(B \text{ red})$

$$f_A(p, p_1) = 3pp_1 + (1 - p)(1 - p_1)$$

$$f_B(p, p_1) = 2\{p(1 - p_1) + (1 - p)p_1\}$$

Marble Game

To find the equilibrium point

$$\begin{aligned}(\partial/\partial p)f_A(p, p_1) &= 3p_1 - (1 - p_1) = 4p_1 - 1 \text{ so } p_1 = 1/4 \\ (\partial/\partial p_1)f_B(p, p_1) &= 2(1 - 2p) \text{ so } p = 1/2\end{aligned}$$

Much research has been directed to repeated games versus the one-shot game and is too large a topic to discuss further.

What are the drawbacks of the equilibrium analysis?