Föreläsning 5: Spelteori

Låt oss börja med ett motiverande exempel!

Exempel 1 (Medelvärdesspelet)**.** *Alla n spelare väljer ett heltal mellan 0 och 100. Sedan berkänas halva medelvärdet av alla gissningar $m = \frac{1}{2}\frac{1}{n}\sum_{i=1}^{n} x_i$ och den som kom närmast medelvärdet vinner! Om lika delas vinsten.*

Spelteori används för beslutsfattande. Ursprungligen studerades s.k. "zero-sum games" en perons förlust är någon annans vinst. Används mycket inom ekonomi, politik, psykologi och biologi. Relevanta frågor är ofta bästa strategin, då man har begränsad information. Poker.

Två stora namn inom spelteori är John von Neumann och John Nash (A Beautiful Mind).

Definitionen av spel är följande. Ett spel består av N spelare som var och en har en uppsättningar handlingar eller strategier i strategimängden s_i, att välja från. Strategirummet ges av mängden $S = s_1 \times s_2 \times \ldots \times s_N$. Till varje spelare associerar vi en vinstfunktion $\pi_i : S \to \mathbb{R}$, som beskriver vinsten för spelare i, som funktion av alla spelares strategier.

I medelvärdesspelet är $s_i = \{0, 1, 2, \ldots, 100\}$, och en vald strategi för spelare 1 är exempelvis $x_1 = 36$. Vinstfunktionen $\pi_1 = 1$ om x_1 var närmast medel $m = \frac{1}{n}\sum_{i=1}^{N} x_i$, och 0 annars.

Antag att vi känner till alla vinstfunktioner, finns det då någon uppsättning strategier som är bäst (det lönar sig ej att byta) för alla spelare? En sådan punkt i S är en sorts jämviktspunkt.

Definition 1 (Nash-jämvikt)**.** *Låt (S, π) vara en spel med n spelare, där s_i är strategimängden för spelare i, och $\pi(x) = (\pi_1(x), \pi_2(x), \ldots, \pi_n(x))$ är vinstfunktionen i punkten $x = (x_1, x_2, \ldots, x_n) \in S$, där x_i är vald strategi för spelare i. En strategiprofil $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ kallas för Nash-jämvikt om det för alla spelare i gäller att:*

$$\forall i, x_i \in s_i : \pi_i(x_1^*, \ldots, x_i^*, \ldots, x_n^*) \geq \pi_i(x_1^*, \ldots, x_i, \ldots, x_n^*)$$

Alltså för varje spelare i gäller att: det lönar sig ej (vinsten blir mindre) att byta strategi från x_i^* till någon annan strategi x_i. Detta brukar kallas rationellt, då man gör det som är bäst för sig själv.

Exempel 2 (Fångarnas dilemma)**.** *Antag att två medlemmar i ett kriminellt gäng är anhållna. Båda kommer förhöras av polis, men har ingen chans att kommunicera med* *varandra. Varje fånge har två val, att skvallra (defect) eller att vara tyst (cooperate). Om båda skvallrar får de ett 3 års fängesestraff, om båda är tysta får de 1 års straff, och om en skvallrar på den andra får den som skvallrar frisläppt och den som var tyst får 5 års fängelse. Vad ska varje fånge göra?*

Man kan beskriva detta i en matris

$$P = \begin{array}{cc} & \begin{array}{cc} C & \quad D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (3,3) & (0,5) \\ (5,0) & (1,1) \end{pmatrix} \end{array}$$

Där elementen i matrisen beskriver radspelaren (första elementet) och kolonnspelarens (andra elementet) vinst. Exempelvis beskriver $(0,5)$ *att om en spelare med strategi* C *möter en spelare med strategi* D *vinner* C *spelaren 0 och* D *spelaren 5.*

Pay-off matrisen ovan motsvarar förkortat straff. Kortast tänkbara straff får man om man skvallrar på sin kompanjon som inte skvallrar.

Vilka strategier är Nash-jämvikt? I fallet (C,C) *är det fördelaktigt att byta till* D, *och i* (C,D) *är det fördelaktigt att byta till* D. *Däremot är* (D,D) *Nash-jämvikt, men inte optimalvinst! Den rationella strategin är* D.

Exempel 3 (Modifierat fängelsedilemma)**.** *Undersök nu det modifierade spelet som har följande Payoff matris:*

$$P = \begin{array}{cc} & \begin{array}{cc} C & \quad D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (5,5) & (0,3) \\ (3,0) & (1,1) \end{pmatrix} \end{array}$$

Detta är mer intressnat då bästa strategin beror på vad moståndaren spelar (som man ju inte vet!).

Detta kan utvidgas till upprepat fängelsedilemma, där man spelar mot samma person mer än en gång. Hur kan man då välja strategi? (Efter 3 steg i spelet har man ju tillgång till motståndarens beslut de 3 första stegen. Hur kan man använda detta till sin fördel?)

Public goods game

Antag att invånarna (N) i en by går samman för att bygga en bro som sparar resväg. Deltagande i bygget är frivilligt men gynnar alla i byn. En strategi ges av $x_i \in [0,1]$

som beskriver insatsen av invånare i, och vinsten är då $\pi_i(x_i) = \alpha \frac{\sum_{j=1}^{N} x_j}{N} - x_i$. Här är $\alpha > 1$ en multiplikativ faktor som beskriver vinsten av samarbetet. Den största totala vinsten ges då alla ger maximal insats: $x_i = 1$ för alla i, och är då $\pi = \alpha - 1$. Antag att i denna situation ändrar spelare j sin strategi till $x_j < 1$, då fås individuell vinst $\pi_j \approx \alpha - x_j > \alpha - 1 = \pi_i$, alltså fås större vinst! Om detta sprider sig minskar x_i efter hand och det visar sig att $x_i = 0$ för alla spelare i är Nash-jämvikt för spelet.

Allmänningens Dilemma

Beskriver konflikten mellan individens och det allmännas bästa. Vi tänker oss en gemensam resurs (äng, fiskbestånd, etc.) som utnyttjas av ett samhälle och vars kvalitet minskar med användningen.

Vi tänker oss N spelare vars strategi motsvarar nyttjandegraden $x_i \in [0, 1]$. Nyttan för individ i ges av $\pi_i(x_1, \ldots, x_n) = x_i \left(1 - \sum_{j=1}^{N} x_j\right)$. Första termen är nyttjandegrad och andra termen motsvarar kvalitet (som minskar i tankt med totala nyttjandet).

Om det bara finns en individ kommer vinsten vara $\pi_1 = x_1(1 - x_1)$ som har maximum då $x_1 = \frac{1}{2}$ och ger vinsten $\pi_i\left(\frac{1}{2}\right) = \frac{1}{4}$, men om flera spelare är inblandade fungerar detta inte. Antag nu att $\sum_{j \neq i}^{N} x_j = T < 1$, hur ska då spelare x_i spela för att maximera sin avkastning? Genom att maximera sin vinst $\pi_i = x_i(1 - T - x_i)$, vars maximum ges av $x_i^* = \frac{1-T}{2}$. Om alla spelare spelar på detta sätt kan ingen enskild spelare öka sin avkastning.

$$x^* = \frac{1 - T}{2} = \frac{1 - \sum_{j \neq i} x^*}{2} = \frac{1 - (N - 1)x^*}{2}$$

så $x^* = \frac{1}{N+1}$ är Nash-jämvikt. När jämvikten råder är avkastningen

$$\pi_i = x_i \left(1 - \sum_{j=1}^{N} x_j\right) = \frac{1}{N + 1} \left(1 - \frac{N}{N + 1}\right) = \frac{1}{(1 + N)^2}$$

Total vinst för alla spelare ges då av

$$\pi = \sum_{j=1}^{N} \pi_j = \frac{N}{(N + 1)^2} \approx \frac{1}{N}$$

Är detta en tragedi? För en invidid var vinsten $\frac{1}{4} \gg \frac{1}{N}$ för stora N.

Kontinuerlig beskrivning (Evolutionary games)

Då man har en stor population spelare kan det vara mer lämpligt att beskriva andelen spelare x_i som spelar strategi i. Spelare har inte heller konstant fitness utan fitness beror ofta på frekvensen av strategier i populationen. Om man antar att strategibyte sker med en hastighet som är proportionell mot skillnaden mellan individuell fitness och populationens medelfitness fås ODE systemet:

$$\frac{dx_i}{dt} = x_i(\pi_i - \bar{\pi})$$

där $\bar{\pi} = \sum_{i=1}^{N} x_i \pi_i$ som är medelavkastningen i populationen. Detta innebär att förändringen är proportionell mot skillnaden mellan fitness hos individ i och populationens genomsnittliga fitness. Vi undersöker fängelsedilemmat som vi undersökte tidigare:

Exempel 4 (Fängelsedilemmat igen)**.** *Antag att vi har en stor population av individer som spelar fängelse-spelet. Nu beskriver vi istället* proportioner *spelare med strategi C och strategi D. Låt Pay-off matrisen till spelet ges av*

$$P = \begin{array}{c c} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (3,3) & (0,5) \\ (5,0) & (1,1) \end{pmatrix} \end{array}$$

och låt x beteckna proportionen spelare som använder strategin C (cooperate). Då vet vi att andelen som spelar D (defect) är $1 - x$, eftersom summan skall vara 1. Vinstfunktionen för respektive strategi är

$$\pi_C = 3x + 0(1-x) = 3x$$
$$\pi_D = 5x + 1(1-x) = 4x + 1$$

Vidare ges medelfitness av

$$\sum_{i=1}^{N} x_i \pi_i = x\pi_C + (1-x)\pi_D$$
$$= 3x^2 + (1-x)(4x+1) = 3x^2 + 4x + 1 - 4x^2 - x = -x^2 + 3x + 1$$

och vi har då att

$$\frac{dx}{dt} = x(3x + x^2 - 3x - 1) = x(x^2 - 1)$$

Vi kan undersöka stabilitet genom att söka noll-ställe till derivatan. Detta ger oss två punkter, $x = 0$ och $x = 1$. Stabilitet undersöks genom att undersökta hur en liten avvikelse förändras i tid. Undersöks jämvikterna kan man se att $x = 0$ är Nash-jämvikt (det lönar sig inte att byta).

I allmänna spel med två spelare

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

kan man ha följande fyra fall

- Domnians: $a > c$ och $b > d$
- Samexistens: $a < c$ och $b > d$
- Bistabilitet: $a > c$ och $b < d$
- Neutralitet: $a = c$ och $b = d$