

Casinomodellen

Begeleider: Lodewijk Kallenberg

Een speler gaat het casino in met een bepaald bedrag, zeg B euro's, en speelt daar herhaald hetzelfde spel. Bij dit spel heeft de speler een winstkans p (dan krijgt hij zijn inzet verdubbeld terug) en verlieskans $1 - p$ (dan verlist hij zijn inzet).

Het doel is om een van tevoren bepaald bedrag, zeg N , te bereiken. De speler stopt als hij het streefbedrag N heeft bereikt of als hij blut is.

De vraagstelling luidt: welke strategie kiest de speler, d.w.z. welke inzet a kiest hij als hij i euro's bezit met $a \in \{1, 2, \dots, i\}$, om de kans om het spel te eindigen met het streefbedrag N te maximaliseren.

Voor de oplossing van dit probleem moet het volgende worden gedaan:

1. Het probleem wordt gemodelleerd als een Markov beslissingsketen.¹
2. Toon aan dat dit een zogenaamd transiënt model is en stel de optimaliteitsvergelijking op.
3. Toon aan dat de optimaliteitsvergelijking een unieke oplossing heeft.
4. Bepaal de optimale strategie voor:
 - (a) $0 < p < \frac{1}{2}$;
 - (b) $p = \frac{1}{2}$;
 - (c) $\frac{1}{2} < p < 1$.
5. Voer een aantal simulaties uit.

¹Bij Markov beslissingsketens hebben we niet te maken met één Markov keten, maar met een aantal Markov ketens. Formeel is dit model als volgt:

- S is een eindige *toestandsruimte*, zeg $S = \{0, 1, 2, \dots, N\}$;
- in toestand $i \in S$ wordt een actie uit een eindige *actieverzameling* $A(i)$ gekozen;
- indien in toestand i actie $a \in A(i)$ wordt gekozen, dan gebeurt het volgende: er is een *directe opbrengst* $r_i(a)$ en met *overgangskans* $p_{ij}(a)$ is het systeem op het volgende beslissingstijdstip in toestand j .

Een (*deterministische*) strategie f is een beslisregel: het geeft aan welke actie in toestand i wordt gekozen; de actie die in toestand i wordt gekozen noteren we met $f(i)$. Iedere strategie genereert een Markov keten met overgangsmatrix $P(f)$, waarbij $[P(f)]_{ij} = p_{ij}(f(i))$ voor alle $i, j \in S$.