Riddler Solitair

Riddler is een puzzelcolumn in het befaamde blog *FiveThirtyEight* van Nate Silver. Elke week worden in deze column problemen gegeven die liggen op het terrein van de wiskunde, logica en kansrekening. Twee puzzels worden elke week gepresenteerd: de meer eenvoudige Riddler Express en de meer uitdagende Riddler Classic. Overmoedig geworden door het feit dat eerdere puzzels van mij – overigens ontleend aan mijn columns in *STAtOR* – waren verschenen in Riddler 538, stuurde ik onlangs een puzzel op die onverwacht een groot succes werd met verrassende oplossingen van adepten van Riddler 538. Dit wil ik de trouwe lezers van mijn *STAtOR*-column niet onthouden. De puzzel die verscheen als Riddler Classic luidde als volgt.

> Riddler solitaire is played with 11 cards: an ace, a two, a three, a four, a five, a six, a seven, an eight, a nine, a 10 and a joker. Each card is worth its face value in points, while the ace counts for 1 point. To play a game, you shuffle the cards so they are randomly ordered, and then turn them over one by one. You start with 0 points, and as you flip over each card your score increases by that card's points, as long as the joker hasn't shown up. The moment the joker appears, the game is over and your score is 0. The key is that you can stop any moment and walk away with a nonzero score.
>
> What strategy maximizes your expected number of points?
>
> Extra credit: With an optimal strategy, how many points would you earn on average in a game of Riddler solitaire?

De bepaling van de optimale stopregel is niet erg lastig, dit in tegenstelling tot de bepaling van de maximale gemiddelde score per spel. Veel inzenders kwamen op het idee om te bekijken wat het effect is op de huidige score zonder joker wanneer niet gestopt wordt maar nog één kaart wordt gepakt. De redenering gaat ruwweg als volgt. Stel dat je huidige score p punten is en dat je tot nu c kaarten hebt gepakt. In deze situatie zijn er nog $10-c$ niet-joker kaarten in het spel die een totale waarde van $55-p$ punten hebben, waarbij elk van deze kaarten een gemiddelde waarde van $(55-p)/(10-c)$ punten heeft. De volgende kaart die je pakt is dan de joker met kans $1/(11-c)$ en is een niet-joker kaart met kans $(10-c)/(11-c)$. Dus de verwachte toename van je huidige score van p punten is dan

$$T = \frac{10-c}{11-c} \times \frac{55-p}{10-c} = \frac{55-p}{11-c},$$

terwijl de verwachte afname van je huidige score gelijk is aan

$$A = \frac{p}{11-c}.$$

De verwachte toename T is alleen dan groter dan de verwachte afname A als $55-p>p$, oftewel als $p<27{,}5$. Dit leidt tot de conclusie om alleen door te gaan als je huidige score minder dan 28 punten is en te stoppen zodra je score 28 of meer punten is. Dit is inderdaad de optimale stopregel zoals met Markov beslissingstheorie voor optimale stopproblemen kan worden bewezen (de zogeheten *one-stage-look-ahead rule* is optimaal voor stopproblemen met de eigenschap dat het proces in de verzameling van ongunstige toestanden blijft als het daar eenmaal in beland is). Meer algemeen, voor het geval van N niet-joker kaarten met respectievelijke waarden 1 tot en met N geeft bovenstaande redenering dat het optimaal is om te stoppen zodra je score ¼ $N(N+1)$ of meer is en anders door te gaan. Opmerkelijk is dat de optimale strategie alleen afhangt van het aantal punten dat al vergaard is en niet van het aantal kaarten dat al getrokken is.

Veel inzenders van een oplossing voor de puzzel gebruikten Monte Carlo simulatie om de maximale gemiddelde score per spel te vinden en sommigen vonden daarbij het verrassende en opmerkelijke resultaat dat onder de optimale stopregel de kans om de joker te pakken gelijk is aan 50%. Eén inzender toonde aan dat deze kans inderdaad precies gelijk is aan 50% door op slimme wijze de computer alle mogelijke permutaties van de 11 kaarten te laten doorlopen en zo de kansverdeling van de eindscore onder de optimale stopregel te berekenen. Waar ik zelf dacht aan een dynamische programmering recursie met een multi-dimensionale toestand als alternatief voor simulatie, was er ook een inzending met een andere aanpak om zonder gebruik van simulatie de maximale gemiddelde score per spel te bepalen. Daartoe werd voor het geval van 11 kaarten eerst opgemerkt dat voor de optimale stopregel met een stopwaarde van 28 punten tenminste 4 kaarten en ten hoogste 7 kaarten nodig zijn, waarbij de eindscore niet meer dan 37 (=27+10) punten kan zijn. De verwachtingswaarde kan dan vervolgens berekend worden met de formule

$$\sum_{c=4}^{7} \sum_{p=28}^{37} p \, \frac{a_c(p)}{\binom{10}{c}} \, \frac{11-c}{11},$$

waarbij $a_c(p)$ wordt gegeven door het aantal combinaties van c verschillende niet-joker kaarten zodat de som van de waarden van de kaarten gelijk aan p is en minder dan 28 is als één van deze c kaarten verwijderd zou worden. De waarden van de $a_c(p)$ werden door aftelling bepaald. Dit leidde tot de verwachtingswaarde 15,453 voor de eindscore onder de optimale stopregel (de standaardeviatie van de eindscore is dan 15,547).

Ten slotte nog enkele opmerkingen over de uitdagende puzzel waarop zo veel interessante reacties binnenkwamen. Sommige inzenders pasten een combinatie van Monte Carlo simulatie en regressieanalyse toe. Zo werd voor het geval van 11 kaarten de kans op een eindscore nul als functie van het stoppunt s benaderd door de regressielijn

$$0{,}01587\,s + 0{,}0556$$

en werd voor het algemene geval van N niet-joker kaarten met respectievelijke waarden 1 tot en met N de maximale verwachtingswaarde van de eindscore benaderd door de regressiecurve

$$0{,}1251N^2 + 0{,}2879N + 0{,}061.$$

met de respectievelijke benaderingswaarden 0,49996 voor s=28 en 15,45 voor N=10.

De verrassende *trade-off* tussen de kans van 50% op een eindscore nul en de verwachtingswaarde van de eindscore onder de optimale stopregel lijkt ook van toepassing te zijn voor het algemene geval van N kaarten (voor N=2 en N=3 is de bewering simpel te verifiëren). Het zou interessant zijn dit verder te onderzoeken. Tot nu toe is verondersteld dat er één joker onder de kaarten is. Wat zijn de resultaten bij twee jokers? Het is dan simpel na te gaan dat het gemiddeld aantal punten per spel maximaal is door te stoppen zodra het aantal vergaarde punten 1/6 $N(N+1)$ of meer is als N het aantal niet joker kaarten is met waarden 1 tot en met N. Maar hoe zit het met de andere resultaten? Wat gebeurt als er onder de kaarten meerdere kaarten zijn met dezelfde waarde? Vragen te over. Leuke werkstukken voor studenten.

Henk Tijms is emeritus-hoogleraar operations research aan de Vrije Universiteit en auteur van diverse leerboeken over operations research en kansrekening.
E-mail: h.c.tijms@xs4all.nl