Als Oranje zich aan de wetten van het toeval houdt, dan haalt het de tweede ronde van het WK. **Hans van Maanen** introduceert de Grote Competitie-Simulator en laat zien hoe u daar de uitslagen van competities mee kunt voorspellen.

Hoe groot is de kans dat Nederland in de finale van het wereldkampioenschap voetbal komt?

Inleiding

Hoe groot is, om bij het begin te beginnen, de kans dat Nederland de eerste ronde overleeft? Alle kenners zijn het er over eens dat Nederland in poule C met onaangenaam sterke tegenstanders – Argentinië, Ivoorkust en Servië-Montenegro – te maken krijgt. De verschillen zijn, met andere woorden, minimaal. Zelfs als het Nederlandse team duidelijk het beste in zijn poule is, bestaat de kans dat het, door botte pech, toch derde of vierde in de poule wordt en naar huis kan. Als Nederland 'objectief' tweede is, wordt die kans zelfs akelig groot.

We zullen eens een poging doen het poulesysteem met wat kansrekening te benaderen. Iets dergelijks is al vaker gedaan voor toernooien zoals Wimbledon, waarbij steeds paarsgewijze winnaars worden aangewezen en een slechte loting al in de eerste ronde tot uitschakeling kan leiden, maar hoe zit het met de waarschijnlijkheden in een halve competitie, zoals die in de eerste ronde van het wereldkampioenschap wordt gespeeld? Hierbij speelt elk team in de poule eenmaal tegen elk ander team, in totaal dus drie wedstrijden. We gaan met behulp van de computer een Grote Competitie-Simulator bouwen.

Laten we beginnen met vier teams, A, B, C en D. Als A de beste is en D de slechtste, en het toeval speelt geen rol van betekenis, dan is het duidelijk dat de poule zal eindigen zoals in tabel 1.

Tabel 1: Afgetekende krachtsverhoudingen, de rol van het toeval is nul

	gespeeld	gewonnen	gelijk	verloren	punten
A	3	3	0	0	9
B	3	2	0	1	6
C	3	1	0	2	3
D	3	0	0	3	0

Dat kan gebeuren; bijvoorbeeld bij dit kampioenschap in de poules D en E, waar de winnaars bijna bij voorbaat duidelijk zijn. In poule C, met Nederland, zijn de krachtsverschillen kleiner en zal het toeval een grotere rol spelen. Naarmate de ploegen meer aan elkaar zijn gewaagd, worden toevallige factoren als blessures, scheidsrechterlijke dwalingen en weersgesteldheid belangrijker. Voor een eerste benadering kunnen we bijvoorbeeld heel eenvoudig aannemen dat A met 67 procent zekerheid van B zal winnen, met 22 procent kans gelijkspeelt, en met 11 procent kans zal verliezen. Voor B geldt hetzelfde ten opzichte van C, en zo ook C versus D. Dat klinkt allemaal niet eens zo gek.

Een dobbelsteen biedt een illustratie van de kansen. D is het slechtst, en kan alleen 1, 2 of 3 gooien. C is al wat beter, en gooit 2, 3 of 4. B komt altijd uit op 3, 4 of 5, en A is het sterkst en werpt altijd 4, 5 of 6. Je zou ook kunnen zeggen: D maakt gemiddeld 2 doelpunten per wedstrijd, A 5. Als rnd(n) een functie is die een willekeurig geheel getal tussen 1 en n geeft, dan wordt het verwachte aantal doelpunten voor team i gegenereerd met

$$r_i = \mathrm{rnd}(3) + 4 - i$$

Om het wat realistischer te maken, kan overal 1 vanaf worden getrokken, maar dat maakt verder niet uit. De spreiding voor alle teams is trouwens gelijk aan $\sqrt{\frac{2}{3}} = 0,82$. Als, om een eerste voorbeeld te nemen, A tegen C speelt, ontstaat de situatie van tabel 2.

Tabel 2: Mogelijke uitslagen wedstrijden tussen A en C

C\A	4	5	6
2	4 – 2	5 – 2	6 – 2
3	4 – 3	5 – 3	6 – 3
4	4 – 4	5 – 4	6 – 4

Van de negen mogelijke uitslagen is er slechts één een gelijkspel voor C: 4 - 4. C is dus vrijwel kansloos tegen A. Maar als A tegen B speelt, wordt het een stuk aardiger, zoals tabel 3 laat zien.

Tabel 3: Mogelijke uitslagen wedstrijden tussen A en B

B\A	4	5	6
3	4 – 3	5 – 4	6 – 3
4	4 – 4	5 – 4	6 – 4
5	4 – 5	5 – 5	6 – 5

Nu kan er van de negen wedstrijden een voor A verloren gaan met een doelpunt verschil, en zijn er twee kansen op een gelijkspel. Dat past dus inderdaad mooi op ons model: $\frac{6}{9}$ winst, $\frac{2}{9}$ gelijk, $\frac{1}{9}$ verlies.
D zal altijd van A verliezen; zelfs als D alle geluk van de wereld heeft, komt het team niet verder dan drie doelpunten, en zelfs als A alles tegenzit, komt het niet lager dan vier doelpunten.
Voor B tegen C, B tegen D en C tegen D zijn de tabellen precies gelijk, want we hebben aangenomen dat de verschillen tussen de vier teams even groot zijn. Dat is overigens niet noodzakelijk, zoals we later zullen zien. We kunnen de zaak zo gesofisticeerd maken als we willen.

Deze vier teams gaan nu een halve competitie spelen. Telkens wordt de uitslag door de gegeven formule bepaald. Een gewonnen wedstrijd telt voor drie punten, een gelijkspel voor een punt, en een verlies voor nul punten.
Hoeveel kans heeft A nu om kampioen van de poule te worden? En B?
Het is niet zo moeilijk om hiervoor een computerprogramma te schrijven en de teams een stuk of miljoen competities te laten spelen. Om de zaak wat te vereenvoudigen, hebben we aangenomen dat alle competities eenduidig beslist zijn, zonder gezeur over gelijke standen en doelsaldo's, maar ook daarin zou eenvoudig voorzien kunnen worden. Tabel 4 geeft het resultaat van de competitiesimulatie.

Tabel 4: de eindstand na een Monte-Carlosimulatie met d=1, s=0,82

	kans 1e plaats	kans 2e plaats	kans 3e plaats	kans 4e plaats
A	85,3	14,3	0,4	0,0
B	14,3	70,4	14,9	0,4
C	0,4	14,9	70,4	14,3
D	0,0	0,4	14,3	85,3

De kansen zien er alleszins behoorlijk uit voor team A. De kans dat zij doorgaan naar de volgende ronde is 99,6 procent. Voor team B is er ook niet zoveel aan de hand, al is er toch nog een kans van vijftien procent dat het naast de prijzen valt. Evenzo is, spiegelbeeldig, de kans voor C om toch door te gaan naar de volgende ronde 15 procent. De regelmaat is evident.

Voor ons model hebben we de kleinst mogelijke invloed van het toeval genomen. Als D alleen 1 of 2 gooit, C 2 of 3, B 3 of 4 en team A alleen 4 of 5, wordt A met 100 procent zekerheid eerste en B met 100 procent zekerheid tweede.
Wat gebeurt er als we de rol van het toeval laten toenemen? Daarnet was het systeem dat elk team een willekeurig getal tussen 1 en 3 krijgt, en A er 3 punten bij mocht tellen, B 2 punten, C 1 punt en D 0 punten. De bonus was telkens 1 punt, de spreiding kwam uit op 0,82.
Nu krijgen de teams een getal tussen 1 en 6, en tellen er volgens hetzelfde systeem bonuspunten bij. De spreiding is nu 1,7 – de rol van het toeval neemt toe en er is zelfs kans dat team D team A verslaat. Hoe gaat de eindstand er dan uitzien? Dat toont tabel 5.

Tabel 5: de eindstand na een Monte-Carlosimulatie met d=1, s=1,7

	kans 1e plaats	kans 2e plaats	kans 3e plaats	kans 4e plaats
A	64,0	26,3	8,2	1,5
B	26,3	42,4	23,1	8,2
C	8,2	23,1	42,4	26,4
D	1,5	8,2	26,3	64,0

Voor het sterkste team is er nog weinig aan de hand: 90 procent kans om door te gaan. De heren van team B moeten zich echter wel wat zorgen gaan maken: ook al zijn ze officieel de op een na sterksten, er is een kans van 31 procent dat ze achter het net vissen.
Als de uitslag wordt gesimuleerd door een willekeurig getal tussen 1 en 9 plus de bonus van telkens 1 punt (dan is de spreiding 2,6 punten), heeft A een kans van 82 procent om bij de eerste twee te eindigen, B nog maar 63 procent. Een op de acht keer (12,8 procent) zal C zelfs kampioen van de poule worden – aangezien er acht poules zijn, zou dat dit wereldkampioenschap met een waarschijnlijkheid van 67 procent gebeuren...

Wie het over een 'poule des verderfs' heeft, heeft het eigenlijk over een 'poule des geluks'. Hoe kleiner de onderlinge verschillen, hoe groter de rol van het toeval.
En dat is precies het probleem in poule C. Nederland en Argentinië zijn ongetwijfeld even sterk (in de zin dat in Nederland op Nederland gegokt wordt, in Argentinië op Argentinië; daar moet geld mee te verdienen zijn). Het zwakste team, Ivoorkust, moet echter toch wel in staat geacht worden deze teams te verslaan. Bij de meest conservatieve schatting, met de minste rol van het toeval, worden de kansen voor de teams om de volgende ronde te bereiken daarmee als in tabel 6.

Tabel 6: De eindstand na een Monte-Carlosimulatie met Nederland en Argentinië even sterk

	kans 1e plaats	kans 2e plaats	kans 3e plaats	kans 4e plaats
NED.	47,1	39,0	11,8	2,1
ARG.	47,1	39,0	11,8	2,1
SER.	5,1	16,8	49,4	28,7
IVO	0,8	5,2	27,0	67,0

Nederland en Argentinië hebben, samengevat, een kans van 86 procent om bovenin de poule te eindigen, en toch een niet verwaarloosbare kans van veertien procent om snel weggestuurd te worden. Ook bij een veel grotere rol van het toeval (met voor Ivoorkust een kans van 1 op 3 om ons te verslaan) blijven onze kansen volgens de simulatie boven de zeventig procent. Dat zit dus wel goed: Nederland haalt in ieder geval de volgende ronde. Het enige wat ze hoeven te doen, is zich aan de wetten van het toeval houden.

Overigens, nu we eenmaal zo ver zijn gevorderd, let niets ons de zaak nog verder te verfijnen. In plaats van een uniforme verdeling van de doelpunten, zouden we ook kunnen kiezen voor bijvoorbeeld een poisson-verdeling (doelpunten zijn helaas zeldzame gebeurtenissen). We kunnen ook de evident sterkste uit de poule bijvoorbeeld 6 in plaats van 3 punten extra geven op het random getal, en kijken wat er dan gebeurt. We kunnen zelfs het aantal ploegen uitbreiden tot achttien, en op grond van competitie-uitslagen uit het verleden of de mening van deskundigen gaan zoeken naar een meer empirische grondslag voor de Grote Competitie-Simulator van het betaalde voetbal. Daar moet geld mee te verdienen zijn.

Hans van Maanen
wetenschapsjournalist
Amsterdam

Verschenen:

Titel: *Elke positieve actie begint met critiek*
Auteur: Sacha la Bastide-van Gemert
Uitgever: Verloren, 2006
ISBN: 90-6550-912-7
Prijs: € 34,00