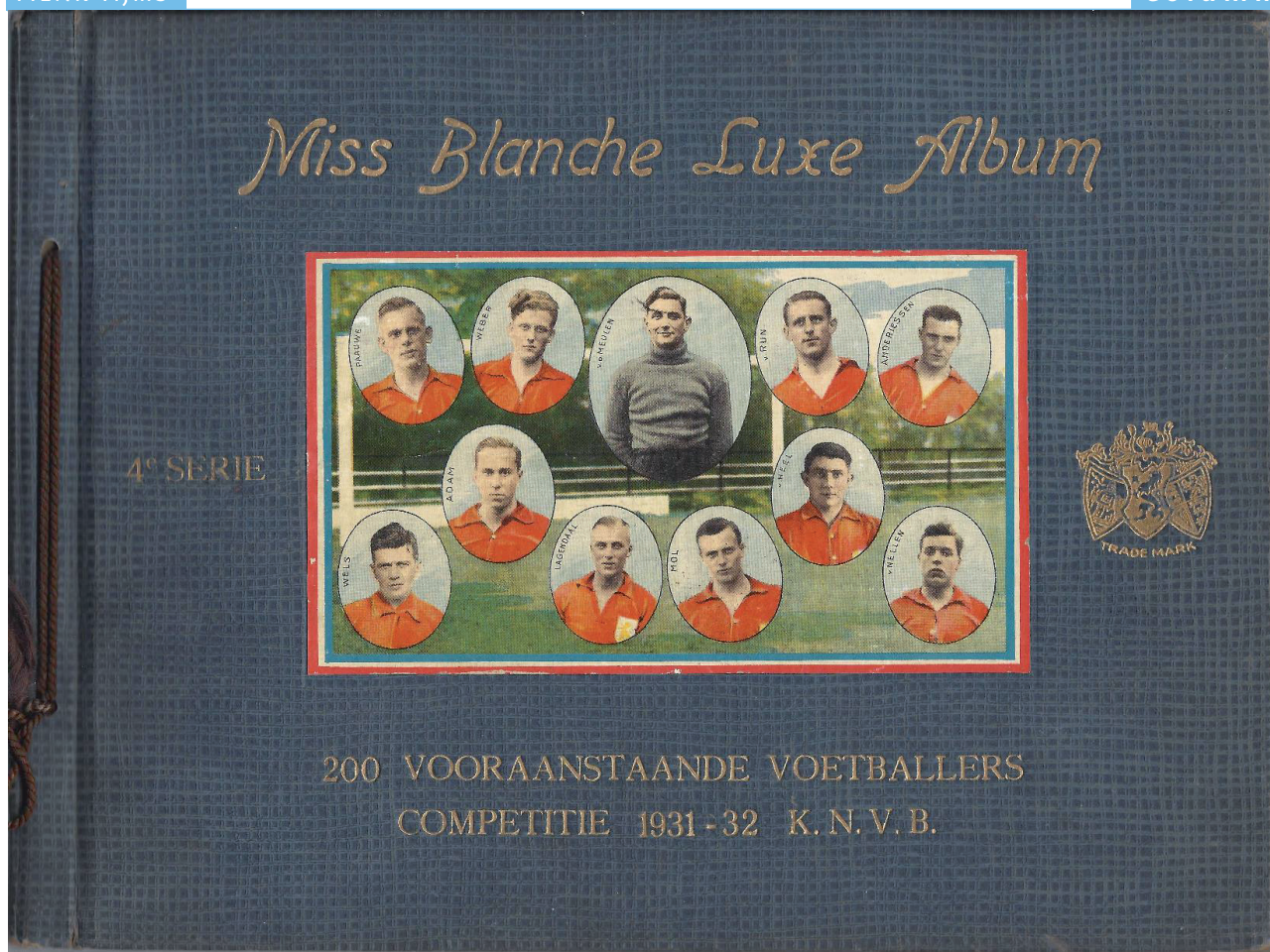




# VOETBALPLAATJES

## wie heeft ze niet verzameld?

Voor het wereldkampioenschap voetbal in Qatar hebben Albert Heijn, Jumbo en andere supermarktketens klaarblijkelijk afgezien van acties met plaatjes van de voetballers die acte de présence geven op dit besmette wereldkampioenschap voetbal in Qatar. Het verzamelen van voetbalplaatjes is wereldwijd een rage onder de jeugd. Menig lezer van *STATOR* zal zich vroeger ook aan deze verslavende activiteit overgegeven hebben. Wereldberoemd zijn de voetbalalbums van de gebroeders Guiseppe en Benito Panini. In 1961 brachten ze vanuit hun simpele krantenkiosk in de Italiaanse stad Modena hun eerste album uit met plaatjes van alle voetballers die in dat jaar actief waren in de Italiaanse voetbalcompetitie Serie A. Dit was het begin van een wereldwijd imperium dat de gebroeders Panini opbouwden en dat hen schatrijk maakte. Overigens waren de Panini-albums niet de eerste

in hun soort. In Nederland werden jaren dertig en veertig van de vorige eeuw al diverse albums met plaatjes van sporthelden uitgebracht door het legendarische sigarettenmerk Miss Blanche van de Vittoria Egyptian Cigarette Company uit Rotterdam, een sigarettenmerk dat ook de inspiratie vormde voor de psychologische roman *Miss Blanche* uit 2004 van Nelleke Noordervliet over een Rotterdamse tabakswinkelier die zijn leven ziet veranderen door een toevallige ontmoeting met een kettingrokende Turkse schone. De Miss-Blanche-albums zijn heden ten dage ware *collectors items*, zeker het album dat in 1932 werd uitgebracht met plaatjes van tweehonderd vooraanstaande voetballers uit de eerste en tweede klas van de Nederlandse voetbalcompetitie 1931-1932, waaronder de befaamde internationals Leo Halle, Puck van Heel en Bas Paauwe. Fabrikanten van margarinemerken als Leeu-

wenzegel en Planta ('pukkeltjes-affaire') deden in de jaren vijftig ook al pogingen om met plaatjes van bekende Nederlandse sporters de verkoop van hun producten te stimuleren. Nog steeds ben ik de trotse bezitter van het complete Leeuwenzegelalbum *Nederlandse Sport Succes-sen* uit 1953 met onder meer plaatjes van heroïsche sporthelden als Kees Broekman, Wim van Est en Piet Kraak.

Voor het Panini-album voor het Europees kampioenschap voetbal 2020 moesten 678 stickers worden verzameld. De stickers konden worden gekocht in pakketjes van zes verschillende stickers. Voor het geval dat je niet kunt ruilen met andere verzamelaars zijn interessante vragen: hoeveel pakketjes heb je nodig om met een kans van 50% of meer een complete collectie stickers te verkrijgen en hoeveel is het benodigde aantal stickers als je een succeskans van 90% of meer wilt hebben? Onder de aanname dat de stickers uniform verdeeld zijn over de pakketjes, geeft Monte Carlo simulatie met 100 000 runs als antwoord dat 780 respectievelijk 993 pakketjes nodig zijn. Een analytische aanpak is ook mogelijk met behulp van een absorberende Markovketen. De Markovketen heeft de 679 toestanden  $i = 0, 1, \dots, 678$ , waarbij toestand  $i$  betekent dat je tot nu toe  $i$  verschillende stickers verzameld heb. Toestand 0 is de begintoestand en toestand 678 is een absorberende toestand. De 1-staps overgangskansen in de Markovmatrix  $P$  zijn simpel en worden gegeven door hypergeometrische kansen. De kansverdeling van het aantal benodigde pakketjes voor een complete verzameling van stickers vind je door de matrix  $P$  herhaald met zichzelf te vermenigvuldigen: de kans dat je meer dan  $n$  pakketjes nodig hebt is 1 minus het  $(0, 678)$ de element van het  $n$ -voudig matrix product van  $P$  met zichzelf.

De berekeningen voor het *coupon collector's* probleem met stickers in dezelfde frequenties zijn veel simpeler als je bij aankoop maar één sticker krijgt. Dan kun je de kansverdeling van het aantal benodigde aankopen voor een complete verzameling van stickers berekenen met de bekende inclusie-exclusie formule uit de kansrekening (of met een absorberende Markovketen) en kunnen voor de verwachtingswaarde  $m_s$  en de variantie  $v_s$  van het aantal benodigde aankopen bij  $s$  verschillende stickers de expliciete formules

$$m_s = s \sum_{k=1}^s \frac{1}{k} \text{ en } v_s = s^2 \sum_{k=1}^s \frac{1}{k^2} - s \sum_{k=1}^s \frac{1}{k}$$

gegeven worden. Deze expliciete formules berusten op het feit dat het aantal aankopen nodig om van verschil-

lende stickers naar  $i+1$  verschillende stickers te gaan een geometrische kansverdeling met succeskans  $(s-i)/s$  heeft. Voor  $s$  voldoende groot kunnen  $\sum_{k=1}^s 1/k$  en  $\sum_{k=1}^s 1/k^2$  worden benaderd door  $\ln(s) + \gamma$  en  $\pi^2/6$ , waarbij  $\gamma = 0,5772\dots$  de Euler-Mascheroni constante is.

De aanname dat de stickers random verdeeld zijn over de pakketjes en dus met dezelfde frekwenties voorkomen, is cruciaal in bovenstaande berekeningen. Een empirische studie van Tim Paulden, 'Soccer fans in shiny sticker shock' in het Britse statistiek tijdschrift *Significance* Vol.18(2021)3, leidt tot grote twijfels of deze aanname vervuld is geweest bij het Panini-album voor het wereldkampioenschap voetbal in 2018. Verbazingwekkend is dit voor mij niet. Ook bij het album *Voetbalsterren* uit 1969 van de Utrechtse uitgever Vanderhout gaan hardnekkige geruchten dat de uitgever van sommige voetballers (waaronder Johan Cruyff) minder plaatjes heeft laten drukken om de verzamelaars hongerig te houden.

Als de frequentie waarmee een sticker voorkomt niet voor elke sticker hetzelfde is, dan wordt de situatie veel ingewikkelder. Voor het geval dat de stickers alleen per stuk verkocht worden, dan kunnen de verwachtingswaarde en de standaardafwijking van het aantal benodigde aankopen voor een complete verzameling nog wel analytisch berekend worden. Dit gaat met numerieke integratie van twee oneigenlijke integralen die met de 'Poissonization' truc van A.N. Kolmogorov verkregen worden. Zo vind je bijvoorbeeld de waarden 61,217 en 35,976 voor de verwachtingswaarde en de standaardafwijking van het aantal benodigde worpen met twee zuivere dobbelstenen totdat elk van de elf mogelijke uitkomsten 2,...,12, voor de som verschenen is. Een approximatie voor de kans dat meer dan  $n$  worpen nodig zijn om deze elf uitkomsten te verkrijgen is 1 minus het product van de kansen  $1 - e^{-np_k}$  over  $k=2, \dots, 12$ , waarbij  $p_k$  de kans is op de som  $k$  in één worp. Deze simpele approximatie is algemeen toepasbaar voor het *coupon collector's* probleem met ongelijke kansen, ook voor het probleem met meerdere coupons per aankoop zoals bij het Panini-album.

HENK TIJMS is emeritus-hoogleraar operations research aan de Vrije Universiteit en auteur van diverse leerboeken over operations research en kansrekening, waaronder het boek *Basic probability: What every math student should know* (World Scientific Press, 2021, 2e druk).  
E-mail: h.c.tijms@xs4all.nl