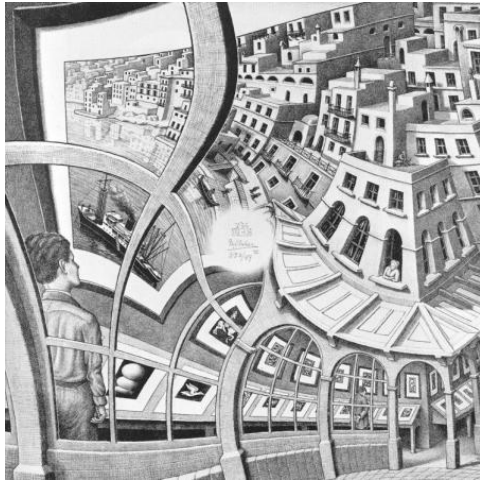


Hoe moeten we gaten in gegevensbestanden opvullen? Ik liep tegen dit probleem aan bij het analyseren van de aantallen partijen die tijdens open toernooien door dames gespeeld worden. In dit type bestanden ontbreken vrijwel altijd partijen omdat er door dames byes opgenomen worden of omdat ze zich op een gegeven moment uit het toernooi terugtrekken. In deze column onderzoek ik wat we kunnen doen om dit type gaten in een gegevensbestand op te vullen.

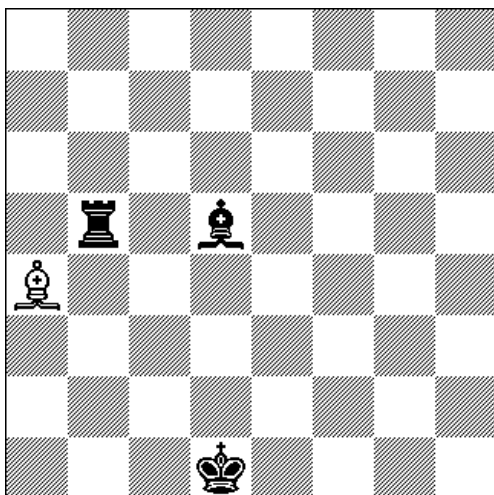


Figuur 1. Prentententoonstelling, Escher, 1956.

Een variatie op het probleem dat ik schetste is de litho de 'Prentententoonstelling' van M.C. Escher uit 1956, zie figuur 1. Over deze litho schreef Escher: "Er komt onder andere een jonge man op voor die, belangstellend staat te kijken naar een prent aan de wand van een tentoonstelling, waar hij zelf op getekend staat." In het midden van deze litho bevindt zich een wit gat en de vraag die men zich kan stellen is waarom Escher dit gat niet opgevuld heeft. Het antwoord op die vraag is dat Escher simpelweg vastliep.

Figuur 2. Print Gallery, Lenstra en de Smit, 2003.

Het waren Hendrik Lenstra en Bart de Smit van de Universiteit Leiden die er in 2002 in slaagden om de wiskunde achter deze litho te ontmaskeren en met behulp van een computeralgoritme het gat op te vullen, zie figuur 2. Tijdens de lezingencyclus van Studium Generale van de Technische Universiteit Delft legde Bart de Smit eind november 2004 aan een zaal bomvol belangstellenden uit hoe zij een en ander voor elkaar gekregen hadden. Op de website Escher en het Droste effect zijn animatie filmpjes te vinden die een en ander op bijzonder aardige wijze uit de doeken doen.



Figuur 3. Smullyan: 'Waar staat de witte koning?'

Een totaal andere variatie op het 'ontbrekende partijen van dames' probleem is afkomstig van de logicus Raymond Smullyan en betreft een retrograde probleem, zie figuur 3. Smullyan vertelt dat de witte koning, Haroun al Rashid, zich onzichtbaar gemaakt heeft en de vraag die Smullyan ons stelt is 'Waar staat de witte koning?'. Zo op het eerste, tweede en derde gezicht is dit een onoplosbaar probleem. Mocht u het na een vierde poging op willen geven dan kunt u op pagina 123 van Smullyan's boek The Chess Mystery of the Arabian Knights (1981) de oplossing vinden.

We zien dat we met wiskunde en logisch nadenken dit soort problemen te lijf kunnen gaan. Dat geldt ook voor het probleem van de niet door dames tijdens een open toernooi gespeelde partijen. In tabel 1 geef ik relevante data van het Isle of Man Open 2018. Aan dit toernooi deden $K=28$ dames mee die tijdens $n=9$ rondes $n \cdot K=252$ partijen hadden moeten spelen. Van deze 252 partijen werden er 13 partijen niet gespeeld. Het totale aantal deelnemers aan dit open toernooi was $N=165$. Met de hypergeometrische distributie, klik [hier](#) voor een nadere toelichting, heb ik de kansen $p(k)$ op het treffen van k dames, met k een getal tussen de 0 en 9, berekend.

n (aantal rondes) = 9			N (aantal deelnemers) = 165				K (aantal dames) = 28		
k	p(k) HypGeo- Distr.	Verwacht =N*p(k)	Feitelijk F(k)	aantal partijen =k*F(k)	Evenredig E(k)	aantal partijen =k*E(k)	"Laagste χ^2 " X(k) Chi-Sq test	aantal partijen =k*X(k)	χ^2
0	0.179	29.545	34	0	31	0	33	0	0.40
1	0.350	57.716	65	65	63	63	58	58	0.00
2	0.291	47.948	39	78	41	82	43	86	0.51
3	0.135	22.205	16	48	18	54	20	60	0.22
4	0.038	6.308	8	32	8	32	8	32	0.45
5	0.007	1.138	2	10	3	15	2	10	2.32
6	0.001	0.130	1	6	1	6	1	6	
7	0.000	0.009	0	0	0	0	0	0	
8	0.000	0.000	0	0	0	0	0	0	
9	0.000	0.000	0	0	0	0	0	0	
	1.000	165	165	239	165	252	165	252	3.91
door dames niet gespeelde partijen				13		0		0	
Totaal aantal partijen (n*K)				252		252		252	

Tabel 1. Isle of Man Open 2018: Het toevoegen van de niet door dames gespeelde partijen.

Ik heb de 13 ontbrekende partijen op twee manieren aan de tabel toegevoegd. Het toevoegen van één niet gespeelde partij heeft tot gevolg dat een van de schakers een niveau hoger terecht komt, dus van niveau k naar niveau $k+1$ opschuift. De eerste methode betreft het evenredig verdelen van de ontbrekende partijen hetgeen neerkomt op het verlagen resp. toevoegen van [3, 5, 3, 1, 1] partijen aan de k -niveaus van 0 tot en met 4 respectievelijk 1 tot en met 5. De tweede methode behelst het streven naar de "laagste χ^2 " (chi-kwadraat) waarde. Laagste omdat ik deze waarde met de hand berekend heb en er wellicht een nog lagere χ^2 waarde bestaat.

Welke van deze twee methodes geniet de voorkeur? Bestaan er nog andere methoden? Wie kon ik om raad vragen? Ik stuurde een email naar clubgenoot Ton de Waal en vroeg hem naar zijn mening. Ton is een wiskundige die in 2003 aan de Erasmus Universiteit Rotterdam gepromoveerd is op het onderwerp Processing of Erroneous and Unsafe Data. Verwerking van foutieve en onveilige data. Hij moest meer licht op dit probleem kunnen werpen.

Ton's reactie op mijn vraag: "Hoe ik het zelf zou doen, hangt af van de data die je ter beschikking hebt. Als je van iedereen per ronde weet tegen wie hij of zij heeft gespeeld (of die ronde niet heeft gespeeld), kan je natuurlijk veel meer dan wanneer je alleen maar jouw tabel hebt. Zelf houd ik me vooral bezig met imputatiemethoden. Bij imputatiemethoden probeer je individuele gevallen te schatten, in dit geval iets als "heeft een specifieke persoon in een specifieke ronde tegen een dame gespeeld". Op individueel niveau gaat dat natuurlijk fout. Op geaggregeerd niveau gaat het vaak wel redelijk goed. Als je het meerdere keer doet, dan kan je de variatie tussen de uitkomsten gebruiken om iets over de nauwkeurigheid van de schattingen te zeggen. Als je alleen jouw tabel hebt, dan zou ik er diep over na moeten denken hoe zo'n imputatiemethode er dan uit zou kunnen zien want ik zou op de een of andere manier,

bijvoorbeeld, wel mee willen nemen dat er één persoon is die zes keer tegen een dame heeft gespeeld en twee personen die vijf keer tegen een dame hebben gespeeld. Kortom, het is een moeilijk probleem en ik kan alleen maar blaffen.”

Blaffen? De volgende twee citaten van Vonnegut en Zappa zijn afkomstig uit Ton's proefschrift.

Vonnegut: *“Tiger got to hunt, Bird got to fly; Man got to sit and wonder: ‘Why, why, why?’. Tiger got to sleep, Bird got to land; Man got to tell himself he understand.”* Uit: Cat's Cradle (1963).

Zappa: *“Evelyn, a dog, having undergone, Further modification, Pondered the significance of short-person behaviour, In pedal-depressed panchromatic resonance, And other highly ambient domains... Arf, she said.”* Uit: Evelyn, A Modified Dog (1975).

Ton: “Kurt Vonnegut en Frank Zappa zijn mijn grote helden. Zij hebben zelfs mijn kijk op het leven in het algemeen, en het schaakspel in het bijzonder, beïnvloed. Vooral de gedachte dat het leven (en zeker het schaakspel) onbegrijpelijk is en dat iedere poging er iets van te begrijpen hopeloos is. Je kunt alleen maar doen alsof je het begrijpt en doorgaan met leven (het citaat van Vonnegut uit mijn proefschrift) of blaffen als een hond als je met een moeilijk probleem wordt geconfronteerd (het citaat van Zappa). Mijn schaakanalyses zijn slechts geblaf na afloop van een partij. Zolang er niet wordt geblunderd, kan ik er weinig zinnigs over zeggen.”

Volgens de componist en schaker Sergei Prokofiev was Johann Sebastian Bach de Emanuel Lasker van de muziek. Als ik deze vergelijking doortrek zou ik Frank Zappa de Gyula Breyer van de muziek willen noemen. De hypermoderne Breyer ontketende een schaakrevolutie en Zappa ontketende met zijn Mothers of Invention een muziekrevolutie. Ton de Waal schreef mij dat Frank Zappa hem aan Jan Timman doet denken. Zappa's muziek is eclectisch en technisch goed uitgevoerd en dit geldt ook voor Timman's schaken. Zappa fuseerde allerlei muziekstijlen – klassiek, blues, jazz, rock, pop – en Timman leverde bijdragen aan diverse schaakgebieden – wedstrijdschaak, eindspelcompositie, schaakjournalistiek. Speciaal voor Jan drie nummers van Frank: A Little Green Rosetta en Jewish Princess en DC Boogie.

Het bekendste boek van Kurt Vonnegut is 'Slaughterhouse-Five' (1969). Dit boek, over de waanzin van oorlog, is gebaseerd op zijn eigen ervaringen in de tweede wereldoorlog. Zo overleeft Vonnegut in 1945 de vuurstorm die Dresden vernietigt door drie verdiepingen onder de grond in een koelcel van het abattoir waar hij als krijgsgevangene tewerkgesteld is te schuilen. Duizenden sterven en Vonnegut leeft. Zijn reactie: “So it goes.”

Bijzonder is het schaakverhaal dat Andrew Leonard in Playing Chess with Kurt Vonnegut (2007) vertelt. Andrew is twaalf jaar oud als Vonnegut hem uitdaagt voor een partijtje schaak.

Figuur 4. Vonnegut Schaak

Andrew Leonard: “Een bijzonder leuke ervaring was Vonnegut tijdens Thanksgiving op zijn hurken op de vloer te zien zitten voor een partijtje schaak. In een opwelling stelde hij voor om de stukken op een andere manier op te stellen. Waarom moeten de pionnen voorop, als offerlammeren die op het punt staan in het abattoir van de frontlinies te worden vermorzeld, deze machteloze leenmannen van de machtigen? Laten we de feodale heren uit hun vossenholen dwingen en voorop laten gaan in het tumult! Hij stelde voor de pionnen op de achterste rij te plaatsen en de ridders, bisschoppen, koningen en koninginnen op de voorste rij!”

