

Hoofdstuk 2

Een schatting voor flesjes

IONICA SMEETS

Heineken verkoopt zijn bier in terug te geven flesjes. Het blijkt verrassend lastig om te schatten hoeveel van die flesjes er nog ergens bij een klant staan. De studiegroep oppert een nieuw model én een andere manier van steekproeven nemen.

Een Delftse studentenvereniging wilde eens als grap het hok van hun nieuwe bestuur compleet volbouwen met lege bierflesjes. De ingenieurs hadden natuurlijk zó uitgerekend hoeveel flesjes ze nodig hadden en ze sloegen aan het verzamelen. Met een busje reden ze langs lokale supermarkten om lege flesjes te kopen tegen vergoeding van het statiegeld. Hoewel dit niet geheel volgens de regels was, vonden de supermarkten het wel een mooie grap en ze werkte allemaal graag mee. Behalve dan één manager, die geen zin had om mee te doen aan dit soort flauwe grappen. Natuurlijk kwamen de studenten een week later precies bij deze manager de duizenden lege flesjes terug brengen.

Deze grap had lang niet overal ter wereld gewerkt. Wij kennen een statiegeldsysteem, maar in veel Afrikaanse landen is er een vol-voor-leeg-systeem. Daar ruilen klanten hun lege flesjes in voor volle. De enige manier om aan bier te komen is daar dus om al flesjes te hebben. Daardoor houden mensen een veel grotere voorraad lege flesjes achter de hand dan hier in Nederland. Als een Afrikaanse student voor een feestje extra bier wilt kopen, dan moet hij ergens al die flesjes hebben staan.

Heerlijk helder hergebruik. De vraag die Heineken naar de studiegroep bracht gaat precies over die voorraad lege flesjes. Als Heineken een nieuw ontwerp flesje op de markt brengt, dan moet het bedrijf schatten hoeveel flesjes er in totaal op de markt zijn. Al die lege flesjes moeten dan immers omgeruild worden. In het verleden bleken er soms veel meer flesje op te duiken dan verwacht.



Figuur 2.1: Altijd fijn om een goede voorraad in huis te hebben.

Wat is een betere manier om te schatten hoeveel flesjes er in de verschillende markten zijn? Als Ashish Aggarwal deze vraag op maandag presenteert bij de studiegroep, vliegt er onmiddellijk een hand omhoog in de zaal: “Kunt u niet gewoon tellen hoeveel flesjes u verkoopt en daarvan aftrekken hoeveel flesjes u terugkrijgt?” Aggarwal legt uit dat dit niet zo eenvoudig is: een deel van de flesjes breekt en zal nooit meer ingeleverd worden. Daarnaast worden in Afrika lege bierflesjes op allerlei andere plekken gebruikt: voor opslag en transport van petroleum en soms zelfs door andere bierbrouwers. Overigens merkt hij op dat de vraag niet alleen geldt voor flesjes, maar ook voor kratten en fusten. Maar voor de eenvoudigheid zullen we het steeds over flesjes hebben.

Hoe lang is een flesje weg? Heineken vermoedt dat de schatting voor het aantal flesjes in de markt af zal hangen van de omlooptijd van een flesje. Wat is een efficiënte manier om met een steekproef die omlooptijd te bepalen? Hierbij is het probleem lastiger doordat er niet in elk seizoen evenveel bier verkocht wordt en er een verschil is tussen de omlooptijden van bijvoorbeeld café of supermarkten. Het meten van de omlooptijden is op zichzelf niet zo moeilijk: op elke fles staat een houdbaarheidsdatum, waarmee te schatten is wanneer het flesje op de markt kwam. Door die houdbaarheidsdatum te noteren als een flesje terugkomt, is er een redelijke schatting te maken van hoe lang dit flesje rondzwierf.

Joris Bierkens, onderzoeker aan de Radboud Universiteit Nijmegen, koos voor Team Heineken omdat hij het probleem interessant vond. De beschikbare gegevens over de omlooptijd bleken niet helemaal perfect: “Ze waren ten eerste geanonimiseerd omdat de echte gegevens vertrouwelijk zijn. Dus werkten we aan landen als Beer Republic

en Ciderland. Daarnaast waren ze iets onhandig verzameld: Soms telde een krat als één datapunt, soms telden alle vierentwintig flessen in een krat als datapunten. Ons advies aan Heineken is dan ook om voortaan één fles per krat te gebruiken bij een steekproef.”

Daarnaast zijn de steekproeven steeds in dezelfde maand gehouden - net na het hoogseizoen, terwijl de wiskundigen liever zien dat er elk jaar een andere maand wordt gebruikt. Dit is dan ook gelijk aanbeveling twee. Toch is uit de beschikbare gegevens al te zien hoe groot de verschillen in omloop zijn. Zoals Ashish Aggarwal maandag in de openingspresentatie aangaf: “De gemiddelde terugkeertijd is dertig dagen, maar sommige flesjes komen pas na zes jaar terug.”

Niet elke fles is hetzelfde. De wiskundigen maken verschillende modellen voor het aantal de bierflesjes in de markt. Ze beginnen met het onderscheiden van verschillende soorten flesjes die er op tijdstip t in de markt zijn:

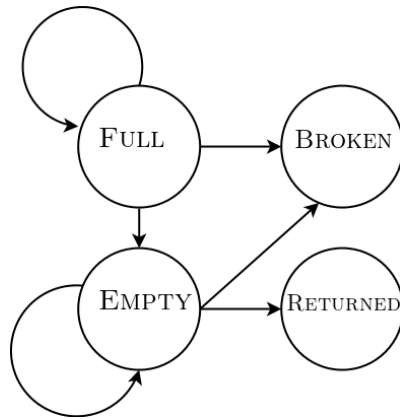
- $R(t)$ zijn de flesjes die op korte termijn Retour komen,
- $S(t)$ zijn de flesjes die Slapen. Ze staan ergens opgeborgen, of worden voor iets anders gebruikt en komen pas terug als er een nieuwe fles op de markt komt,
- $B(t)$ zijn de flesje die gebroken zijn. Ze zijn kapot of permanent ergens anders in gebruik en zullen nooit terugkomen.

Het totaal aantal flesjes in de markt op een bepaald tijdstip is de som van deze drie variabelen: $R(t) + S(t) + B(t)$. Heineken is vooral geïnteresseerd in een schatting voor het aantal slapende flesjes $S(t)$.

De hele week piekeren de wiskundigen hoe je het aantal slapende en gebroken flessen kunt schatten. Ze nemen aan dat een vast percentage van de flesjes gebroken wordt, maar hoe weet je hoe groot dit percentage is? Bierkens beschrijft dat de doorbraak op donderdag kwam met het eenvoudige, slimme idee om aan te nemen dat het aantal slapende flesjes niet meer verandert als de markt stabiel is: “Daarmee is te berekenen hoeveel flesjes er gebroken zijn.” Als er namelijk geen flesjes meer verdwijnen naar de slaapstand, is het verschil tussen wat je terugkrijgt en wat je verkoopt het aantal gebroken flesjes. Met deze gegevens komen ze op een percentage van ongeveer 2% gebroken flesjes per omloopcyclus. Dat betekent dat er pakweg bij elke tweede krat een flesje sneuvelt.

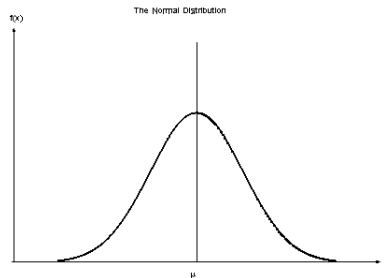
De levencyclus van één flesje Het eerste model dat de wiskundigen maken is vooral geschikt voor landen met een statiegeldsysteem, hierin beschrijven de ze cyclus van één flesje.

Uit dit eenvoudige model ontwikkelt het team een model voor de stroom van flesjes. Hierbij nemen ze aan aan dat de flesjes zich onafhankelijk van elkaar gedragen. In dit model is elk van de mogelijke overgangen bij benadering normaal verdeeld. Het



Figuur 2.2: Een vol flesje kan vol blijven, breken of leeggedronken worden. Een leeg flesje kan blijven staan (slapen), breken of teruggebracht worden.

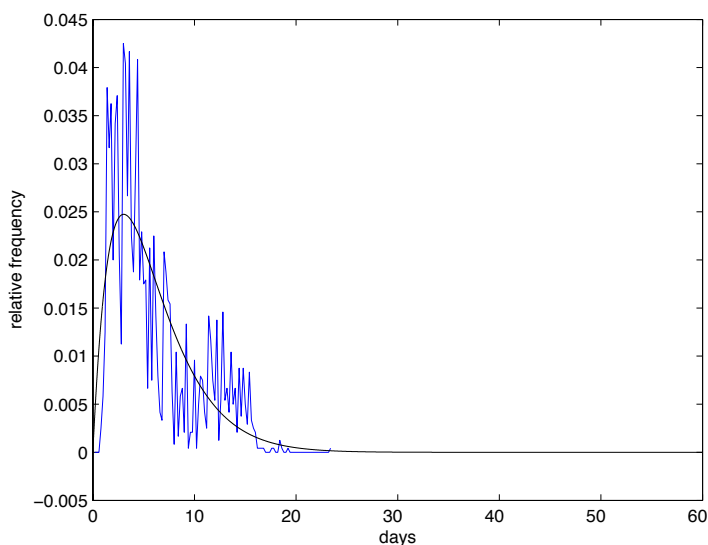
aantal volle flesjes dat er per tijdseenheid breekt heeft dan een kansverdeling van de vorm als hieronder.



Figuur 2.3: Deze normale verdeling komt op heel veel plekken voor: bijvoorbeeld bij lichaamslengte van volwassen mannen of meetfouten. De meeste waarden zitten in de buurt van het gemiddelde (μ) en de uitzonderingen allebei de kanten op worden steeds zeldzamer.

Uit dit model leiden de wiskundigen een stelsel differentiaalvergelijkingen af. Hieruit kunnen ze concluderen dat het voor het model niet uitmaakt hoeveel flesjes er aan het begin in de markt zijn. Wel is het belangrijk om te weten hoeveel flesjes er gemiddeld in het verleden de markt in gingen, waarbij vooral het recente verleden belangrijk is. Het aantal flesjes dat er terugkomt zal een vaste fractie zijn van het aantal lege flesjes.

Verborgen variabelen. De wiskundigen leggen daarna uit hoe met filters en de stelling van Bayes informatie over verborgen variabelen tevoorschijn kan worden getoerd. Om hun model compleet te maken, beschrijven ze hoe je de nodige parameters kunt schatten. De tijd voordat een flesje terugkomt, is de tijd die het nodig heeft om leeg gedronken te raken, plus de tijd die nodig is voor een lege flesje wordt ingeleverd. De studiegroep kiest voor deze tijd een hypoexponentiele verdeling, zie de figuur hieronder. Soms levert deze schatting een grafiek die niet bij de gegevens past. Dat nadeel blijkt een voordeel: als dit gebeurt weet je dat je beter een ander model kunt gebruiken. De studiegroep geeft een een lijst van alternatieve modellen en verwijzingen naar literatuur daarover.



Figuur 2.4: De gegevens van Heineken over hoe vaak flesjes na een bepaalde tijd terugkomen met de hypoexponentiele verdeling die het beste bij deze gegevens past. Hier lijkt het model een keurige beschrijving te geven van de meetdata.

Tenslotte eindigt het rapport met aanbevelingen voor de steekproeven die de omlooptijd van flesjes meten. De wiskundigen geven aan hoe je kunt berekenen hoeveel flesjes er nodig zijn om een bepaalde betrouwbaarheid te halen in de meting. Ze adviseren Heineken om bij de meting twee testen te doen om helder te krijgen wat er wel en niet moet worden meegenomen in een toekomstig rekenmodel. De eerste test moet uitwijzen of de seizoenen invloed hebben op de omlooptijd. Als mensen meer bier drinken in de zomer, komen de flesjes dan sneller terug? Of juist niet? Dit is met een relatief eenvoudige meting te testen en met dezelfde techniek kan Heineken ook andere effecten meten, bijvoorbeeld lange-termijn ontwikkelingen in de verkoop. Een

tweede test zou zijn om te kijken of verschillende distributiekkanalen een ander soort omlooptijd hebben. Misschien brengen cafés wel altijd eerst de nieuwste flesjes terug, omdat de oudere flesjes achterin hun opslagruimte staan. De extra testen zullen zorgen voor nóg betere modellen.

Team Heineken: *Joris Bierkens, Markus Heydenreich, Sindo Nunez Queija, Patrick van Meurs, Floske Spijksma, Jan Tuitman*