

Fouten

Overal waar tekst verwerkt wordt, worden fouten gemaakt. Het gaat dan om slecht lopende zinnen, spelfouten, typefouten, enzovoort. Het is vrijwel onmogelijk om er voor te zorgen dat in lange teksten zoals een boek, krant of tijdschrift geen enkele fout voorkomt. Een manier om het aantal fouten te verminderen is de tekst te laten controleren door een zogenoemde **screeener**. Een screener is iemand die fouten in een tekst probeert te ontdekken. Ook screeners kunnen fouten over het hoofd zien. In deze opgave bekijken we een model waarin fouten in teksten met een bepaalde kans worden ontdekt.

Screeener Chris heeft al vaak teksten gecontroleerd. De kans dat een fout in de tekst door hem wordt ontdekt is gelijk aan 0,8. Deze kans geldt voor elke fout in de tekst. We noteren deze kans als p_C , dus: $p_C = 0,8$.

Veronderstel dat er in een tekst 52 fouten zitten.

- 4p **16** Bereken de kans dat screener Chris minstens 40 van de 52 fouten ontdekt.

De screeners Chris en Dieuwke controleren volkomen onafhankelijk van elkaar een tekst waarin 375 fouten zitten. Dieuwke heeft iets minder ervaring dan Chris. De kans p_D dat een fout in de tekst door haar wordt ontdekt is 0,72, dus voor Dieuwke geldt $p_D = 0,72$. Voor Chris geldt weer $p_C = 0,8$.

- 3p **17** Bereken hoeveel fouten naar verwachting zowel door Chris als ook door Dieuwke zullen worden gevonden.

Een document bevat 64 fouten. Omdat het een belangrijk document is maakt men gebruik van vier goede screeners, die volkomen onafhankelijk van elkaar de tekst controleren. De kans dat een van deze screeners een fout over het hoofd ziet, is voor iedere screener voor elke fout gelijk aan 0,15.

- 4p **18** Bereken de kans dat alle 64 fouten worden ontdekt.

Bij het screenen van een tekst is vooraf natuurlijk niet bekend hoeveel fouten er in voorkomen. Ook is niet altijd bekend hoe groot de kans is dat een fout door een screener wordt ontdekt. De wiskundige Polya vroeg zich af of je na screening van een tekst kunt weten hoeveel fouten er naar verwachting nog in de tekst zitten. Polya bedacht dat je daar achter kunt komen door twee screeners in te schakelen. Hij nam daarbij aan dat de screeners onafhankelijk van elkaar een fout wel of niet ontdekken.

Elke screener maakt een lijst van de door hem ontdekte fouten. Daaruit kan een lijst worden gemaakt van de gemeenschappelijk gevonden fouten. Dat zijn de fouten die de screeners allebei hebben ontdekt. Niet-ontdekte fouten zijn fouten die door geen van beide screeners zijn ontdekt.

Polya heeft hiervoor de volgende formule opgesteld:

$$\text{het verwachte aantal niet-ontdekte fouten} = \frac{(N_A - N_G) \cdot (N_B - N_G)}{N_G}$$

In deze formule is:

N_A = het aantal fouten dat door screener A is gevonden,

N_B = het aantal fouten dat door screener B is gevonden

en

N_G = het aantal gemeenschappelijk gevonden fouten.

We bekijken nu de volgende situatie: in een tekst heeft screener A een aantal fouten gevonden. Screener B heeft alle fouten gevonden die A al gevonden had, plus nog een aantal extra fouten.

- 3p **19** Leg uit dat in deze situatie het verwachte aantal niet-ontdekte fouten volgens de formule nul is.

De screeners Frits en Laura controleren een tekst. Frits vindt 90 fouten en Laura vindt 88 fouten. Het aantal gemeenschappelijk gevonden fouten is 66. Nu kan met behulp van de formule van Polya het verwachte aantal fouten in de tekst berekend worden.

- 4p **20** Bereken op grond van de resultaten van de screening het totaal aantal fouten dat naar verwachting in de tekst zit.