8 Poisson verdeling

Een andere handige discrete kansverdeling is de Poisson verdeling. De Poisson stochastische variabele geeft het aantal gevallen van gebeurtenissen weer, die hier ook weer aantal keren success noemen. Het grote verschil met de binomiale verdeling is dat er bij de Poisson verdeling geen sprake is van een vast aantal pogingen. Bij de Poisson verdeling gaat het om het aantal keren success in een bepaald tijdsinterval of specifieke plaats. Een voorbeeld hiervan is het aantal, dagelijkse ongelukken op een specifiek stuk snelweg. Het Poisson experiment heeft de volgende eigenschappen:

1. het aantal keren success dat plaatsvindt in elk willekeurig interval is onafhankelijk van het aantal keren success in elk ander willekeurig interval
2. de kans op success in een interval is even groot in intervallen van dezelfde grootte
3. de kans op success is evenredig aan de grootte van het interval
4. de kans op meer dan 1 success in een interval ligt dichter bij 0 naarmate het interval kleiner wordt

Als algemene regel geldt dat een Poisson stochastische variabele staat voor het aantal keren dat een relatief zeldzame gebeurtenis die willekeurig en onafhankelijk plaatsvindt. Bijvoorbeeld het aantal mensen dat bij een restaurant binnenkomen is geen Poisson verdeling, want mensen komen meestal in groepen naar een restaurant en dat is in strijd met de eerste eigenschap van het Poisson experiment, te weten het ontbreken van onafhankelijkheid.

De formule voor de Poisson kansverdeling:

Als X een Poisson stochastische variabele is, dan is de kans dat deze een waarde aanneemt van x:

$$P(X = x) = p(x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

waar μ het gemiddelde aantal keren success is in het bepaalde interval of de specifieke plaats en e is de basis van het natuurlijke algoritme (ongeveer 2,71828).

Voorbeeld 8[a]

Robbert is eigenaar van een bedrijf dat websites maakt en hij merkt dat er af en toe spelfouten op zijn websites voorkomen, ondanks dat deze vaak worden nagelezen voordat ze online worden gezet. Na wat analyseren komt hij er achter dat er elke 100 webpagina's (geen websites!) 1,5 spelfout zit. Hij selecteert willekeurig 100 webpagina's en berekent de kans op 0 spelfouten in deze 100 pagina's.

$$P(X=0)=\frac{e^{-\mu}\mu^{x}}{x!}=\frac{e^{-1,5}1,5^{0}}{0!}=\frac{(2,71828)^{-1,5}(1)}{1}=0,2231$$

Wat nou als dit voorbeeld iets interessanter wordt gemaakt?

Voorbeeld 8[b]

Robbert besluit in plaats van 100 pagina's 400 pagina's te nemen. Aangezien het hier om 400 pagina's gaat geldt het gemiddelde van 1,5 fout niet meer. Immers, de 1,5 fout gold per 100 pagina's. Dat houdt in dat in 400 (4x100) pagina's er gemiddeld 6 (4x1,5) fout zit. Als men dus wil weten hoe groot de kans is op 0 fouten in 400 pagina's, berekent men:

$$P(X=0)=\frac{e^{-\mu}\mu^{x}}{x!}=\frac{e^{-6}6^{0}}{0!}=\frac{(2.71828)^{-6}(1)}{1}=0,002479$$

Als Robbert nu de kans wil berekenen op 5 of minder fouten in 400 pagina's dan wordt dit als volgt gedaan:

$$P(X\leq 5)=p(0)+p(1)+p(2)+p(3)+p(4)+p(5)$$

Als men de formule nu verder invult krijgt men $p(0) = 0,002479$, $p(1) = 0,01487$, $p(2) = 0,04462$, $p(3) = 0,08924$, $p(4) = 0,1339$ en $p(5) = 0,1606$. Als men deze bij elkaar optelt krijgt men:

$$P(X\leq 5)=0,4457$$