

## 7 Binomiaal verdelingen

Binomiale verdelingen zijn het resultaat van zogenaamde binomiale experimenten. Binomiale experimenten hebben de volgende eigenschappen:

1. Een binomiaal experiment bestaat uit een vast aantal pogingen. Het aantal pogingen wordt genoteerd als  $n$ .
2. Elke keer zijn er 2 mogelijke uitkomsten. De ene uitkomst noemt men success, dan andere uitkomst failure.
3. De kans op success is  $p$ . De kans op failure is  $1-p$ .
4. De pogingen zijn onafhankelijk wat inhoudt dat de uitkomst van een keer geen invloed heeft op andere pogingen.

De stochastische variabele staat voor het aantal keren success in  $n$  pogingen. Deze variabele noemt men ook wel de binomiale stochastische variabele.

Een kort voorbeeld:

Quirine gooit 10 keer met een 1 euromunt. Het aantal pogingen  $n$  is dus 10. Quirine wil zo vaak mogelijk "kop" gooien, dus ze beschouwt "kop" gooien als success. Als er sprake is van een zuivere munt, dan is de kans op "kop" 50%, dus  $p = 0,5$ . Ook kan er geconstateerd worden dat de pogingen onafhankelijk zijn, daar een worp met de munt geen invloed heeft op een vorige of volgende worp.

### 7.1 Binomiale stochastische variabele

Zoals al eerder vermeld is de binomiale stochastische variabele het aantal keren success in  $n$  pogingen van het experiment. Deze variabele kan waarden aannemen als  $0, 1, 2, \dots, n$ , kortom de stochastische variabele is discreet. Men moet in staat kunnen zijn te berekenen welke kans gerelateerd is aan elke waarde.

Als men een kansboom maakt met bij iedere fase twee takken (uitkomsten) die staan voor success en failure, is het duidelijk dat om de kans te berekenen op  $x$  keer success met  $n$  pogingen er voor elke success in de reeks vermenigvuldigd dient te worden met  $p$ . Als er  $x$  keren success zijn dan zijn er dus ook  $n-x$  keren failure. Elke failure in de reeks dient vermenigvuldigd te worden met  $1-p$ .

Kortom, de kans voor elke reeks takken (in de kansboom) die staan voor  $x$  keer succes en  $n-x$  keer failure is

$$p^x(1-p)^{n-x}$$

In een kansboom kunnen combinaties zich op verschillende manieren voordoen. Met een munt is bijvoorbeeld de combinatie "kop en munt" mogelijk als "kop-munt" of "munt-kop". Met de onderstaande formule kan men het aantal takkenreeksen berekenen met  $x$  keer success en  $n-x$  keer failure:

$$C_x^n = \frac{n!}{x!(n-x)!}$$

Als de twee bovenstaande formules gecombineerd worden, kan men het volgende concluderen:

De kans op  $x$  keer success in een binominaal experiment met  $n$  pogingen en met kans op success =  $p$  is

$$P(X = x) = p(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

#### Voorbeeld 7<sup>a</sup>

Karel gaat zonder te leren naar een multiple choice tentamen. Er zijn 10 vragen en elke vraag heeft 5 mogelijke antwoorden. Dus  $n = 10$  en  $p = 0,2$ .

De kans op niet een goed antwoord is dan

$$P(X = 0) = \frac{10!}{0!(10-0)!} (0,2)^0 (1-0,2)^{10-0}$$

$$P(X = 0) = 0,1074$$

De kans op bijvoorbeeld 2 goede antwoorden is dan

$$P(X = 2) = \frac{10!}{2!(10-2)!} (0,2)^2 (1-0,2)^{10-2}$$

$$P(X = 2) = 0,3020$$

### 7.2 Cumulatieve kans

Het kan voorkomen dat men op zoek is naar de kans op meerdere waarden tegelijk, in plaats van de kans op geen of 2 goede antwoorden, zoals in voorbeeld 7<sup>a</sup>. Het volgende voorbeeld geeft zo'n situatie weer:

#### Voorbeeld 7<sup>b</sup>

Karel, uit voorbeeld 7<sup>a</sup>, weet dat hij 50% van het tentamen goed moet hebben wil hij het tentamen halen. Alles onder de 50% is dus ongewenst. Bij 0 tot 4 goede antwoorden zal het tentamen niet halen. Uit het vorige voorbeeld weet men:

$$P(X = 0) = 0,1074 \text{ en } P(X = 2) = 0,3020$$

Wat men wil berekenen is  $P(X \leq 4)$  wat niets anders is dan de volgende som:

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

$$P(X \leq 4) = 0,1074 + 0,2684 + 0,3020 + 0,2013 + 0,0881 = 0,9672$$

Met andere woorden, de kans is 96,72% dat Karel z'n tentamen niet haalt als hij zonder te leren alles gokt.

### 7.3 Gemiddelde en variantie van een binomiale verdeling

Statistici hebben voor het gemiddelde, de variantie en de standaarddeviatie van de binomiale stochastische variabele standaard formules ontwikkeld:

$$\mu = np$$

$$\sigma^2 = np(1 - p)$$

$$\sigma = \sqrt{np(1 - p)}$$

*Voorbeeld 7<sup>c</sup>*

Wat zou het gemiddelde cijfer en de standaarddeviatie zijn in een klas vol met Karels (dus leerlingen die hun tentamen niet leren)?

$$\mu = np = 10(0,2) = 2$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{10(0,2)(1-0,2)} = 1,26$$