

Samenvatting Statistiek

Bert De Saffel

2017-2018

Inhoudsopgave

I	Theorie	3
1	Kansrekenen	4
1.1	Bewerkingen	4
1.2	Rekenregels	4
1.3	Combinatieleer	5
1.4	Regel van Bayes	5
2	Beschrijvende statistiek	7
2.1	Populaties en steekproeven	7
2.2	Veranderlijken	7
2.3	Discrete veranderlijken	8
2.3.1	Ordenen van gegevens	8
2.3.2	Plaatsparameters en spreidingsparameters	8
2.4	Continue veranderlijken	9
2.4.1	Ordenen van gegevens	9
2.4.2	Plaatsparameters en spreidingsparameters	9
2.4.3	Momenten van een steekproef	10
3	Verdelingsfuncties van een populatie	11
3.1	Kansfunctie	11
3.2	Dichtheidsfunctie	11
3.3	Plaatsparameters	13
3.4	Spreidingsparameters	15
3.5	Momenten	16
3.5.1	De momentenfunctie	17
3.6	De ongelijkheid van Chebychev	20
4	Discrete verdelingen	21
4.1	Uniforme discrete verdeling	21
4.2	Bernoulli experiment	22
4.3	Binomiale Verdeling	22
4.3.1	Momentenfunctie	23
4.3.2	Rekursieformule	23
4.4	Geometrische verdeling	23
4.5	Hypergeometrische verdeling	23
4.6	Poisson verdeling	25

5	Continue verdelingen	26
5.1	Uniforme continue verdeling	26
5.2	Exponentiële verdeling	26
5.3	Normale verdeling	26
5.3.1	Eigenschap	26
5.4	Lineaire combinatie	27
II	Oefeningen	28
6	Kansrekenen	29
7	Verdelingsfunctie van een populatie	36
8	Discrete verdelingen	41
9	Continue verdelingen	45
10	Herhalingsoefeningen	48

Deel I

Theorie

Hoofdstuk 1

Kansrekenen

De kans op een gebeurtenis A:

$$P(A) = \frac{\# \text{ gunstige gevallen}}{\text{tot } \# \text{ mogelijkheden}}$$

Altijd geldig:

$$0 \leq P(A) \leq 1$$

1.1 Bewerkingen

1. $P(A \cap B)$: kans op gebeurtenis A en gebeurtenis B
2. $P(A \cup B)$: kans op gebeurtenis A of gebeurtenis B
3. $P(A|B)$: Voorwaardelijke kans (lees: Wat is de kans op A indien B waar is). A en B zijn onafhankelijk indien $P(A|B) = P(A)$ of $P(B|A) = P(B)$

1.2 Rekenregels

1. Optellingswet: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
2. Vermenigvuldigingswet: $P(A \cap B) = P(A)P(B)$
3. Complementswet: $P(\overline{A}) = 1 - P(A)$
4. $P(A \cap \overline{B}) = P(A) - P(A \cap B)$

Voorbeeld: Wat is $P(A \cup B \cup C)$?

Uitwerking

$$\begin{aligned} &P(A \cup B \cup C) \\ &= P(A \cup B) + P(C) - P[(A \cup B) \cap C] \\ &= P(A) + P(B) - P(A \cap B) + P(C) - P[(A \cap B) \cup (B \cap C)] \\ &= P(A) + P(B) + P(C) - P(A \cap B) - [P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)] \\ &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \end{aligned}$$

1.3 Combinatieleer

De permutatie(volgorde is van belang):

$$P_n = n!$$

(Op hoeveel verschillende manieren kunnen we 5 studenten plaatsen op 5 stoelen **OF** het aantal verschillende manieren om n elementen te ordenen)

De combinatie(volgorde is niet van belang):

$$C_n^p = \binom{n}{p} = \frac{n!}{p!(n-p)!}$$

(Op hoeveel verschillende manieren kunnen we 2 studenten kiezen uit 5 **OF** het aantal verschillende manieren om p elementen te kiezen uit n .)

Voorbeeld: In een vaas zitten 8 rode, 4 witte en 3 blauwe knikkers. Wat is de kans om bij trekken met teruglegging exact 2 rode knikkers te nemen bij het trekken van 5 knikkers?

Uitwerking

$$P(2R) = \frac{8}{15} \cdot \frac{8}{15} \cdot \frac{7}{15} \cdot \frac{7}{15} \cdot \frac{7}{15} \cdot C_5^2$$

Het getal $\frac{8}{15}$ stelt de kans voor om een rode knikker te trekken. Het getal $\frac{7}{15}$ stelt de kans voor om geen rode knikker te trekken. Er wordt vermenigvuldigt met C_5^2 aangezien de 2 rode knikkers elk van de 5 plaatsen kunnen innemen.

$$\begin{aligned} &= C_5^2 \left(\frac{8}{15}\right)^2 \left(\frac{7}{15}\right)^3 \\ &= 10 \cdot \frac{64}{225} \cdot \frac{343}{3375} = 0.29 = 29\% \text{ kans} \end{aligned}$$

1.4 Regel van Bayes

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^n P(A_i)P(B|A_i)}$$

Voorbeeld: In een jaszak zitten 2 munten: één normaal(N) en één vervalst(V) (munt aan elke zijde). Een muntstuk wordt aselekt gekozen en gegooit. Munt komt bovenaan te liggen. Wat is de kans dat dit het normale muntstuk is? Dit muntstuk wordt opnieuw gegooit. Terug munt. Wat is nu de kans dat dit het normale muntstuk is?

Uitwerking

$$\begin{aligned} P(N) &= \frac{1}{2} \\ P(V) &= \frac{1}{2} \end{aligned}$$

De kans dat het munt of kop is bij het normale muntstuk is 0.5.

$$P(M/N) = P(K/N) = \frac{1}{2}$$

De kans dat het munt is bij het valse muntstuk is 1.

$$P(M/V) = 1$$

Regel van Bayes:

$$\begin{aligned} P(N/M) &= \frac{P(N)P(M/N)}{P(M/N)P(N) + P(M/V)P(V)} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2}} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3} \end{aligned}$$

Indien muntstuk opnieuw wordt gegooid:

$$\begin{aligned} P(N/2M) &= \frac{P(N)P(2M/N)}{P(2M/N)P(N) + P(2M/V)P(V)} \\ &= \frac{\frac{1}{2} \cdot (\frac{1}{2} \cdot \frac{1}{2})}{(\frac{1}{2} \cdot \frac{1}{2}) \cdot \frac{1}{2} + 1 \cdot \frac{1}{2}} \\ &= \frac{\frac{1}{4}}{\frac{5}{4}} = \frac{1}{5} \end{aligned}$$

Hoofdstuk 2

Beschrijvende statistiek

2.1 Populaties en steekproeven

- **Populatie:** Een verzameling van elementen
- **Steekproef:** Een deelverzameling van de populatie. Deze moet representatief zijn (bv 3/4 van de populatie is vrouw. In de steekproef moet ook 3/4 vrouw zijn)

2.2 Veranderlijken

- **Stochastische variabele:** Kan waarden aannemen elk met een bepaalde kans.
- **Kwalitatieve variabele:** Geen numerieke waarden (bloedgroep, studierichting)
- **Kwantitatieve variabele:** Numerieke waarden (gewicht, leeftijd)
- **Ordinale variabele:** Er kan een bepaalde orde toegekend worden (zeer goed, goed, matig, voldoende, slecht en zeer slecht)
- **Nominale variabele:** Er kan geen bepaalde orde toegekend worden (studierichting)
- **Onafhankelijke variabele:** De ene meting beïnvloedt de andere niet (gooien van een dobbelsteen)
- **Afhankelijke variabele:** De ene meting beïnvloedt de andere (grootte per geslacht)
- **Discrete variabele:** Heeft enkel discrete waarden (1, 0.4, 7, 9, 1/100)
- **Continue variabele:** Heeft continue waarden ($\sqrt{2}$, $[-10, 10]$)

2.3 Discrete veranderlijken

2.3.1 Ordenen van gegevens

Stel n het aantal elementen van de populatie of van een steekproef en k het aantal verschillende waarden dat de veranderlijke kan aannemen.

$$\sum_{i=1}^k n_i = n$$

- **Absolute frequentie:** Het aantal keer dat een bepaalde waarde aangenomen wordt.
- **Relatieve frequentie:** De verhouding van de absolute frequentie tot het aantal elementen

$$f_i = \frac{n_i}{n} \quad \text{en} \quad \sum_{i=1}^k f_i = 1$$

- **Cumulatieve absolute frequentie:**

$$N_i = \sum_{j=1}^i n_j$$

- **Cumulatieve relatieve frequentie:**

$$F_i = \sum_{j=1}^i f_j$$

- **Frequentietabel:** Voorstelling van alle frequenties.

Een toets bij 204 studenten, gequoteerd op 10, leverde volgende resultaten:

score	2	3	4	5	6	7	8	9	10
n_i	1	5	16	21	39	44	50	26	2

De frequentietabel wordt:

score	2	3	4	5	6	7	8	9	10
n_i	1	5	16	21	39	44	50	26	2
f_i	0.005	0.025	0.078	0.103	0.191	0.216	0.245	0.127	0.01
N_i	1	6	22	43	82	126	176	202	204
F_i	0.005	0.029	0.108	0.211	0.402	0.618	0.863	0.990	1

2.3.2 Plaatsparameters en spreidingsparameters

- **Steekproefgemiddelde:** De som van alle waarnemingen gedeeld door het totaal aantal waarnemingen

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i$$

- **Modus:** De waarde met de grootste absolute frequentie.

- **Mediaan:** Na ordening van de waarnemingen, de middelste waarde als n oneven is en het gemiddelde van de middelste twee als n even is.
- **Variantie:** Het rekenkundig gemiddelde van de **kwadraten van de afwijkingen** van de waarnemingsgetallen ten op zichte van hun rekenkundig gemiddelde:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

- **Standaardafwijking:** De vierkantswortel uit de steekproefvariantie.

$$\sigma = \sqrt{s^2} = s$$

2.4 Continue veranderlijken

2.4.1 Ordenen van gegevens

Het ordenen van continue veranderlijken wordt gedaan met klassen.

- **Spreidingsbreedte (range):** Het verschil tussen het grootste en het kleinste waarnemingsgetal.
- **Aantal klassen:** Het aantal klassen is de vierkantswortel uit het aantal elementen:

$$k = \sqrt{n}$$

- **Klassebreedte (Δx):** De breedte van elke klasse. Indien elke klasse een verschillende breedte heeft wordt Δx_i gebruikt om de i -de klasse aan te duiden.
- De frequenties zijn analoog aan discrete veranderlijken

2.4.2 Plaatsparameters en spreidingsparameters

- **Het rekenkundig gemiddelde:**

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k n_i c_i$$

met c_i het klassemidden van de i -de klasse.

- **De modale klasse:** De klasse met de grootste absolute frequentie.
- **De mediale klasse:** De klasse met de laagste waarden waarin de cumulatieve relatieve frequentie minstens 0.5 bedraagt
- **De variantie:**

$$s^2 = \frac{1}{n-1} \sum_{i=1}^k n_i (c_i - \bar{x})^2$$

- **Standaardafwijking:**

$$\sigma = \sqrt{s^2} = s$$

2.4.3 Momenten van een steekproef

_TODO: epic

Hoofdstuk 3

Verdelingsfuncties van een populatie

Twee soorten verdelingsfuncties:

- Kansfunctie: bij discrete kansen (sommeren)
- Dichtheidsfunctie: bij continue kansen (integreren)

3.1 Kansfunctie

De kansfunctie $f(x)$:

$$\begin{aligned} f(x_i) &= P(X = x_i) && \text{met} && i = 1, 2, \dots, k \\ f(x) &= 0 && \text{als} && k \notin \{x_1, \dots, x_n\} \end{aligned}$$

De som van alle kansen is gelijk aan 1:

$$\sum_{i=1}^k f(x_i) = 1$$

De cumulatieve distributiefunctie wordt gegeven door:

$$F(x) = P(x \leq t) = \sum_{x_i \leq t} f(x_i) dx$$

3.2 Dichtheidsfunctie

De som van alle kansen is gelijk aan 1:

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

Aangezien we integreren:

$$\int_a^a f(x) dx = 0$$

Hieruit volgt:

$$\begin{aligned} P(a \leq x \leq b) \\ &= P(a < x \leq b) \\ &= P(a < x < b) \\ &= P(a \leq x < b) \end{aligned}$$

De cumulatieve distributiefunctie wordt gegeven door:

$$F(x) = P(x \leq t) = \int_{-\infty}^t f(x) dx$$

Voorbeeld: Stel $f(x) = \begin{cases} a + b(x-1)^2 & 0 \leq x \leq 2 \\ 0 & \text{anders} \end{cases}$.

Bepaal a en b zodat f een geldige dichtheidsfunctie is. Bepaal ook $F(t)$.

Uitwerking

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= 1 \Leftrightarrow \int_0^2 a + b(x-1)^3 dx = 1 \\ &\Leftrightarrow ax + b \frac{(x-1)^4}{4} \Big|_0^2 = 1 \\ &\Leftrightarrow 2a + \frac{b}{4}(1-1) = 1 \\ &\Leftrightarrow a = \frac{1}{2} \end{aligned}$$

Is $f(x) \geq 0 \forall x$?

Is $\frac{1}{2} + b(x-1)^3 \geq 0$? ($x \in [0, 2]$)

$$\Rightarrow x \in [0, 2] \Rightarrow -1 \leq (x-1)^3 \leq 1$$

Als $b > 0$:

$$\begin{aligned} -b &\leq b(x-1)^3 \leq b \\ \frac{1}{2} - b &\leq b(x-1)^3 + \frac{1}{2} \leq b + \frac{1}{2} \end{aligned}$$

$f(x)$ zal ≥ 0 indien $\frac{1}{2} - b \geq 0$

$$\Rightarrow b \leq \frac{1}{2}$$

Als $b < 0$:

$$\begin{aligned} -b &\geq b(x-1)^3 \geq b \Rightarrow b \leq b(x-1)^3 \leq -b \\ \Rightarrow b + \frac{1}{2} &\leq b(x-1)^3 + \frac{1}{2} \leq -b + \frac{1}{2} \end{aligned}$$

$f(x)$ zal ≥ 0 indien $\frac{1}{2} + b \geq 0$

$$\Rightarrow b \geq -\frac{1}{2}$$

$$\begin{aligned}
F(t) &= \int_0^t \frac{1}{2} + b(x-1)^3 dx \quad (t \in [0, 2]) \\
&= \frac{1}{2}x + b \frac{(x-1)^4}{4} \Big|_0^t \\
&= \frac{1}{2}t + \frac{b}{4}[(t-1)^4 - 1]
\end{aligned}$$

Bovendien is $F(t) = 0$ als $t < 0$ en is $F(t) = 1$ als $t > 2$

3.3 Plaatsparameters

Van een variabele x kennen we de dichtheidsfunctie $f(t)dt$ die de kans bepaalt dat x een waarde in het interval $[t, t + dt]$ aanneemt. Voor een willekeurige functie $g(x)$ definieert men de verwachte waarde als volgt:

- Discreet:

$$E[g(x)] = \sum_{i=1}^k g(x_i) f(x_i)$$

- Continu:

$$\int_{-\infty}^{+\infty} g(x) f(x) dx$$

Voorbeeld: Wat is de gemiddelde waarde bij het gooien van een dobbelsteen

Uitwerking

$$\begin{aligned}
\mu &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \\
&= 3.5
\end{aligned}$$

Voorbeeld: Bepaal de gemiddelde winst als bij kop 3 euro wordt betaald en bij munt 5 euro. Het munstuk is vervalst zodat $P(k) = 2P(M)$

Uitwerking

We maken een stelsel op met 2 formules. De extra formule is de som van alle kansen die gelijk is aan 1.

$$\begin{cases} P(K) = 2P(M) \\ P(K) + P(M) = 1 \end{cases}$$

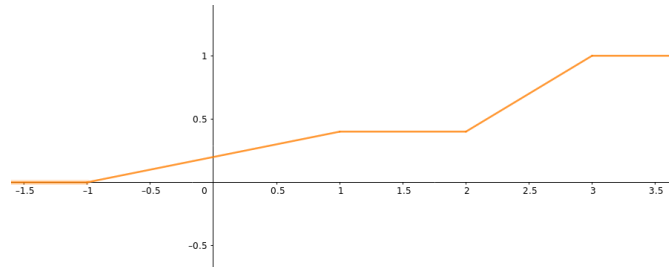
Hieruit wordt afgeleid dat

$$\begin{aligned}
P(K) &= \frac{2}{3} \\
P(M) &= \frac{1}{3}
\end{aligned}$$

De winstfunctie wordt gedefinieerd als g met $g(K) = 3$ en $g(M) = 5$.

$$\begin{aligned} E[g(x)] &= g(K) \cdot P(K) + g(M) \cdot P(M) \\ &= 3 \cdot \frac{2}{3} + 5 \cdot \frac{1}{3} \\ &= 2 + \frac{5}{3} \\ &= \frac{11}{3} \end{aligned}$$

Voorbeeld: Bereken de gemiddelde winst x met $F(x)$ zoals grafisch weergegeven



Uitwerking

$$\begin{aligned} \mu &= E[x] \\ &= \int_{-1}^1 x \cdot f(x) dx + \int_2^3 x \cdot f(x) dx \end{aligned}$$

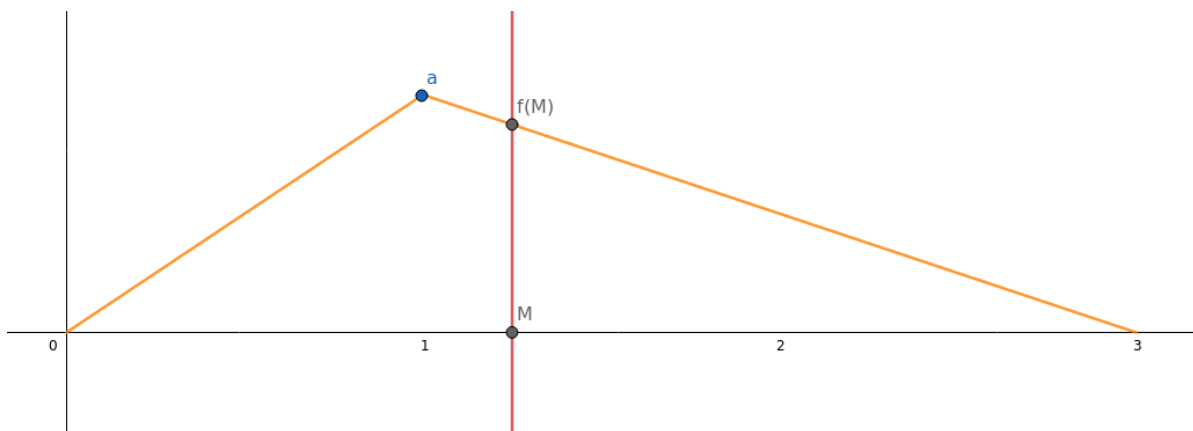
De functie kan gedefinieerd worden als:

$$F(x) = \begin{cases} 0.2(x+1) & x \in [-1, 1] \\ 1 + 0.6(x-3) & x \in [2, 3] \\ 0 & x < -1 \\ 1 & x > 3 \\ 0.4 & x \in [1, 2] \end{cases}$$

We weten dat $f(x) = \frac{dF(x)}{dx}$. We leiden $F(x)$ af voor de intervallen $[-1, 1]$ en $[2, 3]$

$$\begin{aligned} \mu &= \int_{-1}^1 x \cdot \frac{1}{5} dx + \int_2^3 x \cdot \frac{3}{5} dx \\ &= \left. \frac{x^2}{10} \right|_{-1}^1 + \left. \frac{3x^2}{10} \right|_2^3 \\ &= \frac{1}{10} \left[(1-1) + 3(9-4) \right] \\ &= \frac{15}{10} = \frac{3}{2} \end{aligned}$$

Voorbeeld: Bereken de mediaan van x



Uitwerking

De oppervlakte van het rechterdeel moet gelijk zijn aan 50%. We gebruiken de formule van de oppervlakte van een driehoek.

$$\begin{aligned}\frac{(3-M) \cdot f(M)}{2} &= \frac{1}{2} \\ \Leftrightarrow (3-M) \cdot f(M) &= 1\end{aligned}$$

De som van alle kansen moet gelijk zijn aan 1.

$$\begin{aligned}\int_0^3 f(x) dx &= 1 \Leftrightarrow \frac{a}{2} + a = 1 \\ \Leftrightarrow a &= \frac{2}{3}\end{aligned}$$

De rechte door $(2, \frac{2}{3})$ en $(3, 0)$ wordt gegeven door

$$y = -\frac{1}{3}(x - 3)$$

Herschrijf:

$$f(M) = -\frac{1}{3}(M - 3)$$

Nu kan de formule van de oppervlakte van de driehoek terug gebruikt worden:

$$\begin{aligned}(3-M)\left(-\frac{1}{3}(M-3)\right) &= 1 \\ \Leftrightarrow M &= 3 \pm \sqrt{3} \\ \Leftrightarrow M &= 3 - \sqrt{3} \quad (\text{aangezien de functie niet hoger gaat dan } 3)\end{aligned}$$

3.4 Spreidingsparameters

De variantie:

- **Continu:**

$$\sigma^2 = V[x] = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$$

- **Discreet:**

$$\sigma^2 = V[x] = \sum_i (x_i - \mu)^2 f(x_i)$$

De standaardafwijking blijft $\sigma = \sqrt{\sigma^2}$

Voorbeeld: Bewijs $V[ax + b] = a^2 V[x]$

Uitwerking

We bewijzen enkel voor een continue dichtheidsfunctie, maar is analoog aan discrete kansfuncties.

$$\begin{aligned} V[ax + b] &= \int_{-\infty}^{+\infty} (ax + b - a\mu + b)^2 f(x) dx \\ &= \int_{-\infty}^{+\infty} a^2 (x - \mu)^2 f(x) dx \\ &= a^2 \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx \\ &= a^2 V[x] \end{aligned}$$

Voorbeeld: Bewijs $V[x] = E[x^2] - \mu^2$

Uitwerking

$$\begin{aligned} V[x] &= E[(x - \mu)^2] \\ &= E[x^2 - 2x\mu + \mu^2] \\ &= E[x^2] - E[2x\mu] + E[\mu^2] \\ &= E[x^2] - 2\mu E[x] + \mu^2 E[1] \\ &= E[x^2] - 2\mu^2 + \mu^2 \\ &= E[x^2] - \mu^2 \end{aligned}$$

3.5 Momenten

Het moment van de orde k ten opzichte van het punt c:

$$E[(x - c)^k]$$

Er bestaan 2 gevallen voor c.

- $c = 0$: $\mu'_k = E[x^k]$
- $c = \mu$: $\mu_k = E[(x - \mu)^k]$

Een aantal voorbeelden:

1. $\mu'_0 = E[x^0] = E[1] = 1$
2. $\mu'_1 = E[x^1] = \mu$
3. $\mu_0 = E[(x - \mu)^0] = E[1] = 1$
4. $\mu_1 = E[(x - \mu)^1] = E[x] - \mu = 0$
5. $\mu_2 = E[(x - \mu)^2] = V[x] = \sigma$

3.5.1 De momentenfunctie

De momentenfunctie wordt gedefinieerd als:

$$M(t) = E[e^{tx}]$$

- **Discreet:**

$$M(t) = \sum_i e^{tx_i} f(x_i)$$

- **Continu:**

$$M(t) = \int_{-\infty}^{+\infty} e^{tx} f(x) dx$$

Wat ook altijd geldig is:

$$E[x^k] = \left. \frac{d^k M(t)}{dt^k} \right|_{t=0}$$

Bewijs:

$$\begin{aligned} \frac{dM(t)}{dt} &= \left. \frac{d}{dt} \int_{-\infty}^{+\infty} e^{tx} f(x) dx \right|_{t=0} \\ &= \left. \int_{-\infty}^{+\infty} x e^{tx} f(x) dx \right|_{t=0} \\ &= \int_{-\infty}^{+\infty} x f(x) dx \\ &= E[x] \\ &= \mu \end{aligned}$$

$$\begin{aligned} \frac{d^2 M(t)}{dt^2} &= \left. \frac{d}{dt} \int_{-\infty}^{+\infty} x e^{tx} f(x) dx \right|_{t=0} \\ &= \left. \int_{-\infty}^{+\infty} x^2 e^{tx} f(x) dx \right|_{t=0} \\ &= \int_{-\infty}^{+\infty} x^2 f(x) dx \\ &= E[x^2] \end{aligned}$$

$$\begin{aligned}\frac{d^k M(t)}{dt^k} &= \frac{d}{dt} \int_{-\infty}^{+\infty} x^k f(x) dx \\ &= E[x^k]\end{aligned}$$

Voorbeeld: Bereken het gemiddelde en variantie van de geometrische verdeling

Uitwerking

i : # herhalingen van een experiment tot dat een gebeurtenis A voor de eerste keer optreedt en p de kans op succes.

$$f(i) = (1 - p)^{i-1} p$$

Controle dat dit een geldige kansfunctie is:

$$\begin{aligned}\sum_{i=1}^{+\infty} f(i) &= 1 \\ \Rightarrow \sum_{i=1}^{+\infty} (1 - p)^{i-1} p &= 1 \\ \Rightarrow p \sum_{i=1}^{+\infty} (1 - p)^{i-1} &= 1 \quad \text{reeks convergeert naar } \frac{1}{1 - (1 - p)} \\ \Rightarrow p \frac{1}{1 - (1 - p)} &= 1 \\ \Rightarrow 1 &= 1\end{aligned}$$

Het gemiddelde:

$$\mu = \sum_{i=1}^{+\infty} f(i) = \sum_{i=1}^{+\infty} (1-p)^{i-1} p$$

\Rightarrow moeilijk te sommeren, dus gebruik alternatief

$$\mu = \left. \frac{dM}{dt} \right|_{t=0}$$

$$\begin{aligned} M(t) &= \sum_{i=1}^{+\infty} e^{it} f(i) \\ &= \sum_{i=1}^{+\infty} e^{it} (1-p)^{i-1} p \\ &= \sum_{i=1}^{+\infty} \left(e^t (1-p) \right)^{i-1} p e^t \quad \text{aangezien} \quad e^{it} = e^{t(i-1)} e^t \\ &= p e^t \sum_{i=1}^{+\infty} \left(e^t (1-p) \right)^{i-1} \\ &= \frac{p e^t}{1 - e^t (1-p)} \\ \mu &= \left. \frac{dM}{dt} \right|_{t=0} \\ &= \left. \frac{p e^t (-e^t (t-p)) - [1 - e^t (1-p)] p e^t}{[1 - e^t (1-p)]^2} \right|_{t=0} \\ &= \left. \frac{p e^t}{[1 - e^t (1-p)]^2} \right|_{t=0} \\ &= \left. \frac{p}{[1 - 1 - p]^2} \right|_{t=0} \\ &= \frac{1}{p} \end{aligned}$$

De variantie:

$$\begin{aligned}
 \sigma^2 &= E[x^2] - \mu^2 \\
 &= \frac{d^2 M}{dt^2} \Big|_{t=0} - \left(\frac{1}{p}\right)^2 \\
 &= \frac{pe^t(t - e^t(1-p))^2 - 2pe^t(1 - e^t(1-p))}{[1 - e^t(1-p)]^4} \Big|_{t=0} - \left(\frac{1}{p}\right)^2 \\
 &= \frac{p^3 - 2p^2(-(1-p))}{p^4} - \left(\frac{1}{p}\right)^2 \\
 &= \frac{p + 2(1-p)}{p^2} - \left(\frac{1}{p}\right)^2 \\
 &= \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 \\
 &= \frac{1-p}{p^2}
 \end{aligned}$$

3.6 De ongelijkheid van Chebychev

Twee ongelijkheden:

$$\begin{aligned}
 P(|x - \mu| \geq k\sigma) &\leq \frac{1}{k^2} \\
 P(|x - \mu| < k\sigma) &\geq 1 - \frac{1}{k^2}
 \end{aligned}$$

Gebruik enkel deze ongelijkheid als er GEEN informatie is over de verdelingsfunctie. _____

Voorbeeld: Beschouw een x met $\mu = 1$ en $\sigma^2 = \frac{1}{5}$ en een symmetrische verdeling rond μ . Geef een begrenzing voor $P(0 < x < 2)$ en $P(x \geq 1 + \frac{2}{\sqrt{5}})$

Uitwerking

$$\begin{aligned}
 P(0 < x < 2) &= P(P(|x - 1| < 1)) \\
 &= P(P(|x - \mu| < \sqrt{5}\sigma)) \\
 &\geq 1 - \frac{1}{5} \\
 &\geq \frac{4}{5}
 \end{aligned}$$

$$\begin{aligned}
 P(x \geq 1 + \frac{2}{\sqrt{5}}) &= P(x - \mu \geq \frac{2}{\sqrt{5}}) \\
 &= \frac{1}{2} P(|x - \mu| \geq \frac{2}{\sqrt{5}}) \\
 &= \frac{1}{8}
 \end{aligned}$$

Hoofdstuk 4

Discrete verdelingen

4.1 Uniforme discrete verdeling

Elke kans is gelijk.

Voorschrift:

$$f(i) = \frac{1}{n} \quad i \in \{1, 2, 3, \dots, n\}$$

Gemiddelde:

$$\begin{aligned}\mu &= \sum_{i=1}^n i \frac{1}{n} \\ &= \frac{1}{n} \sum_{i=1}^n i \\ &= \frac{1}{n} (1 + 2 + 3 + \dots + (n-1) + n) \\ &= \frac{1}{n} \left(\frac{n(n+1)}{2} \right) \\ &= \frac{n+1}{2}\end{aligned}$$

Variantie:

$$\begin{aligned}\sigma^2 &= E[x^2] - \mu^2 \\ &= \left(\sum_{i=1}^n i^2 \frac{1}{n} \right) - \left(\frac{n+1}{2} \right)^2 \\ &= \frac{1}{n} \left(\frac{n(n+1)(2n+1)}{6} \right) - \left(\frac{n+1}{2} \right)^2 \\ &= \left(\frac{(n+1)(2n+1)}{6} \right) - \left(\frac{n+1}{2} \right)^2 \\ &= (n+1) \left(\frac{(2n+1)}{6} - \frac{1}{4} \right) \\ &= \frac{(n+1)}{12} (4n+2-3n-3) \\ &= \frac{n^2-1}{12}\end{aligned}$$

4.2 Bernoulli experiment

Enkel twee uitkomsten: Succes of geen succes. Success heeft een kans p , geen succes heeft een kans $1 - p$.

Het gemiddelde:

$$\begin{aligned}\mu &= E[x] \\ &= \sum_{i=1}^2 x_i f(x_i) \\ &= 0(1 - p) + 1p \\ &= p\end{aligned}$$

De variantie:

$$\begin{aligned}\sigma^2 &= V[x^2] = E[x^2] - \mu^2 \\ &= \left[\sum_{i=1}^2 x_i^2 f(x_i) \right] - p^2 \\ &= 0^2(1 - p) + 1^2(p) - p^2 \\ &= p(1 - p)\end{aligned}$$

4.3 Binomiale Verdeling

Hoeveel successen zijn er indien n experimenten uitgevoerd worden waarbij elk experiment een kans p heeft om te slagen. De experimenten MOETEN onafhankelijk zijn.

Indien x binomiaal verdeeld: $x : \text{bin}(n, p)$

Voorschrift:

$$f(i) = P(x = i) = C_n^i p^i (1 - p)^{n-i}$$

Voorbeeld: Een arts voert op één dag drie operaties uit die elk slagen met kans 0.9. Wat is de kans op twee succesvolle operaties die dag?

Uitwerking

$x : \text{bin}(3, 0.9)$

$$\begin{aligned}f(2) &= P(x = 2) \\ &= C_3^2 (0.9)^2 (1 - 0.9) \\ &\approx \frac{1}{4}\end{aligned}$$

Gemiddelde: np

Variantie : $np(1 - p)$

4.3.1 Momentenfunctie

$$\begin{aligned} M(t) &= E[e^{ti}] \\ &= \sum_{i=0}^n e^{ti} C_n^i p^i (1-p)^{n-i} \\ &= \sum_{i=0}^n e^{ti} C_n^i \left[(pe^t)^i (1-p)^{n-i} \right] \\ &= (1-p+pe^t)^n \end{aligned}$$

4.3.2 Recursieformule

$$\begin{aligned} f(i+1) &= C_n^{i+1} p^{i+1} (1-p)^{n-(i+1)} \\ &= \frac{n!}{(i+1)!(n-i-1)!} p^{i+1} (1-p)^{n-i-1} \frac{p}{1-p} \\ &= \frac{n!}{i!(n-i)!(i+1)!} p^i (1-p)^{n-i-1} \frac{p}{1-p} \\ &= \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i-1} \cdot \frac{n-i}{i+1} \frac{p}{1-p} \\ &= f(i) \cdot \frac{(n-i)p}{(i+1)(1-p)} \end{aligned}$$

4.4 Geometrische verdeling

Een discrete veranderlijke i waarmee we het aantal keer weergeven dat we een experiment moeten herhalen totdat gebeurtenis A zich voordoet, heeft een geometrische verdeling.

Voorschrift:

$$f(i) = (1-p)^{i-1} p$$

Voorbeeld: We wensen een gepaste persoon te vinden voor een openstaande betrekking. We testen meerdere kandidaten totdat een kandidaat geschikt bevonden wordt. De kans op 'niet geschikt' is 10%. Wat is de kans dat na het testen van vier kandidaten pas de gepaste kandidaat gevonden werd?

Uitwerking

i : # experimenten totdat 1ste succes voorkomt en p : 0.9

$$f(4) = (1-p)^3 p = (0.1)^3 0.9$$

4.5 Hypergeometrische verdeling

N elementen waarvan M de eigenschap A bezitten. We trekken n elementen. Wat is de kans dat i van de n elementen eig. A bezitten?

$$f(i) = \frac{C_M^i C_{N-M}^{n-i}}{C_N^n}$$

Voorbeeld: De variabele $p = 0.3$ stelt de kans voor op een geschikte kandidaat. De variabele i stelt het aantal kandidaten voor totdat er een geschikte kandidaat optreedt en is geometrisch verdeeld. Het kost 90 euro om een kandidaat op geschiktheid te testen.

Bepaal met 75% kans het interval voor de totale kost(TK) om een geschikte kandidaat te vinden.

Uitwerking

$$TK = 90i$$

We maken gebruik van de ongelijkheid van Chebychev:

$$P(|TK - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

Gemiddelde:

$$\begin{aligned}\mu &= E[TK] \\ &= 90E[i] \\ &= 90 \frac{10}{3} \\ &= 300\end{aligned}$$

Standaarddeviatie:

$$\begin{aligned}\sigma &= \sqrt{\sigma^2} \\ &= \sqrt{V[TK]} \\ &= \sqrt{90^2 V[i]} \\ &= \sqrt{90^2 \frac{1-0.3}{0.3^2}} \\ &= \sqrt{63000} \\ &\approx 251\end{aligned}$$

We willen een bovengrens van 75%.

$$1 - \frac{1}{k^2} = 0.75 \Rightarrow k = 2$$

De ongelijkheid wordt:

$$P(|TK - 300| < 502) \geq 75\% = P(300 - 502 < TK < 300 + 502) \geq 75\%$$

Da totale kost bedraagt met een kans van 75% 802 euro

4.6 Poisson verdeling

Definitie:

$$f(i) = P(x = i) = e^{-\lambda} \frac{\lambda^i}{i!}$$

Controle geldige kansfunctie:

$$\begin{aligned} \sum_{i=0}^{+\infty} e^{-\lambda} \frac{\lambda^i}{i!} &= e^{-\lambda} \sum_{i=0}^{+\infty} \frac{\lambda^i}{i!} && \text{reeks convergeert naar } e^{\lambda} \\ &= e^{\lambda-\lambda} \\ &= e^0 \\ &= 1 \end{aligned}$$

Hoofdstuk 5

Continue verdelingen

5.1 Uniforme continue verdeling

$$f(x) = \frac{1}{b-a} \text{ met } x \in \text{ uitkomsteninterval } = [a, b]$$

5.2 Exponentiële verdeling

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$$

5.3 Normale verdeling

Dichtheidsfunctie normale verdeling:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}}$$

x is normaal verdeeld $= x : N(\mu, \sigma)$

5.3.1 Eigenschap

Als x normaal is verdeeld dan is $y = ax + b$ ook normaal verdeeld. m.a.w. $x : N(\mu, \sigma)$ en $y = ax + b$ dan is $y : N(a\mu + b, |a|\sigma)$

Bewijs:

$$\begin{aligned}P(c \leq y \leq d) &= P(c \leq ax + b \leq d) \\&= P\left(\frac{c-b}{a} \leq x \leq \frac{d-b}{a}\right) \\&= \int_{\frac{c-b}{a}}^{\frac{d-b}{a}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx \\&\text{stel } x = \frac{y-b}{a} \\&\text{dan } \frac{d-b}{a} = \frac{y-b}{a} \Rightarrow y = d \\&\text{en } \frac{c-b}{a} = \frac{y-b}{a} \Rightarrow y = c \\&\text{en } dx = \frac{1}{a} dy \\&= \frac{1}{a} \int_c^d \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(y-b-a\mu)^2}{\sigma^2}} dy \\&= \int_c^d \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(y-\mu_y)^2}{\sigma^2}} dy \\&= \int_c^d f_y(y) dy\end{aligned}$$

5.4 Lineaire combinatie

Als $x_i : N(\mu_i, \sigma_i)$ ($i = 1, 2, \dots, n$) onafhankelijke variabelen dan is

$$y = \sum_{i=1}^n a_i x_i + b : N\left(\sum_{i=1}^n a_i \mu_i + b, \sqrt{\sum_{i=1}^n a_i^2 \sigma_i^2}\right)$$

Voorbeeld: Stel $x : N(1, 3)$ $y : N(2, 2)$

Bereken $x - 2y$

Uitwerking

$$\begin{aligned}x - 2y &: N(1 - 2 \cdot 2, \sqrt{1^2 \cdot 3^2 + 2^2 \cdot 2^2}) \\&= N(-3, 5)\end{aligned}$$

Deel II

Oefeningen

Hoofdstuk 6

Kansrekenen

2. Een geldstuk is vervalst zodat kop dubbel zoveel kan voorkomen als munt. Als het geldstuk drie keer geworpen wordt, wat is de kans om juist 2 keer munt te hebben?

Oplossing

Er zijn twee evenementen te definieëren. Kop gooien **K** en munt gooien **M**. Kop gooien kan twee keer zoveel voorkomen als munt gooien.

$$P(K) = 2P(M)$$

We kunnen gebruik maken van het feit dat de som van alle kansen gelijk is aan 1.

$$\begin{aligned} P(K) + P(M) &= 1 \\ \Leftrightarrow 2P(M) + P(M) &= 1 \\ \Leftrightarrow 3P(M) &= 1 \\ \Leftrightarrow P(M) &= \frac{1}{3} \end{aligned}$$

dus

$$P(K) = \frac{2}{3}$$

Aangezien de gebeurtenis munt twee keer moet voorkomen moet kop dus slechts één maal voorkomen.

$$P(2M \cap K)$$

Er moet rekening gehouden worden met de verschillende combinaties:

$$\begin{aligned} &P((M \cap M \cap K) \cup (M \cap K \cap M) \cup (K \cap M \cap M)) \\ &= P(M \cap M \cap K) \cup P(M \cap K \cap M) \cup P(K \cap M \cap M) \\ &= P(M)P(M)P(K) + P(M)P(K)P(M) + P(K)P(M)P(M) \\ &= \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \\ &= 3\left(\frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3}\right) \\ &= 3\left(\frac{2}{27}\right) = \frac{6}{27} = \frac{2}{9} \end{aligned}$$

3. Een dobbelsteen is vervalst zodat de kans dat een gegeven aantal ogen geworpen wordt evenredig is met het aantal ogen. Is A de gebeurtenis een even aantal te gooien, B de gebeurtenis een priemgetal te gooien en C de gebeurtenis een oneven getal te gooien,

1. Bepaal $P(A)$, $P(B)$ en $P(C)$
2. Bereken de kans dat men een even getal of een priemgetal gooit.
3. Bereken de kans dat men een even getal gooit dat geen priemgetal is.
4. Bereken de kans dat men een oneven getal of een priemgetal gooit.

Oplossing

De kans kan als formule worden voorgesteld.

$$P(i) = ip$$

waarbij p een willekeurig getal is. We weten dat de som van alle kansen gelijk is aan 1.

$$\sum_{i=1}^6 P(i) = 1p + 2p + 3p + 4p + 5p + 6p = 1$$

Los op naar p

$$\begin{aligned} 21p &= 1 \\ p &= \frac{1}{21} \end{aligned}$$

De deeloplossingen:

1.

$$\begin{aligned} P(A) &= \frac{2}{21} + \frac{4}{21} + \frac{6}{21} = \frac{12}{21} = \frac{4}{7} \\ P(B) &= \frac{2}{21} + \frac{3}{21} + \frac{5}{21} = \frac{10}{21} \\ P(C) &= 1 - P(A) = 1 - \frac{4}{7} = \frac{3}{7} \end{aligned}$$

2.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{4}{7} + \frac{10}{21} - \frac{2}{21} = \frac{20}{21}$$

3.

$$P(A \cap \overline{B}) = P(A) - P(A \cap B) = \frac{4}{7} - \frac{2}{21} = \frac{10}{21}$$

4.

$$P(B \cup C) = P(B) + P(C) - P(B \cap C) = \frac{10}{21} + \frac{3}{7} - \left(\frac{3}{21} + \frac{5}{21}\right) = \frac{11}{21}$$

4. A en B zijn verschijnselen met $P(A) = 0.1$, $P(B) = 0.5$. Bepaal $P(A \cup B)$, $P(\overline{A})$, $P(\overline{A} \cap B)$ indien a) de verschijnselen elkaar uitsluiten en b) ze onafhankelijk zijn.

Oplossing

(a)

$$\begin{aligned}P(A \cup B) &= P(A) + P(B) - P(A \cap B) = \frac{1}{10} + \frac{1}{2} - 0 = \frac{3}{5} \\P(\bar{A}) &= 1 - P(A) = 1 - \frac{1}{10} = \frac{9}{10} \\P(\bar{A} \cap B) &= P(B) - P(A \cap B) = \frac{1}{2} - 0 = \frac{1}{2}\end{aligned}$$

(b)

$$\begin{aligned}P(A \cup B) &= P(A) + P(B) - P(A \cap B) = \frac{1}{10} + \frac{1}{2} - \frac{1}{20} = \frac{11}{20} \\P(\bar{A}) &= 1 - P(A) = 1 - \frac{1}{10} = \frac{9}{10} \\P(\bar{A} \cap B) &= P(B) - P(A \cap B) = \frac{1}{2} - \frac{1}{20} = \frac{9}{20}\end{aligned}$$

6. Bereken voor een familie van 3 kinderen de kans op a) 3 jongens en b) 2 jongens en 1 meisje

Oplossing

$$\begin{aligned}P(J): \text{Kans op een jongen} &= \frac{1}{2} \\P(M): \text{Kans op een meisje} &= \frac{1}{2}\end{aligned}$$

(a)

$$\begin{aligned}P(J \cap J \cap J) &= P(J)P(J)P(J) \\&= (P(J))^3 \\&= \left(\frac{1}{2}\right)^3 \\&= \frac{1}{8}\end{aligned}$$

(b)

$$\begin{aligned}&P(J \cap J \cap M) + P(J \cap M \cap J) + P(M \cap J \cap J) \\&= 3P(J \cap J \cap M) \\&= 3(P(J))^3 \quad (\text{aangezien } P(J) = P(M)) \\&= 3\left(\frac{1}{2}\right)^3 = \frac{3}{8}\end{aligned}$$

7. Een paar onvervalste dobbelstenen worden geworpen. Wat is de kans dat de som van de ogen een totaal van minstens 8 vertoont.

Oplossing

$$\begin{aligned}
 A &= \text{som van de ogen} \geq 8 \\
 P(A) &= \{2, 6\} \cup \{3, 5\} \cup \{3, 6\} \cup \dots \cup \{6, 6\} \\
 P(A) &= \frac{15}{36} = \frac{5}{12}
 \end{aligned}$$

11. Gegeven 3 kasten A, B en C. Elke kast heeft een aantal laden die ofwel een goudstuk(G), ofwel een zilverstuk(Z) ofwel niets bevatten(N) en dit als volgt:

A: [G—G—G—Z]

B: [G—Z—Z]

C: [G—G—N—Z—Z]

Men kiest willekeurig één van de kasten, opent daarvan at random één lade, en grijpt het muntstuk(indien mogelijk). Wat is de kans dat men kast A heeft uitgekozen indien men een goudstuk heeft genomen?

Oplossing

De kans om eender welke lade te pakken:

$$P(A) = P(B) = P(C) = \frac{1}{3}$$

De kans om een goudstuk te pakken afhankelijk van de lade:

$$\begin{aligned}
 P(G|A) &= \frac{3}{4} \\
 P(G|B) &= \frac{1}{3} \\
 P(G|C) &= \frac{2}{5}
 \end{aligned}$$

De kans dat lade A gekozen werd indien het een goudstuk genomen is:

$$\begin{aligned}
 P(A|G) &= \frac{P(A)P(G|A)}{P(A)P(G|A) + P(B)P(G|B) + P(C)P(G|C)} \\
 &= \frac{\frac{1}{3} \cdot \frac{3}{4}}{\frac{1}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{5}} \\
 &= \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{9} + \frac{2}{15}} \\
 &= \frac{45}{89}
 \end{aligned}$$

16. Wanneer men het waarheidsserum toedient aan een schuldig persoon is het voor 90% betrouwbaar en aan een onschuldig persoon is het voor 99% betrouwbaar. Als een verdachte gekozen wordt uit een groep, waarvan 5% reeds een misdrijf begaan hebben, wat is de kans dat die persoon niet schuldig is als het waarheidsserum schuldig aanwijst?

Oplossing

2 Gebeurtenissen:

1. PS: De persoon is schuldig
2. WS: Het waarheidsserum wijst schuldig aan

De kans dat de persoon niet schuldig is als het waarheidsserum schuldig aanwijst:

$$\begin{aligned}
 P(\overline{PS}|WS) &= \frac{P(\overline{PS})P(WS|\overline{PS})}{P(\overline{PS})P(WS|\overline{PS}) + P(PS)P(WS|PS)} \\
 &= \frac{\frac{19}{20} \cdot \frac{1}{100}}{\frac{19}{20} \cdot \frac{1}{100} + \frac{1}{20} \cdot \frac{9}{10}} \\
 &= \frac{\frac{19}{2000}}{\frac{19}{2000} + \frac{9}{200}} \\
 &= \frac{\frac{19}{2000}}{\frac{19}{2000} + \frac{90}{2000}} \\
 &= \frac{\frac{19}{2000}}{\frac{109}{2000}} \\
 &= \frac{19}{109}
 \end{aligned}$$

17. Uit een spel van 52 kaarten trekt men willekeurig maar tegelijkertijd vijf kaarten. Bereken de kans dat

1. het vijf zwarte kaarten zijn,
2. het drie heren en twee vrouwen zijn,
3. er tenminsten één aas bij is,
4. er ten hoogste één harten bij is.

Oplossing

1.

$$\begin{aligned}
 P(5Z) &= \frac{C_{26}^5}{C_{52}^5} \\
 &= \frac{\frac{26!}{5!(26-5)!}}{\frac{52!}{5!(52-5)!}} \quad (\text{vereenvoudigen voor rekenmachine}) \\
 &= \frac{26!}{21!} \cdot \frac{47!}{52!} \\
 &= \frac{26 \cdot 25 \cdot 24 \cdot 23 \cdot 22}{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48} \\
 &= 0.025
 \end{aligned}$$

2.

$$P(3H \cap 2V) = \frac{C_4^3 \cdot C_4^2}{C_{52}^5}$$

Meestal zullen ze vragen 'geef de correcte uitdrukking' om geen tijd te verliezen aan banaal rekenwerk.

3.

$$P(\text{minstens 1 aas}) = 1 - P(\text{geen aas}) = 1 - \frac{C_4^0 \cdot C_{48}^5}{C_{52}^5}$$

4.

$$\begin{aligned} P(\text{ten hoogste 1 hart}) &= P(0 \text{ hart} \cap \overline{5 \text{ hart}}) \cup P(1 \text{ hart} \cap \overline{4 \text{ hart}}) \\ &= \frac{C_{39}^5}{C_{52}^5} + \frac{C_{13}^1 \cdot C_{39}^4}{C_{52}^5} \end{aligned}$$

20. Bepaal de kans om minstens één maal zes te gooien bij 4 worpen met een dobbelsteen. Bepaal de kans om minstens één maal dubbel zes te gooien bij 24 worpen met 2 dobbelstenen.

Oplossing

$$\begin{aligned} P(\text{minstens één 6 bij 4 worpen}) &= 1 - P(\text{geen 6 bij 4 worpen}) \\ &= 1 - \left(\frac{5}{6}\right)^4 \end{aligned}$$

$$\begin{aligned} P(\text{minstens één keer dubbel 6 bij 24 worpen}) &= 1 - P(\text{geen dubbel 6 bij 24 worpen}) \\ &= 1 - \left(\frac{35}{36}\right)^{24} \end{aligned}$$

21. Bepaal de kans om met de belgische lotto a)drie cijfers b)vier cijfers en c)zes cijfers goed te hebben.

Oplossing

$$P(x) = \frac{C_6^x \cdot C_{42-x}^{6-x}}{C_{42}^6}$$

x in te vullen met 3, 4 of 6.

23. Een urne bevat 10 ballen waarvan n rode. De kans om twee rode ballen te trekken op drie trekkingen met terugleggen is 1.08 keer de kans om twee rode ballen te trekken op drie trekkingen zonder terugleggen. Bepaal n en sluit de triviale gevallen uit.

Oplossing

Twee gebeurtenissen:

1. A: 2 rode ballen trekken op 3 trekkingen met teruglegging
2. B: 2 rode ballen trekken op 3 trekkingen zonder teruglegging

Verder geldt dat $P(A) = 1.08P(B)$ en $2 \leq n \leq 10$

De kans op B:

$$\begin{aligned}
 P(B) &= \frac{C_n^2 C_{10-n}^1}{C_{10}^3} \\
 &= \frac{\frac{n!}{2!(n-2)!} \frac{(10-n)!}{1!(9-n)!}}{\frac{10!}{3!7!}} \\
 &= \frac{n!}{2!(n-2)!} \frac{(10-n)!}{(9-n)!} \frac{3!7!}{10!} \\
 &= \frac{3n(n-1)(10-n)}{10 \cdot 9 \cdot 8}
 \end{aligned}$$

De kans op geen rode bal:

$$\begin{aligned}
 P(R) &= \frac{n}{10} \\
 \Rightarrow P(\bar{R}) &= 1 - \frac{n}{10} = \frac{10-n}{10}
 \end{aligned}$$

De kans op A:

$$\begin{aligned}
 P(A) &= C_3^2 P(R)^2 P(\bar{R}) \\
 &= 3 \left(\frac{n}{10} \right)^2 \left(\frac{10-n}{10} \right) \\
 &= \frac{3n^2(10-n)}{10^3}
 \end{aligned}$$

Berekening van n

$$\begin{aligned}
 P(A) &= 1.08P(B) \\
 \Leftrightarrow \frac{3n^2(10-n)}{10^3} &= \frac{3n(n-1)(10-n)}{10 \cdot 9 \cdot 8} \\
 \Leftrightarrow \frac{n}{10^2} &= 1.08 \frac{n-1}{9 \cdot 8} \\
 \Leftrightarrow \frac{n}{100} &= 1.08 \frac{n-1}{72} \\
 \Leftrightarrow 72n &= 108(n-1) \\
 \Leftrightarrow 72n &= 108n - 108 \\
 \Leftrightarrow 108n - 72n &= 108 \\
 \Leftrightarrow 36n &= 108 \\
 \Leftrightarrow n &= 3
 \end{aligned}$$

Hoofdstuk 7

Verdelingsfunctie van een populatie

- 3. Ga na of de volgende functi $F(x)$ een cumulatieve distributiefunctie kan zijn. Indien ja, bepaal de corresponderende dichtheidsfunctie $f(x)$

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-x^2} & x > 0 \end{cases}$$

Oplossing

Voer de drie controles uit:

1. De functie is nooit negatief.
2. De functie is nooit dalend.
3. De limiet naar $+\infty$ is 1.

1. Controle functie nooit negatief:

$$\begin{aligned} 1 - e^{-x^2} &> 0 \\ \Leftrightarrow -e^{-x^2} &> -1 \\ \Leftrightarrow e^{-x^2} &< 1 \\ \Leftrightarrow e^{-x^2} &> e^0 \\ \Leftrightarrow -x^2 &> 0 \end{aligned}$$

2. Controle functie nooit dalend:

$$F'(x) = 2xe^{-x^2}$$

De afgeleide is altijd positief, dus daalt de functie nooit.

3. Controle limiet naar $+\infty$ is gelijk aan 1.

$$\begin{aligned} \lim_{t \rightarrow +\infty} F(t) &= \lim_{t \rightarrow +\infty} 1 - e^{-x^2} \\ &= \lim_{t \rightarrow +\infty} 1 - \frac{1}{e^{x^2}} \\ &= 1 - 0 = 1 \end{aligned}$$

De kansfunctie hebben we al berekent in stap 2. $f(x) = 2xe^{-x^2}$

-
- 5. Bepaal C zodat de volgende functie een dichtheidsfunctie is. Bepaal de corresponderende cumulatieve distributiefunctie

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{elders} \end{cases}$$

Oplossing

Bepalen C:

$$\begin{aligned} \int_{-\infty}^{+\infty} C(4x - 2x^2) dx &= 1 \\ \Leftrightarrow 2C \int_0^2 2x - x^2 dx &= 1 \\ \Leftrightarrow 2C \left[\frac{2x^2}{2} - \frac{x^3}{3} \right]_0^2 &= 1 \\ \Leftrightarrow 2C \left[x^2 - \frac{x^3}{3} \right]_0^2 &= 1 \\ \Leftrightarrow 2C \left[\left(4 - \frac{8}{3} \right) - \left(0 - \frac{0}{3} \right) \right] &= 1 \\ \Leftrightarrow 2C \left(\frac{4}{3} \right) &= 1 \\ \Leftrightarrow C &= \frac{3}{8} \end{aligned}$$

Dus:

$$f(x) = \begin{cases} \frac{3}{8}(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{elders} \end{cases}$$

De cumulatieve distributiefunctie:

$$F(x) = \begin{cases} 0 & x \leq 2 \\ 0 + \frac{3}{8} \int_0^x (4t - t^2) dt = \frac{3}{8} \left(\frac{4x^2}{2} - \frac{2x^3}{3} \right) & 0 < x < 2 \\ 1 & x \geq 2 \end{cases}$$

- 6. Een vaas bevat 4 rode en 6 witte ballen. Men neemt 3 ballen zonder teruglegging. Als x het aantal getrokken rode ballen voorstelt bepaal dan:

1. de dichtheidsfunctie van x
2. de grafiek van de dichtheidsfunctie en de corresponderende verdelingsfunctie
3. de gemiddelde waarde en de modus

4. de variantie
5. $P(x \geq 1)$

Oplossing

1. $x = \#$ getrokken rode ballen.

De kans om $x = 0, 1, 2$ of 3 rode ballen te trekken kan voorgesteld worden als :

$$f(x) = \frac{C_4^x C_6^{3-x}}{C_{10}^3}$$

Bijgevolg zijn de kansen:

x	0	1	2	3
$f(x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

2. *-TODO: grafiek*

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{6} & 0 \leq x \leq 1 \\ \frac{1}{6} + \frac{1}{2} = \frac{2}{3} & 1 \leq x \leq 2 \\ \frac{2}{3} + \frac{3}{10} = \frac{29}{30} & 2 \leq x \leq 3 \\ \frac{29}{30} + \frac{1}{30} = 1 & x \geq 3 \end{cases}$$

3. Het gemiddelde:

$$\begin{aligned} \mu &= \sum_{i=0}^3 x_i f(x_i) \\ &= 0f(0) + 1f(1) + 2f(2) + 3f(3) \\ &= 0 + \frac{1}{6} + 2\frac{2}{3} + 3\frac{29}{30} \\ &= 1.2 \end{aligned}$$

De modus is 1.

4. De variantie:

$$\begin{aligned} \sigma^2 &= \left[\sum x_i^2 (f(x_i)) \right] - \mu^2 \\ &= 0^2 f(0) + 1^2 f(1) + 2^2 f(2) + 3^2 f(3) - \mu^2 \\ &= 0.56 \end{aligned}$$

- 5.

$$P(x \geq 1) = 1 - P(x < 1) = 1 - F(0) = 1 - \frac{1}{6} = \frac{5}{6}$$

- 10. Voor een gokspel met drie onvervalste dobbelstenen bedraagt de inzet steeds 5 euro. Indien iemand juist één 6 werpt krijgt hij zijn inzet terug, indien juist twee stenen een 6 vertonen krijgt hij 10 euro terug en indien de drie dobbelstenen een 6 vertonen krijgt hij 15 euro. Wat is de gemiddelde winst (of verlies)?

Oplossing

De verwachte waarde van een functie:

$$E[g(x)] = \sum g(x_i)f(x_i)$$

met $g(x)$ = de winst met een inzet van 5 euro

en $f(x)$ = de kans om x aantal zessen te hebben.

$x = \#6$ bij het werpen van een dobbelsteen De functie f kan geschreven worden als:

$$f(x) = C_3^x \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{3-x}$$

Dus:	<table border="0"><tr><td>x</td><td>0</td><td>1</td><td>2</td><td>3</td></tr><tr><td>f(x)</td><td>$\left(\frac{5}{6}\right)^3$</td><td>$3\frac{5^2}{6^3}$</td><td>$\frac{15}{6^3}$</td><td>$\frac{1}{6^3}$</td></tr><tr><td>g(x)</td><td>-5</td><td>0</td><td>5</td><td>10</td></tr></table>	x	0	1	2	3	f(x)	$\left(\frac{5}{6}\right)^3$	$3\frac{5^2}{6^3}$	$\frac{15}{6^3}$	$\frac{1}{6^3}$	g(x)	-5	0	5	10	Hieruit volgt:
x	0	1	2	3													
f(x)	$\left(\frac{5}{6}\right)^3$	$3\frac{5^2}{6^3}$	$\frac{15}{6^3}$	$\frac{1}{6^3}$													
g(x)	-5	0	5	10													

$$\begin{aligned} E[g(x)] &= f(0)g(0) + f(1)g(1) + f(2)g(2) + f(3)g(3) \\ &= -2.5 \end{aligned}$$

Er is dus gemiddeld een verlies van -2,5 euro

-
- 13. Een toevalsveranderlijke x heeft een gemiddelde $\mu = 12$, een dispersie $\sigma = 3$ en zijn dichtheidsfunctie is niet gekend.

Bepaal een ondergrens voor: $P(6 < x < 18)$ en voor $P(3 < x < 21)$

Oplossing

$$\begin{aligned} P(6 < x < 18) &= P(6 - \mu < x - \mu < 18 - \mu) \\ &= P(-6 < x - \mu < 6) \\ &= P(-2\sigma < x - \mu < 2\sigma) \\ &= P(|x - \mu| < 2\sigma) \\ &= 1 - \frac{1}{2^2} \\ &= 1 - \frac{1}{4} \\ &= \frac{3}{4} \end{aligned}$$

$$\begin{aligned} P(3 < x < 21) &= P(3 - \mu < x - \mu < 21 - \mu) \\ &= P(-9 < x - \mu < 9) \\ &= P(-3\sigma < x - \mu < 3\sigma) \\ &= P(|x - \mu| < 3\sigma) \\ &= 1 - \frac{1}{3^2} \\ &= 1 - \frac{1}{9} \\ &= \frac{8}{9} \end{aligned}$$

-
- 16. Onderstel dat het aantal producten in een fabriek, aangemaakt gedurende één week, een stochastische veranderlijke is met gemiddelde $\mu = 50$. Deze veranderlijke heeft een symmetrische verdeling t.o.v. μ
 1. Bepaal een bovengrens voor de kans dat de productie van een bepaalde week minstens 75 zal dragen.
 2. Indien bovendien de variantie gekend is ($\sigma^2 = 25$) wat is dan een ondergrens voor de waarschijnlijkheid dat de productie van een bepaalde week strikt tussen 40 en 60 zal liggen.

Oplossing

1.

$$\begin{aligned}
 P(x \leq \mu - k\sigma) &= P(x \geq \mu + k\sigma) \leq \frac{1}{2k^2} \\
 \Rightarrow p(x \geq 75) &= P(x - \mu \geq 25) = \frac{1}{2k^2} \\
 &= \frac{1}{2\left(\frac{25}{\sigma}\right)^2} \\
 &= \frac{\sigma^2}{1250}
 \end{aligned}$$

2.

$$\begin{aligned}
 \sigma &= \sqrt{\sigma^2} = \sqrt{25} = 5 \\
 P(40 < x < 60) &= P(-10 < x - \mu < 10) \\
 &= P(-2\sigma < x - \mu < 2\sigma) \\
 &= 1 - \frac{1}{2^2} \\
 &= \frac{3}{4}
 \end{aligned}$$

Hoofdstuk 8

Discrete verdelingen

8. Het zelfmoordpercentage in een Amerikaanse staat bedraagt per maand 1 per 100 000 inwoners.

1. Wat is de kans dat van 400 000 inwoners in deze staat er 8 of meer zelfmoorden plaatsgrijpen in een bepaalde maand?
2. Wat is de kans dat er tenminste 2 maanden in het jaar zullen zijn met minstens 8 zelfmoorden?

Oplossing

1. Gem # zelfmoorden bij 100 000 inw/maand = 1
Gem # zelfmoorden bij 400 000 inw/maand = 4
Dit is poisson verdeeld met $\lambda = 4$ dus $f(i) = e^{-4} \frac{4^i}{i!}$

$$\begin{aligned} P(x \geq 8) &= 1 - p(x < 8) \\ &= 1 - \sum_{i=0}^7 e^{-4} \frac{4^i}{i!} \\ &= 1 - e^{-4} \sum_{i=0}^7 \frac{4^i}{i!} \\ &\approx 0.0511 \end{aligned}$$

2.
 - n = 2
 - i = # maanden met > 8 zelfmoorden
 - p = 0.0511

$$\begin{aligned} P(i \geq 2) &= 1 - P(i < 2) \\ &= 1 - \left(C_{12}^0 p^0 (1-p)^{12} + C_{12}^1 p^1 (1-p)^{11} \right) \\ &\approx 0.1227 \end{aligned}$$

Extra oefening. Het aantal olietankers dat per dag een bepaalde haven binnenvaart is poisson verdeeld met gemiddelde 2. De haven kan ten hoogste 3 olietankers per dag verwerken. Als er 2 olietankers per dag binnenvaren dan worden ze doorgestuurd.

1. Bepaal de kans dat er op een dag olietankers worden bediend.
2. Bepaal het gemiddelde aantal olietankers dat per dag bediend wordt.

Oplossing

1.
 - x_i : # olietankers die binnenvaren
 - x : poisson verdeeld
 - $\lambda = 2 \Rightarrow f(i) = e^{-2} \frac{2^i}{i!}$

$$\begin{aligned}
 P(x > 3) &= 1 - P(x \leq 3) \\
 &= 1 - \sum_{i=0}^3 e^{-2} \frac{2^i}{i!} \\
 &= 1 - e^{-2} \left(1 + 2 + 2 + \frac{8}{6}\right) \\
 &\approx 0.1429
 \end{aligned}$$

2. y : # olietankers die bediend worden $\Rightarrow E[y] = \sum x_i f(x_i)$

$$\begin{aligned}
 &0f(0) + 1f(1) + 2f(2) + 3f(3) + 3P(x > 3) \\
 &= 0 + 1 \cdot 2e^{-2} + 2 \cdot 2e^{-2} + 3 \cdot \frac{8}{6}e^{-2} + 3 \cdot 0.1429 \\
 &\approx 1.7821
 \end{aligned}$$

13. Een machine produceert bouten waarvan er 2% defect zijn. Wat is de kans dat bij 50 bouten ten hoogste 2 defect zijn?

Oplossing

- $n = 50$
- $p = 0.02$
- i : # bouten dat defect zijn
- i : binomiaal verdeeld

$$\begin{aligned}
 p(i \leq 2) &= \sum_{i=0}^2 C_{50}^i p^i (1-p)^{50-i} \\
 &= 0.9216
 \end{aligned}$$

Kan ook opgelost worden met de benadering van poisson want $n = 50 \geq 50$ en $p = 0.02 \leq 0.1$ dus $\lambda = np = 1$ en $f(i) = e^{-1} \frac{1}{i!}$

$$\begin{aligned}
 P(i \leq 2) &= \sum_{i=0}^2 e^{-1} \frac{1}{i!} \\
 &\approx 0.9197
 \end{aligned}$$

17. De kans dat een geïnfecteerd persoon sterft aan een ademhalingsinfectie is 0.002. Bepaal de kans dat er minder dan 5 personen van een groep van 2000 geïnfecteerde personen zullen sterven. Gebruik de ongelijkheid van Chebychev en interpreteer het interval $]\mu - 2\sigma, \mu + 2\sigma[$

Oplossing

De verdeling is binomiaal verdeeld maar we gebruiken de poisson benadering (+ bewijs).

- $n = 2000 \geq 50$
- $p = 0.002 \leq 0.1$
- $\lambda = np = 2000 \cdot (0.002) = 4$
- $f(i) = e^{-4} \frac{4^i}{i!}$

1.

$$P(i < 5) = \sum_{i=0}^4 e^{-4} \frac{4^i}{i!} \approx 0.6228$$

2. Het interval $]\mu - 2\sigma, \mu + 2\sigma[$:

$$\begin{aligned} P(|x - \mu| < k\sigma) &\geq 1 - \frac{1}{k^2} \\ \Leftrightarrow P(-k\sigma + \mu < x < k\sigma + \mu) &\geq 1 - \frac{1}{k^2} \\ \mu = \lambda = 2 \rightarrow \sigma = \sqrt{\lambda} = 2 \\ \text{dus }]4 - 2 \cdot 2, 4 + 2 \cdot 2[&\Rightarrow]0, 8[\\ \Leftrightarrow P(0 < x < 8) &\geq 1 - \frac{1}{k^2} \\ &= 0.75 \end{aligned}$$

18. Uit een kaartspel met 52 kaarten wordt 1 kaart getrokken en teruggelegd.

1. Wat is de kans dat men bij de vierde beurt voor het eerst een aas trekt
2. Wat is de kans dat men minstens zes beurten nodig heeft om een aas te trekken

Oplossing

Geometrische verdeling gebruiken

1. $x = 4$ en $p = \frac{1}{13}$

$$\begin{aligned} f(4) &= \left(\left(1 - \frac{1}{13}\right)^{4-1} \cdot \frac{1}{13} \right) \\ &= \frac{\left(1 - \frac{1}{13}\right)^3}{13} \\ &\approx 0.0605 \end{aligned}$$

2.

$$P(\text{minstens 6 beurten nodig om een aas te trekken}) = \left(\frac{12}{13}\right)^5 \\ \approx 0.6702$$

Extra oefening. Het aantal fouten dat een bepaalde secretaresse op een blad typt is poisson verdeeld met gemiddelde 1. De secretaresse typt 50 bladzijden. Wat is de kans dat de eerste fout op bladzijde 3 staat.

Oplossing

- x: eerste blad met typfout $\rightarrow P(x = 3) = (1 - p)^2 p$
- y: poisson met $\lambda = 1 \rightarrow f(i) = e^{-1} \frac{1^i}{i!}$

$$P(y = 1) = 1 - f(0) = 1 - \frac{1}{e} \\ P(x = 3) = \left(\frac{1}{e}\right)^2 \left(1 - \frac{1}{e}\right) \\ \approx 0.0855$$

Hoofdstuk 9

Continue verdelingenen

11. Een Geigerteller levert gemiddeld 30 tellen per minuut in de omgeving van een radio-actief materiaal. Stel dat het aantal tellen per minuut Poisson verdeeld is. Bepaal de kans dat

1. er juist 32 tellen zijn,
2. er tussen 23 en 35 tellen zijn, grenzen niet inbegrepen.

Oplossing

Stel:

- x : # tellen
- x is Poisson verdeeld met $\lambda = 30$
- $f(i) = \frac{e^{-30}(30)^i}{i!}$

1. $P(i = 32) = \frac{e^{-30}(30)^{32}}{32!} = ?$ (teveel werk met rekenmachine)

We lossen door gebruik te maken van de limietstelling:

$$\begin{aligned}\lambda = 30 \geq 15 &\Rightarrow x : N(30, \sqrt{30}) \\ P(x = 32) &\approx P(31,5 < x < 32,5) \\ &= P\left(\frac{31,5 - 30}{\sqrt{30}} < z < \frac{32,5 - 30}{\sqrt{30}}\right) \\ &= P(0,27 < z < 0,46) \\ &= P(0 < z < 0,46) - P(0 < z < 0,27) \\ &= 0,1772 - 0,1064 \\ &= 0,0708\end{aligned}$$

2.

$$\begin{aligned}P(23 < x < 35) &\approx P(23,5 < x < 34,5) \\ &= P\left(\frac{23,5 - 30}{\sqrt{30}} < z < \frac{34,5 - 30}{\sqrt{30}}\right) \\ &= P(-1,19 < z < 0,82) \\ &= P(0 < z < 0,82) - P(0 < z < 1,19) \\ &= 0,2939 + 0,3830 \\ &= 0,6769\end{aligned}$$

Extra: Een speler gooit met een dobbelsteen. Bepaal de kans dat hij bij 90 worpen tenminste 39 keer een 5 of een 6 gooit. Geef de exacte formule en bereken met de limiet-stelling.

Oplossing

Stel:

- x : # keer een vijf of zes.
- kans op $x = p = 1/3$
- x is binomiaal verdeeld
- $f(i) = C_n^i p^i (1-p)^{n-i}$

De exacte formule wordt:

$$P(x \geq 39) = 1 - P(x \leq 38) = 1 - \sum_{i=0}^{38} C_{90}^i \left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^{90-i}$$

We lossen dit op met de limietstelling 2 van het formularium. Bewijs:

- $np = 90 \cdot \frac{1}{3} = 30 \geq 5$
- $n(1-p) = 90 \cdot \frac{2}{3} = 60 \geq 5$

De variabele x benadert een normale verdeling.

$$x : N(np, \sqrt{np(1-p)}) = N(30, \sqrt{20})$$

$$\begin{aligned} P(x \geq 39) &\approx P(x \geq 38,5) \\ &= P\left(z \geq \frac{38,5 - 30}{\sqrt{20}}\right) \\ &= P(z \geq 1,9) \\ &= 0.5 - 0.4713 \\ &= 0.0287 \end{aligned}$$

13. Stel dat $x : t(20df)$ verdeeld is. Bepaal

1. a zodat $P(x > a) = 0.3$
2. b zodat $P(|x| > b) = 0.2$

Oplossing

1. $1 - 0.3 = 0.7$

$$a = t_{0.7}(20df) = 0.533$$

2. $|x| > b \rightarrow x < -b$ en $x > b$ Dus $P(x < -b) + P(x > b) = 0.2$ maar wegens symmetrie beschouwen we $P(x > b) = 0.1$ Hieruit volgt: $b = t_{0.9}(20df) = 1.32$

14. Stel $x : N(0, 3)$ en $y : \chi^2(16df)$ verdeeld met x en y onafhankelijk. Bepaal $a < 0$ zodat $P(ax \geq \sqrt{y}) = 0.3$.

Oplossing

We weten dat de t -verdeling kan omschreven worden als $t(vdf) = \frac{z}{\sqrt{y/v}}$ met $z : N(0, 1)$ en $y : \chi^2(vdf)$. In ons geval herschrijven we de kans eerst als $P(ax \geq \sqrt{y}) = P\left(\frac{x}{\sqrt{y}} \leq \frac{1}{a}\right)$ zodat we de definitie van de t -verdeling kunnen toepassen. Eerst moeten we x normeren $\rightarrow \frac{x-0}{3}$ is $N(0, 1)$. Hieruit volgt dat

$$\frac{x/3}{\sqrt{y/16}}$$

$t(16 \text{ df})$ verdeeld is. De kans wordt:

$$P\left(\frac{4}{3} \frac{x}{\sqrt{y}} \leq \frac{4}{3} \frac{1}{a}\right)$$

Via het formularium kan afgeleidt worden dat $-\frac{4}{3} \frac{1}{a} = t_{0.7}(16df) = 0.535 \Rightarrow a \approx -2.4922$

15. Stel dat $x : F(10, 15df)$ verdeeld is. Bepaal a en b zodat $P(x < a) = 0.95$ en $P(x > b) = 0.95$

Oplossing

$a = F_{0.05}(10, 15df) = 2.54$ en $b = F_{0.95}(10, 15df) = \frac{1}{F_{0.05}(15, 10df)} = \frac{1}{2.85} \approx 0.3509$

18. Stel dat $x : \chi^2(350df)$ verdeeld is, bepaald $P(x > 390)$

Oplossing

Een stelling dat NIET op het formularium staat zegt dat als $v \geq 30$, dan $z = \frac{x-v}{\sqrt{2v}}$ is $N(0, 1)$ Hieruit volgt direct dat $\frac{x-350}{\sqrt{700}} N(0, 1)$ verdeeld is. De kans wordt:

$$P(x > 390) = P\left(z > \frac{390 - 350}{\sqrt{700}}\right) = P(z > 1.51) = 0.5 - 0.4345 = 0.0655$$

Hoofdstuk 10

Herhalingsoefeningen

1. Een zak bevat 10 rode, 6 groene en 4 blauwe bollen. Men trekt gelijktijdig 7 bollen uit de zak. Wat is de kans dat er minstens 1 groene en minstens 1 blauwe bal getrokken wordt?

Oplossing

Twee gebeurtenissen:

1. M1G: Minstens 1 groene bal getrokken
2. M1B: Minstens 1 blauwe bal getrokken

De kans om minstens 1 groene en 1 blauwe bal te trekken:

$$\begin{aligned}P(M1G \cap M1B) &= P(\overline{M1G} \cup \overline{M1B}) \\&= 1 - P(\overline{M1G} \cup \overline{M1B}) \\&= 1 - P(0G \cup 0B) \\&= 1 - [P(0G) + P(0B) - P(0G \cap 0B)] \\&= 1 - \left[\frac{C_4^0 C_{16}^7}{C_{20}^7} + \frac{C_6^0 C_{14}^7}{C_{20}^7} - \frac{C_4^0 C_6^0 C_{10}^7}{C_{20}^7} \right] \\&= 1 - \left[\frac{C_{16}^7}{C_{20}^7} + \frac{C_{14}^7}{C_{20}^7} - \frac{C_{10}^7}{C_{20}^7} \right] \\&= 1 - \frac{1}{C_{20}^7} \left(C_{16}^7 + C_{14}^7 - C_{10}^7 \right)\end{aligned}$$

10. Stel dat $x : N(0, 3)$ verdeeld, $y : N(2, 3)$ verdeeld en $z : \chi^2(7d.f.)$ verdeeld zijn (x, y en z onafhankelijk). Bepaal

1. de kritische waarde a zodat $P((y - 2)^2 > a^2 z) = 0.9$
2. $P(x^2 + y^2 + 9z > 4y + 26)$

Oplossing

1. We weten dat $y : N(2, 3)$, dus $\frac{y-2}{3} : N(0, 1)$. Per definitie volgt dat $\frac{(y-2)^2}{9} : \chi^2(1d.f.)$. Aangezien zowel $\frac{(y-2)^2}{9}$ en z χ^2 verdeeld zijn, kan hieruit de Fischer verdeling afgeleid worden door:

$$\frac{\frac{(y-2)^2}{9}}{\frac{1}{z/7}} = \frac{7}{9} \cdot \frac{(y-2)^2}{z} \text{ is } F(1, 7d.f.)$$

De kans wordt:

$$0.9 = P((y-2)^2 > a^2 z) = P\left(\frac{7}{9} \cdot \frac{(y-2)^2}{z} > \frac{7}{9} a^2\right)$$

Hieruit volgt dat $\frac{7}{9} a^2 = F_{0,9}(1, 7d.f.) = \frac{1}{F_{0,1}(7, 1d.f.)} = \frac{1}{58,91}$. Hieruit volgt dat $a = \sqrt{\frac{9}{7 \cdot 58,91}} = \pm 0.15$

2. We moeten van zowel x , als y een χ^2 verdeling maken zodat we deze kunnen optellen met z .

$$x : N(0, 3) \rightarrow \frac{x-0}{3} : N(0, 1) \rightarrow \frac{x^2}{9} : \chi^2(1d.f.)$$

We weten al van de vorige opgave dat $\frac{(y-2)^2}{9}$ een χ^2 verdeling is (1 d.f.).

We passen de lineaire combinatie toe:

$$\frac{x^2}{9} + \frac{(y-2)^2}{9} + z \text{ is } \chi^2(1 + 1 + 7df)$$

De kans wordt:

$$\begin{aligned} P(x^2 + y^2 + 9z > 4y26) &= P(x^2 + y^2 - 4y + 4 + 9z > 26 + 4) \\ &= P(x^2 + (y-2)^2 + 9z > 30) \\ &= P\left(\frac{x^2}{9} + \frac{(y-2)^2}{9} + z > \frac{30}{9}\right) \\ &= 1 - 0.05 = 0.95 \end{aligned}$$

Uitleg bij $1 - 0.05 = 0.95$. We weten dat we bezig zijn met een chi-kwadraat verdeling met 9 vrijheidsgraden. Als je de waarde $30/9 = 3.33...$ opzoekt op het formularium bij 9 vrijheidsgraden zie je dat deze waarde behoort bij $\chi_{0,05}^2$. Dus de oppervlakte links van $\frac{30}{9}$ zou 0.05 zijn, maar we willen de oppervlakte rechts van $\frac{30}{9}$ dus we doen $1 - 0.05 = 0.95$

11. Het aantal afgestudeerden industrieel ingenieur is normaal verdeeld met gemiddelde 400 en standaarddeviatie 40. Het aantal arbeidsplaatsen dat voor hen beschikbaar is, is ook normaal verdeeld met gemiddelde 450 en standaarddeviatie 20. Bepaal de kans dat er studenten zijn die geen job vinden.

Oplossing

Stel:

- s : # afgestudeerden
- s is normaal verdeeld $\rightarrow s : N(400, 40)$
- p : # plaatsen
- p is normaal verdeeld $\rightarrow p : N(450, 20)$

Het aantal studenten die geen job vinden kan voorgesteld worden als het aantal plaatsen min het aantal studenten. Vermits deze twee normaal verdeeld zijn is de resulterende verdeling ook

normaal verdeeld. Stel x het aantal studenten die geen job vinden dan is $x = p - s$ en $x : N(450 - 400, \sqrt{40^2 + 20^2}) = N(50, 20\sqrt{5})$. De kans wordt:

$$\begin{aligned} P(x \leq 0) &= P\left(\frac{x - 50}{20\sqrt{5}} \leq \frac{0 - 50}{20\sqrt{5}}\right) \\ &= P(z \leq -1.12) \\ &= 0.5 - 0.3686 \\ &= 0.1314 \end{aligned}$$

14. Stel dat de oefeningenpunten (op 20) voor statistiek normaal verdeeld zijn. Als 5% van de studenten meer dan 15 behaalt en 20% minder dan 7, bereken dan het gemiddelde en de standaarddeviatie.

Oplossing

Stel:

- x : Het aantal punten op de test voor een student
- x is normaal verdeeld $\rightarrow x : N(\mu, \sigma)$
- $P(x > 15) = 0.05$
- $P(x < 7) = 0.2$

We kunnen beide kansen normeren en dan op het formularium opzoeken voor welke waarde van a deze kans waar is. We krijgen dan twee vergelijkingen in μ en σ .

$$0.05 = P\left(\frac{x - \mu}{\sigma} > \frac{15 - \mu}{\sigma}\right) = P\left(z > \frac{15 - \mu}{\sigma}\right) = 0.5 - P\left(0 < z < \frac{15 - \mu}{\sigma}\right)$$

We weten dus dat $P\left(0 < z < \frac{15 - \mu}{\sigma}\right)$ gelijk moet zijn aan 0.45 zodat deze vergelijking klopt. We moeten de waarde a vinden zodat $P(0 < z < a) = 0.45$, daarna stellen we $a = \frac{15 - \mu}{\sigma}$. De waarde 0.45 staat echter niet op het formularium. De waarden 0.4505 en 0.4495 staan er wel op dus we zeggen dat $0.45 = \frac{0.4505 + 0.4495}{2}$. De kritische waarde 0.4505 opzoeken komt neer op 1.65 en 0.4495 komt neer op 1.64. Hieruit volgt dat $a = \frac{1.65 + 1.64}{2} = 1.645 = \frac{15 - \mu}{\sigma}$.

Analoog voor de tweede kans:

$$0.2 = P\left(\frac{x - \mu}{\sigma} < \frac{7 - \mu}{\sigma}\right) = P\left(z < \frac{7 - \mu}{\sigma}\right)$$

Via dezelfde werkwijze komen we $\frac{7 - \mu}{\sigma} = -0.84$ uit.

Dit leidt tot het volgende stelsel:

$$\begin{cases} \frac{15 - \mu}{\sigma} = 1.645 \\ \frac{7 - \mu}{\sigma} = -0.84 \end{cases}$$

Met eenvoudig rekenwerk bekomt men $\mu = 9.7$ en $\sigma = 3.27$

16. Om te achterhalen of een persoon een bepaalde ziekte heeft, wordt een bloedtest genomen. Voor de personen die inderdaad ziek zijn, detecteert de bloedtest in 99% van de gevallen de ziekte; echter voor de personen die niet ziek zijn blijkt te bloedtest in 3% van de gevallen wel ten onrechte

de ziekte te detecteren. Als je weet dat 1% van de bevolking de ziekte heeft, bepaal dan de kans dat een persoon de ziekte heeft in geval de bloedtest dit aangeeft.

Oplossing

Twee gebeurtenissen:

1. PZ: De persoon is ziek
2. BZ: De bloedtest toont ziek

De kans dat een persoon ziek is indien de bloedtest dit aangeeft:

$$\begin{aligned} P(PZ|BZ) &= \frac{P(PZ)P(BZ|PZ)}{P(PZ)P(BZ|PZ) + P(PZ)P(\overline{BZ}|PZ)} \\ &= \frac{\frac{1}{100} \cdot \frac{99}{100}}{\frac{1}{100} \cdot \frac{99}{100} + \frac{99}{100} \cdot \frac{3}{100}} \\ &= \frac{\frac{99}{100^2}}{\frac{99}{100^2} + \frac{297}{100^2}} \\ &= \frac{99}{396} \end{aligned}$$
