Oefeningen Statistiek 2de Kandidatuur Industrieel Ingenieur (Hogeschool Gent)

5 juni 2003

Inhoudsopgave

About

Dit document bevat mijn nota's en mijn uitgewerkte oefeningen voor het vak "Statistiek", gedoceerd in de tweede kandidatuur industrieel ingenieur optie informatica.
De auteur is niet verantwoordelijk voor enige onvolmaaktheden/tekortkomingen in dit document. Gemaakt met emacs in LaTeX onder Debian Linux. Bronnen beschikbaar op aanvraag, voor meer informatie:
elie@de-brauwer.be.

1 Kansrekenen

1. Bewijs de gelijkheid: $P\left(\left(A \wedge B\right)/C\right) = P\left(A/\left(B \wedge C\right)\right) \cdot P\left(B/C\right)$.
   We vertrekken van het linkerlid:

$$\begin{aligned}
P\left(\left(A \wedge B\right)/C\right) &= \frac{P\left(A \wedge B \wedge C\right)}{P\left(C\right)} \\
&= \frac{P\left(A/\left(B \wedge C\right)\right)}{P(C)} \\
&= \frac{P\left(A/\left(B \wedge C\right)\right) \cdot P\left(B \wedge C\right)}{P(C)} \\
&= \frac{P\left(A/\left(B \wedge C\right)\right) \cdot P\left(B/C\right) \cdot P(C)}{P(C)} \\
&= P\left(A/\left(B \wedge C\right)\right) \cdot P\left(B/C\right)
\end{aligned}$$

   Vermits het linkerlid kan herleid worden tot het rechterlid volgt hieruit de gelijkheid[1].

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 twee keer munt te hebben ?
   We stoppen $P(kop) = 2 \cdot P(munt)$ in $\sum P = 1$.

$$\begin{aligned}
P(kop) + \cdot P(munt) &= 1 \\
3 \cdot P(munt) &= 2
\end{aligned}$$

   Hieruit volgt dat $P(munt) = \frac{1}{3}$ en $P(kop) = \frac{2}{3}$.

$$\begin{aligned}
P(juisttweekeermunt) &= P(kmm) + P(mkm) + P(mmk) \\
&= 3 \cdot \left(P(k) \cdot P(m) \cdot P(m)\right) \\
&= 3 \cdot \frac{2}{27} \\
&= \frac{2}{9}
\end{aligned}$$

3. Een dobbelsteen is vervalst zodat de kans dat een gegeven aantal ogen geworpen wordt evenreedig is met het aantal ogen. Is A de gebeurtenis een even getal te gooien, B de gebeurtenis een priemgetal te gooien (1 is géén priemgetal) en C de gebeurtenis een oneven getal te gooien,

   (a) Bepaal $P(A)$, $P(B)$ en $P(C)$.

$$\sum P = 1$$

[1] De omzettingen zijn gebaseerd op de vermenigvuldigingswet waaruit volgt dat $P(B/A) = \frac{P(A \wedge B)}{P(A)}$

$$\begin{aligned} P(oog) + 2 \cdot P(oog) \ldots 6 \cdot P(oog) &= 1 \\ 21 \cdot P(oog) &= 1 \\ P(oog) &= \frac{1}{21} \end{aligned}$$

- $P(A)$

$$\begin{aligned} P(A) &= P(2) + P(4) + P(6) \\ &= P(oog)\,(2 + 4 + 6) \\ &= \frac{12}{21} \\ &= \frac{4}{7} \end{aligned}$$

- $P(B)$

$$\begin{aligned} P(B) &= P(2) + P(3) + P(5) \\ &= P(oog)\,(2 + 3 + 5) \\ &= \frac{10}{21} \end{aligned}$$

- $P(C)$

$$\begin{aligned} P(C) &= P(1) + P(3) + P(5) \\ &= \frac{3}{7} \\ &= 1 - P(A) \\ &= P(\overline{A}) \end{aligned}$$

(b) Bereken de kans $P(D)$ dat men een even getal of een priemgetal gooit .

$$\begin{aligned} P(D) &= P(A \vee B) \\ &= P(A) + P(B) - P(A \wedge B) \\ &= \frac{4}{7} + \frac{10}{21} - P(2) \\ &= \frac{20}{21} \\ &= 1 - P(1) \\ &= P(\overline{1}) \end{aligned}$$

(c) Bereken de kans $P(E)$ dat men een even getal gooit dat geen priemgetal is.

$$\begin{aligned} P(E) &= P(4) + P(6) \\ &= \frac{10}{21} \\ &= P(A \wedge \overline{B}) \end{aligned}$$

(d) Bereken de kans $P(F)$ dat men een oneven getal of een priemgetal gooit.

$$\begin{aligned} P(F) &= P(B) \vee P(C) \\ &= P(B) + P(C) - P(B \wedge C) \\ &= \frac{10}{21} + \frac{3}{7} - \frac{8}{21} \\ &= \frac{11}{21} \end{aligned}$$

Deze stap is eveneens gelijk aan $P(\overline{E})$. Want deze zijn tot elkaar te herleiden mits te steunen op de wetten van de Morgan.

4. A en B zijn verschijnselen met $P(A) = 0.1)$,$P(B) = 0.5$. Bepaal,

(a) Wanneer beide verschijnselen elkaar uitsluiten.

i. $P(A \vee B)$

$$\begin{aligned} P(A \vee B) &= P(A) + P(B) - P(A \wedge B) \\ &= P(A) + P(B) - 0 \\ &= 0.6 \end{aligned}$$

ii. $P(\overline{A})$

$$\begin{aligned} P(\overline{A}) &= 1 - P(A) \\ &= 0.9 \end{aligned}$$

iii. $P(\overline{A} \wedge B)$

$$\begin{aligned} P(\overline{A} \wedge B) &= P(B) \\ &= 0.5 \end{aligned}$$

(b) Wanneer beide verschijnselen onafhankelijk zijn.

i. $P(A \vee B)$

$$\begin{aligned} P(A \vee B) &= P(A) + P(B) - P(A \wedge B) \\ &= P(A) + P(B) - P(A) \cdot P(B) \\ &= 0.55 \end{aligned}$$

ii. $P(\overline{A})$

$$\begin{aligned} P(\overline{A}) &= 1 - P(A) \\ &= 0.9 \end{aligned}$$

iii. $P(\overline{A} \wedge B)$

$$\begin{aligned} P(\overline{A} \wedge B) &= P(\overline{A}) \cdot P(B) \\ &= 0.45 \end{aligned}$$

5. De verschijnselen A,B en C sluiten[2] elkaar twee per twee uit. Daarbij is $P(A) = 0.2$, $P(B) = 0.1$ en $P(C) = 0.4$. Bepaal,

   (a) $P(A \vee B \vee C)$

   $$\begin{aligned} P(A \vee B \vee C) &= P(A) + P(B) + P(C) \\ &= 0.2 + 0.1 + 0.4 \\ &= 0.7 \end{aligned}$$

   (b) $P\left(\overline{A} \wedge (B \vee C)\right)$

   $$\begin{aligned} P\left(\overline{A} \wedge (B \vee C)\right) &= P(B \vee C) \\ &= 0.1 + 0.4 \\ &= 0.5 \end{aligned}$$

   (c) $P\left(\overline{B \vee \overline{C}}\right)$

   $$\begin{aligned} P\left(\overline{B \vee \overline{C}}\right) &= 1 - P(B \vee \overline{C}) \\ &= 1 - \left(P(B) + P(\overline{C}) - P(B \wedge \overline{C})\right) \\ &= 1 - (0.1 + 0.6 - 0.1) \\ &= 0.4 \end{aligned}$$

6. Bereken voor een familie van drie kinderen de kans[3] op

   (a) drie jongens

   $$\begin{aligned} P(A) &= P(j) \cdot P(j) \cdot P(j) \\ &= \frac{1}{8} \end{aligned}$$

   (b) twee jongens en 1 meisje

   $$\begin{aligned} P(B) &= P(jjm) + P(jmj) + P(mjj) \\ &= 3 \cdot P(jjm) \\ &= \frac{3}{8} \end{aligned}$$

[2] Dit wil zeggen dat de onderlinge deelverzamelingen van de kansen ledig zijn

[3] Stel dat de kans op een jongen gelijk is aan de kans op een meisje

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

$$\begin{aligned} P(A) &= 1 - (P(1,1->6) + P(2,2->5) + P(3,3->4)) \\ &= 1 - \frac{1}{36} \cdot (11+7+3) \\ &= 1 - \frac{21}{36} \\ &= \frac{5}{12} \end{aligned}$$

8. Een eerste zak bevat vier witten en vijf zwarte ballen en een tweede zak bevat zes witte en vier zwarte ballen. Een bal wordt uit de eerst zak getrokken en ongezien in de tweede gelegd.

   (a) Wat is de kans $P(A)$ dat een bal, nu getrokken uit de tweede zak, zwart is ?

   - Algemeen
     De kans dat $P(B1 = W) = \frac{4}{9}$, dit is ook de kans dat er 7 witte en 4 zwarte ballen in de tweede zak komen te zitten. Anderzijds is de kans dat $P(B2 = Z) = \frac{5}{9}$, eveneens de kans dat er 6 witte en 5 zwarte ballen in de tweede zak komen te zitten.
   - Op een intuïtieve methode.
     Vermits in elke zak zeker vier zwarte ballen zitten heb je zowieso $\frac{4}{11}$ kans dat je een zwarte hebt, nu wordt deze kans vermeerd met de kans dat de extra bal eveneens zwart was, dus:

     $$\frac{4}{11} + \frac{5}{9} \cdot \frac{1}{11} = \frac{41}{99}$$

   - Op een statistisch meer verantwoorde methode.

     $$\begin{aligned} P(B2 = Z) &= P((B2 = Z) \wedge (B1 = W)) \\ &\quad + P((B2 = Z) \wedge (B1 = Z)) \\ &= P((B2 = Z)/B1 = W) \cdot P(B1 = W) \\ &\quad + P((B2 = Z)/B1 = Z) \cdot P(B1 = Z) \\ &= \frac{4}{11} \cdot \frac{4}{9} + \frac{5}{11} \cdot \frac{5}{9} \\ &= \frac{41}{99} \end{aligned}$$

   (b) Als de bal getrokken uit de tweede zak zwart is, wat is de kans dat de eerstgetrokken bal wit was ?

Gebaseerd op $P(A \wedge B) = P(A/B) \cdot P(B) = P(B/A) \cdot P(A)$.

$$\begin{aligned} P(B1 = W/B2 = Z) &= \frac{P\left((B1 = W) \wedge (B2 = Z)\right)}{P\left(B2 = Z\right)} \\ &= \frac{\frac{4}{11} \cdot \frac{4}{9}}{\frac{41}{99}} \\ &= \frac{16}{41} \end{aligned}$$

Dit vraagstuk is ook oplosbaar via de Regel van Bayes.

9. Voor een koppel (1 man, 1 vrouw) is de kans dat de man een bepaalde televisieshow bekijkt 0.4, de kans dat de vrouw kijkt is 0.5. De kans dat de man kjikt, als de vrouw kijkt is 0.7. Bepaal:

   - De kans dat het koppel de show bekijkt.

     $$\begin{aligned} P(M \wedge V) &= P(V) \cdot P(M/V) \\ &= 0.5 \cdot 0.7 \\ &= 0.35 \end{aligned}$$

   - De kans dat de vrouw kijkt als de man kijkt.

     $$\begin{aligned} P(V/M) &= \frac{P(V \wedge M)}{P(M)} \\ &= \frac{0.35}{0.4} \\ &= 0.875 \end{aligned}$$

   - De kans dat minstens één van beiden kijkt.

     $$\begin{aligned} P(M \vee V) &= P(M) + P(V) - P(V \wedge M) \\ &= 0.4 + 0.5 - 0.35 \\ &= 0.55 \end{aligned}$$

10. De kans dat een man na 25 jaar samenwonen nog zal leven is $\frac{3}{5}$, voor zijn vrouw is dat $\frac{2}{3}$. Bereken de kans dat na 25 jaar (in de veronderstelling dat het verschijnsels 'in leven blijven' onafhankelijk[4] zijn van elkaar)[5]:

    - Beiden nog leven.

      $$P(ML \wedge VL) = P(VL/ML) \cdot P(ML)$$

[4] Dus $P(A/B) = P(A)$

[5] ML=man leven, MD=man dood, VL=vrouw leven, VD=vrouw dood

$$
\begin{aligned}
&= P(VL) \cdot P(ML) \\
&= \frac{2}{3} \cdot \frac{3}{5} \\
&= \frac{2}{5}
\end{aligned}
$$

- Alleen de man nog leeft.

$$
\begin{aligned}
P(ML \wedge VD) &= P(ML/VD) \cdot P(VD) \\
&= P(ML) \cdot P(VD) \\
&= \frac{3}{5} \cdot \frac{1}{3} \\
&= \frac{1}{5}
\end{aligned}
$$

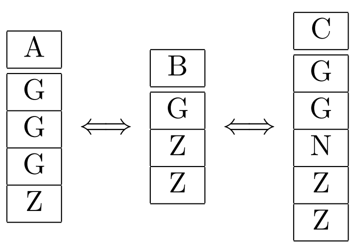
- Alleen de vrouw nog leeft.

$$
\begin{aligned}
P(VL \wedge MD) &= P(VL/MD) \cdot P(MD) \\
&= P(VL) \cdot P(MD) \\
&= \frac{2}{3} \cdot \frac{2}{5} \\
&= \frac{4}{15}
\end{aligned}
$$

- Ten minste één van beide nog leeft.

$$
\begin{aligned}
P(VL \vee ML) &= P(VL) + P(ML) - P(VL \wedge ML) \\
&= \frac{2}{3} + \frac{3}{5} - \frac{2}{5} \\
&= \frac{13}{15}
\end{aligned}
$$

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



Men kiest willekeurig één van de katen, opent daarvan at random één lade, en grijpt het muntstuk (indien mogelijk).

Wat is de kans dat men kast A heeft uigekozen indien men een goudstuk genomen heeft ?

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

12. Gegeven: een kast met vijf laden die al dan niet munstuk bevatten (G = goudstuk, Z = zilverstuk, N = niets), volgens onderstaande figuur.

G
G
N
Z,G
Z,G,G

Indien met willekeurig een lade opentrekt en daaruit willekeurig een muntstuk neemt (indien mogelijk), bereken dan de kans om een goudstuk te nemen.

$$
\begin{aligned}
P(G) &= P(G/S1) \cdot P(S1) + P(G/S2) \cdot P(S2) + P(G/S3) \cdot P(S3) \\
&\quad + P(G/S4) \cdot P(S4) + P(G/S5) \cdot P(S5) \\
&= 1 \cdot \frac{1}{5} + 1 \cdot \frac{1}{5} + 0 + \frac{1}{2} \cdot \frac{1}{5} + \frac{2}{3} \cdot \frac{1}{5} \\
&= \frac{2}{5} + \frac{1}{10} + \frac{2}{15} \\
&= \frac{19}{30}
\end{aligned}
$$

13. Twee onvervalste dobbelstenen worden gegooid. Wat is de kans dat beide vlakken een even aantal ogen vertonen, als de som van de ogen 8 is?
Er kan 8 op 5 manieren[6](nl. $\{(6,2),(2,6),(5,3),(3,5),(4,4)\}$) gegooid worden, dus de kans dat beide bovenvlakken even zijn is simpelweg $\frac{3}{5}$.

[6]Het geval waar beide dobbelstenen met vier naar boven liggen wordt maar 1 maal geteld

14. Als A en B verschijnselen zijn met $P(\overline{A}) = 0.3$,$P(B) = 0.4$ en $P(A \wedge \overline{B}) = 0.5$, bereken dan:

    - $P(A)$

    $$\begin{aligned} P(A) &= 1 - P(\overline{A}) \\ &= 0.7 \end{aligned}$$

    - $P(A \wedge B)$
    Eerst leiden we nog enkele waarden af uit de gegevens:

    $$\begin{aligned} P(A \wedge \overline{B}) &= P(A/\overline{B}) \cdot P(\overline{B}) \\ P(A/\overline{B}) &= \frac{P(A \wedge \overline{B})}{P(\overline{B})} = \frac{5}{6} \end{aligned}$$

    Of ook

    $$\begin{aligned} P(A \wedge \overline{B}) &= P(\overline{B}/P(A)) \cdot P(A) \\ P(\overline{B}/A) &= \frac{P(A \wedge \overline{B})}{P(A)} = \frac{5}{7} \\ P(B/A) &= 1 - P(\overline{B}/A) = \frac{2}{7} \end{aligned}$$

    Dus:

    $$\begin{aligned} P(A \wedge B) &= P(B/A) \cdot P(A) \\ &= \frac{2}{7} \cdot \frac{7}{10} \\ &= \frac{2}{10} \end{aligned}$$

    - $P(A \vee B)$

    $$\begin{aligned} P(A \vee B) &= P(A) + P(B) - P(A \wedge B) \\ &= 0.7 + 0.4 - 0.2 \\ &= 0.9 \end{aligned}$$

15. Een urne bevat 7 rode en 3 zwarte ballen. Een bal wordt at random uit de urne getrokken en vervangen door een bal van de andere kleur. Er wordt een tweede bal getrokken. Als de tweede bal rood is, wat is dan de (voorwaardelijke) kans dat de eerstgetrokken bal ook rood was?[7]

    $$P(1R/2R) = \frac{P(1R \wedge 2R)}{P(2R)}$$

[7] Regel van Bayes: $P(A_j/B) = \frac{P(A_j) \cdot P(B/A_j)}{\sum_{n=1}^{n} P(A_i) \cdot P(B/A_i)}$

$$\begin{aligned}
&= \frac{P(2R/1R) \cdot P(1R)}{P(2R/1R) \cdot P(1R) + P(2R/1Z) \cdot P(1Z)} \\
&= \frac{\frac{6}{10} \cdot \frac{7}{10}}{\frac{6}{10} \cdot \frac{7}{10} + \frac{8}{10} \cdot \frac{3}{10}} \\
&= \frac{42}{66} = \frac{7}{11}
\end{aligned}$$

16. Wanneer men een 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 al het waarheidsserum schuldig aanwijst ?
Afleidingen uit de gegevens:

$$\begin{aligned}
W &= \text{Serum wijst schuldig} \\
S &= \text{Persoon is schuldig} \\
P(W/S) &= 90\% \\
P(\overline{W}/\overline{S}) &= 99\% \\
P(S) = 5\%
\end{aligned}$$

Hieruit volgt dat:

$$\begin{aligned}
P(\overline{S}/W) &= \frac{P(\overline{S} \wedge W)}{P(W)} \\
&= \frac{P(W/\overline{S}) \cdot P(\overline{S})}{P(W/\overline{S}) \cdot P(\overline{S}) + P(W/S) \cdot P(S)} \\
&= \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.05 \cdot 0.90} \\
&= \frac{95}{545} = \frac{19}{109}
\end{aligned}$$

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

- Het vijf zwarte kaarten zijn.

$$\begin{aligned}
P(5Z) &= \frac{C_{26}^5}{C_{52}^5} \\
&= 0.02531
\end{aligned}$$

- Het drie heren en twee vrouwen zijn.

$$\begin{aligned}
P(3H2V) &= \frac{C_4^3 \cdot C_4^2}{C_{52}^5} \\
&= 0.0000092344
\end{aligned}$$

- Er tenminste één aas bij is.

$$\begin{aligned} P(1A) &= 1 - P(\overline{A}) \\ &= 1 - \frac{C_{48}^{5}}{C_{52}^{5}} \\ &= 0.341158 \end{aligned}$$

- Er ten hoogste één harten bij is.

$$\begin{aligned} P(\leq 1\heartsuit) &= P(1\heartsuit) + P(\overline{\heartsuit}) \\ &= \frac{C_{13}^{1} \cdot C_{39}^{4}}{C_{52}^{5}} + \frac{C_{39}^{5}}{C_{52}^{5}} \\ &= 0.41142 + 0.221534 \\ &= 0.632953 \end{aligned}$$

18. Uit een spel van 32 kaarten trekt men willekeurig 8 kaarten, bereken de kans dat onder deze kaarten juist 4 van dezelfde soort zijn. Samenstelling van deze 32 kaarten ? C_{52}^{32} ?

19. De waarschijnlijkheid voor het sluiten van de $i$de relais in de onderstaande circuits is gegeven door $P_i, i = 1, 2, 3, 4, 5$. Indien alle relais onafhankelijk functioneren, wat is dan de waarschijnlijkheid dat er stroom vloeit tussen A en B ? Ze functioneren onafhankelijk dus $P(a) \wedge P(b) = P(A) \cdot P(B)$

$$\begin{aligned} P(stroom) &= P\left(\left(\left(P\left(1\right) \wedge P\left(2\right)\right) \vee \left(P\left(3\right) \wedge P\left(4\right)\right)\right) \cdot P\left(5\right)\right) \\ &= \left(p_1p_2 + p_3p_4 - p_1p_2p_3p_4\right) \cdot p_5 \end{aligned}$$

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

- Minsten éénmaal zes

$$\begin{aligned} P(minstens1keer6) &= 1 - P(\overline{6}) \\ &= 1 - \left(\frac{5}{6}\right)^4 \\ &= 0.517747 \end{aligned}$$

- Minstens dubbel zes

$$\begin{aligned} P(dubbel6) &= 1 - P(\overline{dubbelzes}) \\ &= 1 - \left(\frac{35}{36}\right)^{24} \\ &= 0.491404 \end{aligned}$$

21. Bepaal de kans om met de Belgische Lotto:

    - Drie cijfers goed te hebben

    $$\begin{aligned} P(3) &= \frac{C_6^3 \cdot C_{36}^3}{C_{42}^6} \\ &= 0.027222 \end{aligned}$$

    - Vier cijfers goed te hebben

    $$\begin{aligned} P(4) &= \frac{C_6^4 \cdot C_{36}^2}{C_{42}^6} \\ &= 0.001801 \end{aligned}$$

    - Zes cijfers goed te hebben

    $$\begin{aligned} P(6) &= \frac{C_6^6}{C_{42}^6} \\ &= 0.000000190 \end{aligned}$$

22. Een eerste urne bevat vier rode en een zwarte bal. Een tweede bevat twee witte ballen. Drie ballen worden at random getrokken uit de eerste urne en ongezien in de tweede gelegd. Daarna worden vier ballen at arandom uit de tweede urne getrokken en in de eerste gelegd. Bereken de kans dat na deze operatie de zwarte bal in de eerste urne ligt.

    $$\begin{aligned} P(zwartineerste) &= 1 - P(zwartintweede) \\ &= 1 - P((1 = Z) \wedge (2 \neq Z)) \\ &= 1 - P(1 = Z) \cdot P((2 \neq Z)/(1 = Z)) \\ &= 1 - \frac{C_1^1 \cdot C_4^2}{C_5^3} \cdot \frac{C_4^4}{C_5^4} \\ &= 0.88 \end{aligned}$$

23. Een urne bevat 10 ballen waarvan n rode. De kans om twee rode ballen te trekken op drie trekkinngen 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.

    $$3 \cdot \frac{n}{10} \cdot \frac{n}{10} \cdot \frac{10-n}{10} = 1.08 \cdot \frac{C_n^2 \cdot C_{10-n}^1}{C_{10}^3}$$

    Dit is een verelijking met 1 onbekende (namelijk n), na uitwerking volgt dat $n = 3$.

2 Beschrijvende Statistiek

Geen oefeningen op dit hoofdstuk

3 Kans-, dichtheids- en verdelingsfunctie van een populatie

1. Bepaal voor $\frac{x-\mu}{\sigma}$ alsook de verwachtingswaarde $E\left[\frac{x-\mu}{\sigma}\right]$ en de variantie $V\left[\frac{x-\mu}{\sigma}\right]$.

   - $E\left[\frac{x-\mu}{\sigma}\right]$

$$
\begin{aligned}
E\left[g(x)\right] &= \int g(x)f(x)dx \\
&= \int \frac{x-\mu}{\sigma}f(x)dx \\
&= \int \frac{x}{\sigma}f(x)dx - \int \frac{\mu}{\sigma}f(x)dx \\
&= \frac{1}{\sigma}\int xf(x)dx - \frac{\mu}{\sigma}\int f(x)dx \\
&= \frac{\mu}{\sigma} - \frac{\mu}{\sigma} \\
&= 0
\end{aligned}
$$

   - $E\left[\frac{x-\mu}{\sigma}\right]$ op een tweede manier

$$
\begin{aligned}
E\left[\frac{x-\mu}{\sigma}\right] &= E\left[\frac{x}{\sigma} - \frac{\mu}{\sigma}\right] \\
&= \frac{1}{\sigma}E\left[x\right] - E\left[\frac{\mu}{\sigma}\right] \\
&= \frac{\mu}{\sigma} - \frac{\mu}{\sigma} \\
&= 0
\end{aligned}
$$

   Gebaseerd op:

$$
\begin{aligned}
E[kx] &= kE[x] \\
E[x+y] &= E[x] + E[y]
\end{aligned}
$$

   - $V\left[\frac{x-\mu}{\sigma}\right]$

$$
\begin{aligned}
V\left[\frac{x-\mu}{\sigma}\right] &= V\left[\frac{x}{\sigma}\right] \\
&= \frac{1}{\sigma^2}V[x] \\
&= \frac{\sigma^2}{\sigma^2} \\
&= 1
\end{aligned}
$$

Gebaseerd op:

$$\begin{aligned} V[kx] &= k^2E[x] \\ V[x+a] &= V[x] \end{aligned}$$

2. Een continue toevalsveranderlijke x neemt waarden aan tussen 2 en 5. De dichtheidsfunctie is gegeven door $f(x) = a(1+x)$. Bepaal:

   - de constante a

$$\begin{aligned} x \in [2,5] &\Leftrightarrow \int_2^5 f(x)dx = 1 \\ \int_2^5 f(x)dx &= \int_2^5 (a + ax)\, dx \\ &= \left[ax + \frac{ax^2}{2}\right]_2^5 \\ &= a\left(5 + \frac{25}{2} - 2 - \frac{4}{2}\right) \\ &= a\frac{27}{2} \end{aligned}$$

   Vermits $a\frac{27}{2} = 1$ moet $a = \frac{2}{27}$

   - $P(x < 4)$

$$\begin{aligned} P(x<4) &= \frac{2}{27}\int_2^4 (1+x)dx \\ &= \frac{2}{27}\left[\frac{(1+x)^2}{2}\right]_2^4 \\ &= \frac{2}{27}\left(\frac{25}{2} - \frac{9}{2}\right) \\ &= \frac{16}{27} \end{aligned}$$

   - de cumulatieve dichtheidsfunctie $F(x)$

$$\begin{aligned} F(x) &= \frac{2}{27}\int_2^x (1+x)dx \\ &= \frac{2}{27}\left[\frac{(1+x)^2}{2}\right]_2^x \\ &= \frac{2}{27}\left(\frac{(1+x)^2}{2} - \frac{3^2}{2}\right) \\ &= \frac{2}{27}\left(\frac{x^2}{2} + x - 4\right) \end{aligned}$$

Deze $F(x)$ geldt enkel binnen het opgegeven in het interval, voor dit interval is $F(x)$ gegeven door $y = 0$ en erna is $y = 1$.

- de gemiddelde waarde, modus en mediaan
  - gemiddelde waarde

$$\begin{aligned}\mu &= \int_{-\infty}^{+\infty} x \cdot f(x)dx \\ &= \frac{2}{27}\int_2^5 \left(x^2 + x\right) dx \\ &= \frac{2}{27}\left[\frac{x^3}{3} + \frac{x^2}{2}\right]_2^5 \\ &= \frac{11}{3}\end{aligned}$$

  - de modus
    De modus is het maximum van de dichtheidsfunctie, in dit geval treedt het maximum op voor een waarde van $x = 5$
  - Mediaan

$$\begin{aligned}F(Me) &= \frac{1}{2} \\ \frac{1}{2} &= \frac{2}{27}\left(\frac{x^2}{2} + x - 4\right)\end{aligned}$$

    Een van de oplossingen van deze vierkantsvergelijking is te verwerpen omdat deze niet in het het domein ligt waarover x continu is, in dit geval is $Me = 3.74342$

- de variantie

$$\begin{aligned}\sigma^2 &= \mu_2 - \mu^2 \\ &= \int x^2 f(x)dx - \mu^2 \\ &= \int_{-\infty}^{+\infty} x^2 \cdot f(x)dx - \frac{121}{9} \\ &= \frac{13}{8}\end{aligned}$$

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

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

Deze functie is een cumulatieve distributiefunctie want in limiet streeft deze naar 1 en de afgeleide is steeds groter (zie onder) dan 0 in het interval dat we hier behandelen. Deze dichtheidsfunctie is gegeven door de afgeleide van $F(x)$:

$$F'(x) = f(x) = \begin{cases} 0 & x \leq 0 \\ 2xe^{-x^2} & x > 0 \end{cases}$$

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

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{2} & 1 \leq x < 0 \\ \frac{5}{6} & 2 \leq x < 3 \\ 1 & 3 \leq x \end{cases}$$

In limiet gaat deze functie naar 1, deze functie is niet-dalend over het gehele interval en deze functie is ook altijd groter dan of gelijk aan nul, de corresponderende dichtheidsfunctie f(x) wordt dan:

$$F'(x) = f(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{2} & 1 \leq x < 0 \\ \frac{1}{3} & 2 \leq x < 3 \\ \frac{1}{6} & 3 \leq x \end{cases}$$

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

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

We bepalen C zodat de bepaalde integraal 1 is.

$$\begin{aligned} \int_0^2 C(4x - 2x^2) &= 1 \\ \left[\frac{2Cx^2}{2} - \frac{2Cx^3}{3}\right]_0^2 &= 1 \\ 8C - \frac{16}{3}C &= 1 \\ C &= \frac{3}{8} \end{aligned}$$

De bijhorende cumulatieve distributiefunctie is dan gegeven door:

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{3}{4}(x^2 - \frac{x^3}{3}) & 0 < x < 2 \\ 1 & x \geq 2 \end{cases}$$

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:

   - de dichtheidsfunctie van x

     $$\begin{aligned} f(0) &= P(x=0) \\ &= \frac{6\cdot 5\cdot 4}{10\cdot 9\cdot 8} \\ &= \frac{1}{6} \end{aligned}$$

     $$\begin{aligned} f(1) &= P(x=1) \\ &= 3\cdot\frac{6\cdot 5\cdot 4}{10\cdot 9\cdot 8} \\ &= \frac{1}{2} \end{aligned}$$

     Hier deed men maal drie omdat deze situatie op drie mogelijke manieren gevormd kan worden nl. RWW,WRW en WWR. Bij het geval $P(x=2)$ is dit om een analoge reden.

     $$\begin{aligned} f(2) &= P(x=2) \\ &= 3\cdot\frac{6\cdot 4\cdot 3}{10\cdot 9\cdot 8} \\ &= \frac{3}{10} \end{aligned}$$

     $$\begin{aligned} f(3) &= P(x=3) \\ &= \frac{4\cdot 3\cdot 2}{10\cdot 9\cdot 8} \\ &= \frac{1}{30} \end{aligned}$$

   - de grafiek van de dichtheidsfunctie en de corresponderende verdelingsfunctie
     Bij het opstellen van de dichtheidsfunctie dient men er rekening mee te houden dat het hier om een discrete toevalsveranderlijke gaat, dus deze grafiek herleidt zich tot 4 puntjes.
   - de gemiddelde waarde en de modus
     Omdat we hier met een discrete toevalsveranderlijke werken is de gemiddelde waarde nu niet gegeven door $\mu = \int x\cdot f(x)dx$ maar door $\mu = \sum x_i f(x_i)$.

     $$\begin{aligned} \mu &= 0\cdot f(0) + 1\cdot f(1) + 2\cdot f(2) + 3\cdot f(3) \\ &= 1\cdot\frac{1}{2} + 2\cdot\frac{3}{10} + 3\cdot\frac{1}{30} \\ &= \frac{12}{10} = 1,2 \end{aligned}$$

De modus is gegeven door het punt waar de kansfunctie de hoogste waarde bereikt, in dit geval is dit P=1.

- de variantie

$$\begin{aligned} \sigma^2 &= \mu_2 - \mu^2 \\ &= \frac{1}{2} + 4 \cdot \frac{3}{10} + 9 \cdot \frac{1}{30} - \frac{144}{100} \\ &= \frac{14}{25} \end{aligned}$$

- $P(x \geq 1)$

$$\begin{aligned} P(x \geq 1) &= 1 - P(x = 0) \\ &= 1 - \frac{1}{6} \\ &= \frac{5}{6} \end{aligned}$$

7. Bepaal de gemiddelde waarde en de variantie van de toevalsveranderlijke x met verdelingsfunctie

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \sqrt{x} & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}$$

- Bepalen dichtheidsfunctie
  Vermits hier de verdelingsfunctie gegeven is moeten we deze verdelingsfunctie eerst afleiden om hieruit de dichtheidsfunctie te kunnen bepalen.

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{2}\frac{1}{\sqrt{x}} & 0 < x \leq 1 \\ 0 & x > 1 \end{cases}$$

- μ

$$\begin{aligned} \mu &= \int_0^1 x \cdot f(x) dx \\ &= \frac{1}{2} \int_0^1 \frac{x}{\sqrt{x}} dx \\ &= \frac{1}{2} \left[\frac{2 \cdot x^{\frac{3}{2}}}{3} \right]_0^1 \\ &= \frac{1}{3} \end{aligned}$$

- σ^2

$$\begin{aligned}
\sigma^2 &= V[x] \\
&= E[x^2] - \mu^2 \\
&= \int_0^1 x^2 \cdot f(x)dx - \mu^2 \\
&= \int_0^1 x^2 \cdot \frac{1}{2\sqrt{x}}dx - \mu^2 \\
&= \frac{1}{2}\left[\frac{2 \cdot x^{\frac{5}{2}}}{5}\right]_0^1 - \mu^2 \\
&= \frac{1}{5} - \frac{1}{9} \\
&= \frac{4}{45}
\end{aligned}$$

8. Speler A gooit met een onvervalste dobbelsteen. Als hij een 6 gooit betaalt hij aan speler B 12 euro. Als hij geen 6 gooit betaalt hij aan speler B 3 euro. Hoeveel betaalt A gemiddeld aan B ?

$$\begin{aligned}
f(12) &= P(x = 12) \\
&= P(6) \\
&= \frac{1}{6}
\end{aligned}$$

$$\begin{aligned}
f(3) &= P(x = 3) \\
&= P(\overline{6}) \\
&= \frac{5}{6}
\end{aligned}$$

Nu bepalen we aan de hand van deze gegevens het gemiddelde:

$$\begin{aligned}
\mu &= 12 \cdot \frac{1}{6} + 3 \cdot \frac{5}{6} \\
&= 4.50
\end{aligned}$$

9. Een speler werpt met 3 munstukken. Hij wint 5 euro als hij 3 keer kop gooit, hij wint 3 euro als hij 2 keer kop voorkomt, hij wint 1 euro bij 1 keer kop en hij verliest 15 euro als bij geen enkel munstuk kop verschijnt. Wat is de gemiddelde winst ? Hoeveel moet de speler betalen bij *geen kop* opdat het spel eerlijk zou zijn ?

$$\begin{aligned}
f(5) &= P(x = 5) \\
&= P(3k) \\
&= \frac{1}{8}
\end{aligned}$$

$$\begin{aligned} f(3) &= P(x=3) \\ &= P(2k) \\ &= \frac{3}{8} \end{aligned}$$

$$\begin{aligned} f(1) &= P(x=1) \\ &= P(k) \\ &= \frac{3}{8} \end{aligned}$$

$$\begin{aligned} f(-15) &= P(x=-15) \\ &= P(0k) \\ &= \frac{1}{8} \end{aligned}$$

Bepalen gemiddelde:

$$\begin{aligned} \mu &= 5 \cdot \frac{1}{8} + 3 \cdot \frac{3}{8} + 1 \cdot \frac{3}{8} - 15 \cdot \frac{1}{8} \\ &= \frac{1}{4} \end{aligned}$$

Nu stellen we x gelijk aan de inzet die we moeten geven opdat het gemiddelde 0 zou zijn.

$$\begin{aligned} 0 &= 5 \cdot \frac{1}{8} + 3 \cdot \frac{3}{8} + 1 \cdot \frac{3}{8} + x \cdot \frac{1}{8} \\ x &= -17 \end{aligned}$$

10. Voor een gokspel met drie onvervalste dobbelstenen bedraagt de inzet steeds 5 euro. Indien iemand juist een 6 werpt krijgt hij zijn inzet (5 euro) 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) ?

$$\begin{aligned} f(0) &= P(6) \\ &= 3 \cdot \frac{1 \cdot 5 \cdot 5}{6 \cdot 6 \cdot 6} \\ &= \frac{25}{72} \end{aligned}$$

$$\begin{aligned} f(5) &= P(2maal6) \\ &= 3\frac{1 \cdot 1 \cdot 5}{6 \cdot 6 \cdot 6} \\ &= \frac{15}{216} \end{aligned}$$

$$\begin{aligned} f(10) &= P(3maal6) \\ &= \frac{1}{6\cdot 6\cdot 6} \\ &= \frac{1}{216} \end{aligned}$$

$$\begin{aligned} f(-5) &= P(\overline{6}) \\ &= \left(\frac{5}{6}\right)^3 \\ &= \frac{125}{216} \end{aligned}$$

Deze waarden kloppen want $\sum f(i) = 1$. Nu bepalen we het gemiddelde

$$\begin{aligned} \mu &= 5\cdot\frac{15}{216} + 10\cdot\frac{1}{216} - 5\cdot\frac{125}{216} \\ &= -2.50 \end{aligned}$$

De speler verliest gemiddeld 2.50 euro.

11. Wat heeft meest kans? Bij een worp met 4 dobbelstenen minstens een zes werpen, of bij 24 worpen met twee dobbelstenen ten minste een keer dubbele zes werpen.

$$\begin{aligned} P_1 &= 1 - P(\overline{6}) \\ &= 1 - \frac{5\cdot 5\cdot 5\cdot 5}{6\cdot 6\cdot 6\cdot 6} \\ &= 0.5177 \end{aligned}$$

$$\begin{aligned} P_2 &= 1 - P(\overline{dub6}) \\ &= 1 - \left(\frac{35}{36}\right)^{24} \\ &= 0.491404 \end{aligned}$$

12. Bepaal de momentenfunctie van de binomiale verdeling, met

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

voor $i = 0, 1, \ldots, n$.

13. Bepaal de momentenfunctie van de uniform continue verdeling, met dichtheidsfunctie $f(x) = \frac{1}{b-a}, \forall x \in [a, b]$ en $f(x) = 0$ buiten dit interval.

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

$$
\begin{aligned}
&= \int_a^b \left(\frac{1}{b-a}\right) e^{tx} dx \\
&= \frac{1}{b-a}\frac{1}{t}\left[e^{tx}\right]_{x=a}^{x=b} \\
&= \frac{1}{b-a}\cdot\frac{1}{t}\cdot\left(e^{bt}-e^{at}\right)
\end{aligned}
$$

14. Een toevalsverandelijke 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 $P(3 < x < 21)$.

    - $P(6 < x < 18)$
      We lossen dit op door gebruik te maken van de ongelijkheid van Chebychev[8]

      $$
      \begin{aligned}
      P(6 < x < 18) &= P(-6 < x-\mu < 6) \\
      &= P(|x-\mu| < 6) \\
      &= 1 - P(|x-\mu| > 6) \\
      k = 2 \quad &: \quad P(|x-\mu| > 6) \leq \frac{1}{4}
      \end{aligned}
      $$

      We weten echter dat $P(|x-\mu| > 6) \leq \frac{1}{4}$ dus moet de gevraagde kans alsook $1 - P(|x-\mu| > 6) \geq \frac{3}{4}$.
    - $P(3 < x < 21)$

      $$
      \begin{aligned}
      P(3 < x < 21) &= P(-9 < x-\mu < 9) \\
      &= P(|x-\mu| < 9) \\
      &= 1 - P(|x-\mu| > 9) \\
      k = 3 \quad &: \quad P(|x-\mu| > 9) \leq \frac{1}{9}
      \end{aligned}
      $$

      De gevraagde kans is dus $\geq 1 - \frac{1}{9}$ of $\geq \frac{8}{9}$

15. Als de verdelinng van de IQ's van de studenten van een klas een gemiddelde $\mu = 120$ en een dispersie $\sigma = 8$ heeft, bepaal dan een interval dat ten minste 75% van de IQ's bevat.

    $$
    \begin{aligned}
    P(|x-\mu| > k\sigma) &\leq \frac{1}{k^2} \\
    P(|x-\mu| > k\sigma) &\leq \frac{1}{4} \\
    P(|x-\mu| > 16) &\leq \frac{1}{2^2}
    \end{aligned}
    $$

[8]Chebychev: $P(|x-\mu| > k\sigma) \leq \frac{1}{k^2}$

Na uitwerking van Chebychev (waar we rekening houden met het complement van de gebeurtenis) volgt dat:

$$|x - \mu| \quad > \quad 16$$

Het interval is gegeven door $[\mu - 16, \mu + 16]$ of $[104, 136]$.

16. Onderstel dat x een stochastische veranderlijke is met gemiddelde en variantie beide gelijk aan 20. Wat kan gezegd worden over $P(0 < x < 40)$?

$$\begin{aligned}
P(0 < x < 40) &= P(-20 < x - \mu < 20) \\
&= P(|x - \mu| < 20) \\
&= 1 - P(|x - \mu| > 20) \\
k = \sqrt{20} &: \quad P(|x - \mu| > 20) \leq \frac{1}{20}
\end{aligned}$$

Hieruit volgt dat het complement een kans heeft $\geq \frac{19}{20}$ of ≥ 0.95.

17. Onderstel dat het aantal producten in een fabriek, aangemaakt gedurende een week, een stochastische veranderlijke is met gemiddelde $\mu = 50$. Deze veranderlijke heeft een symmetrische verdeling tov μ.

   - Bepaal een bovengrens voor de kans dat de productie van een bepaalde week minstens 75 zal bedragen
   Vermits de dispersie niet gegeven is moeten we het resultaat in functie van de dispersie (σ) uitdrukken.

$$\begin{aligned}
2 \cdot P(x > 75) &= P\left(\frac{25 < x < 75}{2}\right) \\
&= P\left(-25 < x - \mu < 25\right) \\
&= P\left(-25 < x - \mu < 25\right) \\
&= P\left(|x - \mu| > 25\right)
\end{aligned}$$

Hieruit volgt dat $k\sigma = 25$ of dat $k = \frac{25}{\sigma}$. Nu zegt Chebychev.

$$\begin{aligned}
P\left(\frac{|x - \mu| > 25}{2}\right) &\leq \frac{1}{k^2} \\
2P(x > 75) &\leq \frac{\sigma^2}{25^2} \\
P(x > 75) &\leq \frac{1250}{\sigma^2}
\end{aligned}$$

- 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?

$$\begin{aligned} P(40 < x < 60) &= P(-10 < x - \mu < 10) \\ &= P(|x - \mu| < 10) \end{aligned}$$

Hieruit volgt dat $k\sigma = 10$ dus $k = \frac{10}{5} = 2$.

$$\begin{aligned} P(|x - \mu| < 10) &\leq 1 - \frac{1}{k^2} \\ P(|x - \mu| < 10) &\leq 1 - \frac{1}{4} \\ P(|x - \mu| < 10) &\geq \frac{3}{4} \end{aligned}$$

18. Een uurwerkfabrikant verkoopt zijn uurwerken (40 euro) aan 60 euro het stuk als er geen garantie gegeven wordt. Als er garantie gegeven wordt, dit is *het uurwerk vervangen, indien het binnen twee jaar stuk is*, verkoopt de fabrikant het tegen 68 euro. De compagnie verkoopt 100 000 uurwerken per jaar en een test lever een gemiddelde levensduur van 3.5 jaar met een dispersie van een 0.5 jaar. Is de verkoop met garantie meer winstgevend, indien men veronderstelt dat:

    - de verdeling symmetrisch is
    - de verdeling niet symmetrisch is

$$\begin{aligned} W_{\overline{g}} &= 100000(60 - 40) \\ &= 2000000 \\ W_g &= 100000((68 - 400 \cdot P(\overline{defect})) + (68 - 2 \cdot 40) \cdot P(defect)) \end{aligned}$$

Nu bepalen we of $W_g > W_{\overline{g}}$ aan de hand van Chebychev.

$$\begin{aligned} (28 \cdot (1 - P(defect) - 12 \cdot (P(defect)) &\ ?\geq\ 20 \\ 8 &\ ?\geq\ 40 \cdot P(defect) \\ \frac{1}{5} &\ ?\geq\ P(defect) \\ P(|x - \mu| > k\sigma) &\leq \frac{1}{k^2} \\ P(0 < x \leq 2) &= P(defect) \\ P(x < \mu - k\sigma \vee x > \mu + k\sigma) &\leq \frac{1}{k^2} \end{aligned}$$

$$\begin{aligned} P(x < 2 \vee x > 5) &\leq \frac{1}{9} \\ P(0 < x < 2) \leq P(x < 2 \vee x > 5) &\leq \frac{1}{9} \end{aligned}$$

We stelden $k = 3$. Vermits de kans dat het defect zeker kleiner is dan $\frac{1}{9}$ en we slechts $\frac{1}{5}$ nodig hadden om verkoopt met garantie winstgevender te maken is verkoop met garantie dus winstgevender.

4 Discrete verdelingen

1. Van diskettes, geproduceerd door een zeker bedrijf weet men dat de kans op een defect 0.01 bedraagt en dit onafhankelijk van elkaar. Het bedrijf verkoopt de diskettes in dozen van 10 en geeft geld terug-grantie dat er ten hoogste 1 van de diskettes in de doos defect is. Welk percentage dozen wordt teruggegeven? Indien iemand 3 dozen koopt, wat is dan de waarschijnlijkheid dat hij juist 1 doos terugbrengt ?
   De veranderlijke x stelt het aan defecte diskettes per doos voor
   x kent een binomiale verdeling
   P(doos wordt teruggebracht) = P(x¿1)
   n = 10
   p = 0,01

$$\begin{aligned} P(x > 1) &= 1 - P(x = 0) - P(x = 1) \\ &= 1 - \binom{10}{0} p^0 (1-p)^{10} - \binom{10}{1} p^1 (1-p)^9 \\ &= 1 - 0.99^{10} - 10 \cdot 0.01 \cdot 0.99^9 \\ &= 0.004266 \end{aligned}$$

   In het tweede deel van het probleem koopt iemand drie dozen.
   y = aantal teruggebrachte dozen
   y is terug binomiaal verdeeld
   n = 3
   p = 0.004266

$$\begin{aligned} P(y = 1) &= \binom{3}{1} p^1 \cdot (1-p)^2 \\ &= 0.01269 \end{aligned}$$

2. Wat is de kans om driemaal 2 te gooien bij 5 worpen met een onvervalste dobbelsteen ?
   x = het aantal maal dat men 2 gooit
   n = 5
   p = $\frac{1}{6}$

$$\begin{aligned} P(driemaal2) &= \binom{5}{3} p^3 \cdot (1-p)^2 \\ &= 0.3215 \end{aligned}$$

3. Een basketbalspeler heeft 75% kans om raak te scoren. Wat is de kans dat hij 2 maal scoort bij de volgende 4 pogingen ?

x = aantal rake pogingen
n = 4
p = 0.75

$$\begin{aligned} P(2keerraak) &= \binom{4}{2} 0.75^2 \cdot 0.25^2 \\ &= 0.210938 \end{aligned}$$

4. Een communicatiekanaal brengt de digits 1 en 0 over. Als gevolg van ruis wordt de overgeseinde digit verkeerd ontvangen met waarschijnlijkheid 0.2. Onderstel dat een belangrijk bericht bestaande uit juist 1 binaire digit moet overgebracht worden. Om de kans op fouten te reduceren seint met 00000 (voor 0) en 11111 (voor 1). Om het overgeseinde bericht te decoderen past men de volgende strategie toe: het bericht word tals 1 geïnterpreteerd als er meer 1 dan 0 voorkomt en het bericht wordt als 0 geïnterpreteerd als er meer 0 dan 1 voorkomt. Wat is de kans dat het bericht verkeerd gedecodeerd wordt ?
x = aantal verkeerd ontvangen digits
x kent een binomiale verdeling
p = 0.2

$$\begin{aligned} P(foutief) &= P(x=3) + P(x=4) + P(x=5) \\ &= \binom{5}{3} 0.2^3 \cdot 0.8^2 + \\ &\quad \binom{5}{4} 0.2^4 \cdot 0.8^1 + \\ &\quad \binom{5}{5} 0.2^5 \cdot 0.8^0 \\ &= 0.05792 \end{aligned}$$

5. Een multiple-choice examen bestaat uit 15 vragen elk met 4 mogelijke antwoorden waarvan slechts 1 correct is. Wat is de kans dat met puur gissen er tussen 5 en 10 (grenzen inbegrepen) antwoorden correct zijn.
x = aantal correcte vragen
n = 15
p = 0.25

$$\begin{aligned} P(5 \leq x \leq 10) &= P(x=5) + P(x=6) + P(x=7) \\ &\quad +P(x=8) + P(x=9) + P(x=10) \\ &= \binom{15}{5} 0.25^5 \cdot 0.75^{10} \end{aligned}$$

$$
\begin{aligned}
& +\binom{15}{6}0.25^{6}\cdot 0.75^{9} \\
& +\binom{15}{7}0.25^{7}\cdot 0.75^{8} \\
& +\binom{15}{8}0.25^{8}\cdot 0.75^{7} \\
& +\binom{15}{9}0.25^{9}\cdot 0.75^{6} \\
& +\binom{15}{10}0.25^{10}\cdot 0.75^{5} \\
= \; & 0.313399
\end{aligned}
$$

6. Van de producten uit een fabriek is 10% defect. Men neemt een steekproef van 10 stuks. Bereken de kans dat in een steekproef voorkomen:

   - Geen defecte exemplaren
     x = aantal defecte, binomiaal verdeeld
     n = 10
     p = 0.10

     $$
     \begin{aligned}
     P(\overline{defect}) &= \binom{10}{0}0.10^{0}\cdot 0.90^{10} \\
     &= 0.348678
     \end{aligned}
     $$

   - Ten hoogste 2 defecte exemplaren
     x = aantal defecte, binomiaal verdeeld
     n = 10
     p = 0.10

     $$
     \begin{aligned}
     P(0of1of2defect) &= 1-P(x=3)-P(x=4)-P(x=5) \\
     & \quad -P(x=6)-P(x=7)-P(x=8) \\
     & \quad -P(x=9)-P(x=10) \\
     &= 1-\binom{10}{3}0.10^{3}\cdot 0.90^{7} \\
     & \quad -\binom{10}{4}0.10^{4}\cdot 0.90^{6} \\
     & \quad -\binom{10}{5}0.10^{5}\cdot 0.90^{5} \\
     & \quad -\binom{10}{6}0.10^{6}\cdot 0.90^{4} \\
     & \quad -\binom{10}{7}0.10^{7}\cdot 0.90^{3}
     \end{aligned}
     $$

$$
\begin{aligned}
& -\binom{10}{8}0.10^8 \cdot 0.90^2 \\
& -\binom{10}{9}0.10^9 \cdot 0.90^1 \\
& -\binom{10}{10}0.10^{10} \cdot 0.90^0 \\
= \ & 0.929809
\end{aligned}
$$

- Ten minste 3 defecte exemplaren
  x = aantal defecte, binomiaal verdeeld
  n = 10
  p = 0.10

$$
\begin{aligned}
1 - P(0 of 1 of 2 defect) &= 1 - (P(x=0) + P(x=1) \\
&\quad +P(x=2)) \\
&= 1 - \left(\binom{10}{0}0.10^0 \cdot 0.90^{10}\right. \\
&\quad +\binom{10}{1}0.10^1 \cdot 0.90^9 \\
&\quad \left.+\binom{10}{2}0.10^2 \cdot 0.90^8\right) \\
&= 0.070191
\end{aligned}
$$

- Alle exemplaren defect zijn
  x = aantal defecte, binomiaal verdeeld
  n = 10
  p = 0.10

$$
\begin{aligned}
P(alles defect) &= P(x=10) \\
&= \binom{10}{10}0.1^10 \\
&= 0.0000000001
\end{aligned}
$$

7. Je speelt tien keer met de Belgische lotto (je kiest 6 cijfers uit 42)

   - Wat is de kans om 5 of 6 cijfers juist te hebben bij 1 combinatie

$$
\begin{aligned}
P(5 \vee 6) &= \frac{\binom{6}{5}\binom{36}{1} + \binom{6}{6}}{\binom{42}{6}} \\
&= \frac{\binom{6}{5}\binom{36}{1} + 1}{\binom{42}{6}} \\
&= 0.000041367
\end{aligned}
$$

Dit is eveneens de p uit die we nodig hebben in de binomiale verdeling

- Wat is de kans om minstens 1 keer 5 of 6 cijfers juist op de 10 combinaties
  x = aantal keer we 5 of 6 juiste cijfers hebben
  n = 1
  p = 0.000041367
  x is binomiaal verdeeld

$$
\begin{aligned}
P(minstens1) &= 1 - P(x=0) \\
&= 1 - \binom{10}{0} 0.000041367^0 \cdot (1 - 0.00004167)^{10} \\
&= 0.00041359
\end{aligned}
$$

- Wat is de kans om minder dan 3 cijfers juist te hebben bij 1 combinatie

$$
\begin{aligned}
P(<3) &= \frac{\binom{6}{0}\binom{36}{6} + \binom{6}{1}\binom{36}{5} + \binom{6}{2}\binom{36}{4}}{\binom{42}{6}} \\
&= 0.97093534
\end{aligned}
$$

8. Het zelfmoordpercentage in een Amerikaanse staat bedraagt 1 per 100000 inwoners per maand.

   - Wat is de kans dat van 400000 inwoners in deze staat er 8 of meer zelfmoorden plaatsgrijpen in een bepaalde maand?
     x = aantal zelfmoorden
     x kent een binomiale verdeling
     p = $\frac{1}{100000}$
     n = 400 000

$$
\begin{aligned}
P(x \geq 8) &= 1 - (P(x=0) + P(x=1) + \cdots P(x=7)) \\
&= 1 - \left(\binom{400000}{0} \cdot (1 - \frac{1}{100000}^{400000}) \right. \\
&\quad + \binom{400000}{1} \frac{1}{100000}^{1} \cdot (1 - \frac{1}{100000})^{399999} \\
&\quad + \binom{400000}{2} \frac{1}{100000}^{2} \cdot (1 - \frac{1}{100000})^{399998} \\
&\quad + \binom{400000}{3} \frac{1}{100000}^{3} \cdot (1 - \frac{1}{100000})^{399997} \\
&\quad + \binom{400000}{4} \frac{1}{100000}^{4} \cdot (1 - \frac{1}{100000})^{399996} \\
&\quad + \binom{400000}{5} \frac{1}{100000}^{5} \cdot (1 - \frac{1}{100000})^{399995}
\end{aligned}
$$

$$
\begin{aligned}
& +\binom{400000}{6}\frac{1}{100000}^{6}\cdot(1-\frac{1}{100000})^{399994} \\
& +\binom{400000}{7}\frac{1}{100000}^{7}\cdot(1-\frac{1}{100000}^{399993}\Big) \\
= \; & 1-(0.018315+0.073262+0.146525+0.195367 \\
& +0.195368+0.156294+0.104196+0.05954) \\
= \; & 0.051133
\end{aligned}
$$

- Wat is de kans dat er tenminste 2 maanden in het jaar zullen zijn met minstens 8 zelfmoorden ?
  x = aantal maanden met minstens acht zelfmoorden
  x kent een biniomiale verdeling
  n = 12
  p = 0.051133

$$
\begin{aligned}
P(x\geq 2) &= 1-P(x=0)-P(x=1) \\
&= 1-\binom{12}{0}0.051133^{0}\cdot(1-0.051133)^{12} \\
&\quad -\binom{12}{1}0.051133^{1}\cdot(1-0.051133)^{11} \\
&= 0.12286
\end{aligned}
$$

9. Onderstel dat het gemiddeld aantal ongelukken dat wekelijks plaatsgrijpt op een bepaald stuk snelweg 3 bedraagt. Bereken de kans dat er in een bepaalde week minstens 1 ongeluk is.
   x is het aantal ongelukken
   x kent een Poisson verdeling[9]
   $\lambda = 3$

$$
\begin{aligned}
P(x\leq 1) &= 1-P(x=0) \\
&= 1-e^{-3} \\
&= 0.950213
\end{aligned}
$$

10. In een bepaald groot gebied zijn er gemiddeld 6 tyfonen per jaar. Bepaal de kans dat in een gegeven jaar:

    - Meer dan 3 tyfonen voorkomen
      x is het aantal tyfonen
      x is Poisson verdeeld
      $\lambda = 6$

$$
P(x>3) = 1-P(x=0)-P(x=1)-p(x=2)-p(x=3)
$$

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

$$\begin{aligned} &= 1 - e^{-6}\frac{e^{-6}\cdot 6^1}{1} - \frac{e^{-6}\cdot 6^2}{2!} - \frac{e^{-6}\cdot 6^3}{3!} \\ &= 0.848796 \end{aligned}$$

- Er 6 tot 8 tyfonen zijn (grenzen inbegrepen)
  Zelfde parameters als voorgaande

$$\begin{aligned} P(6 \leq x \leq 8) &= P(x=6) + P(x=7) + P(x=8) \\ &= \frac{e^{-6}\cdot 6^6}{6!} + \frac{e^{-6}\cdot 6^7}{7!} + \frac{e^{-6}\cdot 6^8}{8!} \\ &= 0.401558 \end{aligned}$$

11. Beschouw een experiment dat bestaat uit het tellen van α-deeltjes dat vrijkomt in een tijdsinterval van een seconde bij een gram van een zeker radioactief materiaal. Gemiddeld komen er 3.2 dergelijk α-deeltjes vrij. Bepaal de kans dat er niet meer dan 2 α-deeltjes zullen verschijnen.
    x is het aantal α-deeltjes die gedurende 1 seconde uit 1 gram vrijkomen
    x is Poisson verdeeld
    $\lambda = 3.2$

$$\begin{aligned} P(x \leq 2) &= P(x=0)P(x=1) + P(x=2) \\ &= e^{-3.2} + e^{-3.2}\cdot 3.2 + \frac{e^{-3.2}\cdot 3.2^2}{2!} \\ &= 0.379904 \end{aligned}$$

12. De kans om een defect te hebben aan een zeker product is 0.1. Bepaal de kans dat in een populatie van 10 dergelijke producten (onafhankelijk) er ten hoogste 1 defect is. Los op met de poisson en de binomiaal distributie en vergelijk.
    x = aantal defecten
    x is binomiaal verdeeld
    p = 0.1
    n = 10

$$\begin{aligned} P(x \leq 1) &= P(x=0) + P(x=1) \\ &= \binom{10}{0} 0.1^0 \cdot 0.9^{10} + \binom{10}{1} 0.1^1 \cdot 0.9^9 \\ &= 0.736099 \end{aligned}$$

Om dit probleem op te lossen met een Poisson verdeling gebruiken we de limietstelling die zegt: *Voor $n \to \infty$ en $p \to 0$ nadert de binomiale verdeling naar de Poisson verdeling met $\lambda = np$.*
x= aantal defecten

x is Poison verdeeld
$\lambda = 0.1 * 10 = 1$

$$\begin{aligned} P(x \leq 1) &= P(x=0) + P(x=1) \\ &= e^{-1} + e^{-1} \cdot 1 \\ &= 0.735759 \end{aligned}$$

De oplossing met de Poisson verdeling benadert de uitkomst van de binomiale verdeling tot 2 cijfers na de komma nauwkeurig.

13. Een machine produceert bouten waarvan er 2% defect zijn. Wat is de kans dat bij 50 bouten ten hoogste 2 defect zijn ? Mag je hier benaderen door een poisson verdeling ? Bereken met deze distributie de kans.
x = aantal defecte bouten
x is binomiaal verdeeld
p = 0.02
n = 50

$$\begin{aligned} P(x \leq 2) &= P(x=0) + P(x=1) + P(x=2) \\ &= \binom{50}{0} 0.02^0 \cdot 0.98^{50} + \binom{50}{1} 0.02^1 \cdot 0.98^{49} \\ &\quad + \binom{50}{2} 0.02^2 \cdot 0.98^{48} \\ &= 0.921572 \end{aligned}$$

Vermits p redelijk klein is en n relatief groot kunnen we dit eveneens benaderen door een Poisson verdeling met $\lambda = 0.02 * 50 = 1$

$$\begin{aligned} P(x \leq 2) &= P(x=0) + P(x=1) + P(x=2) \\ &= e^{-1} + e^{-1} + \frac{e^{-1} \cdot 1^2}{2} \\ &= 0.919699 \end{aligned}$$

14. Tussen 14u en 16u is het gemiddeld aantal gespreksoproepen in een bepaalde telefooncentrale 2.5 per minuut. Bepaalde kans dat er gedurende een welbepaalde minuut juist 2 oproepen zijn. Bepaal de kans dat er gedurende 5 minuten meer (>) dan 8 oproepen zijn.
x = het aantal oproepen per minuut
x is Poisson verdeeld
$\lambda = 2.5$

$$\begin{aligned} P(x = 2) &= \frac{e^{-2.5} \cdot 2.5^2}{2!} \\ &= 0.256516 \end{aligned}$$

y = het aantal oproepen per vijf minuten
$\lambda = 2.5 * 5 = 12.5$

$$\begin{aligned}
P(x > 8) &= 1 - P(x=0) - P(x=1) - \cdots - P(x=8) \\
&= 1 - \frac{e^{-12.5} \cdot 12.5^0}{0!} + \frac{e^{-12.5} \cdot 12.5^1}{1!} \\
&\quad + \frac{e^{-12.5} \cdot 12.5^2}{2!} + \frac{e^{-12.5} \cdot 12.5^3}{3!} + \frac{e^{-12.5} \cdot 12.5^4}{4!} \\
&\quad + \frac{e^{-12.5} \cdot 12.5^5}{5!} + \frac{e^{-12.5} \cdot 12.5^6}{6!} + \frac{e^{-12.5} \cdot 12.5^7}{7!} \\
&\quad + \frac{e^{-12.5} \cdot 12.5^8}{8!} \\
&= 0.875387
\end{aligned}$$

15. Wat is de kans dat bij het gooien met twee onvervalste dobbelstenen ten minste 3 worpen nodig zijn alvorens de som van de ogen van de dobbelstenen gelijk is aan 8?
x = aantal keren alvorens de som 8 is
x kent een geometrische verdeling
$p = \frac{5}{36}$

$$\begin{aligned}
P(x \geq 3) &= 1 - P(x=1) - P(x=2) \\
&= 1 - \frac{5}{36} - \left(1 - \frac{5}{36}\right)\frac{5}{36} \\
&= 0.741512
\end{aligned}$$

Merk op dat P(x=0) hier niet bestaat ! De voorwaarde kan niet voldaan zijn alvorens er gegooid werd.

16. Een roulette bestaat uit 38 vakjes waarvan er 18 zwart, 18 rood en 2 groen zijn.

- Wat is de kans dat er tenminste 4 spelbeurten nodig zijn om voor de eerste keer groen te treffen ?
x = aantal keren alvorens groen
p kent een geometrische verdeling
$p = \frac{1}{19}$

$$\begin{aligned}
P(x \geq 4) &= 1 - P(x=1) - P(x=2) - P(x=3) \\
&= 1 - \left(1 - \frac{1}{19}\right)^0 \cdot \frac{1}{19} - \left(1 - \frac{1}{19}\right)^1 \cdot \frac{1}{19} \\
&\quad - \left(1 - \frac{1}{19}\right)^2 \cdot \frac{1}{19} \\
&= 0.85027
\end{aligned}$$

- Wat is de kans dat een oneven aantal spelbeurten nodig zijn om voor de eerste keer groen te treffen ?

$$\begin{aligned} P(x\%2=1) &= \frac{1}{19}\cdot\sum_{i=0}^{\infty}\frac{18}{19}^{2i+1} \\ &= \frac{18}{37} \end{aligned}$$

17. De kans dat een geïnfecteerd persoon sterft aan een ademhalingsinfectie is 0.002. Bepaal de kans dat er minder van 5 personen van een goep van 2000 geïnfecteerde personen zullen sterven. Gebruik de ongelijkheid van Chebychev en interpreteer het interval $]\mu-2\sigma,\mu+2\sigma[$.
    x = het aantal personen die sterven
    x is binomiaal verdeeld
    p = 0.002
    n = 2000

$$\begin{aligned} P(x<5) &= P(x=0)+P(x=1)+\cdots+P(x=4) \\ &= \binom{2000}{0}0.002^{0}\cdot 0.998^{2000}+ \\ &\quad \binom{2000}{1}0.002^{1}\cdot 0.998^{1999}+ \\ &\quad \binom{2000}{2}0.002^{2}\cdot 0.998^{1998}+ \\ &\quad \binom{2000}{3}0.002^{3}\cdot 0.998^{1997}+ \\ &\quad \binom{2000}{4}0.002^{4}\cdot 0.998^{1996} \\ &= 0.628984 \end{aligned}$$

    Chebychev zit in de kast

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

    - Wat is de kans dat men bij de vierde beurt voor het eerst een aas trekt ?
      x = aantal beurten
      x is geometrische verdeeld
      $p=\frac{4}{52}$

$$\begin{aligned} P(x=4) &= \left(\frac{48}{52}\right)^{3}\cdot\frac{4}{52} \\ &= 0.060502 \end{aligned}$$

- Wat is de kans dat men minstens zes beurten nodig heeft om een aas te trekken ?

$$
\begin{aligned}
P(x<6) &= 1-P(x=1)-P(x=2)-\cdots-P(x=5) \\
&= 1-\left(\frac{48}{52}\right)^0\cdot\frac{4}{52}-\left(\frac{48}{52}\right)^1\cdot\frac{4}{52}-\left(\frac{48}{52}\right)^2\cdot\frac{4}{52}- \\
&\quad \left(\frac{48}{52}\right)^3\cdot\frac{4}{52}-\left(\frac{48}{52}\right)^4\cdot\frac{4}{52} \\
&= 0.670177
\end{aligned}
$$

5 Continue verdelingen

De in dit hoofdstuk bekomen oplossingen zijn tot op 3 cijfers na de komma nauwkeurig in vergelijking met de gesuggereerde oplossingen in de cursus. Dit komt omdat ik interpolatie toepaste vanaf a waarden die slechts nauwkeurg zijn op 1 cijfer na de komma (dus voor 0.155 interpolatie tussen a=0.1 en a=0.2 ipv interpolatie tussen 0.15 en 0.16).

1. Bussen arriveren stipt aan een bepaalde halte om de 15 minuten van 7 uur 's morgens (dus om 7u, 7u15, $\cdots$). Indien een passagier bij de halte aankomt op een bepaald tijdstip dat uniform verdeeld is tussen 7u en 7u30, bereken.
   x = aankomsttijd van de persoon
   x kent een uniforme verdeling

$$\begin{aligned} \int_a^b k dx &= 1 \\ k &= \frac{1}{b-a} \\ k &= \frac{1}{30-0} \\ k &= \frac{1}{30} \\ \int_0^{30} \frac{1}{30} dx &= 1 \end{aligned}$$

   De kansfunctie is dus:

$$f(x) = \begin{cases} \frac{1}{30}; 7u < x \leq 7u30 \\ 0; elders \end{cases}$$

   - de kans dat hij minder van 5 minuten op de bus moet wachten
     De kans dat hij minder dan 5 minuten moet wachten is dus gelijk aan de kans dat hij arriveert tussen 7u10 en 7u15 plus de kans dat hij arriveert tussen 7u25 en 7u30.

$$\begin{aligned} P(<5) &= P(7u10 < x \leq 7u15) + P(7u25 < x \leq 7u30) \\ &= \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx \\ &= \frac{5}{30} + \frac{5}{30} \\ &= \frac{1}{3} \\ &= 0.3333 \end{aligned}$$

- de kans dat hij minstens 12 minuten op de bus moet wachten
  De kans dat hij minstens 12 minuten moet wachten is de kans dat hij arriveert tussen 7u en 7u03 plus de kans dat hij arriveert tussen 7u15 en 7u18.

$$\begin{aligned} P(\geq 12) &= P(7u < x \leq 7u03) + P(7u15 < x \leq 7u18) \\ &= \int_0^3 \frac{1}{30} dx + \int_{15}^{18} \frac{1}{30} dx \\ &= \frac{3}{30} + \frac{3}{30} \\ &= \frac{1}{5} \\ &= 0.2 \end{aligned}$$

2. De tijd (in uur) nodig om een bepaalde machine te herstellen is exponentieel verdeeld met $\mu = 2$.
   x = reparatietijd
   x kent een exponentiële verdeling
   $\mu = \theta = 2$
   De kansfunctie is hier gegeven door:

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}} = \frac{1}{2} e^{-\frac{x}{2}}$$

- Wat is de kans dat de reparatietijd meer dan 2 uur in beslag neemt ?

$$\begin{aligned} P(x > 2) &= \int_2^{+\infty} \frac{1}{2} e^{-\frac{x}{2}} dx \\ &= -\left[e^{-\frac{x}{2}} \right]_2^{+\infty} \\ &= -\left(0 - \frac{1}{e} \right) \\ &= \frac{1}{e} \\ &= 0.367879 \end{aligned}$$

- Wat is de kans dat de reparatietijd minstens 10 uur bedraagt als je weet dat het zeker meer is dan 9 uur ?

$$\begin{aligned} P(x > 10/x > 9) &= \frac{P(x > 10) \wedge P(x > 9)}{P(x > 9)} \\ &= \frac{P(x > 10)}{P(x > 9)} \\ &= \frac{\int_{10}^{+\infty} \frac{1}{2} e^{-\frac{x}{2}} dx}{\int_9^{+\infty} \frac{1}{2} e^{-\frac{x}{2}} dx} \end{aligned}$$

$$\begin{aligned}
&= \frac{\left[e^{-\frac{x}{2}}\right]_{10}^{+\infty}}{\left[e^{-\frac{x}{2}}\right]_{9}^{+\infty}} \\
&= \frac{e^{-5}}{e^{-\frac{9}{2}}} \\
&= e^{-\frac{1}{2}} \\
&= \frac{1}{\sqrt{e}} \\
&= 0.606531
\end{aligned}$$

3. Het aantal kilometer dat een auto kan rijden met versleten batterij is exponentieel verdeeld met gemiddelde 15000 km. Indien iemand een trip van 7500 km wil maken met een dergelijke versleten batterij, wat is dan de kans dat hij in staat zal zijn die trip af te maken zonder de batterij te moeten vervangen ?
   x = aantal kilometer die men kan afleggen met een versleten batterij
   x kent een exponentiële verdeling
   $\mu = \theta = 15000$
   De kansfunctie is hier gegeven door:

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}} = \frac{1}{15000} e^{-\frac{x}{15000}}$$

$$\begin{aligned}
P(x > 7500) &= \int_{7500}^{+\infty} \frac{1}{15000} e^{-\frac{x}{15000}} dx \\
&= -\left[e^{-\frac{x}{15000}}\right]_{7500}^{+\infty} \\
&= e^{-\frac{7500}{15000}} \\
&= e^{-\frac{1}{2}} \\
&= \frac{1}{\sqrt{e}} \\
&= 0.606531
\end{aligned}$$

4. Veronderstel x is N(10,6) verdeeld [10]. Bereken:
   - Algemeen:
     We zetten eerst om naar een genormeerde normale verdeling door van de grenzen μ af te trekken en te delen door σ.

[10] N(10,6), normale verdeling met $\mu = 10$ en $\sigma = 6$

- $P(x > 5)$

$$\begin{aligned}
P(x > 5) &= P(\frac{x-10}{6} > \frac{5-10}{6}) \\
&= P(\frac{x-10}{6} > \frac{-5}{6}) \\
&= P(\frac{x-10}{6} > -0.83333) \\
&= 0.5 + P(0 < z < 0.83333) \\
&= 0.5 + P(0 < z < 0.8) \\
&\quad + (P(0 < z < 0.9) + P(0 < z < 0.8)) \cdot \frac{1}{3} \\
&= 0.5 + 0.2881 + \frac{0.3159 - 0.2881}{3} \\
&= 0.797367
\end{aligned}$$

- $P(4 \leq x \leq 16)$

$$\begin{aligned}
P(4 \leq x \leq 16) &= P(\frac{4-10}{6} \leq x \leq \frac{16-10}{6}) \\
&= P(-1 \leq z \leq 1) \\
&= 2 \cdot P(0 \leq z \leq 1) \\
&= 2 \cdot 0.3413 \\
&= 0.6826
\end{aligned}$$

- $P(x < 8)$

$$\begin{aligned}
P(x < 8) &= P(\frac{x-10}{6} < \frac{8-10}{6}) \\
&= P(z < -\frac{1}{3}) \\
&= 0.5 - P(0 < z < \frac{1}{3}) \\
&= 0.5 - P(0 < z < 0.3) - (P(0 < z < 0.3) \\
&\quad - P(0 < z < 0.4)) \cdot \frac{1}{3} \\
&= 0.5 - 0.1179 - (0.1554 - 0.1179) \cdot \frac{1}{3} \\
&= 0.3696
\end{aligned}$$

- $P(x < 20)$

$$\begin{aligned}
P(x < 20) &= P(\frac{x-10}{6} < \frac{20-10}{6}) \\
&= P(z < \frac{10}{6})
\end{aligned}$$

$$\begin{aligned} &= 0.5 + P(0 < z < 1.6) + (P(0 < z < 1.7) \\ &\quad -P(0 < z < 1.6)) \cdot \frac{2}{3} \\ &= 0.5 + 0.4452 + (0.4554 - 0.4452) \cdot \frac{2}{3} \\ &= 0.952 \end{aligned}$$

- $P(x \geq 16)$

$$\begin{aligned} P(x \geq 16) &= P(\frac{x-10}{6} \geq \frac{16-10}{6}) \\ &= P(z \geq 1) \\ &= 0.5 - P(z < 1) \\ &= 0.5 - 0.3413 \\ &= 0.1587 \end{aligned}$$

5. Stel $x_1N(2,3); x_2N(7,5); x_3N(-5,4)$ met $x_1; x_2; x_3$ twee per twee onafhankelijk. Bepaal voor $x = x_1 + x_2 + x_3$:

- Algemeen
  x is normaal verdeeld met gemiddeld 2+7-5 en met variantie $\sqrt{9+25+16}$ of:

$$\begin{aligned} x &; N(2+7-5, \sqrt{9+25+16}) \\ x &; N(4, 5\sqrt{2}) \end{aligned}$$

- $P(x \leq 13)$

$$\begin{aligned} P(x \leq 13) &= P(\frac{x-4}{5\sqrt{2}} \leq \frac{13-4}{5\sqrt{2}}) \\ &= P(z \leq 1.27) \\ &= 0.5 + P(0 < z < 1.2) \\ &\quad +(P(0 < z < 1.3) - P(0 < z < 1.2)) \cdot 0.73 \\ &= 0.5 + 0.3849 + (0.4032 - 0.3849) \cdot 0.73 \\ &= 0.8983 \end{aligned}$$

- $P(x > 7)$

$$\begin{aligned} P(x > 7) &= P(\frac{x-4}{5\sqrt{2}} > \frac{7-4}{5\sqrt{2}}) \\ &= P(z > 0.424264) \\ &= 0.5 - P(0 < z < 0.424264) \\ &= 0.5 - P(0 < z < 0.4) \end{aligned}$$

$$\begin{aligned}
&\quad -(P(0<z<0.5)-P(0<z<0.4))\cdot\frac{1}{4}\\
&= 0.5-0.1554-(0.1915-0.1554)\cdot\frac{1}{4}\\
&= 0.3356
\end{aligned}$$

6. De jaarlijkse regenval in cm in een bepaalde streek is normaal verdeeld met gemiddelde $\mu = 40$ en dispersie $\sigma = 4$. Wat is de waarschijnlijkheid dat in 2 van de volgende vier jaar de regenval minstens 50 cm zal bedragen ? (Onderstel de regenval in verschillende jaren onafhankelijk).
   x = aantal regenval in cm
   $x; N(40,4)$
   We bepalen eerst de kans dat er meer dan 50 cm neerslag valt:

$$\begin{aligned}
P(x>50) &= P(\frac{x-40}{4} > \frac{50-40}{4})\\
&= P(z>2.5)\\
&= 0.5-P(z<2.5)\\
&= 0.5-0.4938\\
&= 0.0062
\end{aligned}$$

   y = aantal jaren met meer dan 50 cm regen
   y kent een binomiale verdeling met $n = 4$

$$\begin{aligned}
P(x=2) &= \binom{4}{2}\cdot 0.0062^2\cdot(1-0.0062)^2\\
&= 0.000228
\end{aligned}$$

7. Veronderstel x is N(2,2) verdeeld. Bepaal de kritische waarde a zodat P(x > -a) = 20%.

$$\begin{aligned}
P(x>-a) &= 0.2\\
P(\frac{x-2}{2} > \frac{a+2}{-2}) &= 0.2\\
P(z > \frac{a+2}{-2}) &= 0.2\\
0.5 - P(z < \frac{a+2}{-2}) &= 0.2\\
P(z < \frac{a+2}{-2}) &= 0.3
\end{aligned}$$

   Uit raadpleging en interpolatie a.d.h.v de tabel volgt:

$$\frac{a+2}{-2} = 0.8428$$

$$\begin{aligned} a+2 &= -2\cdot 0.8428 \\ a &= -2\cdot 0.8428 - 2 \\ a &= -3.685 \end{aligned}$$

8. Het IQ van laatstejaarsstudenten is N(110,8). De school wil een speciale cursus inrichten voor 10% van de studenten met de hoogste IQ-score. Wat is de laagste IQ-score die in aanmerking komt voor deze speciale cursus?
   x = de iqscore
   x is normaal verdeeld

$$\begin{aligned} P(x>a) &= 0.1 \\ P(\frac{x-110}{8} > a-1108) &= 0.1 \\ 0.5 - P(z < \frac{a-110}{8}) &= 0.1 \\ P(z < \frac{a-110}{8}) &= 0.4 \end{aligned}$$

   Uit tabel volgt :

$$\begin{aligned} \frac{a-110}{8} &= 1.28251 \\ a &= 120.26 \end{aligned}$$

9. Een firma produceert lampen waarvan de levensduur in uren N(800,40) verdeeld is. Indien 100 lampen at random uitgetest worden, hoeveel zullen er een levensduur hebben die ligt tussen 778 en 834 uur ?
   Werkwijze: we bepalen het aantal lampen die langer meegaan dan 778 uur, alsook deze die langer meegaan dan 834 uur. Het verschil van deze twee waarden is de kans dat een lamp tussen de 778 en de 834 uur meegaat.

$$\begin{aligned} P(x>778) &= P(z > \frac{778-800}{40}) \\ &= P(z > -0.55) \\ &= 0.5 + P(0 < z < 0.55) \\ &= .70865 \end{aligned}$$

$$\begin{aligned} P(x>834) &= P(z > \frac{834-800}{40}) \\ &= P(z > 85) \\ &= 0.5 - P(0 < z < 0.85) \\ &= 0.198 \end{aligned}$$

$$\begin{aligned} P(778 < x < 834) &= 0.70865 - 0.198 \\ &= 0.51065 \end{aligned}$$

Vermits 51% van de lampen het tussen de 778 en 834 uur houden en we 100 lampen testen kunnen we niet anders dan besluiten dat 51 lampen het tussen de 778 en 834 uur zullen uithouden

10. Twee dobbelstenen worden 180 keer gegooid. Wat is de kans dat een totaal van 7:
Dit is geen continue verdeling maar een discrete binomiale verdeling. We kunnen deze echter benaderen door een Poisson verdeling of een Normale verdeling door gebruik te maken van de limietstellingen.[11]

    - tenminste 25 keer wordt gegooid ?
    We herleiden deze verdeling naar $N(np, \sqrt{np(1-p)})$ of $N(30, 5)$, hierbij is x de veranderlijke die het aantal voorstelt dat 7 gegooid wordt. We moeten hier wel rekening houden met de continuïteitscorrectie !

    $$\begin{aligned} P(x \geq 24.5) &= P(z \geq \frac{24.5-30}{5}) \\ &= P(z \geq -1.1) \\ &= 0.5 + P(z < 1.1) \\ &= 0.8643 \end{aligned}$$

    - juist 30 keer wordt gegooid ?

    $$\begin{aligned} P(29.5 < x < 30.5) &= P(\frac{29.5-30}{5} < x < \frac{30.5-30}{5}) \\ &= P(-0.1 < z < 0.1) \\ &= 2 * P(z < 0.1) \\ &= 0.0796 \end{aligned}$$

11. Een Geigerteller levert gemiddeld 30 tellen per minuut in de omgeving van een radioactief materiaal. Stel dat het aantal tellen per minuut Poisson verdeeld is. Bepaal de kans dat [12]

    - er juist 32 tellen zijn

    $$P(31.5 < x < 32.5) = P(\frac{31.5-30}{\sqrt{30}} < z < \frac{32.5-30}{\sqrt{30}})$$

[11] Als x binomiaal verdeeld is met parameters n en p, dan nadert deze verdeling naar de normale verdeling $N(np, \sqrt{np(1-p)}$ als zowel $np \geq 5$ en $n(1-p) \geq 5$

[12] We maken gebruik van een andere limietstelling: Als x poisson verdeeld met parameter λ, dan nadert deze verdeling naar de normale verdeling $N(\lambda, \sqrt{\lambda})$ als λ voldoende groot is.

$$
\begin{aligned}
&= P(0.274 < z < 0.456) \\
&= P(z < 0.456) - P(z < 0.274) \\
&= 0.17345 - 0.10786 \\
&= 0.065
\end{aligned}
$$

- er tussen 23 en 35 tellen zijn, grenzen niet inbegrepen

$$
\begin{aligned}
P(23.5 < x < 34.5) &= P(\frac{23.5 - 30}{\sqrt{30}} < z < \frac{34.5 - 30}{\sqrt{30}}) \\
&= P(-1.187 < z < 0.822) \\
&= P(z < 1.187) + P(z < 0.822) \\
&= 0.3822 + 0.2942 \\
&= 0.6764
\end{aligned}
$$

12. Stel dat x $\chi^2(24d.f.)$ verdeel en y $\chi^2(5d.f.)$ verdeeld is en dat x en y onafhankelijk zijn. Bepaal

- a zodat $P(x \leq a) = 0.95$
  a = 36.4 volgt rechtstreeks uit de tabel
- $P(x \geq 19)$
  P(x¡19) = 0.25, dit volgt rechtstreeks uit te tabel, om het gevraagde te bekomen nemen we hiervan het complement wat resulteert in een kans van 0.75.
- b zodat $P(33.2 \leq x \leq b) = 0.095$

$$
\begin{aligned}
P(33.2 \leq x \leq b) &= 0.095 \\
P(33.2 \leq x) - P(x \geq b) &= 0.095 \\
1 - P(x \leq 33.2) - 1 + P(x \leq b) &= 0.095 \\
-0.90 + P(x \leq b) &= 0.095 \\
P(x \leq b) &= 0.995
\end{aligned}
$$

  Hieruit volgt dat b gelijk moet zijn aan 45.6
- $P(x + y) > 16$
  De som van x en y is $\chi^2(29d.f.)$

$$
P(x + y > 16) = 0.975
$$

  Deze waarde is eveneens rechtstreeks afkomstig uit de tabel.

13. Stel dat x $t(20d.f.)$ verdeeld is. Bepaal

- a zodat $P(x > a) = 0.3$
  a moet gelijk zijn aan 0.533

- b zodat $P(|x| > b) = 0.2$ b moet gelijk zijn aan 1.32

14. Stel x $N(0,3)$ en y $\chi^2(16d.f.)$ verdeeld en x en y onafhankelijk. Bepaal $a < 0$ zodat zodat $P(ax \geq \sqrt{y}) = 0.3$.
$x : N(0,3)$
$y : \chi^2(16d.f.)$
We gaan de uitdrukking $ax \geq \sqrt{y}$ herleiden naar een combinatie van x en y waarvan we de verdeling kennen. We weten eveneens dat een t verdeling gegeven is door $\frac{N}{\sqrt{\frac{\chi}{d.f}}}$. Deze vorm lijkt sterkt op hetgeen we reeds hebben. De t verdeling zou in dit geval $\frac{\frac{x-0}{3}}{\sqrt{\frac{y}{16}}}$ zijn.

$$\begin{aligned} P\left(ax \geq \sqrt{y}\right) &= 0.3 \\ P\left(\frac{4x}{3\sqrt{y}} \leq \frac{4}{3a}\right) &= 0.3 \end{aligned}$$

In de vorige stap werd de gelijkheid omgedraaid, want $a < 0$, ook gebruiken we de symmetrie om een waarde afleesbaar uit de tabellen te bekomen.

$$\begin{aligned} P\left(\frac{4x}{3\sqrt{y}} \leq \frac{-4}{3a}\right) &= 0.7 \\ t_{0,7}\left(16d.f.\right) &= 0.535 \\ \frac{-4}{3a} &= 0.535 \\ a &= -2.4922 \end{aligned}$$

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

- $P(x < a) = 0.95$ [13]

$$\begin{aligned} \alpha &= 0.05 \\ a &= 2.54 \end{aligned}$$

- $P(x > b) = 0.95$ [14]

$$\begin{aligned} P\left(\frac{1}{x} < \frac{1}{b}\right) &= 0.95 \\ P\left(\frac{1}{x} > \frac{1}{b}\right) &= 0.05 \end{aligned}$$

[13] Rechtstreeks uit te tabel te halen

[14] Niet in de tabel, dus omvormen, we weten dat $\frac{1}{x}$ een Fischer verdeling kent met 15,10 d.f.

$$\begin{aligned} \frac{1}{b} &= 2.85 \\ b &= \frac{1}{2.85} \end{aligned}$$

16. Controleer de formule $F_\alpha(1, \nu d.f.) = \left(t_{1-\frac{\alpha}{2}}\right)^2$ voor $\alpha = 0.1$ en $\nu = 3$.

17. Als x $\chi^2(5d.f.)$ verdeeld en y $\chi^2(10d.f.)$ verdeeld is (x en y zijn onafhankelijk), bepaal de kritische waarde a zodat $P(x > ay) = 0.005$
$x : \chi^2(5d.f.)$
$y : \chi^2(10d.f.)$
We herleiden dit probleem tot een Fischer verdeling met 5,10 d.f.

$$\begin{aligned} P\left(x > ay\right) &= 0.05 \\ P\left(\frac{x}{y} > a\right) &= 0.05 \\ P\left(\frac{\frac{x}{5}}{\frac{y}{10}} > 2a\right) &= 0.05 \\ 2a &= 3.33 \\ a &= 1.667 \end{aligned}$$

18. Stel dat x $\chi^2(350d.f.)$ verdeeld is, bepaal $P(x > 390)$.

6 Schattingstheorie

1. Bij het meten van een reactietijd noteert een psycholoog een standaarddeviatie van 0.05 seconden. Hoe groot moet het aantal metingen zijn om aan te nemen, met een betrouwbaarheid van 95%, respectievelijk 99%, dat de fout op het schatten van de gemiddelde reactietijd minder dan 0.01 seconde bedraagt (veronderstel $n \geq 30$).
   $\sigma_0 = 0.05$
   x is de reactietijd, deze kent een normale verdeling $N(\mu_0, \sigma_0)$. Deze normalizeren we tot:

   $$Z \quad : \quad \frac{\overline{x} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}}$$

   Z is verdeeld volgens $N(0,1)$.

   - Met 95% betrouwbaarheid
     We zoeken dus n zodat geldt:

     $$P(-0.01 < \overline{x} - \mu_0 < 0.01) \quad = \quad 0.95$$

     Om dit te bereiken baseren we ons eerst op de normale verdeling:

     $$\begin{aligned}
     P\left(-Z_{0.475} < Z < Z_{0.475}\right) &= 0.95 \\
     P\left(-1.96 < Z < 1.96\right) &= 0.95 \\
     P\left(-1.96 < \frac{\overline{x} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} < 1.96\right) &= 0.95 \\
     P\left(-1.96 \cdot \frac{\sigma_0}{\sqrt{n}} < \overline{x} - \mu_0 < 1.96 \cdot \frac{\sigma_0}{\sqrt{n}}\right) &= 0.95 \\
     \left|1.96 \cdot \frac{\sigma_0}{\sqrt{n}}\right| &\leq 0.01 \\
     \left(1.96 \cdot \frac{0.05}{0.01}\right)^2 &\leq n \\
     n &\geq 97
     \end{aligned}$$

   - Met 99% betrouwbaarheid

     $$\begin{aligned}
     P\left(-Z_{0.495} < Z < Z_{0.495}\right) &= 0.99 \\
     P\left(-2.575 < Z < 2.575\right) &= 0.99 \\
     P\left(-2.575 < \frac{\overline{x} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} < 2.575\right) &= 0.99 \\
     P\left(-2.575 \cdot \frac{\sigma_0}{\sqrt{n}} < \overline{x} - \mu_0 < 2.575 \cdot \frac{\sigma_0}{\sqrt{n}}\right) &= 0.99
     \end{aligned}$$

$$\begin{aligned}
\left|2.575 \cdot \frac{\sigma_0}{\sqrt{n}}\right| &\leq 0.01 \\
\left(2.575 \cdot \frac{0.05}{0.01}\right)^2 &\leq n \\
n &\geq 166
\end{aligned}$$

2. Men meet de reactietijd van 5 personen. Deze reactietijd is benaderd normaal verdeeld. De resultaten zijn: $0.28, 0.3, 0.33, 0.31$.

   - Bepaal het 95% en het 99% betrouwbaarheidsinterval voor de gemiddelde reactietijd.
     We bepalen dus het 95% betrouwbaarheidsinterval voor μ_0 terwijl σ_0 onbekend is. Verder bepalen de het steekproefgemiddelde en de steekproefdeviatie. De gemiddelde waarde uit de steekproef[15] is $\overline{x} = 0.298$. De steekproefvariantie [16] is $s^2 = 0.00057$. Hieruit halen we t welke een student of t-verdeling kent met 4 vrijheidsgraden:

     $$t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

     Nu bepalen we het 95% interval via:

     $$\begin{aligned}
     P\left(-t_{0.975} < t < t_{0.975}\right) &= 0.95 \\
     P\left(-2.78 < t < 2.78\right) &= 0.95 \\
     P\left(-2.78 < \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}} < 2.78\right) &= 0.95 \\
     P\left(-2.78 \cdot \frac{s}{\sqrt{n}} + \mu_0 < \overline{x} < 2.78 \cdot \frac{s}{\sqrt{n}} + \mu_0\right) &= 0.95 \\
     P\left(0.268317 < \overline{x} < 0.327683\right) &= 0.95
     \end{aligned}$$

     Nu bepalen we het 99% interval via:

     $$\begin{aligned}
     P\left(-t_{0.995} < t < t_{0.995}\right) &= 0.99 \\
     P\left(-4.60 < t < 4.60\right) &= 0.99 \\
     P\left(-4.60 < \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}} < 4.60\right) &= 0.99 \\
     P\left(-4.60 \cdot \frac{s}{\sqrt{n}} + \mu_0 < \overline{x} < 4.60 \cdot \frac{s}{\sqrt{n}} + \mu_0\right) &= 0.99 \\
     P\left(0.248885 < \overline{x} < 0.347115\right) &= 0.99
     \end{aligned}$$

[15] $\overline{x} = \frac{1}{n}\left(x_1 + \ldots + x_n\right)$
[16] $s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$

- Bepaal het 95% betrouwbaarheidsinterval voor de standaarddeviatie.
  We weten dat $\frac{(n-1)\cdot s^2}{\sigma_0^2}$ $\chi^2(n-1d.f.)$-verdeeld is dus:

$$\begin{aligned}
P\left(\chi^2_{0.025} < \frac{(n-1)\cdot s^2}{\sigma_0^2} < \chi^2_{0.975}\right) &= 0.95 \\
P\left(0.484 < \frac{(n-1)\cdot s^2}{\sigma_0^2} < 11.1\right) &= 0.95 \\
P\left(\frac{0.484}{(n-1)s^2} < \frac{1}{\sigma_0^2} < \frac{11.1}{(n-1)s^2}\right) &= 0.95 \\
P\left(\frac{(n-1)s^2}{11.1} < \sigma_0^2 < \frac{(n-1)s^2}{0.484}\right) &= 0.95 \\
P\left(0.014332 < \sigma_0 < 0.068635\right) &= 0.95
\end{aligned}$$

3. Groep A, bestaande uit 50 patiënten, werd behandeld met een nieuw type slaappil. Groep B, 100 patiënten, kreeg een klassiek slaapmiddel. Voor groep A was de gemiddelde slaapduur $\overline{x}_a = 7.82u$ met een standaarddeviatie $s_A = 0.24u$. De resultaten voor groep B waren $\overline{x}_b = 6.75u$ en $s_B = 0.3u$. Bepaal het 95% betrouwbaarheidsinterval voor het verschil in gemiddelde slaapduur veroorzaakt door de twee types slaappil.
   We zoeken de het verschil tussen twee gemiddeldes waarbij de varianties ongekend zijn. We baseren ons dus op:

$$t = \frac{\overline{x}_1 - \overline{x}_2 - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

   Dit kent een $t(n_1 + n_2 - 2d.f.)$ en bovendien is:

$$\begin{aligned}
s_p^2 &= \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \\
s_p^2 &= 0.079273 \\
s_p &= 0.28155
\end{aligned}$$

   Dus:

$$\begin{aligned}
P\left(-t_{0.975} < t < t_{0.975}\right) &= 0.95 \\
P\left(-1.96 < t < 1.96\right) &= 0.95 \\
P\left(-1.96 < \frac{\overline{x}_1 - \overline{x}_2 - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < 1.96\right) &= 0.95 \\
P\left(0.974419 < |\mu_1 - \mu_2| < 1.16558\right) &= 0.95
\end{aligned}$$

4. Bepaal het 99% betrouwbaarheidsinterval voor de standaarddeviatie van de grootte van de studenten, als je weet dat voor een steekproef van 16 studenten de standaarddeviatie 6 cm bedraagt (veronderstel een normale distributie).

$$\begin{aligned}
P\left(-\chi_{0.005} < \frac{(n-1)s^2}{\sigma_0^2} < \chi_{0.995}\right) &= 0.99 \\
P\left(4.6 < \frac{(n-1)s^2}{\sigma_0^2} < 32.8\right) &= 0.99 \\
P\left(\frac{4.6}{(n-1)s^2} < \frac{1}{\sigma_0^2} < \frac{32.8}{(n-1)s^2}\right) &= 0.99 \\
P\left(\frac{(n-1)s^2}{4.6} > \sigma_0^2 > \frac{(n-1)s^2}{32.8}\right) &= 0.99 \\
P\left(4.05751 < \sigma_0 < 10.8347\right) &= 0.99
\end{aligned}$$

5. De diameter van ringen wordt normaal verdeeld verondersteld. De standaarddeviatie van twee steekproeven van elke 10 ringen bedraagt respectievelijk $s_1 = 0.042$ cm en $s_2 = 0.035$ cm. Bepaal het 98% betrouwbaarheidsinterval voor de verhouding van de varianties $\frac{\sigma_2^2}{\sigma_1^2}$.

$$\begin{aligned}
P\left(F_{0.99}(9, 9d.f.) < \frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{s_1^2}{s_2^2} < F_{0.01}(9, 9d.f.)\right) &= 0.98 \\
P\left(\frac{1}{F_{0.01}} \cdot \frac{s_2^2}{s_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < F_{0.01} \cdot \frac{s_2^2}{s_1^2}\right) &= 0.98 \\
P\left(\frac{1}{5.35} \cdot \frac{s_2^2}{s_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < 5.35 \cdot \frac{s_2^2}{s_1^2}\right) &= 0.98 \\
P\left(0.129803 < \frac{\sigma_2^2}{\sigma_1^2} < 3.71528\right) &= 0.98
\end{aligned}$$

6. Twee steekproeven met respectievelijk 6 en 8 metingen, afkomstig uit een normale distributie, hebben dezelfde steekproefvariantie. Bepaal het 90% betrouwbaarheidsinterval voor $\frac{\sigma_2^2}{\sigma_1^2}$.

$$\begin{aligned}
P\left(F_{0.95}(5, 7d.f.) < \frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{s^2}{s^2} < F_{0.05}(5, 7d.f.)\right) &= 0.90 \\
P\left(\frac{1}{F_{0.05}(7, 5d.f.)} < \frac{\sigma_2^2}{\sigma_1^2} < F_{0.05}(5, 7d.f.)\right) &= 0.90 \\
P\left(0.204918 < \frac{\sigma_2^2}{\sigma_1^2} < 3.970\right) &= 0.90
\end{aligned}$$

7. Bij een steekproef van 150 studenten zijn er 50 die met auto naar school komen. De totale schoolbevolking is 8000. Bepaal het 95% betrouwbaarheidsinterval voor het aantal studenten dat met de auto naar school komt. Zijn 3000 parkeerplaatsen voldoende ?
   x is het aantal studenten dat met de auto komt en kent een binomiale verdeling waarbij $p_0 = \frac{1}{3}$, en waarbij $n = 150$.

$$\begin{aligned}
P\left(-Z_{0.475} < \frac{p-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} < Z_{0.475}\right) &= 0.95 \\
P\left(-1.96 < p - p_0 < 1.96\right) &= 0.95 \\
P\left(-1.96 \cdot \sqrt{\frac{p_0(1-p_0)}{n}} < p - p_0 < 1.96 \cdot \sqrt{\frac{p_0(1-p_0)}{n}}\right) &= 0.95 \\
P\left(-0.07544 + p_0 < p < 0.07544 + p_0\right) &= 0.95 \\
P\left(0.257893 < p < 0.408774\right) &= 0.95
\end{aligned}$$

   Het aantal studenten dat in totaal met de auto komt ligt in het interval:

$$\begin{aligned}
&[3000 \cdot 0.257893, 3000 \cdot 0.408774] \\
&[2063, 3270]
\end{aligned}$$

8. Het gewicht van 15 mensen (veronderstel het gewicht is normaal verdeeld) geeft volgende resultaten:

$$65, 70, 74, 63, 88, 54, 76, 81, 62, 64, 69, 86, 91, 83, 54$$

   Bepaal het 95% voorspellingsinterval voor het gewicht.
   n.t.k.

7 Testen van Hypothesen

1. Het IQ van twee groepen van 16 studenten wordt gemeten; men bekomt $\overline{x}_1 = 107$, $s_1 = 10$, $\overline{x}_2 = 112$, $s_2 = 8$ (veronderstel dat de populatie normaal verdeeld is met vergelijkbare σ). Is er op 95% en 90% betrouwbaarheid voldoende reden om te besluiten dat $\mu_2 > \mu_1$?
   Eerst bepalen we $s_p^2 = 82$, hieruit volgt dat $s_p = 9.05539$. Daaruit berekenen we dat $t_{ber} = -1.56175$ (t verdeling 30 d.f.). Dit geldt wanneer de nulhypthese geldt: $H_0 : mu_1 = \mu_2$. De alternatieve hypothese is $H_1 : \mu_2 > \mu_1$.
   95%
   $$t_{0.05}(30d.f.) \quad = \quad -1.7$$
   Dus hier ligt t_{ber} in het aanvaardingsgebied. H_0 aanvaarden Voor 90%:
   $$t_{0.10}(30d.f.) \quad = \quad -1.31$$
   Hieruit volgt dat $\mu_2 > \mu_1$ Dus hier ligt t_{ber} in het verwerpingsgebied. H_1 aanvaarden.

2. Een toestel vult flesjes met vloeistof. De inhoud van de flesjes is normaal verdeeld en de stadaarddeviatie bedraagt 0.25. Een steekproef van 25 flesjes levert $s = 0.32$. Is de variantie significant groter geworden op 0.05 niveau ?
   $H_0 : \sigma = \sigma_0$
   $H_1 : \sigma > \sigma_0$
   $$\begin{aligned}\chi^2_{ber} &= \frac{(n-1)s^2}{\sigma_0} \\ \chi^2_{ber} &= \frac{(24)0.32^2}{0.25} \\ \chi^2_{ber} &= 39.3216\end{aligned}$$
   Dit is een enkelvoudige proef dus vergelijkgen we met $\chi^2_{0.95}(24d.f.) = 36.5$, hieruit volgt dat $\chi^2_{ber} > \chi^2_{0.95}$, we verwerpen dus H_0, de variantie is dus toegenomen.

3. Een fabrikant van kaarsen beweert dat de brandduur van zijn kaarsen normaal verdeeld is met $\mu = 5$ uur. Een klant wil veel kaarsen kopen als de fabrikant de waarheid spreekt. Hij koopt er 50 en berekent $\overline{x} = 4.9$ uur en $s = 0.25$ uur. Zal de steekproef de klant overtuigen, met 95% betrouwbaarheid, om over te gaan op een massale aankoop ?
   $H_0 : \mu_0 = 5$
   $H_1 : \mu_0 < 5$
   $$\begin{aligned}t_{ber} &= \frac{4.9 - 5}{\frac{0.25}{\sqrt{50}}} \\ t_{ber} &= -2.8284\end{aligned}$$

Dit vergelijkingen we met $t_{0.05}(49d.f.) = -1.675$, hieruit volgt dat $t_{ber} < t_{0.05}$, dus we verwerpen H_0.

4. Bij een ondervraging statistiek halen 12 studenten uit een eerste groep gemiddeld 78% met standaarddeviatie 6%, terwijl 15 studenten uit een tweede groep gemiddeld 70% halen met een standaarddeviatie van 8%. (Veronderstel populaties normaal verdeeld).

   - Kunnen we zeggen met een betrouwbaarheid van 90% dat beide varianties gelijk zijn ?
     $H_0 : \sigma_1 = \sigma_2$
     $H_1 : \sigma_1 \neq \sigma_2$

     $$\begin{aligned} f &= \frac{s_1^2}{s_2^2} \\ &= 0.5625 \end{aligned}$$

     Nu vergelijken we $F_{ber}(11, 14)$ met de grenzen van deze tweezijdige voorwaarde, we testen dus of:

     $$\begin{aligned} F_{0.95}(11, 14d.f.) < & \quad F_{ber} & < F_{0.05}(11, 14d.f.) \\ \frac{1}{F_{0.05}(14, 11d.f.)} < & \quad F_{ber} & < F_{0.05}(11, 14d.f.) \\ 0.3636 < & \quad F_b er & < 2.55 \end{aligned}$$

     Deze voorwaarde is voldaan dus we aanvaarden H_0
   - Is groep 1 beter dan groep 2, met een betrouwbaarheid van 95% ?
     $H_0 : \mu_1 = \mu_2$
     $H_1 : \mu_1 > \mu_2$

     $$\begin{aligned} s_p &= 0.071889 \\ t_{ber} &= 2.8733 \\ t_{0.95}(11, 14d.f.) &= 1.71 \end{aligned}$$

     We verwerpen H_0 en nemen H_1 aan.

5. Met het volgende experiment wil men weten of een nieuwe methode voor de productie van synthetische diamant de moeite waard is. Onderstel dat het karaat van diamanten normaal verdeeld is. Men fabriceert 6 diamanten volgens deze methode en bekomt de volgende gegevens: 0.46 0.61 0.52 0.48 0.57 0.54

- Test eenzijdig voor $\alpha = 0.1$ en $\alpha = 0.2$ de hypothese: het gemiddelde karaat is gelijk aan 0.5. Bepaal het significantieniveau α_{ber}.

$$\begin{aligned} \overline{x} &= 0.53 \\ s^2 &= 0.00312 \\ s &= 0.055857 \end{aligned}$$

  $H_0 : \mu = 5$
  $H_1 : \mu < 0.5$

$$\begin{aligned} t_{ber} &= 1.31559 \\ t_{0.90}(5d.f.) &= 1.48 \\ t_{0.80}(5d.f.) &= 0.92 \end{aligned}$$

  Hieruit volgt dat als $\alpha = 0.1$ we H_0 aanvaarden en dat als $\alpha = 0.2$ we H_0 verwerpen.

- Met de klassieke methode bekomt men voor het karaat van 8 diamanten: 0.49 0.53 0.62 0.58 0.49 0.55 0.61 0.57 Ga na of de methodes significant verschillen op 95% (veronderstel dat de standaarddeviaties ongeveer gelijk zijn). Bepaal het significantieniveau.
  $H_0 : \mu_1 = \mu_2$
  $H_1 : \mu_1 \neq \mu_2$

$$\begin{aligned} \overline{x}_2 &= 0.555 \\ s_2^2 &= 0.002457 \\ s_2 &= 0.04957 \\ s_p &= 0.52281 \\ t_{ber} &= -0.885422 \\ t_{0.025}(12d.f.) &= -2.18 \end{aligned}$$

  H_0 aanvaarden.

6. Oef 6

7. Voor de punten van 60 studenten, bekomt men een gemiddelde van 11.2 met een standaarddeviatie van 0.9. Kan met met 95% betrouwbaarheid stellen dat het gemiddelde kleiner is dan 12 ?
   $H_0 : \mu = 12$
   $H_1 : \mu < 12$

$$\begin{aligned} n &= 60 \\ \overline{x} &= 11.2 \\ s &= 0.9 \end{aligned}$$

$$\begin{aligned} t_{ber} &= -6.8853 \\ -t_{0.95}(59d.f.) &= -1.67 \end{aligned}$$

We verwerpen H_0, de bewering was dus waar.

8. In een groep van 200 jongeren zijn er 120 die roken. De theoretische kans p_0 is niet gekend. Kan men met ee nbetrouwbaarheid van 99% stellen dat meer dan de helft van de jongeren rookt ?
$H_0 : p_0 = \frac{1}{2}$
$H_0 : p_0 > \frac{1}{2}$

$$\begin{aligned} z_{ber} &= \frac{\frac{3}{5} - \frac{1}{2}}{\sqrt{\frac{\frac{1}{2}.\frac{1}{2}}{200}}} \\ &= 2.82843 \\ Z_{0.490} &= 2.32667 \end{aligned}$$

H_0 verwerpen, het gestelde was dus waar.

9. Men gooit 100 keer met een munstuk en bekomt 60 keer kruis. Is het muntstuk vervalst met $\alpha = 0.05$? Bepaal het significantieniveau.
$H_0 : p_0 - \frac{1}{2}$
$H_1 : P_0 \neq \frac{1}{2}$

$$\begin{aligned} z_{ber} &= \frac{\frac{3}{5} - \frac{1}{2}}{\sqrt{\frac{\frac{1}{2}.\frac{1}{2}}{100}}} \\ &= 2 \\ Z_{0.475} &= 1.96 \end{aligned}$$

We verwerpen H_0, het muntstuk is inderdaad vervalst.

8 Regressie

9 Herhalingsoefeningen

9.1 Reeks1: hoofdstukken 1,2,3,4,5

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 én minstens 1 blauwe bal getrokken wordt ? (Gebruik de eigenschappen van de kansrekening om deze berekeningen te vereenvoudigen).
   Eerst gaan we over op het complement:
   $$1 - \left(\overline{groen \wedge blauw}\right)$$
   Nu passen we de wetten van de Morgan toe:
   $$1 - \left(\overline{groen} \vee \overline{blauw}\right)$$
   Nu passen we de definitie van OF toe en werken we verder uit:
   $$\begin{array}{l} 1 - \overline{groen} - \overline{blauw} + \overline{groen} \wedge \overline{blauw} \\ 1 - \dfrac{C_{14}^7}{C_{20}^7} - \dfrac{C_{16}^7}{C_{20}^7} + \dfrac{C_{10}^7}{C_{20}^7} \\ 0.809701 \end{array}$$

2. Een variable x heeft kansfuntie:
   $$f(x) = \begin{cases} a(x^2 - 2x) & \text{als } x \in [0,2] \\ 0 & \text{anders} \end{cases}$$
   - Bepaal a en toon aan dat de kansfunctie symmetrisch is tegenover het gemiddelde.
     Eerst bepalen we de a:
     $$\begin{aligned} \int_0^2 a \cdot \left(x^2 - 2x\right) dx &= 1 \\ a\left[\frac{x^3}{3} - \frac{2x^2}{2}\right]_0^2 &= 1 \\ a\left[\frac{x^3}{3} - x^2\right]_0^2 &= 1 \\ a\left(\frac{8}{3} - 4\right) &= 1 \\ a\left(-\frac{4}{3}\right) &= 1 \\ a &= -\frac{3}{4} \end{aligned}$$
     Om de symmetrie te kunnen afleiden moeten we de scheefheid van de functie kunnen bepalen, dit is mogelijk op twee manieren,

enerzijds via de momentenfunctie anderzijds via de definitie [17] van de scheefheid, ik verkies hier om het via de definitie te doen. Hiervoor hebben we echter het gemiddelde en de standaarddeviatie nodig. Deze dienen we eerste te bepalen.

$$\begin{aligned}
\mu &= E[x] \\
&= -\frac{3}{4}\int_0^2 x^3 - 2x^2 dx \\
&= -\frac{3}{4}\left[\frac{x^4}{4} - \frac{2x^3}{3}\right]_0^2 \\
&= -\frac{3}{4}\left(4 - \frac{16}{3}\right) \\
&= -\frac{3}{4} \cdot -\frac{4}{3} \\
&= 1
\end{aligned}$$

Nu weten we dat het gemiddelde in het midden van het interval ligt, dus er is al een grote kans dat we van symmetrie kunnen spreken, nu zoeken we de dispersie door eerst de variantie te bepalen. [18]

$$\begin{aligned}
\sigma^2 &= V[x] \\
&= E\left[x^2\right] - \mu^2 \\
&= -\frac{3}{4}\int_0^2 x^4 - 2x^3 dx - 1 \\
&= -\frac{3}{4}\left[\frac{x^5}{5} - \frac{x^4}{2}\right]_0^2 - 1 \\
&= -\frac{3}{4}\left(\frac{32}{5} - 8\right) - 1 \\
&= \frac{6}{5} - 1 \\
\sigma^2 &= \frac{1}{5} \\
\sigma &= \frac{1}{\sqrt{5}}
\end{aligned}$$

Nu we dit alles kennen bepalen we de scheefheid:

$$S = \left[\left(\frac{x-\mu}{\sigma}\right)^3\right]$$

[17] $S = \left[\left(\frac{x-\mu}{\sigma}\right)^3\right]$

[18] Het is perfect mogelijk de symmetrie aan te tonen zonder de dispersie eerste te bepalen maar we kunnen dit op dit punt nog zeker weten zonder voorkennis

$$
\begin{aligned}
&= \left[\left(\frac{x-1}{\sigma}\right)^3\right] \\
&= -\frac{3}{4\cdot\sigma^3}\int_0^2 (x-1)^3\cdot\left(x^2-2x\right)dx \\
&= -\frac{3}{4\cdot\sigma^3}\int_0^2 \left(x^3-3x^2+3x-1\right)^3\cdot\left(x^2-2x\right)dx \\
&= -\frac{3}{4\cdot\sigma^3}\int_0^2 \left(x^5-5x^4+9x^3-7x^2+2x\right)dx \\
&= -\frac{3}{4\cdot\sigma^3}\left[\frac{x^6}{6}-x^5+\frac{9x^4}{4}-\frac{7x^3}{3}+x^2\right]_0^2 \\
&= -\frac{3}{4\cdot\sigma^3}\left(\frac{32}{3}-32+36-\frac{56}{3}+4\right) \\
&= -\frac{3}{4\cdot\sigma^3}\cdot(8-8) \\
S &= 0
\end{aligned}
$$

Hieruit volgt de symmetrie.

- Bereken $P(x \leq 0.5)$ exact.

$$
\begin{aligned}
P(x\leq 5) &= -\frac{3}{4\cdot\sigma^3}\int_0^{\frac{1}{2}}\left(x^2-2x\right)dx \\
&= -\frac{3}{4}\left[\frac{x^3}{3}-x^2\right]_0^{\frac{1}{2}} \\
&= -\frac{3}{4}\left(\frac{1}{24}-\frac{1}{4}\right) \\
&= -\frac{3}{4}\cdot-\frac{5}{24} \\
&= \frac{5}{32}
\end{aligned}
$$

- Bepaal een boven of ondergrens voor $P(x \leq 0.5)$ met Chebychev. Uit Chebychev volgt:

$$
\begin{aligned}
P\left(|x-\mu|\geq k\sigma\right) &\leq \frac{1}{k^2} \\
P\left(|0.5-1|\geq k\sigma\right) &\leq \frac{1}{k^2} \\
P\left(|-0.5|\geq k\sigma\right) &\leq \frac{1}{k^2}
\end{aligned}
$$

Hieruit bepalen we de waarde voor k:

$$0.5 \geq k\cdot\sigma$$

$$\frac{1}{2} \geq k \cdot \frac{1}{\sqrt{5}}$$
$$\frac{\sqrt{5}}{2} \geq k$$

Nu we k kennen volgt hieruit dat de kans die uit Chebychev volgt gelijk is aan: $P \leq \frac{4}{5}$, echter gaat Chebychev over een symmetrisch interval, dus is de hier bepaalde kans de kans dat er een afwijking van meer dan 0.5 tov het gemiddelde bestaat. We moeten deze kans dus nog delen door twee (want de verdeling is symmetrisch). Dus $P \leq \frac{2}{5}$.

3. Uit een poisson verdeelde populatie met $\lambda = 2$ worden 8 steekproefwaarden genomen. Noem i het aantal steekproefwaarden dat strikt groter is dan 2. Gevraagd:

   - de verdeling van i
     De veranderlijke i in dit geval kent een binomiale verdeling met $n = 8$. p volgt uit:

     $$\begin{aligned} P(x > 2) &= 1 - P(x = 0) - P(x = 1) - P(x = 2) \\ &= 1 - \frac{e^{-2} \cdot 2^0}{0!} - \frac{e^{-2} \cdot 2^1}{1!} - \frac{e^{-2} \cdot 2^2}{2!} \\ &= 0.323324 \end{aligned}$$

   - de gemiddelde waarde en de variantie van i

     $$\begin{aligned} \mu &= np \\ &= 0.323324 \cdot 8 \\ &= 2.58659 \\ \sigma^2 &= n \cdot p(1 - p) \\ &= 8 \cdot 0.323324(1 - 323324) \\ &= 1.75028 \\ \sigma &= 1.32298 \end{aligned}$$

4. Men trekt 5 kaarten, met terugleggen, uit een kaartspel met 32 kaarten. Interpreteer met Chebychev de kans om voor het aantal harten een waarde te bekomen in het interval $]\mu - 2\sigma, \mu + 2\sigma[$. Bereken ook de exacte kans.
   Eerst bepalen we de parameters van deze binomiale verdeling.

   $$\begin{aligned} \mu &= np \\ &= \frac{5}{4} \\ \sigma^2 &= np(1 - p) \end{aligned}$$

$$\begin{aligned} &= 5 \cdot \frac{1}{4} \cdot \frac{3}{4} \\ &= \frac{15}{16} \\ \sigma &= \sqrt{\sigma^2} \\ &= \frac{\sqrt{15}}{4} \end{aligned}$$

Met deze kennis proberen we Chebychev toe te passen.

$$\begin{array}{ccc} \mu - 2\sigma < & x & < \mu + 2\sigma \\ -2\sigma < & x - \mu & < 2\sigma \end{array}$$

We zullen dus het complement van Chebychev moeten toepassen:

$$\begin{aligned} P\left(|x-\mu| < 2\sigma\right) &\geq 1 - \frac{1}{4} \\ &\geq \frac{3}{4} \\ P\left(0 < x \leq 3\right) &\geq \frac{3}{4} \end{aligned}$$

De exacte kans is gegeven door:

$$\begin{aligned} P\left(0 < x \leq 3\right) &= 1 - P(4\heartsuit) - P(5\heartsuit) \\ P\left(0 < x \leq 3\right) &= 1 - C_5^4 \left(\frac{1}{4}\right)^4 \left(\frac{3}{4}\right)^1 - C_5^5 \left(\frac{1}{4}\right)^5 \left(\frac{3}{4}\right)^0 \\ P\left(0 < x \leq 3\right) &= 0.984375 \end{aligned}$$

5. Men trekt 1 kaart uit een kaartspel met 32 kaarten. Interpreteer met Chebychev de kans dat het aantal pogingen alvorens een eerste aas te trekken (met terugleggen) groter of gelijk is aan $\mu + \frac{3}{2}\sigma$. Bepaal ook de exacte kans.
We bepalen de parameters van deze geometrische verdeling.

$$\begin{aligned} \mu &= \frac{1}{p} \\ \mu &= 8 \\ \sigma^2 &= \frac{1-p}{p^2} \\ \sigma^2 &= \frac{1-\frac{1}{8}}{\left(\frac{1}{8}\right)^2} \\ \sigma^2 &= 56 \\ \sigma &= 2 \cdot \sqrt{14} \end{aligned}$$

Chebychev geeft:

$$\begin{aligned} P\left(|x-\mu| \geq \frac{3}{2}\sigma\right) &= \frac{1}{\left(\frac{3}{2}\right)^2} \\ P\left(x \geq 20\right) &= 0.44444 \end{aligned}$$

Deze oplossing konden we bekomen omdat $\mu - \frac{3}{2}\sigma$ buiten de mogelijk oplossingen zit, dus alle kansen zitten aan de rechterkant, omdat deze waarde negatief is en bovendien niet mogelijk in dit geval.
Vervolgens bepalen we de exacte kans.

$$\begin{aligned} P(x \geq 20) &= 1 - P(x=1) - P(x=2) - \cdot P(x=19) \\ &= 1 - 0.125 - 0.109375 - 0.095703 - 0.08374 \\ &\quad -0.073273 - 0.064114 - 0.056099 - 0.049087 \\ &\quad -0.042951 - 0.037582 - 0.032884 - 0.028774 \\ &\quad -0.025177 - 0.02203 - 0.019276 - 0.016867 \\ &\quad -0.014758 - 0.012914 - 0.011299 \\ &= 0.08 \end{aligned}$$

Het komt wel zoiets uit... somehow

6. De gemiddelde levensduur van een televisie is 10 jaar. Interpreteer met Chebychev de kans dat de levensduur van een televisie ligt in het interval $]\mu - \frac{4}{3}\sigma, \mu + \frac{4}{3}\sigma[$. Bereken ook de exacte kans.
We bepalen de parameters van deze Exponentiële verdeling:

$$\begin{aligned} \mu &= \theta \\ &= 10 \\ \sigma^2 &= \theta^2 \\ &= 100 \\ \sigma &= 10 \end{aligned}$$

Uit Chebychev volgt:

$$\begin{aligned} P\left(|x-\mu| < \frac{4}{3}\sigma\right) &\geq 1 - \frac{1}{\left(\frac{4}{3}\right)^2} \\ P\left(|x-10| < \frac{4}{3}\cdot 10\right) &\geq 0.4375 \\ P\left(10 - \frac{4}{3}\cdot 10 < x < 10 + \frac{4}{3}\cdot 10\right) &\geq 0.4375 \\ P\left(0 \leq x < \frac{70}{3}\right) &\geq 0.4375 \end{aligned}$$

De exacte kans is gegeven door:

$$\begin{aligned} P\left(x \leq \frac{70}{3}\right) &= \int_0^{\frac{70}{3}} \frac{1}{\theta} \cdot e^{-\frac{x}{\theta}} dx \\ &= -\left[e^{-\frac{x}{10}}\right]_0^{\frac{70}{3}} \\ &= 1 - e^{-\frac{7}{3}} \\ &= 0.903028 \end{aligned}$$

7. Een veranderlijke is $N(2,3)$-verdeeld. Men neemt 4 onafhankelijke steekproefwaarden x_1, x_2, x_3, x_4 en bepaalt hieruit $y = \frac{3x_1+x_2+2x_3+4x_4}{10}$. Geef voor y een interval, symmetrisch om zijn gemiddelde zodat men 84% kans heeft dat y tot dit interval behoort.
   y is dus een lineaire combinatie van normale verdelingen en zal dus ook normaal verdeeld zijn, we bepalen de parameters van y:

$$\begin{aligned} \mu &= \sum a_i \mu_i \\ &= 2\left(\frac{3}{10} + \frac{1}{10} + \frac{2}{10} + \frac{4}{10}\right) \\ &= 2 \cdot 1 \\ &= 2 \\ \sigma^2 &= \sum a_i^2 \cdot \sigma_i^2 \\ &= 9\left(\frac{9}{100} + \frac{1}{100} + \frac{4}{100} + \frac{16}{100}\right) \\ &= \frac{27}{10} \\ \sigma &= \sqrt{\frac{27}{10}} \end{aligned}$$

   y is dus $N\left(2, \sqrt{\frac{27}{10}}\right)$-verdeeld. Nu bepalen we de grens voor 42% uit de tabel van de genormeerd normale verdeling:

$$\begin{aligned} -Z_{42} < & \frac{x-\mu}{\sigma} < Z_{42} \\ -1.41 < & \frac{x-\mu}{\sigma} < 1.41 \\ -1.41 < & \frac{x-2}{\sqrt{\frac{27}{10}}} < 1.41 \\ -1.41 \cdot \sqrt{\frac{27}{10}} < & \; x - 2 < 1.41 \cdot \sqrt{\frac{27}{10}} \\ 2 - 1.41 \cdot \sqrt{\frac{27}{10}} < & \; x < 2 + 1.41 \cdot \sqrt{\frac{27}{10}} \\ -0.316866 < & \; x < 4.31687 \end{aligned}$$

8. De veranderlijke i is binomiaal verdeeld met $n = 19$ en $p = 0.5$. Bepaal:

   - $P(i \leq 8)$ exact (met zo weinig mogelijk rekenwerk)

$$\begin{aligned} P(i \leq 8) &= P(i=0) + P(i=1) + \ldots + P(x=8) \\ &= C_{19}^{0} \cdot 0.5^{0} \cdot 0.5^{19} + \ldots C_{19}^{8} \cdot 0.5^{8} \cdot 0.5^{11} \\ &= 0.5^{19} \left(C_{19}^{0} + \ldots + C_{19}^{8} \right) \\ &= 0.323803 \end{aligned}$$

   - $P(i \leq 8)$ benaderend (met gepaste limietstelling)
     We benaderen dus de normale verdeling met $N\left(np, \sqrt{np(1-p)}\right)$ of in dit geval $N\,(9.5, 2.17945)$. We normeren de grens

$$\begin{array}{cc} P(i \leq 8.5) & N\,(9.5, 2.17945) \\ \dfrac{8.5 - 9.5}{2.17945} & N(0,1) \\ -0.45883 & N(0,1) \end{array}$$

     Dankzij de tabel kunnen we de kans bepalen

$$\begin{aligned} P(i \leq 8.5) &= 1 - 0.5 - Z_{0.45883} \\ &= 1 - 0.5 - 0.17648 \\ &= 0.32352 \end{aligned}$$

   - de kritische waarde a zodat $P(i \geq a) \approx 0.03$ (exact en benaderend).
     Benaderend:

$$\begin{aligned} P(i \geq a) &\approx 0.03 \\ P(i \leq a) &\approx 0.5 + 0.47 \end{aligned}$$

     Tabel raadplegen

$$a \approx 0 + 1.881$$

     Uit de genormeerd normale halen en terugbrengen naar de initiële verdeling:

$$\begin{aligned} a &\approx 1.881 \cdot \sigma + \mu \\ a &\approx 13.6 \\ a &\approx 14 \end{aligned}$$

9. Stel dat x $\chi^2(20d.f.)$ verdeeld, y $N(1,2)$ verdeeld en z genormeerd normaal verdeeld zijn (x, y en z zijn onafhankelijk), bepaal de kritische waarde a zodat: $P(4x + (y-1)^2 + 4z^2 > a) = 0.95$

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 zijn onafhankelijk). Bepaal:

    - de kritische waarde a zodat: $P((y-2)^2 > a^2 z) = 0.95$
    - $P(x^2 + y^2 + 9z > 4y + 26)$

11. Het aantal afgestudeerden industieel ingenenieur is normaal verdeeld met gemiddeld 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.
    x is het aantal afgestudeerde industrieel ingenieurs $N(400, 40)$
    y is het aantal jobs $N(450, 20)$
    x-y is het aantal jobs dat niet ingevuld raakt $N(-50, \sqrt{40^2 + 20^2})$.

$$\begin{aligned} P(x-y>0) &= ? \\ Z_{\frac{0+50}{20\cdot\sqrt{5}}} &= 0.5-? \\ Z_{1.11803} &= 0.5 - 0.36818 \\ P(x-y>0) &= 0.13182 \end{aligned}$$

12. Een vaas bevat 6 rode, 10 witte en 4 zwarte ballen. Een bal wordt willekeurig uit de vaas getrokken en als volgt vervangen door één van een andere kleur: rood → zwart; wit → zwart; zwart → rood.
    Dan wordt een tweede bal getrokken. Gegeven dat de tweede bal wit is, wat is dan de (voorwaardelijke) kans dat de eerste getrokken bal zwart was.
    Regel van Bayes:

$$\begin{aligned} P(1Z/2W) &= \frac{P(2Z/1W)\cdot P(1W)}{P(2Z/1W)\cdot P(1Z) + P(2Z/1Z)\cdot P(1W) + P(2Z/1R)\cdot P(1R)} \\ P(1Z/2W) &= \frac{\frac{1}{2}\cdot\frac{1}{5}}{\frac{1}{2}\cdot\frac{1}{5} + \frac{9}{20}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{3}{10}} \\ P(1Z/2W) &= \frac{\frac{1}{10}}{\frac{1}{10} + \frac{9}{40} + \frac{3}{20}} \\ P(1Z/2W) &= \frac{4}{19} \end{aligned}$$

13. Een firma produceert lampen van een bepaald type met een gemiddelde levensduur van 1600 uur. Hoe groot meot een steekproef minstens

zijn opdat de kans dat het steekproefgemiddelde (wat de levensduur betreft) minstens 1500 uur zou zijn, meer dan 0.95 zou bedragen. (Gebruik de centrale limietstelling).
Via de centrale limietstelling weten we dat het steekproefgemiddelde normaal verdeeld is met $N\left(\mu_0, \frac{\sigma_0}{\sqrt{n}}\right)$.

$$\begin{aligned}
P\left(\overline{x} > 1500\right) &= 0.95 \\
P\left(x > \frac{1500-1600}{\frac{160}{\sqrt{n}}}\right) &= 0.45 \\
\frac{1500-1600}{\frac{160}{\sqrt{n}}} &= 1.645 \\
n &= \left(\frac{1.645 \cdot 160}{-100}\right)^2 \\
n &= 7
\end{aligned}$$

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.
Uit het eerste gegeven volgt:

$$\begin{aligned}
P(x > 15) &= 0.05 \\
P\left((x > \frac{15-\mu}{\sigma}\right) &= 0.45 \\
\frac{15-\mu}{\sigma} &= 1.645
\end{aligned}$$

Uit het tweede gegeven volgt:

$$\begin{aligned}
P(x < 7) &= 0.20 \\
P\left((x < \frac{7-\mu}{\sigma}\right) &= 0.30 \\
\frac{7-\mu}{\sigma} &= -0.84 \\
\sigma &= \frac{7-\mu}{-0.84}
\end{aligned}$$

Wanneer we dit substitueren in de eerste formule bekomen we $\mu = 9.70423$, en $\sigma = 3.21932$

15. Een spel met inzet verloopt als volgt: een computer genereert volledig willekeurig (discreet uniform) één van de cijfers 0,1,2,3,4,5,6,7,8,9. Een speler gokt naar het juiste cijfer. Bij elke gok moet er 0.2 euro betaald worden. Er wordt verstandig gegokt (nooit twee keer hetzelfde cijfer)

dus er zijn hoogstens 10 gokpogingen. Indien de eerte gok direct juist is wordt 1 euro uitbetaald, de tweede gok juist levert 0,9 euro op,...,de tiende levert 0,1 euro op. Wat is de verwachte winst of verlies ?

$$\begin{aligned} W &= \frac{1}{10}(0.8 + 0.5 + 0.2 - 0.1 - 0.4 - 0.7 - 1.0 - 1.3 - 1.6 - 1.9) \\ W &= -0.55 \end{aligned}$$

Gemiddelde verliest men 0.55 euro.

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 personene die niet ziek zijn blijkt de 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.
Met behulp van de regel van Bayes: (Z=ziek, T=test is positief)

$$\begin{aligned} P(Z/T+) &= \frac{P(T+/Z) \cdot P(Z)}{P(T+/Z) \cdot P(Z) + P(T+/\overline{Z}) \cdot P(\overline{Z})} \\ &= \frac{0.99 \cdot 0.01}{0.99 \cdot 0.01 + 0.03 \cdot 0.99} \\ &= 0.25 \end{aligned}$$

9.2 Reeks2: hoofdstukken 6,7

1. Stel een 95% betrouwbaarheidsinterval op voor de verhouding $\frac{\sigma_2^2}{\sigma_1^2}$ van twee normaal verdeelde populaties. Een steekproef uit de eerste populatie van waarden geeft een steekproefvariantie van 16 , een steekproef uit de tweede populatie van 16 waarden geeft een steekproefvariantie van 12.

$$\begin{aligned} F_{0.975}(9, 15d.f.) < &\quad \tfrac{\sigma_2^2}{\sigma_1^2} \cdot \tfrac{s_1^2}{s_2^2} &< F_{0.025}(9, 15d.f.) \\ \frac{1}{F_{0.025}}(15, 9d.f.) \cdot \frac{s_2^2}{s_1^2} < &\quad \tfrac{\sigma_2^2}{\sigma_1^2} &< F_{0.025}(9, 15d.f.) \cdot \frac{s_2^2}{s_1^2} \\ \frac{1}{3.77} \cdot \frac{12}{16} < &\quad \tfrac{\sigma_2^2}{\sigma_1^2} &< 3.12 \cdot \frac{12}{16} \\ 0.198939 < &\quad \tfrac{\sigma_2^2}{\sigma_1^2} &< 2.34 \end{aligned}$$

2. Bij een medisch onderzoek onder bejaarden stelt men vast dat van de 260 onderzochte personen er 12 nierklachten hebben.

- Bepaal het 95% betrouwbaarheidsinterval voor de kans op nierklachten bij bejaarden.

$$
\begin{array}{rcl}
-z_{0.475} < & \frac{p-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} & < z_{0.475} \\
-1.96 \cdot \sqrt{\frac{p_0\,(1-p_0)}{n}} + p < & p_0 & < 1.96 \cdot \sqrt{\frac{p_0\,(1-p_0)}{n}} + p \\
0.0267 < & p_0 & < 0.093
\end{array}
$$

- Bepaal de grens a van het 95% betrouwbaarheidsinterval $[0, a]$ van het aantal bejaarden met nierklachten.

3. Uit een populatie die $N(\mu, 3)$ verdeeld is neemt men een steekproef van omvang 9. Men bekomt $\overline{x} = 26.5$ en s=3.3. Kan je op 99% betrouwbaarheid beweren dat $\mu = 24$? Test dit tweezijdig als éénzijdig.
   Tweezijdig:

$$
\begin{aligned}
H_0 &: \mu = 24 \\
H_1 &= \mu \neq 24 \\
z_{ber} &= \frac{26.5 - 24}{\frac{3}{\sqrt{9}}} \\
&= 2.5 \\
z_{0.495} &= 2.575
\end{aligned}
$$

   We aanvaarden H_0
   Éénzijdig:

$$
\begin{aligned}
H_0 &: \mu = 24 \\
H_1 &= \mu > 24 \\
z_{ber} &= \frac{26.5 - 24}{\frac{3}{\sqrt{9}}} \\
&= 2.5 \\
z_{0.490} &= 2.326
\end{aligned}
$$

   We verwerpen H_0

4. Men meet de maximale temperatuur, die normaal verdeeld is, in de noordelijke zone en de zuidelijke zone van een land. Een steekproef van 21 metingen in de noordelijke zone geeft $\overline{x}_1 = 15.6$ en $s_1^2 = 9.5$. Een steekproef van 25 metingen in de zuidelijke zonde geeft $\overline{x}_2 = 19.8$ en $s_2^2 = 13.6$.
   Gevraagd:

- Zijn de varianties gelijk met $\alpha = 0.05$?

$$
\begin{aligned}
H_0 &: \frac{\sigma_2^2}{\sigma_1^2} = 1 \\
H_1 &: \frac{\sigma_2^2}{\sigma_1^2} \neq 1 \\
f_{ber} &= \frac{s_1^2}{s_2^2} \\
&= 0.698529
\end{aligned}
$$

De varianties zijn gelijk indien:

$$
\begin{array}{rcl}
f_{0.975}(20, 24d.f.) < & f_{ber} & < f_{0.025}(20, 24d.f.) \\
\dfrac{1}{f_{0.025}(24, 20d.f.)} < & f_{ber} & < f_{0.025}(20, 24d.f.) \\
0.41 < & 0.698529 & < 2.33
\end{array}
$$

H_0 aanvaarden we, we mogen de varianties gelijk beschouwen.

- Kan men besluiten dat $\mu_2 - \mu_1 > 2$ op 95% en op 99% betrouwbaarheid ?