Hypothese Toetsen

Inhoudsopgave

Hypothese toetsing

De Poisson verdeling

Herziene uitgave 2013 voor wiskunde D vwo 6, 40 slu

Colofon

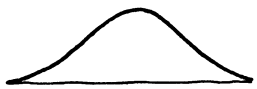
Auteurs: Leon van den Broek, Maris van Haandel, Dolf van den Hombergh, Aafke Piekaar, Daan van Smaalen

Illustraties: Wilson Design, Uden

Distributie: Iddink voortgezet onderwijs bv, Postbus 14, 6710 BA Ede

Website: www.wageningse-methode.nl

1 Discreet en continu



1 Verkeersdoden
Het aantal verkeersdoden in Gelderland is de laatste jaren normaal verdeeld met gemiddelde 105 en standaardafwijking 15. Hoe groot is in een jaar de kans op minder dan 100 verkeersdoden? We volgen twee berekeningen.
1) normcdf(0,100,105,15)
2) "minder dan 100" is hetzelfde als "hoogstens 99", dus normcdf(0,99,105,15).
a. Welke antwoorden geven deze berekeningen?

De kans op precies 100 verkeersdoden is nogal klein.
b. Weet jij een manier om die kans te berekenen?

Anneke berekent de kans op precies 100 verkeersdoden zo: normcdf(100,100,105,15).
c. Welk antwoord vindt ze? Kan dat goed zijn?

De kans dat het aantal verkeersdoden groter is dan 99 en kleiner is dan 100 is natuurlijk 0. Anneke berekent deze kans met normcdf(99,100,105,15).
d. Welk antwoord vindt ze?

We hebben dus een probleem. Dat wordt veroorzaakt doordat de stochast "het aantal verkeersdoden in een jaar" alleen *gehele* waarden kan aannemen. We benaderen hem met een normale verdeling, die ook *andere* waarden kan aannemen.

Stochasten, dus grootheden die bij een kansexperiment allerlei waarden kunnen aannemen, afhankelijk van het toeval, zijn te onderscheiden in twee soorten: discrete en continue.

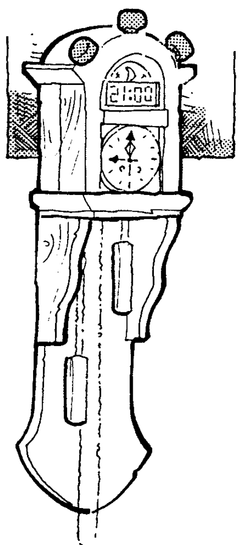
Een **discrete stochast** verandert trapsgewijs. De mogelijke uitkomsten zijn een rij losse punten op de getallenlijn. Bij twee opvolgende uitkomsten kan geen enkele waarde daartussen als uitkomst optreden.

Voorbeelden
- Kansexperiment: je stopt zes verschillende brieven in zes verschillend geadresseerde enveloppen.
  X = het aantal brieven dat in de juiste envelop zit.
- Kansexperiment: je gooit met een dobbelsteen tot je voor het eerst een zes gegooid hebt.
  Y = het aantal worpen dat daarvoor nodig is.

2 **a.** Welke waarden kan *X* aannemen?
b. En welke waarden kan *Y* aannemen?

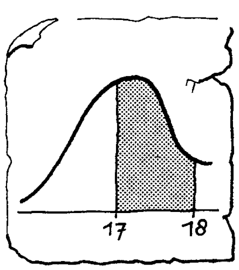
Een **continue stochast** verandert traploos. De mogelijke uitkomsten zijn een interval op de getallenlijn. Bij elk tweetal uitkomsten kan elke waarde daartussen ook als uitkomst optreden.

Voorbeelden

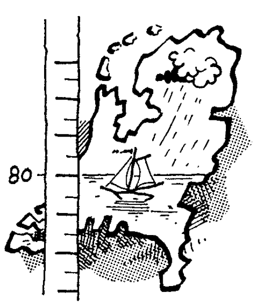


- Kansexperiment: je kiest een tomaat uit een grote partij tomaten.
  G = het gewicht van die tomaat.
- Kansexperiment: je neemt de tijd op waarin iemand de 100 meter loopt.
  H = de tijd die je klokt.
- Kansexperiment: je kijkt op een willekeurig moment naar een klok.
  T is de tijd die de klok aangeeft (in minuten na het vorige hele uur.

3 Hoe, discreet of continu, is *T* verdeeld en welke waarden kan *T* aannemen
a. bij een gewone (analoge) klok?
b. bij een digitale klok (die niet de seconden aangeeft)?



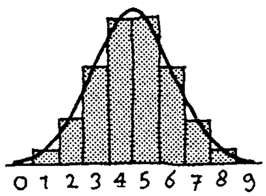
Continue en discrete stochasten worden grafisch verschillend weergegeven. Bij een discrete stochast teken je een kanshistogram, bij een continue stochast een vloeiende kromme.
Bij een kanshistogram is de oppervlakte van de staven een maat voor de bijbehorende kans en is de som van die oppervlaktes 1. De oppervlakte van twee staven, bijvoorbeeld de staven bij 17 en 18, is gelijk aan de kans dat de uitkomst 17 of 18 is.
Bij een continue stochast is de totale oppervlakte onder de bijbehorende kromme gelijk aan 1. De oppervlakte onder de kromme op een interval, bijvoorbeeld dat met eindpunten 17 en 18, is gelijk aan de kans dat een uitkomst in dat interval (dus tussen 17 en 18) ligt.



4 *X* is het aantal keren kop bij een worp met vijf munten.
a. Wat is het verschil tussen P(X = 3) en P($2 < X < 4$)?

Y is de jaarlijkse hoeveelheid neerslag in Nederland. *Y* is normaal verdeeld met gemiddelde μ = 80 cm en standaardafwijking σ = 15 cm.
b. Wat is het verschil tussen P($Y = 78$) en P($77 < Y < 79$) ?

In de praktijk worden de gemeten waarden van een continue stochast vrijwel altijd afgerond. Een jaarlijkse neerslag van 78 cm betekent dan dat de hoeveelheid neerslag tussen 77,5 en 78,5 cm ligt.



5 X is het aantal keren kop bij negen worpen met een munt.
a. Is *X* een discrete of een continue stochast?
b. Ga na: E(X)= 4,5 en Sd(X) = 1,5 .
c. Bereken op vier cijfers na de komma: P(X = 6).

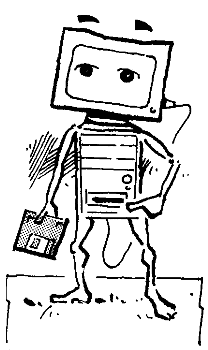
Hiernaast is het kanshistogram van *X* getekend. Erbij is de kromme getekend van de normaal verdeelde stochast *U* die het best bij *X* past. Dat wil zeggen: E(U) = E(X) = 4,5 en Sd(U) = Sd(X) = 1,5.
d. Bereken ook P($5{,}5 < U < 6{,}5$) en ga na dat dit een redelijke benadering is van P(X = 6).

e. Bereken P($X < 6$) en P($U < 5{,}5$) en ga na dat de uitkomsten ongeveer gelijk zijn.
f. Bereken P($X \leq 6$) en P($U \leq 6{,}5$) en ga na dat de uitkomsten ongeveer gelijk zijn.

Laat *X* een discrete stochast zijn die alleen gehele waarden aanneemt en laat *U* de continue benadering zijn van X. Dan geldt:

- P(X = 6) $\approx$ P($5{,}5 \leq U \leq 6{,}5$)
- P($3 \leq X \leq 8$) $\approx$ P($2{,}5 \leq U \leq 8{,}5$)
- P($3 < X \leq 8$) = P($4 \leq X \leq 8$) $\approx$ P($3{,}5 \leq U \leq 8{,}5$)
- P($3 \leq X < 8$) = P($3 \leq X \leq 7$) $\approx$ P($2{,}5 \leq U \leq 7{,}5$)

Als je *U* gebruikt om kansen voor *X* te benaderen moet je de waarden dus corrigeren met 0,5. Dit heet de **continuïteitscorrectie**.



Kies het item *Binomiaal / Normaal* in het computerprogramma *Kans en Simulatie* van De Wageningse Methode. Dat vergelijkt een binomiale stochast met zijn normale benadering.

6 Laat X en U zijn zoals in het theoriehok op de vorige bladzijde.
Neem over en vul de ontbrekende getallen in:
a. $P(X < 7) \approx P\ (U \leq \_\ )$
b. $P(X \geq 9) \approx P(U \geq \_\ )$
c. $P(5 < X < 12) \approx P(\ \_ \leq U \leq \_\ )$
d. $P(10 > X \geq 2) \approx P(\ \_ \geq U \geq \_\ )$

7 Terug naar opgave **1**. Het aantal verkeersdoden in Gelderland is de laatste jaren bij benadering normaal verdeeld met gemiddelde 105 en standaardafwijking 15.
a. Hoe groot is in een jaar de kans op minder dan 100 verkeersdoden?
b. Bereken de kans op precies 100 verkeersdoden.

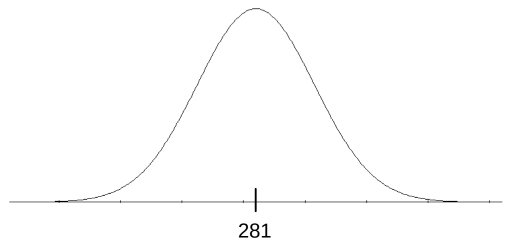
8 In een diepvriespak lekkerbekjes zitten volgens de verpakking 4 tot 6 wijtingfilets in beslag die samen een gewicht hebben van 300 gram. Neem aan dat het gewicht *Y* in zo'n pak normaal verdeeld is met een gemiddelde van 300 gram en een standaardafwijking van 15 gram.
a. Waarom zal de standaardafwijking in dit gewicht waarschijnlijk groter zijn dan de standaardafwijking in het gewicht van bijvoorbeeld een pak suiker?
b. Hoe groot is de kans dat het gewicht *Y* meer dan 10% afwijkt van de 300 gram die op de verpakking staat?

Pim koopt drie van deze diepvriespakken met een totaal gewicht van *T* gram.
c. Hoe groot zijn E(*T*), Var(*T*) en Sd(*T*)?
d. Hoe groot is de kans dat dit totale gewicht *T* meer dan 10% afwijkt van het te verwachten totale gewicht E(*T*)?
e. Kun je verklaren waarom de kans op deze afwijking van 10% duidelijk kleiner is dan de kans op een afwijking van 10% bij één pak?

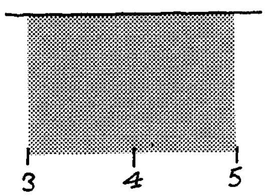
Neem aan dat er evenveel pakken zijn met 4 filets, met 5 filets en met 6 filets. Coen koopt ook drie diepvriespakken lekkerbekjes.
f. Hoe groot is de kans:
- dat er in één pak 4, in één pak 5 en in één pak 6 filets zitten?
- dat hij in totaal 12 filets koopt?
- dat hij in totaal 15 filets koopt?

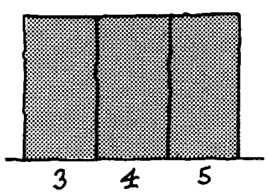
9 **De duur van een zwangerschap**

Als een vrouw zwanger is, wordt berekend op welke dag de geboorte valt te verwachten. Dat is de dag waarop de vrouw is "uitgeteld". 4 % van de geboortes vindt inderdaad plaats op de dag dat de vrouw is uitgeteld.
Een zwangerschap duurt gemiddeld 281 dagen. We veronderstellen dat de duur normaal verdeeld is.



a. Bereken de standaardafwijking van de duur van de zwangerschap.
b. Bereken het percentage zwangerschappen dat langer dan 295 dagen duurt.





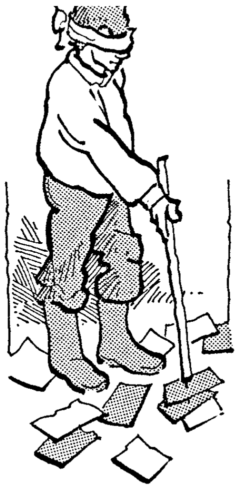
Overzichtsvragen

1 De discrete stochast X kan alleen de waarden 3, 4 en 5 aannemen. De continue stochast Y kan alle waarden tussen 3 en 5 aannemen. Beide kansverdelingen zijn uniform (dat wil zeggen dat alle mogelijke waarden evenveel kans hebben); zie nevenstaande plaatjes.
a. Welk plaatje hoort bij X en welk bij Y?

b. Bereken $P(X = 3)$ en $P(Y = 3)$.
c. Bereken $P(X < 5)$ en $P(Y < 5)$.
d. Bereken $P(4 \leq X \leq 5)$ en $P(4 \leq Y \leq 5)$.

2 De binomiale stochast X met 25 herhalingen en succeskans 0,4 laat zich goed benaderen door de normale stochast U.
a. Wat zijn het gemiddelde en de standaardafwijking van U?
b. Vul in: $P(X \leq 10) \approx P(U$ _____)
$P(X < 10) \approx P(U$ _____)
$P(X = 10) \approx P($ ___ U ___)

2 Wie heeft gelijk?



10 Populariteit van het kabinet

Volgens een artikel in een landelijk dagblad staat dat nog maar 40% van de bevolking achter het huidige kabinet. Volgens de minister-president klopt dat niet; hij is van mening dat dat percentage hoger is.

We willen achterhalen of die 40% juist is. Daartoe houden we op straat een kleine enquête onder 50 willekeurig gekozen voorbijgangers. (In de praktijk is het moeilijk ervoor te zorgen dat de geënquêteerden echt willekeurig worden gekozen. Aan dit soort problemen gaan we hier voorbij.) We leggen hen de volgende vraag voor: "staat u achter het kabinet?". We tellen het aantal mensen dat "ja" antwoordt. Dat aantal noemen we X.

a. Wie ben jij geneigd gelijk te geven als $X = 18$? En als $X = 30$?

Stel dat de 40% juist is.

b. Wat is dan de verwachtingswaarde $E(X)$?

c. Wat is de kans dat dan $X \geq 27$?

d. Veronderstel dat bij de enquête 29 mensen "ja" antwoorden. Wat denk jij dan van het percentage 40%?

11 Van exportkwaliteit

Een eierhandelaar heeft een belangrijke exportorder in de wacht gesleept. Hij moet nu op korte termijn veel eieren met een gewicht van minstens 65 gram zien te kopen. Een boer beweert dat 70% van zijn eieren voor export geschikt is. De handelaar vertrouwt die 70% niet. Hij bekijkt de partij eieren eens goed en beweert dat die 70% te hoog is.

Om hun dilemma te doorbreken, besluiten de boer en de handelaar 50 eieren te wegen. Het aantal eieren in de steekproef dat minder dan 65 gram weegt, noemen we Y.

a. Wie ben jij geneigd gelijk te geven als $Y = 13$? En als $Y = 27$?

Stel dat de boer gelijk heeft.

b. Wat is dan de verwachtingswaarde van Y?

Onder de aanname dat de handelaar gelijk heeft, ken je de kans p niet dat een ei lichter dan 65 gram is. Dan kun je $E(Y)$ niet berekenen; je weet alleen maar dat $E(Y) > 15$.

De boer en de handelaar spreken van tevoren af bij welke waarden van Y de handelaar gelijk krijgt: als $Y \geq 20$.

c. Bereken de kans dat $Y \geq 20$ terwijl de boer gelijk heeft.

Helaas voor de boer, in de steekproef had Y de waarde 24. De handelaar heeft de eieren niet gekocht.

12 Reden tot ongerustheid?

Onrust in Weurt

In het dorpje Weurt bij Nijmegen heerst grote onrust over een volgens de bevolking onrustbarend hoog aantal gevallen van kanker onder de 2600 inwoners.

Een op verzoek van de bewoners gehouden onderzoek van de GGD regio Nijmegen heeft de onrust alleen maar aangewakkerd. De GGD constateerde dat in de periode 1989-1992 bij mannen in Weurt 50 procent meer gevallen van kanker voorkwamen dan het landelijk gemiddelde. Er waren 33 gevallen van kanker geconstateerd, terwijl op basis van het landelijk gemiddelde 22 gevallen te verwachten waren.

Weurt (gemeente Beuningen) is aan drie kanten omgeven door industrieterreinen, waar een vuilverbrandingsoven, een ijzergieterij en andere zware industrie dagelijks hun afvalstoffen lozen. Volgens de bewoners zijn de fabrieken verantwoordelijk voor de kankergevallen en steeds meer voorkomende neus-, keel- en oogklachten.

Lees bovenstaand artikel uit *NRC-Handelsblad* van 19 januari 1995 (ingekort). Je mag aannemen dat de helft van Weurts bevolking mannelijk is.

a. Leg uit dat je uit het artikel kunt afleiden dat onder normale omstandigheden het percentage kankergevallen onder mannen ongeveer 1,7% is.

Stel dat in Weurt de kans op kanker even groot is als in de rest van Nederland, dus 0,017. Je kunt de mannelijke bevolking van Weurt dan beschouwen als een groep van 1300 willekeurige mannen. Het aantal kankergevallen in zo'n groep noemen we X.

X is binomiaal verdeeld. Hieronder staat een tabel voor deze kansen met x = 25 t/m 36.

x	$P(X \leq x)$	x	$P(X \leq x)$
25	0,7722	31	0,9731
26	0,8287	32	0,9829
27	0,8749	33	0,9895
28	0,9112	34	0,9936
29	0,9387	35	0,9963
30	0,9589	36	0,9979

b. Hoe groot is de kans op $X \geq 33$?

c. Heeft de bevolking van Weurt reden tot ongerustheid? Toelichten.

Je kunt de kansen uit de tabel ook normaal benaderen.

d. Benader met behulp van de normale verdeling $P(X \leq 29)$. Denk aan de continuïteitscorrectie.

In opgave **10**, **11** en **12** heb je drie probleemsituaties gezien. In alle drie staan er twee mogelijkheden (of hypothesen) tegenover elkaar:

- Er niets speciaals aan de hand is. Dat is de zogenaamde **nulhypothese**; die wordt genoteerd met H_0.
- Er wel iets (afwijkends) aan de hand is. Dat is de **alternatieve hypothese**; die wordt genoteerd met H_1.

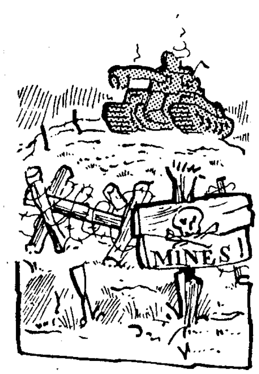
In opgave **10** over de populariteit van het kabinet is
H_0: het percentage dat achter het kabinet staat is 40%,
H_1: het percentage dat achter het kabinet staat is meer dan 40%.
Als we het percentage dat achter het kabinet staat aangeven met p, kunnen we de hypothesen zo formuleren:
H_0: $p = 0{,}4$ en H_1: $p > 0{,}4$.

13 Formuleer zo ook H_0 en H_1 voor opgave **11** en **12**.

Theorie
We hebben de volgende situatie. Iemand doet een bewering, een ander twijfelt aan de juistheid daarvan. Een **hypothesetoets** is een procedure om te beslissen wie gelijk krijgt. Daarbij heb je

- twee hypothesen: de **nulhypothese** H_0 en de **alternatieve hypothese** H_1,
- een **toetsingsgrootheid**; dat is het aantal X dat geteld wordt (of een gewicht dat gemeten wordt, of ...),
- een criterium dat zegt bij welke waarden van X de nulhypothese wordt verworpen. Deze waarden vormen het zogenaamde **kritieke gebied**.

Het kritieke gebied wordt zo bepaald dat de kans dat X een waarde aanneemt in het kritieke gebied kleiner is dan een vooraf afgesproken α. Deze α heet **significantieniveau** (of **onbetrouwbaarheidsdrempel**). Voor α neemt men vaak 0,05, 0,01 of zelfs 0,005, afhankelijk van hoe zwaarwegend de beslissing is.



Voorbeeld
We nemen α = 0,05. Bij opgave **10** is het aantal X dat achter het kabinet staat binomiaal verdeeld. We veronderstellen dat H_0 waar is. Dan is de succeskans $p = 0{,}4$ en het aantal herhalingen $n = 50$. Het kritieke gebied bestaat uit de waarden die groter dan of gelijk aan g zijn, zo dat $P(X \geq g) < 0{,}05$.
binomcdf(50,0.4,25) = 0,94265... en binomcdf(50,0.4,26) = 0,96859... . Het kritieke gebied is dus {27, 28, ... , 50}.

14 We nemen $\alpha = 0{,}05$.
Bepaal het kritieke gebied bij opgave **11** en **12**.

15 Bliksmaak
Sommige dranken zijn verkrijgbaar in blikjes en in flesjes. Coen beweert dat hij de bliksmaak herkent. Die bewering willen we toetsen: tien keer wordt hem op een presenteerblaadje twee glazen aangeboden. Het ene glas is gevuld met drank uit een flesje, het andere met dezelfde drank uit een blikje. Uit beide glazen neemt Coen een slok en zegt welk glas de drank uit blik bevat.
a. Definieer bij deze hypothesetoets de toetsingsgrootheid. Hoe is de toetsingsgrootheid verdeeld?

De kans dat hij een goed glas aanwijst, noemen we *p*. Misschien bluft Coen en kan hij de bliksmaak helemaal niet herkennen. Dan gokt hij maar. Dat geval nemen we als H_0, omdat dan *p* bekend is.
b. Wat zijn dus de twee hypothesen H_0 en H_1 die hier getoetst worden?

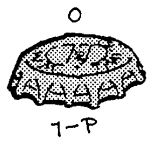
Laat *X* het aantal keren zijn dat Coen terecht het glas met de drank uit een blikje aanwijst.
c. Wat is onder H_0 de verwachtingswaarde van *X*?
d. Wat is het kritieke gebied? Kies $\alpha = 10\%$.

Bij de test bleek dat Coen bij acht van de tien drankjes het glas met de drank uit blik aanwees.
e. Concluderen we dat Coen bliksmaak kan herkennen?
f. Zou de conclusie hetzelfde zijn geweest bij $\alpha = 5\%$?

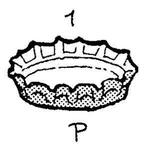
De redenering
Gegeven is een toetsingsgrootheid *X* en een significantieniveau α.
Er zijn twee hypothesen: H_0 en H_1.
Neem aan dat H_0 juist is.
Bepaal dan het kritieke gebied. Als H_0 waar is, kan *X* weliswaar een waarde in het kritieke gebied aannemen, maar het is wel (erg) onwaarschijnlijk dat dat gebeurt; de kans daarop is namelijk kleiner dan α.
We nemen een steekproef, houden een enquête of voeren een experiment uit. Als de waarde van *X* toch in het kritieke gebied blijkt te liggen, verwerpen we H_0 (en accepteren dus H_1).

Als *X* in het kritieke gebied zit, verwerp je H_0. Waarschijnlijk doe je dat terecht, maar dat is niet helemaal zeker. Het is dus mogelijk dat je een *verkeerde* beslissing neemt. Vandaar de term *kritiek* gebied.

16 Kop of munt



Bij het begin van een wedstrijd moet een munt opgegooid worden om uit te maken welke ploeg voor de wind mag spelen. Toevallig heeft niemand een munt bij de hand. Iets wat er op lijkt, een kroonkurk, is snel gevonden.
Maar het is de vraag of we een kroonkurk wel als munt kunnen gebruiken. Is de kans *p* op stand 1 (rand omhoog, zie tekening) wel even groot als de kans op stand 0 (rand naar beneden)?



a. Wat kun je zeggen van de grootte van kans *p* als je er vanuit gaat dat de kroonkurk inderdaad als munt gebruikt kan worden ("er is niets bijzonders aan de hand")? Dat is H_0.

Een kroonkurk is lang niet zo symmetrisch als een munt. Het kan best zijn dat hij vaker in stand 1 dan in stand 0 komt, of juist net andersom. Als dat zo is, is $p \neq 0,5$. Dat is H_1.

We gaan ook nu weer via een statistische toets beslissen of de kroonkurk wel of niet als munt gebruikt kan worden. We gooien de kroonkurk 50 keer op en tellen het aantal keren X dat de kroonkurk in stand 1 valt.
b. Wat zou jij beslissen als $X = 27$? En als $X = 44$? En als $X = 8$?

Merk op dat H_0 nu wordt verworpen bij waarden van X die sterk afwijken van $E(X)$, zowel afwijkingen naar boven als naar beneden.

We kiezen $\alpha = 0,05$. Het kritieke gebied bestaat uit twee stukken, elk met kans 0,025.
c. Wat is het kritieke gebied?
d. Stel dat de kroonkurk 18 keer in stand 1 valt.
Verwerp je dan de mening H_0 (en accepteer je dus H_1)?
e. En welke mening zou je accepteren als $X = 34$?
En als $X = 15$?

H_0 wordt in opgave **16** verworpen als X een erg grote waarde heeft, maar ook als X een erg kleine waarde heeft. Dat komt omdat $p \neq 0,5$ de alternatieve hypothese is. We spreken hier van een **tweezijdige toets**.

Toetsen waarbij H_0: $p = 0,3$ en H_1: $p > 0,3$ of waarbij H_0: $p = 0,65$ en H_1: $p < 0,65$ zijn **éénzijdige toetsen**.

17 In de vorige opgave hebben we gezien dat je, bij een significantieniveau van 5%, op grond van 18 keer stand 1 bij vijftig worpen met de kroonkurk H_0 niet kunt verwerpen: 18 zit niet in het kritieke gebied.
Men besluit voor alle zekerheid nog een serie van 50 worpen te doen. Het resultaat is nu: 20 keer stand 1. Dus weer geen resultaat waarbij H_0 verworpen wordt (bij α = 5%). Maar het valt wel op dat de afwijking van 25 (het onder H_0 te verwachten aantal) dezelfde kant op is.
De twee proeven samen vormen een serie van 100 worpen waarbij 38 keer stand 1 gegooid is.
a. Bereken het (tweezijdige) kritieke gebied voor de serie van 100 worpen (weer bij $\alpha = 0{,}05$).
b. Ga na dat op grond van het resultaat "38 keer stand 1" H_0 nu wel verworpen kan worden.

18 Mens erger je niet
Bij een spelletje *Mens erger je niet* heeft Harrie flink verloren. Volgens hem ligt dat aan de dobbelsteen; die zou niet helemaal in orde zijn. Hij had bij dat spelletje opvallend weinig zessen gegooid, terwijl zijn vriendinnetje Mady juist erg vaak een zes gooide. Harrie heeft op school net voor het eerst van een hypothesetoets gehoord en besluit die kennis meteen te gebruiken.
a. Harrie besluit een tweezijdige toets op te stellen. Ben jij het daarmee eens?
b. Welke hypothesen stelt Harrie op?

Hij gooit 100 keer met de dobbelsteen en telt daarbij op het aantal keren zes. Dat aantal noemen we Y.
c. Wat is het kritieke gebied bij $\alpha = 0{,}10$?

Na de 100 worpen blijkt hij 23 keer een zes te hebben gegooid.
d. Wordt de hypothese H_1 geaccepteerd?

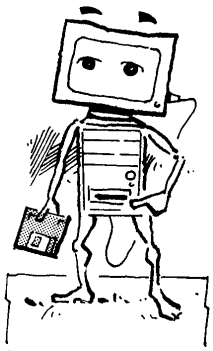
19 Bij een hypothesetoets is de toetsingsgrootheid Y binomiaal verdeeld met $n = 70$.

H_0: $p = 0{,}40$, H_1: $p \neq 0{,}40$ en $\alpha = 0{,}05$.
a. Wat is het kritieke gebied?
Wordt de hypothese H_1 geaccepteerd bij het steekproefresultaat $Y = 36$?

H_0: $p = 0{,}40$, H_1: $p > 0{,}40$ en $\alpha = 0{,}05$.
b. Wat is het kritieke gebied?
Wordt de hypothese H_1 geaccepteerd bij het steekproefresultaat $Y = 36$?

H_0: $p = 0{,}40$, H_1: $p < 0{,}40$ en $\alpha = 0{,}05$.

c. Wat is het kritieke gebied?
Wordt de hypothese H_1 geaccepteerd bij het steekproef-resultaat $Y = 21$?



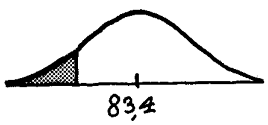
Het computerprogramma *De Normale Verdeling* op de schijf Statistiek van De Wageningse Methode rekent bij een toets het kritiek gebied in de verschillende gevallen uit.

Overzichtsvraag

1 A en B zijn het niet met elkaar eens over een succeskans p. A beweert dat $p = \frac{1}{4}$.
Om het geschil te beslechten, nemen ze een steekproef van grootte 100. Als significantie-niveau nemen ze 5%.
Bepaal het kritieke gebied in elk van de volgende gevallen.
a. B beweert dat $p < \frac{1}{4}$.
b. B beweert dat $p > \frac{1}{4}$.
c. B beweert dat $p \neq \frac{1}{4}$.

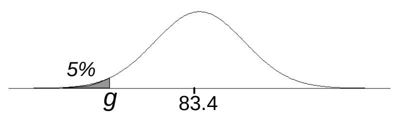
3 Toetsen van hypothesen

In deze paragraaf gaan we het toetsen van hypothesen toepassen op normaal verdeelde stochasten met gegeven standaardafwijking.



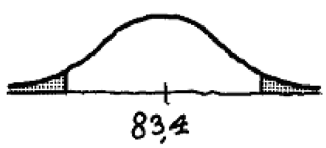
20 Bij een hypothesetoets is de toetsingsgrootheid X normaal verdeeld met $\sigma = 4{,}6$ en met onbekende μ.
Er zijn twee hypotheses: H_0: $\mu = 83{,}4$ en H_1: $\mu < 83{,}4$
a. Bereken onder aanname van H_0 de kans dat $X \leq 74{,}5$.
b. Neem $\alpha = 0{,}05$. Wordt bij het steekproefresultaat $X = 74{,}5$ de nulhypothese verworpen?

Om het kritieke gebied te vinden, willen we weten voor welke "grenswaarde" g geldt: $P(X \leq g) = 0{,}05$ onder H_0.



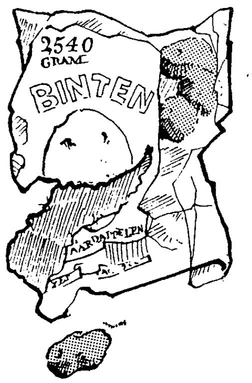
Die waarde van g bereken je op de GR als volgt:
DISTR, 3:invNorm(0.05,83.4,4.6) , ENTER
c. Wat is het kritieke gebied?

21 De toetsingsgrootheid X is normaal verdeeld met $\sigma = 4{,}6$ en met onbekende μ.
Er zijn twee hypotheses: H_0: $\mu = 83{,}4$ en H_1: $\mu > 83{,}4$
a. Wat is het kritieke gebied bij $\alpha = 0{,}05$?
b. Wordt de hypothese H_1 geaccepteerd bij het steekproefresultaat $X = 90{,}7$?



22 De toetsingsgrootheid X is normaal verdeeld met $\sigma = 4{,}6$ en met onbekende μ.
Er zijn twee hypotheses: H_0: $\mu = 83{,}4$, H_1: $\mu \neq 83{,}4$.
a. Wat is het kritieke gebied bij $\alpha = 0{,}05$?
b. Wordt de hypothese H_1 geaccepteerd bij het steekproefresultaat $X = 92{,}7$?

23 Zakken aardappelen
Zakken met 2,5 kg aardappelen bevatten natuurlijk zelden precies 2500 gram. Ontevreden klanten beweren dat er vaak te weinig in zit. Een leverancier beweert dat in zijn zakken van 2,5 kg gemiddeld 2540 gram aardappelen zit met een standaardafwijking van 80 gram.



Een consumentenvereniging doet een onderzoek. In verschillende winkels worden in totaal 40 van die zakken gekocht. De totale inhoud T van die zakken wordt gewogen. We mogen aannemen dat T normaal verdeeld is.
a. Als de leverancier het bij het rechte eind heeft, dus als er niets bijzonders aan de hand is, wat is dan het te verwachten totale gewicht?
b. Heb je hier te maken met een éénzijdige of met een tweezijdige toets? Waarom?
c. Welke twee hypothesen staan tegenover elkaar?
d. Hoe groot is, als H_0 juist is, Sd(T)?
e. Bepaal het kritieke gebied bij significantieniveau 5%.

De totale inhoud van de 40 zakken bleek 100,84 kg te zijn.
f. Worden de ontevreden klanten in het gelijk gesteld bij een significantieniveau van 5%?

24 Kunstmest

In deze opgave gaan we uit van een jaar van 365 dagen. In zo'n jaar telt januari 31, februari 28, maart 31 en april 30 dagen. We nummeren de dagen van het jaar vanaf 1 januari; 1 februari heeft dan nummer 32.

Voor de bemesting van grasland gebruikt men stikstofkunstmest. Uit onderzoek is gebleken dat de eerste bemesting in het voorjaar het hoogste rendement geeft als men direct na het bereiken van een temperatuursom (T-som) van 200 °C strooit.
De T-som is de som van de gemiddelde etmaaltemperaturen vanaf 1 januari. De gemiddelde etmaaltemperatuur per dag wordt telkens de volgende ochtend berekend en bij de vorige T-som opgeteld. Zodra de T-som meer dan 200 is, worden de boeren hiervan via de radio op de hoogte gebracht. De dag waarop dit gebeurt, noemen we de melddag.
Uit gegevens over lange tijd blijkt dat het nummer van de melddag bij benadering normaal verdeeld is met een gemiddelde van 105 en een standaardafwijking van 10.
a. Bereken de kans dat de melddag een dag in april is.

De mest moet beslist droog bewaard worden. Boeren en tussenhandelaren nemen deze daarom niet in voorraad. Zodra de melddag is aangebroken, wordt de mest bij kunstmestfabriek KF besteld. KF moet daar rekening mee houden. Bij het opstellen van een voorlopig jaarschema in december wenst KF dat het risico van een onvoldoende voorraad stikstofkunstmest op de melddag kleiner is dan 1%.

b. Bereken de uiterste datum die KF in het voorlopig jaarschema kan opnemen voor het op peil zijn van de voorraad kunstmest.

Eindexamen Wiskunde A 1993, tweede tijdvak, gedeeltelijk

25 Taaltest

Met deze test wordt onderzocht of iemand iets afweet van een bepaalde taal. De test bestaat uit tien vragen. Bij iedere vraag zijn er drie woorden in de vreemde taal gegeven met de bijbehorende Nederlandse woorden, maar die staan in een willekeurige volgorde.

Bijvoorbeeld:

vreemde taal	Nederlands
pour	kat
chat	voor
pauvre	arm

De proefpersoon moet bij elke vraag de juiste volgorde van de woorden aangeven door één van de zes mogelijke volgordes te kiezen. Hij krijgt dan zoveel punten als er woorden goed geplaatst zijn. Als hij het doet zoals in het voorbeeld (en de vreemde taal is Frans), dan krijgt hij 1 punt.

X_1 is het aantal punten dat bij de eerste vraag gescoord wordt.

Neem even aan (H_0) dat de proefpersoon niets van de vreemde taal weet.

a. Ga na dat dan $P(X_1 = 1) = \frac{1}{2}$.

b. Geef de kansverdeling van X_1 en ga na dat $E(X_1) = 1$ en $Var(X_1) = 1$.

Laat T het totaal aantal punten zijn dat de proefpersoon behaalt bij de tien vragen, dus $T = X_1 + X_2 + ... + X_{10}$.

c. Bereken $P(T=0)$, $P(T=1)$, $P(T=28)$ en $P(T=30)$.

d. Bereken $E(T)$ en $Var(T)$.

e. Waarom is T bij benadering normaal verdeeld?

f. Bereken de normale benadering van $P(T=15)$.

We willen met een hypothesetoets onderzoeken of de proefpersoon wel of niet iets van de taal afweet.

g. H_0: $\mu = 10$ is hierbij de nulhypothese.
Wat is de alternatieve hypothese H_1?

De proefpersoon scoort in totaal 17 punten.

h. Is deze score bij een significantieniveau van 5% voldoende om H_0 (de proefpersoon weet niets van de vreemde taal af) te verwerpen?

26 Omzet

Het bedrag dat in een week bij de kassa's van een supermarkt binnenkomt, is in zes van de tien weken meer dan € 40.000.
Neem aan dat de wekelijkse omzet X normaal verdeeld is met standaardafwijking € 6515.
a. Bereken E(X).

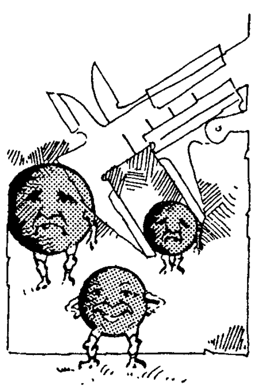
De eigenaar van een andere winkel beweert dat zijn omzet gemiddeld € 45.000 per week is met een standaardafwijking van € 5.000.
Een accountant constateert dat in de afgelopen vier weken de totale omzet € 168.743,36 was en vermoedt dat de gemiddelde weekomzet lager is dan € 45.000.
b. Is er, bij een significantieniveau van 5%, voldoende reden om de bewering van de eigenaar te verwerpen?

27 Regen

Volgens de VVV van het eiland Texel regent het daar in de zomer maar op 15% van de dagen. Anja gaat daar drie weken kamperen.
X is het aantal dagen dat het regent van de eenentwintig dagen dat Anja op Texel kampeert.

Neem aan dat X binomiaal verdeeld is.
Veronderstel dat de VVV gelijk heeft (dat is H_0).
a. Wat zijn dan E(X) en Sd(X)?
b. Waarom is het eigenlijk twijfelachtig of X wel binomiaal verdeeld is?

Neem aan dat X bij benadering normaal verdeeld is met standaardafwijking 1,6.
c. Als het 8 dagen van Anja's vakantie regent, is dat dan voldoende reden om de beweerde 15% van de VVV te verwerpen, bij significantieniveau 5%?



28 Kogeltjes

In een kogellagerfabriek worden stalen kogeltjes gemaakt die een diameter tussen 3,98 en 4,02 mm dienen te hebben. Door middel van twee zeven worden de kogeltjes gesorteerd. Het is inmiddels bekend dat de diameter van de kogeltjes normaal verdeeld is met standaardafwijking 0,008 mm, maar het gemiddelde μ is onbekend.
Van een aselecte steekproef van 250 kogeltjes werden er 14 uitgezeefd omdat ze te dun of te dik waren.

a. Neem aan dat $\mu = 4,005$.

Bereken de kans dat een aselect gekozen kogeltje een te grote of te kleine diameter heeft.
b. Wordt op grond van de uitkomst van de steekproef de hypothese H_0: $\mu = 4{,}005$ verworpen? Neem als significantieniveau $\alpha = 0{,}05$.

Naar: WS bulletin, januari 1982

29 De reactietijd van een automobilist is de tijd die hij nodig heeft om te reageren op iets wat hij ziet gebeuren. Voor automobilisten van 18 tot en met 40 jaar is de reactietijd normaal verdeeld met gemiddelde 0,4 seconde en standaardafwijking 0,05 seconde.

a. Bereken hoeveel tijd 90% van de 30-jarige automobilisten nodig hebben om te reageren.

Het is bekend dat de reactietijd toeneemt na alcoholgebruik.
Een 30-jarige automobilist beweert echter dat zijn reactietijd niet verandert na beperkt alcoholgebruik.
Men besluit een test te doen. Hierbij blijkt zijn reactietijd 0,49 seconde te zijn.
b. Onderzoek of deze reactietijd aanleiding geeft om, met een significantieniveau van 5%, de bewering van de automobilist te verwerpen.

30 Een bosbouwproject zoekt beleggers. We nemen aan dat de opbrengst van een hectare normaal verdeeld is. De projectleider beweert dat een hectare nieuwe aanplant na 20 jaar 625 kuub hout oplevert, met standaardafwijking 75 kuub.
Een potentiële belegger betwijfelt dat. Hij beweert dat de opbrengst van een hectare na 20 jaar 400 kuub is met een standaardafwijking van 60 kuub.
Om het meningsverschil op te lossen spreken de projectleider en de belegger het volgende af.
Volgend jaar wordt er weer van een hectare geoogst. Als de opbrengst groter is dan een zekere waarde g, krijgt de projectleider gelijk, anders de belegger

Neem $g = 490$.
a. Bereken de kans dat de projectleider gelijk heeft, maar de belegger gelijk krijgt.
b. Bereken de kans dat de belegger gelijk heeft, maar de projectleider gelijk krijgt.

c. Vind g zo dat de kans dat de kans dat de belegger onterecht gelijk krijgt gelijk is aan de kans dat de projectontwikkelaatr ten onechte gelijk krijgt.

Overzichtsvraag

1 De levensduur van gloeilampen is normaal verdeeld met gemiddelde 1200 branduren en standaardafwijking 160 branduren.

Een willekeurige gloeilamp ging al na 904 branduren stuk.

a. Is dat voldoende reden om het gemiddelde van 1200 branduren te verwerpen, bij significantieniveau 5%?

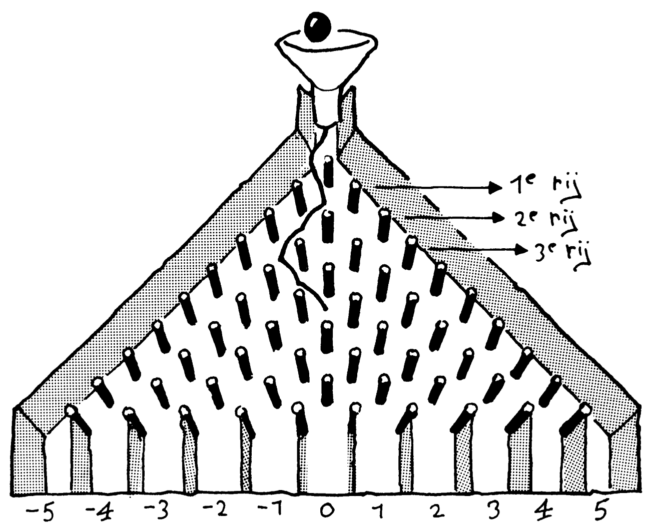
Vier willekeurige gloeilampen hadden een gemiddelde levensduur van 1084 branduren.

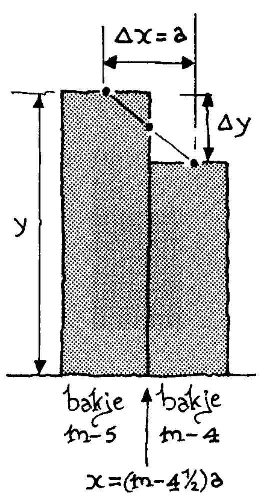
b. Is dat voldoende reden om het gemiddelde van 1200 branduren te verwerpen, bij significantieniveau 5%?

4 Van binomiaal naar normaal

De standaardnormale kromme is verwant aan de grafiek van de functie $y = e^{-x^2}$. Het is verrassend dat de e-macht ineens opduikt in de kansrekening. Wat heeft de functie $y = e^x$ te maken met de binomiaalcoëfficienten $\binom{n}{m}$?

We leggen het verband daartussen aan de hand van het Galtonbord met 10 rijen pinnen (maar dat is niet gemakkelijk).





We nummeren de bakjes; het middelste krijgt nummer 0. De breedte van de bakjes noemen we *a*. Elke keer als een balletje een pin raakt, springt het $\frac{1}{2}a$ naar rechts of $\frac{1}{2}a$ naar links.

Als een balletje *m* keer naar rechts gaat en dus 10–*m* keer naar links, komt het in bakje *m*–5; de kans daarop is: $\binom{10}{m} \cdot (\frac{1}{2})^{10}$.

Als een balletje *m*–1 keer naar rechts gaat en dus 9–*m* keer naar links, komt het in bakje *m*–4; de kans daarop is: $\binom{10}{m+1} \cdot (\frac{1}{2})^{10}$.

De grens tussen de twee bakjes ligt bij $x = (m-4\frac{1}{2})a$. Daar is de *y*-waarde van de kanspolygoon het gemiddelde van de kansen op *m* en *m*+1 successen: $\frac{1}{2} \cdot \left(\binom{10}{m} + \binom{10}{m+1}\right) \cdot (\frac{1}{2})^{10}$.

31 a. Ga na dat $\binom{10}{m+1} = \frac{10-m}{m+1} \cdot \binom{10}{m}$.

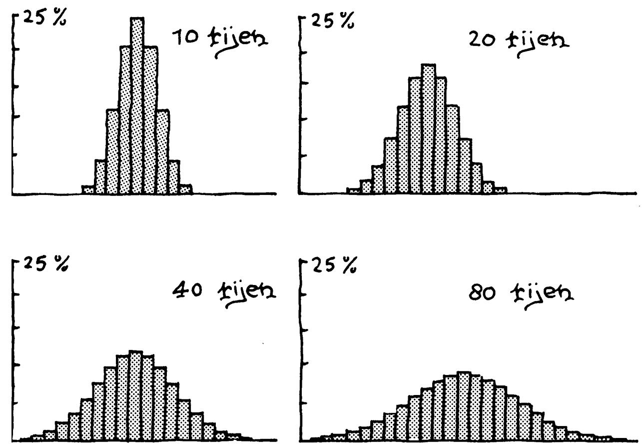
b. Laat zien dat bijgevolg de y-waarde bij $x = (m-4\frac{1}{2})a$ gelijk is aan $\frac{1}{2} \cdot \frac{11}{m+1} \cdot \binom{10}{m} \cdot (\frac{1}{2})^{10}$.

c. Laat zien dat de richtingscoëfficiënt van het lijnstukje van de kanspolygoon is: $\frac{\Delta y}{\Delta x} = \frac{1}{a} \cdot \frac{9-2m}{m+1} \cdot \binom{10}{m} \cdot (\frac{1}{2})^{10}$.

d. Ga na dat geldt: $\frac{\Delta y}{\Delta x} = -\frac{4}{11a^2} \cdot x \cdot y$.

Bij een Galtonbord met n rijen vinden we: $\frac{\Delta y}{\Delta x} = -\frac{4}{(n+1)a^2} \cdot x \cdot y$.

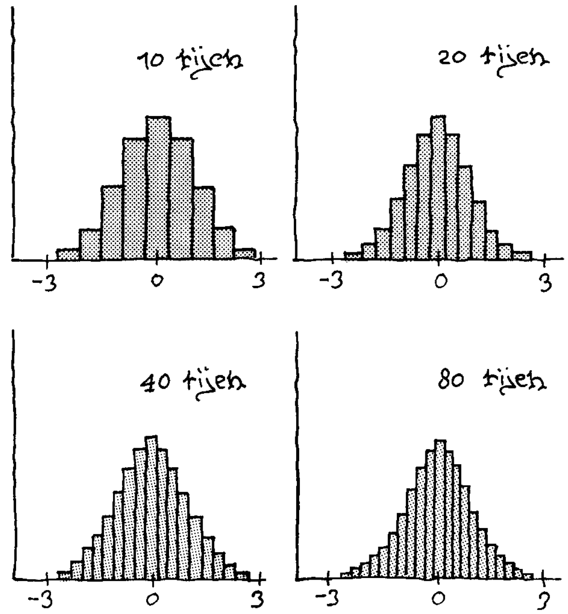
Bekijk nu de verdelingen voor n = 10, 20, 40 en 80. Ze worden steeds breder en lager.

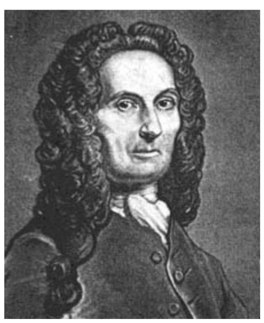


Om de normale verdeling in beeld te krijgen, moeten we de bakjes bij toenemende n steeds smaller maken. (Dit komt op hetzelfde neer als op de GR het window bijstellen.) We kiezen $a = \frac{2}{\sqrt{n+1}}$.

De betrekking tussen de richtingscoëfficiënt $\frac{\Delta y}{\Delta x}$ en de coördinaten van het punt (x,y) wordt dan:

$$\frac{\Delta y}{\Delta x} = -x \cdot y$$





Abraham de Moivre
(1667 - 1754)

Omstreeks 1720 ontdekte De Moivre de normale kromme als limietvorm voor de verdeling van het aantal keren kop bij een groot aantal worpen met een munt. Zijn wiskundige afleiding werd in die tijd niet opgemerkt.

Voor grote waarden van *n* kunnen we dit redelijk vervangen door $\frac{dy}{dx} = -x \cdot y$.

We zoeken een functie *y* waarvan de afgeleide gelijk is aan -*x* maal de functie zelf. En zo'n functie is wel te vinden: $y = e^{-\frac{1}{2}x^2}$. (*)

32 $y = c \cdot e^{-\frac{1}{2}x^2}$ waarbij c een constante is.

Laat zien dat geldt: $\frac{dy}{dx} = -x \cdot y$.

Hierna moet alleen nog de constante *c* zo gevonden worden dat de oppervlakte onder de grafiek gelijk is aan 1. Het blijkt dat $c = \frac{1}{\sqrt{2\pi}}$.

(*) $\frac{dy}{dx} = -x \cdot y$ is een zogenaamde differentiaalvergelijking. Hoe je hiervoor de oplossingsfuncties *y* kunt vinden, leer je in het hoofdstuk Continue Dynamische Modellen.

Antwoorden

Paragraaf 1 Discreet en continu

1 **a.** 0,3694 , 0,3446
b. ?
c. 0 ; nee
d. 0,0249

2 **a.** 0,1,2,3,4 en 6
b. Alle positieve gehele getallen.

3 **a.** Continu
b. Discreet

4 **a.** Er is geen verschil: $P(X = 3) = P(2 < X < 4)$.
b. $P(Y = 78) \approx 0$, $P(77 < Y < 79) \neq 0$.

5 **a.** Een discrete stochast.
b. $E(X) = 9 \cdot \frac{1}{2} = 4{,}5$, $Sd(X) = \sqrt{9 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 1{,}5$
c. binompdf(9,0.5,6) = 0,1641
d. normalcdf(5.5,6.5,4.5,1.5) = 0,1613
e. 0,7461 , 0,7475
f. 0,9102 , 0,9088

6 **a.** $P(U \leq 6{,}5)$
b. $P(U \geq 8{,}5)$
c. $P(5{,}5 \leq U \leq 11{,}5)$
d. $P(9{,}5 \geq U \geq 1{,}5)$

7 **a.** 0,3569
b. 0,0252

8 **a.** Een stuk filet weegt veel meer dan een korrel suiker. Met suiker is het gewicht dus nauwkeuriger af te passen.
b. $P(Y < 270) + P(Y > 330) \approx 0{,}0455$
c. $E(T) = 900$, $Var(T) = 675$, $Sd(T) \approx 25{,}98$
d. $P(T < 810) + P(T > 990) \approx 0{,}00053$
e. Een te licht pak wordt waarschijnlijk gecompenseerd door een te zwaar pak.
f. $\frac{2}{9}$, $\frac{1}{27}$, $\frac{7}{27}$

9 **a.** Zij U de continue benadering van de duur van de zwangerschap. $P(U < 281{,}5) = 0{,}52$.
normalcdf(0,281,σ) = 0,52 geeft $\sigma = 9{,}97$.
b. normalcdf(295.5,10000,281,9.97) ≈ 0,0729, dus 7,3% van de zwangerschappen duurt langer dan 295 dagen.

Paragraaf 2 Wie heeft gelijk?

10 a. Het dagblad ; de minister-president.
b. 20
c. 0,0314
d. Ik denk dan dat die 40% fout is.

11 a. De boer ; de handelaar.
b. 15
c. 0,0848

12 a. Normaal zijn 22 kankergevallen onder 1300 mannen. Dat is 22 : 1300 · 100% ≈ 1,7%
b. 1 – 0,9829 = 0,0171
c. Ja, want als er niets aan de hand is, is een aantal van 33 of meer kankergevallen erg onwaarschijnlijk.
d. E(X) = 22,1 , Sd(X) ≈ 4,661
normcdf(0,29.5,22.1,4.661) ≈ 0,9438

13 Bij opgave **11**:
H_0: de kans op een zwaar genoeg ei is 0,7.
H_1: de kans op een zwaar genoeg ei is kleiner dan 0,7.
Bij opgave **12**:
H_0: de kans op een kankerpatiënt is 0,017.
H_1: de kans op een kankerpatiënt is groter dan 0,017.

14 Bij opgave **2**: { 21,22 ... , 50}
Bij opgave **3**: {31,32, ... , 1300}

15 a. Het aantal keer dat Coen het juiste glas aanwijst. Binomiaal.
b. $p = \frac{1}{2}$; $p > \frac{1}{2}$
c. E(X) = 5
d. {8,9,10}
e. Ja
f. Nee

16 a. $p = \frac{1}{2}$
b. De kroonkurk kan wel als munt gebruikt worden ; niet ; niet.
c. {0,1, ... ,17,33, ... ,49,50}
d. Nee
e. H_1 ; H_1

17 a. {0,1, ... ,39,61, ... ,99,100}
b. 38 zit in het kritieke gebied.

18 a. Ja, want je moet onbevooroordeeld toetsen.
b. H_0: $p = \frac{1}{6}$; H_1: $p \neq \frac{1}{6}$
c. {0,1, ... , 10,24, ... 99,100}
d. Nee

19 a. {0,1, ... , 19,37, ... ,69,70} ; nee
b. {36,37, ... ,70} ; ja
c. {0,1,... , 20} ; nee

Paragraaf 3 Toetsen van hypothesen

20 a. 0,0265
b. Ja
c. De verzameling getallen ≤ 75,83.

21 a. De verzameling getallen ≥ 90,97.
b. Nee.

22 a. De verzameling getallen ≤ 74,38 of ≥ 92,42.
b. Ja

23 a. E(T) = 101,6 kg
b. Een eenzijdige toets, want de consumentenbond beweert dat er minder in zit.
c. H_0: $\mu = 101{,}6$, H_1: $\mu < 101{,}6$
d. $Sd(T) = \sqrt{0{,}256} \approx 0{,}5060$
e. De verzameling getallen ≤ 100,77.
f. Nee

24 a. Laat X het nummer van de melddag zijn. De melddag valt in april als 91 ≤ X ≤ 120.
P(91≤X≤120) = normcdf(90.5,120.5,105,10) ≈ 0.8659
b. Dagnummer 81, dus 22 maart.

25 a. Er zijn zes mogelijke rijtjes van 3 woorden. Bij drie daarvan staat één woord op de juiste plaats.
b. $P(X_1 = 0) = \frac{1}{3}$, $P(X_1 = 1) = \frac{1}{2}$, $P(X_1 = 2) = 0$, $P(X_1 = 3) = \frac{1}{6}$
$E(X_1) = \frac{1}{3} \cdot 0 + \frac{1}{2} \cdot 1 + \frac{1}{6} \cdot 3 = 1$
$Var(X_1) = \frac{1}{3} \cdot (0-1)^2 + \frac{1}{2} \cdot (1-1)^2 + \frac{1}{6} \cdot (3-1)^2 = 1$
c. $(\frac{1}{3})^{10}$, $10 \cdot \frac{1}{2} \cdot (\frac{1}{3})^9$, $10 \cdot \frac{1}{2} \cdot (\frac{1}{6})^9$, $(\frac{1}{6})^{10}$
d. E(T) = 10 , Var(T) = 10
e. Centrale limietstelling: als onafhankelijke stochasten bij elkaar opgeteld worden, dan gaat de verdeling steeds meer op de normale verdeling lijken, naarmate het aantal stochasten toeneemt.
f. normcdf(14.5,15.5,10,√10) = 0,0364
g. H_1: $\mu > 10$
h. Kritieke gebied is de verzameling getallen ≥ 15,2. Ja dus. (Met continuïteitscorrectie 15,7)

26 a. $\Phi(\frac{40000-\mu}{6515}) = 0{,}4$ geeft E(X) = € 41.651

b. $T = X_1 + X_2 + X_3 + X_4$, E(T) = 180.000 , Sd(T) = 10.000
H_0: μ = 180.000 , H_1: μ < 180.000
Kritieke gebied is de verzameling getallen ≤ 163.551. Nee dus.

27 a. E(X) = 3,15 , Sd(X) = 1,636
b. Het weer van de ene dag is van invloed op dat van de volgende dag. De onafhankelijkheid is dus niet gewaarborgd.
c. H_0: $\mu = 3{,}15$, H_1: $\mu > 3{,}15$
Kritieke gebied is de verzameling getallen ≥ 5,78. Ja dus.
(Met continuïteitscorrectie: 6,28)

28 a. 0,0313
b. Nee, want $P(X \geq 14 \mid n = 250,\ p = 0{,}0313) = 0{,}0271 > \frac{1}{2}\alpha = 0{,}025$

29 a. invNorm(0.90,0.4,0.05) = 0,464 sec.
b. P(X ≥ 0,49 | μ=0.4 , σ=0.05) = 0,369 < 5%, dus is er voldoende aanleiding om de bewering van de automobilist te verwerpen.

30 Zij *X* de opbrengst van de hectare.
a. $P(X < 490 \mid \mu = 625, \sigma = 75) = 0{,}0359$
b. $P(X \geq 490 \mid \mu = 400 , \sigma = 60 = 0{,}0668$
c. Maak een tabel van
$P(X < g \mid \mu = 625, \sigma = 75) - P(X \geq g \mid \mu = 400 , \sigma = 60)$
voor *g* vanaf 490:

490	0,03088
491	0,02768
492	0,02451
493	00,02137
494	0,01825
495	0,01515
496	0,01208
497	0,00903
498	0,00600
499	0,00299
500	0
501	-0.00300
502	-0,0059
503	-0,0089
504	-0,0118

Paragraaf 4 Van binomiaal naar normaal

31 a. $\binom{10}{m+1} = \frac{10!}{(m+1)!\cdot(9-m)!} = \frac{10!\cdot(10-m)}{(m+1)\cdot m!\cdot(10-m)\cdot(9-m)} =$

$\frac{10-m}{m+1}\cdot\frac{10!}{m!\cdot(9-m)!} = \frac{10-m}{m+1}\cdot\binom{10}{m}$

b. $y = \frac{1}{2}\left(1 + \frac{10-m}{m+1}\right)\cdot\binom{10}{m}\cdot(\frac{1}{2})^{10} =$

$\frac{1}{2}\left(\frac{m+1}{m+1} + \frac{10-m}{m+1}\right)\cdot\binom{10}{m}\cdot(\frac{1}{2})^{10} = \frac{1}{2}\cdot\frac{11}{m+1}\cdot\binom{10}{m}\cdot(\frac{1}{2})^{10}$

c. $\Delta x = a$ en $\Delta y = \binom{10}{m+1}\cdot(\frac{1}{2})^{10} - \binom{10}{m}\cdot(\frac{1}{2})^{10} =$

$\left(\binom{10}{m+1} - \binom{10}{m}\right)\cdot(\frac{1}{2})^{10} = \left(\frac{10-m}{m+1} - \frac{m+1}{m+1}\right)\cdot\binom{10}{m}\cdot(\frac{1}{2})^{10} =$

$\frac{9-2m}{m+1}\cdot\binom{10}{m}\cdot(\frac{1}{2})^{10}$

Dus $\frac{\Delta y}{\Delta x} = \frac{1}{a}\cdot\frac{9-2m}{m+1}\cdot\binom{10}{m}.(\frac{1}{2})^{10}$.

d. Vervang in c. – $9-2m$ door $\frac{-2}{a}$ x en $\frac{1}{m+1}\cdot\binom{10}{m}.(\frac{1}{2})^{10}$ door $\frac{2}{11}y$

32 Kettingregel: $\frac{d}{dx}\, c\cdot e^{-\frac{1}{2}x^2} = c\cdot e^{-\frac{1}{2}x^2}\cdot -\frac{1}{2}\cdot 2x =$

$-x\cdot c\cdot e^{-\frac{1}{2}x^2} = -x\cdot y$.

Poisson Verdeling

1 Wachten

Een winkel heeft gemiddeld per uur 10 klanten.
Wat is de kans dat in de komende vijf minuten geen klant komt?

Dit is een voor de handliggende vraag. Met de kansrekening die je tot nu toe geleerd hebt kun je deze vraag niet beantwoorden. We gaan proberen grip te krijgen op deze vraag. We nemen aan dat de klanten onafhankelijk van elkaar in de winkel komen. Ook nemen we aan dat elk moment van binnenkomst voor elk van de klanten even waarschijnlijk.

1 We nummeren de tien klanten die gemiddeld per uur in de winkel komen: 1, 2, 3, … , 10; bijvoorbeeld op volgorde van lichaamslengte (de kleinste krijgt nummer 1), of op grond van iets anders dat niets met hun aankomsttijden te maken heeft.
a. Wat is de kans dat klant nummer 1 niet in de komende vijf minuten arriveert?
b. Wat is de kans dat klant nummer 7 niet in de komende vijf minuten arriveert?
c. Wat is de kans dat geen van de tien klanten in de komende vijf minuten arriveren?

Klaar is Kees? Misschien denk je dat hiermee de vraag beantwoord is. Maar er was gegeven dat er *gemiddeld* tien klanten per uur zouden komen. Dat is iets anders dan dat er elk uur *precies* tien klanten komen. En waarom kijken we per uur. Je zou het gegeven ook kunnen vervangen door: "er komen gemiddeld twintig klanten per twee uur".

2 **a.** Bereken uitgaande van 20 klanten per twee uur wat de kans is op geen klanten in de komende vijf minuten.
b. Vergelijk je antwoord met dat van de vorige opgave.

Je antwoorden zijn niet gelijk, maar verschillen ook weer niet zo heel veel.

3 **a.** Wat is - onder de aanname dat er (precies) tien klanten per uur komen – de kans op precies 1 klant in de komende vijf minuten?
b. En op precies twee klanten de komende vijf minuten?

Onder de aanname dat er (precies) tien klanten per uur komen, is het aantal klanten in de komende vijf minuten binomiaal verdeeld met parameters *n* = 10 en "succes"-kans $p = \frac{1}{12}$.

Dat is een goede reden om nog eens naar de binomiale verdeling te kijken, en wel zonder GR.

Gegeven: er zullen onafhankelijk van elkaar 10 klanten komen tussen 14:00 en 15:00 uur.
We gaan zonder Grafische Rekenmachine de kans berekenen dat 3 van de 10 klanten tussen 14:15 en 14:30 uur komen. Zodoende wordt de theorie nog eens duidelijk.

4 Het aantal klanten dat tussen 14:15 en 14:30 uur arriveert is een typisch voorbeeld van een binomiaal verdeelde stochast.
Eén deelexperimentje dat hier speelt is: een klant komt in de winkel.
Die komt tussen 14:15 en 14:30 uur met kans ¼: de zogenaamde succeskans, meestal *p* geheten.
Er zijn 10 klanten: het experimentje wordt dus 10 keer herhaald: *n* = 10.
Het aantal klanten dat tussen 14:15 en 14:30 uur komt noemen we *X*.

a. Welke waarden kan *X* aannemen?

We nummeren de klanten 1 t/m 10.
b. Wat is de kans dat klant 2, 6 en 9 tussen 14:15 en 14:30 uur komen en alle andere niet?

Deze mogelijkheid noteren we met: NJNNNJNNJN.
c. Hoeveel rijtjes zijn er met 3 J's en 7 N's?

Al die rijtjes hebben dezelfde kans; die kans heb je in onderdeel b berekend.
d. Wat is dus P(*X*=3), de kans dat er drie klanten tussen 14:15 en 14:30 uur komen?

De afleiding van de binomiale verdeling
In opgave 4 werd gevraagd naar de kans op 3 successen, als het aantal herhalingen *n* = 20 is en de succeskans $p = \frac{1}{4}$. De kans op één speciaal rijtje met 3 successen (en dus 7 mislukkingen) is $(\frac{1}{4})^3 \cdot (\frac{3}{4})^7$.

Er zijn $\binom{10}{3}$ van zulke rijtjes.

Dus is de kans op 5 successen: $\binom{10}{3} \cdot (\frac{1}{4})^3 \cdot (\frac{3}{4})^7$.

5 Jan zit in een klas van 25 leerlingen.
a. Bereken zonder gebruik te maken van de GR de kans dat er precies twee klasgenoten zijn die in dezelfde week jarig zijn als Jan. Je kunt aan het eind wel een gewoon rekenmachientje gebruiken.
b. Controleer je antwoord op de GR.

6 Een vliegtuig heeft 200 zitplaatsen. Van de geboekte vluchten komt 5% van de passagiers niet opdagen. Daarom verkoopt de maatschappij niet 200 maar 204 zitplaatsen.
a. Bereken zonder GR de kans dat er precies één zitplaats te weinig zal blijken te zijn.
b. Controleer die kans op de GR.

Algemeen
Laat X het aantal successen zijn bij een binomiaal kansexperiment met n herhalingen, elk met succeskans p, dan is $P(X=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$, voor k = 0, 1, 2, …., n.

7 **a.** De formule laat zich vereenvoudigen in het geval $p = ½$. Hoe?
b. Welke afspraak moet je maken voor 0^0 om de formule ook te laten gelden in het geval $p = 0$?

8 Het binomium van Newton luidt: $(a+b)^n =$
$$\binom{n}{0} \cdot a^0 \cdot b^n + \binom{n}{1} \cdot a^1 \cdot b^{n-1} + \binom{n}{2} \cdot a^2 \cdot b^{n-2} + \ldots + \binom{n}{n} \cdot a^n \cdot b^0$$
Hierin zijn a en b willekeurige reële getallen en is n een positief geheel getal.
In de Appendix C wordt deze formule afgeleid.
a. Hoe ziet de formule eruit in de gevallen $n = 1$ en $n = 2$?
b. Leg uit dat uit het binomium van Newton volgt:
$$2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n}$$
c. Neem $b = 1 - a$. Wat is de uitkomst van:
$$\binom{n}{0} \cdot a^0 \cdot (1-a)^n + \binom{n}{1} \cdot a^1 \cdot (1-a)^{n-1} + \ldots + \binom{n}{n} \cdot a^n \cdot (1-a)^0\,?$$
Wat betekent dat in termen van een binomiaal kansexperiment?

De naam "binomiale verdeling" komt van het binomium van Newton. En die formule heeft zijn naam te danken aan het feit dat $a+b$ een tweeterm (binomium) is. De formule zegt dus hoe een macht van een tweeterm kan worden uitgewerkt.

De vraag waarmee we deze paragraaf begonnen is van het volgende type:

gegeven een *gemiddeld* aantal per tijdseenheid,
gevraagd de kans op een zeker aantal tijdens een tijdsinterval.

Dit type komt vaak voor. Bijvoorbeeld:

- het aantal brandmeldingen op een dag als er gemiddeld 523 zijn per jaar
- het aantal telefoontjes dat de belastingdienst in een uur krijgt als er gemiddeld 444 per dag binnenkomen
- het aantal dodelijke ongevallen van fietsers op een dag als er gemiddeld 120 per jaar zijn
- het aantal typfouten op een bladzijde, als er gemiddeld 292 zijn in een boek van 400 bladzijden.

9 Verzin zelf ook twee (heel) andere voorbeelden van dit type.

"Al onze medewerkers zijn in gesprek; u wordt zo spoedig mogelijk geholpen."

Wachten is irritant en meestal tijdverspilling. In totaal wacht een mens een heel jaar van zijn leven: op de bus, op een telefoontje, in de file, voor het verkeerslicht, voor de kassa, op de krant, op het weekend, bij een telefonische hulpdienst.

Geen wonder dat wachten een hot item is in de moderne maatschappij. Wachttijdtheorie is een onderdeel van de kansrekening. Centrale vragen daarin zijn:

- als er gemiddeld vijf klanten per uur komen, hoeveel mag je er dan in het komende kwartier verwachten?
- als je gemiddeld 20 minuten moet wachten op een lift, hoe groot is dan de kans dat je meer dan 1 uur op een lift moet wachten.
- hoeveel kassa's moeten er zijn om de wachttijd minder dan 5 minuten te houden?

2 De Poissonverdeling

Stel dat er in een winkel gemiddeld 10 klanten per uur komen.
Wat is dan de kans dat er in het komende kwartier precies 3 klanten komen.

Dit is het centrale thema van deze paragraaf. Wij gaan die kans precies berekenen

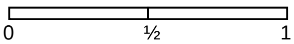
10 **a.** Waarom is het belangrijk voor de winkelier (bedrijfsleiding) om te weten wat de kans is op 3 klanten in een kwartier?
b. Hoe groot schat jij de kans op 3 klanten in een kwartier als er gemiddeld 10 per uur komen? Het gaat erom of je een idee hebt; je kunt natuurlijk onmogelijk zomaar die kans berekenen.

11 Stel dat je weet dat er komend uur precies 10 klanten komen. (Hoe je dat te weten bent gekomen doet er nu even niet toe.) Maar je hebt geen idee wanneer ze in die periode van een uur zullen komen: elk moment is voor elk van de klanten even waarschijnlijk.
Ze komen onafhankelijk van elkaar.
a. Wat betekent het dat ze onafhankelijk van elkaar komen?
b. Wat is de kans dat er 8 in het eerste halfuur komen en de andere 2 in het tweede halfuur?

Stel dat de winkelier geen personeel heeft en stel dat elke klant precies 5 minuten nodig heeft om geholpen te worden.
c. Hoe groot schat jij de kans dat geen enkele klant hoeft te wachten? Toelichten.

Sommige winkels proberen de wachttijden bewust klein te houden. Door deze service willen ze klanten winnen en vasthouden. Het is duidelijk dat het kunnen inschatten van wachttijden voor deze winkels erg belangrijk is. In dit hoofdstuk willen we zicht krijgen op deze problematiek.

We bekijken een periode van 1 uur als geheel en opgesplitst in twee periodes van ½ uur.
$P_1(n)$ is de kans op precies n klanten in dat uur,
$P_{½}(n)$ is de kans op n klanten in een halfuur,



12 Er geldt $P_1(0) = (P_{½}(0))^2$. Immers,
de kans op 0 klanten in een uur =
de kans op 0 klanten in het eerste half uur én 0 klanten in het tweede half uur =
de kans op 0 klanten in het eerste half uur ×
de kans op 0 klanten in het tweede half uur.

a. Druk zo ook $P_1(1)$ uit in $P_{½}(0)$ en $P_{½}(1)$.
b. Druk $P_1(3)$ uit in $P_{½}(0)$, $P_{½}(1)$, $P_{½}(2)$ en $P_{½}(3)$.

Als er 7 klanten in het eerste halfuur komen én 0 in het tweede halfuur, dan komen er 7 klanten in het hele uur én die 7 komen allemaal in de eerste helft van dat uur.
Dit vertalen we in kansen:

$$P_{½}(7) \times P_{½}(0) = P_1(7) \times (½)^7. \qquad (1)$$

13 a. Een zelfde redenering geeft:

$$P_{½}(6) \times P_{½}(1) = P_1(7) \times \_\_\_. \text{ Vul in.} \qquad (2)$$

$P_1(7)$ kennen we (nog) niet, maar die kans kunnen we weg laten vallen. Vergelijk maar de gelijkheden (1) en (2).

b. Leg uit dat hieruit volgt dat $P_{½}(7) = \frac{1}{7} \cdot \frac{P_{1/2}(1)}{P_{1/2}(0)} \cdot P_{½}(6)$.
c. Druk zo ook $P_{½}(n)$ uit in $P_{½}(n-1)$.
d. Controleer of de formule uit onderdeel **c** klopt voor het geval $n = 1$.

14 Als we $P_{½}(0)$ en $P_{½}(1)$ zouden kennen, zouden we alle kansen $P_{½}(n)$ kennen. Het gaat trouwens niet om de kansen $P_{½}(0)$ en $P_{½}(1)$ zelf, maar om hun verhouding $\frac{P_{1/2}(1)}{P_{1/2}(0)}$. Die verhouding noemen we λ.

a. Druk achtereenvolgens $P_{½}(1)$, $P_{½}(2)$, $P_{½}(3)$, $P_{½}(4)$ en $P_{½}(5)$ uit in λ en $P_{½}(0)$. (Gebruik opgave **13c**.)
b. Geef een formule voor $P_{½}(n)$, uitgedrukt in λ.

Als je $P_{½}(0)$ zou kennen, zou je λ kunnen uitrekenen (en omgekeerd, want de som van alle kansen $P_{½}(n)$ is 1.
Daarvoor moet je wel oneindig veel kansen optellen en dat is geen sinecure. We vinden:
$P_{½}(0) + P_{½}(1) + P_{½}(2) + P_{½}(3) + P_{½}(4) + \ldots =$
$P_{½}(0) + \lambda \cdot P_{½}(0) + \lambda^2/2! \cdot P_{½}(0) + \lambda^3/3! \cdot P_{½}(0) + \lambda^4?4! \cdot P_{½}(0) + \ldots =$
$P_{½}(0) \cdot (\ 1 + \lambda + \lambda^2/2! + \lambda^3/3! + 0{,}1 \cdot \lambda^4/4! + \ldots\ ) =$
$P_{½}(0) \cdot e^{\lambda}$.
En dit moet dus 1 zijn.

Dat de oneindige som $1 + \lambda + \lambda^2/2! + \lambda^3/3! + \lambda^4/4! + \ldots = e^\lambda$ wordt uitgelegd in appendix A.

15 **a.** Neem $P_{½}(0) = 0{,}1$; dan is dus 10% kans dat er in een half uur geen klanten komen in de winkel.
Hoe groot is λ dan?

b. Laat zien dat algemeen geldt: $\lambda = \ln \frac{1}{P_{1/2}(0)}$.

Een voorval, bijvoorbeeld een klant komt binnen, kan optreden of niet. We nemen aan dat de voorvallen onafhankelijk van elkaar optreden. We tellen het aantal keer dat een voorval optreedt in een zekere tijdsperiode. Dat aantal noemen we *X*.

De verhouding $\dfrac{\text{kans op 1 voorval}}{\text{kans op 0 voorvallen}}$ noemen we λ.

Dan is de kans op *k* voorvallen in die tijdsperiode

$$P(X{=}k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \text{ voor } k = 0, 1, 2, 3, \ldots$$

We zeggen dat *X* Poissonverdeeld is met parameter λ.

16. We gaan de verwachtingswaarde van een Poissonverdeelde stochast *X* met parameter λ uitrekenen.
Die verwachtingswaarde is per definitie:
$E(X) = 0 \cdot P(X{=}0) + 1 \cdot P(X{=}1) + 2 \cdot P(X{=}2) + 3 \cdot P(X{=}3) + 4 \cdot P(X{=}4) + \ldots$
Dit is een oneindige som.

a. Laat zien dat

$$E(X) = e^{-\lambda} \cdot \lambda \cdot \left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \frac{\lambda^4}{4!} + \ldots\right).$$

b. Bewijs dat $E(X) = \lambda$; gebruik de reeks in appendix A.

X is het aantal keer dat een voorval optreedt in een zekere tijdsperiode. De voorvallen treden onafhankelijk van elkaar op.
Zeg dat het voorval gemiddeld λ keer in de tijdsperiode optreedt (dat is dus de verwachtingswaarde).
Dan is *X* **Poissonverdeeld** met parameter λ:

$$P(X{=}k) = \frac{\lambda^k}{k\,!} \cdot e^{-\lambda} \text{ voor } k = 0, 1, 2, 3, \ldots$$

17 Het aantal klanten per uur is X en heeft parameter λ.
Het aantal klanten per half uur is Y en heeft parameter ½λ.

In opgave **12a** heb je gezien dat $P(X=0) = (P(Y=0))^2$, zonder de kansen van de Poissonverdeling te gebruiken.
a. Controleer dat de gelijkheid geldt met de Poisson-kansen $P(X=0)$ en $P(Y=0)$.

In opgave **12b** heb je gezien dat $P(X=1) = 2 \cdot P(Y=0) \cdot P(Y=1)$, zonder de kansen van de Poissonverdeling te gebruiken.
b. Controleer dat de gelijkheid geldt met de Poisson-kansen $P(X=1)$, $P(Y=0)$ en $P(Y=1)$.

Terugblik

We zijn geïnteresseerd in de kansverdeling van het aantal voorvallen in een uur als er gemiddeld λ voorvallen per uur plaatsvinden.
We nemen aan dat de voorvallen onafhankelijk van elkaar plaatsvinden.
We hebben gezien:

- Het aantal voorvallen per uur hangt samen met het aantal voorvallen per half uur.
- De kans op 7 voorvallen in een uur gelijk is aan 1/7 × λ de kans op 6 voorvallen in een uur, enz.
- De kans op 7 voorvallen in een uur is 1/7! × λ × de kans op 0 voorvallen in een uur.
- Omdat alle kansen samen 1 zijn, kon de kans op 0 voorvallen in een uur worden uitgerekend: die is $e^{-\lambda}$.

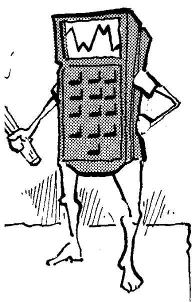
Als er per uur gemiddeld 24 klanten (onafhankelijk van elkaar) in een winkel komen, komen er natuurlijk gemiddeld 12 klanten per half uur.
Daar zal niemand aan twijfelen. Dat kun je ook formeel bewijzen. Als volgt.
Het aantal klanten X dat in een uur komt is Poisson-verdeeld met parameter $\lambda = 24$.
Het aantal klanten Y in een halfuur is ook Poisson-verdeeld. We gaan bewijzen dat de parameter van Y 12 is.
Splits een uur op in twee halve uren. Het aantal klanten in het eerste halfuur noemen we Y_1 en in het tweede halfuur Y_2.
Dan $X = Y_1 + Y_2$; dus $E(X) = E(Y_1) + E(Y_2)$.
Omdat $E(Y_1) = E(Y_2)$, volgt hieruit dat $E(Y) = ½ \cdot 24$ en dat is de parameter.

Als het aantal klanten in 1 uur Poissonverdeeld is met gemiddelde λ, dan is het aantal klanten in ½ uur Poissonverdeeld met gemiddelde ½ λ.

Algemeen Het aantal klanten in *t* uur Poissonverdeeld met gemiddelde $t \cdot \lambda$.

18 In een jaar komen er gemiddeld 123 brandmeldingen voor in een zekere stad.
a. Bereken de kans dat er morgen precies twee brandmeldingen zijn.

Er vallen in Nederland jaarlijks gemiddeld 810 fietsdoden in het verkeer.
b. Bereken de kans dat er morgen drie fietsdoden vallen.



Op de GR (TI84)
2ND DISTR

- poissonpdf (λ,*k*) geeft de kans op *k* voorvallen bij een Poissonverdeling met parameter λ.
- poissoncdf (λ,k) geeft de kans op hoogstens *k* voorvallen bij een Poissonverdeling met parameter λ.

19 **a.** Controleer je antwoorden op opgave **17** met de GR.

Gemiddeld worden er elke dag in Nederland 496 baby's geboren.
b. Wat is de kans dat er morgen niet meer dan 450 baby's geboren worden in Nederland.

Er vallen in Nederland jaarlijks gemiddeld 810 fietsdoden in het verkeer.
c. Bereken de kans dat er in een jaar minder dan 800 fietsdoden vallen in het verkeer.

20 Gegeven zijn twee Poissonverdeelde stochasten *X* en *Y* die onafhankelijk zijn van elkaar, met parameter λ en μ. Dan is *X*+*Y* ook Poissonverdeeld en wel met parameter λ+μ. We gaan dit bewijzen.
a. Ga na dat $P(X+Y=10) =$
$P(X=0)\cdot P(Y=10) + P(X=1)\cdot P(Y=9) + P(X=2)\cdot P(Y=8) + \ldots + P(X=10)\cdot P(Y=0)$.
b. Vul hierin de Poissonkansen in en laat zien dat de som gelijk is aan $\frac{(\lambda+\mu)^{10}}{10!} \cdot e^{-(\lambda+\mu)}$.
Tip: Binomium van Newton; zie appendix B.

c. Ga na dat hiermee de belofte van deze opgave is ingelost.

Intuïtief is het resultaat van de vorige opgave duidelijk: als de ene winkel gemiddeld λ klanten per uur krijgt en de andere winkel μ klanten, dan krijgen ze samen λ+μ klanten per uur.

21 Twee winkels zijn elkaars concurrenten. De aantallen klanten *X* en *Y* die deze per uur krijgen zijn Poisson-verdeeld met gemiddelden respectievelijk 1 en 2.

Een klant gaat naar een van de twee winkels.
a. Wat is, denk je, de kans dat hij naar de eerste winkel gaat?

Waarschijnlijk heb je bij onderdeel **a** intuïtief de juiste kans gegeven. We gaan die kans berekenen.Bedenk dat:
P(de winkels krijgen samen 1 klant) · P(die ene klant gaat naar de 1ste winkel) =
P(de 1ste winkel krijgt 1 klant en de 2de winkel krijgt 0 klanten) =
P(de 1ste winkel krijgt 1 klant) · P(de 2de winkel krijgt 0 klanten).

b. Schrijf de drie Poissonkansen op:
P(de winkels krijgen samen 1 klant in een uur),
P(de 1ste winkel krijgt 1 klant in een uur) en
P(de 2de winkel krijgt 0 klanten in een uur).
c. Bereken hieruit P(de ene klant gaat naar de 1ste winkel).
Had je in onderdeel **a** dit antwoord?

Stel je weet dat er in totaal 7 klanten naar de winkels gaan, maar je weet van geen van de klanten in welk van de twee winkels ze binnen zullen gaan.
d. Bereken de kans dat er 3 naar de eerste winkel gaan (en dus 4 naar de tweede).

Algemeen
Twee winkels zijn elkaars concurrenten. De aantallen klanten *X* en *Y* die deze per uur krijgen zijn Poisson-verdeeld met gemiddelden respectievelijk λ en μ. Stel je weet dat er in totaal *n* klanten naar de winkels gaan, maar je weet van geen van de klanten naar welke van de twee winkels ze gaan.
Dan is het aantal klanten dat naar de eerste winkel gaat binomiaal verdeeld met *n* herhalingen en succeskans $\frac{\lambda}{\lambda+\mu}$.

Speciaal bij zeldzame gebeurtenissen speelt de Poissonverdeling een belangrijke rol.

22 Een zekere ziekte is zeer zeldzaam: één op de honderdduizend mensen heeft haar. De kans dat iemand de ziekte heeft is dus 0,00001.
De ziekte is niet overdraagbaar en ook niet erfelijk bepaald. We mogen het hebben-van-de-ziekte voor een individu dus onafhankelijk beschouwen van het al dan niet hebben-van-de-ziekte van andere personen.

Er zijn 800 duizend Amsterdammers. Het aantal Amsterdammers dat de ziekte heeft noemen we X.
a. Wat is de verwachtingswaarde van X?

Elke Amsterdammer heeft kans 0,00001 om de ziekte te hebben, onafhankelijk van zijn stadsgenoten. Dus X is binomiaal verdeeld met parameters n = 800000 en p = 0,00001.
b. Bereken P(X=6).

De stochast Y is Poissonverdeeld met parameter $\lambda = 8$.
c. Bereken P(Y=6).

Als het goed is, heb je in b en c (nagenoeg) hetzelfde antwoord gekregen. Hoe kun je dat begrijpen?
We gaan daartoe P(X=6) uitrekenen zonder GR:
$P(X=6) = \binom{800000}{6} \cdot 0{,}00001^6 \cdot 0{,}99999^{799994}$.
$$\binom{800000}{6} = \frac{800000 \cdot 799999 \cdot 799998 \cdot 799997 \cdot 799996 \cdot 799995}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} \approx \frac{800000^6}{6!}$$
en dus $\binom{800000}{6} \cdot 0{,}00001^6 \approx \frac{8^6}{6!}$
d. Ga dat na.
e. Ga na dat $0{,}99999^{799994} \approx 0{,}99999^{800000}$, omdat de twee getallen een factor $0{,}99999^6$ verschillen en die is nagenoeg 1.
f. Ga na dat $0{,}99999^{800000} = (1 - \frac{8}{800000})^{800000}$.
De laatste uitdrukking is nagenoeg gelijk aan e^{-8}. Er geldt namelijk:
$(1 - \frac{x}{n})^n \approx e^{-x}$ als n groot is en x relatief klein. Ook geldt dan: $(1 + \frac{x}{n})^n \approx e^x$. Het bewijs van deze stelling staat in appendix C.

g. Laat zien dat uit de onderdelen **d**, **e** en **f** volgt dat P(X=6) ≈ P(Y=6).

X is binomiaal verdeeld met parameters n en p, waarbij p klein en n groot is,
Y is Poissonverdeeld met parameter $\lambda = p \cdot n$.
Dan zijn $P(X=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$ en $P(Y=k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$ nagenoeg gelijk.
Dus hebben X en Y nagenoeg dezelfde kansverdeling.

Op blz 41 wordt deen de verdelingen vergeleken in het geval $\lambda = 2$, $p = 0{,}02$ en $n = 100$.

Opmerking
Rekenmachines hebben een beperkt rekendomein. Op mijn machine moet de parameter n bij de binomiale verdeling kleiner dan 1 miljoen zijn. Als n groter dan 1 miljoen is (en p klein) brengt de Poissonverdeling uitkomst.

23 De zeldzame ziekte van de vorige opgave komt voor bij 1 op de 100000 mensen. Nederland telt 17 miljoen mensen.
Bereken de kans dat er in Nederland ten hoogste 185 mensen met de ziekte zijn.


Simeon-Denis Poisson
(1781-1840)

In 1837 introduceerde de Franse wiskundige Poisson een benadering van de binomiale kansverdeling. Hij bekeek de al geruime tijd bekende binomiale verdeling voor gevallen waarbij het aantal herhalingen n heel groot is en de succeskans p heel klein. Hij liet zien dat de binomiale kans op k successen $P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$ nagenoeg gelijk is aan $e^{-\lambda} \cdot \lambda^k / k!$, waarin $\lambda = n \cdot p$ de verwachtingswaarde van X is.
Dat is precies wat we in opgave **22** hebben laten zien.

Voor deze verdeling hoef je de kans op succes per kansexperiment niet te weten. Als verwachtingswaarde λ gebruik je het gemiddelde aantal "successen" bij een aantal series van n kansexperimenten.
Poisson besteedde niet meer dan één bladzijde aan zijn ontdekking.

Ludwig von Bortkiewicz

De Duitse wiskunde L. von Bortkiewicz was de eerste die het belang van de Poissonverdeling onderkende. Van hem is het volgende voorbeeld afkomstig.
Hij bekeek het aantal doden per jaar onder de Pruisische cavaleristen door een trap van een paard. Er zijn inderdaad veel herhalingen met een zeer kleine "succes"-kans (een dodelijke trap van een paard).

Het veelvuldig optreden van Poissonkansen in het dagelijks leven is niet zo vreemd als je dit voorbeeld bekijkt. In veel praktische situaties is er sprake van een zeer groot aantal uitvoeringen en een zeer kleine (onbekende) succeskans.

24 von Bortkiewicz telde in 14 regimenten over 20 jaren (1875 t/m 1894) in totaal 196 doden door een trap van een paard. Hij concludeerde dat per regiment per jaar het aantal door een trap van een paard gedode cavaleristen Poissonverdeeld was met verwachtinggswaarde 0,7.
Bereken de kansen dat in een regiment in een jaar 0, 1 en 2 doden vielen door een trap van een paard.

25 De lotto is een kansspel. De speler kruist zes nummers aan uit 1 t/m 45 en kiest een van de zes mogelijke kleuren. De notaris trekt ook zes nummers en een kleur; als die nummers precies hetzelfde zijn aan de zes die de speler koos, en bovendien de kleur hetzelfde is, wint hij de jackpot.
a. Ga na dat de kans daarop $2{,}05 \cdot 10^{-8}$ is.

Stel dat er in een week 1 miljoen lottoformulieren worden ingevuld. De jackpot valt als iemand de juiste zes nummers én de juiste kleur heeft gekozen.

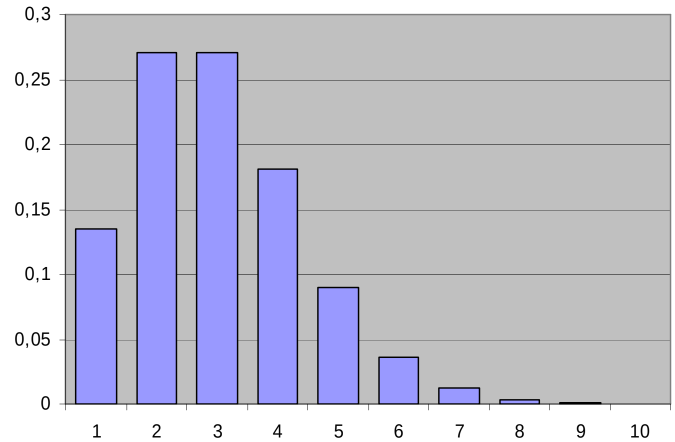
b. Bereken de kans dat de jackpot in een zekere week niet valt.

Als meer dan een speler alles goed heeft, moeten de gelukkigen de jackpot delen.
c. Bereken de kans dat de jackpot moet worden gedeeld.

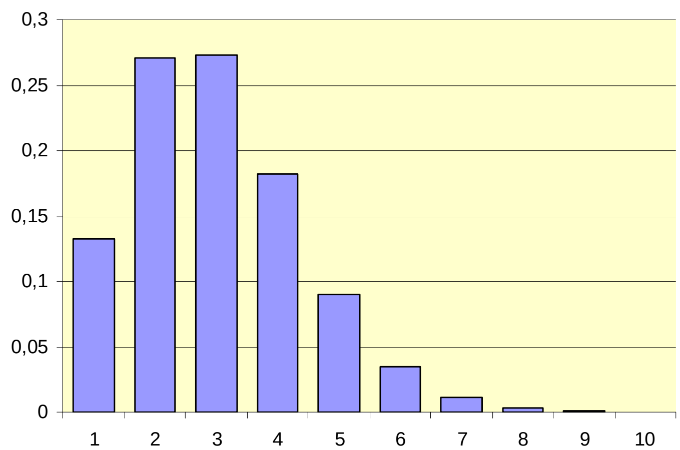
Vergelijk de kansverdelingen met dezelfde verwachtingswaarde:

- Poisson met λ = 2
- Binomiaal met n = 100 en p = 0,02

Poisson λ = 2



Binomiaal *p* =0,02, *n* =100



3 Wanneer komt de volgende klant?

Vraag
Het aantal klanten is Poissonverdeeld met gemiddelde 3 per uur.
Wat is de kans dat de eerste klant ten minste 0,5 uur op zich laat wachten?

Berekening
Het aantal klanten per half uur is Poissonverdeeld met gemiddelde ½ · 3 = 1,5.
De eerste aankomsttijd (in uren) noemen we *T*; dat is de tijd die het duurt voordat de eerste klant binnenkomt.
$P(T \geq 0{,}5)$ = P(0 klanten in de eerste 0,5 uur) =
$\frac{1{,}5^0}{0!} \cdot e^{-1{,}5} \approx 0{,}2231$.

26 We gaan verder met de context van bovenstaande vraag.
a. Bereken de kans dat de eerste klant meer dan ½ uur, maar minder dan 1½ uur op zich laat wachten.
b. Bereken de kans dat de derde en vierde klant in het derde kwartier komen.
Tip: Dan moeten in het eerste halfuur 2 klanten komen en in het daarop volgende kwartier ook 2.

27 Het aantal klanten *X* is Poissonverdeeld met gemiddelde λ per uur.
T is de tijd die het duurt voordat de eerste klant komt.
a. Welke waarden kan *T* aannemen?
b. Is *X* discreet of continu verdeeld? En *T*?
c. Bewijs dat $P(T \leq t) = 1 - e^{-\lambda t}$, voor alle $t > 0$.
d. Controleer de formule in onderdeel c voor de "randgevallen" $t = 0$ en t nadert tot oneindig.

Als een aantal "successen" *X* Poissonverdeeld is met gemiddelde λ en *T* is de tijdsduur dat je op het eerste succes moet wachten,
dan $P(T \leq t) = 1 - e^{-\lambda t}$.

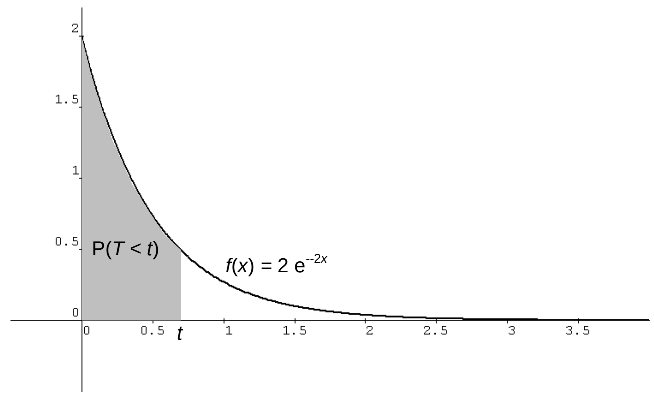
Definitie
Een stochast *T* heet **exponentiëel verdeeld** met parameter λ als *T* alle positieve getallen als waarde kan aannemen en $P(T \leq t) = 1 - e^{-\lambda t}$ voor alle $t > 0$.

28 **a.** Teken de grafiek van de functie $F(t) = 1 - e^{-\lambda t}$ als functie van t, voor enkele waarden van λ in één figuur.
F heet de **verdelingsfunctie** van T.

b. Teken de grafiek van de afgeleide $f = F'$ voor dezelfde waarden van λ, ook in één figuur.
f heet de **dichtheidsfunctie** van T.

c. Leg uit: $P(T \le t) = \int_0^t \lambda e^{-\lambda x}\, dx$.

Stel dat een winkelier gemiddeld λ klanten per uur krijgt. De kans dat hij hoogstens t uur hoeft te wachten voordat de eerste klant komt, is dus de oppervlakte onder de grafiek van de functie $f(x) = \lambda e^{-\lambda x}$, links van t.



Iets dergelijks heb je al eerder ontmoet: ook bij een normaalverdeelde stochast X zijn de kansen $P(X < t)$ de oppervlakte onder een grafiek, namelijk van de klok-kromme.

29 Study probeert elke vrijdagavond een lift te krijgen om het weekend bij haar ouders door te brengen. De helft van de keren duurt het minder dan 30 minuten voordat ze een lift krijgt.
a. Wat is de kans dat ze op een vrijdag binnen 5 minuten een lift krijgt?
b. En wat is de kans dat ze na 1 uur nog geen lift heeft.

30 Op een vrijdagavond heeft Study zonder succes al 40 minuten staan liften. Ze zegt bij zichzelf: "Ik sta hier nu al veertig minuten, terwijl ik gemiddeld niet meer dan een half uur op een lift hoef te wachten. Ik zal nu dus wel snel een lift krijgen."
Geef commentaar op Study's gedachte.

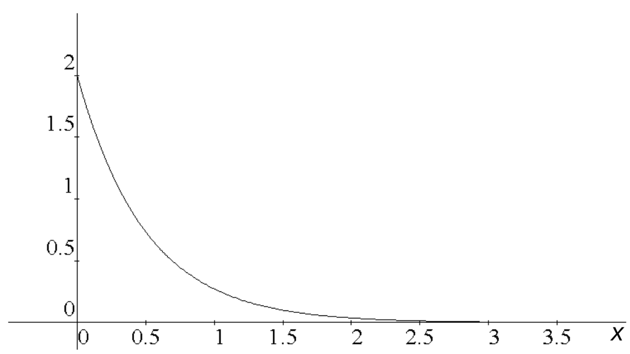
De exponentiële verdeling heeft geen geheugen. Dat betekent het volgende.
Als je al bijvoorbeeld 10 minuten (zonder succes) hebt staan liften, wordt de kans dat je een lift krijgt daar niet groter (of kleiner) door.

31 Choice speelt *Mens erger je niet.* Ze heeft op het ogenblik geen enkele pion op het speelbord. Pas als ze "een zes" (dat is 6 ogen) gegooid heeft, mag ze een pion op het bord zetten. De eerste vijf keer dat ze aan de beurt is, werpt ze 2, 2, 4, 1 en 5 ogen.
a. Wat is de kans dat ze in de volgende beurt 6 ogen werpt?
b. Heeft de dobbelsteen een geheugen?

Definitie
We zeggen dat een stochast X *geen geheugen* heeft, als
$P(X > a+b) = P(X > a) \cdot P(X > b)$ voor alle getallen a en b.
Ga na dat dit een verstandige definitie is.

32 Een exponentieel verdeelde stochast heeft geen geheugen.
Bewijs dat.
Tip: Pas de definitie op bladzijde 42 toe.

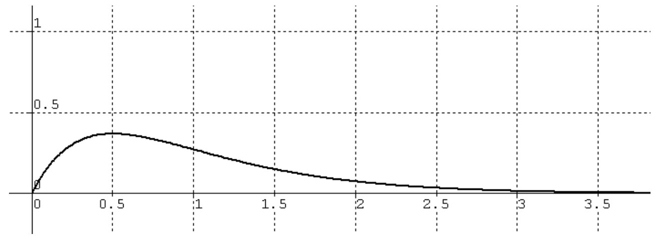
33 Hieronder staat de grafiek van de dichtheidsfunctie $f(x) = 2\,e^{-2x}$ van een exponentieel verdeelde stochast T met parameter 2.



Om de verwachtingswaarde van T te berekenen, moet je $x \cdot f(x)$ intereren van 0 tot oneindig: $\int_0^\infty x \cdot 2e^{-2x}\,dx$.

In appendix D wordt toegelicht dat je zo de verwachtingswaarde van de continue stochast T met parameter 2 vindt.

Hieronder staat de grafiek van $x \cdot 2\, e^{-2x}$.



a. Maak een schatting van E(T

4 Appendix

A De e-macht als oneindige som

We gaan bewijzen dat de oneindige som
$1 + x + x^2 / 2 + x^3 / 6 + x^4 / 24 + x^5 / 120 + \ldots$
gelijk is aan e^x, voor elk getal x.

35 Bekijk de functie
$y_5(x) = 1 + x + x^2 / 2 + x^3 / 6 + x^4 / 24 + x^5 / 120$.
a. Bereken de afgeleide $y_5'(x)$.

$y_5'(x)$ lijkt veel op $y_5(x)$. Deze twee functies verschillen slechts één term, namelijk $x^5 / 120$. En als $|x| \leq 1$ is dat verschil kleiner dan 0,01.
b. Ga dat na.

Bekijk de functie
$y_{50}(x) = 1 + x + x^2 / 2 + x^3 / 6 + x^4 / 24 + \ldots + x^{50} / 50!$
c. Bereken de afgeleide $y_{50}'(x)$.

$y_{50}'(x)$ lijkt veel op $y_{50}(x)$. Deze twee functies verschillen slechts één term, namelijk $x^{50} / 50!$. En als $|x| \leq 17$ is dat verschil kleiner dan 0,0011.
d. Ga dat na.

Bekijk de eindige som
$1 + x + x^2 / 2 + x^3 / 6 + x^4 / 24 + \ldots + x^n / n!$
Dat is een functie y_n van x.
e. Wat is het verschil tussen $y_n'(x)$ en $y_n(x)$?

Voor elk getal x geldt:
De noemer n! groeit op den duur (veel) sneller dan de teller x^n.
Daarom wordt het verschil tussen $y_n'(x)$ en $y_n(x)$ voor grote waarden van n willekeurig klein.
Voor de oneindige som
$y(x) = 1 + x + x^2 / 2 + x^3 / 6 + x^4 / 24 + \ldots$ geldt dus:
$y'(x) = y(x)$.
De functie y is zijn eigen afgeleide!
En die functies kennen we: $y(x) = c \cdot e^x$, waarbij c een willekeurige constante is.

We moeten nog de waarde van c bepalen. Dat lukt omdat we $y(0)$ kunnen uitrekenen,
f. Wat is de waarde van c?

Ga na dat we hiermee onze belofte hebben ingelost:
$1 + x + x^2 / 2 + x^3 / 6 + x^4 / 24 + x^5 / 120 + \ldots = e^x$.

B Het binomium van Newton

36 **a.** Schrijf zonder haakjes en met zo weinig mogelijk termen: $(x+y)^2$, $(x+y)^3$ en $(x+y)^4$.

Hoe gaat dit verder? Zit er enig systeem in? Wat zal $(x+y)^{10}$ opleveren? We gaan dat op verschillende manieren uitvinden.

Manier1: haakjes uitwerken
Het is duidelijk dat $(1+x)^7$een veelterm is van graad 7 dus een som van de vorm:
$\ldots + \ldots\cdot x + \ldots\cdot x^2 + \ldots\cdot x^3 + \ldots\cdot x^4 + \ldots\cdot x^5 + \ldots\cdot x^6 + \ldots\cdot x^7$.
Om hier beter over te kunnen praten geven we de coëfficiënten (de getallen op de puntjes) namen
$a_0 + a_1\cdot x + a_2\cdot x^2 + a_3\cdot x^3 + a_4\cdot x^4 + a_5\cdot x^5 + a_6\cdot x^6 + a_7\cdot x^7$.

Je weet waarschijnlijk bij voorbaat wel wat a_0 en a_7 zijn?
b. Wat zijn die?

Stel dat je de uitwerking van $(1+x)^7$ kent. Dan volgt daaruit de uitwerking van $(1+x)^8$ als volgt.
$(1+x)^8 = (1+x)\cdot(1+x)^7 = 1\cdot(1+x)^7 + x\cdot(1+x)^7 =$
$a_0 + a_1\cdot x + a_2\cdot x^2 + a_3\cdot x^3 + a_4\cdot x^4 + a_5\cdot x^5 + a_6\cdot x^6 + a_7\cdot x^7 +$
$x\cdot(a_0 + a_1\cdot x + a_2\cdot x^2 + a_3\cdot x^3 + a_4\cdot x^4 + a_5\cdot x^5 + a_6\cdot x^6 + a_7\cdot x^7) =$
$a_0 + (a_0+a_1)\cdot x + (a_1+a_2)\cdot x^2 + (a_2+a_3)\cdot x^3 + (a_3+a_4)\cdot x^4 +$
$(a_4+a_5)\cdot x^5 + (a_5+a_6)\cdot x^6 + (a_6+a_7)\cdot x^7 + a_8\cdot x^8$.
Als je dus de coëfficiënten van de uitwerking van $(1+x)^7$ kent, ken je die ook van de uitwerking van $(1+x)^8$: je moet gewoon steeds twee opvolgende coëfficiënten van de uitwerking van $(1+x)^7$ optellen.
Om aan de uitwerking van $(1+x)^7$ te komen, moet je het volgende schema invullen.

1

$1 + 1\,x$

$1 + \ldots\,x + 1\,x^2$

$1 + \ldots\,x + \ldots x^2 + 1\,x^3$

$1 + \ldots\,x + \ldots x^2 + \ldots x^3 + 1\,x^4$

$1 + \ldots\,x + \ldots x^2 + \ldots x^3 + \ldots x^4 + 1\,x^5$

$1 + \ldots\,x + \ldots x^2 + \ldots x^3 + \ldots x^4 + \ldots x^5 + 1\,x^6$

$1 + \ldots\,x + \ldots x^2 + \ldots x^3 + \ldots x^4 + \ldots x^5 + \ldots x^6 + 1\,x^7$

c. Vul de coëfficiënten van bovenaf in.

Dit schema is een oude bekende. Met precies hetzelfde schema maak je namelijk de driehoek van Pascal.

$$\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 2 \quad 1 \\
1 \quad 3 \quad 3 \quad 1 \\
1 \quad 4 \quad 6 \quad 4 \quad 1 \\
1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1 \\
1 \quad 7 \quad 21 \quad 35 \quad 35 \quad 21 \quad 7 \quad 1 \\
1 \quad 8 \quad 28 \quad 56 \quad 70 \quad 56 \quad 28 \quad 8 \quad 1 \\
1 \quad 9 \quad 36 \quad 84 \quad 126 \quad 126 \quad 84 \quad 36 \quad 9 \quad 1 \\
1 \quad 10 \quad 45 \quad 120 \quad 210 \quad 252 \quad 210 \quad 120 \quad 45 \quad 10 \quad 1
\end{array}$$

Kennelijk is de coëfficiënt van bijvoorbeeld x^5 in de uitwerking van $(1+x)^7$ gelijk aan $\binom{7}{5} = 21$.

d. Wat is de coëfficiënt van x^4 in de uitwerking van $(1+x)^{10}$?

Conclusie

$$(1+x)^{10} = \binom{10}{0}x^0 + \binom{10}{1}x^1 + \binom{10}{2}x^2 + \binom{10}{3}x^3 + \binom{10}{4}x^4 + \binom{10}{5}x^5 + \binom{10}{6}x^6 + \binom{10}{7}x^7 + \binom{10}{8}x^8 + \binom{10}{9}x^9 + \binom{10}{10}x^{10}$$

afgekort:

$$(1+x)^{10} = \sum_{k=0}^{10}\binom{10}{k}x^k$$

e. Controleer of de eerste en de laatste term van deze som juist zijn.

Manier2: met differentiëren

37 Bekijk $(1+x)^{10} = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3 + a_4 \cdot x^4 + a_5 \cdot x^5 + a_6 \cdot x^6 + a_7 \cdot x^7 + a_8 \cdot x^8 + a_9 \cdot x^9 + a_{10} \cdot x^{10}$.
We gaan linker- en rechterlid vier keer differentiëren.
De vierde afgeleide van het linkerlid is $10 \cdot 9 \cdot 8 \cdot 7 \cdot (1+x)^6$

a. Schrijf de vierde afgeleide van het rechterlid op.

Nu vullen we voor x in beide vierde afgeleides 0 in en vereenvoudigen de uitkomsten.

b. Wat zijn die uitkomsten?

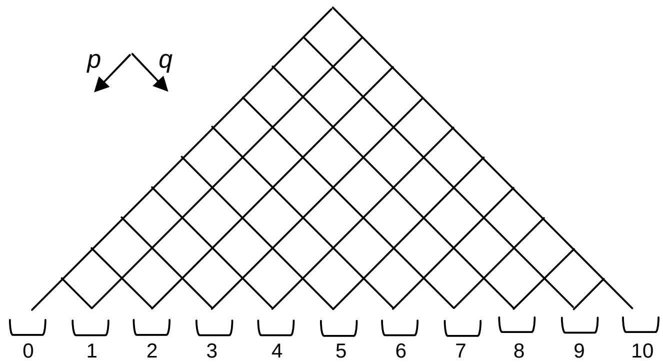
c. Ga na dat uit het voorgaande volgt dat $a_4 = \frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1}$

d. Ga na dat $\frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = \binom{10}{4}$

We hebben hetzelfde resultaat gevonden als in opgave 37.

Manier3: met kansrekening

38 We bekijken het asymmetrische Galtonbord het 10 rijen. Voor een kogeltje is de kans om naar links te vallen p en de kans om naar rechts te vallen $q = 1-p$.
We nummeren de bakjes van links naar rechts: 0, 1, 2, …, 10.



a. Wat is de kans dat een kogeltje in bakje met nummer k terecht komt?

b. Leg uit: $1 = \sum_{k=0}^{10} \binom{10}{k} p^k q^{10-k}$

Als je beide leden van de formule in onderdeel b vermenigvuldigt met a^{10}, krijg je $a^{10} = \sum_{k=0}^{10} \binom{10}{k} (ap)^k (aq)^{10-k}$.

c. Leg dat uit.

Vervang ap door x en aq door y. Dan krijg je :
$(x+y)^{10} = \sum_{k=0}^{10} \binom{10}{k} x^k y^{10-k}$.

d. Leg dat uit.

Voor alle getallen x en y en alle gehele exponenten n geldt:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} .$$

Deze formule staat bekend als het **binomium van Newton.**

$x+y$ is een tweeterm (binomus); vandaar de naam.

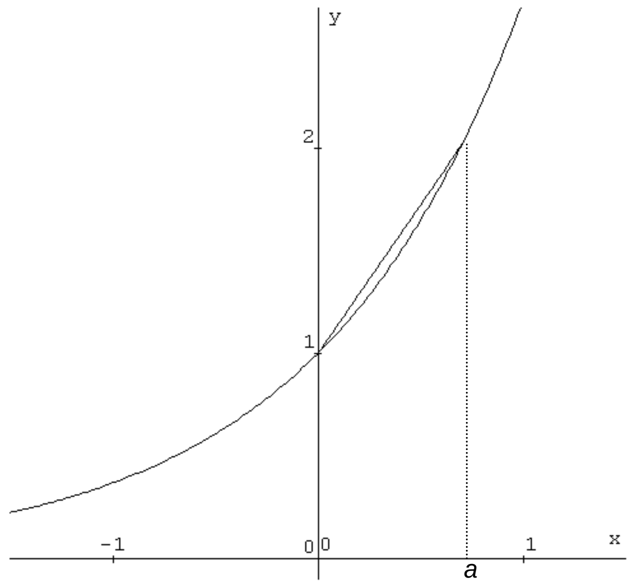
39 **a.** Ga na dat het binomium van Newton en het resultaat van opgave 38 met elkaar in overeenstemming zijn.
b. Ga na dat de uitkomsten van opgave 38a speciale gevallen zijn van het binomium van Newton.

40 **a.** Bewijs uit het binomium van Newton dat $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

b. Welke uitkomst levert het binomium van Newton voor $\sum_{k=0}^{n} (-1)^k \cdot \binom{n}{k}$?

C De e-macht als limiet

41 We gaan het getal e benaderen.
Ons uitgangspunt is: $\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^x = \mathrm{e}^x$.
Bijgevolg is $\mathrm{e}^a \approx 1 + a$. En dit klopt beter naarmate a dichter bij 0 ligt.
a. Leg dat uit, bijvoorbeeld aan de hand van de figuur hieronder.



Kies $a = \frac{x}{n}$. Dan krijg je $\mathrm{e}^{\frac{x}{n}} \approx 1 + \frac{x}{n}$. Dit klopt des te beter als n groot is en x niet al te groot.
Hieruit volgt dat $\mathrm{e}^x \approx (1 + \frac{x}{n})^n$.
b. Leg dat uit.

Door in deze formule voor x de waarde 1 te kiezen en voor n een groot getal, kun je e benaderen.
c. Doe dat.

Voor alle getallen x geldt:
$$\mathrm{e}^x \approx (1 + \tfrac{x}{n})^n$$
en dit klopt beter naarmate n groter is.

D De verwachtingswaarde van *T*

Uit opgave 28c volgt dat P($t < T < t+\Delta t$) = $\int_t^{t+\Delta t} 2 \cdot e^{-2x}\,dx$.

Hieruit volgt:
P($t < T < t+\Delta t$) ≈ $2 \cdot e^{-2t} \cdot \Delta t$ voor kleine waarden van Δt
.

De verwachtingswaarde van een discrete stochast X (die de waarden 0, 1, 2, 3, … aanneemt bereken je als volgt:
vermenigvuldig de waarden die hij kan aannemen met de bijbehorende kansen en tel de producten op:
0 × P(X=0) + 1 × P(X=1) + 2 × P(X=2) +3 × P(X=3) + …

Dat doen we nu ook bij de continue stochast T:
tel de producten t × P($t < T < t+\Delta t$) op.
Laat vervolgens Δt tot 0 naderen en je krijgt als verwachtings waarde:

$$E(T) = \int_0^{\infty} t \cdot 2 \cdot e^{-2x}\,dt .$$

5 Antwoorden

Paragraaf 1 Wachten

1 **a.** $\frac{11}{12}$

b. $\frac{11}{12}$

c. $(\frac{11}{12})^{10} \approx 0{,}4189$

2 **a.** $(\frac{23}{24})^{20} \approx 0{,}4269$

b. Die verschillen (een beetje).

3 **a.** binompdf(10, $\frac{1}{12}$, 1) = 0,3808 of $10 \cdot \frac{1}{12} \cdot (\frac{11}{12})^9$

b. binompdf(10, $\frac{1}{12}$, 2) = 0,1558 of $45 \cdot (\frac{1}{12})^2 \cdot (\frac{11}{12})^8$

4 **a.** 0, 1, 2, …, 10

b. $(\frac{1}{4})^3 \cdot (\frac{3}{4})^7 \approx 0{,}0021$

c. $\binom{10}{3} = \frac{10!}{3! \cdot 7!} = 120$

d. $\binom{10}{3} \cdot (\frac{1}{4})^3 \cdot (\frac{3}{4})^7 \approx 120 \cdot 0{,}0021 \approx 0{,}250$

5 **a.** Het aantal klasgenoten van Jan dat in dezelfde week jarig is, is binomiaal verdeeld met $p = \frac{1}{52}$ en $n = 24$.

De gevraagde kans is $\binom{24}{2} \cdot (\frac{1}{52})^2 \cdot (\frac{51}{52})^{22} \approx 0{,}0666$

b. binompdf(24,1/52,2)

6 **a.** Het aantal reizigers dat niet komt opdagen, is binomiaada en.10797(,)-4.079878 Tm [( )-[3

c. $\binom{n}{0} \cdot a^0 \cdot (1-a)^n + \binom{n}{1} \cdot a^1 \cdot (1-a)^{n-1} + \ \ldots + \binom{n}{n} \cdot a^n \cdot (1-a)^0$ is dan gelijk aan 1.
Vat a op als de kans op succes. Dan zegt de formule dat de som van de kansen op 0, 1, 2, …, n successen 1 is.

Paragraaf 2 De Poissonverdeling

10 a. Daar moet hij het personeel op aanpassen. Klanten die moeten wachten verliest hij anders.
b. ??

11 a. Ze hebben niet afgesproken samen naar de winkel te gaan. Als een klant binnenkomt is dat van geen invloed op de aankomsttijd van een andere klant.

b. $\binom{10}{8} \cdot (\frac{1}{2})^8 \cdot (\frac{1}{2})^2 = 0{,}04395$ of binompdf(10,0,5,8)

c. Heel klein. Er is niet zo veel speling: Als een volgende klant precies komt als de vorige net weg is, is er maar 10 minuten in het uur over.

12 a. 0 klanten in het hele uur betekent dat er 0 klanten komen in het eerste halfuur en daarna ook 0 klanten in het tweede halfuur. De kansen op de twee laatstgenoemde gebeurtenissen zijn gelijk; die moet je vermenigvuldigen.
b. $P_1(1) = P_{½}(0) \cdot P_{½}(1) + P_{½}(1) \cdot P_{½}(0) = 2 \cdot P_{½}(0) \cdot P_{½}(1)$
c. Stel je weet dat er in totaal 7 klanten naar de winkels gaan, maar je weet van geen van de klanten in welk van de twee winkels ze binnen zullen gaan.
Stel je weet dat er in totaal 7 klanten naar de winkels gaan, maar je weet van geen van de klanten in welk van de twee winkels ze binnen zullen gaan.
$P_1(3) = P_{½}(0) \cdot P_{½}(3) + P_{½}(1) \cdot P_{½}(2) + P_{½}(2) \cdot P_{½}(1) + P_{½}(3) \cdot P_{½}(0) = 2 \cdot P_{½}(0) \cdot P_{½}(3) + 2 \cdot P_{½}(1) \cdot P_{½}(2)$

13 a. $P_{½}(6) \cdot P_{½}(1) =$
P(6 klanten in het eerste halfuur en 1 in het tweede halfuur) =
P(7 klanten in het hele uur en daarvan 1 in het tweede halfuur) =
$P_1(7) \cdot P_1$(1 van de 7 klanten komt in het eerste halfuur) =
$P_1(7) \cdot 7 \cdot (½)^7$
b. $P_{½}(6) \cdot P_{½}(1) = 7 \cdot P_{½}(7) \cdot P_{½}(0)$
Deel beide leden door $7 \cdot P_{½}(0)$ geeft het gewenste resultaat.

c. $P_{½}(n) = \frac{1}{n} \cdot \frac{P_{1/2}(1)}{P_{1/2}(0)} \cdot P_{½}(n-1)$

d. Voor $n = 1$ zegt de formule: $P_{½}(1) = \frac{P_{1/2}(1)}{P_{1/2}(0)} \cdot P_{½}(0)$.

14 a. $P_{½}(1) = \lambda \cdot P_{½}(0)$
$P_{½}(2) = \frac{1}{2} \cdot \lambda^2 \cdot P_{½}(0)$
$P_{½}(3) = \frac{1}{3} \cdot \frac{1}{2} \cdot \lambda^3 \cdot P_{½}(0)$
$P_{½}(4) = \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \lambda^4 \cdot P_{½}(0)$
$P_{½}(5) = \frac{1}{5} \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \lambda^5 \cdot P_{½}(0)$
b. $P_{½}(n) = \frac{1}{n!} \cdot \lambda^n \cdot P_{½}(0)$

15 a. $0{,}1 \cdot e^{\lambda} = 1$, dus $\lambda = \ln(10) \approx 2{,}3026$
b. $P_{1/2}(0) \cdot e^{\lambda} = 1$, dus $e^{\lambda} = 1/P_1(0)$, dus $\lambda = \ln(1/P_1(0))$.

16 a. $E(X) = \frac{\lambda^1}{1!} \cdot e^{-\lambda} + 2 \cdot \frac{\lambda^2}{2!} \cdot e^{-\lambda} + 3 \cdot \frac{\lambda^3}{3!} \cdot e^{-\lambda} + \ldots$
Haal $e^{-\lambda}$ buiten haakjes en haal één factor λ buiten haakjes.
b. Tussen de haakjes staat e^{λ}. Omdat $e^{-\lambda} \cdot e^{\lambda} = 1$, is $E(X) = \lambda$.

17 a. $P(X=0) = e^{-\lambda}$ en $P(Y=0) = e^{-\frac{1}{2}\lambda}$.
Inderdaad is $e^{-\lambda} = (e^{-\frac{1}{2}\lambda})^2$
b. $P(X=1) = \lambda \cdot e^{-\lambda}$, $P(Y=0) = e^{-\frac{1}{2}\lambda}$ en $P(Y=1) = \frac{1}{2} \lambda e^{-\frac{1}{2}\lambda}$
Inderdaad is $\lambda \cdot e^{-\lambda} = 2 \cdot e^{-\frac{1}{2}\lambda} \cdot \frac{1}{2} \lambda e^{-\frac{1}{2}\lambda}$

18 a. Er zijn gemiddeld 123 / 365 brandmeldingen per dag.
X is Poissonverdeeld met $\lambda = \frac{123}{365}$.
$P(X=2) = (\frac{123}{365})^2 / 2! \cdot e^{\frac{123}{365}} \approx 0{,}04054$
b. Er vallen gemiddeld 810/365 fietsdoden per dag.
X s Poissonverdeeld met $\lambda = \frac{810}{365}$.
$P(X=2) = (\frac{810}{365})^3 / 3! \cdot e^{\frac{810}{365}} \approx 0{,}1980$

19 a. poissonpdf(123/365,2) = 0,04054
poissonpdf(810/365,3) =0,19799
b. poissoncdf(496,450) = 0,0193
c. poissoncdf(810,799) = 0,3580

20 a. *X* en *Y* nemen positieve gehele waarden aan en nul.
Als $X+Y = 10$, is X=0 en Y=10 of X=1 en Y=9 of … X=10 en Y=0. Omdat *X* en *Y* onafhankelijk zijn, volgt het beweerde.
b. $e^{-\lambda} \cdot \frac{\mu^{10}}{10!} e^{-\mu} + \lambda e^{-\lambda} \cdot \frac{\mu^9}{9!} e^{-\mu} + \frac{\lambda^2}{2!} e^{-\lambda} \cdot \frac{\mu^8}{8!} e^{-\mu} + \ldots +$

$\frac{\lambda^{10}}{10!}\, e^{-\lambda}\cdot e^{-\mu} =$

$e^{-\lambda}\cdot e^{-\mu}\left(\frac{\mu^{10}}{10!} + \lambda\cdot\frac{\mu^{9}}{9!} + \frac{\lambda^{2}}{2!}\cdot\frac{\mu^{8}}{8!} + \ldots + \frac{\lambda^{10}}{10!}\right) =$

$e^{-(\lambda+\mu)}\cdot\frac{1}{10!}\cdot\left(\binom{10}{0}\lambda^{0}\mu^{10} + \binom{10}{1}\lambda^{1}\mu^{9} + \binom{10}{2}\lambda^{2}\mu^{8} + \ldots\binom{10}{10}\lambda^{10}\mu^{0}\right)$

$= e^{-(\lambda+\mu)}\cdot\frac{1}{10!}\cdot(\lambda+\mu)^{10}$;

de laatste stap volgens het binomium van Newton.

c. In de laatste uitdrukking herken je (hopelijk) de kans op waarde 10 van een Poissonverdeelde stochast met parameter $\lambda+\mu$.

21 a. Elke klant gaat met 2 keer zo grote kans naar de tweede winkel dan naar de eerste. De kans dat hij dus naar de eerste winkel gaat is

b. De drie kansen zijn:

$\frac{3^1}{1!}\cdot e^{-3} = 3e^{-3}$, $\frac{1^1}{1!}\cdot e^{-1} = e^{-1}$ en $\frac{2^0}{0!}\cdot e^{-2} = e^{-2}$.

c. $3\,e^{-3}\cdot$ P(die ene klant gaat naar de 1ste winkel) = $e^{-1}\cdot e^{-2}$.
Links en rechts vallen de e-machten tegen elkaar weg.
Dus P(die ene klant gaat naar de 1ste winkel) = 1/3.

d. $\binom{7}{3}\cdot\left(\frac{1}{3}\right)^3\cdot\left(\frac{2}{3}\right)^4 \approx 0,2561$

22 a. 8

b. 0,122138

c. 0,122138

d. $\binom{800000}{6} = \frac{800000\cdot 799999\cdot 799998\cdot 799997\cdot 799996\cdot 799995}{6!}$ en 799999, 799998, 799997, 799996, 799995 zijn nagenoeg gelijk aan 800000. Vervolgens is $800000^6\cdot 0{,}00001^6 = (800000\cdot 0{,}00001)^6 = 8^6$.

e. $0{,}99999^{799994}\cdot 0{,}99999^{6} = 0{,}99999^{800000}$ en $0{,}99999^{6} = 0{,}99994\ldots \approx 1$

f. $0{,}99999 = 1 - 0{,}00001 = 1 - \frac{8}{800000}$.

g. Het resultaat van d geeft:

$P(X=6) \approx \frac{8^6}{6!}\cdot\left(1 - \frac{8}{800000}\right)^{800000}$ en dit is precies de Poissonkans op uitkomst 6 bij parameter 8.

23 Het aantal dragers is bij benadering Poissonverdeeld met $\lambda = 170$.
poissoncdf(170,185) = 0,8818

24 X = aantal doden in een regiment in een jaar.
$P(X=0) = e^{-0,7} = 0{,}4966$
$P(X=1) = 0{,}7\cdot e^{-0,7} = 0{,}3476$
$P(X=2) = 0{,}7^2 / 2!\cdot e^{-0,7} = 0{,}1217$

25 a. De kans op zes goede nummers is

$\frac{6}{45} \cdot \frac{5}{44} \cdot \frac{4}{43} \cdot \frac{3}{42} \cdot \frac{2}{41} \cdot \frac{1}{40} = 1{,}2277 \cdot 10^{-7}$. Of $1 / \binom{45}{6}$.

De kans op de goede kleur is $\frac{1}{6}$

Vermenigvuldigd geeft dat $2{,}0462 \cdot 10^{-8}$.

b. Noem het aantal deelnemers met alles goed Y.
Y is Poissonverdeeld met $\lambda = 2{,}05 \cdot 10^{-2}$.
$P(Y{=}0) = e^{-0{,}0205} = 0{,}9797$

c. $P(Y \geq 2) = 1 - (P(Y{=}0) + P(Y{=}1)) \approx 0{,}000207$

Paragraaf 3 Wanneer komt de volgende klant

26 a. Het aantal klanten in de eerste 1½ uur is Poissonverdeeld met $\lambda = 1\tfrac{1}{2} \cdot 3 = 4\tfrac{1}{2}$.
$P(\tfrac{1}{2} < T < 1\tfrac{1}{2}) = P(T > \tfrac{1}{2}) - P(T \geq 1\tfrac{1}{2}) = e^{-1{,}5} - e^{-4{,}5} = 0{,}2231 - 0{,}0111 = 0{,}2120$

b. P(2 klanten in eerste halfuur) · P(2 klanten in derde kwartier) = $\frac{(1{,}5)^2}{2!} e^{-1{,}5} \cdot \frac{(0{,}75)^2}{2!} e^{-0{,}75} = 0{,}2510 \cdot 0{,}1329 = 0{,}0333$.

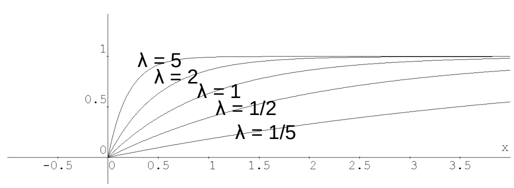
27 a. Alle getallen groter dan 0

b. X is discreet verdeeld, T is continu verdeeld

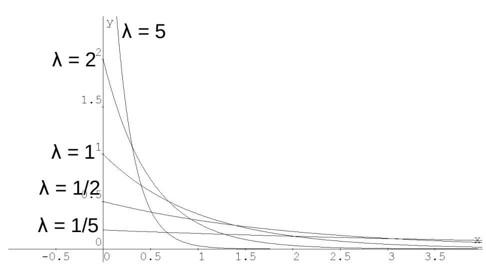
c. Het aantal klanten in de eerste t uur is Poissonverdeeld met parameter $t \cdot \lambda$.
$P(T \leq t) = 1 - P(T > t) = 1 - P(0$ klanten in de eerste t uur$) = 1 - \frac{(t \cdot \lambda)^0}{0!} e^{-t \cdot \lambda} = 1 - e^{-t \cdot \lambda}$.

d. Als $t = 0$, levert de formule $1 - e^0 = 0$ en dat moet ook: de kans dat de eerste klant onmiddellijk komt is 0.
Als t nadert tot oneindig, levert de formule $1 - 0 = 1$ en dat moet ook: de kans dat er ooit een klant zal komen is 1.

29 a.



b.



c. $f(t) = \lambda\, e^{-\lambda t}$.

$F(t) = \int_0^t \lambda e^{-\lambda x} dx + c$ zijn de primitieven van f.

$F(0) = 0$ en $\int_0^0 \lambda e^{-\lambda x} dx = 0$. Dus $c = 0$.

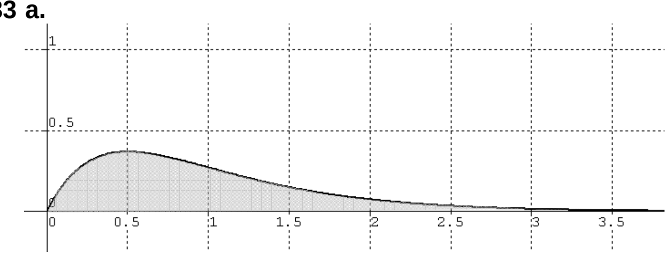
29 **a.** T is de wachttijd in minuten.
$P(T \leq 30)\ 1 - e^{-\lambda \cdot 30} = 0,5$. Hieruit volgt dat $\lambda = 0,0231$.
$P(T \leq 5) = 1 - e^{-0,0231 \cdot 5} = 0,1091$
b. $P(T > 60) = e^{-0,0231 \cdot 60} = 0,2501$

30 Studys gedachte is fout. De situatie na de eerste 40 minuten is precies dezelfde als toen ze begon te liften. De kans op een lift is dus niet beïnvloed door wat er vooraf gebeurd is.

31 **a.** Gewoon $\frac{1}{6}$. De voorgeschiedenis van de dobbelsteen heeft geen invloed op de kans op een 6.
b. Nee.

32 $P(X > a) = e^{-\lambda a}$, $P(X > b) = e^{-\lambda b}$ en $P(X > a+b) = e^{-\lambda(a+b)}$
Inderdaad geldt: $e^{-\lambda a} \cdot e^{-\lambda b} = e^{-\lambda(a+b)}$.

33 **a.**



Hokjes tellen: de oppervlakte is ongeveer 2 hokjes van oppervlakte 0,25. Dus is de oppervlakte ongeveer ½ .

b. $y' = -\,e^{-2x} + -x \cdot -2\,e^{-2x} - ½ \cdot -2\,e^{-2x} = 2x\,e^{-2x}$
c. Als we in de primitieve 0 invullen krijgen we – ½ . Als we een heel groot getal invullen, krijgen we 0.
Dus $E(T) = 0 - -½ = ½$.
d. $1/\lambda$

34 **a.** 40 minuten
b. $1/\lambda = 10$, dus $\lambda = 0,1$
Het aantal auto's X dat een lifter meeneemt in 30 min. is Poissonverdeeld met $\lambda = 3$.
$P(X \geq 4) = 1 - P(X \leq 3) = 1 – 0,6472 = 0,3528$

Paragraaf 4 Appendix

35 a. $y_5'(x) = 1 + x + x^2 / 2 + x^3 / 6 + x^4 / 24$
b. Als $|x| < 1$, dan $|x|^5 < 1$ en $|x^5 /120| < 1/120 < 0{,}01$
c. $\underline{y_{50}'(x) =} 1 + x + x^2 / 2 + x^3 / 6 + x^4 / 24 + \ldots + x^{49} / 49!$
d. Als $|x| < 17$, dan $|x|^{50} < 17^{50}$ en $|x^{50} /50!| \approx 0{,}001095 < 0{,}0111$
e. $x^n / n!$
f. $y(0) = 1 + 0 + 0 + \ldots = 1 = c \cdot e^0$. Dus $c = 1$.

36 a. $(x+y)^2 = x^2 + 2xy + y^2$
$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$
$(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$
b. $a_0 = a_7 = 1$
c. 1

$1 + 1x$
$1 + 2x + 1x^2$
$1 + 3x + 3x^2 + 1x^3$
$1 + 4x + 6x^2 + 4x^3 + 1x^4$
$1 + 5x + 10x^2 + 10x^3 + 5x^4 + 1x^5$
$1 + 6x + 15x^2 + 20x^3 + 15x^4 + 6x^5 + 1x^6$
$1 + 7x + 21x^2 + 35x^3 + 35x^4 + 21x^5 + 7x^6 + 1x^7$

d. 120
e. De eerste term is $\binom{10}{0} = 1$ en dat klopt.

De laatste term is $\binom{10}{10} = 1$ en dat klopt.

37 a. $4{\cdot}3{\cdot}2{\cdot}1\, a_4 + 5{\cdot}4{\cdot}3{\cdot}2\, a_5 x + 6{\cdot}5{\cdot}4{\cdot}3\, a_6 x^2 + 7{\cdot}6{\cdot}5{\cdot}4\, a_7 x^3 + 8{\cdot}7{\cdot}6{\cdot}5\, a_8 x^4 + 9{\cdot}8{\cdot}7{\cdot}6\, a_9 x^5 + 10{\cdot}9{\cdot}8{\cdot}7\, a_{10} x^6$
b. links: $10{\cdot}9{\cdot}8{\cdot}7{\cdot}1^6 = 10 \cdot 9 \cdot 8 \cdot 7$
rechts: $4{\cdot}3{\cdot}2{\cdot}1{\cdot}a_4$
c. Kennelijk is $10 \cdot 9 \cdot 8 \cdot 7 = 4{\cdot}3{\cdot}2{\cdot}1{\cdot}a_4$
Delen door $4{\cdot}3{\cdot}2{\cdot}1$ geeft het resultaat.
d. Vermenigvuldig in $\frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1}$ teller en noemer met

$6{\cdot}5{\cdot}4{\cdot}3{\cdot}2{\cdot}1$. Dan krijg je: $\frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{4 \cdot 3 \cdot 2 \cdot 1 \bullet 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{10!}{4! \cdot 6!} = \binom{10}{4}$

38 a. $\binom{10}{k} \cdot p^k \cdot q^{10-k}$

b. Rechts staat de kansen dat een kogeltje in een bakje komt over alle bakjes opgeteld. Het kogeltje komt zeker in een van de bakjes. Die totale kans is dus 1.

c. $a^{10} \cdot \sum_{k=0}^{10} \binom{10}{k} \cdot p^k \cdot q^{10-k} = \sum_{k=0}^{10} \binom{10}{k}\ a^{10} \cdot p^k \cdot q^{10-k} =$

$\sum_{k=0}^{10} \binom{10}{k} \cdot a^k p^k \cdot a^{10-k} q^{10-k} = \sum_{k=0}^{10} \binom{10}{k} (ap)^k \cdot (aq)^{10-k}$.

d. Het rechterlid is duidelijk. Links komt er te staan: Omdat $p+q = 1$, is $x+y = ap + aq = a(p+q) = a$ en daarmee is het linkerlid verklaard.

39 **a.** Kies in het binomium van Newton y =1 en het resultaat van opgave komt er te staan.

b. Kies n=2: $(x+y)^2 = \binom{2}{0}x^2 + \binom{2}{1}xy + \binom{2}{2}y^2 = x^2 + 2xy + y^2$

Kies n=3: $(x+y)^3 = \binom{3}{0}x^3 + \binom{3}{1}x^2y + \binom{3}{2}xy^2 + \binom{3}{3}y^3 =$

$x^3 + 3x^2y + 3xy^2 + y^3$

Kies n=4: $(x+y)^4 = \binom{4}{0}x^4 + \binom{4}{1}x^3y + \binom{4}{2}x^2y^2 + \binom{4}{3}x^2y^2 + \binom{4}{4}y^4$

$= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$

40 **a.** Kies $x = y = 1$ en de formule komt er te staan.

b. Kies x = -1 en y = 1. Dan levert het binomium van Newton dat de gevraagde som gelijk is aan $(1+-1)^n = 0$.

41 **a.** De afgeleide van e^x voor x=0 is 1. Dus de raaklijn aan de grafiek heeft vergelijking $y = x+1$. Dicht bij 0 zijn $y=e^x$ en $y=x+1$ ongeveer gelijk. Dus $e^a \approx a + 1$ als a dicht bij 0 ligt.

b. Neem van beide leden van $e^{\frac{x}{n}} \approx 1+\frac{x}{n}$ de n-de macht.

c. Kies bijvoorbeeld n = 1000 en x = 1. Dan geeft de formule: $e \approx 1{,}001^{1000}$ en dat is ongeveer 2,7169.
Kies bijvoorbeeld n = 1000 en x = -1. Dan geeft de formule: $e \approx 1 / 0{,}999^{1000}$ en dat is ongeveer 2,7196.
e ligt tussen deze twee waarden in.