4.5 Poissonverdeling

Inleiding

Behalve binomiale en hypergeometrische kansmodellen bestaan er nog vele andere discrete kansmodellen. In deze paragraaf maak je kennis met een ander veel voorkomende discreet kansmodel: het Poissonverdeelde kansmodel.
Dit kansmodel is bedacht door **Siméon Poisson (1781—1840)**.

Figuur 1 Siméon Poisson

Je leert in dit onderwerp

- werken met discrete stochasten die Poisson verdeeld zijn
- de bijbehorende kansverdelingen opstellen
- regels voor de verwachtingswaarde en de standaardafwijking van die stochasten.

Voorkennis

- een kansverdeling opstellen bij een stochast;
- de verwachting, de variantie en standaardafwijking berekenen op basis van een kansverdeling;
- eventueel (maar niet noodzakelijk): de somformule van een rekenkundige reeks
- eventueel (maar niet noodzakelijk): rekenen met het natuurlijke getal e

Verkennen

Opgave V1

Een loterij bestaat uit zeshonderd briefjes met de nummers 1 t/m 600. Je koopt een lootje.

a Hoeveel bedraagt de verwachtingswaarde van het lotnummer?

De loterij verdeelt bij elke trekking flink wat troostprijzen.

In de gemeente PolderWei ontvangt men per trekking gemiddeld 3 van deze troostprijzen.

b Hoe groot is de kans dat PolderWei meer dan 3 troostprijzen ontvangt bij de volgende trekking?

Uitleg

Een klein bedrijf krijgt dagelijks gemiddeld 3 telefoontjes. Hoe groot is de kans dat dit bedrijf op een dag 4 telefoontjes krijgt?

De stochast X staat voor het aantal telefoontjes dat het bedrijf op een dag binnenkrijgt. Ga ervan uit dat er 8 uur op een dag telefoontjes binnen kunnen komen, dat er voldoende medewerkers aanwezig zijn om alles af te handelen en dat de telefoontjes onafhankelijk van elkaar en willekeurig binnenkomen.

- Verdeel de dag in 8 uren. De kans is $\frac{3}{8}$ dat er in een uur een telefoontje binnenkomt. De verwachtingswaarde van het aantal telefoontjes op een dag is $\frac{3}{8} \cdot 8 = 3$.

Neem $n = 8$ en $p = \frac{3}{8}$ en bereken de kans op 4 telefoontjes:

$$\binom{8}{4} \cdot \left(\frac{3}{8}\right)^4 \cdot \left(\frac{5}{8}\right)^4 \approx 0{,}2112$$

$$\text{Var}\,(X) = 8 \cdot \frac{3}{8} \cdot \left(1 - \frac{3}{8}\right) = 1{,}875$$

- Verdeel de dag in 480 minuten. De kans is $\frac{3}{480}$ dat er in een minuut een telefoontje binnenkomt.

  De verwachtingswaarde van het aantal telefoontjes op een dag is $\frac{3}{480} \cdot 480 = 3$.

  Neem $n = 480$ en $p = \frac{3}{480}$ en bereken de kans op 4 telefoontjes:

  $$\binom{480}{4} \cdot \left(\frac{3}{480}\right)^4 \cdot \left(\frac{477}{480}\right)^{476} \approx 0{,}1686$$

  $$\text{Var}\,(X) = 480 \cdot \frac{3}{480} \cdot \left(1 - \frac{3}{480}\right) = 2{,}98125$$

Je kunt de dag in nog kleinere tijdsintervallen verdelen. Hoe kleiner de tijdsintervallen, hoe kleiner de kans op ‘succes’ in zo'n tijdsinterval. In theorie kan er meer dan één keer ‘succes’ zijn in een tijdsinterval, maar deze kans is nog veel kleiner en wordt daarom verwaarloosd. Je hebt nu te maken met n tijdsintervallen, met telkens dezelfde kans p op ‘succes’. Om te bepalen wat de kans is dat er 4 telefoontjes op een dag binnenkomen, moet je n naar oneindig laten gaan en p naar nul, maar zodanig dat $n \cdot p = 3$. Het blijkt dan dat $\text{P}\,(X = 4) \approx 0{,}1680$ en dat $\text{Var}\,(X) = 3$.

Een kansverdeling die voldoet aan deze situatie wordt een Poissonverdeling genoemd. Deze kansverdeling is ontdekt door Siméon Poisson.

Stel dat het bedrijf gemiddeld λ telefoontjes op een dag binnenkrijgt, dan kun je de volgende formule gebruiken:

$\text{P}\,(X = x) = \frac{\lambda^x}{x!} \cdot \text{e}^{-\lambda}$, waarbij $\text{e} = 2{,}718...$

Het getal e kun je net zoals π vinden op de grafische rekenmachine.

Verder geldt: $\text{E}\,(X) = \text{Var}\,(X) = \lambda$

Kansen bij een Poissonverdeling kunnen ook met de grafische rekenmachine berekend worden, bekijk daarvoor het **Practicum**.

Opgave 1

Bekijk de **Uitleg** over een bedrijf dat gemiddeld 3 telefoontjes op een dag binnenkrijgt.

a Controleer met de formule dat $\text{P}\,(X = 4) \approx 0{,}1680$.

b Bereken de kans dat er op een dag 5 telefoontjes binnenkomen. Rond af op vier decimalen.

c Bereken de kans dat er meer dan 1 telefoontje op een dag binnenkomt. Rond af op vier decimalen.

d Voor het bepalen wat de kans is dat het bedrijf in een werkweek 16 telefoontjes binnenkrijgt, kun je ook de formule uit de uitleg gebruiken. Alleen nu is λ gelijk aan het gemiddeld aantal telefoontjes in een werkweek.

Bereken in vier decimalen de kans dat het bedrijf in een werkweek 16 telefoontjes binnenkrijgt.

Opgave 2

Bekijk de **Uitleg** over de Poissonverdeling.

a Verdeel de dag in seconden en benader de kans van $\text{P}\,(X = 4)$.

b Verdeel de dag in n tijdsintervallen en zorg dat $np = 3$ met p de kans op succes.

Toon aan dat de kans op 4 telefoontjes op een dag als je dit binomiaal benadert, gelijk is aan:

$$\text{P}\,(X = 4) = \binom{n}{4} \cdot \left(\frac{3}{n}\right)^4 \cdot \left(1 - \frac{3}{n}\right)^{n-4}$$

c Je mag aannemen dat $\lim\limits_{n \to \infty} \left(1 - \frac{3}{n}\right)^{n-4} = \mathrm{e}^{-3}$.

Toon aan dat $\lim\limits_{n \to \infty} \binom{n}{4} \cdot \left(\frac{3}{n}\right)^4 \cdot \left(1 - \frac{3}{n}\right)^{n-4} = \frac{3^4}{4!} \cdot \mathrm{e}^{-3}$.

d Toon aan dat $\mathrm{Var}(X) = 3$. Gebruik daarvoor de variantie van een binomiaal verdeelde stochast.

Theorie en voorbeelden

Om te onthouden

Een discrete stochast X heeft een **Poissonverdeling** met parameter λ (lambda), als er verwacht wordt dat een bepaalde relatief zeldzame gebeurtenis gemiddeld λ keer voorvalt in een bepaalde tijdsperiode. Voorwaarden zijn wel dat de gebeurtenissen:

- niet tegelijk kunnen optreden
- onafhankelijk van elkaar optreden
- willekeurig voorkomen

De laatste voorwaarde betekent dat als je de tijdsperiode in gelijke delen verdeelt, de kans dat een gebeurtenis in elk tijdsdeel voorvalt dan even groot is.

Een tijdsperiode kan ook een lengte, afstand, oppervlakte, enzovoort zijn.

Er geldt:

- $\mathrm{P}(X = x) = \frac{\lambda^x}{x!} \cdot \mathrm{e}^{-\lambda}$
- $\mathrm{E}(X) = \lambda$ en $\mathrm{Var}(X) = \lambda$

In deze paragraaf mag je ervan uitgaan dat de gebeurtenissen (bij benadering) een Poissonverdeling hebben, tenzij anders vermeld.

Voorbeeld 1

Gemiddeld komen er wereldwijd 60 vulkaanuitbarstingen per jaar voor.

Wat is in drie decimalen de kans dat er in een week twee vulkaanuitbarstingen plaatsvinden?

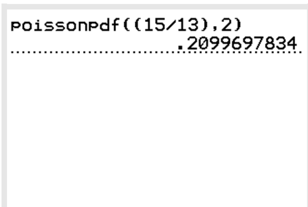
Antwoord

Stochast X is het aantal vulkaanuitbarstingen per week en je mag voor X de Poissonverdeling gebruiken. Opmerking: of je in werkelijkheid X als Poissonverdeling mag beschouwen is afhankelijk van verschillende complexe factoren, waar verder niet op wordt ingegaan.

Als je uitgaat van 52 weken in een jaar, dan is het gemiddelde aantal vulkaanuitbarstingen per week gelijk aan $\lambda = \frac{60}{52} = 1\frac{2}{13}$.

$$\mathrm{P}\left(X = 2 | \lambda = 1\tfrac{2}{13}\right) = \frac{\left(1\frac{2}{13}\right)^2}{2!} \cdot \mathrm{e}^{-1\frac{2}{13}} \approx 0{,}210$$

Dit kun je ook met de grafische rekenmachine uitrekenen.



Figuur 2

Opgave 3

In **Voorbeeld 1** is gegeven dat er gemiddeld 60 vulkaanuitbarstingen per jaar zijn en dat dit Poissonverdeeld is.

a Waarom mag je het gemiddelde per jaar delen door 52 om het gemiddelde per week te krijgen?

b Hoe groot is de kans dat er 7 vulkaanuitbarstingen in een maand zijn?

Opgave 4

Tijdens mooie herfstdagen komen er gemiddeld 7 klanten per uur een ijsje kopen bij een snackbar.

a Waarom kun je het aantal klanten per uur bij deze snackbar tijdens mooie herfstdagen behandelen als een Poissonverdeelde stochast?

b Hoe groot is de kans dat er op een zomerse herfstdag minder dan 7 klanten per uur een ijsje kopen bij de snackbar? Rond af op vier decimalen.

c Hoe groot is de kans dat er tijdens zo'n mooie herfstdag in twee uur tijd 7 klanten een ijsje kopen bij de snackbar? Rond af op vier decimalen.

d Hoe groot is de standaardafwijking van het aantal klanten per uur op een mooie herfstdag? Rond af op twee decimalen.

Voorbeeld 2

Een secretaresse maakt gemiddeld één typefout op een bladzijde tekst.

Bereken de kans dat de secretaresse op een bladzijde meer dan vijf typefouten maakt.

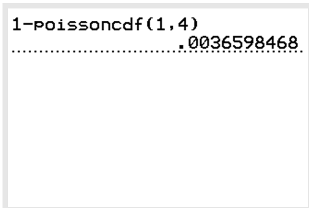
Antwoord

Noem X het aantal typefouten dat de secretaresse op een bladzijde maakt. Je mag voor X de Poissonverdeling gebruiken.

$\text{P}(X > 5 | \lambda = 1) = 1 - \text{P}(X \leq 4)$

Je kunt $\text{P}(X \leq 4)$ met de hand uitrekenen, maar met de grafische rekenmachine gaat dat sneller.

Je vindt $\text{P}(X > 5) \approx 0{,}004$.



Figuur 3

Opgave 5

In **Voorbeeld 2** zie je dat een secretaresse gemiddeld 1 typefout op een bladzijde tekst maakt.

a Bereken in vier decimalen de kans dat de secretaresse minder dan 4 typefouten op een bladzijde maakt.

b Bereken in vier decimalen de kans dat de secretaresse op vijf bladzijden tekst minstens 6 typefouten maakt.

Opgave 6

Van een softwareprogrammeur is bekend dat zij gemiddeld per 500 regels code 3 bugs (softwarefoutjes) veroorzaakt. Haar opdrachtgever geeft haar de opdracht om een stuk software te schrijven dat 500 regels code zal bevatten.

a Hoe groot is de kans dat er helemaal geen bugs in zullen zitten? Rond af op vier decimalen.

b Hoeveel bugs verwacht je als de programmeur 4 stukken code van ieder 500 regels maakt?

c Hoe groot is de kans dat er in totaal meer dan 3 bugs in de 4 stukken code staan? Rond af op vier decimalen.

Verwerken

Opgave 7

Waarom is bij de volgende situaties een Poissonverdeling niet zo geschikt om te gebruiken?

1. Het aantal klanten dat in een restaurant binnenkomt.
2. Het aantal eieren dat een kip legt.

Opgave 8

Bij de balie van de gemeente komen tussen 10:00 en 11:00 uur gemiddeld 6 bezoekers per vijf minuten.

a Bereken de verwachtingswaarde en de standaardafwijking van het aantal bezoekers per vijf minuten tussen 10:00 en 11:00 uur. Rond af op twee decimalen.

b Bereken de kans dat er tussen 10:10 en 10:15 uur precies 6 bezoekers komen. Rond af op vier decimalen.

c Bereken de kans dat er tussen 10:25 en 10:30 uur meer dan 4 bezoekers komen. Rond af op vier decimalen.

Opgave 9

Tussen 1960 en 1980 werden er gemiddeld 2 grienden (een walvissoort) per decennium aan de Nederlandse kusten gevonden.

a Waarom kun je uitgaan van een Poissonverdeelde stochast G voor het aantal grienden per decennium? Welke verwachtingswaarde heeft G?

Beargumenteer je antwoorden.

b Maak de kansverdeling voor stochast G, het aantal grienden dat per decennium aan de Nederlands kust werd gevonden in de jaren tussen 1960 en 1980.

Geef de kansen voor de waarden $G = 0$ tot en met $G = 6$. Bereken $\text{P}(G \geq 7)$.

Opgave 10

De vader van Jelle belegt iedere ochtend een heel stokbrood met een hagelslagsoort waarin ook een aantal gele chocoladesterretjes zitten. Hij maakt er lunchpakketjes voor zijn kinderen van. Iedere ochtend tellen ze met elkaar de gele chocoladesterretjes die op het stokbrood terechtkomen.

Tot hun verbazing bevat het belegde stokbrood op een dag in helemaal geen gele chocoladesterretjes.

a Hoeveel gele chocoladesterretjes strooit vader 's ochtends gemiddeld op het stokbrood als de kans op helemaal geen geel chocoladesterretje kleiner is dan 1%?

In de loop van de tijd blijkt het aantal gele chocoladesterretjes op het stokbrood te dalen naar gemiddeld 2 stuks.

b Hoe groot is de kans dat Jelles vader op een ochtend toch meer dan 4 gele chocoladesterretjes op het stokbrood strooit? Rond af op vier decimalen.

Opgave 11

Het aantal klanten K dat per kwartier aan de balie van een klantenservice verschijnt, is een Poissonverdeelde stochast. De kans dat er in een kwartier slechts 1 klant aan de balie staat, is afgerond 0,0027.
Bereken de kans dat er in een minuut meer dan 1 klant bij de balie komt. Rond af op vier decimalen.

Opgave 12

De toepassing waardoor de Poissonverdeling echt bekendheid kreeg, is uitgevoerd aan het einde van de 19^{e} eeuw. De statisticus Ladislaus Bortkiewicz gebruikte de Poissonverdeling bij zijn onderzoek naar dodelijke ongelukken in het Pruisische leger die veroorzaakt werden door een trap van een paard. In een periode van 20 jaar hebben 10 Pruisische legerkorpsen jaarlijks aan de legerleiding gemeld hoeveel manschappen er waren omgekomen door de trap van een paard. Dat levert deze, ondertussen beroemde, frequentietabel op.

aantal doden door trap van paard per melding	0	1	2	3	4	≥ 5
frequentie	109	65	22	3	1	0

Tabel 1

Toon aan dat deze frequentieverdeling bij benadering overeenkomt met een Poissonverdeling.

Toepassen

Opgave 13: Poissonverdeling als limietgeval

Je hebt gezien dat je de Poissonverdeling kunt zien als een soort limiet geval van de binomiale verdeling. Stel dat stochast X Poissonverdeeld is met parameter λ. Gebruik verder $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \mathrm{e}$.

a Toon aan dat uit de gegeven limiet volgt dat: $\lim_{n\to\infty}\left(1-\frac{x}{n}\right)^n = \mathrm{e}^{-x}$

b Toon aan dat $\mathrm{P}(X = x) = \frac{\lambda^x}{x!}\cdot \mathrm{e}^{-\lambda}$.

Maak gebruik van dat $\lim_{n\to\infty}\left(1-\frac{x}{n}\right)^n = \mathrm{e}^{-x}$.

c Toon aan dat $\mathrm{Var}(X) = \lambda$.

Testen

Opgave 14

De redactie van de dagelijkse krant PolderJournaal is trots op het feit dat ze gemiddeld slechts 2 typefouten per dag maken.

a Bereken, zonder gebruik te maken van de Poissonfuncties op de grafische rekenmachine, de kans dat er een PolderJournaal wordt gedrukt zonder typefout.

Controleer je antwoord door deze kans ook met de Poissonfunctie op de grafische rekenmachine te berekenen.

b Hoe groot is de kans dat er een PolderJournaal wordt gedrukt met meer dan 2 typefouten?

c Hoeveel typefouten zijn er te verwachten in vijf verschillende PolderJournaals?

Opgave 15

In een fabriek voor diepvriespizza's strooit een machine olijven over de pizza's.

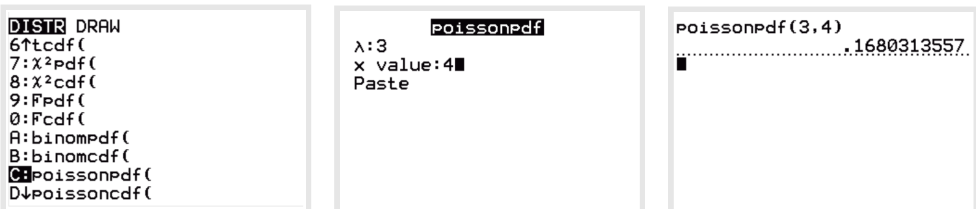
Wat blijkt, na vele steekproeven? Ongeveer één op de 13 pizza's heeft precies 1 olijf van de machine gekregen.

Wat is de standaardafwijking van het aantal olijven per pizza?

Practicum

Je hebt al gezien hoe je met je grafische rekenmachine met de binomiale kansverdeling kunt werken. Er zijn echter nog veel meer soorten kansverdelingen. De zogenaamde ‘normale verdeling’ ga je nog tegenkomen.

Hier hebben we het over de **Poissonverdeling**.
Op dezelfde plaats waar je de binomiale verdeling vindt, tref je ook de Poissonverdeling aan.
De Poissonverdeling heeft echter maar één parameter, namelijk λ.
Hier zie je bijvoorbeeld hoe $\mathrm{P}(X = 4 | \lambda = 3)$ wordt berekend op de TI-84. Op andere machines gaat dit op een vergelijkbare wijze.



Figuur 4