APPENDIX: HET POISSON PROCES

Een stochastisch proces dat onlosmakelijk verbonden is met de Poisson verdeling is het Poisson proces. Dit is een telproces dat het aantal optredens van een bepaalde gebeurtenis telt over de loop van de tijd. De gebeurtenissen kunnen van velerlei aard zijn: de aankomst van klanten bij een bank, het optreden van ernstige aardbevingen, de aankomst van telefoongesprekken bij een telefooncentrale, de binnenkomst van storingen in een electriciteitscentrale, de binnenkomst van noodoproepen bij een alarmcentrale, etc. In de rest van deze paragraaf zullen we voor het optreden in de tijd van gebeurtenissen de terminologie hanteren van aankomsten van klanten. Wanneer is een telproces een Poisson proces? Daarvoor is het noodzakelijk een *onbegrensd grote* populatie van potentiële klanten te veronderstellen, waarbij de klanten zich onafhankelijk van elkaar gedragen. Het aankomstproces van klanten bij een bedieningsfaciliteit heet een *Poisson proces* als het proces de volgende eigenschappen heeft:

a. de klanten komen één voor één aan

b. de aantallen aankomsten in disjuncte tijdsintervallen zijn onafhankelijk van elkaar

c. de kansverdeling van het aantal klanten dat aankomt in een gegeven tijdsinterval heeft een Poisson verdeling waarvan de verwachtingswaarde evenredig is met de lengte van het tijdsinterval.*

Noteren we met

$$\lambda = \text{de verwachtingswaarde van het aantal aankomsten in een tijdsinterval van eenheidslengte,}$$

dan stelt eigenschap c. dat voor elke vaste $t > 0$ geldt

$$P(k \text{ aankomsten in een tijdstinterval van lengte } t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!} \quad \text{voor } k = 0, 1, \ldots \ .$$

*In feite kan de derde aanname verzwakt worden tot de aanname dat de kansverdeling van het aantal aankomsten in een gegeven tijdsinterval alleen afhangt van de lengte van het tijdsinterval en niet van de positie van het interval op de tijdas. Verder moet dan worden aangenomen dat de kans op twee of meer aankomsten in een zeer klein tijdsinterval van lengte h gegeven dat een aankomst in het tijdsinterval heeft plaatsgevonden naar 0 gaat als h naar 0 gaat.

Het getal λ heet de *aankomstintensiteit* van het Poisson proces.

Het Poisson proces geeft een goede modelbeschrijving in veel praktische situaties. De verklaring hiervan is gelegen in het volgende, ruwweg geformuleerde, resultaat. Stel dat op micro niveau sprake is van een *zeer groot aantal* stochastische processen die *onafhankelijk* van elkaar zijn en die elk *spaarzaam* aankomsten genereren. Dan kan aangetoond worden dat op macro niveau de resultante van al deze processen bij goede benadering een Poisson proces is. Dit verklaart bijvoorbeeld waarom het aankomstproces van klanten bij een bank of een postkantoor in veel gevallen goed beschreven wordt door een Poisson proces. Je hebt een zeer grote populatie van potentiële klanten: ieder afzonderlijk persoon gedraagt zich weliswaar volgens een bepaald patroon maar de samenvoeging van al deze typisch onafhankelijke patronen leidt tot een onvoorspelbaar geheel dat zich goed laat beschrijven door een Poisson proces.

De eigenschap b. die het ontbreken van na-effecten in het aankomstproces veronderstelt, is de meest kenmerkende eigenschap van het Poisson proces en kan alleen vervuld zijn als de populatie van klanten onbegrensd groot is Intuitief zal duidelijk zijn dat deze eigenschap impliceert: voor elk gegeven tijdstip t geldt dat de kansverdeling van de tijd tot de eerstvolgende aankomst niet afhangt van hoe lang geleden het is dat voor het laatst een klant binnenkwam. Deze eigenschap staat bekend als de *geheugenloosheidseigenschap* van het Poisson proces. Het Poisson proces met deze geheugenloosheidseigenschap staat lijnrecht tegenover een deterministisch aankomst proces waarin klanten met vaste tussentijden aankomen.

Voorbeeld A.1 Bij het Centraal Station staan groeptaxi's. Een groeptaxi vertrekt zodra vier passagiers zijn ingestopt of tien minuten verstreken is sinds de eerste passagier is ingestapt. Veronderstel dat passagiers aankomen volgens een Poisson proces met een gemiddelde van één passagier per drie minuten.

(a) Je stapt als eerste passagier in. Wat is de kans dat je tien minuten moet wachten tot het vertrek van de taxi?

(b) Je bent als eerste passagier ingestapt en je zit al vijf minuten in de taxi. Inmiddels zijn ook twee andere passagiers ingestapt. Wat is de kans dat je nog vijf minuten moet wachten tot het vertrek van de taxi?

Het antwoord op vraag (a) berust op de constatering dat je 10 minuten moet wachten alleen dan als het aantal passagiers dat de komende 10 minuten aankomt minder dan drie is. Kiezen we de minuut als tijdseenheid, dan is dit aantal Poisson verdeeld en heeft de verwachtingswaarde 10λ met $\lambda = \frac{1}{3}$. Dus

$$\begin{aligned}
&P(\text{je moet 10 minuten wachten})\\
&= P(\text{komende 10 minuten stappen 0, 1 of 2 passagiers in})\\
&= e^{-10\lambda} + e^{-10\lambda}\frac{10\lambda}{1!} + e^{-10\lambda}\frac{(10\lambda)^2}{2!} = 0.3528.
\end{aligned}$$

De beantwoording van vraag (b) berust op de geheugenloosheidseigenschap van het Poisson proces. De wachttijd tot de aankomst van de eerstkomende passagier is exponentieel verdeeld met verwachtingswaarde van drie minuten, ongeacht hoe lang geleden de laatste passagier instapte. De kans dat je nog 5 minuten moet wachten is gelijk aan de kans $e^{-5\lambda} = 0.1889$ dat de komende vijf minuten er geen nieuwe passagier bij komt.

Voorbeeld A.2 Je komt op een random moment tussen 5 uur en kwart over 5 bij de bushalte aan de eerstkomende bus naar huis te nemen. Zowel buslijn 1 als buslijn 3 kun je nemen. Buslijn 1 vertrekt precies op het uur en elke 15 minuten daarna van de halte, terwijl buslijn 3 gemiddeld één keer in de 15 minuten vertrekt waarbij de vertrekmomenten worden bepaald door een Poisson proces. Wat is de kans dat je met buslijn 1 naar huis gaat?

Onder de conditie dat vanaf het moment van je aankomst bij de halte het nog x minuten duurt alvorens bus 1 volgens schema aankomt, is de kans dat je met bus 1 naar huis gaat gelijk aan de kans dat de komende x minuten geen bus 3 aankomt. Deze kans is gelijk aan $e^{-\lambda x}$, ongeacht hoe lang geleden bus 3 voor het laatst aankwam. Hierbij is $\lambda = \frac{1}{15}$. Door te middelen over de uniform verdeelde tijd dat het duurt tot bus 1 aankomt, vind je

$$P(\text{je gaat met bus 1 naar huis}) = \frac{1}{15}\int_0^{15} e^{-\lambda x}\, dx = 1 - \frac{1}{e}(> \frac{1}{2}).$$

Alternatieve definities van het Poisson proces

Bij een Poisson aankomstproces is het aantal aankomsten in een gegeven tijdsinterval *discreet* verdeeld, maar heeft de tijd tussen twee opeenvolgende aankomsten een *continue* kansverdeling en wel de exponentiële verdeling.*

*Het Poisson proces heeft zowel een discrete component (Poisson verdeling voor aantal aankomsten) als een continue component(exponentiële verdeling voor tussenaankomsttijden). Deze zaken worden nog wel eens door elkaar gehaald. Een illustratieve situatie is de volgende. Uit fysische experimenten is bekend dat de emissie van alpha-deeltjes uit radioactief materiaal goed beschreven kan worden door een Poisson proces: het aantal deeltjes dat in een gegeven tijdsinterval uitgestoten wordt is Poisson verdeeld en de tijden tussen de uitstoot van opeenvolgende deeltjes zijn onafhankelijk en exponentieel verdeeld.

Dit is als volgt in te zien:

$$\begin{aligned}&P(\text{tijd tussen twee opeenvolgende aankomsten is groter dan } y)\\&= P(\text{in een interval van de lengte } y \text{ vindt geen aankomst plaats})\\&= e^{-\lambda y} \quad \text{voor elke } y > 0.\end{aligned}$$

Bij een Poisson aankomstproces met aankomstintensiteit λ heeft de stochastische tijd T tussen twee opeenvolgende aankomsten dus de kansverdelingsfunctie

$$P(T \leq y) = 1 - e^{-\lambda y} \quad \text{voor } y \geq 0.$$

Dit is een continue verdelingsfunctie met als afgeleide de exponentiële kansdichtheidsfunctie $\lambda e^{-\lambda y}$. De verwachtingswaarde van de exponentiële verdeling is $\frac{1}{\lambda}$. Een verwachtingswaarde van $\frac{1}{\lambda}$ voor de tussenaankomsttijd is in overeenstemming met de betekenis van de aankomstintensiteit λ. Vanwege eigenschap b. van het Poisson proces zal het niemand verbazen dat de tijden tussen de aankomsten van opeenvolgende klanten onafhankelijk van elkaar zijn. De volgende alternatieve definitie van het Poisson proces kan dan worden gegeven:

Een aankomstproces van klanten is een *Poisson proces* dan en slechts dan als de tussenaankomsttijden van klanten onderling *onafhankelijke* stochastische variabelen zijn met eenzelfde *exponentiële* kansverdeling.

Deze definitie is de basis voor de simulatie van aankomsttijdstippen in een Poisson aankomstproces met intensiteit λ . Je doet elke keer met de inversiemethode trekkingen uit de exponentiële kansverdelingsfunctie $F(t) = 1-e^{-\lambda t}$. Dit doe je door bij een random getal u de tussentijd t op te lossen uit de vergelijking $F(t) = 1 - e^{-\lambda t} = u$, oftewel $t = -\frac{1}{\lambda}\ln(1 - u)$.

Uit de tweede alternatieve definitie van het Poisson proces leiden we vervolgens een derde alternatieve definitie af die voor toepassingen veelal de meest bruikbare is. Daartoe hebben we het begrip "faalintensiteit" van een positieve stochastische variabele nodig. Dit begrip is het simpelst uit te leggen als we veronderstellen dat de positieve stochast X de levensduur van een machine voorstelt, waarbij we tevens veronderstellen dat X exponentieel verdeeld is met verwachtingswaarde $\frac{1}{\mu}$. Dan geldt, ongeacht de waarde van de leeftijd x van de machine

$$\begin{aligned}&P(\text{machine van leeftijd } x \text{ faalt in het komend tijdje } \Delta x)\\&\approx \mu\Delta x + o(\Delta x) \qquad \text{for } \Delta x \to 0,\end{aligned}$$

waarbij $o(\Delta x)$ een algemene verzamelnaam is voor een restterm die verwaarloosbaar klein is ten opzichte van Δx als Δx zelf heel klein is. Een machine met een exponentiële levensduur heeft dus een *constante* faalintensiteit μ ("used is good as new"). De afleiding van dit resultaat is als volgt. Op grond van $P(X \leq x) = 1 - e^{-\mu x}$ voor $x > 0$ geldt

$$\begin{aligned} P(X \leq x + \Delta x \mid X > x) &= \frac{P(x < X \leq x + \Delta x)}{P(X > x)} \\ &= \frac{e^{-\mu x} - e^{-\mu(x+\Delta x)}}{1 - e^{-\mu x}} = 1 - e^{-\mu \Delta x}. \end{aligned}$$

Uit de reeksontwikkeling $1 - e^{-u} = 1 - u + \frac{u^2}{2!} - \frac{u^3}{3!} + \cdots$ volgt

$$1 - e^{-\mu \Delta x} = \mu \Delta x - \frac{(\mu \Delta x)^2}{2!} + \frac{(\mu \Delta x)^3}{3!} - \cdots = \mu \Delta x + o(\Delta x),$$

waarbij $o(\Delta x)$ de wiskundige notatie is voor elke functie van Δx die verwaarloosbaar klein is ten opzichte van Δx als Δx zelf heel klein is, dwz $o(\Delta x)/\Delta x \to 0$ als $\Delta x \to 0$. Op grond van bovenstaande kunnen we nu de derde alternatieve definitie van het Poisson proces geven:

Een aankomstproces $\{N(t), t \geq 0\}$ van klanten, waarbij $N(t)$ het totale aantal aankomsten tot tijdstip t aangeeft, heet een *Poisson proces* met aankomstintensiteit λ als:

a. De aantallen aankomsten in disjuncte tijdsintervallen onafhankelijk van elkaar zijn,

b. Voor elke $t \geq 0$, $P(N(t + \Delta t) - N(t) = 1) = \lambda \Delta t + o(\Delta t)$ als $\Delta t \to 0$,

c. Voor elke $t \geq 0$, $P(N(t + \Delta t) - N(t) \geq 2) = o(\Delta t)$ als $\Delta t \to 0$.

Deze definitie heeft als bijkomend voordeel dat de definitie eenvoudig uitbreidbaar is naar de stuatie waarin de aankomstintensiteit van klanten tijdsafhankelijk is.

Samenvoegen en splitsen van Poisson processen

In toepassingen komt vaak de situatie voor dat twee Poisson processen samengevoegd worden of dat een Poisson proces uitgedund wordt. Stel dat een call-center telefonische verzoeken voor informatie afhandelt voor twee totaal verschillende bedrijven. Gesprekken voor het ene bedrijf A komen binnen volgens een Poisson proces met aankomstintensiteit λ_A en onafhankelijk daarvan komen gesprekken voor het andere bedrijf B binnen volgens een Poisson proces met aankomstintensiteit λ_B. De samenvoeging van deze twee aankomstprocessen is dan een Poisson proces met aankomstintensiteit $\lambda_A + \lambda_B$. Verder kan bewezen worden dat een random gekozen gesprek met kans $\frac{\lambda_A}{\lambda_A+\lambda_B}$ een gesprek voor bedrijf A is en met kans $\frac{\lambda_B}{\lambda_A+\lambda_B}$ een gesprek voor bedrijf B.

Als voorbeeld voor het splitsen van een Poisson proces beschouw de situatie dat het optreden in de tijd van aardbevingen in een bepaald gebied beschreven wordt door een Poisson proces met intensiteit λ. Veronderstel dat de sterktes van de aardbevingen onafhankelijk van elkaar zijn, waarbij de sterkte van een aardbeving met kans p geclassificeerd wordt als een zware en met kans $1 - p$ als een niet-zware. Het kan dan bewezen worden dat het optreden in de tijd van zware aardbevingen beschreven wordt door een Poisson proces met intensiteit λp en van niet-zware aardbevingen door een Poisson proces met intensiteit $\lambda(1 - p)$. Dit resultaat is niet verrassend. Wel verrassend is het resultaat dat de twee Poisson processen die de zware en niet-zware aardbevingen beschrijven *onafhankelijk* van elkaar zijn.

Voorbeeld A.3 Een stuk radioactief materiaal stoot α-deeltjes uit volgens een Poisson proces met een gemiddelde van 0.84 deeltje per seconde. Een Geiger teller registreert, onafhankelijk van elkaar, elk uitgestoten deeltje met kans 0.95. In een 10-seconden periode zijn 12 deeltjes geregistreerd. Wat is de kans dat meer dan 15 deeltjes uitgestoten werden in die periode?

Het proces dat de uitstoot van niet-geregistreerde deeltjes beschrijft is een Poisson proces met een intensiteit van $0.84 \times 0.05 = 0.0402$ deeltje per seconde en dit proces is onafhankelijk van het Poisson proces dat de uitstoot van geregistreerde deeltjes beschrijft. Dus het aantal deeltjes dat door de Geiger teller gemist werd in de 10-seconde periode is Poisson verdeeld met verwachtingswaarde $10 \times 0.0402 = 0.402$, ongeacht het aantal uitgestote deeltjes dat geregistreerd werd. De gevraagde kans is de kans dat in een gegeven periode van 10 seconde meer dan drie niet-geregistreerde deeltjes worden uitgestoten en is dus gelijk aan $1 - \sum_{j=0}^{3} e^{-0.402} \frac{(0.402)^j}{j!} = 0.00079$.

Niet-homogeen Poisson proces

In veel praktische situaties is de aankomstintensiteit van klanten tijdsafhankelijk. De alternatieve definitie van het Poisson proces, waarin aankomsten worden geteld door zeer kleine tijdsintervallen te beschouwen, laat zich direct uitbreiden naar het geval van tijdsafhankelijke aankomsten. Een telproces $\{N(t), t \geq 0\}$, waarbij $N(t)$ het totale aantal aankomsten tot tijdstip t aangeeft, heet een *niet-homogeen Poisson proces* met aankomstintensiteitsfunctie $\lambda(t)$ als:

a. De aantallen aankomsten in disjuncte tijdsintervallen onafhankelijk van elkaar zijn,

b. Voor elke $t \geq 0$, $P(N(t+\Delta t)-N(t)=1)=\lambda(t)\Delta t+o(\Delta t)$ als $\Delta t \to 0$,

c. Voor elke $t \geq 0$, $P(N(t+\Delta t)-N(t)\geq 2)=o(\Delta t)$ als $\Delta t \to 0$.

Hierbij is $o(h)$ een generieke notatie voor elke functie $f(h)$ met de eigenschap dat $f(h)$ sneller naar nul gaat dan h zelf als h naar nul gaat (bijvoorbeeld $f(h)=h^2$ is een $o(h)$ functie). Preciezer gesteld, een functie $f(h)$ is een $o(h)$ functie voor $h \to 0$ als $\lim_{h\to 0} f(h)/h = 0$.

Voor een niet-homogeen Poisson aankomstproces $\{N((t), t \geq 0\}$ kan worden aangetoond dat voor alle t_1, $t_2 \geq 0$ met $t_2 > t_1$ geldt dat het totale aantal aankomsten in het tijdsinterval (t_1, t_2) een *Poisson verdeling* heeft met verwachtingswaarde $M(t_2)-M(t_1)$, waarbij

$$M(t)=\int_0^t \lambda(x)dx \quad \text{voor } t \geq 0.$$

De functie $M(t)$ geeft dus het verwachte aantal aankomsten in $(0,t)$.

Voorbeeld B.4 Bij een pinautomaat komen klanten aan volgens een niet-homogeen Poisson proces. Nemen we zes uur 's ochtends als $t=0$ en het uur als tijdseenheid dan komen klanten tussen zes uur 's ochtends en 12 uur 's middags aan met intensiteit $\lambda(t)=7.5t$ voor $0 \leq t \leq 6$ en tussen 12 uur 's middags en zes uur de volgende ochtend met intensiteit $\lambda(t)=45-2.5(t-6)$ voor $6 \leq t \leq 24$. Wat is de kansverdeling van het totale aantal klanten dat per etmaal van de pinautomaat gebruik maakt? Het antwoord is dat dit aantal een Poisson verdeling heeft met verwachtingswaarde

$$\int_0^{24} \lambda(t)dt = \int_0^6 7.5tdt + \int_6^{24} [45-2.5(t-6)]dt = 540.$$

Stel dat elke klant 50 euro pint met kans $\frac{3}{4}$ en 100 euro met kans $\frac{1}{4}$. Wat is de verwachting van het totale bedrag dat per etmaal wordt gepind? Noem een klant voor 50 euro een type 1 klant en een klant voor 100 euro een type 2 klant. De type 1 klanten komen dan aan volgens een niet-homogeen Poisson proces met intensiteit $\frac{3}{4}\lambda(t)$ en de type 2 klanten volgens een niet-homogeen Poisson proces met intensiteit $\frac{1}{4}\lambda(t)$ (waarom?). Het totale aantal keer dat 50 euro per etmaal gepind wordt is dan Poisson verdeeld met verwachtingswaarde $\frac{3}{4} \times 540 = 405$ en het totale aantal keer dat 100 euro geïnd wordt is Poisson verdeeld met verwachtingswaarde $\frac{1}{4} \times 540 = 135$. Het totale bedrag dat per etmaal gepind wordt, heeft dus verwachtingswaarde $\mu = 50 \times 405 + 100 \times 135 = 33\,750$ euro

Voor het geval de aankomstintensiteitsfunctie $\lambda(t)$ begrensd in $t \geq 0$ is, kan een interessante alternatieve beschrijving van het niet-homogeen Poisson aankomstproces gegeven worden. Kies een vast getal λ met $\lambda \geq \lambda(t)$ voor alle $t \geq 0$. Construeer nu als volgt een telproces:

1. Aankomsten vinden plaats volgens een Poisson proces met intensiteit λ.

2. Een aankomst die plaatsvindt op tijdstip s wordt geaccepteerd met kans $\lambda(s)/\lambda$ en wordt verworpen anders.

Het telproces $\{N(t), t \geq 0\}$ met $N(t)$ het aantal geaccepteerde aankomsten tot tijdstip t is dan een niet-homogeen Poisson proces met aankomstintensiteitsfunctie $\lambda(t)$. Dit is als volgt in te zien. De kans dat het Poisson proces een aankomst genereert in $(t, t + \Delta t)$ is $\lambda \Delta t + o(\Delta t)$ voor Δt klein en de kans dat deze aankomst geaccepteerd wordt is dan $\lambda(t)/\lambda$, oftewel de kans dat in $(t, t + \Delta t)$ een geaccepteerde aankomst plaatsvindt is gelijk aan $(\lambda \Delta t) \times (\lambda(t)/\lambda) + o(\Delta t) = \lambda(t)\Delta t + o(\Delta t)$. Deze constructie van een niet-homogeen Poisson proces is nuttig voor het simuleren van aankomsten in een niet-homogeen Poisson proces.