Vrije Universiteit Amsterdam
Opleiding Wiskunde – Vak Poisson Processen

Poisson Processen

Arno Weber

email: aeweber@cs.vu.nl

Januari 2003

Inhoudsopgave

1. Computersimulaties

In dit verslag houden we ons bezig met het Poisson-proces. We doen enkele experimenten, zowel op de computer als in de dagelijkse praktijk, en verdiepen ons in de theorie achter het model van zo'n proces.

Om te beginnen simuleren we een Poisson-proces op de computer, gebruikmakend van het pakket *Maple7*. We halen eerst uit de theorie even het volgende in herinnering. Wanneer we een dergelijk proces met *intensiteit* λ laten lopen over het tijdsinterval $[0, t]$, weten we dat het aantal gebeurtenissen in dit interval $Pois(\lambda t)$ verdeeld is en de lengte van een tussentijd $Exp(\lambda)$ verdeeld is. Als we onze λ en onze t hebben gekozen, kunnen we een Poisson-proces simuleren door *Maple* een aantal realisaties uit de $Exp(\lambda)$-verdeling te laten uitvoeren. De uitkomst van zo'n realisatie is uiteraard een getal groter dan 0, en deze beschouwen we als een tussentijd. We weten dat de som van alle tot nu toe gegeven tussentijden het tijdstip van de laatste gebeurtenis geeft. We stoppen daarom met het uitvoeren van realisaties wanneer deze totale som t overschrijdt. Het aantal gebeurtenissen dat binnen $[0, t]$ heeft plaatsgevonden is dan gelijk aan het aantal uitgevoerde realisaties minus 1, en heeft een $Pois(\lambda t)$-verdeling.

Het doel van zo'n simulatie zal zijn de parameter λ (die we al kennen) te schatten en te kijken of deze schatting in de buurt komt van de echte λ.
We zullen nu twee manieren bespreken waarop we λ kunnen schatten.

1. We weten dat $1/\lambda$ de verwachting van een $Exp(\lambda)$ verdeelde stochast is. We willen daarom bij het schatten van λ een "goede schatter" voor deze verwachting betrekken. In het geval dat *Maple* n realisaties heeft uitgevoerd, beschikken we over een realisatie van een steekproef $(X_1, X_2, \ldots, X_n)$ uit de $Exp(\lambda)$-verdeling. De tussentijden zijn immers onafhankelijk en hebben dezelfde verdeling. Onze schatter voor λ zal nu zijn: $1/\overline{X_n}$, met $\overline{X_n}$ het steekproefgemiddelde. Onze motivatie hiervoor is dat dit de *meest aannemelijke schatter* voor λ is. Dat laatste zullen we even bewijzen:

   Aangenomen dat alle $X_i > 0$ wordt de *likelihood-functie* gegeven door:

$$L(\lambda, X_1, \ldots, X_n) = \prod_{i=1}^{n} \lambda e^{-\lambda X_i}$$

Deze maximaliseren we naar λ door de *log-likelihood* te bepalen en zijn afgeleide naar λ gelijk aan nul te stellen:

$$\ell(\lambda) = \sum_{i=1}^{n} \log \lambda e^{-\lambda X_i} = \sum_{i=1}^{n} (\log \lambda - \lambda X_i) = n \log \lambda - \lambda \sum_{i=1}^{n} X_i, \text{ dus}$$

$$\dot{\ell}(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} X_i = 0, \text{ dus de schatter wordt:}$$

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\overline{X_n}}$$

□

2. We weten dat λ de verwachting van een $Pois(\lambda)$ verdeelde stochast is. Als we ons tijdsinterval $[0, t]$ kiezen met t een natuurlijk getal, en hierover een Poisson-proces simuleren, kunnen we dit tijdsinterval opsplitsen in t deelintervallen $[k-1, k]$, $k = 1, 2, \ldots, t$ van lengte 1. Het aantal gebeurtenissen in zo'n interval is nu $Pois(\lambda)$ verdeeld. Het is de moeite waard hier even het volgende over op te merken. Wanneer we op een geheeltallig tijdstip overspringen naar een nieuw deelinterval, is het restant van de tussentijd die op dat moment aan het verlopen is, de wachttijd tot de eerste gebeurtenis binnen dit nieuwe deelinterval. Het blijkt dat deze wachttijd $Exp(\lambda)$ verdeeld is, hoewel het een gedeelte is van een tijdsperiode waarvan we dachten dat deze ook $Exp(\lambda)$ verdeeld is. Meer hierover in het tweede onderdeel. Het is nu duidelijk dat nadat we *Maple* aan het werk hebben gezet, we de beschikking hebben over t onafhankelijke realisaties uit de $Pois(\lambda)$-verdeling. Met andere woorden, we beschikken over een steekproef $(Y_1, Y_2, \ldots, Y_t)$ uit de $Pois(\lambda)$-verdeling. Onze schatter voor λ zal nu het steekproefgemiddelde $\overline{Y_t}$ zijn. Overigens hoeven we bij de berekening hiervan niet van ieder deelinterval afzonderlijk het aantal gebeurtenissen te bepalen, maar delen we gewoon het totaal aantal gebeurtenissen door t. Wederom beweren we dat deze schatter de *meest aannemelijke schatter* voor λ is. Het bewijs gaat op dezelfde manier als we eerder hebben gezien:

$$L(\lambda, Y_1, \ldots, Y_t) = \prod_{i=1}^{t} e^{-\lambda} \frac{\lambda^{Y_i}}{Y_i!} \quad (\text{alle } Y_i \geq 0)$$

$$\ell(\lambda) = \sum_{i=1}^{t} \log e^{-\lambda} \frac{\lambda^{Y_i}}{Y_i!} = \sum_{i=1}^{t} (-\lambda) + \sum_{i=1}^{t} (\log \lambda^{Y_i}) - \sum_{i=1}^{t} (\log Y_i!) =$$

$$-\lambda t + \log \lambda \Big(\sum_{i=1}^{t} Y_i \Big) - \sum_{i=1}^{t} (\log Y_i!)$$

$$\dot{\ell}(\lambda) = -t + \frac{\sum_{i=1}^{t} Y_i}{\lambda} = 0, \quad \text{dus} \quad \hat{\lambda} = \frac{\sum_{i=1}^{t} Y_i}{t} = \overline{Y_t}$$

$\square$

Aldus zijn er twee simulaties uitgevoerd, één met $\lambda = 1$ en één met $\lambda = 5$, en beide met $t = 10$. De uitvoer van *Maple* staat in tabel 1 en tabel 2 in de bijlage bij dit verslag. Daarin vinden we ook voor beide simulaties het aantal gebeurtenissen dat in $[0, t]$ heeft plaatsgevonden en de twee schattingen voor λ, op basis van de twee eerder beschreven schatters, met hun procentuele afwijking ten opzichte van de echte λ.
Over de schattingen zelf valt weinig te melden behalve dat ze in beide gevallen redelijk in de buurt van λ liggen. Wel vallen de volgende punten op:

- De twee schattingen liggen bij de tweede simulatie ($\lambda = 5$) dichter bij elkaar dan bij de eerste simulatie ($\lambda = 1$). Dit is achteraf bezien wel begrijpelijk. Het is namelijk zo dat het gedeelte van het tijdsinterval dat ligt tussen de laatste gebeurtenis en t niet bij beide schattingen "meedoet" Bij de schatting op basis van de $Exp(\lambda)$-verdeling delen we het totaal aantal gebeurtenissen door de totale lengte van alle tussentijden (dat is het tijdstip van de laatste gebeurtenis). Maar bij de schatting op basis van de $Pois(\lambda)$-verdeling delen we het totaal aantal gebeurtenissen door t. De lengte van het deelinterval dat tussen de laatste gebeurtenis en t ligt veroorzaakt dus het verschil tussen beide schattingen. Echter deze lengte verwachten we voor een vaste $t \in \mathbb{N}$ kleiner naarmate λ groter is, en dus is het niet zo verwonderlijk dat het verschil tussen de schattingen kleiner is in het geval $\lambda = 5$.

- In het geval $\lambda = 5$ hebben de schattingen een kleinere procentuele afwijking dan in het geval $\lambda = 1$. Ook dat is niet zo vreemd. Immers, voor de schatting op basis van de $Exp(\lambda)$-verdeling hebben we in het tweede geval meer waarnemingen ter beschikking. Dat laatste is ook wel te verwachten, omdat de intensiteit λ in dat geval veel groter is. Dat de schatting op basis van de $Pois(\lambda)$-verdeling in het tweede geval ook beter is, is nu eenvoudig te verklaren door het feit dat beide soorten schattingen voor grote λ's dicht bij elkaar liggen (met grote kans).
  Het is dus al met al niet zo vreemd dat we bij vaste $t \in \mathbb{N}$, betere schattingen voor hogere λ's krijgen.

2. Wachttijd-paradox

Veronderstel dat we een Poisson-proces met intensiteit λ laten lopen over een onbegrensd tijdsinterval (d.w.z. over $[0, \infty)$). We nemen ergens op dit interval een tijdstip t vast en zijn geïnteresseerd in de lengte van het tussentijdsinterval waar t in valt. Dat is het tijdsverschil tussen de laatste gebeurtenis vóór t en de eerstvolgende gebeurtenis ná t. De *wachttijd-paradox* zegt onder andere dat de lengte van deze tussentijd niet $Exp(\lambda)$ verdeeld is, zoals je misschien zou verwachten. In dit onderdeel gaan we deze paradox d.m.v. computersimulaties illustreren.

Allereerst geven we nog even de kansverdelingen die hierbij een rol spelen. Het tijdsinterval tussen de laatste gebeurtenis voor t en t zelf noemen we U_t. Het tijdsinterval tussen t en de eerstvolgende gebeurtenis na t noemen we V_t. De totale lengte van het tussentijdsinterval waar t in ligt noemen we L_t. Merk op dat $L_t = U_t + V_t$. Zie afbeelding 3 in de bijlage.

Het blijkt dat V_t een $Exp(\lambda)$-verdeling bezit. Dit betekent dat we het nulpunt van het proces net zo goed kunnen kiezen op een plaats waar het Poisson-proces al aan de gang is. Het resulterende proces is toch weer een Poisson-proces, want de wachttijd tot de eerste gebeurtenis blijft $Exp(\lambda)$ verdeeld.

$$f_{V_t}(v) = \begin{cases} \lambda e^{-\lambda v} & \text{als } v > 0 \\ 0 & \text{als } v \leq 0 \end{cases}$$

De kansverdeling van U_t is gedeeltelijk continu en gedeeltelijk discreet. We kunnen een Poisson-proces “omkeren”, waardoor we voor U_t een $Exp(\lambda)$-verdeling verwachten. Dit is slechts gedeeltelijk het geval. Omdat het Poisson-proces een *begin* heeft, namelijk het tijdstip 0, kan U_t niet groter dan t worden. U_t is daarom met positieve kans gelijk aan t. De verdeling die U_t uiteindelijk heeft heet een *afgeknotte exponentiële verdeling*.

$$f_{U_t}(u) = \begin{cases} \lambda e^{-\lambda u} & \text{als } 0 < u < t \\ 0 & \text{als } u \leq 0, u > t \end{cases}$$

$$P(U_t = t) = e^{-\lambda t}$$

Tot slot blijkt L_t een absoluut-continue verdeling te bezitten.

$$f_{L_t}(x) = \begin{cases} \lambda^2 x e^{-\lambda x} & \text{als } x \leq t \\ \lambda(1 + \lambda t) e^{-\lambda x} & \text{als } x > t \end{cases}$$

We willen nu de wachttijd-paradox illustreren d.m.v. computersimulaties. Daartoe gaan we, net als in het vorige onderdeel, realisaties uit een exponentiële verdeling door *Maple* laten uitvoeren. We kiezen $\lambda = 1/3$ en kiezen als vastgelegd tijdstip $t = 4$. We stoppen zodra het totaal van alle realisaties $t = 4$ heeft overschreden. Als dat gebeurt is, beschikken we over een realisatie van V_4, van U_4 en van L_4.
Op deze manier zijn er 20 realisaties uitgevoerd die te vinden zijn in tabel 4. Deze realisaties zijn onderling onafhankelijk waardoor je kunt zeggen dat we over steekproeven beschikken uit de verdelingen van V_4, U_4 en L_4. Om nu uitspraken te doen over de verdelingen van deze stochasten, maken we onder andere gebruik van de schatter voor λ op basis van het steekproefgemiddelde, die in het eerste onderdeel besproken is. De volgende punten illustreren de wachttijdparadox en laten zien dat de resultaten van de simulaties inderdaad duiden op de gegeven verdelingen voor V_4 en U_4.

1. Met de 20 realisaties van L_4 kunnen we laten zien dat voor deze stochast niet echt een $Exp(1/3)$-verdeling in aanmerking komt. Immers, de schatting voor λ uit het eerste onderdeel (1/Steekproefgemiddelde) levert 0,215, hetgeen wat aan de lage kant is.

2. V_4 lijkt wel een $Exp(1/3)$-verdeling te hebben. Want dezelfde schatting voor λ die we ook voor L_4 hebben uitgeprobeerd levert hier 0,391, wat redelijk in de buurt van 1/3 ligt. Verder valt het op dat er veel lage waarden (onder de 3) voor V_4 gegeven zijn. Dit duidt op "veel dichtheid" voor lage waarden. We zien ook dat andere waarden minder voorkomen naarmate ze hoger zijn, wat duidt op een "daling van de dichtheidsfunctie". Om deze redenen lijkt V_4 redelijk exponentieel verdeeld.

3. Bij de realisaties van U_4 valt meteen op dat er een aantal gelijk aan $t = 4$ zijn. De fractie waarvoor dit geldt is 6/20 wat aardig in de buurt van het getal exp(-4/3) ligt, hetgeen $P(U_4 = 4)$ zou moeten zijn. Verder zien we, als we naar de overige waarden kijken, dat er ook hier veel "lage" waarden zijn en dat het aantal daalt naarmate de waarden groter worden. Een exponentiële verdeling voor deze overige waarden lijkt dus om dezelfde reden als bij V_3 voor de hand liggend. En daarmee lijkt de verdeling van U_4 op een afgeknotte exponentiële verdeling.
   Het mag overigens duidelijk zijn dat we, vanwege de afknotting, in dit geval niet zo eenvoudig de parameter λ van deze verdeling kunnen schatten. We laten dat dan ook zitten.

We hebben nu laten zien dat de gegeven verdelingen redelijk overeenstemmen met de computersimulaties en hebben de wachttijd-paradox hopelijk voldoende geïllustreerd.

3. Poisson-experiment

Het derde onderdeel van dit verslag gaat niet over computersimulaties, maar over een experiment uit de dagelijkse praktijk. Ik heb op zaterdag 1 februari langs een doorgaand fietspad in Maarssen tijdstippen genoteerd waarop er fietsers voorbij kwamen. Dit een half uur lang van elf uur tot half twaalf 's morgens. Ondanks de hevige sneeuwval kwamen er toch flink wat fietsers voorbij, en heeft het uiteindelijk 55 metingen opgeleverd. Het doel is nu om na te gaan in hoeverre dit proces te beschrijven is d.m.v. een Poisson-proces, waarbij we het voorbijgaan van een fietser als een *gebeurtenis* beschouwen. We spreken de volgende punten af:

- Groepen fietsers (zoals gezinnen en groepjes kinderen) beschouwen we als één. Fietsers in zo'n groep rijden namelijk niet "onafhankelijk van elkaar" voorbij, maar rijden juist met opzet vlak achter of naast elkaar, waardoor een belangrijk kenmerk van het Poisson-proces al niet meer op zou gaan: de onvoorspelbaarheid.

- De fietsers van beide richtingen worden meegeteld.

- Naast fietsers worden ook mensen op andere rijtuigen zoals scooters meegerekend. Deze noemen we voor het gemak ook gewoon fietsers. Voetgangers worden niet meegerekend.

- Als tijdseenheid kiezen we 100 seconden , dat rekent namelijk het makkelijkst. Het tijdsinterval waarover ik heb gemeten is dus [0,18], immers, ik heb 30 minuten lang gemeten.

- De waarnemingen zijn in eerste instantie de tijdstippen waarop de fietsers voorbijkwamen. Echter, we willen de waarnemingen net zoals in het eerste onderdeel kunnen opvatten als realisaties uit een exponentiële verdeling. De tijdstippen van de gebeurtenissen zijn daarom niet opgenomen in dit verslag, maar wel alle tussentijden. Daartoe zijn aan de hand van de waargenomen tijdstippen de tussentijden in seconden berekend (rekening houdend met het feit dat er 60 seconden in een minuut zitten). Vervolgens zijn deze aantallen seconden steeds gedeeld door 100, zodat de tussentijden in de afgesproken tijdseenheid worden gegeven. (Overigens moeten we nog maar zien of de tussentijden inderdaad exponentieel verdeeld zijn.)

We beschikken dus over de tussentijden in seconden/100. Deze tussentijden beschouwen we vanaf nu als de waarnemingen en zijn te vinden in tabel 5, het zijn er 55.

Nu is de vraag in hoeverre dit proces overeenkomt met het Poisson-proces. Om daarachter te komen gaan we eerst kijken wat de parameter λ dan zou moeten zijn. Hierbij gebruiken we ook de schatters die in het eerste onderdeel besproken zijn. Allereerst schatten we λ op basis van het steekproefgemiddelde van de 55 waarnemingen die $Exp(\lambda)$ verdeeld zouden moeten zijn. Dit leidt tot het volgende resultaat:

$$\hat{\lambda}_1 = \frac{\text{Aantal metingen}}{\text{Totaal van de tussentijden}} = \frac{55}{17,93} \approx 3,067$$

De volgende criteria zullen nu laten zien dat het Poisson-proces inderdaad een redelijke beschrijving geeft.

1. In plaats van naar de verwachting van de tussentijden te kijken kunnen we ook kijken naar de *variantie* van de tussentijden. We weten dat een goede schatter van variantie bij een steekproef $(X_1, X_2, \ldots, X_n)$ uit een bepaalde verdeling gegeven wordt door de *steekproefvariantie*:

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2$$

   Verder weten we dat de variantie van een $Exp(\lambda)$ verdeelde stochast gegeven wordt door $1/\lambda^2$, zodat we de volgende schatting voor λ kunnen maken (de verschrikkelijke rekenpartij zal ik de lezer besparen):

$$\hat{\lambda}_2 = \sqrt{\frac{1}{\text{Steekproefvariantie}}} \approx 3,184$$

   hetgeen redelijk in de buurt van onze eerste schatting komt.

2. Door het totale tijdsinterval te verdelen in 18 disjuncte deelintervallen van lengte 1, beschikken we (hopelijk) ook over een realisatie van een steekproef uit de $Pois(\lambda)$-verdeling en kunnen we λ schatten zoals beschreven in onderdeel 1, de tweede manier. Dit levert:

$$\hat{\lambda}_3 = \frac{\text{Totaal aantal waargenomen gebeurtenissen}}{\text{Aantal deelintervallen}} = \frac{55}{18} \approx 3,056$$

   Dit zit heel dicht bij de eerste schatting van λ, maar uit onderdeel 1 weten we al dat dit kan als λ groot is t.o.v. het tijdsinterval.

3. Ook in het geval van de Poisson-verdeling kunnen we ook kijken naar de variantie, welke in het geval van een $Pois(\lambda)$-verdeling precies λ is. We schatten d.m.v. de steekproefvariantie en besparen de lezer ook hier de lange rekenpartij:

$$\hat{\lambda}_4 = \text{Steekproefvariantie} \approx 2,291$$

Deze schatting wijkt enigszins af van de overige schattingen, maar de afwijking is wel beperkt (ongeveer 25% t.o.v. $\hat{\lambda}_1$).

4. Als laatste criterium laten we nog even zien dat de verdeling van een tussentijd inderdaad lijkt op een exponentiële verdeling. We splitsen daartoe alle tussentijdmetingen op al naar gelang ze in een bepaald grootte-interval vallen, en tellen voor ieder van zo'n interval het aantal metingen dat daarin valt. Het resultaat is tabel 6. We zien hier dat veel metingen in het interval 0,01 t/m 0,20 terecht komen, wat duidt op "veel dichtheid" in de buurt van 0. Ook zien we dat het aantal metingen in de hogere zones afneemt, wat duidt op een "daling van de dichtheidsfunctie". In gedachten zien we het dichtheidsgrafiekje van de exponentiële verdeling verschijnen.

Met de bovenstaande criteria is hopelijk voldoende gemotiveerd waarom het Poisson-proces inderdaad een redelijke beschrijving geeft voor het proces van voorbijgaande fietsers over het fietspad.

Tot slot nog een opmerking. Tijdens het experiment zijn de fietsers van beide richtingen meegeteld. Deze hadden echter ook allebei apart genomen kunnen worden, en we konden dan onderzoeken of het allebei Poisson-processen zijn, eventueel met verschillende parameters (wanneer bijvoorbeeld de ene plaats meer bezoekers trekt dan de andere). In dat geval zou het proces dat we nu hebben bekeken opgevat kunnen worden als de "vereniging" van twee Poisson-processen. Daarover gaat het laatste onderdeel van dit verslag.

4. Vereniging van twee processen

In dit laatste onderdeel bewijzen we nog een stelling over Poisson-processen. Stel we kiezen een tijdsinterval $[0, t]$ en laten hierover een Poisson-proces $M(t)$ met intensiteit λ lopen. Tegelijkertijd laten we over dit zelfde interval, onafhankelijk van $M(t)$, nóg een Poisson-proces $N(t)$ met intensiteit μ lopen. Zie afbeelding 7 in de bijlage. We construeren aan de hand van deze twee processen een nieuw proces $O(t)$ door te kijken naar de *vereniging* van de processen $M(t)$ en $N(t)$. Wat we precies bedoelen met deze *vereniging* wordt geïllustreerd in afbeelding 7. We beweren nu dat het proces $O(t)$ ook een Poisson-proces is met intensiteit ν waarvoor geldt: $\nu = \lambda + \mu$. Voordat we dit aantonen formuleren we eerst twee stellingen die we verderop zullen bewijzen.

Stelling 1: Het minimum van twee onafhankelijke exponentieel verdeelde stochasten is weer exponentieel verdeeld, met als parameter de som van de parameters.

Stelling 2: De som van twee onafhankelijke Poisson verdeelde stochasten is weer Poisson verdeeld, met als parameter de som van de parameters.

Om nu te laten zien dat $O(t)$ een Poisson-proces is, bewijzen we eerst dat de tussentijden onafhankelijk en $Exp(\lambda + \mu)$ verdeeld zijn.
Daartoe merken we op dat iedere tussentijd begint op het moment dat er een gebeurtenis in één van beide processen heeft plaatsgevonden. Er zijn nu twee mogelijkheden. De eerste mogelijkheid is dat deze tussentijd eindigt op het moment dat er in hetzelfde proces weer een gebeurtenis plaatsvindt. Het is bekend dat de lengte van het tijdsinterval tussen twee gebeurtenissen in hetzelfde proces een exponentiële verdeling heeft, dus onze tussentijd heeft dat in dit geval ook. De tweede mogelijkheid is dat de tussentijd eindigt op het moment dat er in het andere proces een gebeurtenis plaatsvindt. Vanwege de wachttijd-paradox geldt dat de tijd tussen een vast gekozen tijdstip (in dit geval het beginpunt van de tussentijd) en de eerstvolgende gebeurtenis exponentieel verdeeld is. Onze tussentijd is dat in dit geval dus ook.
We kunnen dus concluderen dat onze tussentijd het minimum is van twee onafhankelijke exponentieel verdeelde stochasten, met parameters λ en μ. Vanwege stelling 1 is de tussentijd dus $Exp(\lambda + \mu)$ verdeeld.
Verder merken op dat de tussentijden in $O(t)$ onderling onafhankelijk zijn. Dit volgt direct uit het feit dat de tussentijden in $M(t)$ en $N(t)$ allemaal onderling onafhankelijk zijn.

Een andere belangrijke eigenschap van $O(t)$ die we willen aantonen is dat het aantal gebeurtenissen in $[0, t]$ $Pois((\lambda + \mu)t)$ verdeeld is. Uit de theorie weten we dat dit al volgt uit het feit dat de tussentijden onafhankelijk en $Exp(\lambda + \mu)$ verdeeld zijn. We kunnen dit echter ook rechtstreeks bewijzen door gebruik te maken van stelling 2. Het aantal gebeurtenissen dat in $O(t)$ heeft plaatsgevonden op tijdstip t is immers gelijk aan de som van deze aantallen in $M(t)$ en $N(t)$, die resp. $Pois(\lambda t)$ en $Pois(\mu t)$ verdeeld zijn..

We hebben nu laten zijn dat $O(t)$ een Poisson-proces is met intensiteit $\nu = \lambda + \mu$, en hebben daarbij gebruik gemaakt van twee stellingen. Het resteert nog deze twee stellingen te bewijzen.

Bewijs van Stelling 1:

Zij $X \sim Exp(\lambda)$:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{als } x > 0 \\ 0 & \text{als } x \leq 0 \end{cases}$$

en $Y \sim Exp(\mu)$:

$$f_Y(y) = \begin{cases} \mu e^{-\mu y} & \text{als } y > 0 \\ 0 & \text{als } y \leq 0 \end{cases}$$

We zullen echter gebruik maken van de verdelingsfuncties van X en Y die eenvoudig met behulp van de gegeven dichtheden te bepalen zijn (maar die we ook uit het hoofd kennen).

Neem aan: X en Y zijn onderling onafhankelijk.

Definieer: $Z = \min\{X, Y\}$, dan geldt:

$$P(Z < z) = P(X < z \text{ of } Y < z) = P(X < z) + P(Y < z) - P(X < z \text{ én } Y < z) =$$

$$1 - e^{-\lambda z} + 1 - e^{-\mu z} - (1 - e^{-\lambda z})(1 - e^{-\mu z}) = 1 - e^{-(\lambda+\mu)z} \text{ als } z > 0 \text{ en } 0 \text{ als } z \leq 0.$$

Dus:

$$f_Z(z) = \begin{cases} (\lambda + \mu) e^{-(\lambda+\mu)z} & \text{als } z > 0 \\ 0 & \text{als } z \leq 0 \end{cases}$$

$\square$

Bewijs van Stelling 2:

Zij $X \sim Pois(\lambda)$:

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

en $Y \sim Pois(\mu)$:

$$P(Y = y) = e^{-\mu}\frac{\mu^y}{y!}, \text{ voor } y = 0, 1, 2, \ldots$$

Neem aan: X en Y zijn onderling onafhankelijk.
Definieer: $Z = X + Y$, en neem een $z \in \{0, 1, 2, \ldots\}$. Dan geldt vanwege de convolutieformule en de binomiaalformule van Newton, dat:

$$P(Z = z) = P(X + Y = z) = \sum_{x=0}^{\infty} P(X = x;\ Y = z - x) =$$

$$\sum_{x=0}^{z} P(X = x)P(Y = z - x) = \sum_{x=0}^{z} \frac{e^{-\lambda}\lambda^x}{x!} \cdot \frac{e^{-\mu}\mu^{z-x}}{(z-x)!} =$$

$$e^{-(\lambda+\mu)} \sum_{x=0}^{z} \frac{\lambda^x \mu^{z-x}}{x!(z-x)!} = \frac{e^{-(\lambda+\mu)}}{z!} \sum_{x=0}^{z} \binom{z}{x} \lambda^x \mu^{z-x} = \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^z}{z!}$$

$\square$

Bijlage