Der Poissonprozess

Kim Boström

14. Oktober 2005

1 Prozess ohne Gedächtnis

Nehmen wir einen beliebigen Prozess an, der Ereignisse in der Zeit erzeugt, welche statistisch unabhängig voneinander sind. Dies können die Zerfälle einer radioaktiven Substanz, das Platzen von Popkorn in einer Pfanne, die geworfenen Sechsen eines Würfels, die Ankunft von Buslinien an einer Haltestelle, oder die Aktionspotentiale einer neuronalen Verbindung sein. Statistische Unabhängigkeit bedeutet, dass jedes Ereignis unabhängig von den vorherigen Ereignissen mit der gleichen Wahrscheinlichkeit eintritt, man spricht auch von einem *gedächtnislosen Prozess*. Statistische Unabhängigkeit ist z.B. immer dann gewährleistet, wenn die Einzelereignisse voneinander unabhängige Ursachen haben.

Denken wir uns einen diskreten Zeitschritt Δt und sei dann q die Wahrscheinlichkeit dass ein bestimmtes Ereignis innerhalb von Δt genau einmal stattfindet. Δt muss also genügend klein sein, damit das Ereignis darin wirklich nur einmal stattfindet, es handelt sich bei Δt also um eine Art *Regenerationsdauer* unterhalb derer das Ereignis nicht erneut stattfinden kann.

Offenbar ist $p = 1 - q$ die Wahrscheinlichkeit dass dieses Ereigniss in Δt *nicht* stattfindet. Folglich ist

$$P(n\Delta t) = p^n = (1-q)^n \tag{1}$$

die Wahrscheinlichkeit dass innerhalb von n Zeitschritten der Länge Δt, also innerhalb der Zeitspanne $t = n\Delta t$, kein Ereignis stattfindet. Umgeschrieben finden wir

$$n = \frac{t}{\Delta t}, \tag{2}$$

welches in (1) eingesetzt zu der Formel

$$P(t) = (1-q)^{t/\Delta t} \tag{3}$$

führt. Da q nun die Wahrscheinlichkeit für genau ein Ereignis innerhalb der Zeitspanne Δt ist, ist die Anzahl von Ereignissen innerhalb der Zeitspanne $t = n\Delta t$ typischerweise, d.h.

im Grenzfalle $n \to \infty$, gegeben durch $N(t) = nq$. Folglich ist die *Ereignisrate* λ, also die mittlere Anzahl von Ereignissen pro verstrichener Zeit, gegeben durch

$$\lambda = \frac{N(t)}{t} = \frac{nq}{n\Delta t} = \frac{q}{\Delta t}. \tag{4}$$

Eine Rate hat die Dimension "Ereignisse pro Zeit" und falls sie sich auf das Eintreten eines einzelnen Ereignisses bezieht (was im Falle $t = \Delta t$ ja gegeben ist), dann wird daraus "Wahrscheinlichkeit pro Zeit". Setzen wir (4) in (1) ein, so ergibt sich

$$P(t) = (1 - \lambda\Delta t)^{t/\Delta t}. \tag{5}$$

Nehmen wir nun an, dass die Regenerationsdauer Δt unendlich klein ist, dass also der Prozess sich nach jedem eingetroffenen Ereignis sofort wieder erholt und ein weiteres Ereignis stattfinden kann. Dann ergibt sich im Grenzwert $\Delta t \to 0$, wie man durch Vergleich mit der bekannten Formel

$$\lim_{N\to\infty} \left(1 - \frac{x}{N}\right)^N = e^x \tag{6}$$

und Setzen von $N = 1/\Delta t$ erkennen kann, das exponentielle Gesetz

$$P(t) = e^{-\lambda t}, \tag{7}$$

mit $P(t)$ als *Nichteintrittswahrscheinlichkeit* und λ als *Ereignisrate*. Im Falle von radioaktiven Atomen nennt man $P(t)$ die Überlebenswahrscheinlichkeit und λ die Zerfallsrate. Wie wir sehen gilt der exponentielle Zusammenhang aber ganz allgemein für die Statistik von zeitlich voneinander unabhängigen Ereignissen. $P(t)$ ist die Wahrscheinlichkeit für das *Nichteintreten* eines Ereignisses innerhalb der Zeitspanne t. Nach Konstruktion ist daher $t \geq 0$.

2 Wartezeiten: Die Exponentialverteilung

Jetzt ist es natürlich besonders interessant nach der mittleren *Wartezeit* zwischen den Ereignissen zu fragen. Gerade wenn man an der Bushaltestelle steht, interessiert einen diese Zahl ganz besonders. Sei X die Zufallsgröße für die Wartezeit, d.h. X hat den Wertebereich $[0, \infty)$ und eine bestimmte Wahrscheinlichkeitsdichte $\rho_X(t)$. Offenbar ist die Wahrscheinlichkeit dafür dass X einen Wert $> t$ hat gleich der Überlebenswahrscheinlichkeit $P(t)$, denn die gibt ja gerade die Wahrscheinlichkeit an, dass das Ereignis bis zum Zeitpunkt t noch nicht eingetroffen ist, also

$$P(X > t) = P(t) = e^{-\lambda t}. \tag{8}$$

Daher ist die *komplementäre* Wahrscheinlichkeit dass nämlich das Ereignis bis t eingetroffen ist, gegeben durch

$$\begin{aligned} P(X \leq t) &= 1 - P(X > t) && (9) \\ &= 1 - P(t) && (10) \\ &= 1 - e^{-\lambda t}. && (11) \end{aligned}$$

Diese Wahrscheinlichkeit $P(X \leq 0)$ ist eine *kumulative Wahrscheinlichkeitsverteilung* und hat die Dichte

$$\rho(t) = \frac{dP(X \leq t)}{dt} = \lambda e^{-\lambda t}, \tag{12}$$

so dass nämlich

$$P(X \leq t) = \int_0^t dt'\, \rho(t'). \tag{13}$$

Wir haben also mit $\rho(t)$ die Wahrscheinlichkeitsdichte der Wartezeit X gefunden. Die *mittlere Wartezeit* ist folglich gegeben durch

$$\begin{aligned} \tau = \langle X \rangle &= \int_0^\infty dt\, t\, \rho(t) && (14) \\ &= \lambda \int_0^\infty dt\, t\, e^{-\lambda t} && (15) \\ &= \frac{1}{\lambda} \int_0^\infty du\, u\, e^{-u} = \frac{1}{\lambda}\Gamma(2) && (16) \\ &= \frac{1}{\lambda}, && (17) \end{aligned}$$

wobei wir die Darstellung

$$\Gamma(x) = \int_0^\infty dt\, t^{x-1}\, e^{-t} \tag{18}$$

für die Gammafunktion verwendet haben, für die $\Gamma(n) = (n-1)!$ für natürliche Zahlen n erfüllt ist. Also erfüllen die mittlere Wartezeit τ und die Ereignisrate λ die Relation

$$\tau\lambda = 1, \tag{19}$$

was ja doch auch mehr als verständlich ist. Schließlich sagt mir λ ja gerade, wieviele Ereignisse pro Zeiteinheit eintreten und also ist der Kehrwert davon die mittlere Wartezeit zwischen den Ereignissen. Im Falle eines regelmäßigen Prozesses wäre λ die Frequenz und τ das feste Zeitintervall. Der Poissonprozess ist also die Verallgemeinerung eines regelmäßigen Prozesses zu einem maximal unregelmäßigen Prozess, da nun die Einzelereignisse statistisch ganz voneinander unabhängig sind.

Zusammenfassend: **Die Wartezeit eines Poissonprozesses ist exponentiell verteilt**. Es ist exponentiell unwahrscheinlich, dass zwei aufeinanderfolgende Ereignisse weit auseinanderliegen.

3 Mehrfachereignisse: Die Gammaverteilung

Nun interessieren wir uns für die Wahrscheinlichkeit dass in einem gegebenen Zeitintervall der Länge t eine gewisse Anzahl k an Ereignissen stattfindet. Die Wartezeit T_k bis zum Auftreten des k-ten Ereignisses ist natürlich die Summe der einzelnen Wartezeiten,

$$T_k = X_1 + \ldots + X_k. \tag{20}$$

Wie lautet nun die Verteilung der Zufallsgröße T_k? Jedes der X_i hat die Verteilung $\rho(t) = \lambda e^{-\lambda t}$, wie wir im vorherigen Abschnitt herausgefunden haben. Aus der Theorie wissen wir, dass die Summe Z zweier Zufallsgrößen X und Y so verteilt ist, dass ihre Dichte gerade die *Faltung* der beiden Dichten von X und Y ist. Genauer: Wenn $\rho_X(x)$ und $\rho_Y(y)$ die auf ganz $\mathbb{R}$ definierten Dichten der Zufallsgrößen X und Y sind, dann ist die Dichte der Zufallsgröße

$$Z = X + Y \tag{21}$$

gegeben durch

$$\rho_Z(z) = (\rho_X * \rho_Y)(z) \equiv \int dx\, \rho_X(x)\rho_Y(z-x). \tag{22}$$

Im vorliegenden Falle handelt es sich bei T_k um die k-fache Summe derselben Zufallsgröße X. Damit ist die resultierende Dichte $\rho_k(t)$ gegeben durch die k-fache Faltung der Ursprungsdichte $\rho(t)$ mit sich selbst,

$$\rho_k(t) = \rho^{*k}(t). \tag{23}$$

Zur Berechnung dieser vielen Faltungen müssen wir $\rho(t)$, das ja nur für $t \geq 0$ definiert ist, auf ganz $\mathbb{R}$ erweitern. Wir ergänzen es durch die Heaviside-Funktion $\theta(t)$, die ja auf der negativen Halbachse verschwindet: $\rho(t) = \theta(t)\lambda e^{-\lambda t}$. Die zweifache Faltung von $\rho(t)$ mit sich selbst ergibt

$$\begin{aligned}
\rho^{*2}(t) &= (\rho * \rho)(t) = \int dt'\, \rho(t')\rho(t-t') && (24)\\
&= \int dt'\, \theta(t')\theta(t-t')\lambda e^{-\lambda t'}\lambda e^{-\lambda(t-t')} && (25)\\
&= \lambda^2 e^{-\lambda t}\int dt'\, \theta(t')\theta(t-t') && (26)\\
&= \lambda^2 e^{-\lambda t}\, t\, \theta(t) && (27)
\end{aligned}$$

Die nächste Faltung ergibt

$$\begin{aligned}
\rho^{*3}(t) &= (\rho^{*2} * \rho)(t) = \int dt'\, \rho^{*2}(t')\rho(t-t') && (28)\\
&= \int dt'\, \theta(t')\theta(t-t')\lambda^2 e^{-\lambda t}\, t\, \lambda e^{-\lambda(t-t')} && (29)\\
&= \lambda^3 e^{-\lambda t}\int dt'\, \theta(t')\theta(t-t')\, t' && (30)\\
&= \lambda^3 e^{-\lambda t}\frac{1}{2}\, t^2\, \theta(t). && (31)
\end{aligned}$$

Nach k-facher Anwendung der Faltung ergibt sich demnach

$$\rho^{*k}(t) = t^{k-1}\frac{\lambda^k e^{-\lambda t}}{(k-1)!}\theta(t). \tag{32}$$

Wir schließen also: Das Zeitintervall T_k für das Auftreten von genau k Ereignissen ist verteilt mit der Dichte

$$\rho_k(t) = t^{k-1}\frac{\lambda^k e^{-\lambda t}}{(k-1)!}, \tag{33}$$

wobei wir nun wieder implizit $t \geq 0$ angenommen haben. Nun fragen wir wieder nach der *kumulativen Verteilung*, also der Wahrscheinlichkeit dass bis zum Zeitpunkt t genau k Ereignisse aufgetreten sind. Diese beträgt

$$\begin{aligned} P(T_k \leq t) &= \int_0^t dt'\, \rho_k(t') && (34) \\ &= \int_0^t dt'\, t'^{k-1}\frac{\lambda^k e^{-\lambda t'}}{(k-1)!}. && (35) \end{aligned}$$

Benutzen wir die Darstellung

$$\gamma(x,t) = \int_0^t ds\, s^{x-1}e^{-s} \tag{36}$$

für die sogenannte *unvollständige Gammafunktion*, aus welcher sich die Gammafunktion durch

$$\Gamma(x) = \lim_{t\to\infty}\gamma(x,t) \tag{37}$$

ergibt, so erhalten wir für die kumulative Wahrscheinlichkeitsverteilung von T_k den Audruck

$$P(T_k \leq t) = \frac{\gamma(k,\lambda t)}{\Gamma(k)}. \tag{38}$$

Angesichts dieser vielen Gammas können wir berechtigterweise und zusammenfassend sagen: **Die Wartezeit von Mehrfachereignissen ist gammaverteilt.**. Für $k = 1$ ergibt sich

$$P(T_1 \leq t) = \frac{\gamma(1,\lambda t)}{\Gamma(1)} = \int_0^t dt'\, \lambda\, e^{-\lambda t'} = 1 - e^{-\lambda t}, \tag{39}$$

mit der Dichte

$$\rho(t) = \lambda e^{-\lambda t}. \tag{40}$$

Wir erhalten also für den Spezialfall $k = 1$ wieder das exponentielle Gesetz des vorherigen Abschnitts.

Schließen wir diesen Abschnitt mit einem kleinen lustigen Detail: Die Zufallsgröße

$$\hat{\tau} := \frac{T_k}{k} \tag{41}$$

ist ein unverzerrter Schätzer für die mittlere Wartezeit τ, denn

$$\begin{aligned}
\langle\hat{\tau}\rangle &= \langle\frac{T_k}{k}\rangle = \langle\frac{T_1 + \ldots + T_k}{k}\rangle && (42)\\
&= \frac{\langle T_1\rangle + \ldots + \langle T_k\rangle}{k} && (43)\\
&= \frac{k \cdot 1/\lambda}{k} = \frac{1}{\lambda} && (44)\\
&= \tau, && (45)
\end{aligned}$$

und

$$\Delta^2\hat{\tau} = \frac{1}{k\lambda^2} \to 0 \tag{46}$$

für $k \to \infty$. Wegen $T_k = X_1 + \ldots + T_k$ entspricht T_k/k ganz einfach dem *arithmetischen Mittelwert* von k Einzelmessungen der Wartezeit. Schließlich würden wir ja auch in einem realen Experiment die mittlere Wartezeit abschätzen durch

$$\overline{\tau} = \frac{\text{Wartezeit bis } k \text{ Ereignisse eintreten}}{k}, \tag{47}$$

welches sich für hohe k beliebig gut der tatsächlichen mittleren Wartezeit τ nähert.

4 Ereignishäufigkeit: Die Poissonverteilung

Leider ist uns bei aller Untersuchung des Poissonprozesses eines noch nicht begegnet: die Poissonverteilung selbst. Wo kommt die nun her?

Lösen wir uns von der Wartezeit und wenden wir uns der *Häufigkeit* von Ereignissen innerhalb eines gegebenen Zeitintervalls t zu. Dazu interpretieren wir die Funktion $P(T_k \leq t)$ um: Wenn die Wartezeit für k Ereignisse unterhalb von t liegt, dann treten ganz offenbar *mindestens* k Ereignisse innerhalb von t ein. Stellt N_t die Anzahl der eingetretenen Ereignisse innerhalb der Zeitspanne t dar, so gilt also die Äquivalenz

$$T_k \leq t \quad \Leftrightarrow \quad N_t \geq k, \tag{48}$$

und mithin folgt ganz unmittelbar

$$P(N_t \geq k) = P(T_k \leq t). \tag{49}$$

Nun können wir ganz bequem zur Beschreibung der Zufallsgröße N_t übergehen. Aus (49) und (38) schließen wir

$$P(N_t \geq k) = \frac{\gamma(k, \lambda t)}{\Gamma(k)}. \tag{50}$$

Für natürliche Zahlen k hat die unvollständige Gammafunktion eine Reihendarstellung, nämlich

$$\gamma(k, \lambda t) = \Gamma(k)\left(1 - \sum_{l=0}^{k-1} \frac{e^{-\lambda t}(\lambda t)^l}{l!}\right), \tag{51}$$

und somit finden wir

$$P(N_t \geq k) = 1 - \sum_{l=0}^{k-1} \frac{e^{-\lambda t}(\lambda t)^l}{l!}. \tag{52}$$

Die komplementäre Wahrscheinlichkeit dass nämlich *weniger* als k Ereignisse innerhalb von t stattfinden, beläuft sich dann auf

$$\begin{aligned} P(N_t < k) &= 1 - P(N_t \geq k) && (53) \\ &= \sum_{l=0}^{k-1} \frac{e^{-\lambda t}(\lambda t)^l}{l!}. && (54) \end{aligned}$$

Ersetzen wir k durch $k+1$ dann ist $P(N_t < k+1) = P(N_t \leq k)$ welches wir auch umbenennen in

$$P_k(t) \equiv P(N_t \leq k), \tag{55}$$

und also

$$P_k(t) = \sum_{l=0}^{k} \frac{e^{-\lambda t}(\lambda t)^l}{l!}. \tag{56}$$

Dies ist wieder eine *kumulative* Wahrscheinlichkeit, diesmal aber für die diskrete Zufallsgröße N_t. Die zugehörige Wahrscheinlichkeitsdichte

$$p_k(t) \equiv P(N_t = k) \tag{57}$$

finden wir durch Differenzenbildung an der Stelle k,

$$\begin{aligned} p_k(t) &= P_k(t) - P_{k-1}(t) && (58) \\ &= \sum_{l=0}^{k} \frac{e^{-\lambda t}(\lambda t)^l}{l!} - \sum_{l=0}^{k-1} \frac{e^{-\lambda t}(\lambda t)^l}{l!} && (59) \\ &= \frac{e^{-\lambda t}(\lambda t)^k}{k!}. && (60) \end{aligned}$$

Dies ist nun also tatsächlich die lang ersehnte *Poissonverteilung*! Fassen wir also zusammen: **Die Häufigkeit unabhängiger Ereignisse innerhalb eines gegebenen Zeitintervalls ist poissonverteilt**.

Als donnernden Abschluss berechnen wir noch den Erwartungswert der Größe N_t:

$$\begin{aligned} \langle N_t \rangle &= \sum_{k=0}^{\infty} k\, p_k(t) = \sum_{k=0}^{\infty} k\, \frac{e^{-\lambda t}(\lambda t)^k}{k!} && (61) \\ &= \sum_{k=1}^{\infty} k\, \frac{e^{-\lambda t}(\lambda t)^k}{k!} = \sum_{k=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{(k-1)!} && (62) \\ &= \sum_{k=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^{k+1}}{k!} = \lambda t\, e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} && (63) \\ &= \lambda t\, e^{-\lambda t} e^{\lambda t} = \lambda t. && (64) \end{aligned}$$

Also ist in der Tat die mittlere Häufigkeit der Ereignisse pro Zeiteinheit gleich der Rate,

$$\frac{\langle N_t \rangle}{t} = \lambda. \tag{65}$$

Wegen

$$\langle N_t^2 \rangle = \sum_{k=0}^{\infty} k^2 \, p_k(t) = \sum_{k=0}^{\infty} k^2 \, \frac{e^{-\lambda t} (\lambda t)^k}{k!} \tag{66}$$

$$= \sum_{k=1}^{\infty} k \, \frac{e^{-\lambda t} (\lambda t)^k}{(k-1)!} = e^{-\lambda t} t \frac{d}{dt} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)!} \tag{67}$$

$$= e^{-\lambda t} t \frac{d}{dt} \sum_{k=0}^{\infty} \frac{(\lambda t)^{k+1}}{k!} = e^{-\lambda t} t \frac{d}{dt} \left\{ \lambda t \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \right\} \tag{68}$$

$$= e^{-\lambda t} t \frac{d}{dt} \left\{ \lambda t e^{\lambda t} \right\} = \lambda^2 t^2 + \lambda t. \tag{69}$$

ergibt sich für die Varianz

$$\Delta^2 N_t = \langle N_t^2 \rangle - \langle N_t \rangle^2 \tag{70}$$

$$= \lambda^2 t^2 + \lambda t - \lambda^2 t^2 = \lambda t, \tag{71}$$

und also gilt für Poissonprozesse **Der Erwartungswert ist gleich der Varianz**,

$$\langle N_t \rangle = \Delta^2 N_t. \tag{72}$$

Es läßt sich zeigen, dass dies sogar ausschließlich für Poissonprozesse gilt, es handelt sich also um eine *definierende Eigenschaft*.