Kapitel 3

Poisson-Verteilung, Poisson-Approximation

Definition 3.1. *Eine $\mathbb{N}_0$-wertige Zufallsvariable X heißt* ***Poisson-verteilt*** *zum Parameter $\lambda > 0$ ($X \sim \mathrm{Pois}(\lambda)$), wenn für alle $k \in \mathbb{N}_0$*

$$\mathbf{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

gilt.

Bemerkung. *Dass es sich um eine Wahrscheinlichkeitsverteilung handelt, ergibt sich dank der Exponentialreihe:*

$$\sum_{k=0}^{\infty}\mathbf{P}(X = k) = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1.$$

Bevor wir uns damit beschäftigen, in welchen Situationen die Poisson-Verteilung auftritt, beweisen wir eine mathematisch besonders nützliche Eigenschaft.

Lemma 3.2. *Seien X_1, X_2 unabhängig, $X_1 \sim \mathrm{Pois}(\lambda_1)$, $X_2 \sim \mathrm{Pois}(\lambda_2)$. Dann gilt:*

$$X_1 + X_2 \sim \mathrm{Pois}(\lambda_1 + \lambda_2).$$

Die Summe unabhängiger Poisson-verteilter Zufallsvariablen ist also wieder Poisson-verteilt, und der Parameter der Summe ist die Summe der Parameter.

Beweis. $X_1 + X_2$ ist sicher $\mathbb{N}_0$-wertig, und es gilt

$$\begin{aligned}\mathbf{P}(X_1 + X_2 = k) &= \sum_{l=0}^{k}\mathbf{P}(X_1 = l)\mathbf{P}(X_2 = k-l) \overset{(*)}{=} \sum_{l=0}^{k} e^{-\lambda_1}\frac{\lambda_1^l}{l!}e^{-\lambda_2}\frac{\lambda_2^{k-l}}{(k-l)!}\\ &= e^{-\lambda_1-\lambda_2}\frac{1}{k!}\sum_{l=0}^{k}\frac{k!}{l!(k-l)!}\lambda_1^l\lambda_2^{k-l} = e^{-\lambda_1-\lambda_2}\frac{(\lambda_1+\lambda_2)^k}{k!}.\end{aligned}$$

Dabei haben wir an der Stelle $(*)$ ausdrücklich die Unabhängigkeit von X_1 und X_2 benutzt. Ohne diese Zusatzbedingung ist dieser Schritt im allgemeinen nicht zulässig. $\square$

Die praktische Relevanz der Poisson-Verteilung ergibt sich aus folgendem Satz, aus dem folgt, dass viele zufällige Vorgänge in guter Näherung durch Poisson-verteilte Zufallsvariablen beschrieben werden können.

Satz 3.3 (Poissonapproximation)**.** *Seien $X_1, \dots, X_n$ unabhängige $\{0,1\}$-wertige Zufallsvariablen mit*

$$\mathbf{P}(X_i = 1) = 1 - \mathbf{P}(X_i = 0) = p_i.$$

Sei $S = X_1 + \dots + X_n$ und Z Zufallsvariable mit $Z \sim \mathrm{Pois}(\lambda)$ wobei $\lambda = p_1 + \dots + p_n$. Dann gilt für $A \subset \mathbb{N}_0$:

$$|\mathbf{P}(S \in A) - \mathbf{P}(Z \in A)| \leq \sum_{i=1}^{n} p_i^2.$$

Bemerkung. *Falls die $X_1, \dots, X_n$ u.i.v. sind, also $p_i = p$ für ein p, so ist $S \sim \mathrm{Bin}(n,p)$. Für $Z \sim \mathrm{Pois}(\lambda)$ mit $\lambda = np$ gilt wegen $\sum_{i=1}^n p_i^2 = \frac{\lambda^2}{n}$*

$$\sup_{A \subset \mathbb{N}_0} |\mathbf{P}(S \in A) - \mathbf{P}(Z \in A)| \leq \frac{\lambda^2}{n} = np^2.$$

Korollar 3.4. *Sei $0 \leq p(n) \leq 1$ und $0 \leq \lambda$ mit $np(n) \to \lambda$. Sei $S_n \sim \mathrm{Bin}(n, p(n))$ und $Z \sim \mathrm{Pois}(\lambda)$. Dann gilt*

$$\sup_{A \subset \mathbb{N}_0} |\mathbf{P}(S_n \in A) - \mathbf{P}(Z \in A)| \to 0 \quad \textit{für } n \to \infty.$$

Beweis. Sei $Z_n \sim \mathrm{Pois}(\lambda_n)$ mit $\lambda_n = np(n)$.
Zu zeigen ist:

$$\sup_{A \subset \mathbb{N}_0} |\mathbf{P}(Z_n \in A) - \mathbf{P}(Z \in A)| \to 0.$$

Wegen $|x^k - y^k| \leq k|x-y||x^{k-1} + y^{k-1}|$ für beliebige $x, y \in \mathbb{R}$ und $k \in \mathbb{N}$ (benutzt an der Stelle $(*)$) gilt:

$$\begin{aligned}
\sup_{A \subset \mathbb{N}_0} \left| \sum_{k \in A} e^{-\lambda_n} \frac{\lambda_n^k}{k!} - e^{-\lambda} \frac{\lambda^k}{k!} \right| &\leq \sum_{k=0}^{\infty} \left| e^{-\lambda_n} \frac{\lambda_n^k}{k!} - e^{-\lambda} \frac{\lambda^k}{k!} \right| \\
&\leq \sum_{k=0}^{\infty} \left| e^{-\lambda_n} - e^{-\lambda} \right| \frac{\lambda_n^k}{k!} + e^{-\lambda} \sum_{k=0}^{\infty} \left| \frac{\lambda_n^k - \lambda^k}{k!} \right| \\
&\overset{(*)}{\leq} \left| e^{-\lambda_n} - e^{-\lambda} \right| e^{\lambda_n} + e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda_n^{k-1} + \lambda^{k-1}}{k!} |\lambda_n - \lambda| \\
&\to \left| e^{\lambda} - e^{\lambda} \right| \cdot e^{\lambda} + e^{-\lambda} |\lambda - \lambda| \left(e^{\lambda} + e^{\lambda} \right) = 0
\end{aligned}$$

für $n \to \infty$. $\square$

In typischen Anwendungen geht es oft um eine große Zahl $n \gg 1$ von Individuen, Partikeln o.ä., von denen jedes eine geringe Wahrscheinlichkeit p_i hat, in einem gegebenen Zeitraum und unabhängig von allen anderen eine bestimmte Aktion (Anfrage an eine Hotline stellen, ein gegebenes Geschäft aufsuchen, radioaktiv zerfallen, ...) zu initiieren:

$$\text{f. } i = 1, \dots, n : \quad \mathbf{P}(i \text{ initiiert Aktion}) = p_i = 1 - \mathbf{P}(i \text{ initiiert Aktion nicht}).$$

Beschreiben wir also das zufällige Initiieren der Aktion durch i durch die $\{0, 1\}$-wertige Zufallsvariable $X_i \sim \mathrm{Bin}(1, p_i)$ (1 für "i initiiert die fragliche Aktion", 0 für "i initiiert die Aktion nicht"), so ist die Anzahl der initiierten Aktionen (Anrufe, ankommende Kundschaft, Zerfälle, ...) im gegebenen Zeitraum die Summe S_n der X_i, also im speziellen Fall $p_1 = \dots = p_n = p$ nach Beispiel 12 $\mathrm{Bin}(n, p)$-verteilt. Da Binomialkoeffizienten für große n sehr schnell so groß werden, dass sie auch mit modernen Rechnern nicht mehr sinnvoll berechnet werden können, sind wir heilfroh, dass wir die Binomialverteilung mit großen n und kleinen p durch die Poissonverteilung approximieren können.

Beispiel 15 (Radioaktiver Zerfall)**.** *Ein Stück Erz enthalte $0{,}24\,mg$ des Isotops ${}^{238}U$, das entspricht ca. $6 \cdot 10^{17}$ Uran-238-Kernen. Von diesen hat jeder eine Wahrscheinlichkeit von ca. $4{,}92 \cdot 10^{-18}$, innerhalb einer Sekunde radioaktiv zu zerfallen. Wir haben es also mit der Situation $n = 6 \cdot 10^{17}$, $p = 4{,}92 \cdot 10^{-18}$ zu tun, und die Zahl der in einer Sekunde zu erwartenden Zerfälle ist dann $S \sim \mathrm{Bin}(6 \cdot 10^{17}, 4{,}92 \cdot 10^{-18})$. Es ist nicht leicht, die Wahrscheinlichkeit $\mathbf{P}(S \geq 2)$, dass innerhalb einer Sekunde mindestens 2 Zerfälle stattfinden, zu berechnen.*

Die Bemerkung vor dem Korollar besagt nun, dass wir S durch $Z \sim \mathrm{Pois}(\lambda)$ mit $\lambda = np \approx 2{,}95$ ersetzen dürfen und dabei einen Fehler von höchstens

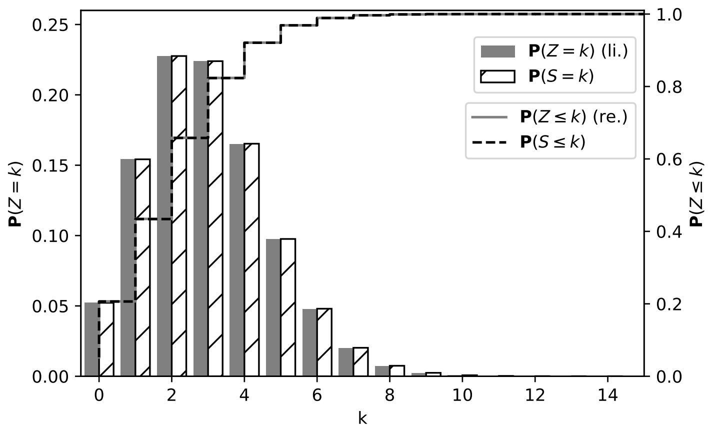
$$|\mathbf{P}(S \geq 2) - \mathbf{P}(Z \geq 2)| \leq \frac{\lambda}{n} = np^2 \approx 1{,}45 \cdot 10^{-17}$$

machen werden.

Wir berechnen also

$$\begin{aligned} \mathbf{P}(Z \geq 2) &= 1 - \mathbf{P}(Z \leq 1) = 1 - \mathbf{P}(Z = 0) - \mathbf{P}(Z = 1) \\ &= 1 - e^{-\lambda} \left(\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} \right) = 1 - e^{-2{,}95} 3{,}95 \approx 0{,}793 \end{aligned}$$

und machen einen Fehler von höchstens $1{,}45 \cdot 10^{-17}$.



Wir bemerken noch, dass wir auch ein größeres Stück Erz mit, sagen wir, 24 mg *Uran-238 hätten betrachten können. Dann hätten wir* $n \approx 6 \cdot 10^{19}$, $\lambda \approx 295$ *und einen Fehler von höchstens* $1,45 \cdot 10^{-15}$ *gehabt. Allerdings wäre dann*

$$\mathbf{P}(Z \leq 1) = e^{-295}(1 + 295) \approx 2,26 \cdot 10^{-126}$$

gewesen. Hier hätten wir dann eher nach der Wahrscheinlichkeit für $Z \geq 200$ *fragen sollen, denn wir hätten (mit Hilfe eines Computers)*

$$1 - \mathbf{P}(Z \geq 200) = \mathbf{P}(Z \leq 199) \approx 1,82 \cdot 10^{-9}$$

erhalten. Das wäre bei einem Fehler von höchstens[1] $1,45 \cdot 10^{-15}$ *schon wieder aussagekräftig.*

Beispiel 16 (Qualitätskontrolle)**.** *Sei*

$$X_i = \begin{cases} 1 & \textit{falls Artikel } i \textit{ fehlerhaft,} \\ 0 & \textit{sonst} \end{cases}$$

mit $\mathbf{P}(X_i = 1) = 0,015$.
Annahme: $X_1, \ldots, X_n$ *sind unabhängig identisch verteilt.*

Packe Paket mit n *Artikeln mit der Forderung*

$$q = \mathbf{P}(\textit{ Paket enthält 100 fehlerfreie Artikel }) \geq 0,9.$$

Frage: Wie groß muss n *sein?*
Lösung: $q = \mathbf{P}(X_1 + \ldots + X_n \leq n - 100) \approx \mathbf{P}(Z \leq n - 100) = \sum_{k=0}^{n-100} e^{-np} \frac{(np)^k}{k!}$, *wo*

[1] wohlgemerkt: für die schlimmstmögliche Menge – für $\{A = Z^{-1}(\{0, 1, \ldots, 199\})\}$ dürfte der tatsächliche Fehler noch deutlich kleiner sein.

$Z \sim \text{Pois}(np)$. Dabei ist nun n nicht fixiert, $p = 0{,}015$ hingegen schon.
Wähle

$$n = \min\{m : \sum_{k=0}^{m-100} e^{-mp} \frac{(mp)^k}{k!} \geq 0{,}9\}.$$

Für die Summe gibt es folgende Ergebnisse:

$$m = 100 : \ 0{,}22; \qquad m = 101 : \ 0{,}55; \qquad m = 102 : 0{,}80; \qquad m = 103 : \ 0{,}93.$$

Der Approximationsfehler ist $\leq np^2 \approx 0{,}023$.

Also ist das Ergebnis ist $n = 103$.

Wir sind noch den Beweis des Satzes schuldig, auf dem unsere bisherigen Überlegungen aufbauten. Dazu benötigen wir noch folgendes Lemma.

Lemma 3.5. *Seien U, V zwei Zufallsvariablen auf dem gleichen Wahrscheinlichkeitsraum. Dann gilt für jede Menge A:*

$$|\mathbf{P}(U \in A) - \mathbf{P}(V \in A)| \leq \mathbf{P}(U \neq V)$$

Man beachte, dass die rechte Seite nur Sinn ergibt, wenn U und V auf dem gleichen Wahrscheinlichkeitsraum definiert sind, die linke Seite hingegen immer.

Beweis. Es gilt:

$$|\mathbf{P}(U \in A) - \mathbf{P}(V \in A)| = |\mathbf{P}(U \in A, U \neq V) - \mathbf{P}(V \in A, U \neq V)| \leq \mathbf{P}(U \neq V).$$

□

Beweis von Satz 3.3. Die Beweisidee beruht auf sogenannten starken Approximationen (coupling technique). Wir konstruieren Zufallsvariablen Z' und S' auf demselben Wahrscheinlichkeitsraum (bisher waren die Wahrscheinlichkeitsräume, auf denen Z und S definiert waren, weder spezifiziert noch im allgemeinen bekannt), wobei Z' dieselbe Verteilung wie Z hat und S' dieselbe Verteilung wie S. Dann werden sich in der Konstruktion Z' und S' mit hoher Wahrscheinlichkeit nur wenig unterscheiden. In unserem Fall genauer: $\mathbf{P}(Z' \neq S')$ ist klein.
Die Zufallsvariablen Z' und S' werden mit Hilfe der $\{-1\} \cup \mathbb{N}_0$-wertigen Zufallsvariablen $U_1, \ldots, U_n$ konstruiert.
Seien $U_1, \ldots, U_n$ unabhängig mit

$$\begin{aligned} \mathbf{P}(U_i = -1) &= e^{-p_i} - (1 - p_i), \\ \mathbf{P}(U_i = 0) &= 1 - p_i, \\ \mathbf{P}(U_i = k) &= e^{-p_i} \frac{p_i^k}{k!} \quad \text{für } k = 1, 2, \ldots \end{aligned}$$

Dies ergibt Sinn wegen

1. $e^{-p_i} - (1-p_i) \geq 0$, da $f(x) = e^x$ konvex ist,
2. $\sum_{j=-1}^{\infty} \mathbf{P}(U_i = j) = 1$.

Nun definieren wir:

$$X_i' = \mathbf{1}_{\{U_i \neq 0\}} = \begin{cases} 1 & \text{falls } U_i \neq 0; \\ 0 & \text{sonst;} \end{cases}$$

$$Z_i' = U_i \mathbf{1}_{\{U_i \geq 0\}} = \begin{cases} U_i & \text{falls } U_i \geq 1; \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt: $\mathcal{L}(X_i') = \mathcal{L}(X_i)$, $\mathcal{L}(Z_i') = \text{Pois}(p_i)$. Also

$$\begin{aligned} S' &= X_1' + \ldots + X_n' \overset{Bsp.12}{\sim} \mathcal{L}(S), \\ Z' &= Z_1' + \ldots + Z_n' \overset{Lemma\ 3.2}{\sim} \text{Pois}(\lambda) = \mathcal{L}(Z) \end{aligned}$$

mit $\lambda = \sum_{i=1}^{n} p_i$.
Die Aussage folgt nun aus Lemma 3.5 und der Ungleichung

$$\mathbf{P}(S' \neq Z') \leq \sum_{i=1}^{n} p_i^2.$$

Zum Beweis dieser Ungleichung zeige zunächst $\mathbf{P}(X_i' \neq Z_i') \leq p_i^2$:

$$\begin{aligned} \mathbf{P}(X_i' \neq Z_i') &= \mathbf{P}(U_i = -1 \text{ oder } U_i \geq 2) \\ &= 1 - \mathbf{P}(U_i = 0) - \mathbf{P}(U_i = 1) \\ &= 1 - (1-p_i) - e^{-p_i} p_i \\ &= p_i(1 - e^{-p_i}) \leq p_i^2 \end{aligned}$$

wegen $e^{-p_i} - (1-p_i) \geq 0$, siehe oben. Hieraus erhält man:

$$\mathbf{P}(S' \neq Z') \leq \mathbf{P}(\exists i : X_i' \neq Z_i') \leq \sum_{i=1}^{n} \mathbf{P}(X_i' \neq Z_i') \leq \sum_{i=1}^{n} p_i^2.$$

□