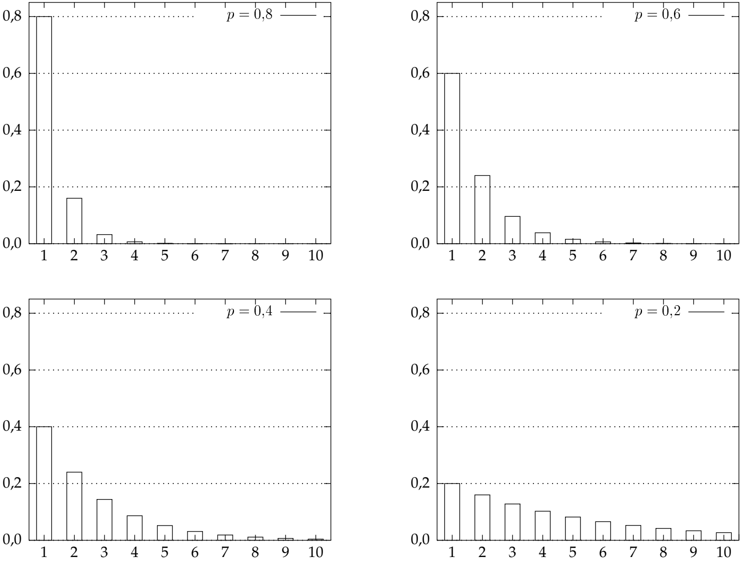


Dichte der geometrischen Verteilung

Sei X wieder geometrisch verteilt mit Erfolgswahrscheinlichkeit p. Dann ist $\Pr[X = k]$ die Wahrscheinlichkeit, dass wir bei einem binären Experiment mit Erfolgswahrscheinlichkeit p genau in der k-ten unabhängigen Wiederholung das erste Mal erfolgreich sind.

Wie groß ist die Wahrscheinlichkeit $\Pr[X > y + x \mid X > x]$?

Da bei den ersten x Versuchen kein Erfolg eintrat, stellen wir uns vor, dass das „eigentliche" Experiment erst ab dem $(x + 1)$-ten Versuch beginnt. Die Zeit bis zum ersten Erfolg bei diesem neuen Experiment nennen wir X'. Damit $X > y + x$ gilt, muss $X' > y$ gelten. Es ist intuitiv, dass X' wieder geometrisch verteilt ist mit Erfolgswahrscheinlichkeit p, dass also für $x, y \in \mathbb{N}$ gilt:

$$\Pr[X > y + x \mid X > x] = \Pr[X' > y]. \tag{6}$$

Formal gilt

$$\Pr[X > x] = \sum_{i=x+1}^{\infty} (1-p)^{i-1} p = (1-p)^x p \cdot \sum_{i=0}^{\infty} (1-p)^i$$
$$= (1-p)^x p \cdot \frac{1}{1-(1-p)} = (1-p)^x \,,$$

sowie

$$\begin{aligned}\Pr[X > y + x \mid X > x] &= \frac{\Pr[X > y+x, X > x]}{\Pr[X > x]} \\ &= \frac{\Pr[X > y+x]}{\Pr[X > x]} \\ &= (1-p)^{y+x} \cdot (1-p)^{-x} = (1-p)^y \\ &= \Pr[X > y] \,.\end{aligned}$$

Diese Eigenschaft nennt man *Gedächtnislosigkeit*, da eine geometrisch verteilte Zufallsvariable gewissermaßen vergisst, dass sie schon x Misserfolge hinter sich hat und sich deshalb zum Zeitpunkt $y + x$ genauso verhält wie ursprünglich zur Zeit y.

Warten auf den n-ten Erfolg.

Wir betrachten n unabhängige Zufallsvariablen $X_1, \ldots, X_n$, die jeweils geometrisch verteilt sind mit Parameter p, und bestimmen die Dichte der Zufallsvariablen $Z := X_1 + \cdots + X_n$. Damit bezeichnet Z also die Anzahl der Versuche bis zum n-ten erfolgreichen Experiment (einschließlich).

Falls $Z = z$ ist, so werden also genau n erfolgreiche und $z - n$ nicht erfolgreiche Experimente durchgeführt. Dafür gibt es genau $\binom{z-1}{n-1}$ Möglichkeiten, von denen jede mit Wahrscheinlichkeit $p^n(1-p)^{z-n}$ eintritt. Es gilt also

$$f_Z(z) = \binom{z-1}{n-1} \cdot p^n(1-p)^{z-n}\,.$$

Die Zufallsvariable Z nennt man negativ binomialverteilt mit Ordnung n.

Das Coupon-Collector-Problem

In manchen Branchen legen Firmen den Verpackungen ihrer Produkte oft kleine Bilder oder andere Gegenstände bei, um den Käufer zum Sammeln anzuregen. Wenn es insgesamt n verschiedene solche Beilagen gibt, wie viele Packungen muss man im Mittel erwerben, bis man eine vollständige Sammlung besitzt? Hierbei nehmen wir an, dass bei jedem Kauf jede Beilage mit gleicher Wahrscheinlichkeit auftritt.

Bezeichne

- X die Anzahl der zu tätigenden Käufe und
- Phase i die Schritte vom Erwerb der $(i-1)$-ten Beilage (ausschließlich) bis zum Erwerb der i-ten Beilage (einschließlich).

Sei etwa $n = 4$, und seien die Beilagen mit den Zahlen $1, 2, 3, 4$ identifiziert. Ein Experiment ist z.B.:

$$\underbrace{2}_{1}, \underbrace{2, 1}_{2}, \underbrace{2, 2, 3}_{3}, \underbrace{1, 3, 2, 3, 1, 4}_{4}\ .$$

Beobachtung:

Phase i endet genau dann, wenn wir eine der $n - i + 1$ Beilagen erhalten, die wir noch nicht besitzen.

Somit ist X_i geometrisch verteilt mit Parameter $p = \frac{n-i+1}{n}$ und es gilt $\mathbb{E}[X_i] = \frac{n}{n-i+1}$.

Damit folgt aber sofort

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i=1}^{n} \mathbb{E}[X_i] \\ &= \sum_{i=1}^{n} \frac{n}{n-i+1} \\ &= n \cdot \sum_{i=1}^{n} \frac{1}{i} = n \cdot H_n,\end{aligned}$$

wobei $H_n := \sum_{i=1}^{n} \frac{1}{i}$ die n-te harmonische Zahl bezeichnet. Da $H_n = \ln n + O(1)$, folgt $\mathbb{E}[X] = n \ln n + O(n)$.

5.4 Poisson-Verteilung

Die Poisson-Verteilung kann verwendet werden, um die Anzahl von Ereignissen zu modellieren, welche mit konstanter Rate und unabhängig voneinander in einem Zeitintervall auftreten.

Eine Poisson-verteilte Zufallsvariable X mit dem Parameter $\lambda \geq 0$ hat den Wertebereich $W_X = \mathbb{N}_0$ und besitzt die Dichte

$$f_X(i) = \frac{e^{-\lambda}\lambda^i}{i!} \quad \text{für } i \in \mathbb{N}_0.$$

f_X ist eine zulässige Dichte, da

$$\begin{aligned} \sum_{i=0}^{\infty} f_X(i) &= \sum_{i=0}^{\infty} \frac{e^{-\lambda}\lambda^i}{i!} \\ &= e^{-\lambda} \cdot e^{\lambda} = 1\,. \end{aligned}$$

Für den Erwartungswert erhalten wir

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{i=0}^{\infty} i \cdot \frac{e^{-\lambda}\lambda^i}{i!} \\
&= \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} \\
&= \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^{i}}{i!} \\
&= \lambda e^{-\lambda} e^{\lambda} = \lambda\,.
\end{aligned}
$$

Da

$$
\begin{aligned}
\mathbb{E}[X(X-1)] &= \sum_{i=0}^{\infty} i(i-1) \cdot \frac{e^{-\lambda}\lambda^i}{i!} \\
&= \lambda^2 e^{-\lambda} \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} \\
&= \lambda^2 e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^{i}}{i!} \\
&= \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2
\end{aligned}
$$

und

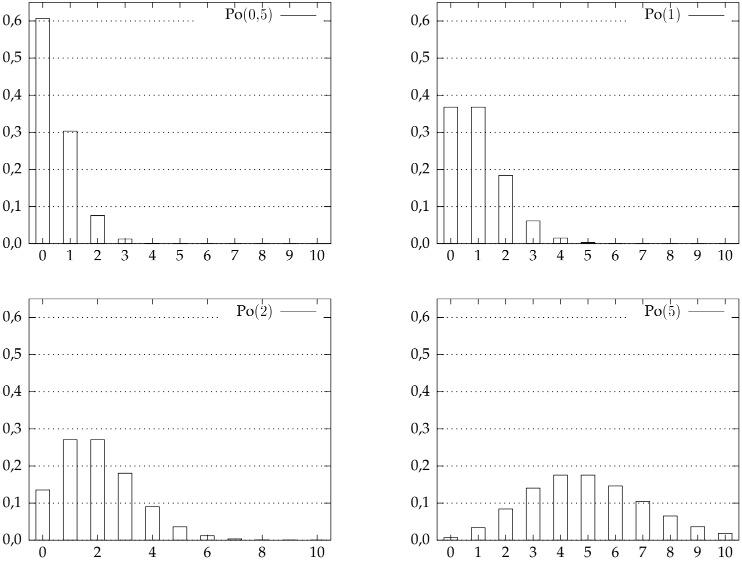
$$
\begin{aligned}
&\mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 \\
&= \mathbb{E}[X^2] - \mathbb{E}[X] + \mathbb{E}[X] - \mathbb{E}[X]^2 = \mathrm{Var}[X]\,,
\end{aligned}
$$

folgt

$$\mathrm{Var}[X] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda. \tag{7}$$

Dafür, dass eine Zufallsvariable X Poisson-verteilt mit Parameter λ ist, schreiben wir auch

$$X \sim \mathrm{Po}(\lambda).$$



Dichte der Poisson-Verteilung

5.4.1 Poisson-Verteilung als Grenzwert der Binomialverteilung

Wir betrachten eine Folge von binomialverteilten Zufallsvariablen X_n mit $X_n \sim \mathrm{Bin}(n, p_n)$, wobei $p_n = \lambda/n$. Für ein beliebiges k mit $0 \leq k \leq n$ ist die Wahrscheinlichkeit, dass X_n den Wert k annimmt, gleich

$$
\begin{aligned}
b(k; n, p_n) &= \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} \\
&= \frac{(n \cdot p_n)^k}{k!} \cdot \frac{n^{\underline{k}}}{n^k} \cdot (1 - p_n)^{-k} \cdot (1 - p_n)^n \\
&= \frac{\lambda^k}{k!} \cdot \frac{n^{\underline{k}}}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^{-k} \cdot \left(1 - \frac{\lambda}{n}\right)^n .
\end{aligned}
$$

Wir betrachten nun $n \to \infty$ und erinnern uns, dass

$$\lim_{n\to\infty} \frac{n^{\underline{k}}}{n^k} = 1,$$

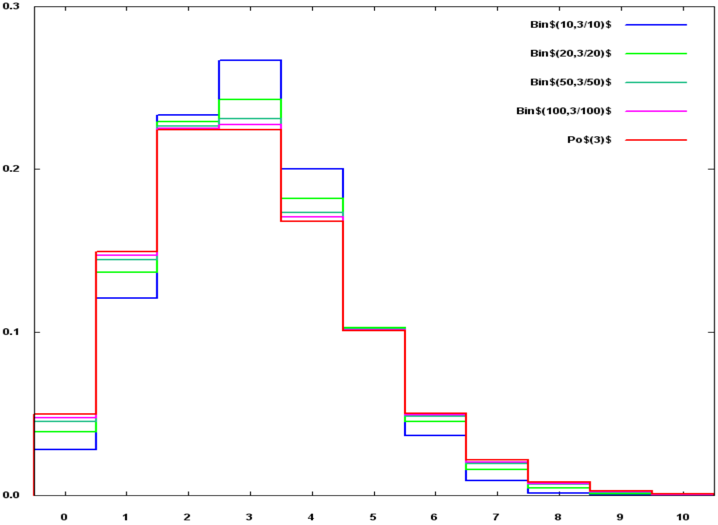
$$\lim_{n\to\infty} (1 - \frac{\lambda}{n})^{-k} = 1, \text{ und}$$

$$\lim_{n\to\infty} (1 - \frac{\lambda}{n})^{n} = e^{-\lambda}.$$

Damit folgt

$$\lim_{n\to\infty} b(k; n, p_n) = \lim_{n\to\infty} \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} = e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

Die Wahrscheinlichkeit $b(k; n, p_n)$ konvergiert also für $n \to \infty$ gegen die Wahrscheinlichkeit, dass eine Poisson-verteilte Zufallsvariable mit Parameter λ den Wert k annimmt. Insgesamt folgt somit, dass die Verteilung einer Zufallsvariablen $X \sim \mathrm{Bin}(n, \lambda/n)$ sich für $n \to \infty$ der Poisson-Verteilung $\mathrm{Po}(\lambda)$ annähert.



Vergleich von Binomial- und Poisson-Verteilung

Ist also n im Vergleich zu λ hinreichend groß, so kann man die Poisson-Verteilung als Approximation der Binomialverteilung verwenden.

Diese Tatsache wird manchmal auch als Gesetz seltener Ereignisse bezeichnet, da die Wahrscheinlichkeit eines einzelnen Treffers $p_n = \lambda/n$ relativ klein sein muss, wenn die Approximation gute Ergebnisse liefern soll.

Die folgenden Voraussetzungen müssen erfüllt sein, damit die Annahme der Poisson-Verteilung gerechtfertigt ist:

- Die Ereignisse treten nie zur gleichen Zeit auf.
- Die Wahrscheinlichkeit, dass ein Ereignis in einem (kleinen) Zeitintervall auftritt, ist proportional zur Länge des Intervalls.
- Die Anzahl der Ereignisse in einem festen Zeitintervall hängt nur von dessen Länge ab, nicht aber von der Lage auf der Zeitachse.
- Wenn man zwei disjunkte Zeitintervalle betrachtet, so sind die Anzahlen der Ereignisse in diesen Zeiträumen voneinander unabhängig.

Beispiel 58

Wir wollen wissen, wie oft eine bestimmte Gegend im Durchschnitt von einer Naturkatastrophe (z.B. Vulkanausbruch) getroffen wird. Aus Statistiken entnehmen wir, dass so ein Ereignis im Mittel 10^{-4}-mal pro Jahr auftritt. Wir interessieren uns nun für die Wahrscheinlichkeit, dass die Region in einem Jahr mehr als einmal von einem solchen Unglück heimgesucht wird.

Die Voraussetzungen scheinen erfüllt zu sein, die Anzahl X der Katastrophen durch eine Poisson-Verteilung mit Parameter $\lambda = 10^{-4}$ zu modellieren.

Damit gilt

$$\begin{aligned}\Pr[X \geq 2] &= 1 - \Pr[X = 0] - \Pr[X = 1] = 1 - e^{-\lambda} - \lambda e^{-\lambda} \\ &\approx 1 - 0{,}999900005 - 0{,}000099990 = 5 \cdot 10^{-9}.\end{aligned}$$

Summe von Poisson-verteilten Zufallsvariablen

Satz 59

Sind X und Y unabhängige Zufallsvariablen mit $X \sim \mathrm{Po}(\lambda)$ und $Y \sim \mathrm{Po}(\mu)$, dann gilt

$$Z := X + Y \sim \mathrm{Po}(\lambda + \mu)\,.$$

Beweis:

$$
\begin{aligned}
f_Z(z) &= \sum_{x=0}^{\infty} f_X(x) \cdot f_Y(z-x) = \sum_{x=0}^{z} \frac{e^{-\lambda}\lambda^x}{x!} \cdot \frac{e^{-\mu}\mu^{z-x}}{(z-x)!} \\
&= e^{-(\lambda+\mu)} \cdot \frac{(\lambda+\mu)^z}{z!} \cdot \sum_{x=0}^{z} \frac{z!}{x!(z-x)!} \left(\frac{\lambda}{\lambda+\mu}\right)^x \left(\frac{\mu}{\lambda+\mu}\right)^{z-x} \\
&= e^{-(\lambda+\mu)} \cdot (\lambda+\mu)^z \frac{1}{z!} \cdot \sum_{x=0}^{z} \binom{z}{x} p^x (1-p)^{z-x},
\end{aligned}
$$

wobei $p := \frac{\lambda}{\lambda+\mu}$.
Da die Summe gleich 1 ist, folgt

$$
f_Z(z) = e^{-(\lambda+\mu)} \cdot (\lambda+\mu)^z \frac{1}{z!}\,.
$$

$\square$

Erläuterungen und Beispiele zur Poisson-Verteilung

- In der Wikipedia finden sich ein paar weitere Details und Beispiele.
- Eine Anwendung der Poisson-Verteilung auf die Fußball-Bundesliga wird in einem Artikel präsentiert, der im Spektrum der Wissenschaft, Heft Juni 2010, erschienen ist.

6. Abschätzen von Wahrscheinlichkeiten

6.1 Die Ungleichungen von Markov und Chebyshev

Satz 60 (Markov-Ungleichung)

Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$

Äquivalent dazu:

$$\Pr[X \geq t \cdot \mathbb{E}[X]] \leq 1/t\,.$$

Beweis:

$$\begin{aligned}
t \cdot \Pr[X \geq t] &= t \cdot \sum_{x \in W_X,\, x \geq t} \Pr[X = x] \\
&\leq \sum_{x \in W_X,\, x \geq t} x \cdot \Pr[X = x] \\
&\leq \sum_{x \in W_X} x \cdot \Pr[X = x] \\
&= \mathbb{E}[X]\,.
\end{aligned}$$

$\square$

Alternativer Beweis:

Es gilt

$$\mathbb{E}[X] = \mathbb{E}[X|X < t] \cdot \Pr[X < t] + \mathbb{E}[X|X \geq t] \cdot \Pr[X \geq t]\,.$$

Wegen $\mathbb{E}[X|X < t] \cdot \Pr[X < t] \geq 0$ und $\mathbb{E}[X|X \geq t] \geq t$ folgt sofort

$$\mathbb{E}[X] \geq t \cdot \Pr[X \geq t]\,.$$

Die Markov-Ungleichung ist nach Andrey Andreyevich Markov (1856–1922) benannt, der an der Universität von St. Petersburg bei Chebyshev studierte und später dort arbeitete. Neben seiner mathematischen Tätigkeit fiel Markov durch heftige Proteste gegen das Zaren-Regime auf, und nur sein Status als vermeintlich harmloser Akademiker schützte ihn vor Repressalien durch die Behörden. Im Jahr 1913 organisierte er parallel zum dreihundertjährigen Geburtstag der Zarenfamilie Romanov eine Feier zum zweihundertjährigen Geburtstag des Gesetzes der großen Zahlen (s.u.).

Die folgende Abschätzung ist nach Pavnuty Lvovich Chebyshev (1821–1894) benannt, der ebenfalls an der Staatl. Universität in St. Petersburg wirkte.

Satz 61 (Chebyshev-Ungleichung)

Sei X eine Zufallsvariable, und sei $t \in \mathbb{R}$ mit $t > 0$. Dann gilt

$$\Pr[|X - \mathbb{E}[X]| \geq t] \;\leq\; \frac{\mathrm{Var}[X]}{t^2}\,.$$

Äquivalent dazu:

$$\Pr[|X - \mathbb{E}[X]| \geq t\sqrt{\mathrm{Var}[X]}] \;\leq\; 1/t^2\,.$$

Beweis:

Wir stellen fest, dass

$$\Pr[|X - \mathbb{E}[X]| \geq t] \;=\; \Pr[(X - \mathbb{E}[X])^2 \geq t^2]\,.$$

Setze

$$Y := (X - \mathbb{E}[X])^2\,.$$

Dann gilt $\mathbb{E}[Y] = \mathrm{Var}[X]$, und damit mit der Markov-Ungleichung:

$$\Pr[|X - \mathbb{E}[X]| \geq t] \;=\; \Pr[Y \geq t^2] \;\leq\; \frac{\mathbb{E}[Y]}{t^2} \;=\; \frac{\mathrm{Var}[X]}{t^2}\,.$$

□