

- In einem bekannten Experiment beobachteten die Physiker Rutherford und Geiger den Zerfall einer radioaktiven Substanz.
- Sie studierten die Emission von  $\alpha$ -Teilchen eines radioaktiven Präparates in  $n = 2608$  Zeitabschnitten von 7.5 Sekunden.
- Die folgende Tabelle gibt die Versuchsergebnisse wieder.
- Hierbei gibt  $k$  die Anzahl der Emissionen wieder,  $r_k$  die relativen Häufigkeiten und  $\pi_\lambda(k)$  die Poisson-Wahrscheinlichkeit zum Parameter  $\lambda = 3.87$ .

# Der Poissonsche Punktprozess

$k$	$r_k$	$\pi_\lambda(k)$
0	0.0219	0.0208
1	0.0778	0.0807
2	0.1469	0.1561
3	0.2013	0.2015
4	0.2040	0.1949
5	0.1564	0.1509
6	0.1047	0.0973
7	0.0533	0.0538
8	0.0173	0.0260
9	0.0103	0.0112
10	0.0038	0.0043
11	0.0015	0.0015
12	0	0.0005
13	0.0004	0.0002
14	0.0004	$4 \times 10^{-5}$

- Wir konstruieren ein mathematisches Modell für auf einer Zeitachse zufällig eintretende Vorkommnisse.
- Die Zeitachse sei  $(0, \infty)$ , und die „Vorkommnisse“ seien einfach zufällige Punkte auf dieser Achse.
- Die Konstruktion eines unterliegenden Wahrscheinlichkeitsraumes ist leider etwas aufwendig und soll hier einfach weggelassen werden (wir glauben hier einfach mal, dass man das kann).
- Ist  $I = (t, t + s]$  ein halboffenes Intervall, so bezeichnen wir mit  $N_I$  die zufällige Anzahl der Punkte in  $I$ .
- $N_I$  ist also eine Zufallsgröße mit Werten in  $\mathbb{N}_0$ .
- Statt  $N_{(0,t]}$  schreiben wir auch einfach  $N_t$ .

An unser Modell stellen wir eine Anzahl von Bedingungen (P1) bis (P5), die für Anwendungen oft nur teilweise realistisch sind.

- (P1) Die Verteilung von  $N_I$  hängt nur von der Länge des Intervalls  $I$  ab. Man bezeichnet das auch als (zeitliche) Homogenität des Punktprozesses.
- (P2) Sind  $I_1, I_2, \dots, I_k$  paarweise disjunkte Intervalle, so sind  $N_{I_1}, N_{I_2}, \dots, N_{I_k}$  unabhängige Zufallsgrößen.
- (P3) Für alle  $I$  (stets mit endlicher Länge) existiert  $EN_I$ .
- (P4) Es existiert ein Intervall  $I$  mit  $P(N_I > 0) > 0$ .

- Für kleine Intervalle ist die Wahrscheinlichkeit dafür, dass überhaupt ein Punkt in diesem Intervall liegt, klein. Es gilt:
- $P(N_I \geq 1) = \sum_{k=1}^{\infty} P(N_I = k) \leq \sum_{k=1}^{\infty} kP(N_I = k) = EN_I$  und demzufolge

$$P(N_{(t, t+\varepsilon]} \geq 1) \leq \lambda\varepsilon \quad \text{für alle } t, \varepsilon \geq 0.$$

- Unsere letzte Forderung besagt, dass sich je zwei Punkte separieren lassen, es also keine Mehrfachpunkte gibt.
- Dazu sei für  $T > 0$

$$D_T(\omega) := \inf_{t, s \leq T} \{|t - s| : |N_t - N_s| \geq 1\}$$

- Dann besagt unsere Forderung (P5):

$$P(D_T \leq \alpha_n) \rightarrow_{n \rightarrow \infty} 0 \quad (\text{P5})$$

für jede Nullfolge  $\alpha_n$  und jedes endliche  $T$ .

## Satz

Sind (P1) bis (P5) erfüllt, so sind für alle  $t, s \geq 0$  die Zufallsgrößen  $N_{(t, t+s]}$  Poisson-verteilt mit Parameter  $\lambda s$ .

## Satz (Zentraler Grenzwertsatz)

Für jedes  $n \in \mathbb{N}$  seien die Zufallsvariablen  $X_1, \dots, X_n$  unabhängig und besitzen dieselbe Verteilung mit  $\mathbb{E}[X_i^2] < \infty$ . Bezeichnet  $\mu := \mathbb{E}X_i$  und  $\sigma^2 := \mathbb{V}X_i > 0$ ,  $1 \leq i \leq n$ , dann gilt für alle  $a, b \in \mathbb{R}$  mit  $a < b$

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}\left(\left\{a \leq \frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n\sigma^2}} \leq b\right\}\right) &= \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx \\ &= \int_a^b \varphi_{0,1}(x) dx. \end{aligned}$$

- Es geht also im Zentralen Grenzwertsatz um die Konvergenz der Verteilung richtig skalierteter Summen von i.i.d. ZVen.
- Es läge sicherlich nahe davon zu sprechen, dass eine Folge von Verteilungen  $P_n$  gegen eine Grenzverteilung  $P_0$  konvergiert, falls

$$P_n(\{x\}) \xrightarrow{n \rightarrow \infty} P_0(\{x\}) \quad \forall x \in \mathbb{R}$$

bzw.

$$F_n(x) := P_n((-\infty, x]) = \sum_{y \leq x} P_n(y) \rightarrow F_0(x)$$

für eine geeignete (Verteilungs-) Funktion  $F_0$  gilt.

- Hierbei ist  $F$  die Verteilungsfunktion einer Zufallsvariablen  $X$ , falls gilt

$$P(X \leq a) = F(a).$$

- Das folgende Beispiel zeigt, dass diese Begriffsbildung nicht das Gewünschte liefert.

## Definition

Eine Folge von Verteilungsfunktionen  $F_n$  auf  $\mathbb{R}$  heißt verteilungskonvergent gegen  $F_0$ , falls  $F_0$  eine Verteilungsfunktion ist, d. h. falls gilt

- a)  $F_0$  ist monoton wachsend;
- b)  $F_0$  ist rechtsseitig stetig;
- c)  $\lim_{x \rightarrow -\infty} F(x) = 0$  und  $\lim_{x \rightarrow \infty} F(x) = 1$

und falls

$$F_n(x) \rightarrow F_0(x)$$

für alle  $x$ , in denen  $F_0$  stetig ist, gilt. Ist  $F_0$  die Verteilungsfunktion einer Wahrscheinlichkeit  $P_0$  auf  $\mathbb{R}$ , so schreiben wir

$$P_n \xrightarrow{\mathcal{D}} P_0.$$

## Satz

Es seien  $(X_n)_n$  Zufallsvariablen mit

$$X_n \rightarrow X_0 \quad \text{in Wahrscheinlichkeit.}$$

Dann konvergiert  $P_{X_n}$ , die Verteilung von  $X_n$ , in Verteilung gegen  $P_{X_0}$ .

## Bemerkung

Die Umkehrung gilt in der Regel nicht, z.B. sei  $X$  eine Zufallsvariable mit

$$P(X = 1) = P(X = -1) = \frac{1}{2}.$$

$(X_n)$  sei eine Folge von Zufallsvariablen mit

$$X_{2n} = X \quad \text{und} \quad X_{2n+1} = -X \quad \forall n \in \mathbb{N}.$$

Da  $P_{X_n} = P_X$  für alle  $n \in \mathbb{N}$  gilt, ist  $X_n$  natürlich verteilungskovngent gegen  $X$ . Andererseits gilt

$$P(|X_{2n+1} - X| \geq 1) = 1 \quad \forall n \in \mathbb{N}.$$

## Satz

Es seien  $P_n$  diskrete Wahrscheinlichkeitsverteilungen über  $\mathbb{R}$  und  $F_0 : \mathbb{R} \rightarrow \mathbb{R}$  differenzierbar, monoton wachsend mit

$$\lim_{x \rightarrow -\infty} F_0(x) = 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} F_0(x) = 1.$$

Gilt dann

$$\lim_{n \rightarrow \infty} \sum_{x \in \mathbb{R}} f(x) P_n(\{x\}) = \int_{-\infty}^{\infty} f(x) F_0'(x) dx,$$

so für alle stetigen Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit existenten Limiten  $\lim_{x \rightarrow \pm\infty} f(x)$ , so ist  $P_n$  verteilungskonvergent und es gilt

$$F_n \rightarrow F_0 \quad \text{in Verteilung.}$$

Mithilfe dieses Satzes können wir den Zentralen Grenzwertsatz

## Korollar

Für jedes  $n \in \mathbb{N}$  seien  $X_1, \dots, X_n$  unabhängige und zum Parameter  $p$  Bernoulli-verteilte Zufallsvariablen. Dann gilt für alle  $a < b$ :

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left\{a \leq \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \leq b\right\}\right) = \int_a^b \varphi_{0,1}(x) dx.$$

## Bemerkungen

- Die Funktion  $x \rightarrow \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$  heißt auch *Gaußsche Glockenkurve*, wegen des glockenförmigen Verlaufs ihres Graphen.
- Die Integrale  $\int_a^b \varphi(x) dx$  sind nicht in geschlossener Form mit Hilfe von Polynomen, rationalen Funktionen, Wurzelausdrücken oder elementaren transzendenten Funktionen (wie  $\sin$ ,  $\cos$ ,  $\exp$ , etc.) darstellbar.
- Es gilt offenbar für  $a < b$

$$\int_a^b \varphi(x) dx = \int_{-\infty}^b \varphi(x) dx - \int_{-\infty}^a \varphi(x) dx = \Phi(b) - \Phi(a),$$

wobei wir  $\Phi(y) := \int_{-\infty}^y \varphi(x) dx$  gesetzt haben.

## Bemerkungen

- *Wie nicht anders zu erwarten ist, gilt*

$$\int_{-\infty}^{\infty} \varphi(x) dx = 1.$$

- *Der Beweis, den man üblicherweise in der Analysis für diese Tatsache gibt, benutzt Polarkoordinaten.*
- *Wir geben hier einen Beweis, der sich darauf stützt, dass wir den Satz von de-Moivre-Laplace schon kennen.*
- *Es gilt für  $a \in \mathbb{R}$ :*

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - np}{\sqrt{npq}} \leq a\right) = \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-x^2/2} dx,$$

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - np}{\sqrt{npq}} \geq a\right) = 1 - \Phi(a).$$

## Korollar

*Erüllt eine Folge von unabhängigen, identisch verteilten Zufallsvariablen die Bedingungen des Zentralen Grenzwertsatzes, so gilt auch das Schwache Gesetz der Großen Zahlen.*