## Stochastischer Prozess

Ein *stochastischer Prozess* ist das Modell eines Zufallsexperiments, das sich in der Zeit entwickelt und eine Folge numerischer Werte erzeugt. *Beispiele:*

(a) Die täglichen Börsenkurse einer Aktie zu einem bestimmten Zeitpunkt

(b) Die Anzahl der während eines Fussballspiels geschossenen Tore

(c) Die Anzahl der Autos, die jede Stunde an einem Messpunkt vorbeifahren.

## Bernoulli-Prozess

Beim *Bernoulli-Prozess* betrachten wir eine Folge unabhängiger, bernoulli-verteilter Zufallsvariablen $X_1$, $X_2$, $X_3$, …; d. h:

$X_i \colon \Omega \to \{1, 0\}$ für $i = 1, 2, 3, \ldots$

mit $P(X_i = 1) = p$ und $P(X_i = 0) = 1 - p$

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | … | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |

Die Anzahl der Zeitschritte $T$ bis (und mit) erstem Erfolg („1“) ist *geometrisch verteilt* mit

$P(T = k) = (1-p)^{k-1}p \quad \left[\mathrm{E}(T) = \frac{1}{p},\ \mathrm{Var}(T) = \frac{1-p}{p^2}\right]$

Betrachten wir eine festes Intervall der Länge $n$, so ist die Anzahl $S$ der darin beobachteten Erfolge („1“) *binomialverteilt* mit

$$P(S = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \left[\mathrm{E}(S) = np,\ \mathrm{Var}(S) = np(1-p)\right]$$

## Poisson-Prozess

Wir betrachten jetzt eine Zufallsvariable $X$, die beschreibt, wie oft sich ein bestimmtes (seltenes) Ereignis innerhalb einer Zeitspanne $\tau$ ereignet, wobei diese Ereignisse mit einer festen mittleren Rate $\lambda$ (Intensität) und zu beliebigen Zeitpunkten auftreten. Die folgenden Bedingungen charakterisieren diese Art von Ereignissen:

1. Die Wahrscheinlichkeit, dass sich zu einem Zeitpunkt mehr als ein Ereignis ereignet, ist vernachlässigbar.

2. Die Wahrscheinlichkeit, ein Ereignis während eines kurzen Zeitraums zu beobachten, ist proportional zur Länge des Zeitraums.

3. Die Wahrscheinlichkeit $p_X(k, \tau)$, $k$ Ereignisse in einem Intervall der Länge $\tau$ zu beobachten, ist in allen Intervallen der Länge $\tau$ gleich. (Zeithomogenität)

4. Die Anzahl der Ereignisse, die sich in einem Intervall $[t, t + \tau]$ ereignen, ist unabhängig von der Anzahl der Ereignisse, die sich in anderen, nicht überlappenden Intervallen ereignen. (Unabhängigkeit)

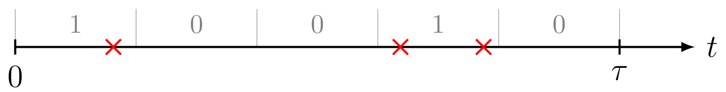
### Beispiele

- Anzahl Druckfehler auf einer Buchseite
- Anzahl Blitzeinschläge pro Fläche und Jahr
- Anzahl radioaktiver Zerfälle einer Substanz pro Zeiteinheit (Messdauer $\ll$ Halbwertszeit)
- Anzahl der pro Stunde an einem Schalter ankommenden Kunden

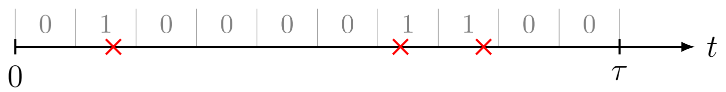
### Vom Bernoulli-Prozess zum Poisson-Prozess

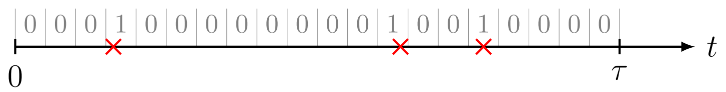
Während der Zeitspanne $\tau$ erwarten wir im Mittel $\lambda$ Ereignisse.

Unterteilen wir das Intervall $[0, \tau]$ in $n$ Teilintervalle gleicher Länge, können wir den oben beschriebenen Prozess näherungsweise durch einen Bernoulli-Prozess ersetzen.




Indem wir die Anzahl $n$ der Teilintervalle erhöhen, können wir vermeiden, dass zwei Ereignisse in das gleiche Teilintervall fallen.




Bei einem Bernoulli-Prozess der Länge $n$ beträgt der Erwartungswert der Anzahl Erfolge $\mathrm{E}(S) = np$. Dies soll nun der Intensität $\lambda$ des Poisson-Prozesses entsprechen.

Somit gilt $\lambda = np$ und $p = \dfrac{\lambda}{n}$

### Beispiel

Die folgende Tabelle enthält die Wahrscheinlichkeiten $B(n, p, k)$ für Bernoulli-Prozesse der Länge $n = 10, 10^2, 10^3, 10^4$, wobei die Wahrscheinlichkeit $p$ so gewählt wird, dass der Erwartungswert der Anzahl Erfolge konstant $\lambda = 2 = p/n$ beträgt.

| $k$ | $B\left(10, \frac{2}{10}, k\right)$ | $B\left(10^2, \frac{2}{10^2}, k\right)$ | $B\left(10^3, \frac{2}{10^3}, k\right)$ | $B\left(10^4, \frac{2}{10^4}, k\right)$ |
|---|---|---|---|---|
| 0 | 0.1074 | 0.1326 | 0.1351 | 0.1353 |
| 1 | 0.2684 | 0.2707 | 0.2707 | 0.2707 |
| 2 | 0.3020 | 0.2734 | 0.2709 | 0.2707 |
| 3 | 0.2013 | 0.1823 | 0.1806 | 0.1805 |
| 4 | 0.0881 | 0.0902 | 0.0902 | 0.0902 |
| 5 | 0.0264 | 0.0353 | 0.0360 | 0.0361 |
| ... | ... | ... | ... | ... |

*Vermutungen:* (a) $B(n, p/n, k)$ ist konvergent für $n \to \infty$
(b) $B(n, p/n, n) \to 0$ für $n \to \infty$

### Die Poisson-Verteilung als Limes der Binomialverteilung

$$p_B(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \qquad \text{ersetze } p \text{ durch } \frac{\lambda}{n}$$

$$= \frac{n(n-1)\ldots(n-k+1)}{k!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n}{n} \cdot \frac{n-1}{n} \ldots \frac{n-k+1}{n} \cdot \frac{\lambda^k}{k!} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k}$$

Ist $k$ konstant und strebt $n \to \infty$, so folgt:

$$\frac{n}{n} \to 1, \ \frac{n-1}{n} = \frac{n}{n} - \frac{1}{n} \to 1, \ldots, \frac{n-k+1}{n} = \frac{n}{n} - \frac{k-1}{n} \to 1$$

$$\left(1 - \frac{\lambda}{n}\right)^n \to \mathrm{e}^{-k} \quad \text{und} \quad \left(1 - \frac{\lambda}{n}\right)^{-k} \to 1$$

$$B(n, p = \lambda/n, k) \overset{n\to\infty}{\longrightarrow} p_X(k, \lambda) = p_\lambda(k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$$

### Zusammenfassung

Die Wahrscheinlichkeitsfunktion $p_\lambda(k) = \dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$ modelliert die Anzahl von Ereignissen, die bei einer konstanten Intensität $\lambda$ unabhängig voneinander in einem festen Zeitintervall oder räumlichen Gebiet eintreten.

### Der Erwartungswert der Poisson-Verteilung

gut zu wissen: $\displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!} = \mathrm{e}^x$ und $\displaystyle\sum_{k=0}^{n} ky = \sum_{k=1}^{n} ky$

$$\mathrm{E}(X) = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} k\frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} (k+1)\frac{\lambda^{k+1}}{(k+1)!}$$

$$= \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k \lambda}{k!} = \lambda \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda \mathrm{e}^{-\lambda} \mathrm{e}^{\lambda} = \lambda$$

**Varianz der Poisson-Verteilung**

$$E(X^2) = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{\infty} (k+1)^2 \frac{\lambda^{k+1}}{(k+1)!} e^{-\lambda}$$

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

$$= \lambda\big( \mathrm{E}(X) + e^{-\lambda} e^{\lambda} \big) = \lambda^2 - \lambda$$

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$$

**Beispiel (Vergleich einer empirischen Verteilung)**

Aus den Unterlagen von 10 Kavallerieregimentern über einen Zeitraum von 20 Jahren (also 200 Regimentsjahre) ermittelte der Ökonom und Statistiker Ladislaus von Bortkiewicz (1868–1931) für die Anzahlen $k$ der im Laufe eines Jahres pro Regiment auftretenden Todesfälle durch Hufschlag folgende Häufigkeiten $n_k$.

| $k$ | 0 | 1 | 2 | 3 | 4 | $\geq 5$ |
|---|---|---|---|---|---|---|
| $n_k$ | 109 | 65 | 22 | 3 | 1 | 0 |

(a) Berechne die durchschnittliche Anzahl der Toten durch Hufschlag pro Regimentsjahr.

$$\frac{1}{200}(0 \cdot 109 + 1 \cdot 65 + 2 \cdot 22 + 3 \cdot 3 + 4 \cdot 1) = 0.61$$

(b) Verwende das Resultat von (a) als Erwartungswert $\lambda$ für eine poissonverteilte Zufallsvariable und schätze damit die absoluten Häufigkeiten für die Anzahl der Toten durch Hufschlag.

$\lambda = 0.61$

| $k$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $p_\lambda(k)$ | 0.543 | 0.331 | 0.101 | 0.021 | 0.003 |
| $\hat{n}_k = 200 p_\lambda(k)$ | 108.6 | 66.2 | 20.2 | 4.2 | 0.6 |

(c) Vergleiche die tatsächliche (empirische) Verteilung mit der geschätzten Verteilung der Anzahl der Toten durch Hufschlag.

Die Übereinstimmung ist „erstaunlich gut“.