Einführung in die stochastischen Prozesse

Vorlesung mit Übung von

Sonja Greven und David Rügamer
Institut für Statistik
Ludwig-Maximilians-Universität München

Sommersemester 2017

Termine

- Vorlesung (Prof. Dr. Sonja Greven): Montag 12.15 - 13.50 (A 022) und ca. jeden zweiten Mittwoch, 10.15-11.50 (A 017)

- Übung (David Rügamer):
  ca. jeden zweiten Mittwoch, 10.15-11.45 (A 017)

- Genaue Zeiten auf der Moodleseite der Veranstaltung:
  `https://moodle.lmu.de/course/view.php?id=1351`

- Sprechstunde nach Vereinbarung

Literatur & Unterlagen

- Folien und Formelsammlung sind auf der Moodleseite der Veranstaltung.
  Die Folien beruhen ursprünglich auf dem Vorlesungsskript, was ebenfalls mit Stand 2015 auf der Seite zu finden ist, sind aber aktueller und fehlerfreier.

Primärliteratur:

- Fahrmeir, L., Kaufmann, H. und Ost, F.: Stochastische Prozesse. Hanser Verlag München, 1981. (Kopiervorlage auf Anfrage erhätlich).

Sekundärliteratur:

- Billingsley, P.: Probability and measure. Wiley, New York (2nd ed.), 1986
- Karlin, S. und Taylor, H. M.: A first course in stochastic processes. Academic Press, 1975.
- Guttorp, P.: Stochastic Modeling of Scientific Data. Chapman & Hall, 1995.

Scheinerwerb

Klausur am Ende des Semesters.

- Vorschlag: Freitag, 11.08.2017, von 10 - 12 Uhr
- 90 Minuten
- Zugelassene Hilfsmittel: Formelsammlung mit handschriftlichen Notizen auf den Vorderseiten, Taschenrechner und Wörterbuch
- Inhalte von Vorlesung **und** Übung sind relevant. Die Hausübungen stellen eine zusätzliche Übungs- und Vertiefungsmöglichkeit dar.
- Eine Anmeldung per Moodle ist notwendig.

ARSnova

- ARS steht für 'Audience Response System' und wir werden es für Publikumsfragen und Feedback benutzen.
- Browser-basiert, keine Installation und keine Registrierung notwendig. Anonym.
- Link oder QR-Code: `https://arsnova.eu/mobile/#id/32819119`

Ziel der Vorlesung

- Grundannahme vieler statistischer Verfahren: Stochastische Unabhängigkeit der Beobachtungen.

- Stochastische Prozesse: Modelle zur statistischen Analyse bei abhängigen Daten.

- In dieser Vorlesung:
  - Überwiegend zeitliche Abhängigkeit.

    (Räumliche Abhängigkeit $\Rightarrow$ Vorlesung Räumliche Statistik).

    Unterschied zur Zeitreihenanalyse (Zeit diskret, Zustandsraum stetig):
    Oft diskrete Daten in stetiger oder diskreter Zeit.
  - Vorstellung verschiedener Modellklassen von stochastischen Prozessen.
  - Theoretische Grundlagen.
  - Inferenzkonzepte.
  - Beispiele aus verschiedenen Anwendungsbereichen.

Stochastische Prozesse in der Statistik

Stochastische Prozesse sind in vielen Bereichen der Statistik wichtig:

- Sie bilden die Grundlage für statistische Modelle für abhängige Daten z.B.
  - in der räumlichen Statistik (z.B. Punktprozesse, Markov-Zufallsfelder)
  - in den Bereichen Zeitreihenanalyse und Longitudinaldaten
  - in der Lebenszeit- und Ereigniszeitanalyse (Punkt- und Zählprozesse)
  - in weiteren Bereichen, z.B. genetische Statistik, Finanzstatistik, ...
- Sie spielen eine Rolle in der bayesianischen Inferenz
  - bei der Wahl von geeigneten Prioris (z.B. random walks)
  - bei der Inferenz mit Markov Chain Monte Carlo (MCMC)-Verfahren.
- Sie spielen ebenfalls eine Rolle in der asymptotischen Statistik.

Überblick

Kapitel 1: Einführung und Beispiele

- Einführende Beispiele
- Erste Definition stochastischer Prozesse
- Einige spezielle stochastische Prozesse
- Einführung erster Grundbegriffe

Kapitel 2: Grundbegriffe der allgemeinen Theorien stochastischer Prozesse

- Definitionen stochastischer Prozesse
- Verteilung eines stochastischen Prozesses
- Eigenschaften stochastischer Prozesse

Kapitel 3: Markov-Ketten

Kapitel 4: Diskrete Markov-Prozesse

Kapitel 5: Erneuerungs- und Semi-Markov-Prozesse

Ausblick: Spezielle stochastische Prozesse

Kapitel 6: Martingale

Kapitel 7: Punkt- und Zählprozesse

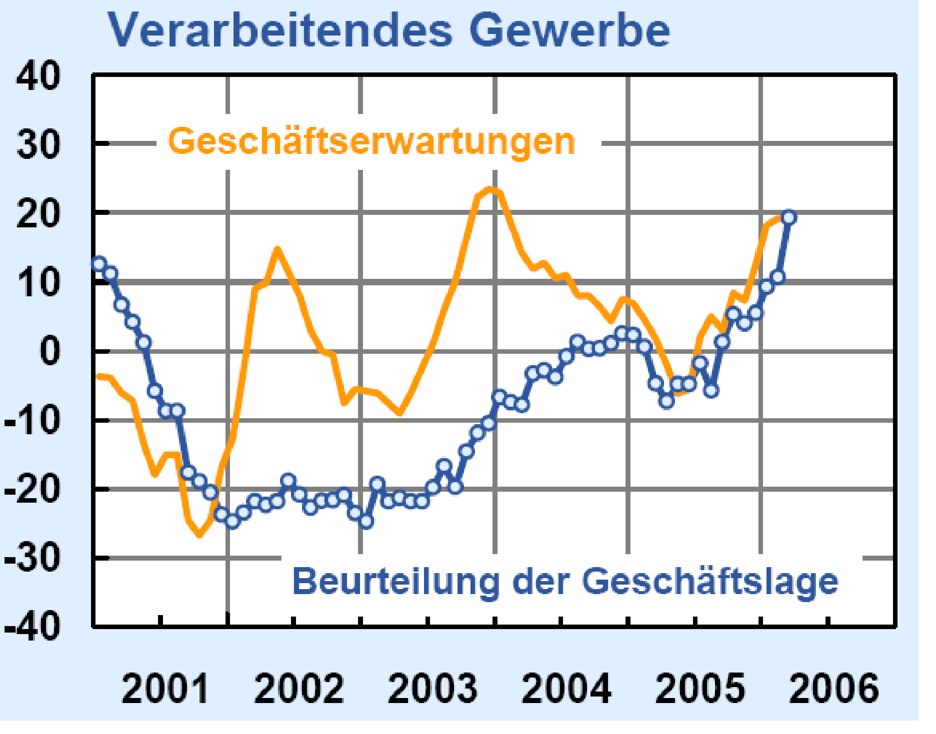
Kapitel 8: Markov-Prozesse mit stetigem Zustands- und Parameterraum

wird von Dr. Holger Fink nach dem SoSe als Blockkurs voraussichtlich vom 21.-24.8.17 angeboten.

Kapitel 1
Einführung und Beispiele

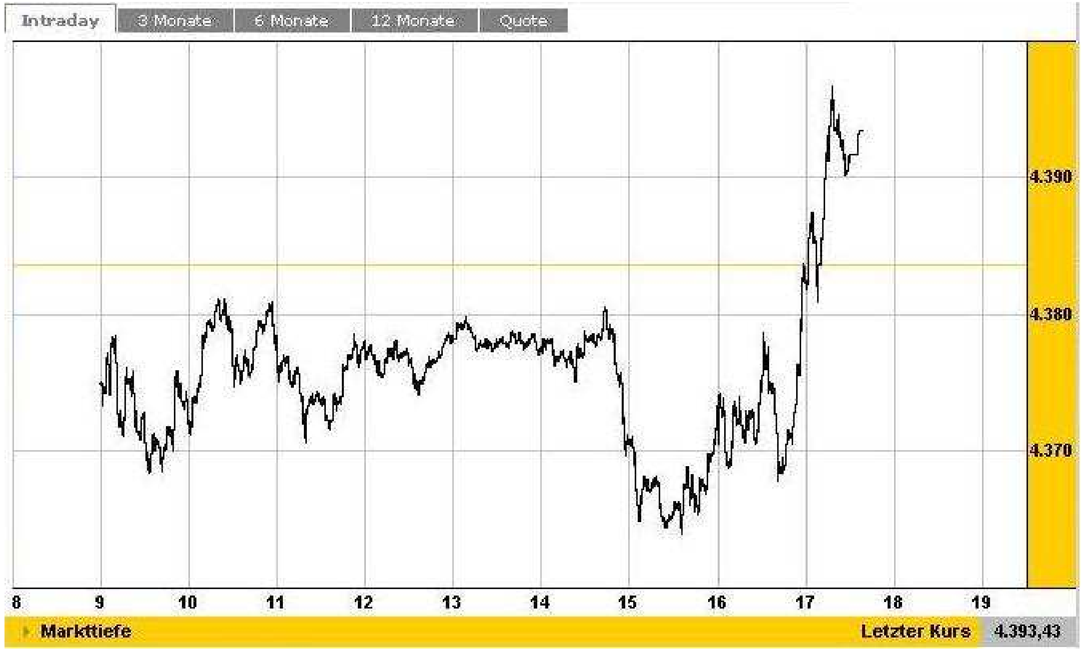
1.1 Einführende Beispiele und erste Definition

- Rein zeitliche Beispiele:
  - IFO Geschäftsklima-Index



Zeit: diskret; Zustandsraum: stetig.

– Aktienkurse (Inter-Tagesdaten)

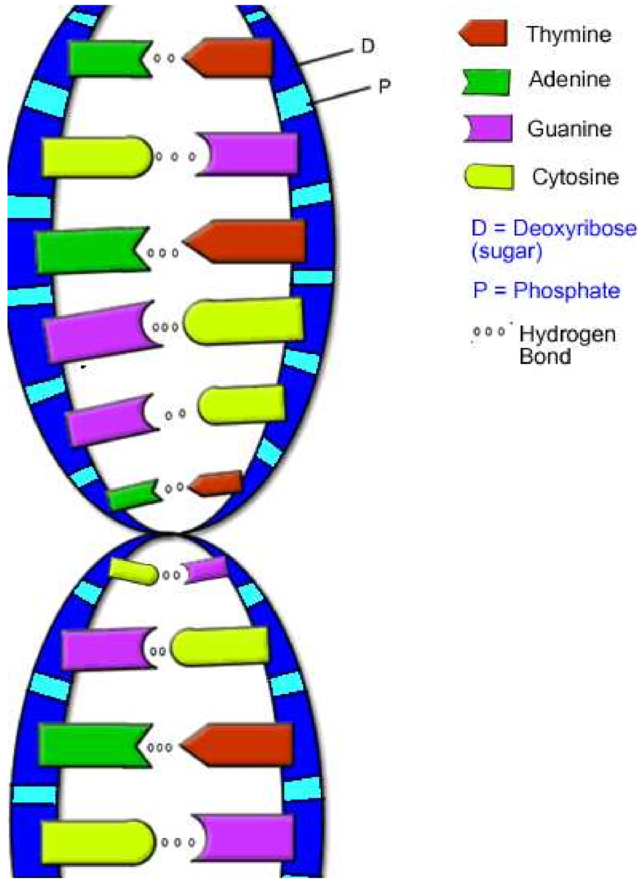


Zeit: stetig; Zustandsraum: stetig.

– Individuelle Arbeitslosigkeitsverläufe

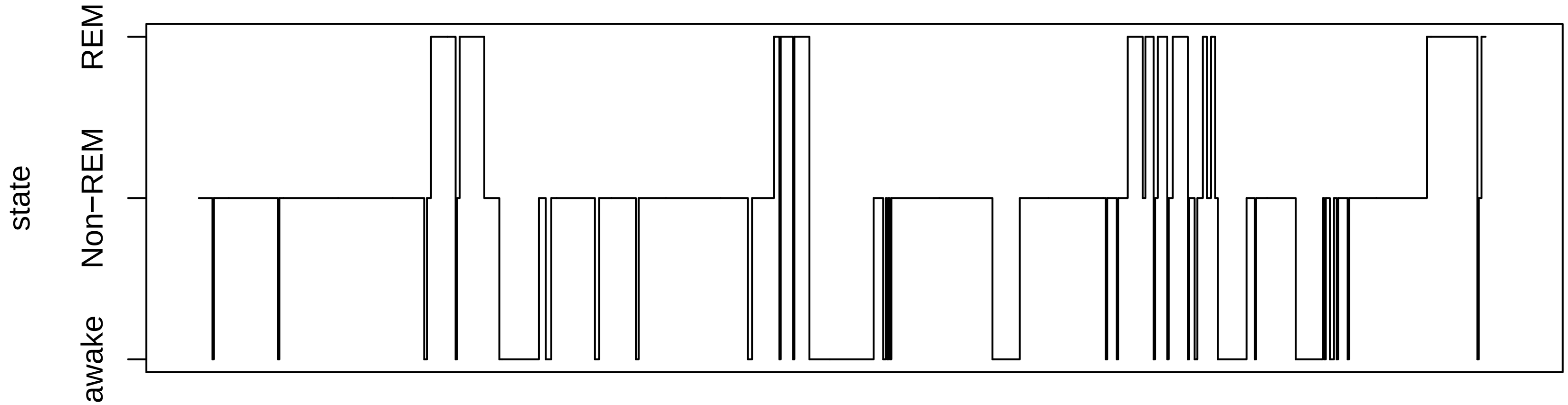
Zeit: diskret; Zustandsraum: binär (ja/nein) bzw. kategorial (zusätzlich Studierende, Elternzeit, . . .).

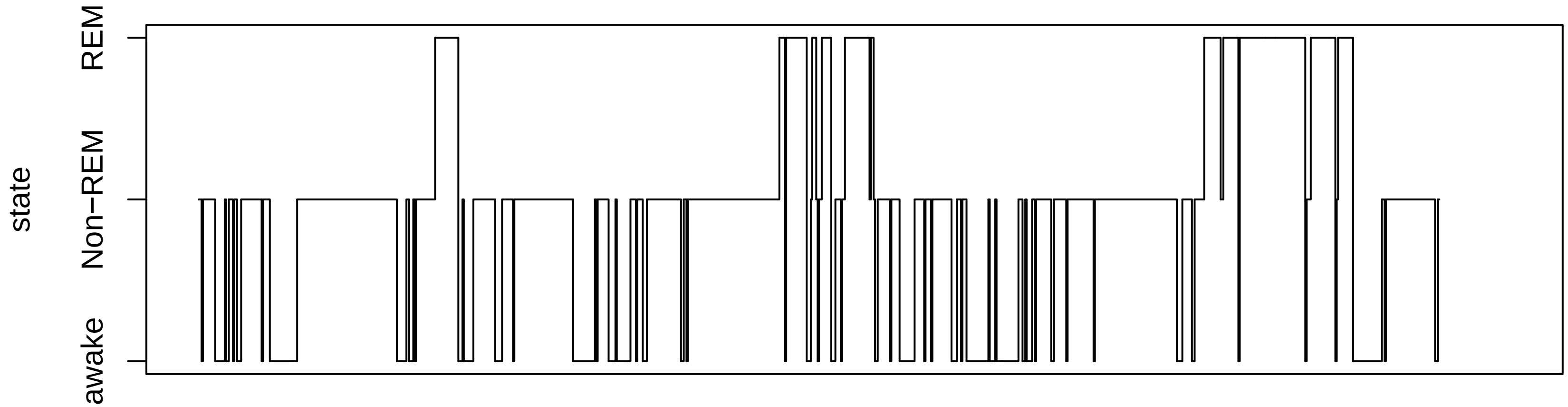
– DNA-Sequenzen



"Zeit": diskret; Zustandsraum: diskret.

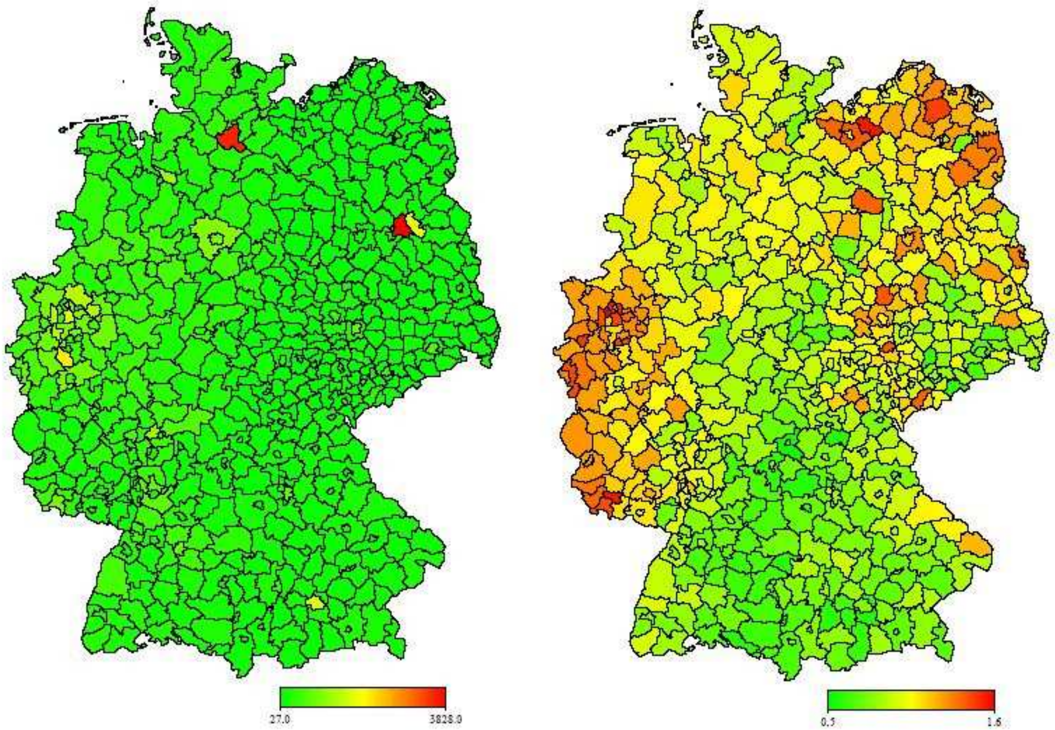
– Schlafzustände





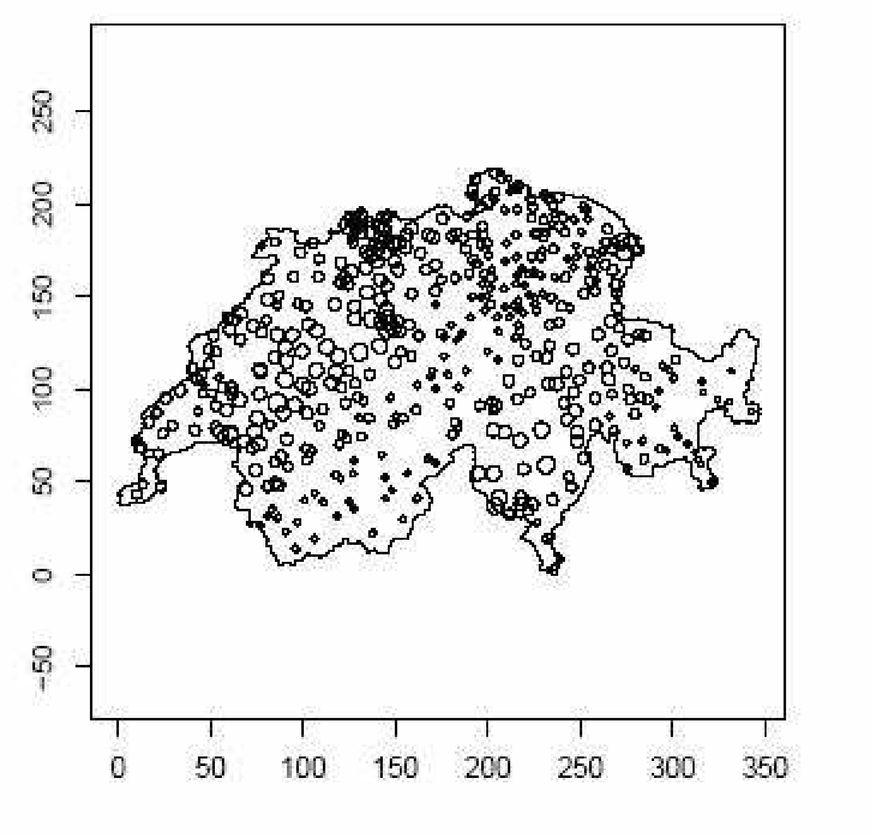
Zeit: stetig; Zustandsraum: diskret.

- Rein räumliche Beispiele:
  - Krebsatlas



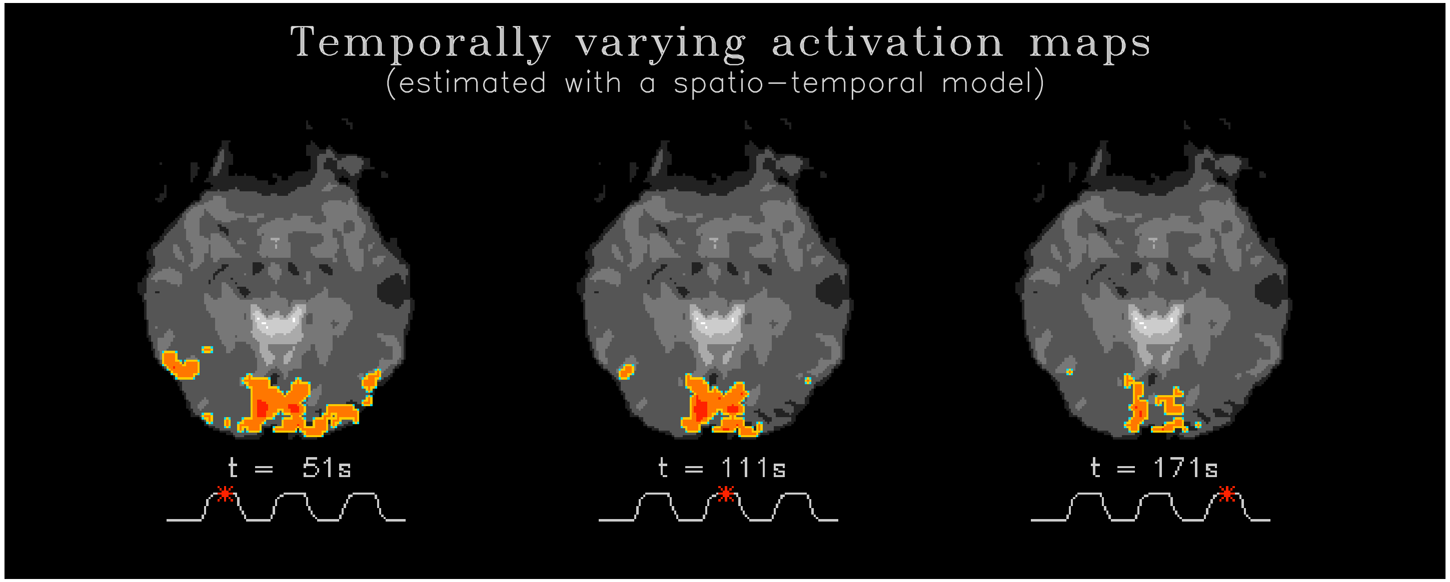
Raum: diskret; Zustandsraum: diskret.

– Schweizer Regenfalldaten



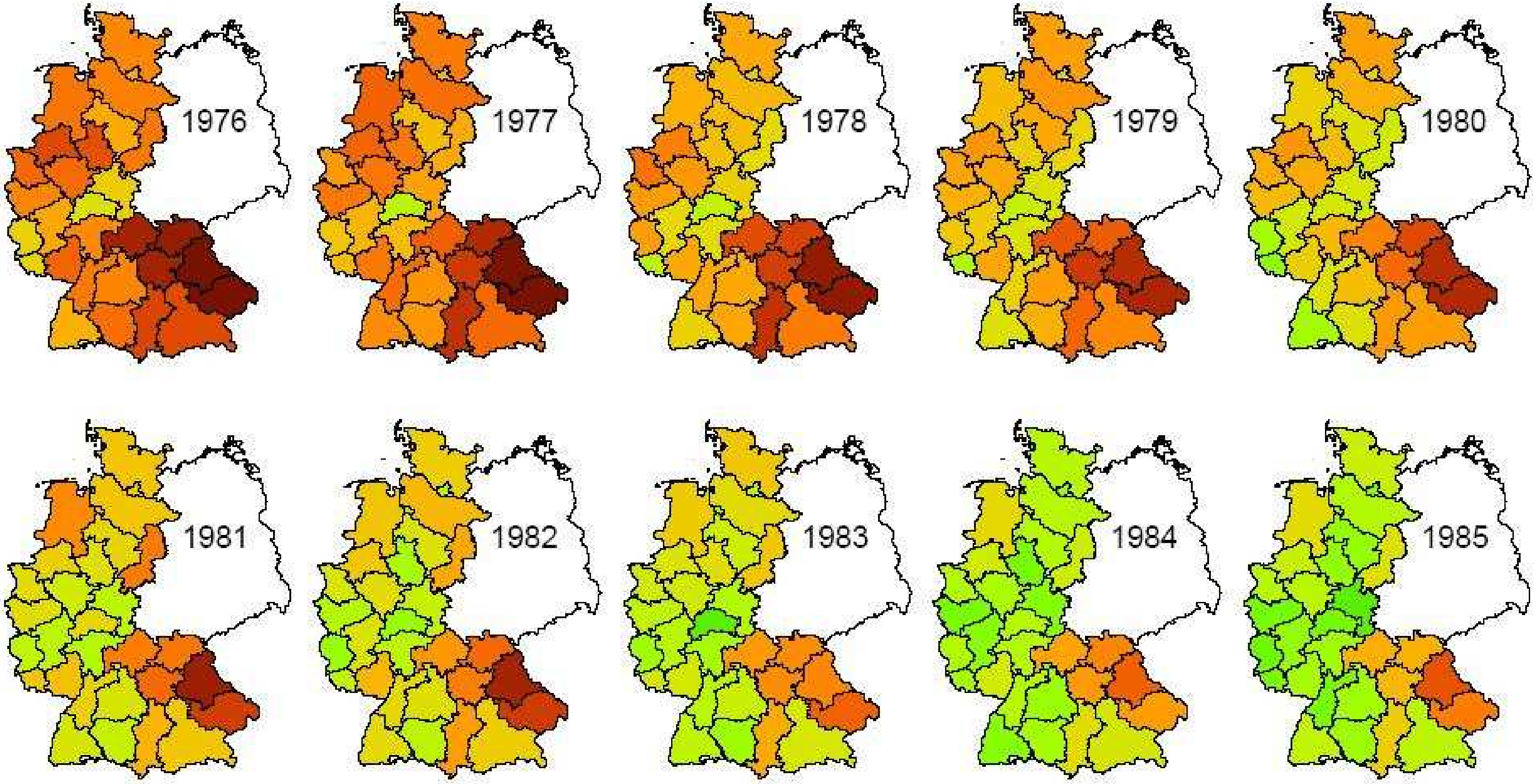
Raum: stetig; Zustandsraum: stetig.

- Räumlich-zeitliche Beispiele:
  - fMRI-Daten (funktionelle Magnetresonanztomographie)



Zeit: diskret, Raum: diskret, Zustandsraum: stetig.

– Krebsatlas



Zeit: diskret; Raum: diskret; Zustandsraum: diskret.

- Allgemein: Realisierungen von Zufallsvariablen $X_t,\ t \in T$, mit einer geeigneten Indexmenge T, z.B. mit

$$\begin{array}{ll} T \subseteq \mathbb{N}_0,\ \mathbb{R}_+ & \text{(Zeit)}, \\ T \subseteq \mathbb{Z}^2 & \text{(räumliches Gitter)}, \\ T \subseteq \mathbb{R}^2 & \text{(Raum, kontinuierliche Teilmenge)}, \\ T \subseteq \mathbb{Z}^2 \times \mathbb{N}_0 & \text{(Raum–Zeit)}. \end{array}$$

- Definition: Stochastischer Prozess als Familie von Zufallsvariablen

  Eine Familie $\{X_t,\ t \in T\}$ von Zufallsvariablen heißt **stochastischer Prozess**. Die Indexmenge T heißt **Parameterraum**, der Wertebereich S der Zufallsvariablen heißt **Zustandsraum**.

- Genauer nimmt man an:

  $X_t:\ (\Omega, \mathcal{F}, P) \longrightarrow (S, \mathcal{S})$ Zufallsvariable für alle t

  $\mathcal{S}$ σ–Algebra auf S, $(S, \mathcal{S})$ Meßraum

  $(\Omega, \mathcal{F}, P)$ Wahrscheinlichkeitsraum

  $\Rightarrow \quad X = \{\Omega, \mathcal{F}, P, \{X_t,\ t \in T\}\}$ stochastischer Prozess

1.2 Spezielle Stochastische Prozesse

1.2.1 Irrfahrten (Random Walks)

- Stochastische Modellierung von einfachen Spielsituationen. Man betrachte zwei Spieler A und B, die in jeder Spielrunde mit Wahrscheinlichkeit p bzw. q (bzw. r) gewinnen (bzw. unentschieden spielen). X_t bezeichne den Gewinn von A nach t Spielrunden.

- Irrfahrten können ebenfalls als diskrete Idealisierung von Kursverläufen betrachtet werden (siehe Abschnitt zum Wienerprozess).

- Namensgebend ist die Vorstellung einer desorientierten Person, die sich in jedem Schritt mit gewisser Wahrscheinlichkeit nach rechts bzw. links bewegt bzw. stehenbleibt.

- Irrfahrten sind wichtiger Baustein oder Spezialfall von komplexeren statistischen Modellen, z.B. in der räumlichen Statistik oder Zeitreihenanalyse, und werden auch zur Konstruktion vonPrioris in der bayesianischen Statistik verwendet.

- Diskrete einfache Irrfahrt auf der Geraden:
  - Start in $X_0 = 0$ (Zeit $t = 0$).
  - Bewegung: $(t = 1, 2, \ldots)$

    Ein Schritt nach rechts mit Wahrscheinlichkeit p

    Ein Schritt nach links mit Wahrscheinlichkeit q

    Verbleiben im Zustand mit Wahrscheinlichkeit r

    $p + q + r = 1.$
  - X_t Position nach t-tem Schritt.

- Alternative Darstellung:
  - $\{Z_t,\, t = 1, 2, ...\}$ i.i.d. Folge von Zufallsvariablen für t-ten Schritt.
  - $Z_t \in \{-1, 0, 1\}$

$$\begin{aligned} P(Z_t = -1) &= q \\ P(Z_t = 0) &= r \\ P(Z_t = 1) &= p \end{aligned}$$

  - Definition von X_t:

$$\begin{aligned} X_0 &= 0 \\ X_t &= Z_1 + Z_2 + \ldots + Z_t \end{aligned}$$

bzw.

$$X_t = X_{t-1} + Z_t, \quad t \geq 1$$

- Formale Definition:
  - Die Folge $X = \{X_t,\ t \in \mathbb{N}_0\}$, mit $X_t = X_{t-1} + Z_t$ und $X_0 = 0$, heißt **(einfache) Irrfahrt** auf der Geraden.
  - $\{X_t,\ t \in \mathbb{N}_0\}$ ist ein stochastischer Prozess mit Parameterraum $T =$ und Zustandsraum $S =$.
  - Zugrundeliegender Ergebnisraum Ω mit Ergebnissen ω:

    ω Folge der Realisierungen von $Z_1, Z_2, \ldots$, z.B.

$$
\begin{aligned}
\omega &= (1, 1, -1, 1, 1, 0, 0, 1, -1, \ldots) \\
&= (Z_1(\omega) = 1, Z_2(\omega) = 1, Z_3(\omega) = -1, \ldots)
\end{aligned}
$$

    $\Rightarrow \Omega = \{-1, 0, 1\}^{\mathbb{N}}$.
  - Die durch (t, X_t) gebildete Treppenfunktion heißt **Pfad**, Trajektorie oder Realisierung des Prozesses.

- Spezialfälle:

$$r = 0 \qquad \text{kein Unentschieden}$$

$$p = q = \frac{1}{2} \qquad \text{faires Spiel, symmetrische Irrfahrt}$$

- Modifikation: Absorbierende Schranken

Interpretation: Spieler A mit Anfangskapital $a > 0$

Spieler B mit Anfangskapital $b > 0$

X_t Gewinn von A nach t Spielen

$\{X_t = -a\}$ Ruin von A ("Gambler's ruin")

$\{X_t = b\}$ Ruin von B.

Ziel: Berechnung von Ruin- bzw. Gewinnwahrscheinlichkeiten.

Gewinnwahrscheinlichkeit bei $p = q = \frac{1}{2}$ ist

$$\text{für } A: \quad P_A = \frac{a}{a+b} \quad \text{und für } B: \quad P_B = \frac{b}{a+b}$$

Allgemeinere Gewinnwahrscheinlichkeiten $\Rightarrow$ Übung.

- Markov–Eigenschaft der diskreten Irrfahrt:

$$\begin{aligned} &P(X_{t+1} = j \quad | \quad X_t = i, X_{t-1} = i_{t-1}, X_{t-2} = i_{t-2}, \ldots, X_0 = i_0) = \\ &P(X_{t+1} = j \quad | \quad X_t = i). \end{aligned}$$

Interpretation: Die Zukunft X_{t+1} ist bei bekannter Gegenwart X_t (bedingt) unabhängig von der Vergangenheit $X_{t-1}, \ldots, X_1$.

Hierbei ist das Wort 'bedingt' wesentlich: $\{X_{t+1} = j\}$ und $\{X_{t-1} = i_{t-1}, \ldots, X_1 = i_1\}$ sind i.d.R. abhängige Ereignisse bzw. $X_{t+1}, X_{t-1}, \ldots, X_1$ abhängige Zufallsvariablen.

Jedoch gilt bei der Markoveigenschaft: $X_{t+1} \perp (X_{t-1}, \ldots, X_1)|X_t$, d.h. bedingte Unabhängigkeit gegeben X_t.

- Ungerichtete (äquivalente) Form der Markov–Eigenschaft:

$$P(X_t = i|X_{s \neq t}) = P(X_t = i|X_{t+1} = i_{t+1}, X_{t-1} = i_{t-1})$$

- Zwei–dimensionale symmetrische Irrfahrt

$$
\begin{aligned}
X_t &= (X_{1t}, X_{2t}) \in \mathbb{Z}^2 = S \\
X_{t+1} &= X_t + Z_t \ ; \\
Z_t &\in \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right)
\end{aligned}
$$

mit

$$
P\left(Z_n = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \ldots = P\left(Z_n = \begin{pmatrix} 0 \\ -1 \end{pmatrix}\right) = \frac{1}{4}
$$

- Allgemeine Irrfahrt (Random Walk):

  $\{Z_t,\ t \in \mathbb{N}\}$ iid Folge, Z_t beliebig verteilt.

- Gauß-Irrfahrt:

  $\{Z_t,\ t \in \mathbb{N}\}$ iid $N(0, \sigma^2)$, "Gauß'sches Weißes Rauschen"

Markov–Eigenschaft:

$$f(x_t \mid X_{t-1} = x_{t-1}, \ldots, X_1 = x_1) = f(x_t \mid X_{t-1} = x_{t-1}) \quad \sim \quad N(x_{t-1}, \sigma^2)$$

$$\text{bzw.} \quad X_t \mid X_{t-1}, \ldots, X_1 \sim X_t \mid X_{t-1} \quad \sim \quad N(X_{t-1}, \sigma^2)$$

- Autoregressiver Prozess der Ordnung 1 ($AR(1)$)

$$X_t = \rho X_{t-1} + Z_t, \quad Z_t \text{ i.i.d.}$$

- Autoregressiver Prozess der Ordnung p ($AR(p)$)

$$X_t = \rho_1 X_{t-1} + \ldots + \rho_p X_{t-p} + Z_t, \quad Z_t \text{ i.i.d.}$$

- Gauß-Prozesse, falls Z_t i.i.d. $N(0, \sigma^2)$.

- Markov-Eigenschaft:

  Für $AR(1)$ (Markov-Eigenschaft 1. Ordnung)

$$X_t \mid X_{t-1}, \ldots, X_1 \sim X_t | X_{t-1} \sim N(\rho X_{t-1}, \sigma^2)$$

  Für $AR(p)$ (Markov-Eigenschaft p. Ordnung):

$$X_t \mid X_{t-1}, \ldots, X_1 \sim X_t \mid X_{t-1}, \ldots, X_{t-p} \sim N(\mu_t, \sigma^2)$$

  mit

$$\mu_t = \rho_1 X_{t-1} + \ldots + \rho_p X_{t-p}.$$

  Autoregressive Prozesse bilden wichtige Bausteine in der Zeitreihenanalyse.

- Räumliche Markov-Prozesse

  $\{X_s,\ s = (i,j) \in \mathbb{Z}^2\}$ heißt räumlicher (Gitter-)Prozess oder Zufallsfeld.

  Beispiele:

  - $X_{ij} \in \{0,1\}$ Indikatorvariablen, z.B.

$$X_{ij} = \begin{cases} 1 & \text{archäologischer Fund in } (i,j) \\ 0 & \text{sonst} \end{cases}$$

  - $X_s \in \{0,1,2,\ldots\}$ Zählvariable, z.B. Anzahl von Krebserkrankungen in Landkreisen; irreguläres Gitter.
  - X_{ij} stetig verteilt, z.B. Graustufe/Farbe in der Bildanalyse

- Räumliche Markov–Eigenschaft

$$f(x_{ij} | \{X_{kl},\ (k,l) \neq (i,j)\}) = f(x_{ij} | \{X_{kl},\ (k,l) \sim (i,j)\})$$

  $\sim$: "Nachbar von"

1.2.2 Wiener Prozess

- Stochastisches Modell für die Brown'sche Bewegung (zeitstetige Irrfahrt kleiner Teilchen in homogener, ruhender Flüssigkeit).

  Irrfahrt kommt durch zufälliges Zusammenstoßen mit Molekülen der Flüssigkeit zustande. Moderne Herleitung: N. Wiener, 1923.

  Parameterraum $T = \mathbb{R}_+$.

  Zustandsraum $S = \mathbb{R}$ (bzw. $\mathbb{R}^2$, $\mathbb{R}^3$).

  Bezeichnung: $W = \{W_t,\ t \in \mathbb{R}_+\}$ oder $\{W(t),\ t \in \mathbb{R}_+\}$

- Wichtiger Baustein zur Modellierung von Wertpapierpreisen mit Anwendung in der Optionsbewertung (Black-Scholes-Regel):

  Bezeichne $P_0(t)$ den Preis eines risikolosen Wertpapiers (z.B. Sparguthaben) zum Zeitpunkt t. Bei stetiger Verzinsung mit konstantem Zinssatz r gilt:

$$\begin{aligned} P_0(t) &= P_0(0)e^{rt} \\ \ln P_0(t) &= \ln P_0(0) + rt \end{aligned}$$

$\Longrightarrow$ loglinearer Ansatz für Aktienkurse:

$$\ln P(t) = \ln P(0) + rt + W(t),$$

$W(t)$ regelloser Fehler in stetiger Zeit.

- Wiener-Prozess kann als Grenzfall von Irrfahrten hergeleitet werden:

  $X(t)$ symmetrische diskrete Irrfahrt, $X(0) = 0$.

  Bewegungen zu den Zeitpunkten $n\Delta t$, $n = 1, 2, \ldots$ um $\pm\Delta x$.

$$\begin{aligned} X(t) &= \text{Lage des Teilchens für } t = n\Delta t \\ &= \sum_{k=1}^{n} Z_k \end{aligned}$$

Z_k i.i.d. mit $\begin{cases} P(Z_k = +\Delta x) = \frac{1}{2} \\ P(Z_k = -\triangle x) = \frac{1}{2} \end{cases}$

$$\begin{aligned} E(Z_k) &= 0 \\ Var(Z_k) &= (\Delta x)^2 \end{aligned}$$

$$\Longrightarrow \qquad \begin{aligned} E(X(t)) &= 0 \\ Var(X(t)) &= (\Delta x)^2 n = \underbrace{\frac{(\Delta x)^2}{\Delta t}}_{=:\sigma^2} t \end{aligned}$$

Grenzübergang $n \to \infty$ ($\Delta t \to 0, \Delta x \to 0$), so dass

$$\sigma^2 := \frac{(\triangle x)^2}{\triangle t}, \quad 0 < \sigma^2 < \infty$$

konstant bleibt.

Zentraler Grenzwertsatz

$$\Longrightarrow \quad X(t) \sim N(0, \sigma^2 t) \quad \text{für } n \longrightarrow \infty.$$

- "Herleitung" des Wiener-Prozesses als Grenzfall von Irrfahrten funktioniert auch für allgemeine "symmetrische" Irrfahrten, z.B. mit

$$Z_k \sim N(0, \sigma^2 \Delta t), \quad t = n\Delta t.$$

- Vor dem Grenzübergang gilt
  - Zuwächse $X(t) - X(s)$, $X(v) - X(u)$, mit $u < v < s < t$ sind unabhängig (zusammengesetzt aus getrennten i.i.d. Teilsummen)
  - Zuwächse $X(t+s) - X(s)$ sind stationär (Verteilung hängt nur von der Zeitdifferenz ab):
  $$X(t+s) - X(s) \quad \sim \quad X(t) - X(0) \quad \sim \quad N(0, \sigma^2 t)$$
  - Plausibel: Unabhängigkeit und Stationarität der Zuwächse überträgt sich auf Grenzprozess für $n \longrightarrow \infty$. Exakter Beweis aber schwierig.

  $\Rightarrow$ Axiomatischer Zugang

- Axiomatische Definition des Wiener-Prozesses

  Ein stochastischer Prozess $W = \{W(t), t \in \mathbb{R}_+\}$, $S = \mathbb{R}$, heißt Wiener-Prozess, wenn gilt:

(W1) Die Zuwächse sind normalverteilt und stationär:

$$W(s+t) - W(s) \quad \sim \quad N(0, \sigma^2 t) \quad \text{für alle } s, t \geq 0.$$

(W2) Für alle $0 \leq t_1 < t_2 < \ldots < t_n,\ n \geq 3$ sind die Zuwächse

$$W(t_2) - W(t_1),\ \ldots,\ W(t_n) - W(t_{n-1})$$

unabhängig.

(W3) $W(0) = 0$.

(W4) Die Pfade sind stetig.

(W3) und (W4) werden häufig nur mit Wahrscheinlichkeit 1 gefordert.

- Bemerkungen:
  - Für $\sigma^2 = 1$ heißt der Wiener–Prozess normiert.
  - (W3) ist Anfangsbedingung. Addition von c liefert

$$\tilde{W}(t) = W(t) + c$$

    mit $\tilde{W}(0) = c$.
  - (W1), (W2), (W3) bestimmen die endlich-dimensionalen Verteilungen von

$$W(t_1), W(t_2), \ldots, W(t_n) \quad \forall\, n \geq 1, t_1, \ldots, t_n \in \mathbb{R}_+.$$

  - (W4) folgt nicht aus (W1) – (W3).

    Man müsste sogar zeigen, dass (W4) mit (W1) – (W3) verträglich ist.

- Eindimensionale Verteilungen:

$$W(t) \sim N(0, \sigma^2 t) \quad \forall\, t \in \mathbb{R}_+,$$

da

$$W(t) - W(0) = W(t) \sim N(0, \sigma^2 t)$$

nach (W1) und (W3).

- Bivariate Verteilungen:

$$\begin{pmatrix} W(s) \\ W(t) \end{pmatrix} \sim N(0, \Sigma), \quad 0 \leq s < t$$

mit

$$\Sigma = \sigma^2 \begin{pmatrix} s & s \\ s & t \end{pmatrix}.$$

Für s, t beliebig: $\quad Cov(W(s), W(t)) = \min(t, s)\sigma^2$.

Zugehörige Dichte:

$$f_{s,t}(x_1, x_2) = \frac{1}{2\pi\sigma^2\sqrt{s(t-s)}} \exp\left\{-\frac{1}{2\sigma^2}\left[\frac{x_1^2}{s} + \frac{(x_2-x_1)^2}{t-s}\right]\right\}, \quad s < t$$

- Bedingte Dichten:

$$\begin{aligned} W(s) \mid [W(t) = b] &\sim N\left(\frac{s}{t}b, \sigma^2 \frac{s}{t}(t-s)\right), s < t \\ W(t) \mid [W(s) = a] &\sim N(a, \sigma^2(t-s)) \end{aligned}$$

- Allgemeine endlichdimensionale Verteilungen: Für $0 \leq t_1 < t_2 < \ldots < t_n,\ n \in \mathbb{N}$ gilt

$$(W(t_1), \ldots, W(t_n))' \sim N(0, \sigma^2 \Sigma)$$

$$\Sigma = \begin{pmatrix} t_1 & t_1 & t_1 & \ldots & t_1 \\ t_1 & t_2 & t_2 & \ldots & t_2 \\ t_1 & t_2 & t_3 & \ldots & t_3 \\ \vdots & \vdots & \vdots & & \vdots \\ t_1 & t_2 & t_3 & \ldots & t_n \end{pmatrix}$$

Dichte:

$$\begin{aligned} f_{t_1, \ldots t_n}(x_1, \ldots, x_n) &= f(x_1) f(x_2 \mid x_1) \cdot \ldots \cdot f(x_n \mid x_{n-1}) \\ &= \frac{\exp\left\{-\frac{1}{2\sigma^2}\left[\frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1} + \ldots + \frac{(x_n - x_{n-1})^2}{t_n - t_{n-1}}\right]\right\}}{(\sigma\sqrt{2\pi})^n \sqrt{t_1(t_2 - t_1) \cdot \ldots \cdot (t_n - t_{n-1})}} \end{aligned}$$

Dies zeigt auch die Markoveigenschaft des Wiener Prozess.

- Pfadeigenschaften:
  - Stetig nach (W4).
  - In jedem endlichen Intervall von unbeschränkter Variation:

$$t_0, t_1, \ldots, t_n \text{ Gitter auf } [t_0, t_n], \text{ Schrittweite } \delta_n$$

$$\Longrightarrow \sum_{k=1}^{n} |W(t_k) - W(t_{k-1})| \longrightarrow \infty \,, \quad \text{für} \quad n \longrightarrow \infty$$

  - Mit Wahrscheinlichkeit 1 nirgends differenzierbar. Plausibilitätserklärung:

    Für den Differenzenquotienten gilt

$$\frac{W(t+h) - W(t)}{h} \sim N\left(0, \frac{\sigma^2}{\mid h \mid}\right)$$

    d.h. die Varianz des Differenzenquotienten $\longrightarrow \infty$ für $h \to 0$.

Es gilt sogar

$$P\left\{a \leq \frac{W(t+h) - W(t)}{h} \leq b\right\} \stackrel{h\to 0}{\longrightarrow} 0, \quad \text{[a,b] endl. Intervall}$$

$\Rightarrow$ Differenzenquotient kann nicht gegen endliche Zufallsvariable konvergieren.

1.2.3 Zählprozesse und Poisson-Prozess

- Punkt- und Zählprozesse treten in vielen Bereichen der statistischen Modellierung auf, z.B. in der räumlichen Statistik (Beispiel: Projekt 'Determining high-risk zones for unexploded World War II bombs by using point process methodology', Mahling, Hoehle, Kuechenhoff, 2013, JRSS-C), in der Versicherungsstatistik oder in der Lebenszeit- und Ereigniszeitanalyse.

- Allgemeine Notation für Zählprozesse:
  - $\{S_n, n \in \mathbb{N}\}$ Stochastischer Prozess von Ereigniszeitpunkten, z.B.
    - ◇ Eintreten von Schadensfällen bei KFZ–Versicherung
    - ◇ Todesfälle in klinischer Studie
    - ◇ Ankünfte von Jobs bei Computer, von Kunden bei "Servicestelle", . . .
    - ◇ Kauf eines Produkts
    - ◇ Transaktionen an der Börse

– $\{T_n, n \in \mathbb{N}\}$ Stochastischer Prozess von Verweildauern, Zwischenankunftszeiten, etc. mit ($S_0 := 0$)

$$T_n = S_n - S_{n-1}$$

$\Rightarrow$ **Zählprozess**

$$N(t) = \sum_{n=1}^{\infty} I_{(0,t]}(S_n) = \text{ Anzahl der Ereignisse in } (0, t].$$

– Vereinbarung: $T_n > 0$ für alle $n \in \mathbb{N}$, d.h. keine gleichzeitigen Ereignisse.

– Pfade sind rechtsstetige Treppenfunktionen mit Sprunghöhe 1

- Verhältnis zwischen Zählprozess und Ereigniszeitpunkten:

$$\begin{aligned} S_n \leq t \quad &\Leftrightarrow \quad N(t) \geq n \\ S_n \leq t < S_{n+1} \quad &\Leftrightarrow \quad N(t) = n \\ N(t) = \max_n \{S_n \leq t\} \quad &= \quad \min_n \{S_{n+1} > t\}. \end{aligned}$$

- Mehr zu Zählprozessen in Kapitel 7; hier: Poisson-Prozess als einfachster Zählprozess

- Definition: Zählprozess N mit **unabhängigen und stationären Zuwächsen**:

  N besitzt unabhängige Zuwächse $:\Leftrightarrow \forall\, n \;\forall\, t_0 < t_1 < \ldots < t_n$

$$N(t_1) - N(t_0), \ldots, N(t_n) - N(t_{n-1}) \quad \text{unabhängig}$$

  N besitzt stationäre Zuwächse $:\Leftrightarrow \forall\; 0 \leq t_1 < t_2,\; s > 0$

$$N(t_2) - N(t_1) \text{ und } N(t_2 + s) - N(t_1 + s) \text{ identisch verteilt}$$

- Definition (homogener) **Poisson-Prozess**:

  Ein Zählprozess $N = \{N(t), t \geq 0\}$ heißt (homogener) Poisson-Prozess :$\Leftrightarrow$

  (1) N hat unabhängige und stationäre Zuwächse

  (2) $P\{N(t+h) - N(t) \geq 2\} = o(h)$
  $P\{N(t+h) - N(t) = 1\} = \lambda h + o(h)$

- Einschub Landau-Symbol: $o(h)$ bezeichnet eine Funktion von h, die für $h \downarrow 0$ schneller gegen 0 geht als h, d.h. $\lim_{h \downarrow 0} \frac{o(h)}{h} = 0$. Z.B. $h^{1+\alpha} = o(h)$ für $\alpha > 0$.

  Rechenregeln: $o(h) \pm o(h) = o(h)$.

- Eigenschaft (2) lässt sich ersetzen durch
  - Sprünge der Pfade haben (mit Wkeit 1) die Höhe 1,
  - keine gleichzeitigen Ereignisse,
  - Zählprozess nach unserer Übereinkunft.

- Für einen Poisson-Prozess N gilt

$$P(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad n \in \mathbb{N}_0, \tag{$\star$}$$

d.h.

$$N(t) \sim Po(\lambda t)$$

mit Intensität oder Rate $\lambda > 0$.

Bemerkung: (2) kann auch durch $N(t) \sim Po(\lambda t)$ ersetzt werden.

Beweis:

Sei $p_0(t) := P(N(t) = 0)$, $p_n(t) := P(N(t) = n)$.

Wir zeigen zunächst $p_0(t) = e^{-\lambda t}$.

$$
\begin{aligned}
p_0(t+h) &= P\{N(t+h)=0\} \\
&= P\{N(t)=0, N(t+h)-N(t)=0\} \\
&\overset{(1)}{=} P\{N(t)=0\}\cdot P\{N(t+h)-N(t)=0\} \\
&\overset{(1)}{=} p_0(t)p_0(h) \\
\Rightarrow \frac{p_0(t+h)-p_0(t)}{h} &= p_0(t)\frac{p_0(h)-1}{h}
\end{aligned}
$$

Nach Definition des Poisson-Prosses, (2),

$$p_0(h) = P(N(h)=0) = 1-\lambda h - o(h) - o(h) = 1-\lambda h + o(h)$$

$$\Rightarrow p_0'(t) = \lim_{h\to 0} p_0(t)\frac{-\lambda h + o(h)}{h} = -\lambda p_0(t), \text{ mit } p_0(0) = P(N(0)=0) = 1.$$

Dies ist eine lineare Differentialgleichung für die unbekannte Funktion $p_0(t)$ und besitzt eine eindeutige Lösung. Einsetzen zeigt, dass $e^{-\lambda t}$ die Gleichung inklusive der Anfangsbedingung erfüllt.

Für $n > 0$ gilt

$$\begin{aligned}
p_n(t+h) &= P\{N(t+h)=n\} \\
&= P\{N(t)=n, N(t+h)-N(t)=0\} \\
&+ P\{N(t)=n-1, N(t+h)-N(t)=1\} \\
&+ \underbrace{\sum_{k=2}^{n} P\{N(t)=n-k, N(t+h)-N(t)=k\}}_{o(h)\text{ wegen }(2)} \\
&\overset{(1)}{=} p_n(t)p_0(h) + p_{n-1}(t)p_1(h) + o(h) \\
&\overset{(2)}{=} p_n(t)(1-\lambda h) + p_{n-1}(t)\lambda h + o(h)
\end{aligned}$$

$$\Rightarrow p_n'(t) = \lim_{h\to 0} \frac{p_n(t+h)-p_n(t)}{h} = -\lambda p_n(t) + \lambda p_{n-1}(t), \quad n = 1, 2, \ldots$$

mit $p_0(t) = e^{-\lambda t}$, $p_n(0) = P(N(0)=n) = 0$.
Man verifiziert nun leicht, dass die Wahrscheinlichkeitsfunktion der Poisson-Verteilung (eindeutige) Lösung dieser Differentialgleichung ist. □

- Der Poisson-Prozess ist ein homogener (stationärer) Markov-Prozess, d.h. für $s_0 < ... < s_n < s < t$ und $j \geq i$ gilt

$$
\begin{aligned}
P(N(t) = j \mid N(s) = i, N(s_n) = i_n, \dots, N(s_1) = i_1) &\overset{\text{ME}}{=} P(N(t) = j \mid N(s) = i) \\
&\overset{\text{Hom}}{=} P(N(t-s) = j - i) \\
&\overset{(\star)}{=} \frac{(\lambda(t-s))^{j-i}}{(j-i)!} e^{-\lambda(t-s)}
\end{aligned}
$$

- Der Beweis benutzt nur die Unabhängigkeit [und Stationarität] der Zuwächse.
$\Rightarrow$ Jeder Prozess mit unabhängigen [und stationären] Zuwächsen ist ein [homogener] Markov-Prozess.

Beweis:

$$P\{N(t) = j \mid N(s) = i, N(s_n) = i_n, \ldots, N(s_0) = i_0\}$$

$$\begin{aligned}
&= P\{N(t) - N(s) = j - i \mid N(s) - N(s_n) = i - i_n, \ldots, N(s_0) - N(0) = i_0\} \\
&\overset{(1)}{=} P\{N(t) - N(s) = j - i\} \\
&= P\{N(t) - N(s) = j - i \mid N(s) - N(0) = i\} \\
&= P\{N(t) = j \mid N(s) = i\} \quad \text{(Markoveigenschaft)} \\
&\overset{(1)}{=} P\{N(t-s) - N(0) = j - i\} \quad \text{(Homogenität).}
\end{aligned}$$

Aus ($\star$) folgt die Formel für die Übergangswahrscheinlichkeit. $\square$

- N ist Poisson–Prozess mit Rate $\lambda \Leftrightarrow$ Die Zwischenzeiten T_n sind i.i.d. Ex(λ) verteilt (exponentialverteilt mit Parameter λ).

Beweis:

Wir zeigen nur die "'$\Rightarrow$"' Richtung. Es gilt

$$P\{T_1 > t\} = P\{N(t) = 0\} \overset{(\star)}{=} e^{-\lambda t},$$

d.h. $T_1 \sim$ Ex(λ). Wegen Unabhängigkeit, Stationarität der Zuwächse und ($\star$) gilt

$$\begin{aligned}
P\{T_2 > t \mid T_1 = s\} &= P\{N(s+t) - N(s) = 0 \mid T_1 = s\} \\
&\overset{(1)}{=} P\{N(s+t) - N(s) = 0\} \\
&\overset{(1)}{=} P\{N(t) = 0\} \overset{(\star)}{=} e^{-\lambda t}
\end{aligned}$$

unabhängig von s, also T_1, T_2 i.i.d. Ex(λ). Analog für $n \geq 2$. $\square$

- Für die Wartezeit auf das n-te Ereignis $S_n = T_1 + \ldots + T_n$ gilt die Gammaverteilung

$$\begin{aligned} S_n &\sim Ga(n, \lambda) \\ f_{S_n}(t) &= \frac{e^{-\lambda t} \lambda^n t^{n-1}}{(n-1)!} \end{aligned}$$

mit

$$\begin{aligned} E(S_n) &= \frac{n}{\lambda}, \\ Var(S_n) &= \frac{n}{\lambda^2}, \\ Modus(S_n) &= \frac{n-1}{\lambda}. \end{aligned}$$

- Vorwärtsrekurrenzzeit:

$$V(t) = S_{N(t)+1} - t$$

$$P(V(t) \leq x) = F_{V(t)}(x) = 1 - e^{-\lambda x}$$

d.h. $V(t) \sim Ex(\lambda)$.

Bemerkung: $V(t) \sim Ex(\lambda)$ unabhängig von gewähltem t folgt aus "Gedächtnislosigkeit" der Exponentialverteilung:

$$X \sim Ex(\lambda) \Rightarrow P(X \leq t + x \mid X > t) = P(X \leq x).$$

- Rückwärtsrekurrenzzeit

$$
\begin{aligned}
U(t) &= t - S_{N(t)} \\
P(U(t) = t) &= e^{-\lambda t} \\
P(U(t) \leq x) &= 1 - e^{-\lambda x}, \quad 0 \leq x < t
\end{aligned}
$$

Bemerkung: $U(t) + V(t) = T_{N(t)+1} = S_{N(t)+1} - S_{N(t)}$

- Sampling Paradoxon: $N(t)$ in $S_{N(t)}$ ist zufälliger Index; $T_{(N(t)+1)} \neq T_n$, n fest.

Für fest vorgegebenes t gilt

$$
E(T_{N(t)+1}) = E(U(t)) + E(V(t)) = \frac{1}{\lambda}(1 - e^{-\lambda t}) + \frac{1}{\lambda} > \frac{1}{\lambda} = E(T_n)
$$

Plausibilitätserklärung: Bei fest vorgegebenem t ist $T_{N(t)+1}$ die zufällige Länge der enthaltenen Zwischenzeit. Im Schnitt werden längere Intervalle dabei favorisiert.

- Überlagerung von Poisson-Prozessen

$$\left.\begin{array}{ll} L = \{L(t), t \geq 0\} & \text{PP mit Rate } \lambda \\ M = \{M(t), t \geq 0\} & \text{PP mit Rate } \mu \end{array}\right\} \text{unabhängig}$$

Dann heißt $N = \{N(t), t \geq 0\}$ mit

$$N(t) = L(t) + M(t)$$

Überlagerung von L und M. Es gilt: N ist Poisson-Prozess mit Rate $\lambda + \mu$.

- Zerlegung von Poisson-Prozessen:

  N Poisson-Prozess mit Rate λ. Bei Eintritt eines Ereignisses wird ein binomisches Experiment $P(X = 1) = p$, $P(X = 0) = 1 - p$, unabhängig von N, durchgeführt.

$$\begin{aligned} X = 1 \quad &: \quad \text{Typ–1 Ereignis} \Rightarrow \text{Zählprozess } M \\ X = 0 \quad &: \quad \text{Typ–0 Ereignis} \Rightarrow \text{Zählprozess } L \end{aligned}$$

  Es gilt: M und L sind unabhängige PP mit Raten λp und $\lambda(1 - p)$.

- Typisches Beispiel: Netzwerke, etwa Straßensystem.

- Verallgemeinerung: Räumlicher Poisson-Prozess
  - Anzahl der Ereignisse in einem Gebiet A ist Poisson-verteilt mit Erwartungswert $\lambda \cdot$ Fläche(A).
  - Anzahl der Ereignisse in zwei nicht überlappenden Gebieten A_1 und A_2 sind unabhängig.

- Erweiterung I: Inhomogener Poisson-Prozess

  Ein Zählprozess $N = \{N(t), t \geq 0\}$ heißt nichtstationärer (inhomogener) Poisson-Prozess mit Rate (Intensitätsfunktion) $\lambda(t)$, $t \geq 0$:$\Leftrightarrow$

  (1) N hat unabhängige Zuwächse

  (2) $P(N(t+h) - N(t) \geq 2) = o(h)$
  $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$

Es lässt sich zeigen:

$$
\begin{aligned}
P(N(t+s) - N(t) = n) &= \exp\left(-\int\limits_t^{t+s} \lambda(u)du\right) \frac{\left(\int\limits_t^{t+s} \lambda(u)du\right)^n}{n!} \\
&= \exp(-(\Lambda(t+s) - \Lambda(t)) \frac{(\Lambda(t+s) - \Lambda(t))^n}{n!}
\end{aligned}
$$

mit $\Lambda(t) = \int\limits_0^t \lambda(u)du$ als kumulierte Rate, d.h.

$$
N(t+s) - N(t) \sim Po\left(\int\limits_t^{t+s} \lambda(s)ds\right)
$$

- Erweiterung II: Bewerteter (Compound) Poisson-Prozess

  N Poisson-Prozess, $\{Y_n, n \in \mathbb{N}\}$ von N unabhängige i.i.d. Folge.

  Dann heißt $X = \{X(t), t \geq 0\}$ mit

$$X(t) = \sum_{n=1}^{N(t)} Y_n$$

  bewerteter Poisson-Prozess.

- Beispiele:
  - Schadensprozesse

$$\begin{array}{rl} N(t) & \text{Prozess der Schadensanzahl,} \\ \{Y_n, n \in \mathbb{N}\} & \text{zugehörige Schadenshöhen,} \\ X(t) = \sum\limits_{n=1}^{N(t)} Y_n & \text{Gesamtschaden in } (0, t]. \end{array}$$

– Klassisches Versicherungsmodell von Cramér–Lundberg

$R(t) = c_0 + c_1 t - X(t)$	Risikoprozess,
c_1	Prämienintensität,
$X(t)$	bewerteter Poisson-Prozess mit Intensität λ,
Y_n i.i.d. $\sim F$	Schadenhöhen.

Ziel: Bestimmung der Wahrscheinlichkeiten

$$P(R(t) > 0 \ \forall t) = ?$$

oder

$$P(R(t) \leq 0) \leq ?$$