Stochastische Prozesse – Zusammenfassung

Einführung – Stochastische Prozesse

- $(X_t)_{t\in\mathcal{T}} = \{X_t : t \in \mathcal{T}\}$: stochastische Prozesse sind Familien von Zufallsvariablen, d.h. für jedes $t \in \mathcal{T}$ haben wir eine Zufallsvariable.
- $\mathcal{T} = \{\mathbf{0}, 1, 2, \ldots, N\}$, $\mathcal{T} = \mathbb{N}_0$ und $\mathcal{T} = [0, \infty)$: $\mathcal{T}$ ist die **Menge der betrachteten Zeitpunkte**.
- X_t: gibt Zustand eines Systems zum Zeitpunkt t an. Sind i.d.R. untereinander **abhängig**.
- t: ist ein Zeitpunkt in $\mathcal{T} = [0, \infty)$, kann aber auch ein Arbeitsschritt o. ä. sein.
- $x_0, x_1, x_2, \ldots$: Zustände, **konkrete Realisation** des **Prozesses.**
- $S = \{\mathbf{1}, 2, \ldots, N\}$**: Zustandsraum**, gibt die **Menge** der **möglichen Zustände** des **Systems** an. Alle X_t Werte nehmen einen solchen Zustand an.
- **Trajektorie**: (besonders im stetigen Fall) eine Realisierung des gesamten Prozesses.

Gibt vier verschiedene Arten von stochastischen Prozessen:

- Zeitdiskret und **zustandsdiskret**
- Zeitdiskret und zustandskontinuierlich
- Zeitkontinuierlich und **zustandsdiskret**
- Zeitkontinuierlich und zustandskontinuierlich

Markov-Ketten

Allgemeines:

- (Diskrete) Markov-Kette: ein zeitdiskreter stochastischer Prozess $(X_t)_{t\in\mathbb{N}_0} = (X_n)_{n\in\mathbb{N}_0}$ mit diskretem Zustandsraum S.
- $\mathrm{P}(\mathrm{X_{t+1}} = \mathrm{x_{t+1}} | \mathrm{X_t} = \mathrm{x_t}, \mathrm{X_{t-1}} = \mathrm{x_{t-1}}, \ldots, \mathrm{X_0} = \mathrm{x_0}) = \mathrm{P}(\mathrm{X_{t+1}} = \mathrm{x_{t+1}} | X_t = \mathrm{x_t})$: **Der Prozess hat kein Gedächtnis**.
- Markov-Kette heisst homogen, wenn für alle t gilt: $P(\mathrm{X_{t+1}} = \mathrm{j} | \mathrm{X_t} = \mathrm{i}) = \mathrm{P}(\mathrm{X_1} = \mathrm{j} | \mathrm{X_0} = \mathrm{i})$. D. h. Übergangswahrscheinlichkeiten hängen nicht von der Zeit ab.
- $\mathrm{p}_{ij} = \mathrm{P}(\mathrm{X_{t+1}} = \mathrm{j} | \mathrm{X_t} = \mathrm{i}) = \mathrm{P}(\mathrm{X_1} = \mathrm{j} | \mathrm{X_0} = \mathrm{i})$: Ein-Schritt-Übergangswahrscheinlichkeit vom i-ten zum j-ten Zustand mit p_{ij}: $0 \leq \mathrm{p}_{ij} \leq 1$
- $\boldsymbol{P}$: Übergangsmatrix $P = \begin{pmatrix} p_{11} & \cdots & p_{1N} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NN} \end{pmatrix}$, Zeilensummen müssen sich zu 1 addieren → stochastische Matrix.
- Übergangsdiagramm/-graphen oder Zustandsraumdarstellung: Graph mit Zuständen und Pfeilen als Übergangswahrscheinlichkeit

Zustandsverteilungen:

- Interesse: mit welcher **Wahrscheinlichkeit** sich **Prozess** zu einem **Zeitpunkt** im **Zustand $\boldsymbol{j}$ befindet**, ohne dass man den vorherigen Zustand kennt.
- $\vec{\pi}(t)$ bezeichnen wir auch als **Zustandsverteilung** zum **Zeitpunkt** t bzw. als **Zustandsvektor**.
- $\pi_i(t) = P(\mathrm{X_t} = \mathrm{i})$: Wahrscheinlichkeit, mit der sich der Prozess zum Zeitpunkt t im Zustand i befindet. $\pi_1(t), \ldots, \pi_N(t)$ wird Zeilenvektor $\vec{\pi}(t) = \big(\pi_1(t), \ldots, \pi_N(t)\big)$
- $\vec{\pi}(t)$: **Zustandsverteilung** zum **Zeitpunkt** t bzw. als **Zustandsvektor**, diese **hängt** vom **Startzustand ab.**
- $\vec{\pi}(0) = \big(\pi_1(0), \pi_2(0), \ldots, \pi_N(0)\big)$: Anfangs-/Startverteilung
- $\boldsymbol{P}^t$: t-Schritt-Übergangsmatrix, gibt Übergangswahrscheinlichkeiten zum Zeitpunkt t wieder. $\boldsymbol{P}^0 = \boldsymbol{I}$ Einheitsmatrix
- $\vec{\pi}(t) = \vec{\pi}(0) \cdot \boldsymbol{P}^t$: Gibt Wahrscheinlichkeitsverteilung zum Zeitpunkt t wieder. → %% führt zu Multiplikation/Potenzierung gemäss LA!
- **library(expm); P <- rbind(c(0.1, 0.9, 0), c(0.7, 0, 0.3), c(1, 0, 0)); P5 <- P %^% 5; t(c(1,0,0)) %*% P5**

Klassifikation von Zuständen:

- $i \to j$: j ist von i direkt oder indirekt erreichbar.
- $i \nrightarrow j$: j ist von i direkt oder indirekt **nie** erreichbar.
- $i \leftrightarrow j$: j, i sind wechselseitig erreichbar.
- $p_{ii} = 1$: **absorbierender** Zustand, man verlässt diesen Zustand nie wieder, wenn man ihn einmal erreicht hat. Dieser bildet eigene Klasse und ist rekurrent/reduzibel.
- Zustände lassen sich eindeutig in Klassen unterteilen. i, j liegen genau dann in derselben Klassen, wenn $i \leftrightarrow j$ gilt.
- **Irreduzibilität**: Zustandsraum S besteht nur aus einer Klasse.
- **Reduzibel**: wenn für i, j gilt $i \nrightarrow j$.
- **Rekurrent**: Zustand j ist rekurrent, wenn man bei Start in j sicher wieder nach j zurückkehrt.
- **Transient:** Zustand j ist transient, wenn man bei Start in j mit positiver Wahrscheinlichkeit nicht wieder nach j zurückkehrt.
- **Klasseneigenschaften:** Rekurrent und transient gilt jeweils für die gesamte Klasse.
- R_j: heisst Rückkehrzeit für den Zustand j.
- **Periode eines Zustandes:** Anzahl Schritte, nachdem minimal eine Rückkehr in einen Zustand stattfindet. Alle Zustände in einer Klasse haben dieselbe Periode.
- **Aperiodisch:** wenn Zustand i $p_{ii} > 0$ ist.
- $p_{ii} = 0$: kann periodisch oder aperiodisch sein.
- **Periodisch:** Zustand i braucht $1 > d$ Schritte für die Rückkehr und es muss gelten $p_{ii} = 0$. **summary(kette)**
- **library(markovchain); P <- rbind(c(0.1, 0.9, 0), c(0.7, 0, 0.3), c(1, 0, 0)); (kette <- new("markovchain", states = c("A", "B", "C"), transitionMatrix = P))**

Aufenthaltsdauern für endlichen Zeithorizont:

- Sind an **Aufenthaltsdauer** in einem **bestimmten Zustand** interessiert. → **Wie oft … im Mittel …**
- Hängt von Startverteilung $\vec{\pi}(0)$ ab (reduzible Ketten)
- $Z_j(T)$: Bezeichnet die Anzahl der Aufenthalte im Zustand j im Zeitraum $t = 0, \dots, T$
- **Erwartungswert** der **Aufenthalte** $E\left(Z_j(T)\right) = \vec{\pi}(0) \cdot \left(\sum_{t=0}^{T} \boldsymbol{P}^t\right) \cdot \vec{e}_j' = \vec{\pi}(0) \cdot \boldsymbol{M}(T) \cdot \vec{e}_j'$, wobei $\boldsymbol{M}(\boldsymbol{T})$ = **Aufenthaltsdauermatrix**
- Die **Zeilen**summen der Aufenthaltsmatrix sind $T + 1$, also $\sum_{j=1}^{N} m_{ij} = T + 1$.
- $\vec{N} = \vec{\pi}(0) \cdot \boldsymbol{M}(T)$: Vektor der erwarteten Aufenthaltsdauern, Summe seiner Komponenten $= T + 1$
- **Beispiel:** Falls Maschine zum Zeitpunkt t in Ordnung ist, ist sie das auch zum Zeitpunkt $t + 1$ mit Wahrscheinlichkeit von 90 %. Eine defekte Maschine ist beim nächsten Zeitpunkt mit Sicherheit wieder repariert. Die Maschine funktioniere für $t = 0$. Wir sind interessiert an der erwarteten Anzahl Ausfalltage im Zeitraum $\mathcal{T} = \{0,1,2,3,4\}$
- **library(expm); P <- rbind(c(0.9,0.1), c(1,0)); M4 <- diag(2); for (k in 1:4) M4 <- M4 + (P %^% k); M4; t(c(1,0)) %*% M4 %*% (c(0,1)); rowSums(M4)** # = 4+1

- **Beispiel:** Berechnen der erwarteten Anzahl Aufenthalte in Zustand zwei. Es werden nur Aufenthalte zu den Zeiten $t = 0,1,3,5,6$ gezählt. Start $t = 0$ im Zustand 2.
- Matrix gegeben: **(P <- rbind(c(0.12, 0.88), c(0.8, 0.2))); (M <- diag(2))** #$\boldsymbol{M}^0 = \boldsymbol{I}$; **for(i in c(1,3,5,6)) M <- M + (P %^% i); M; t(c(0, 1)) %*% M %*% c(0, 1)**

Kostenmodelle für endlichen Zeithorizont

Zustandsabhängige Kosten

- Zu jedem Zeitpunkt $t = 0,1,2, \dots, N$ fallen Kosten K_t an, die nur vom Zustand des Prozesses zur Zeit t abhängen.
- $\vec{c} = (c_1, c_2, \dots, c_N)$: Kostenvektor für jeden Zustand j des Zustandsraums S.
- K: Gesamtkosten aus Summe der Einzelkosten.
- $E(K_t) = \vec{\pi}(0) \cdot \boldsymbol{P}^t \cdot \vec{c}'$: **erwartete Kosten** zum **Zeitpunkt** t.
- $E(K) = \vec{\pi}(0) \cdot \left(\sum_{t=0}^{T} \boldsymbol{P}^t\right) \cdot \vec{c}' = \vec{\pi}(0) \cdot \boldsymbol{M}(T) \cdot \vec{c}'$: erwartete **Gesamtkosten bis und mit zum Zeitpunkt** T.
- **Beispiel:** Eine Versicherung soll während einer Reisezeit von 7 Tagen 100 CHF pro Krankheitstag bezahlen. Die Bedingung ist, dass der Versicherungsnehmer gesund ist am Tage des Reiseantritts. Wie teuer muss die Police sein, damit die Versicherung keinen Verlust macht? Eine Person, die heute gesund ist, ist morgen krank mit Wahrscheinlichkeit 0.05. Eine kranke Person ist mit Wahrscheinlichkeit 0.6 am folgenden Tag wieder gesund. $S = \{1,2\}$, 1 = gesund, 2 = krank
- **library(expm); P <- rbind(c(0.95,0.05), c(0.6,0.4)); M4 <- diag(2); for (k in 1:6) M4 <- M4 + (P %^% k); M4; t(c(1,0)) %*% M4 %*% (c(0,100))**

Übergangsabhängige Kosten

- u_{ij}: Kosten u_{ij} entstehen beim Übergang von Zustand i in Zustand j
- $\boldsymbol{U}$: Übergangskostenmatrix $\boldsymbol{U} = \begin{pmatrix} u_{11} & \cdots & u_{1N} \\ \vdots & \ddots & \vdots \\ u_{N1} & \cdots & u_{NN} \end{pmatrix}$
- $\vec{c} = \left(\sum_{j=1}^{N} u_{1j}p_{1j}, \sum_{j=1}^{N} u_{2j}p_{2j}, \dots \sum_{j=1}^{N} u_{Nj}p_{Nj}\right) = \boldsymbol{P} \circ \boldsymbol{U}$: Kostenvektor, Komponenten von $\vec{c}$ sind die **Zeilensummen** der **komponentenweise** Multiplikation von $\boldsymbol{P}$ und $\boldsymbol{U}$.
- $E(K) = \vec{\pi}(0) \cdot \left(\sum_{t=0}^{T-1} \boldsymbol{P}^t\right) \cdot \vec{c}' = \vec{\pi}(0) \cdot \boldsymbol{M}(T-1) \cdot \vec{c}'$: erwartete **Gesamtkosten bis zum Zeitpunkt** T.
- **Beispiel:** Betrachten die gegebene Markov-Kette mit Zustandsraum $S = \{1,2,3\}$ Übergangsmatrix $\boldsymbol{P}$ (siehe unten). Übergang von Zustand 1 zu Zustand 3 koste 4 CHF. Übergang von Zustand 1 zu Zustand 2 und von Zustand 3 zu Zustand 1 koste 9 CHF. Der Übergang von Zustand 3 zu Zustand 2 koste 2 CHF. Die restlichen Übergänge sind kostenfrei. Was sind die erwarteten Kosten bis $T = 12$ bei Start in 3.
- **(P <- matrix(c(0, 0.7, 0.3, 0.6, 0, 0.4, 0.1, 0.85, 0.05), ncol = 3, byrow = T)); (U <- matrix(c(0, 9, 4, 0, 0, 0, 9, 2, 0), ncol = 3, byrow = T));**
- **(c_sch <- c(rowSums(P * U))); (M <- diag(3)) for (i in 1:(12-1)) M <- M + (P %^% i) ; M; t(c(0,0,1)) %*% M %*% c_sch**

Asymptotisches Verhalten:

- Sind am langfristigen Verhalten des Prozesses interessiert. Dies sind die stationären Eigenschaften nach dem Einschwingverhalten.
- **Asymptotische/eindeutige/ergodische Verteilung:** Gibt es eine Zustandsverteilung $\vec{\pi}(\infty)$, so dass unabhängig vom Anfangszustand gilt, dass: $\boldsymbol{\vec{\pi}(\infty) = \lim_{t\to\infty} \vec{\pi}_j(t)}$
- **Stationäre Verteilung:** Eine Zustandsverteilung $\vec{\pi}^*$ heisst **stationär**, wenn gilt: $\vec{\pi}^* = \vec{\pi}^* \cdot \boldsymbol{P}$**.** D.h. wenn **Verteilung** zu einem Zeitpunkt t **stationär** ist, so **ändert** sie sich in den **weiteren Schritten nicht mehr**.
- Da sich eine **asymptotische Verteilung** auch zumindest für sehr grosse t nicht mehr ändern darf, **müssen** asymptotische Verteilungen **stets stationär sein**.
- Jede Markov-Kette hat mindestens eine stationäre Verteilung. Diese ist dann ein potenzieller Kandidat für die asymptotische Verteilung.
- **Eindeutige Grenzverteilung:** $(X_t)_{t\in\mathbb{N}_0}$ besitzt eine eindeutige stationäre Verteilung und eine eindeutige Grenzverteilung, welche gleich der stationären Verteilung ist.
- Jede aperiodische und irreduzible Markov-Kette konvergiert also gegen ihre (eindeutige) stationäre Verteilung, und zwar unabhängig von ihrer Startverteilung.

Stationäre Verteilungen und Eigenwerte:

- jede **aperiodische irreduzible** Markov-Kette besitzt eine eindeutige Grenzverteilung
- Eine **stationäre Verteilung** $\vec{\pi}^*$ ist ein linker Eigenvektor zum Eigenwert 1.
- Für die stationäre Verteilung muss ein rechter Eigenvektor zum Eigenwert 1 von $\boldsymbol{P}'$ berechnet werden.
- Eigenvektor $(\vec{\pi}^*)'$ ist nur dann eine stationäre Verteilung, wenn $\sum_{i=1}^{N} \pi_i^* = 1$ und $\pi_i^* \geq 0$ für alle $i = 1, \dots, N$
- **Eigenvektor zum Eigenwert 1 (diskreter Fall)** muss deshalb überall dasselbe Vorzeichen haben!
- **P <- rbind(c(0.4, 0.6), c(0.9, 0.1)); eigen(t(P)); v <- eigen(t(P))$vectors[,1]; v <- v/sum(v)**
- **Reeller Anteil am Eigenvektor:** **eigen(t(P)); (v <- Re(eigen(t(P))$vectors[,1])); (v <- v/sum(v))**

Asymptotische Besetzung:

- Asymptotische Besetzung $\hat{\pi}_j = \lim_{T\to\infty} \frac{Z_j(T)}{T+1}$ also den Anteil der Besuche in Zustand j für die gesamte Realisation. $\hat{\pi}_j$ wird zu Vektor $\vec{\hat{\pi}}$
- Für aperiodische und irreduzible Markov-Ketten gilt mit Wahrscheinlichkeit 1 $\vec{\hat{\pi}} = \vec{\pi}(\infty)$ d.h. für fast jede Realisation des Prozesses ist die asymptotische Besetzung gleich der eindeutigen Grenzverteilung.
- Bei solchen Ketten gilt also, dass die Zustandshäufigkeiten einer Trajektorie (und zwar fast jeder) den Wahrscheinlichkeiten der Grenzverteilung entsprechen. Wir können also eine Betrachtung im Querschnitt (Beobachtung oder Simulation vieler Trajektorien und Auswertung an einem festen, weit in der Zukunft liegenden, Zeitpunkt t) durch eine Betrachtung im Längsschnitt, also Beobachtung einer einzigen (allerdings genügend langen) Trajektorie ersetzen.

Asymptotische Kostenmodelle:

- Betrachtung Kostenmodelle für Prozesse mit unendlicher Laufzeit.
- Interesse für durchschnittliche Kosten pro Zeitschritt

Asymptotische Durchschnittskosten

- Gehen von **eindeutigen Besetzungsverteilung** $\vec{\pi}$ aus, die **unabhängig** vom **Anfangszustand** ist.
- Wenn es **keine** eindeutige Besetzungsverteilung unabhängig vom Anfangszustand gibt, dann müssen die asymptotischen durchschnittlichen Kosten pro Einheit via Kostenmodelle für **endlichen Horizont** mit einer gegebenen Startverteilung und der $\boldsymbol{M}$ Matrix berechnet werden. **Siehe weiter oben/Beispiel Helpdesk AB4**
- **Durchschnittliche pro Zeitschritt anfallenden Kosten** $\widetilde{K} = \vec{\pi}(\infty) \cdot \vec{c}'$
- Berechnung der **erwarteten Kosten pro Zeitschritt** bei **übergangsabhängigen** Kosten: $E\left(\widetilde{K}\right) = \vec{\pi}(\infty) \cdot \vec{\tilde{c}}'$
- **Beispiel:** Haben gegebene Übergangsmatrix $\boldsymbol{P}$ und Übergangskostenmatrix $\boldsymbol{U}$ (siehe unten): Wie hoch sind (auf lange Sicht) die **durchschnittlichen** Kosten pro Tag?
- **(P <- rbind(c(0.92, 0.05, 0.03), c(0.05, 0.92, 0.03), c(0.05, 0.05, 0.90))); (U <- rbind(c(1000, 900, 800), c(900, 1000, 800), c(1300, 1300, 1500)));**
- **(es <- eigen(t(P))); (pi_stat <- es$vectors[, 1]/sum(es$vectors[, 1])); (c_sch <- c(rowSums(P * U))); t(pi_stat) %*% c_sch**

Diskontierte Gesamtkosten

- Keine Voraussetzung aperiodisch und irreduzibel! → Diskontierung gemäss Modul FUM… #viel schlimm…
- Betrachtung von zustandsabhängigen Kosten.
- K_α: Gesamtkosten mit Diskontierungsfaktor $\alpha \in (0,1)$
- $E(K_\alpha) = \vec{\pi}(0) \cdot \left(\sum_{t=0}^{\infty} \alpha^t \cdot \boldsymbol{P}^t\right) \cdot \vec{c}'$: Erwartungswert mit Startverteilung | $E(K_\alpha) = \vec{\pi}(0) \cdot (\boldsymbol{I} - \alpha \cdot \boldsymbol{P})^{-1} \cdot \vec{c}'$: die **erwarteten diskontierten Gesamtkosten**.
- **P <- rbind(c(1, 0, 0), c(0.35, 0.55, 0.1), c(0.3, 0.1, 0.6)); cost <- c(0, 600, 720); pi_0 <- c(0,1,0); (clv <- t(pi_0) %*% solve(diag(3) - 0.95*P) %*% cost)**

Customer Lifetime Value

- CLV = wichtige Kenngrösse, die messen soll, wie viel Umsatz/Gewinn ein Kunde in der Zukunft noch beitragen wird.
- Geht davon aus, dass ein Kunde in der nächsten Periode mit Wahrscheinlichkeit $r \in (0,1)$ immer noch Kunde ist und mit Wahrscheinlichkeit $1 - r$ kündigt.
- r: retention rate
- Kunde, der einmal gekündigt hat, kann nicht wieder zurückgewonnen werden.
- Verhalten des Kunden kann durch Markov-Kette mit Zustandsraum $S = \{1,2\}$ modelliert werden, wobei 1 = Person noch Kunde, 2 = kein Kunde mehr.
- Die Übergangsmatrix ist: $\boldsymbol{P} = \begin{pmatrix} r & 1-r \\ 0 & 1 \end{pmatrix}$
- C: erhalten festen Betrag (Gewinn/Umsatz) pro Periode vom Kunden
- $\alpha \in (0,1)$: Diskontierungsfaktor $\alpha = \frac{1}{1+z}$, wobei z = Inflationsrate, Zinssatz, o.ä.
- CLV eines Kunden ist erwartete diskontierte Gewinn $CLV = C \cdot \frac{1}{1-\alpha r}$. CLV ist auch dann definiert, wenn entweder r oder α gleich 1 ist, aber nicht wenn beide 1 sind.
- CLV wird gross, wenn $(1 - \alpha r)$ klein
- **Beispiel:** Kunden A bringt uns p.a. CHF 600 CHF ein mit 30 % p.a. Kündigungswahrscheinlichkeit. Der zweite Kunde B bringt uns CHF 720 p.a. ein, Kündigung 36 %.
- **Erwartete Gesamtgewinne ohne Diskontierung:** $CLV_A = C \cdot \frac{1}{1-\alpha r} = 600\, CHF \cdot \frac{1}{1-1\cdot(1-0.3)} = \frac{600\, CHF}{0.3} = 2000\, CHF$, $CLV_B = 720\, CHF \cdot \frac{1}{1-1\cdot(1-0.36)} = \frac{720\, CHF}{0.36} = 2000\, CHF$
- **Erwartete Gesamtgewinne mit Diskontierung 0.95:** $CLV_A = 600\, CHF \cdot \frac{1}{1-0.95\cdot(1-0.3)} = \frac{600 CHF}{0.335} = 1791.04\, CHF$, $CLV_B = 720\, CHF \cdot \frac{1}{1-0.95\cdot(1-0.36)} = \frac{720 CHF}{0.392} = 1836.73 CHF$

Markov Chain Monte Carlo

- Markov-Kette durch Definition der Übergangsmatrix $\boldsymbol{P}$ zu konstruieren, so dass diese die gewünschte Zielverteilung als eindeutige Grenzverteilung $\vec{\pi}^*$ hat.
- Eine Trajektorie als Stichprobe aus der Verteilung $\vec{\pi}^*$ nehmen.
- Gesucht ist beliebige Verteilung $\vec{\pi}$, sodass $\vec{\pi} \cdot \boldsymbol{P} = \vec{\pi}$ gilt. Markov-Kette ist damit aperiodisch und irreduzibel. Dies garantiert, dass $\vec{\pi}$ eine eindeutige Grenzverteilung ist.
- Eine stationäre Zustandsverteilung ist nun eine, die sich nach einem Übergang nicht mehr ändert – d.h. es müssen in jedem Schritt genauso viele Teilchen aus jedem Zustand heraus- wie hereinspringen.
- Detailed-Balance-Bedingung: Es springen nicht nur aus jedem Zustand genauso viele Teilchen hinaus wie hinein, sondern es gilt für alle Paare von Zuständen $i \neq j$, dass in jedem Schritt genauso viele Teilchen von i nach j springen wie von j nach i. Es gilt $\pi_i p_{ij} = \pi_j p_{ji}$
- Konstruieren nun Übergangswahrscheinlichkeiten $\boldsymbol{Q}$ mit Erfüllung der Detailed-Balance-Bedingung. → symmetrische Übergangsmatrix, irreduzibel.
- Grundidee ist, jeweils basierend auf den Übergangswahrscheinlichkeiten aus $\boldsymbol{Q}$ den nächsten Zustand vorzuschlagen, aber den Sprung nur mit einer bestimmten Wahrscheinlichkeit auszuführen (und ansonsten im aktuellen Zustand zu bleiben). Befinden wir uns nun im Zustand i, so schlägt die Markov-Kette $\boldsymbol{Q}$ mit Wahrscheinlichkeit q_{ij} den Zustand j vor. Mit Wahrscheinlichkeit
  - α_{ij} nehmen wir diesen Vorschlag an und springen nach j,
  - $1 - \alpha_{ij}$ nehmen wir ihn nicht an, und bleiben in i.

  Für $i \neq j$ ist die (bedingte) Wahrscheinlichkeit für einen Sprung von i nach j also: $p_{ij} = \alpha_{ij} \cdot q_{ij}$. Argument gilt $\alpha_{ij} = \min\left\{1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}}\right\}$

Metropolis-Hastings-Algorithmus: Eine Markov-Kette mit gewünschter stationärer Verteilung $\vec{\pi}$ kann man wie folgt implementieren: (**siehe auch AB8**)

- Initialisiere x_0 beliebig mit einem der Zustände aus S.
- Generiere ausgehend von x_t einen Vorschlag x^* aus der Markov-Kette mit Übergangsmatrix $\boldsymbol{Q}$.
- Setze mit Wahrscheinlichkeit $\min\left\{1, \frac{\pi_{x^*} \cdot q_{x^*, x_t}}{\pi_{x_t} \cdot q_{x_t, x^*}}\right\} x_{t+1} = x^*$, ansonsten $x_{t+1} = x_t$. Wiederhole Schritte 2 und 3, bis die Realisation der Kette die gewünschte Länge T hat.
- Wenn $\boldsymbol{Q}$ symmetrisch, dann gilt $\alpha_{ij} = \min\left\{1, \frac{\pi_j}{\pi_i}\right\}$
- **Beispiel:** Gegeben sind $\boldsymbol{Q} = \begin{pmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{pmatrix}$ (ist symmetrisch) und stationäre Verteilung $\vec{\pi} = c \cdot (1,2,4,8)$. $\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0.5 & 0 & 1 & 0 \\ 0 & 0.5 & 0 & 1 \\ 0.125 & 0 & 0.5 & 0 \end{pmatrix}$
- $\boldsymbol{P} = \boldsymbol{A} \circ \boldsymbol{Q} = \begin{pmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.25 & 0 & 0.5 & 0 \\ 0 & 0.25 & 0 & 0.5 \\ 0.0625 & 0 & 0.25 & 0 \end{pmatrix}$ komponentenw. Multipliz. Diagonalelemente von $\boldsymbol{P}$ ergänzen, sodass Zeilensumme = 1 ist. $\boldsymbol{P} = \begin{pmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 & 0 \\ 0 & 0.25 & 0.25 & 0.5 \\ 0.0625 & 0 & 0.25 & 0.6875 \end{pmatrix}$
- $\vec{\pi} \cdot \boldsymbol{P} = \vec{\pi} \Rightarrow (1,2,4,8) \cdot \boldsymbol{P} = (1,2,4,8)$ wir haben das Ziel somit erreicht.
- **Beispiel:** Metropolis-Hastings-Algorithmus benutzen, um aus einer A-Posteriori-Verteilung zu simulieren. Anzahl stammt aus Poisson-Verteilung, wobei $\lambda > 0$ von Gamma-Verteilung mit Parameter $\alpha = 6, \beta = 0.4$ annehmen. Gehen Sie davon aus, dass k = 5 Teilchen beobachtet wurden. Vorschlagsdichte aus Standardnorm:

```
posterior <- function(lambda, alpha, beta, k){
  ifelse(lambda > 0, dpois(x = k, lambda = lambda) *
  dgamma(x = lambda, shape = alpha, scale = beta), 0)}
```

```
k <- 5;  alpha <- 6;  beta <- 0.4
vorschlag <- function(x){ rnorm(1, mean = x, sd = 1) }
mcmc <- function(nsim, x_0) {
  trajektorie <- vector("numeric", length = nsim);  x_current <- x_0
  for (i in 1:nsim) {
    trajektorie[i] <- x_current;  x_proposal <- vorschlag(x_current)
    alpha_akzeptanz <- posterior(x_proposal, alpha = alpha,
    beta = beta, k = k) / posterior(x_current, alpha = alpha,
    beta = beta, k = k)
    if (runif(1) < alpha_akzeptanz) x_current <- x_proposal }
trajektorie}
mcmc_ergebnis <- mcmc(nsim = 10000, x_0 = 1)[5001:10000]
```

Zeitkontinuierliche Markov-Prozesse

- Gleiche Definition wie bei diskreten Markov-Prozesse. Zeit ist nun stetig. Zustandsraum ist weiterhin diskret und endlich S.
- Prozess ist wiederum gedächtnislos.
- T_i: Wartezeit, das System bleibt eine gewisse Wartezeit T_i im Zustand i, dann springt es in neuen Zustand k, usw.
- $\boldsymbol{P}$: Übergansmatrix mit Übergangswahrscheinlichkeiten, Zeilensumme ist gleich 1, **aber** Hauptdiagonalen p_{ii} sind immer $p_{ii} = 0$.
- λ_i: Wartezeitparameter, ist Kehrwert der mittleren Wartezeiten $\lambda_i = \frac{1}{E(T_i)}$, Einheit von λ ist $\frac{1}{Zeit}$

Aufenthaltsdauer- und Wartezeitverteilung eines Markov-Prozess

- Gleiche Definition wie bei diskreten Markov-Prozesse. Zeit ist nun stetig. Zustandsraum ist weiterhin diskret und endlich S.
- Verteilung der Aufenthaltsdauer ist Verteilungsfunktion der Exponentialfunktion $Exp\left(\frac{1}{\lambda}\right)$ $\quad F(t) = 1 - e^{-\lambda t}, falls\ t \geq 0,\ sonst\ 0$
- Die nach einer Wartezeit noch verbleibende Wartezeit hat die selbe Verteilung wie die Wartezeit selbst.

Raten- und Generatorenmatrizen

- $\boldsymbol{R}$: Ratenmatrix, Zusammenfassung von Rate λ und Übergangswahrscheinlichkeit $\boldsymbol{P}$ zu einer Matrix $\boldsymbol{R} = (r_{ij}) = \boldsymbol{p}_{ij} \cdot \lambda_i$. Hauptdiagonalen = 0, $r_{ii} = 0$
- Mit der Angabe der Ratenmatrix ist der Markov-Prozess vollständig bestimmt. → Zeilenweise Multiplikation mit λ_i **(R <- P * lambda)**
- Gibt wieder Ratendiagramm mit Angabe der entsprechenden r_{ij}.
- $\boldsymbol{Q}$**:** Generatorenmatrix besteht aus Ratenmatrix $\boldsymbol{R}$ und zusätzlich werden Hauptdiagonalelemente aus $\boldsymbol{R}$ durch die negative Zeilensumme ergänzt. $\mathbf{Q} = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$
- Damit hat die Generatormatrix die Eigenschaft, dass in jeder Zeile die Summe der Matrixelemente gerade 0 ergibt.
- **(R <- rbind(c(0, 0.4, 0), c(0.5, 0, 0.4), c(0, 1, 0))); (Q <- R - diag(rowSums(R))); rowSums(Q) #Kontrolle Summe = 0**

Zustandsverteilungen

- interessieren uns für die zeitabhängige Zustandsverteilung $\vec{\pi}(t)$.
- $\vec{\pi}(t) = \vec{\pi}(0) \cdot \exp(\boldsymbol{Q} \cdot t)$ → **library(expm); Q <- Q; t(c(1/3, 1/3, 1/3)) %*% expm(Q*t)** Funktion expm() in R, nicht exp().

Grenzwertverhalten

- Interessieren uns für asymptotische Verteilung, Langzeitverhalten
- Eindeutige asymptotische Verteilung gilt, wenn Markovprozess irreduzibel ist.
- $\vec{\pi}^* \cdot \boldsymbol{Q} = \vec{0}$: dieses lin. GS muss gelöst werden. So erhalten wir asymptotische Verteilung.
- Zweite Variante ist die Eigenwertanalyse. Eigenvektor von $\boldsymbol{Q}'$ zum Eigenwert **0**. → normieren zu Summe 1. Eigenwert zur Generatorenmatrix ist höchsten 0.
- Eigenwertspektrum von $\boldsymbol{Q}'$ erlaubt Abschätzung, wie schnell der stationäre Zustand erreicht wird. Dies wird durch den (negativen) Eigenwert $\tilde{\lambda}$ beeinflusst, der am nächsten bei 0 liegt. Er liefert einen Anteil, der wie $\exp(-|\tilde{\lambda}|t)$ abfällt und nach einer Faustregel nach einer Zeit $5/|\tilde{\lambda}|$ vernachlässigt werden kann.
- **(Q <- rbind(c(-3/5, 3/5, 0), c(0.5, -1.5, 1), c(0, 2, -2))); eigen(t(Q)); (v <- eigen(t(Q))$vectors[,3]); (v <- v/sum(v)); sum(v) #Kontrolle Summe = 1**

Kostenmodelle asymptotisch

- Analog zum diskreten Fall.
- $E\left(\frac{dK_t}{dt}\right) = \vec{\pi}^* \cdot \vec{c}'$: mittlere Kosten pro Zeiteinheit K_t, Einheit Fr/Zeit.
- $\boldsymbol{U}$**:** Übergangskostenmatrix, übergangsabhängigen Kosten.
- $E\left(\frac{dK_t}{dt}\right) = \vec{\pi}^* \cdot \vec{\tilde{c}}'$, wobei $\vec{\tilde{c}} = R \circ U = (r_{ij}u_{ij})$, komponentenweise Multiplikation von R mit U, dann von Matrix Summe jeder Zeile bilden → ergibt Spaltenvektor.
- **Q <- matrix(c(-5, 1, 3, 1, 10, -11, 0, 1, 0, 0, -3, 3, 5, 5, 5, -15), byrow=T, ncol=4); es <- eigen(t(Q)); ev <- es$vectors[,4]; piStern <- ev / sum(ev)**
- **U <- matrix(1, nrow=4, ncol=4); diag(U) <- 0; R <- Q; diag(R) <- 0 ; piStern %*% rowSums(R * U) #Kosten**
- **Asymptotische Gewinn/Verlustrate: U <- matrix(0, nrow=4, ncol=4) U[,2] <- c(100, 0,100,100) U[,3] <- c(-100, -100, 0,-100); piStern %*% rowSums(R * U)**

Punkt-/Zählprozesse

- Prozesse, die Abfolge von einzelnen Ereignissen beschreiben, auch Strom von Ereignissen genannt.
- Ereignis ist durch seinen Eintrittszeitpunkt gekennzeichnet. Jedes Ereignis hat genau einen solchen Zeitpunkt.
- Werden Aussagen über Zeit zwischen zwei aufeinanderfolgenden Ereignissen machen, sog. Zwischenankunftszeiten = Wartezeit.
- S_n: Eintrittszeitpunkt des n-ten Ereignisses. Setzen weiterhin $S_0 = 0$
- $T_n = S_n - S_{n-1}$: Wartezeit zwischen den Ereignissen.
- $N(t)$: Anzahl der Ereignisse im Intervall $(0, t]$, sog. Zählprozess. Dieser ist ganzzahlig und monoton steigend.
- Es gilt folgende Hierarchie:
- **Punktprozesse:** Werden durch eine Folge von zufälligen Wartezeiten beschrieben.
  - **Erneuerungsprozesse:** Die Wartezeiten sind iid.
    - **Poisson-Prozesse:** Die Wartezeiten sind iid-exponentialverteilt.

Poisson-Prozesse

- Die Wartezeiten (Zeiten zwischen Ereignissen) sind iid-exponentialverteilt. → Gedächtnislosigkeit
- T: Wartezeit, ohne Index T_i, da diese aus derselben Verteilung stammen.
- **Verteilungsfunktion** (Exp(λ)-Verteilung) ist: $F_T(t) = P(T \leq t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$ Ihre Ableitung, die **Dichtefunktion** ist: $f_T(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$
- Der **Erwartungswert** ist $E(T) = \int_{-\infty}^{\infty} t f(t) dt \ = \frac{1}{\lambda}$
- Verteilung $N(t)$ mit Parameter λt: $P(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$
- **Beispiel:** Serverzugriffe mit Rate von 1/6 pro Sekunde. Server stürzt ab, wenn mehr als 3 Zugriffe pro Sekunde erfolgen. Wahrscheinlichkeit Absturz innerhalb nächsten Tages: $P(N(1) > 3) = 1 - \sum_{k=0}^{3} P(N(1) = k) = 1 - \sum_{k=0}^{3} e^{-\lambda t} \frac{(\lambda t)^k}{k!} = 1 - \sum_{k=0}^{3} e^{-\frac{1}{6} \cdot 1} \frac{\left(\frac{1}{6} \cdot 1\right)^k}{k!} = 1 - \sum_{k=0}^{3} e^{-\frac{1}{6}} \frac{\left(\frac{1}{6}\right)^k}{k!} = 2.814753 \ \cdot \ 10^{-5}$
  Nun Wahrscheinlichkeit keinen Absturz in den nächsten 24 Stunden: $24 \cdot 60 \cdot 60 = 86\,400\ s$ $P(Absturz\ innerhalb\ 24\ h) = 1 - P(kein\ Absturz\ innerhalb\ 24\ h)$
  $= 1 - \left(P(kein\ Absturz\ innerhalb\ 1\ s)\right)^{86400} = 1 - \left(1 - P(kein\ Absturz\ innerhalb\ 1\ s)\right)^{86400} = 1 - (1 - 2.814753 \ \cdot \ 10^{-5})^{86400} = 0.9121373$
- $\lambda = \lambda_1 + \lambda_2$: Zwei unabhängige Poisson-Prozesse können addiert werden.
- **Beispiel:** An Haltestelle treffen Busse und S-Bahnen ein. Die Ankunftsprozesse können als Poisson-Prozesse mit Raten $0.3/min$ und $0.1/min$ modelliert werden. Wie lange muss ein Fahrgast im Mittel an der Haltestelle warten? $\lambda = \lambda_1 + \lambda_2 = \frac{0.1}{min} + \frac{0.3}{min} = \frac{0.4}{min}$ → $E(X) = \frac{1}{\lambda} = \frac{1}{0.4} = 2.5\ min$
  Wie hoch ist die Wahrscheinlichkeit, dass er länger als 1 Minute warten muss? $P(N(t) = k) = (N(1) = 0) = e^{-\lambda t} \frac{(\lambda t)^k}{k!} = e^{-0.4 \cdot 1} \frac{(0.4 \cdot 1)^0}{0!} = e^{-0.4} = 0.67032$
  Wie hoch ist die Wahrscheinlichkeit, dass innerhalb einer Minute mindestens zwei Busse fahren?
  $P(N(t) = k) = (N(1) \geq 2) = 1 - P(kein\ Bus) - P(ein\ Bus) = 1 - P(N(1) = 0) - P(N(1) = 1) = 1 - e^{-0.3} \frac{(0.3)^0}{0!} - e^{-0.3} \frac{(0.3)^1}{1!} = 0.0369$
- Wenn Gesamtzahl der in einem Zeitintervall auftretenden Ereignisse bekannt ist, verhalten sich Zeitpunkte wie uniform verteilt:
- Wartezeitsimulation von Poissonprozess bis zum Zeitpunkt $T > 0$:
- Ziehe die Gesamtzahl auftretender Ereignisse im Intervall $(0, T]$ aus einer Poissonverteilung: $k \sim Pois(T\lambda)$
- Ziehe k Zufallszahlen $u_1, u_2, \dots, u_k$ unabhängig aus einer Gleichverteilung auf $(0, T]$
- Ordne die gleichverteilten Zufallszahlen und setze die Eintrittszeitpunkte $s_1, s_2, \dots, s_k$ als $s_i = u_i$ →
- **Beispiel:** Auf einem Server werden durchschnittlich 13 Jobs pro Stunde unabhängig voneinander gestartet. Wie gross ist die Wahrscheinlichkeit, dass innerhalb von 24 Stunden zwei Jobs im Abstand von weniger als 1 Sekunde gestartet werden?

```
k minabst <- function() {
   lambda<-13/60/60
   maxtime <- 24*60*60
   numjobs <- rpois(1, lambda =
      lambda * maxtime)
   jobtimes <- sort(runif(numjobs,
      min = 0, max = maxtime))
   min(diff(jobtimes))
}
x  <-  replicate(10000,  minabst())
mean(x<1)
```

Erneuerungsprozesse

- Allgemeine Erneuerungsprozesse mit iid verteilten Wartezeiten.

Lebensdauerverteilungen Allgemein:

- Wartezeit für nächste Ereignis entspricht der Lebensdauer des entsprechenden Bauteils.
- Lebensdauer hängt vom Alter des Bauteils ab. → nicht gedächtnislos
- $F(t)$: **Verteilungsfunktion der Lebensdauer**; gibt **Wahrscheinlichkeit** an, dass **Bauteil** bis zum Zeitpunkt t **defekt** ist.
- $S(t) = 1 - F(t)$: **Überlebensfunktion**, welcher Anteil der **Bauteile funktionieren** zum Zeitpunkt t **noch**.
- $h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1-F(t)}$: **Ausfallrate**, ist die bedingte Dichte der Lebensdauer zum Zeitpunkt t gegeben die Funktionsfähigkeit zum Zeitpunkt t. Wie lange hält das Bauteil noch, nachdem es bereits bis zum Zeitpunkt t gehalten hat?

Lebensdauerverteilung mit Weibull-Verteilung:

- $F(t) = 1 - S(t) = 1 - \exp\left(-(\lambda t)^\beta\right)$: **Verteilungsfunktion** der **Weibull**-**Verteilung**
- $S(t) = \exp\left(-(\lambda t)^\beta\right)$: **Überlebensfunktion** der **Weibull**-**Verteilung**
- $f(t) = \lambda\beta(\lambda t)^{\beta-1} \cdot e^{-(\lambda t)^\beta}$: **Dichtefunktion** der **Weibull**-**Verteilung**
- $h(t) = \frac{f(t)}{S(t)} = \lambda\beta(\lambda t)^{\beta-1}$: **Ausfallrate** der **Weibull**-**Verteilung**
- $E(T) = \frac{1}{\lambda}\Gamma\left(1 + \frac{1}{\beta}\right)$
- $\beta > 1$: ist $h(t)$ monoton steigend, Bauteile altern und fallen mit der Zeit stärker aus, während $\beta < 1$: ist $h(t)$ monoton fallend, Ausfallrate nimmt mit der Zeit ab.
- dweibull(); pweibull(); qweibull(); rweibull() → shape = β, scale = $1/\lambda$

Langzeitverhalten von Erneuerungsprozessen

- $E(T_n) = \lim\limits_{N\to\infty} \frac{S_N}{N} = \tau$ können mittlere Wartezeit bestimmen, indem wir die Gesamtzeit S_N beim Eintreffen des N-ten Ereignisses durch die Anzahl der Ereignisse N teilen.
- $\lim\limits_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\tau}$ für lange Zeiten ist die Anzahl der Ereignisse geteilt durch die Gesamtlaufzeit der Kehrwert der mittleren Rate zwischen zwei Ereignissen.
- **Preventive Maintenance:** vorbeugende Instandhaltung. **Bsp:** Betrachten Autobatterie mit $[1{,}5]$ gleichverteilter Lebensdauer. Batterie wird ausgewechselt, wenn defekt oder wenn sie nach drei Jahren ausgebaut wird. Können Funktionsdauer T_a als Funktion $g(T)$ der Zufallsvariable T definieren: $g(t) = \begin{cases} t, & t < 3 \\ 3, & t \geq 3 \end{cases}$. $f(t) = \begin{cases} \frac{1}{4}, & t \in [1{,}5] \\ 0, & sonst \end{cases}$ $f(t)$ sei Dichtefunktion der Lebensdauer. Damit ergibt sich der Erwartungswert der Funktionsdauer durch: $E\left(g(T)\right) = \int_0^\infty g(t) \cdot f(t)\, dt = \int_1^3 \frac{t}{4}\, dt + \int_3^5 \frac{3}{4}\, dt = 2.5 = E(T_a)$

Kumulative Prozesse

- $K(t)$: generalisierten Kostenprozess, **Kosten** und **Gewinne laufen** als totale Nettokosten im Intervall $[0, t]$ **auf**.
- $E(T) = \int_{-\infty}^{\infty} t \cdot f_T(t)\, dt$: **erwartete Lebensdauer eines Bauteils**, das Integral bilden von t multipliziert mit der Dichtefunktion $f_T(t)$
- $E\left(K(T)\right) = \int_{-\infty}^{\infty} k(t) \cdot f_T(t)\, dt$: **erwartete Kosten pro Zyklus**, $k(t)$: Kostenfunktion, $f_T(t)$: Dichte der Lebensdauer einer Funktion
- $\lim\limits_{t\to\infty} \frac{K(t)}{t} = \frac{E(K(T))}{E(T_a)}$: asymptotisch **erwartete/mittlere Kosten(-rate) pro Zykluszeit**, Erwartungswert der Zykluskosten geteilt durch den Erwartungswert der Zykluszeiten.
- T_n: Wartezeiten, hier auch Zyklen genannt, $S_n = \sum_{i=10}^{n} T_i$, $K_n = K(S_n) - K(S_{n-1})$, n-te Zyklus des Intervalls $(S_{n-1}, S_n]$
- Der stochastische Prozess $K(t), t \geq 0$ heisst kumulativer Prozess, wenn die Folge $(T_n, K_n), n \geq 1$ eine Folge von iid-bivariaten Zufallsvariablen ist.
- Fortsetzung Bsp. Autobatterie: Ungeplanter Austausch = CHF 150, geplant CHF 50. $E(T_a) = 2.5$. $E(K) = P(T \leq 3) \cdot 150\ CHF + P(T > 3) \cdot 50\ CHF = \frac{1}{2} \cdot 150\ CHF + \frac{1}{2} \cdot 50\ CHF = 100\ CHF$. Mittlere Kosten pro Zykluszeit mit vorzeitiger Erneuerung $\frac{E(K(T))}{E(T_a)} = \frac{100\ CHF}{2.5\ Jahre} = 40 \frac{CHF}{Jahr}$, ohne vorzeitig: $\frac{E(K(T))}{E(T_a)} = \frac{150\ CHF}{3\ Jahre} = 50 \frac{CHF}{Jahr}$ → mit vorzeitig = besser

Stand der Arbeit

SW	Vorlesung	AB/Aufgabe	Skript	ZF
1	☒	☒	☒	☒
2	☒	☒ *NB1*	☒	☒
3	☒	☒	☒	☒
4	☒	☒ + ☒ *NB2*	☒	☒
5	☒	~~☒~~	☒	☒
6	☒	☒ + ☒ *NB3*	☒	☒
7	☒	☒	☒	☒
8	☒	☒ + ☒ *NB4*	☒	☒
9	☒	☒	☒	☒
10	☒	☒ + ☒ *NB5*	☒	☒
11	☒	☒	☒	☒
12	☒	☒ + ☒ *NB6*	☒	☒
13	☒	☒	☒	☒
14	☒	~~☐~~	☒	☒

Modulauflagen

Obligatorische schriftliche Modulschlussprüfung in den Prüfungswochen. Zählt 60 % für die Modulnote. Hilfsmittel: open Book Prüfung

Im Praktikum werden Notebooks (RMarkdown) von Ihnen bearbeitet. Diese zählen zusammen 40 % für die Modulnote.