

107. A04

2023 W

Fragenausarbeitung

## 1. State and prove Bayes' theorem (Satz von Bayes)

Satz von Bayes:

$$P(B_j | A) = \frac{P(B_j) \cdot P(A | B_j)}{\sum_i P(B_i) \cdot P(A | B_i)} = P(A) \quad \text{lt. Satz der totalen Wahrscheinlichkeit}$$

Beweis Satz d. totalen Wahrsch.:

$$\begin{aligned} P(A) &= P(A \cap \Omega) \\ &= P(A \cap \bigcup_{i=1}^n B_i) = P\left(\bigcup_{i=1}^n (A \cap B_i)\right) \stackrel{\text{d-Add.}}{=} \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A \cap B_i) \cdot \frac{P(B_i)}{P(B_i)} \\ &= \sum_{i=1}^n P(B_i) \cdot \frac{P(A \cap B_i)}{P(B_i)} = \sum_{i=1}^n P(B_i) \cdot P(A | B_i) \end{aligned}$$

↑  
bed. Wahrsch. □

Beweis Satz von Bayes:

$$P(B_j | A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(B_j \cap A)}{P(B_j)} \cdot \frac{P(B_j)}{P(A)} = \frac{P(A | B_j) \cdot P(B_j)}{P(A)}$$

wobei:

- bedingte Wahrsch.:  $P(A | B) = \frac{P(A \cap B)}{P(B)}$
- Satz d. vollständ. Wsk.:  $\sum_{i=1}^n P(A | B_i) \cdot P(B_i) = P(A)$

□

2. define the distribution function (Verteilungsfunktion) of a random variable  $X$  (or: of a  $d$ -dimensional random vector  $(X_1, \dots, X_d)$ ) and state its main properties (Grills Skriptum, Abschnitt 2.3). Compute the distribution function of  $X$  if  $X$  has one of the classical distributions

Verteilungsfunktion:  $F_X: \mathbb{R} \rightarrow [0, 1]$ 

$$F_X(x) := P(X \leq x)$$

d-dimensional:

$$F_X(x) = F_{(X_1, \dots, X_d)}(x_1, \dots, x_d) = P(X_1 \leq x_1, \dots, X_d \leq x_d)$$

Eigenschaften:

1)  $0 \leq F(x) \leq 1 \quad \forall x$

4)  $\lim_{x \rightarrow -\infty} F(x) = 0$

2)  $F$  ist monoton steigend

5)  $\lim_{x \rightarrow \infty} F(x) = 1$

3)  $F$  ist rechtsstetig

Anm.:

$$P(X \leq a) = F_X(a)$$

$$P(X < a) = F_X(a-0)$$

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

$$P(a < X < b) = F_X(b-0) - F_X(a)$$

$$P(a \leq X \leq b) = F_X(b) - F_X(a-0)$$

$$P(a \leq X < b) = F_X(b-0) - F_X(a-0)$$

$$P(X = a) = F_X(a) - F_X(a-0)$$

3. Explain the difference between discrete, continuous and "mixed" random variables. Explain what is the density (Dichtefunktion) of a continuous random variable and of a random vector. Give the density of an Exp(lambda), Gaussian, or other classical example. Abschnitt 2.3

- Diskrete ZV: Die Verteilungsfunktion hat Sprungstellen & ist dazwischen konstant.
- Stetige ZV: Die Verteilungsfunktion besitzt keine Sprungstellen, sondern ist stetig.
- gemischte ZV: Die Verteilungsfunktion hat Sprungstellen & ist dazwischen stetig, aber nicht überall konstant.

Dichtefunktion:

- bei stetiger Verteilung mit Verteilungsfkt.:  $F_X: f_X = \frac{\partial F_X(x)}{\partial x}$  ... Dichte

$$F_X(y) = \int_{-\infty}^y f_X(u) du$$

mehrdim.:  $F_X(y) = \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_d} f_X(u_1, \dots, u_d) du_1 \dots du_d$

Fkt  $p(x)$  ist genau dann eine Wskl-funktion, wenn gilt:

$$p(x) \geq 0 \quad \forall x \in \mathbb{R}^d$$

$$\sum_x p(x) = 1$$

f. diskrete ZV

Fkt  $f_X$  ist genau dann eine Dichte funktion, wenn gilt:

$$f_X \geq 0 \quad \forall x \in \mathbb{R}^d$$

$$\int f(x) dx = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_d) dx_1 \dots dx_d = 1$$

f. stetige ZV

4. Explain in formulas the meaning of the following statement: the random variables  $X_1, \dots, X_n$  are independent. Abschnitt 2.3

Unabhängigkeit von ZV:

- Ereignisse sind unabhängig, wenn gilt:  $P(A \cap B) = P(A)P(B)$

stetige ZV:  $X = (X_1, \dots, X_n)$  sind unabhängig, wenn:

$$F_{(X_1, \dots, X_n)}(y_1, \dots, y_n) = \prod_{i=1}^n F_{X_i}(y_i)$$

diskrete ZV:  $P_{(X_1, \dots, X_n)}(y_1, \dots, y_n) = P(X_1 = y_1, \dots, X_n = y_n) = \prod_{i=1}^n P_{X_i}(y_i)$

5. given two random variables  $X, Y$ , be able to define what is the density of  $X$  conditioned to  $Y=y$ , and to compute it in explicit examples. Abschnitt 2.3

Bedingte Verteilung:  $P(X = x_i | Y = y_j) = \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}$

Bedingte Dichte:  $X, Y$  stetig verteilt mit Dichte  $f_{X,Y}$ .

$$f_X(x | Y = y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Bedingte Wahrsch. als Integral der bedingten Dichte:

$$P(X \leq a | Y = y) = \int_{-\infty}^a f_X(x | Y = y) dx$$

6. be able to state the Transformationsatz für Dichten (Abschnitt 2.4) and to apply it, for instance to compute the density of  $X+Y$ ,  $X/Y$  when  $X$  and  $Y$  are independent

Transformationsatz für Dichten:

$X = (X_1, \dots, X_n)$  stetig verteilt mit Dichte  $f_x$ .

$g: \mathbb{R}^n \rightarrow \mathbb{R}^n$  injektiv

$g(X) = Y$  ist dann ebenfalls stetig verteilt mit Dichte

$$f_y(y) = \begin{cases} f_x(g^{-1}(y)) \cdot \frac{1}{\left| \frac{dg}{dx}(g^{-1}(y)) \right|} & \text{wenn } y \in g(\mathbb{R}^n) \\ 0 & \text{sonst} \end{cases}$$

Funktionsdeterminante:

$$\frac{dg}{dx} = \det\left(\left(\frac{dg_i}{dx_j}\right)_{n \times n}\right)$$

bei  $g(x, y) = g_1(x, y), g_2(x, y)$ :

$$\frac{dg}{dx} = \begin{vmatrix} \frac{dg_1}{dx} & \frac{dg_1}{dy} \\ \frac{dg_2}{dx} & \frac{dg_2}{dy} \end{vmatrix}$$

Sei  $X, Y$  unabhängig mit Dichte  $f_x$  &  $f_y$ . Dann ist die Dichte von  $X+Y$ :

$$f_{x+y}(z) = f_x * f_y = \int_{-\infty}^{\infty} f_x(x) \cdot f_y(z-x) dx$$

von  $Z = \frac{X}{Y}$ :

$$f_z(z) = \int_{-\infty}^{\infty} f_y(y) f_x(zy) |y| dy$$

← Faltungsintegral

bei diskrete ZV:

$f$  durch  $p$  ersetzen

↓ Dichte      ↓ Wskl-mass

& statt  $\int_{-\infty}^{\infty}$ :  $\sum_x$

7. define the expectation value/variance of a random variable  $X$  (either continuous, discrete or mixed), and the covariance between  $X$  and  $Y$ . Give (better: prove) the main properties of expected value and variance. Compute the expectation value/variance/etc in classical examples of distributions (Abschnitt 2.5)

Erwartungswert  $E$

$$\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n} \rightarrow E(X)$$

Def.: diskret:  $E(X) = \sum_x x \cdot p_x(x)$

stetig:  $E(X) = \int_{-\infty}^{\infty} x \cdot f_x(x) dx$

gemischt:  $E(X) = \sum_x x \cdot p_x(x) + \int_{-\infty}^{\infty} x \cdot f_x(x) dx$

Eigenschaften:

- 1) Linearität:  $E(aX + b) = a \cdot E(X) + b$
- 2) Additivität:  $E(X + Y) = E(X) + E(Y)$
- 3) Monotonie:  $X \leq Y \Rightarrow E(X) \leq E(Y)$
- 4) wenn  $X, Y$  unabhängig:  $E(XY) = E(X)E(Y)$

Varianz  $V$ :

Def.:  $V(X) = E[(X - E(X))^2]$  ← Maß für Streuung

oder:  $V(X) = E(X^2) - E(X)^2$  (lt. Steinerschen Verschiebungssatz für  $a=0$ )  
bzw. Definition ausmultipliziert

Steinerscher Verschiebungssatz:

für beliebiges  $a \in \mathbb{R}$ :  $E(X - a)^2 = V(X) + (E(X) - a)^2$

Eigenschaften Varianz:

- 1)  $V(X) \geq 0$
- 2)  $V(X) = 0$  genau dann, wenn  $P(X = E(X)) = 1$
- 3)  $V(aX + b) = a^2 \cdot V(X)$
- 4)  $V(X + Y) = V(X) + V(Y)$ , wenn  $X, Y$  unabhängig.

$E(X)$  ... Lageparameter: gibt an, wie groß Werte sind, die  $X$  annimmt.

$V(X)$  ... Streuungsparameter: gibt Abweichungen der einzelnen Werte zueinander an.

\*  $E(X) = 0 \Rightarrow X$  ist zentriert

\*  $E(X^2) = 1 \Rightarrow X$  ist normiert

\*  $E(X) = 0$  &  $V(X) = 1 \Rightarrow X$  ist standardisiert

\*  $\tilde{X} = \frac{X - E(X)}{\sqrt{V(X)}} \Rightarrow X$  zentriert & normiert

↓

$\sqrt{V(X)}$  ... Standardabweichung

## Covarianz Cov:

$$\text{Cov}(X, Y) := E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

$\text{Cov}(X, Y) \dots > 0$ : positiv korreliert

$< 0$ : negativ korreliert

$= 0$ : unkorreliert

Eigenschaften: 1)  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

2)  $\text{Cov}(X, X) = V(X)$

3)  $\text{Cov}(aX + bY, Z) = a \cdot \text{Cov}(X, Z) + b \cdot \text{Cov}(Y, Z)$

4)  $\text{Cov}(X, Y) = 0$ , wenn  $X, Y$  unabhängig

Korrelationskoeffizient:  $\text{corr}(X, Y) = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}$

Satz:  $V(X+Y) = V(X) + 2 \cdot \text{Cov}(X, Y) + V(Y)$

Beweis: Def für  $V(X+Y)$  mit  $E$  einsetzen & ausmultiplizieren

## Cauchy - Schwarz Ungleichung:

Wenn  $E(X^2) < \infty$  &  $E(Y^2) < \infty$ :

$$E(XY) \leq \sqrt{E(X^2)E(Y^2)}$$

Gleichheit, wenn  $\exists a > 0$ ,  
dass gilt:  $aX = Y$ 

Beweis:

$$E\left[\left(\frac{X}{\sqrt{E(X^2)}} - \frac{Y}{\sqrt{E(Y^2)}}\right)^2\right] \geq 0$$

$$E\left[\frac{X^2}{E(X^2)} - 2 \cdot \frac{XY}{\sqrt{E(X^2)E(Y^2)}} + \frac{Y^2}{E(Y^2)}\right] = \frac{E(X^2)}{E(X^2)} - 2 \cdot \frac{E(XY)}{\sqrt{E(X^2)E(Y^2)}} + \frac{E(Y^2)}{E(Y^2)} \geq 0$$

$$2 \geq 2 \cdot \frac{E(XY)}{\sqrt{E(X^2)E(Y^2)}} \quad | : 2$$

$$1 \geq \frac{E(XY)}{\sqrt{E(X^2)E(Y^2)}} \quad \Rightarrow \quad E(XY) \leq \sqrt{E(X^2)E(Y^2)} \quad \square$$

## Markov - Ungleichung:

Sei  $X \geq 0$  &  $\lambda > 0$ . Dann gilt:  $P(X \geq \lambda) \leq \frac{1}{\lambda} E(X)$ Beweis: Def.  $E(X)$ :  $E(X) = \sum_x x \cdot P(X=x)$ daraus folgt:  $E(X) = \sum_{x < \lambda} x \cdot P(X=x) + \sum_{x \geq \lambda} x \cdot P(X=x)$ 

$$E(X) \geq \sum_{x \geq \lambda} x \cdot P(X=x) \geq \sum_{x \geq \lambda} \lambda \cdot P(X=x) \geq \lambda \cdot P(X \geq \lambda)$$

$$\text{Anm.: } \sum_{x \geq \lambda} P(X=x) = P(X \geq \lambda)$$

$$\Rightarrow E(X) \geq \lambda \cdot P(X \geq \lambda) \quad | : \lambda$$

$$P(X \geq \lambda) \leq \frac{1}{\lambda} \cdot E(X)$$

□



Chebyshev - Ungleichung (Spezialfall d. Markov-Ungleichung)

$$P(|X - E(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$$

Beweis:  $Y = (X - E(X))^2$

$$P(|X - E(X)| \geq \lambda) = P(Y \geq \lambda^2)$$

Einsetzen in Markov-Ungleichung:

$$P(Y \geq \lambda^2) \leq \frac{1}{\lambda^2} \cdot E(Y) \quad \text{mit } E(Y) = E((X - E(X))^2) = V(X)$$

$$\Rightarrow P(|X - E(X)| \geq \lambda) \leq \frac{1}{\lambda^2} \cdot V(X)$$

□

9. Define the moment generating function of a random variable (r.v.) (in the discrete and in the continuous case) and give its main properties. Compute it for a given classical example (abschnitt 2.6)

Momente: zB  $E(X)$ ;  $E((X - E(X))^2) = V(X)$ ;  $E(X^n)$   $n \in \mathbb{N}$

- $E(X^n)$  ... n-te Moment (=  $M_n(X)$ )
- $E((X - E(X))^n)$  ... n-te zentrale Moment (=  $m_n(X)$ )

Momentenerzeugende Funktion:  $M_X(t) = E(e^{xt})$

- diskret:  $M_X(t) = \sum_x e^{xt} \cdot p_X(x)$

- stetig:  $M_X(t) = \int_0^{\infty} e^{xt} \cdot f_X(x) dx$

n-tes Moment:  $\left. \frac{d^n}{dt^n} M_X(t) \right|_{t=0} = M_n(X) = E(X^n)$

Eigenschaften:

- 1)  $X, Y$  unabhängig.  $\Rightarrow M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$
- 2) Für  $Y = aX + B$ ,  $a, b \in \mathbb{R}$ :  $M_Y(t) = e^{bt} \cdot M_X(at)$
- 3) Wenn  $m_k = E(X^k)$  endlich, dann ist  $\varphi_X(t)$  (charakteristische Funktion)  $k$ -mal diff. bar und  $\varphi_X^{(k)}(0) = i^k \cdot m_k$
- 4) Wenn  $M_X(t) = M_Y(t)$  & endlich  $\Rightarrow X$  &  $Y$  haben gleiche Verteilung

10. State and prove (using Chebyshev) the weak law of large numbers. State the central limit theorem.

Schwaches Gesetz der großen Zahlen:

- $(X_n, n \in \mathbb{N})$  sei Folge von iid. ZV
- $E(X_n) = \mu$
- $V(X_n) = \sigma^2 < \infty$
- $S_n = X_1 + \dots + X_n$

Dann:  $P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \rightarrow 0 \quad \forall \varepsilon > 0$

Beweis:

Chebyshev-Ungleichung:  $P(|X - \mu| \geq \lambda) \leq \frac{\sigma^2}{\lambda^2}$

Betrachte  $Y = \frac{1}{n} \cdot \sum_{i=1}^n X_i$ ;  $E(Y) = \frac{1}{n} \cdot \sum_{i=1}^n E(X_i) = \frac{1}{n} \cdot n \cdot \mu = \mu$

$V(Y) = V\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot \sum_{i=1}^n V(X_i) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$   
↑ unabhängig

einsetzen in Chebyshev:

$$P(|Y - \mu| \geq \lambda) \leq \frac{\sigma^2}{n \cdot \lambda^2} \xrightarrow{n \rightarrow \infty} 0 \quad \square$$

Zentraler Grenzwertsatz:

- $(X_n, n \in \mathbb{N})$  sei Folge von iid ZV
- $E(X) = \mu$
- $V(X) = \sigma^2$

$\Rightarrow$

Dann ist  $S_n$  näherungsweise normalverteilt mit Mittel  $n\mu$  &  $V(S_n) = n \cdot \sigma^2$ , also es gilt:

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\mu}{\sqrt{n \cdot \sigma^2}} \leq x\right) = \Phi(x)$$

Markovketten in diskreter Zeit:

$X = (X_i)_{i \geq 0}$  ... Zustand zum Zeitpunkt i mit  $X_0, \dots, X_n \in S$  .. Zustandsraum  
 und  $P(X_n = x_n | X_0 = x_0, \dots, X_{n-1} = x_{n-1}) = P(X_n = x_n | X_{n-1} = x_{n-1})$   
 d.h. Zukunft ist unabhängig von Vergangenheit  $\Rightarrow$  MK ist gedächtnislos

$P(X_{n+1} = j | X_n = i)$  ... Übergangswahrscheinlichkeiten einer MK

homogene MK:  $p_{ij} = P(X_{n+1} = j | X_n = i)$

$p_{ij}(t) = P(X_{n+t} = j | X_n = i)$  ... t-stufige Übergangswsk.

Übergangsmatrix:  $P_{a,b} = P(X_1 = b | X_0 = a)$   $a, b \in S$

$$P = \begin{pmatrix} P_{a_1, b_1} & P_{a_1, b_2} & \dots & P_{a_1, b_k} \\ P_{a_2, b_1} & & & \vdots \\ \vdots & & & \vdots \\ P_{a_k, b_1} & & & P_{a_k, b_k} \end{pmatrix}$$

n-stufige Übergangsmatrix:  $P(X_n = b | X_0 = a) = P_{a,b}^{(n)}$

Erwartungswert:  $f \in \mathbb{R}^{|S|}$ ,  $|S| < \infty$

$$E(f(X_n) | X_0 = i) = (P^n \cdot f)(i) = \sum_{j \in S} P^n(i, j) \cdot f(j)$$

$$\text{für } n=1: E(f(X_n) | X_0 = i) = (P \cdot f)(i) = \sum_{j \in S} p(i, j) \cdot f(j)$$

$$p(i, j) = P(X_1 = j | X_0 = i)$$

Eigenschaften: 1) Zeilensummen ergeben 1:  $\sum_j p_{ij} = 1$

2) Alle Einträge nicht-negativ:  $\forall p_{ij} \geq 0$

### Klassenzersetzung:

- Zustände  $i$  &  $j$  kommunizieren, wenn:  $i \leftrightarrow j$   $\Rightarrow$  Äquivalenzrelation, dh Zustandsraum wird in Äquivalenzklassen zerlegt, die wir Klassen nennen.
- $i \rightarrow j$  "j folgt i":  
 $i, j \in S, \exists n \geq 0$ , sodass es möglich ist in  $n$ -Schritten von  $i$  nach  $j$  zu gehen mit pos Wsk.
- Eine Eigenschaft heißt Klasseneigenschaft, wenn sie für alle Zustände oder keinen gilt.
- $i \in S$  ist absorbierend, wenn  $p_{ij} = 1$   
 $\Rightarrow$  absorbierende Klassen kommunizieren nur mit sich selbst.
- Irreduzible MK: eine MK, die nur eine Klasse besitzt

### Klasseneigenschaften:

eine Klasse ist...

... offen: es existiert ein Element außerhalb der Klasse, von diesem kein Pfad zurück in die Klasse führt. sonst:

... abgeschlossen

Periode eines Zustands:

$$d(i) = \text{ggT} \{ t \geq 0 : p_{ii}(t) > 0 \}$$

$\Rightarrow \exists$  ein Pfad der Länge  $n$  von  $i$  nach  $j$

-  $d(i) = 1 \Rightarrow$  Zustand  $i$  ist aperiodisch, sonst periodisch

(Ahm.: Zustände mit Schleife sind aperiodisch)

13. Be able to define recurrence/transience and positive/null recurrence, and to decide transience/recurrence of a state  $x$  in simple examples. Define what is a stationary distribution and explain the relation with positive recurrence and with average return time  $\tau_i$ . Abschnitt 3.3.2

**Rekurrenz:** Zustand  $i \in S$  rekurrent, wenn ein Zustand fast sicher  $\infty$ -oft besucht wird. Sonst transient.  
Gilt diese Eigenschaft  $\forall i \in S$ , dann gilt:  
MK ist rekurrent (transient)

positiv rekurrent:  $E_i(\tau_i) < \infty$   
sonst: nullrekurrent

**Stationäre Verteilung:**

$(\pi_i, i \in \Omega_x)$  heißt stationäre Verteilung, wenn:

- $\pi_i \geq 0$
- $\sum_i \pi_i = 1$
- $\sum_i \pi_i \cdot p_{ij} = \pi_j$

Aus der Existenz der stationären Verteilung folgt die positive Rekurrenz:

$$E_i(\tau_i) < \infty$$

Anm.: Jede irreduzible, endliche MK ist positiv rekurrent:  $\sum_i \pi_i = 1$

Nullrekurrenz kann nur auftreten, wenn Klasse  $\infty$ -viele Zustände hat.

## Markovketten in stetiger Zeit:

### Algorithmus:

- gegeben:
- Übergangsmatrix  $P$
  - $\nu_i; i \in S$  ... Startverteilung auf  $S$
  - $\lambda_i > 0; i \in S$  .. Sprungraten

- 1) Start in  $X_0$  mit Startverteilung  $\nu$
- 2) Falls  $X_0 = i$ , warte Zeit  $T \sim \exp(\lambda_i)$  & springe in nächsten Zustand  $j$  mit Wslk  $p_{ij}$ .
- 3) Wiederhole 2 ausgehend von Zustand  $X_T$

Markov-Eigenschaft: Für MK  $(X_t, t \in \mathbb{N})$  in stetiger Zeit gilt  $\forall n \in \mathbb{N}$ ,  $0 \leq t_1 \leq \dots \leq t_n < \infty$  und  $i_1, \dots, i_n \in S$ :

$$P(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1}) \Leftarrow \text{Gedächtnislosigkeit}$$

### $Q$ -Matrix (Generator):

$$q_{ij} = \begin{cases} \lambda_i p_{ij} & \text{für } i \neq j \\ \lambda_i (p_{ij} - 1) & \text{für } i = j \end{cases}$$

$\Rightarrow$  Diagonale enthält negative, Rest positive Einträge.

MK  $X_t$  bzw. Erzeuger  $Q$  heißt konservativ, wenn

$$\forall i \in S: \quad q_{ij} > -\infty$$

$$\sum_j q_{ij} = 0 \quad \leftarrow \text{Zeilensumme}$$

Eingebettete MK mit "jump-matrix" bzw. Übergangsmatrix  $P$ :

$$P_{ij} = \begin{cases} \frac{q_{ij}}{-q_{ii}} & \text{wenn } q_{ii} < 0, i \neq j \\ 1 & \text{wenn } q_{ii} = 0, i = j \\ 0 & \text{sonst} \end{cases}$$

16. Give the relation between  $Q$  and the transition matrix  $P_t$  of the continuous-time MC. Be able to compute  $P_t$  in simple examples. Explain what is the phenomenon of "explosion"

$$P(X_t = j | X_0 = i) = (e^{Qt})_{ij}$$

Wenn  $Q$  diagonalisierbar, also  $Q = MDM^{-1}$

dann ist  $P(t) = e^{Qt} = MDM^{-1} = M \cdot \begin{pmatrix} e^{t\lambda_1} & & \\ & \ddots & \\ & & e^{t\lambda_d} \end{pmatrix} \cdot M^{-1}$

↑  
Übergangsmatrix

17. Define what is an estimator (Schätzer) and what it means that it is consistent and unbiased (erwartungstreu). Be able to explain how to find estimator of the parameters of a concrete distribution using the method of moments and the Maximum likelihood method. Be able to prove that "Momentschätzer" are consistent, using the law of large numbers.

Stichprobe: Folge von unabhängigen ZV mit Verteilung

Statistik  $T$  ist eine ZV, die aus Stichprobe berechnet werden kann:

$$T = (X_1, \dots, X_n)$$

Schätzer  $(\hat{\theta}_n)$  ist Folge von Statistiken  $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$

Eigenschaften von Schätzern  $\hat{\theta}_n$ :

◦ erwartungstreu:

$$\text{wenn } \forall \theta \in \Theta : \mathbb{E}_\theta(\hat{\theta}_n) = \theta$$

+ effizient, wenn  $\hat{\theta}_n$  die kleinstmögliche Varianz unter allen erwartungstreuen Schätzern hat:

$$\Rightarrow \mathbb{V}_\theta(\hat{\theta}_n) \leq \mathbb{V}_\theta(\tilde{\theta}_n) \quad \forall \theta$$

$$\text{oder: } \lim_{n \rightarrow \infty} \mathbb{V}_\theta(\hat{\theta}_n(X_1, \dots, X_n)) = 0$$

◦ konsistent:

$$\lim_{n \rightarrow \infty} \hat{\theta}_n(X_1, \dots, X_n) = \theta$$

Momentenmethode:

Sei  $\theta$  ein einzelner reeller Parameter. Dann können wir  $\mathbb{E}(X)$  als Fkt von  $\theta$  darstellen:

$$\mathbb{E}_\theta(X) = m(\theta)$$

Umkehrfkt.:  $m^{-1}(\bar{x}_n) = \hat{\theta}_n$  ... stark konsistenter Schätzer = Momentenschätzer



## Maximum-Likelihood Methode:

1) Aufstellen d. Likelihood-Funktion:

$$\text{diskret: } L(X_1, \dots, X_n; \theta) = \prod_{i=1}^n p_{\theta}(X_i)$$

$$\text{stetig: } L(X_1, \dots, X_n; \theta) = \prod_{i=1}^n f_{\theta}(X_i)$$

2) Berechne ML-Schätzer, der  $L(X_1, \dots, X_n; \theta)$  maximiert:

$$\hat{\theta}_n(X_1, \dots, X_n) = \operatorname{argmax} L_n(X_1, \dots, X_n; \theta)$$

$$\Rightarrow \frac{d}{d\theta} L_n = 0 \quad \Leftarrow \text{ableiten, nullsetzen \& nach } \theta \text{ umformen}$$

Oft ist die Log-Likelihood-Funktion leichter abzuleiten:

$$l_n = \log(L_n(X_1, \dots, X_n; \theta))$$

(gleiches Maximum, weil Log streng monoton wachsend)