MatheC4 - Stichpunkte

Patrick Cerny - FAU Erlangen

October 9, 2011

Wer Fehler findet bitte eine Mail an: patrick dot cerny at googlemail dot com

Inhaltsangabe

1 Handout I

:D

2 Handout II

2.1 Begriffe

- **Ergebnis:** Einer von mehreren möglichen Ausgängen eines Experiments
- **Ergebnismenge Ω:** Menge der möglichen Ergebnisse
- **Ereignis:**
  - Aussage über den Ausgang eines Experiments
  - Tritt ein, wenn die Aussage auf das Ergebnis des Experiments zutrifft
  - Ein Ereignis ist eine Teilmenge des Ergebnisraums
- Ω heißt das **sichere Ereignis**
- $\emptyset = \bar{\Omega}$ heißt das **unmögliche Ereignis**
- $\{\omega\}$ heißt ein **Elementarereignis**
- Gilt: $A \cap B = \emptyset \implies$ man schreibt $A + B$ anstelle von $A \cup B$, es gelten die De-Morganschen Regeln
- Eine Funktion $P : \mathcal{A} \to \Re$ auf einer σ-Algebra $\mathcal{A}$ von Teilmengen einer Menge Ω, die Kolmogoroff 1-4 erfüllt, heißt eine **Wahrscheinlichkeitsverteilung** auf $\mathcal{A}$
- Ein Tripel $(\Omega, \mathcal{A}, P)$ heißt ein **Wahrscheinlichkeitsraum**
- Elemente der σ-Algebra $\mathcal{A}$ A_k heißen **Ereignisse**
- $P(A)$ heißt die **Wahrscheinlichkeit des Ereignisses** A

2.2 Axiome von Kolmogoroff:

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $P(A + B) = P(A) + P(B)$
4. Für jede Folge $A_1, A_2, \ldots$ von paarweise disjunkten Mengen $A_k \in \mathcal{A}$ gilt:

$$P(\sum_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k) \tag{1}$$

3 Handout III - LaPlace-Experimente

- Allgemein gilt:

$$P(A) = \frac{|A|}{|\Omega_n^N|} = \frac{K}{N} \tag{2}$$

3.1 Galton-Brett

- **Binärvektor:** $A_k^n = \{\delta = (\delta_1, \delta_2, \ldots, \delta_n); \sum_{i=1}^{n} \delta_i = k\}$ mit $\delta \in \{0, 1\}$

3.2 Urnen-/Schubladenmodelle

- Gedankenmodelle zur Beschreibung von Zufallsexperimenten
- Verwenden von Binärvektoren (z.B. Ziehen von n Kugeln ohne Berücksichtung der Reihenfolge), daher gilt auch: $|\Omega_n^N| = \binom{N}{n}$
- Urnenmodell:

$$P(B_k) = \frac{\overbrace{\binom{K}{k}}^{\text{erste Sorte}} \overbrace{\binom{N-K}{n-k}}^{\text{zweite Sorte}}}{\underbrace{\binom{N}{n}}_{\text{Gesamtanzahl}}} \tag{3}$$

mit k und $n-k$ Einsen der Binärvektoren B_k: $\sum_{i=1}^{K} \delta_i = k$ und $\sum_{i=K+1}^{N} \delta_i = n-k$.

4 Handout IV: Bedingte Wahrscheinlichkeiten

Beeinflusst das Eintreten eines Ereignisses B das eines anderen Ereignisses A?

- **Bedingte Wahrscheinlichkeit:**

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ mit } P(B) > 0 \tag{4}$$

- $P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C)$
- Bei unübersichtlichen Experimenten $\implies$ Vorsicht vor dem **Paradoxon von Bertrand**! $\implies$ Wahrscheinlichkeitsraum definieren
- **Partition (von Ω):** Ereignisse $B_n \in \mathcal{A}$ mit paarweise disjunkten B_n und $\sum_n B_n = \Omega$
- **totale Wahrscheinlichkeit:** $P(A) = \sum_n P(A|B_n)P(B_n)$
- **Formel von Bayes:**

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_n P(A|B_n)P(B_n)}, \text{ für } P(A) > 0 \tag{5}$$

- **A-priori-Wahrscheinlichkeit:** Wahrscheinlichkeit des Eintreten eines Ereignisses in der Ausgangskonfiguration
- **A-posteriori-Wahrscheinlichkeit:** Wahrscheinlichkeit des Eintreten eines Ereignisses nach Einführung einer Bedingung

5 Handout V: Stochastische Unabhängigkeit

5.1 Experimente mit mehrfachen Wiederholungen

- Zwei Ereignisse A und B heißen **stochastisch unabhängig**, wenn gilt:

$$P(A \cap B) = P(A)P(B) \neq 0 \tag{6}$$

und damit auch $\bar{\text{A}}$,B bzw. A,$\bar{\text{B}}$, sowie $\bar{\text{A}}$,$\bar{\text{B}}$

- Zu Produktexperimenten definieren wir **Produkträume**: $(\Omega_n, \mathcal{A}_n, P_n)$ mit jeweils $\Omega_1, \dots, \Omega_n$, $\mathcal{A}_1, \dots, \mathcal{A}_n$ und $P_1, \dots, P_n$
- Kleinste σ-Algebra des Produktraums ist die **Produkt σ-Algebra**: $\bigotimes_{k=1}^n \mathcal{A}_k$
- Analog: **Produktwahrscheinlichkeit P** $= P_1 \otimes P_2 \otimes \cdots \otimes P_n = \bigotimes_{i=1}^n P_i$

5.2 Versuchsreihen

- **Versuchsreihe der Länge n:** Gilt $(\Omega_k, \mathcal{A}_k, P_k) = (\Omega_0, \mathcal{A}_0, P_0) \implies$ n-fache unabhängige Wiederholung eines Zufallsexperiments
- **Bernoulli-Experiment:**
  1. $\Omega_0 = \{0, 1\}$
  2. $\mathcal{A}_0 = 2^{\Omega_0}$
  3. $P\{1\} = p$
- Bernoulli-Experimente $\implies$ Die Summe der Erfolge:

$$P_n(A_k) = \binom{n}{k} p^k q^{n-k} \tag{7}$$

6 Handout VI: Diskrete Wahrscheinlichkeitsräume

- Problem: Was muss man mindestens über P wissen, um P(A) für jedes Ereignis A berechnen zu können?
- Charakterisiert durch $(\Omega, \mathcal{A}, P)$, mit $\mathcal{A} = 2^\Omega$
- P heißt hier **diskrete Wahrscheinlichkeitsverteilung**
- Funktion $f : \Omega \to \Re$, mit $f(\omega) = P\{\omega\}$ heißt **Wahrscheinlichkeitsfunktion von P**
- Es gilt:

$$P\{\Omega\} = 1 \implies \sum_{\omega \in \Omega} f(\omega) = 1, \text{ mit } \Omega \text{ abzählbar} \tag{8}$$

6.1 LaPlace-Verteilung

- $\mathcal{L}(\Omega)$ heißt LaPlace-Verteilung, mit Wahrscheinlichkeitsfunktion $f(\omega) = \frac{1}{|\Omega|}$
- Anwendung: Zufallsexperimente, bei denen jedes Ergebnis die gleiche Chance hat

6.2 Hypergeometrische Verteilung

- $\mathcal{H}(N, K, n)$-Verteilung, mit $f(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$
- Anwendung: Zufallsexperimente, bei denen man die Ergebnisse als Anzahlen von schwarzen Kugeln unter n gezogenen interpretieren kann

6.3 Binomialverteilung

- $\mathcal{B}(k; n, p)$-Verteilung mit $f(k) = \binom{n}{k} p^k q^{n-k}$
- Anwendung: Summe der Erfolge bei einer Bernoulli-Versuchsreihe

Definition 1. *(Gedächtnislosigkeit)*

Für alle $m, n = 0, 1, 2, \ldots$ gilt:

$$P(A_{m+n}) = P(A_m) P(A_n)$$

6.4 Geometrische-Verteilung

- $\mathcal{G}(p)$-Verteilung, mit $f(n) = pq^{n-1}$
- Anwendung: Beschreibt die Wartezeit für das erstmalige Eintreten eines Ereignisses unter der Annahme der Gedächtnislosigkeit
- Interpretation: $p = f(1)$ ist die Wahrscheinlichkeit, dass das Ereignis gleich beim ersten Versuch eintritt

6.5 Poisson-Verteilung

- $\mathcal{P}(k;\mu)$-Verteilung, mit $f(k) = \frac{\mu^k}{k!}e^{-\mu}$ und $\mu = np$
- Anwendung: Beschreibt die Häufigkeit des Eintreten eines Ereignisses, das zu zufälligen Zeitpunkten eintritt
- Streben $n \to \infty$ und $p \to 0$, d.h. $np = \mu = const.$, so gilt: $\mathcal{B}(k;n,p) \approx \mathcal{P}(k;np)$, für $n \gg 1$ und $p \ll 1$

6.6 Erzeugende Funktion

- Erzeugende Funktion $\hat{f}(z) := \sum_{n=0}^{\infty} f(n)z^n, (0 \leq z \leq 1)$, mit P als diskrete Verteilung
- Interpretation: $f(n) = a_n$ ist der Wert der Wahrscheinlichkeitsfunktion an der Stelle n, d.h. für diskrete Verteilungen ist die erzeugende Funktion eine zur Verteilungsfunktion äquivalente Charakterisierung
- Berechnung der Wahrscheinlichkeitsfunktion der Verteilung mit der Erzeugendenfunktion:
  - Gegebene Funktion in eine Potenzreihe entwickeln, mit Hilfe der Taylorreihe
  - $$\hat{f}(z) = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{f}^{(k)}(0)z^k \implies f(k) = \frac{1}{k!}\hat{f}^{(k)}(0) \tag{9}$$
  - oder durch Rückführung auf bekannte Reihen

6.7 Momente diskreter Verteilungen

- Diskreter Wahrscheinlichkeitsraum $(\mathcal{X}, 2^{\mathcal{X}}, P)$, mit $f : \mathcal{X} \to \Re$
- Gesucht: Kenngrößen zur Charakterisierung der allgemeinen Gestalt der Verteilung P
- Wichtigste Kenngrößen: **Mittelwert** und **Varianz**
- **K-te absolute Moment:** $m_k = m_k(P) = \sum_{x \in \mathcal{X}} x^k f(x)$, mit $k = 1$ als **Mittelwert**
- Spezialfall: Für $\mathcal{X} \subset [0, \infty)$ gibt es die **momenterzeugende Funktion:**
  $$M(t) = \sum_{x \in \mathcal{X}} e^{tx} f(x), \text{ mit } t < 1 \tag{10}$$
- Berechnung der abs. Momente: $M^{(K)}(0) := \lim_{t \to 0} M^{(k)}(t) = \ldots = m_k(P)$
- Interpretation: $m_1 = M'(0)$
- **K-ten zentralen Momente:** $\hat{m_k} = \hat{m_k}(P) = \sum_{x \in \mathcal{X}} (x - m_1(P))^k f(x)$, gilt nur, wenn $m_{k-1}(P)$ existiert
- Interpretation: $\hat{m_2} = m_2 - m_1^2$ heißt **Varianz** der Verteilung P (**Steiner'sche Satz**)

6.8 Zufallsvariable

- Zufallsvariable $X : \Omega \to \mathcal{X}$
- Interessante Ereignisse:
  - X nimmt bei der Durchführung des Experiments einen Wert in einer vorgegebenen Menge A an
  - X nimmt bei der Durchführung des Experiments einen Wert $y \in \mathcal{X}$ an
- X heißt **diskrete Zufallsvariable** für alle $A \subset \mathcal{X}$, wenn gilt: $(X \in A) \in \mathcal{A}$ bzw. wenn $\forall y \in \mathcal{X}$ gilt: $(X = y) \in \mathcal{A}$
- Diese Ereignisse werden durch die Urbildmengen von A bzw. $\{y\}$ unter der Abbildung X beschrieben
- Es gilt insbesondere: $(X \in A + B) = (X \in A) + (X \in B)$
- Die Abbildung $P^X : 2^{\mathcal{X}} \to \Re$ definiert durch die Verteilung $P^X(A) = P(X \in A)$ heißt **diskrete Wahrscheinlichkeitsverteilung auf $\mathcal{X}$ mit der Wahrscheinlichkeitsfunktion** $f^X(y) = P(X = y)$
- Gilt: f^X und f^Y stochastisch unabhängig, dann heißt $f^X * f^Y(n) = \sum_{k=0}^{n} f^X(k) f^Y(n-k)$ die **Faltung** der Wahrscheinlichkeitsfunktionen $f^X \text{und} f^Y$ $\to$ gut zur Berechnung von Summen von Zufallsvariable
- Es gilt außerdem: $f^{X+Y} = f^X * f^Y$

6.9 Der Erwartungswert

- Der Mittelwert $\mathcal{E}_P X = \mathcal{E}X = m_1(P^X) = \sum_{y \in \mathcal{X}} y f^X(y)$ der Verteilung von X heißt auch der **Erwartungswert** der Zufallsvariablen X
- Existiert der Erwartungswert einer diskreten Zufallsvariablen X auf einem diskreten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$, so ist: $\mathcal{E}_P X = \sum_{\omega \in \Omega} X(\omega) P\{\omega\}$

7 Handout VII: Geometrische Wahrscheinlichkeiten

- Problem: Überabzählbare Ergebnismengen
- Lösung: Einführung des Konzepts des **Inhalts** $|M|$ einer Menge M
- Messbare Mengen: Länge $|l|$ eines Intervalls ist $|l| = b - a$
- $\Longrightarrow$ Einführung σ-Algebra der n-dimensionalen Borelschen Mengen $\mathcal{B}_n$ (Begriff der **kleinsten σ-Algebra**)
- Abbildung $\lambda : \mathcal{B}_n \to [0, \infty]$ als Maßeinheit für Mengen (**Lebesguesche Maß**)
  - $\lambda(B) \geq 0$
  - $\lambda(l) = |l|$
- Eine Wahrscheinlichkeitsverteilung auf $\mathcal{B}_n$ des $\Re^n$ heißt **n-dimensionale Verteilung**

Definition 2. *(Uniforme Verteilung/Gleichverteilung)*

$$\mathcal{U}(M)\text{-}Verteilung : P(B) = \frac{\lambda(B \cap M)}{\lambda(M)}, (0 < \lambda(M) < \infty; P : \mathcal{B}_n \to \Re)$$

- $\Re : P(A) = \sum_{\omega \in A} f(\omega) \leftrightarrow \Re^n : P(A) = \int_A f(x) dx$

7.1 Lebesgue-Integral

- **Lebesgue-Integral (Treppenfunktion):** $\int f(x)dx := \sum_{k=1}^{m} x_k |I_k|$, mit $f(x) = \sum_{k=1}^{m} x_k 1_{I_k}(x)$
- **L-Integral (nichtnegative Funktion):** $\int f(x)dx := \lim_{k\to\infty} \int f_k(x)dx$
- Kriterien für L-Integrierbarkeit:
  - $0 \leq f_1(x) \leq f_2(x) \leq \ldots$
  - $\lim_{k\to\infty} f_k(x) = f(x)$
  - $\int f_k(x)dx \leq c$
- **Entscheidend:**

$$\lim_{k\to\infty} \int f_k(x)dx = \int \lim_{k\to\infty} f_k(x)dx = \int f(x)dx \tag{11}$$

- Für Borelsche Mengen gilt: $\lambda(B) = \int 1_B(x)dx$, wobei $\int_B f(x)dx := \int 1_B(x)f(x)dx$
- Ist $f(x)$ R-integrierbar auf dem Intervall $[a,b] \subset \Re$ mit $-\infty < a < b < \infty$, dann ist $f(x)$ L-integrierbar auf $[a,b]$:

$$\int_a^b f(x)dx = \int 1_{[a,b]}(x)f(x)dx = \int_{[a,b]} f(x)dx \tag{12}$$

- Uneigentliches Integral: Ist $f(x) \geq 0$ im uneigentlichen Sinne R-integrierbar, so ist sie auch L-integrierbar und es gilt: $\int_{-\infty}^{\infty} f(x)dx = \int f(x)dx$

7.2 Mehrdimensionale Integrale

- **Satz von Fubini:**

$$\int f(x,y)d(x,y) = \int_{R^m} \left(\int_{R^n} f(x,y)dy \right) dx = \int_{R^n} \left(\int_{R^m} f(x,y)dx \right) dy \tag{13}$$

- Zusätzlich können durch den **Parametrisierungssatz** kartesische in z.B. Polarkoordinaten umgerechnet werden (Stichwort: **Funktionaldeterminante**)

8 Handout VIII: Absolutstetige Verteilungen

- **Wahrscheinlichkeitsdichte:**
  - $f(x) \geq 0$, für fast alle x (d.h. Ausnahme: Nullmenge N mit $\lambda(N) = 0$)
  - $\int f(x)dx = 1$
- Mengenfunktion $P : \mathcal{B}_n \to \Re$ definiert durch $P(B) = \int 1_B(x)f(x)dx$, ist eine **n-dimensionale Verteilung**, mit $f : \Re^n \to \Re$ als Wahrscheinlichkeitsdichte
- **Absolutstetige Verteilung:** $P(B) = \int 1_B(x)f(x)dx$
- **Dichte $\mathcal{U}(N)$:**

$$f(x) = \tfrac{1}{\lambda(M)} 1_M(x) = \begin{cases} \frac{1}{\lambda(M)} & \text{, falls } x \in M \\ 0 & \text{, falls } x \notin M \end{cases}$$

- $\implies 1_{A\cap B}(x) = 1_A(x)1_B(x)$ und es folgt: $\int 1_B(x)f(x)dx = \ldots = \frac{\lambda(B\cap M)}{\lambda(M)}$

- $\mathcal{N}(0,1)$**-Verteilung:** $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$
- Es gilt auch: $f(x_1, x_2) = f_1(x_1) f_2(x_2|x_1)$ ist eine zweidimensionale Dichte
- Hängt das Wahrscheinlichkeitsgesetzt x_2 nicht vom Ergebnis des ersten ab, so ist $f_2(x_2|x_1) = f_2(x_2)$
- Analog: Sind zwei Zufallsexperimente unabhängig voneinander durchgeführt, folgt: $f(x_1|x_2) = f_1(x_1) f_2(x_2)$

9 Handout IX: Eindimensionale Verteilungen

- Ist P eine eindim. Verteilung, so heißt die Funktion $F(t) := P(-\infty, t]$ $(F : \Re \to \Re)$ **Verteilungsfunktion** der Verteilung P
- $F(s-0) := P(-\infty, s)$ für $t \to s \implies (-\infty, s] = (-\infty, s) + \{s\}$
- **Exponentialverteilung (**$\mathcal{E}(\lambda)$**-Verteilung):**

$$F(t) = \begin{cases} 0 & \text{, falls } t \leq 0 \\ 1 - e^{\lambda t} & \text{, falls } t > 0 \end{cases}$$

- **Dichte der Exponentialverteilung:**

$$f(x) = \begin{cases} 0 & , x \leq 0 \\ \lambda e^{-\lambda x} & , x > 0 \end{cases}$$

- Ist P absolutstetig, so ist die Verteilungsfunktion Stammfunktion zur Dichte:

$$F(t) = P(-\infty, t] = \int 1_{(-\infty, t]}(x) f(x) dx = \int_{-\infty}^{t} f(x) dx \tag{14}$$

- **Dichte der** $\mathcal{U}[a,b]$**-Verteilung:**

$$f(x) = \begin{cases} 0 & x < a \\ \frac{1}{b-a} & a \leq x \leq b \\ 0 & x > b \end{cases}$$

- **K-ten (absoluten) Momente:** $m_k = m_k(P) = \int x^k f(x) dx$
- **K-ten zentralen Momente:** $\hat{m}_k = \hat{m}_k(P) = \int (x - m_1(P))^k f(x) dx$
- **Mittelwert:** $m_1 = \int x f(x) dx$, **Varianz:** $\hat{m}_2 = \int (x - m_1)^2 f(x) dx$
- **Momenterzeugende Funktion von P:** $M(t) = \int e^{tx} f(x) dx = \int_0^\infty e^{tx} f(x) dx$ (Analog zum diskreten Fall: $M^{(k)}(0) = m_k(P)$)

10 Handout X: Zufallsvariable

10.1 Übersicht

- Zufallsgrößen sind Abbildungen $X : \Omega \to \hat{\Omega}$ von der Ergebnismenge Ω eines Wahrschienlichkeitsraums in eine Menge $\hat{\Omega}$ mit noch zu definierenden Eigenschaften
- Eine Abbildung $X : \Omega \to \hat{\Omega}$ mit $(X \in A) \in \mathcal{A}$ heißt $\mathcal{A}$**-**$\hat{\mathcal{A}}$-messbar für die σ-Algebren $\mathcal{A}$ und $\hat{\mathcal{A}}$

- Eine $\mathcal{A}$-$\hat{\mathcal{A}}$-messbare Abbildung auf $(\Omega, \mathcal{A}, P)$ heißt eine **Zufallsgröße**
- $P^X : \hat{\mathcal{A}} \to \Re$ mit $P^X(A) = P(X \in A)$ heißt **Verteilung der Zufallsgröße X**
- Eine Abbildung $X : \Omega \to \Re$ auf $(\Omega, \mathcal{A}, P)$, die messbar bzgl. $\mathcal{A}$ und $\mathcal{B}$ auf $\Re$ ist, heißt **Zufallsvariable**
- X ist Zufallsvariable auf $(\Omega, \mathcal{A}, P)$, wenn $\forall t \in \Re$ die Urbilder

$$(X \leq t) = (X \in (-\infty, t]) = \{\omega \in \Omega; X(\omega) \leq t\}$$

  in der σ-Algebra $\mathcal{A}$ liegen
- Ist X eine Zufallsvariable, dann existieren die Wahrscheinlichkeiten $P(X \leq t)$ $\forall t \in \Re$ und

$$P(X \leq t) = P(X \in (-\infty, t]) = P^X(-\infty, t] = F^X(t)$$

  die Verteilungsfunktion von X
- **Normalverteilung ($\mathcal{N}(\mu, \sigma^2)$-Verteilung):** $f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t-\mu)^2}$, mit Mittelwert μ und Varianz σ^2
- **Rayleigh-Verteilung:**

$$F(t) = \begin{cases} 0 & , t \leq 0 \\ 1 - e^{-\frac{1}{2}\left(\frac{t}{\beta}\right)^2} & , t > 0 \end{cases}$$

10.2 Funktionen von Zufallsvariablen

- Eine n-dimensionale Zufallsvariable heißt **Zufallsvektor**
- **Kompositionssatz:** Sind $X : \Omega \to \Re^n$ und $G : \Re^n \to \Re^m$ Zufallsvektoren, dann ist auch $Y = G \circ X$ ein Zufallsvektor und Y und G besitzen die gleiche Verteilung: $P^Y = P^G$

11 Handout XI: Funktionen von Zufallsvariablen

11.1 Marginalverteilungen

- **Marginalverteilung von P^X:** Verteilung P^{X_k} der k-ten Komponente eines Zufallsvektors X
- Mit $Z_k(x_1, x_2, \ldots, x_n) = x_k$ lässt sich die k-te Komponente als Komposition darstellen: $X_k = Z_k \circ X$
- Es gilt:

$$P^{X_k}(B) = P^{Z_k}(B) = P^X(Z_k \in B) = \int 1_{Z_k \in B}(x) f(x) dx$$

- Mit Zufallsvektor X gilt für die k-te Marginalverteilung die Dichte (**Marginaldichte**):

$$f_k(x_k) = \underbrace{\int_{\Re} \ldots \int_{\Re}}_{\text{(n-1)-mal}} f(x_1, \ldots, x_n) dx_1 \ldots dx_{k-1} dx_{k+1} \ldots dx_n$$

- **Stochastische Unabhängigkeit:** $f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \ldots f_n(x_n)$, mit f_k als Marginaldichten von f

11.2 Transformationssatz für Dichten

- Problemstellung: Berechnung der Verteilung eines Zufallsvektors $Y = G \circ X$ bzw. die Berechnung der Verteilung des Zufallsvektors G
- Gesucht: Funktion $g(y)$, die der Gleichung $\int_B g(y)dy = \int_{G \in B} f(x)dx$ genügt
- Lösung: **Parametrisierungssatz**: Er besagt, dass wenn: $f(x) = g(G(x))|J_G(x)|$, $g(G(x)) = g(y)$ mit $y = G(x)$ und $x = G^*(y)$ (Umkehrabbildung) nach $g(y)$ aufgelöst werden kann:

$$g(y) = \begin{cases} f(G^*(y))\frac{1}{|J_G(G^*(y))|} & , y \in M^* \\ 0 & , \text{sonst} \end{cases}$$

- Muss die Verteilung einer Abbildung $G_1 : \Re^2 \to \Re^1$ berechnet werden, müssen wir G um einen geeigneten (d.h. Transformationssatzkonform) Zufallsvektor G_2 einführen: $G = (G_1, G_2)$ und anschließend die erste Marginalverteilung P^{G_1} berechnen
- **Affin lineare Transformation:**
  - $y = G(x) = Ax + b$ mit $\left(\frac{\partial G_i}{\partial x_k}(x)\right) = A$, $J_G(x) = det(A)$
  - Ist Matrix A nichtsingulär, d.h. $det(a) \neq 0$, so ist $x = G^*(y) = A^{-1}(y - b)$ $\implies g(y) = \frac{1}{|det(A)|} f(A^{-1}(y - b))$
- **Faltungen:** X_1, X_2 stochastisch unabhängig, dann folgt:

  $g_1(t) = (f_1 * f_2)(t) = \int_{\Re} f_1(t - s)f_2(s)ds$, mit $g_1(t)$ Marginaldichte von g

  Das Integral heißt **Faltungsintegral** und $f_1 * f_2$ die **Faltungsdichte** von f_1 und f_2

12 Handout XII: Erwartungswert und Varianz

12.1 P-Integral

- **Normaldarstellung von X:** $X(\omega) = \sum_k x_k 1_{A_k}(\omega)$, mit $X : \Omega \to \Re$ als **Treppenfunktion**
- **P-Integral von X für Treppenfunktion:** $\int X(\omega)P(d\omega) := \sum_k x_k P(A_k)$, für $\sum_k |x_k|P(A_k) < \infty$
- **Allgemein:** $\int XdP = \int X(\omega)P(d\omega) = \lim_{n\to\infty} \int X_n(\omega)P(d\omega)$
- Rechenregeln:
  - $X(\omega) = 1$ ist P-integrierbar $\implies \int 1dP = 1$
  - X ist P-integrierbar, wenn $|X|(\omega) := |X(\omega)|$ P-integrierbar und es gilt: $|\int XdP| \leq \int |X|dP$
- Besonderheiten:
  - Produktregel erfordert stochastische Unabhängigkeit: $\int(X \cdot Y)dP = \int XdP \cdot \int YdP$
  - Kompositionssatz: $\int Y(\omega)P(d\omega) = \int G(X(\omega))P(d\omega) = \int G(y)P^X(dy)$
  - $m_k(P) = \int x^k P(dx)$ heißt **k-tes Moment** der Verteilung P, mit
    - $m_k(P) = \int y^k f(y)dy$ für absolutstetige Verteilungen und
    - $m_k(P) = \sum x^k f(x)$ für diskrete Verteilungen
- Analog: **K-te zentrale Moment:** $\hat{m}_k(P) = \int(x - m_1(P))^k P(dx)$

12.2 Der Erwartungswert

- **Erwartungswert von X:** $\mathcal{E}X = \mathcal{E}_P X = \int X(\omega)P(d\omega) = m_1(P^X)$
- Allgemein: $\mathcal{E}(X^k) = m_k(P^X)$
- Rechenregeln:
  - $\mathcal{E}1 = 1$
  - $|\mathcal{E}X| \leq \mathcal{E}|X|$
  - $\mathcal{E}(aX + bY + c) = a\mathcal{E}X + b\mathcal{E}Y + c$ und $\mathcal{E}(XY) = (\mathcal{E}X)(\mathcal{E}Y)$
- **Funktionen von Zufallsvariablen:**
  - $\mathcal{E}Y = \int G(X(\omega))P(d\omega) = \int G(x)P^X(dx) = \int G(x)f(x)dx$ oder
  - $\mathcal{E}Y = m_1(P^Y) = \int yg(y)dy$

12.3 Die Varianz

- $\hat{m_2}(P^X) = \int G(x)P^X(dx) = \int Y(\omega)P(d\omega) = \mathcal{E}Y$, mit $G(x) = (x - \mathcal{E}X)^2$
- $\implies \hat{m_2}(P^X) = \mathcal{E}(X - \mathcal{E}X)^2 = \mathrm{var}(X)$ heißt **Varianz** der Zufallsvariablen X
- Rechenregeln:
  - $var(x) = \mathcal{E}(X^2) - (\mathcal{E}X)^2)$
  - $var(aX + b) = a^2 var(X)$
- **Kovarianz:** $cov(X_1, X_2) = \mathcal{E}[(X_1 - \mathcal{E}X_1)(X_2 - \mathcal{E}X_2)]$
- Weitere Rechenregeln:
  - $var(X_1 + X_2) = var(X_1) + 2cov(X_1, X_2) + var(X_2)$
  - $cov(X_1, X_2) = \mathcal{E}(X_1 X_2) - (\mathcal{E}X_1)(\mathcal{E}X_2)$
- X_1 und X_2 stochastisch unabhängig $\implies \mathcal{E}(X_1 X_2) = (\mathcal{E}X_1)(\mathcal{E}X_2)$
- $\implies cov(X_1, X_2) = 0$ sowie $var(X_1 + X_2) = var(X_1) + var(X_2)$
- Rechenregeln, die Dritte:
  - $cov(Y, Y) = \mathcal{E}(Y - \mathcal{E}Y)^2 = var(Y)$
  - $cov(X_1, X_2) = cov(X_2, X_1)$
  - $cov(X_1 + a, X_2 + b) = cov(X_1, X_2)$
  - Bilinearität: $cov(a_1 X_1 + a_2 X_2, Y) = a_1 cov(X_1, Y) + a_2 cov(X_2, Y)$
  - Bilinearität(2): $cov(X, b_1 Y_1 + b_2 Y_2) = b_1 cov(X, Y_1) + b_2 cov(X, Y_2)$
- Durch Bilinearität lässt sich eine cov-Matrix aufspannen: $\sum_{i=1}^m \sum_{k=1}^n a_i b_k cov(X_i, Y_k) = a^\top \mathcal{C}_{XY} b$
- cov-Matrix $\mathcal{C}_X = \begin{pmatrix} cov(X_1, X_1) & cov(X_1, X_2) & \cdots & cov(X_1, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n, X_1) & cov(X_n, X_2) & \cdots & cov(X_n, X_n) \end{pmatrix}$
- C_X ist symmetrisch und positiv definit
- **Korrelationskoeffizient:** $\rho(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{var(X_1)}\sqrt{var(X_2)}}$
- $\rho(X_1, X_2) = \pm 1$ besagt, dass X_1 und X_2 affin linear abhängig sind
- $\rho(X_1, X_2) = 0$ bedeutet, X_1 und X_2 sind **unkorreliert**, d.h. $cov(X, Y) = 0$

13 Handout XIII: Die Gesetze der großen Zahlen

- $S_n = X_1 + \ldots + X_n$ ist die **n-te Partialsumme**, $\bar{X}_n = \frac{1}{n} S_n$ der **n-te statistische Mittelwert** der Folge (X_k)
- Sind X_k unkorreliert, so ist die Kovarianzmatrix eine Diagonalmatrix
- **Schwaches Gesetz der großen Zahlen:** $\lim_{n \to \infty} P(|\bar{X}_n - \mu| \leq \epsilon) = 1$, d.h. $\bar{X}_n$ konvergiert stochastisch gegen $\mu = \mathcal{E} X_k$
- **Starkes Gesetz der großen Zahlen:** $P(\lim_{n \to \infty} \bar{X}_n = \mu) = 1$, d.h. $\bar{X}_n$ konvergiert fast sicher gegen $\mu = \mathcal{E} X_k$
- Das Gesetz der großen Zahlen ist die Verallgemeinerung von: "Die relativen Häufigkeiten konvergieren gegen die Wahrscheinlichkeit."

14 Handout XIV: Die Normalverteilung

- $\mathcal{N}(0,1)$**-Verteilung:**

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

  mit Mittelwert $m_1(P_0) = 0$ und Varianz $\hat{m}_2(P_0) = 1$. Eine $\mathcal{N}(0,1)$-verteilte Zufallsvariable heißt **Gaußsche Einheitsvariable**
- Dichte eines Zufallsvektors $G(x) = Ax + b$: $g(y) = \frac{1}{|det(A)|} f(A^{-1}(y - b))$
- 1-dimensionaler Spezialfall: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$, mit Mittelwert μ und Varianz σ^2, heißt $\mathcal{N}(\mu, \sigma^2)$-Verteilung
- Ist X eine $\mathcal{N}(\mu, \sigma^2)$-verteilte Zufallsvariable, so heißt $E = \frac{1}{\sqrt{\sigma^2}}(X - \mu)$ **Gaußsche Einheitsvariable** bzw. allgemein: $Y = \frac{1}{\sqrt{var(X)}}(X - \mathcal{E}X)$, und dann gilt: $\mathcal{E}Y = 0 \wedge var(Y) = 1$
- Zu jeder $\mathcal{N}(\mu, \sigma^2)$-verteilten Zufallsvariablen X gibt es eine Gaußsche Einheitsvariable E mit $X = \sqrt{\sigma^2} E + \mu$

14.1 Der Zentrale Grenzwertsatz

- **ZGS:** $\lim_{n \to \infty} F_n(t) = \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{s^2}{2}} \, ds$
- **Normierte Partialsummen:** $X_1, X_2, \ldots$ besitzen gleiche Verteilung, mit $\mathcal{E}X_k = \mu \wedge var(X_k) = \sigma^2$, sowie $\mu_n = \mathcal{E}S_n = n\mu \wedge \sigma_n^2 = var(S_n) = n\sigma^2$
- Dann gilt: $S_n^* = \frac{1}{\sqrt{\sigma_n^2}}(S_n - \mu_n) = \frac{1}{\sqrt{n\sigma^2}}(S_n - n\mu)$, d.h. es gilt der ZGS
- **Grenzwertsatz von Moivre und Laplace:** $X_1, X_2, \ldots$ st. unabh. Zufallsvariable, die nur Werte 0 und 1 mit $P(X_k = 1) = p$ und $P(X_k = 0) = 1 - p$ annehmen $\Longrightarrow$ S_n ist binomialverteilt mit $\mathcal{E}S_n = np \wedge var(S_n) = np(1 - p)$
- Dann ist: $S_n^* = \frac{1}{\sqrt{np(1-p)}}(S_n - np)$ und $S_n \leq m \Leftrightarrow S_n^* \leq \frac{m-np}{\sqrt{np(1-p)}}$
- $\Longrightarrow P(S_n \leq m) \approx \Phi\left(\frac{m-np}{\sqrt{np(1-p)}}\right)$, mit Faustregel: $np(1-p) \geq 9$

14.2 Die n-dimensionale Normalverteilung

- Verteilung eines Zufallsvektors E mit st. unabh. $\mathcal{N}(0,1)$-verteilten Komponenten besitzt die Dichte:

$$\varphi_n(y_1, y_2, \ldots, y_n) = \varphi(y_1)\varphi(y_2)\cdots\varphi(y_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}(y_1^2+\ldots+y_n^2)}$$

- Eine Verteilung P_n mit dieser Dichte heißt **n-dimensionale standardisierte Normalverteilung**
- Für Zufallsvektor $X(y) = Ay + b$ mit Dichte

$$f(x) = \frac{1}{|det(A)|}\varphi_n(A^{-1}(x-b)) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sqrt{det(C)}} e^{-\frac{1}{2}(x-b)^\top C^{-1}(x-b)}$$

  heißt **n-dimensionale Normalverteilung** bzw. $\mathcal{N}(b,C)$**-Verteilung**
- Sei $X\mathcal{N}(b,C)$-verteilt. Dann besitzt X eine Darstellung der Form $X = AE + b$
- Zum **Transformationssatz** (analog zum 1-dimensionalen Fall):

$$g(y) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}y^\top(A^\top C^{-1}A)y}$$

  mit $b = \mathcal{E}X$ der Erwartungswert und $C = \mathcal{C}_X$ die Kovarianzmatrix dieses Zufallsvektors
- Ist C eine Diagonalmatrix, so sind die Komponenten $f_n(x_n) = f_k(t) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(t-b_k)^2}{2\sigma_k^2}}$ eines $\mathcal{N}(b,C)$-verteilten Zufallsvektors st. unabh.
- Ist X ein **normalverteilter Zufallsvektor**, so sind seine Komponenten genau dann st. unabh., wenn je zwei verschiedene Komponenten Kovarianz Null besitzen
- **Funktionen von Zufallsvariablen:**
  - Sei $X\mathcal{N}(b,C)$-verteilter Zufallsvektor, dann ist die Zufallsvariable $Y = a_1X_1 + \ldots + a_nX_n + c = A^\top X + c$ normalverteilt mit Mittelwert $\mu = a^\top b + c$ sowie Varianz $\sigma^2 = a^\top Ca$
  - Sei E n-dimensionaler Gaußscher Einheitsvektor und U eine $n\times n$-Orthogonalmatrix, dann ist $H = UE$ ebenfalls ein Gaußscher Einheitsvektor

15 Quellenangabe

Quellenangabe

[1] Inhalt stammt aus den MatheC4-Foliensätzen 2011 (Prof. W. Merz) der FAU-Erlangen