Zusammenfassung zu Stochastik I

Sara Adams

11. Juli 2004

Diese Zusammenfassung basiert auf der Vorlesung
Stochastik I
gehalten im Wintersemester 2003/04
von Prof. Dr. Ludger Overbeck
an der Justus-Liebig Universität Gießen

Inhaltsverzeichnis

1 Stochastik

1.1 Zufallsvariablen

Definitionen

- **Zustandsraum** $\Omega = \{\omega : \omega \in \Omega\}$ Raum der möglichen Zustände
- **Zufallsvariable** $X : (\Omega, \mathcal{A}) \to (M, \mathcal{B})$
- **Urbildmenge** unter B: $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$

1.2 Die σ-Algebra

Definitionen

- Ein System $\mathcal{A} \subseteq$ von Ereignissen heißt **Algebra**, falls gilt:
  - $\Omega \in \mathcal{A}$
  - $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$
  - $A_1, .., A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^n A_i \in \mathcal{A}$
- Algebra $\mathcal{A}$ σ-**Algebra** $\Leftrightarrow [A_i \in \mathcal{A} \Rightarrow \bigcup_{i=1}^\infty A_i \in \mathcal{A}]$ (σ-**Additivität**)
- $\sigma(X) = \{X^{-1}(B) : B \subseteq \mathbb{R}\}$ die von X **erzeugte** σ-**Algebra**

Eigenschaften

- $\mathcal{A} = \{A = X^-1(B) | B \subseteq \mathbb{R}\}$
  - $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$
  - $A_1, .., A_n \in \mathcal{A} \Rightarrow \bigcap A_i \in \mathcal{A}$
- X, Y Zufallsvariablen $\Rightarrow \sigma(X) \subseteq \sigma(X, Y)$

1.3 Verteilung von Zufallsvariablen

Definitionen

- **Verteilung** von X in M: $Q : B \to [0, 1]$
- **Verteilung** von X in Ω: $P : \Omega \to [0, 1]$
- $B_1 \text{x} B_2 = \{(x, y) \in \mathbb{R}^2 : x \in B_1, y \in B_2\}$
- **gemeinsame Verteilung** von X, Y : $Q_{(X,Y)}[B_1 \text{x} B_2] = P\big[\{\omega \in \Omega : X(\omega) \in B_1, Y(\omega) \in B_2\}\big]$
- **Dirac-Verteilung** $\delta_x(A) = 1$ für $x \in A, \delta_x(A) = 0$ für $x \notin A$
- X **Bernoulli-Variable** :$X(\omega) \in \{0, 1\}$, **Erfolgswahrscheinlichkeit** $p = Q[\{1\}] = P[\{\omega \in \Omega : X(\omega) = 1\}],\ Q[\{0\}] = 1 - p$

- Q **Bernoulli-Verteilung** $\Leftrightarrow\ Q = p \cdot \delta_1 + (1 - p) \cdot \delta_0$
- **diskrete Verteilung** $Q \Leftrightarrow p_i = Q[\{a_i\}] = P[\{\omega \in \Omega : X(\omega) = a_i\}], Q = \sum_{i=1}^\infty p_i \cdot \delta_{a_i}$
- **Laplace-Modell**: $M = \{a_1, .., a_k\}, p_i = \frac{1}{k} \Rightarrow Q[B] = \frac{|B|}{|M|}$

Eigenschaften

- $Q[B] = P[X \in B] = P[A],\ A = X^{-1}(B)$
- $x \in \mathbb{R}$
  - $\delta_x(\mathbb{R}) = 1,\ \delta_x(\emptyset) = 0$
  - $\delta(B_1 \cup B_2) = \delta_x(B_1) + \delta_x(B_2),\ \cap B_1 = \emptyset$
  - $\delta_x(\bigcup_{i=1}^\infty) = \sum_{i=1}^\infty \delta_x(B_i), B_i \cap B_j = \emptyset \,\forall\, i \neq j$
  - Die Bernoulli-Verteilung ist normiert und σ-additiv

1.4 Die Wahrscheinlichkeitsverteilung

Definition

- M Messraum, $\mathcal{B} \subseteq\ P(M)$ σ-Algebra. Eine **Wahrscheinlichkeitsverteilung** ist eine auf $\mathcal{B}$ definierte Funktion mit:
  - $Q[M] = 1$
  - $Q[B] \geq 0\ \forall B \in \mathcal{B}$
  - $Q[\bigcup_{i=1}^\infty B_i] = \sum_{i=1}^\infty Q[B_i]\ B_i \cap B_j = \emptyset$

Eigenschaften

- $Q[B^C] = 1 - Q[B]\ \forall B \in \mathcal{B}$
- $Q[\emptyset] = 0$
- $Q[\bigcup_{i=1}^m B_i] = \sum_{i=1}^m Q[B_i]\ B_i \cap B_j = \emptyset\ i \neq j\ i, j = 1, .., m$
- $Q[B_2 \backslash B_1] = Q[B_2] - Q[B_1]\ B_1 \subseteq B_2$
- $Q[B_1 \cup B_2] + Q[B_1 \cap B_2] = Q[B_1] + Q[B_2]$
- allgemeine Additionsformel

$$Q[\bigcup_{i=1}^n B_i] = \sum_{\emptyset \neq T \subseteq \{1,..,n\}} (-1)^{|T|-1} Q[\bigcap_{j \in T} B_j]$$

- $|Q[B_1] - Q[B_2]| \leq |Q[B_1 \Delta B2]\ [B_1 \Delta B_2 = (B_1 \cap B_2^C) \cup (B_2 \cap B_1^C)]$
- Subadditivität: für nichtdisjunkte B_i gilt:

$$Q[\bigcup_{i=1}^n b_i] \leq \sum_{i=1}^n Q[b_i]$$

- σ-Stetigkeit

$$B, B_i \in \mathcal{B},\ B_n \uparrow B\ (B_i \subseteq B_{i+1}, B = \bigcup B_i) \Rightarrow Q[B_n] \uparrow Q[B] \text{ für } n \to \infty$$

$$B, B_i \in \mathcal{B},\ B_n \downarrow B\ (B_i \supset B_{i+1}, B = \bigcap B_i) \Rightarrow Q[B_n] \downarrow Q[B] \text{ für } n \to \infty$$

1.5 Laplace-Modelle

Grundlagen

- $M = \{(i_1, .., in) | i_j \in M_0, j = 1, .., n\} = M_0^n, \quad |M| = |M_0|^n$
- $M = \{(i_1, .., in) | i_j \in M_j, j = 1, .., n\} = M_0^n, \quad |M| = \prod_{i=1}^n |M_i|$
- $M = \{(i_1, .., in) | i_j \in M_0, i_j \neq i_k \text{ für } j \neq k\}$ Permutation, $|M| = n!$

Modelle

- Geordnete Stichprobe ohne Zurücklegen
$|M| = |\{(x_1, .., xk) : x_i \neq xk, 1 \leq x_i \leq n\}| = \frac{n!}{(n-k)!}$
- Geordnete Stichprobe mit Zurücklegen
$|M| = n^k$
- Ungeordnete Stichprobe ohne Zurücklegen
$|M| = |\{\{x_1, .., xk\} \subseteq \{1, .., n\}\}| = \binom{n}{k}$ Binomialkoeffizient C_k^n
- Ungeordnete Stichprobe mit Zurücklegen
$|M| = K_k^n = \binom{n+k-1}{k}$
- Mulitnomialkoeffizient
Es gibt $\frac{n!}{k_i!..k_m!} = \binom{n}{k_1,..,k_m}$ Möglichkeiten n Objekte auf m Gruppen mit jeweils k_i in Gruppe i aufzuteilen, wobei $\sum_{i=1}^m k_i = n$.

1.6 Random Walks

Definition

- **Random Walk** mit Startwert $a : S_i = a + \sum_{j=1}^i X_j, P[X_j = 1] = P[X_j = -1] = \frac{1}{2}$

Eigenschaften

- Nach n Schritten gibt es 2^n verschiedene Pfade.
- Es gibt $\binom{n}{\frac{n+b-a}{2}}$ Pfade von $(0, a)$ nach (n, b).
- $x, y \in N \Rightarrow$
  - $P[\{max_{1\leq i\leq n} S_i \geq x, S_n \geq x + y\}] = P[\{max_{1\leq i\leq n} S_i \geq x, S_n \leq x - y\}]$
  - $P[\{max_{1\leq i\leq n} S_i \geq x, S_n > x + y\}] = P[\{max_{1\leq i\leq n} S_i \geq x, S_n < x - y\}]$
  - $x \in N \Rightarrow P[max_{1\leq i\leq n} S_i \geq x] = 2 \cdot P[S_n > x] + P[S_n = x]$

  - $P[S_1 \neq 0, S_2 \neq 0, .., S_{2n} \neq 0] = P[S_{2n} = 0]$
  - Ballot Theorem: $x \leq n \in N \Rightarrow \exists!\ \frac{x}{n}\binom{n}{\frac{n+x}{2}}$ Pfade von $(0, 0)$ nach (n, x), die immer positiv sind.

1.7 Bedingte Wahrscheinlichkeit

Definitionen

- Seien X,Y Zufallsvaribalen auf $[\Omega, \mathcal{A}, P)$ nach $(M, \mathcal{B})$. Die Wahrscheinlichkeitsverteilung (auf $\mathcal{B}$ mit $A, B \in \mathcal{B}$)

$$Q_0[B] = P[X \in B | Y \in A] = \frac{P[X \in B, Y \in A]}{P[Y \in A]}$$

heißt **bedingte Wahrscheinlichkeit von X gegeben Y** $\in A$

- Für $C, D \in A$ und $P[C] > 0$ ist mit

$$A \ni D \to P[D|C] = \frac{P[D \cap C]}{P[C]}$$

ein neues Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$ definiert.

Eigenschaften

- $P[\bigcap_{i=1}^n A_i] = P[A_1] \cdot P[A_2|A_1] \cdot .. \cdot P[An|A_1 \cap .. \cap A_{n-1}]$
- Regel von der totalen Wahrscheinlichkeit

$$\bigcup_{i=1}^n A_i = \Omega, A_i \cap A_j = \emptyset, P[A_i] > 0 \ \Rightarrow P[B] = \sum_{i=1}^n P[B|A_i] \cdot P[A_i] \ \forall\, B \in \mathcal{B}$$

- Bayessche Regel

$$\bigcup_{i=1}^n A_i = \Omega, A_i \cap A_j = \emptyset, P[A_i] > 0, P[B] > 0 \Rightarrow P[A_j|B] = \frac{P[A_j] \cdot P[B|A_j]}{\sum_{i=1}^n P[A_i] \cdot P[B|A_i]}$$

1.8 Markovketten

Definitionen

- Die Folge $(S_n)_n$ von Zufallsvariblen auf $(\Omega, \mathcal{A}, P)$ hat die **Markoveigenschaft**, falls $\forall n$ gilt:

$$P[S_n \in A | S_1 = x_1, .., S_{n-1} = x_{n-1}] = P[S_n \in A | S_{n-1} = x_{n-1}]$$

mit $x_i \in M_i, A \in \mathcal{B}, M_i$ diskrete Mengen.

- $(S_n)_n$ heißt Markovkette, falls $(S_n)_n$ die Markoveigenschaft erfüllt.
- Eine Folge $(S_n)_n$ von Zufallsvariablen mit Werten in $\{a_1, .., a_m\}$, die markovsch ist, heißt **Markovkette mit Übergangsmatrix** $\Pi = (p(i, j))_{1\leq i,j\leq n}$, falls $P[S_n = a_j | S_{n-1} = a_i] = p(i, j)\ \forall 1 \leq i, j \leq m, \forall n$.

Eigenschaften

- Eine Markovkette mit Übergangsmatrix Π hat zur Zeit n die Verteilung $\vec{P}^n = (P_1^n..P_m^n), p_j^n = P[S_n = j]$ mit $\vec{P}n = \vec{P}^0\Pi^n, \vec{P}^0$ Anfangsverteilung.

1.9 Stochastische Unabhängigkeit

Definitionen

- Die Ereignisse $A_1, .., A_n \in \mathcal{A}$ in $(\Omega, \mathcal{A}, P)$ sind **stochastisch unabhängig**, falls für jedes $T \subseteq \{1, .., n\}$ gilt:
$$P[\bigcap_{i\in T} A_i] = \prod_{i\in T} P[A_i]$$
- Die Zufallsvariblen $X_1, .., Xn$ sind stochastisch unabhängig, falls $\forall B_i \in \mathcal{B}$ gilt:
$$P[X_i \in B_i, 1 \leq i \leq n] = \prod_{i=1}^{n} P[X_i \in B_i]$$
- $X_1, .., Xn$ mit gleicher Verteilung $\Rightarrow$ X_i **identisch verteilt**. X_i unabhängig $\Rightarrow$ X_i **unabhängig identisch verteilt (i.i.d)**

Eigenschaften

- X_i unabhängig mit Werten in $M_i, T_i : M_i \to M_i'$ meßbare Transformationen $\Rightarrow Y_i = T_i(X_i)$ unabhängig

1.10 Faltung von Verteilungen

Eigenschaften

- Seien X_1, X_2 stochastisch unabhängig mit diskreten Verteilungen $Q_{X_j} = \sum_{i=1}^{\infty} p_i^{(j)} \cdot \delta_{a_i^{(j)}}$. Dann gilt:
  - $P[S = x] = \sum_{i=1}^{\infty} p_i^{(1)} \cdot p_{j(i)}^{(2)}, j(i) =$ Index k mit $a_k^{(2)} = x - a_i^{(1)}$, falls existent, sonst ∞ $p_\infty = 0$.
  - $P[S \leq x] = \sum_{i=1}^{\infty} Q_{X_1}(\{a_i^{(1)}\}) \cdot F_{X_2}(x - a_i^{(1)})$ mit $F_{X_2}(y) = P[X_2 \leq y]$
- Seien X_1, X_2 stochastisch unabhängig mit Dichten $f_1, f2$. Dann besitzt $X_1 + X_2$ die Dichte $f_{X_1+X_2}(u) = \int_{-\infty}^{\infty} f_2(v) f_1(u - v)dv = f_1 * f_2(u)$

1.11 Verteilungsfunktion und Quantile

Definitionen

- Verteilungsfunktion: $F : R \to [0, 1], F(x) = P[X \leq x] = Q[(-\infty, x)]$
- X diskret auf $\{a_i\}_{i\in N} \subseteq R, P[X \in \{a_i\}_{i\in N}] = 1 \Rightarrow F(x) = \sum_{i:a_i\leq x} P[X = a_i]$
- X mit Dichte $f \Rightarrow F(x) = \int_{-\infty}^{\infty} f(y)dy,$ F stetig

- F streng monoton wachsend $\Rightarrow F^{-1}(u)$ **u-Quantil** von X, $F(F^{-1}(u)) = u$
- **Quantilsfunktion** $F^{-1}(u) = inf\{x|F(x) \geq u\}, 0 \leq u \leq 1$
- **Median** $F^{-1}(\frac{1}{2})$, **unteres Quantil** $F^{-1}(\frac{1}{4})$, **oberes Quantil** $F^{-1}(\frac{3}{4})$

Eigenschaften

- $x_1 \leq x_2 \Rightarrow F(x_1) \leq F(x_2)$
- F ist rechtsseitig stetig und besitzt linksseitige Limiten $F(x-) = lim_{y\uparrow x}F(y)$
- $lim_{x\to\infty}F(x) = 1,\ lim_{x\to-\infty}F(x) = 0$
- $Q = \sum_{i=1}^{\infty} p_i\delta_{a_i}, a_i < a_{i+1}, g : [0, 1] \to R, g(u) = a_k$ falls $\sum_{i=1}^{k-1} p_i < u \leq \sum_{i=1}^{k} p_i, U \sim U([0, 1]) \Rightarrow X = g(I) \sim Q$
- $X \sim Q, F$ streng monoton steigend, $U \sim U([0, 1]) \Rightarrow X = F^{-1}(U) \sim Q$
- $U \sim U([0, 1]) \Rightarrow X = F^{-1}(U) \sim F$
- $Y = \sigma X + \mu, \sigma > 0 \Rightarrow F_y^{-1}(u) = F_x^{-1}(u) + \mu$

1.12 Integration und Momente

Definitionen

- $\int gdP = \int g(\omega)P[d\omega] = sup\{\int edP|e \leq g, e$ elementar$\}$
- X mit diskreter Verteilung $Q = \sum_{i=1}^{\infty} p_i\delta a_i, T$ Funktion auf $M = \{a_1, a_2, ..\} \Rightarrow T_n(a) = \sum_{i=1}^{n} T(a_i)1_{\{a_i\}}(a) \to \int TdQ(n \to \infty)$
- Q mit Dichte f $\Rightarrow$ $Q[B] = \int 1_B dQ = \int_B f(y)dy = \int 1_B(y)f(y)dy \quad \sum_{i=1}^{m_n} x_i^{(n)} 1_{B_i^{(n)}} \uparrow T(n \to \infty)$
- **m-tes Moment** von X $E[X^m] = \int T^m dQ = \int X^m(\omega)P[d\omega] = \int x^m Q[d\omega]$
- Erwartungswert $E[X]$, Varianz $var(X) = \sigma^2(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$, Standardabweichung $\sigma(X) = \sqrt{\sigma^2(X)}$
- $X, Y : (\Omega, \mathcal{A}, P) \to (M, \mathcal{B}), \bigcup_{i=1}^{\infty} A_i = M \Rightarrow \int T(Y)dP = \sum_{i=1}^{n} \int_{X\in A_i} T(Y)dP, P[X \in A_i] > 0 \Rightarrow E[T(Y)|X \in A_i] = \frac{\int_{X\in A_i} T(y)dP}{P[X\in A_i]}$ Mittelwert von $T(Y)$. falls X nach A_i fällt $\Rightarrow E[T(Y)] = \sum_{i=1}^{n} E[T(Y)|X \in A_i] \cdot P[X \in A_i]$
- X diskret, $A_i = \{x_i\}, X \sim \sum_{i=1}^{\infty} p_i\delta_{x_i}$ $\Rightarrow$ Regressionsfunktion von $T(Y)$ $m(x_i) = E[T(Y)|X = x_i]$
- Fehler/Rauschen $\epsilon : Y = m(X) + (Y - m(X)) = m(X) + \epsilon$ (Unterschied zwischen erwartetem Output und tatsächlichem Output)
- Vorhersageregel $g(X) \Rightarrow E[(Y - g(X))] = \int (Y - g(X))^2 dP$ mittlerer quadratischer Fehler

- $E[X^2], E[Y^2] < \infty \Rightarrow < X, Y >= E[XY] = \int X \cdot Y dP = \int xyQ[dxdy]$
- X Zufallsvariable auf $[0,1], E[X] = \mu \Rightarrow 0 \leq Var(X) \leq \mu(1-\mu) \leq \frac{1}{4}$

Eigenschaften

- $E[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx$
- Transformationssatz $T: R \to R, Y = T(X) \Rightarrow E[Y] = \int Y(\omega) P[d\omega] = \int T(X(\omega)) P[d\omega] = \int T(y) Q_X[dy]$
- $X : (\Omega, \mathcal{A}, P) \to (R, \mathcal{B}), T : (R, \mathcal{B}) \to (R, \mathcal{B}) \Rightarrow E[T(X)] = \int T(X(\omega)) P[d\omega] = \int T(y) Q_X[dy], Q_X[B] = P[X \in B]$, insb. $\int T(y) Q_X[dy] = \int iddQ_y$
- Linearität: $E[X_1], E[X_2] < \infty \Rightarrow E[a \cdot X_1 + b \cdot X_2 + c] = a \cdot E[X_1] + b \cdot E[X_2]$
- Monotonie: $X_1 \leq X_2 \Rightarrow E[X_1] \leq E[X_2]$
- X, Y unabhängig, $E[X] = E[Y] = 0 \Rightarrow E[(X+Y)^3] = E[X^3] + E[Y^3]$
- $E[X^2] < \infty \Rightarrow E[Y] = E[m(X)], E[m(X) \cdot \epsilon] = 0$
- Die optimale Vorhersageregel ist $g = m$.
- $\int |T(y)| Q[dy] < \infty$
  $\Rightarrow \int T(y) Q[dy] = \sum_{i=1}^{\infty} \int_{B_i} T(y) Q[dy], \bigcup_{i=1}^{\infty} = B, B_i$ pw. disjunkt;
  $B_i \downarrow \emptyset \Rightarrow \int_{B_i} T dQ \to 0 (i \to \infty)$
- $X \geq 0$ mit Verteilungsfunktion $F \Rightarrow E[X] = \int_0^{\infty} (1 - F(x)) dx$
- $E[X^2] < \infty \Rightarrow Var(aX + b) = a^2 Var(X), Var(X) \geq 0, Var(X) \leq E[X^2]$
- X, Y unabhängig $\Rightarrow E[XY] = E[X] \cdot E[Y]$
- $S_n = \sum_{i=1}^{n}, X_i$ unabhängig, $\Rightarrow Var(S_n) = \sum_{i=1}^{n} Var(X_i)$
- Tschebyscheff-Ungleichung $P[|X - E[X]| \leq \epsilon] \leq \frac{Var(X)}{\epsilon^2} = \frac{\sigma^2(X)}{\epsilon^2}$
- Markov-Ungleichung $P[|X| \geq \epsilon] \leq \frac{E[|X|^m]}{\epsilon^m} \ \forall \epsilon > 0$
- 3σ-Regel $P[|X - E[X]| \geq 3\sigma] \leq \frac{1}{9}$
-
  - $< X + Y, Z >=< X, Z > + < Y, Z >, < X, Y >=< Y, X >$
  - $< X, X >= \int X^2 dP \geq 0, < aX, Y >= a < X, Y >$
  - $(\int XY dP)^2 =< X, Y >^2 \leq < X, X > \cdot < Y, Y >= \int X^2 dp \cdot \int Y^2 dP$ Cauchy-Schwarz-Ungleichung

1.13 Laplace- und Fourier-Transformierte

Definitionen

- Laplace-Transformierte: $s \to L_X(s) = E[e^{sX}]$
- Wahrscheinlichkeitserzeugende Funktion (WEF): $P[X = j] = p_j, \sum_{j=0}^{\infty} p_j = 1 \Rightarrow f(s) = \sum_{k=0}^{\infty} s^k p_k, s \in [0,1] \Rightarrow p_k = \frac{1}{k!} f^{(k)}(0)$
- Charakteristische Funktion/Fourier-Transformierte:
  $C(s) = C_X(s) = E[e^{isX}] = E[cos(sX)] + iE[sin(sX)]$

Eigenschaften

- $Y = \sigma X + \mu \Rightarrow L_Y(s) = e^{s\mu} L_X(\sigma s)$
- L_X in einer Umgebung von 0 diffbar $\Rightarrow L'_X(0) = E[X]$
- $E[|X|^2] < \infty \Rightarrow C^{(p)}(s) = i^p \int_{\Omega} e^{isX} X^p dP = i^p \int_R e^{isy} y^p Q_x[dy]$
- $E[|X|^2] < \infty \Rightarrow C(s) = \sum_{j=0}^{p} (si)^j \int \frac{X^j}{j!} dP + \int_0^s \frac{(s-x)^{p-1}}{(p-1)!} [C^{(p)}(x) - C^{(p)}(0)] dx$
- $E[|X|^2] < \infty \Rightarrow C(s) = 1 + is \int X dP - \frac{s^2}{2} \int X^2 dP + R_2(s), \ R2(s) \leq \frac{s^2}{2} sup_{0 \leq a \leq 1} \int |e^{iasX} - 1| X^2 dP$
- $s > 0 \Rightarrow \forall \delta \exists \epsilon > 0 : R_2(s) \leq \frac{s^2}{2} (\delta E[X^2] + 2 \int_{\{|X| > \epsilon\}} X^2 dP$
- $X \sim -X \Rightarrow C_X$ reellwertig
- Eindeutigkeitssatz: X, Y reelwertig $\Rightarrow C_X = C_Y \Leftrightarrow Q_X = Q_Y$
- Inversionsformel: C_X integrierbar $\Rightarrow$ X hat Dichte $f, f(x) = \frac{1}{2\pi} \int_{-\infty}^{infty} e^{-iux} C_X(u) du$

1.14 Spezielle Verteilungen

1.14.1 Binomialverteilung

- Die Summe von n unabhängigen Bernoulli-verteilten Zufallsvariablen mit Parameter p hat die Verteilung $P[S_n] = \binom{n}{k} p^k \cdot (1-p)^{n-k}$. Sie heißt **Binomialverteilung** $S_n \sim B(n,p)$
- $X \sim B(n,p) \Rightarrow E[X] = np, Var(X) = np(1-p)$
- $X \sim B(n,p), Y \sim (m,p) \Rightarrow (X+Y) \sim B(n+m,p)$
- $X \sim B(n,p) \Rightarrow C_X(z) = (e^{iz} p + 1 - p)^n$

1.14.2 Poissonverteilung

- X **poissonverteilt** mit Parameter $\mu > 0$ $(X \sim P(\mu))$:
  $P_\mu[X = k] = \frac{\mu^k}{k!} e^{-\mu}, k = 0, 1, ..$
- $X \sim P(\mu) \Rightarrow E[X] = Var(X) = \mu$
- wef$(X) = E[z^X] = e^{\mu(z-1)}$

1.14.3 Gleichverteilung

- X **gleichverteilt** auf [a,b] ($X \sim U([a,b])$), falls Dichte $f(x) = \frac{1}{b-a}$ für $a \leq x \leq b$ und $f(x) = 0$ sonst
- X gleichverteilt auf [0,1]: $0 \leq y_1 \leq y_2 \leq 1 \Rightarrow P[y_1 \leq X \leq y_2] = y_2 - y_1$
- $X \sim U([0,1]) \Rightarrow E[X] = \frac{1}{2}, Var(X) = \frac{1}{12}$

1.14.4 Normalverteilung

- X **normalverteilt** mit Parametern μ, σ^2 ($X \sim N(\mu, \sigma^2)$), falls Dichte $f_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot exp(-\frac{(x-\mu)^2}{2\sigma^2}), \mu \in R, \sigma^2 > 0$
- $X \sim N(\mu, \sigma^2) \Rightarrow P[X \leq x] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} exp(-\frac{(t-\mu)^2}{2\sigma^2})dt$
- $\phi(x) = f_{0,1}(x)$ Dichte der Standardnormalverteilung
- $\Phi(y) = \int_{-\infty}^{y} \phi(x)dx$ Verteilungsfunktion der Standardnormalverteilung
- $\Phi(x) = 1 - \Phi(-x)$
- $X \sim N(0,1) \Rightarrow Y = \sigma X + \mu \sim N(\mu, , \sigma^2), \sigma \in (0, \infty)$
- X_1, X_2 unabhängig, $X_i \sim N(\mu, \sigma^2) \Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
- $X \sim N(\mu, \sigma^2) \Rightarrow E[X] = \mu, Var(X) = \sigma^2$
- $X \sim N(0,1) \Rightarrow C_X(s) = e^{-\frac{s^2}{2}}$

1.14.5 Lognormalverteilung

- X **lognormalverteilt** ($X \sim N(\mu, \sigma^2)) \Rightarrow Y = Exp(X)$
- $P[Y \leq y] = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(\ln(z)-\mu)^2}{2\sigma^2})dz$
- $f(z) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(\ln(z)-\mu)^2}{2\sigma^2})$ für $z > 0$
- $E[X] = exp(\mu + \frac{\sigma^2}{2}), Var(X) = (e^{\sigma^2} - 1)exp(2\mu + \sigma^2)$

1.14.6 Exponentialverteilung

- X **exponentialverteilt** mit Intensitätsrate λ ($X \sim Exp(\lambda)$), falls $f_\lambda(x) = \lambda \cdot e^{-\lambda x}$ für $x \geq 0$, $f_\lambda(x) = 0$ sonst
- Verteilungsfunktion $F_\lambda(x) = P[X \leq x] = 1 - e^{-\lambda x}$
- Survivalfunktion $S_\lambda(x) = P[X \geq x] = e^{-\lambda x}$
- Hazardfunktion $\frac{f_\lambda(x)}{S_\lambda(x)}$
- Gedächtnislosigkeit: $x, y \geq 0 \Rightarrow P[X > x + y | X > x] = P[X > y]$

- $X_i \sim Exp(\lambda)$ $i = 1, .., n$ unabhängig $\Rightarrow M_n = min_{i=1}^n X_i \sim Exp(\sum_{i=1}^n \lambda_i)$
- $X \sim Exp(\lambda) \Rightarrow E[X] = \lambda^{-1}, Var(X) = \lambda^{-2}$
- $F^{-1}(u) = -\frac{\ln(1-u)}{\lambda}$

1.14.7 Gammaverteilung

- X **gammaverteilt**$(X \sim \Gamma(n, \lambda)) : f_{(n,\lambda)}(x) = \frac{1}{(n-1)!}\lambda^n x^{n-1} e^{-\lambda x}$ für $x > 0$
- Gammafunktion $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \alpha > 0$
- $X \sim \Gamma(n, \lambda), Y \sim \Gamma(m, \lambda) \Rightarrow (X + Y) \sim \Gamma(n + m, \lambda)$
- $X \sim \Gamma(\alpha, \lambda), \alpha, \lambda > 0 \Rightarrow f_{(\alpha,\lambda)}(x) = \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x}$
- $E[X] = \frac{1}{\lambda} \cdot \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}$
- $\Gamma(\alpha + 1) = t\Gamma(t), \Gamma(\frac{1}{2}) = \sqrt{\pi}$
- $n \in N \Rightarrow \Gamma(n) = (n-1)!$
- $X \sim \Gamma(\alpha, \lambda) \Rightarrow E[e^{sX}] = \frac{\lambda^\alpha}{(\lambda-s)^\alpha}$

1.14.8 Cauchyverteilung

- X **cauchyverteilt**: $f(x) = \frac{1}{\pi b} \frac{1}{1+(\frac{x-a}{b})^2}$, $b > 0$

1.14.9 Multinomialverteilung

- Zufallsvektor $Y = (Y^1, .., Y^m)$ **multinomialverteilt** zu den Parametern n und $\vec{p} = (p_1, .., p_m)$, $(Y \sim Mult(n, \vec{p}) \Leftrightarrow P[Y^1 = k_1, .., Y^m = k_m] = \binom{n}{k_1, .., k_m} \cdot \prod_{i=1}^m p_i^{k_i}$, $\sum_{i=1}^m k_i = n, \sum_{i=1}^m p_i = 1, p_i \in [0,1], k_i \in \mathbb{N}_0$

1.15 Die Schiefe

Definition

- **Schiefe**: X Zufallsvariable, $\mu = E[X], \sigma^2 = Var(X) \Rightarrow \eta = E[(\frac{X-\mu}{\sigma})^3]$

Eigenschaften

- $\eta = \frac{E[X^3] - 3E[X]E[X^2] + 2E[X]^3}{\sigma^3}$
- $a > 0 \Rightarrow \eta(aX + b) = \eta(X)$
- $S_n = \sum_{i=1}^n X_i, X_i$ iid, $E[X_i] = 0, \Rightarrow \eta(S_n) = \frac{\eta(X_1)}{\sqrt{n}}$
- $S_n \sim B(n,p) \Rightarrow \eta(S_n) = \frac{1-2p}{\sqrt{np(1-p)}}$

2 Grenzwertsätze der Wahrscheinlichkeitstheorie

2.1 Summen unabhängigker Zufallsvariablen

Definitionen

- X_i iid $\Rightarrow \check{S}_n(\omega) = \frac{1}{n}\sum_{i=1}^n x_i = \frac{1}{n}\sum_{i=1}^n X_i(\omega)$
- Eine Folge von Zufallsvariablen $(T_i)_i$ konvergiert gegen T **in Wahrscheinlichkeit/P-stochastisch**, falls $\forall \epsilon > 0 : P[|T_n - T| \geq \epsilon] \to 0 (n \to \infty)$
- **Konvergenz** in $L_2 : E[T_i^2] < \infty, ||T_n - T||_2 \to \infty (n \to \infty)$

Sätze

- Bienaymé: $X_1, .., X_n$ unabhängig, $E[X_i^2] < \infty \Rightarrow Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i)$
- $||\check{S}_n - E[X_1]||_2 = \sqrt{Var(\check{S}_n - E[X_1)} \to 0(n \to \infty)$
- Schwaches Gesetz der Großen Zahl: X_i iid, $E[X_1^2] < \infty \Rightarrow P[\{\omega : |\check{S}_n - E[X_1]| \geq \epsilon\}] \to 0(n \to \infty)$
- (X_i) unabhängig, $S_n = \sum_{i=1}^n X_i \Rightarrow C_{S_n}(s) = \prod_{j=1}^n C_{X_j}(s)$

2.2 Zentraler Grenzwertsatz

Definitionen

- Lindeberg-Bedingung: $\forall \epsilon > 0$:

$$L_n(\epsilon) = \frac{\sum_{j=1}^n \int_{\{|X_j - E[X_j]| > \epsilon\sqrt{\sum_{j=1}^n \sigma_j^2}\}} (X_j - E[X_j])^2 dP}{\sum_{j=1}^n \sigma_j^2} \to 0(n \to \infty)$$

- Feller'sche Bedingung:

$$\lim_{n\to\infty} \frac{max_{1 \leq j \leq n} \sigma(X_j)}{\sigma(\sum_{i=1}^n X_i)} = 0$$

Sätze

- $X_i \sim N(\mu_i, \sigma_i^2) \Rightarrow \frac{\sum_{j=1}^n (X_j - \mu_j)}{\sqrt{\sum_{j=1}^n \sigma_j^2}} \sim N(0,1)$
- Zentraler Grenzwertsatz (ZGWS):
  $(X_i)_i$ Folge unabhängiger Zufallsvariablen, $Var(X_j) = \sigma_j^2 < \infty$, gelte die Lindeberg-Bedingung

$$\Rightarrow \forall x \in R : \lim P[\frac{\sum_{j=1}^n (X_j - E[X_j])}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \leq x] = \Phi(x) = \int_{-\infty}^{infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

  $(X_i)_i$ iid, $\mu = E[X_i, \sigma^2 = Var(X_i)$

$$\Rightarrow L_n(\epsilon) = \sigma^{-2} \int_{\{|X_1 - E[X_1]| \geq \epsilon\sqrt{n\sigma^2}\}} (X_1 - E[X_1])^2 dP \to 0(n \to \infty)$$

2.3 Starkes Gesetz der Großen Zahlen

Definition

-

$$A_n = \{\omega \in \Omega : \check{S}_n(\omega) - E[X_1] \geq \epsilon$$

$$A_\infty = \{\omega \in \Omega : \omega \in A_n \text{für unendlich vielen} \} = \bigcap_{m>1} \bigcup_{n \geq m} A_n$$

Eigenschaften

- Starkes Gesetz der Großen Zahlen: $(X_i)_i$ Folge von iid Zufallsvariablen, $E[X_i] < \infty, \Omega_0 = \{\omega \in \Omega : \lim_{n\to\infty} \check{S}_n(\omega) = E[X_1]\} \Rightarrow P[\Omega_0] = 1$ ($\check{S}_n$ **konvergiert (P)-fast sicher** gegen $E[X_1]$)
- Borel-Cantelli: A_i beliebige Ereignisse, $\sum_{n=1}^\infty P[A_n] < \infty \Rightarrow P[A_\infty] = 0$
- Jensen'sche Ungleichungen: T konkav $\Rightarrow T(E[Y]) \geq E[T(Y)], T$ konvex $\Rightarrow T(E[Y]) \leq E[T(Y)]$

2.4 Empirische Verteilungen

Definitionen

- Stichprobe $(X_1, .., x_n) \Rightarrow Q_n(B) = \frac{1}{n}\sum_{i=1} n 1_B(x_i)$ **Relative Häufigkeit** des Eintretens von B
- $Q_n(\omega) = \sum_{i=1}^n \frac{1}{n}\delta_{X_i(\omega)}$ **Empirische Verteilung** zu $X_1(\omega), .., X_n(\omega), x_i = X_i(\omega)$

Eigenschaften

- $X_1, .., X_n$ iid $\Rightarrow$
  - $E[Q_n(B)] = Q(B)$
  - $Var(Q_n(B)) = \frac{Q(B)(1-Q(B))}{n}$
  - $\lim_{n\to\infty} Q_n(B) = Q(B)$
  - $P[\frac{\sqrt{n}(Q_n(B) - Q(B))}{Q(B)(1-Q(B))} \leq x] \to \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy (n \to \infty) \; \forall x \in R$
- Glivenko-Cantelli: F_n emp. Verteilungsfunktion zu iid $X_1, .., X_n, X_i \sim F$
  $\Rightarrow D_n = sup_{x\in R}|F_n(x) - F(x)| \to 0 \;\; P-$fs

2.5 Der Poisson-Prozeß

Definitionen

- **Poissonprozeß mit Intensität** λ **:**
  $N = (N(t))_{t\geq 0}, \; N(t) = sup\{n : S_n \leq t\}, S_n = \sum_{i=1}^n X_i, X_i \sim Exp(\lambda)$ iid

Eigenschaften

- $n \in N : P[N(t)?n] = e^{\lambda t}\frac{(\lambda t)^n}{n!}(\sim P(\lambda t))$
- $t \downarrow 0 : P[N(t) \geq 2] = 0(t), P[N(t) = 1] = \lambda t + 0(t), P[N(t) = 0] = 1 - \lambda t + 0(t)$
- $P[S_n \leq t, S_{n+1} > t + s] = e^{-\lambda s}e^{-\lambda t}\frac{(\lambda t)^n}{n!}$
- $S_{N(t)+k} - t$ hat die gleiche Verteilung wie S_k, insb. $(S_{N(t)+k} - t) \sim Exp(\lambda)$
- $t, u \geq 0 \Rightarrow N(t-u) - N(t) \sim N(u)$

3 Statistische Methoden

3.1 Maximum-Likelihood-Verfahren

Definition

- **Likelihood-Funktion** $L_n(\theta)(x_1, .., x_n)$:
  - diskret: $L_n = \prod_{i=1}^n p(\theta, x_i)$
  - mit Dichte: $L_n = \prod_{i=1}^n f(\theta, x_i)$
- **Maximum-Likelihood-Estimator** (MLE): MLE=$argmax_{\theta \in \Theta} L_n(\theta)$
- **Kullback-Leiber-Information**:
  - diskret: $K(\theta_0, \theta) = \sum_{i=1}^{\infty} p(a_i, \theta_0) \ln(\frac{p(a_i,\theta_0)}{p(a_i,\theta)})$
  - mit Dichten: $K(\lambda_0, \lambda) = \int_{-\infty}^{\infty} f_{\lambda_0}(x) ln\frac{f_{\lambda_0}(x)}{f_\lambda(x)}dx$
- θ_n **stark konsistent** $\Leftrightarrow \theta_n \to \theta_0 (n \to \infty) P_{\theta_0}$-fs
- **relativer Likelihood-Quotient** $R_n(\theta) = \frac{L_n(\theta)}{L_n(\theta_n)}$
- **Likelihood-Bereich** $I = \{\theta : R_n(\theta) \geq 1 - a\}$
- $I_{jk}(\theta_0) = E_{theta_0}[\frac{dln\left(f(\theta_0,X_1)\right)}{d\theta_j}\frac{dln\left(f(\theta_0,X_1)\right)}{d\theta_k}]$
- **Fischersche Informationsmatrix** $I(\theta_0) = \left(I_{jk}(\theta_0)\right)_{j,k\in\{1,..,n\}}$
- **Chi-Quadrat-Statistik** (χ^2-Statistik: $\sum_{i=1}^m n\frac{(p_j - \hat{p}_j)^2}{p_j}$

Eigenschaften

- $\forall \theta \in \Theta, \theta \neq \theta_0 : K(\theta_0, \theta) > 0$
- $H = \{\theta_1, .., \theta_k\} \Rightarrow \theta_n$ stark konsistener Schätzer
- $I = \{\theta \in \Theta : ln\left(L_n(\theta)\right) \geq ln\left(L_n(\theta_n)\right) + ln(1-a)\} = \{\theta \in \Theta : \frac{1}{2}(\theta - \theta_n)^T\left(\frac{d^2}{d^2\theta}ln\left(L_n(\tilde{\theta})\right)\right)(\theta - \theta_n) \geq ln(1-a)\}$ mit $\tilde{\theta} \in \Theta$

- $I_0 \sim \{\theta_0 \in \Theta : (\theta - \theta_n)^T I(\theta_0)(\theta - \theta_n \leq -\frac{2}{n}ln(1-a)\}$
- $f(\theta, x) = \frac{1}{\sqrt{2\pi\sigma^2}}exp\left(-\frac{1}{2\sigma^2}(x-\theta)^2\right) \Rightarrow I(\theta) = \frac{1}{\sigma^2}$ (Normalvert. mit bekannter Varianz)
- $\forall \theta \in \Theta, \theta \neq \theta_0 : K(\theta_0, \theta) > 0$
- $ln\left(p(\theta, x)\right) = \sum_{i=1}^{m-1} Y^i ln(p_i) + Y^m ln(1 - \sum_{i=1}^{m-1} p_i) \Rightarrow I_{j,k}(\theta) = E[\frac{Y^m \cdot Y^m}{p_m^2}] = \frac{1}{p_m}$ für $j \neq k,\ I_{j,k}(\theta) = \frac{1}{p_j} + \frac{1}{p_m}$ für $j = k \Rightarrow I_0 = \{(p_1, ..p_m) : \sum_{i=1}^m n\frac{(p_j - \hat{p}_j)^2}{p_j} \leq -2ln(1-a)\}$($\hat{p}$ relative Häufigkeit) ($Y \sim Mult(n, \vec{p})$)

3.2 Grundbegriffe der Testtheorie

- H_0 **Hypothese**, H_1 **Gegenhypothese**
- f_1 **Fehler 1. Art**: H_0 ist richtig, wird aber verworfen
- f_2 **Fehler 2. Art**: H_1 ist richtig, wird aber verworfen
- **Entscheidungsregel des Standard-Signifikanztests**: $\exists\ K \subseteq M^n$: Verwerfe $H_0 \Leftrightarrow \vec{X} \in K$, K **kritischer Bereich**
- α **Signifikanzniveau**: $0 < \alpha < 1, P_\theta[\vec{X} \in K] \leq \alpha\ \forall\ \theta \in \Theta_0$
- $\beta(\theta)$ **Gütefunktion** des Tests: $\beta(\theta) = P_\theta[\vec{X} \in K]$
- **Signifikanztest zum Niveau** α : $sup_{\theta \in \Theta_0}\beta(\theta) \leq \alpha$
- $P_\theta[\vec{X} \in K] \leq \alpha\ \forall\ \theta \in \Theta_0 \Leftrightarrow sup_{\theta\in\Theta_0} P_\theta[\vec{X} \in K] = sup_{\theta\in\Theta_0}\beta(\theta) \leq \alpha$

3.3 Neymann-Pearson-Tests für die Normalverteilung

Definitionen

- t: $M^n \to \{0,1\}$ **Test**: $t(\vec{X}) = 1 \Leftrightarrow$ Lehne H_0 ab $\left(t(\vec{X}) = 0 \Leftrightarrow \text{Nehme } H_0 \text{ an}\right)$
- $\tau_\alpha = \left\{t : M^n \to \{0,1\}, P_0[t(\vec{X}) = 1] \leq \alpha\right\}$ Menge aller Tests mt Fehler 1. Art $\leq \alpha$
- $B(a,b)$ **Betafunktion**: $B(a,b) = \int_0^1 t^{a-1} \cdot (1-t)^{b-1}\mathrm{d}t$

Eigenschaften

- **Fundamentalsatz von Neymann-Pearson**: $k \geq 0, M_k = \{\vec{X} : \tilde{f}_1(\vec{X}) > k \cdot \tilde{f}_0(\vec{X})\},\ M_k^+ = \{\vec{X} : \tilde{f}_1(\vec{X}) \geq k \cdot \tilde{f}_0(\vec{X})\}$:
  - $\forall\ \alpha \in (0,1)\ \exists\ k :\ \int_{M_k} \tilde{f}_0(\vec{x})\mathrm{d}\vec{x} \leq \alpha \leq \int_{M_k^+} \tilde{f}_0(\vec{x})\mathrm{d}\vec{x}$
  - $\int t_0(\vec{x}) \cdot \tilde{f}_0(\vec{x})\mathrm{d}\vec{x} = \alpha, t_0(\vec{x}) = 1$ für $\vec{x} \in M_k,\ t_0(\vec{x}) = 0$ für $\vec{x} \notin M_k^+ \Rightarrow \int t_0(\vec{x}) \cdot \tilde{f}_1(\vec{x})\mathrm{d}\vec{x} = sup_{t\in\tau_\alpha} \int t(\vec{x}) \cdot \tilde{f}_1(\vec{x})\mathrm{d}\vec{x}$
- $N_1, .., N_n$ i.i.d $N(0,1), \chi_n^2 = \sum_{i=1}^n N_i^2 \Rightarrow f_{\chi_n^2}(x) = \frac{1}{2^{\frac{n}{2}} \cdot \Gamma(\frac{n}{2})} \cdot x^{\frac{n}{2}-1} \cdot exp(-\frac{x}{2})$ für $x \geq 0,\ f_{\chi_n^2}(x) = 0$ für $x < 0$ (**Chi-Quadrat-Dichte mit n Freiheitsgraden**)
- $B(a,b) = \frac{\Gamma(a)\cdot\Gamma(b)}{\Gamma(a+b)}$