Zusammenfassung Stochastik

§1 Einfache kombinatorische Modelle

Satz: Sei M Ereignismenge der Einzelexperimente („elem. Versuchsausgänge"), Ω Menge aller Versuchsausgänge bei k Wiederholungen des Einzelexperiments. Für die Kardinalität $|\Omega|$ gilt dann, abhängig des Versuchsaufbaus:

- Geordnet, mit Zurücklegen:
  $|\Omega| = |M|^k$
- Geordnet, ohne Zurücklegen:
  $|\Omega| = (|M|)_k$
- Ungeordnet, mit Zurücklegen:
  $|\Omega| = \binom{|M|+k-1}{k-1}$
  (Idee: Codierung des Versuchsausgang als $|M| + k - 1$ bit-Binärwort.)
- Ungeordnet, ohne Zurücklegen:
  $|\Omega| = \binom{|M|}{k}$

Anwendungsbeispiel:

- Lotto:
  $P(\text{(genau) 3 aus 49 richtig}) = \frac{\binom{6}{3}\binom{43}{3}}{\binom{49}{6}}$
  $\hookrightarrow$ MultiHyg
- Kartenspiel: Wkt., beim Skat (10 Karten pro Spieler, insg. 32), genau 2 Buben und genau 2 Asse zu erhalten:
  $p = \frac{\binom{4}{2}\binom{4}{2}\binom{24}{10}}{\binom{32}{10}}$
- Allgemein: Anzahl der Möglichkeiten, eine Menge mit n Elementen auf m Einzelmengen mit jeweils $n_1, n_2, ..., n_m$ Elementen aufzuteilen:
  $\binom{n}{n_1} \cdot \binom{n-n_1}{n_2} \cdot \ldots \cdot \binom{n_m}{n_m} = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_m!}$

§2 Axiomatik

Definition: Sei $\Omega \neq \emptyset$. $\mathfrak{F} \subset \mathcal{P}(\Omega)$ heißt Algebra in Ω, wenn gilt:

1. $\Omega \in \mathfrak{F}$
2. $A \in \mathfrak{F} \Longrightarrow \overline{A} \in \mathfrak{F}$
3. $A, B \in \mathfrak{F} \Longrightarrow A \cup B \in \mathfrak{F}$
   Gilt sogar für jede beliebige Folge $(A_n) \subset \mathfrak{F}$, dass $\bigcup_{n \in \mathbb{N}} A_n \in \mathfrak{F}$, dann heißt $\mathfrak{F}$ auch σ-Algebra in Ω.

Definition: Sei $\mathfrak{F}$ σ-Algebra in $\Omega \neq \emptyset$. Dann nennt man $(\Omega, \mathfrak{F})$ einen messbaren Raum. Ferner heißt eine reelle Funktion $P : \mathfrak{F} \longrightarrow \mathbb{R}$ Wahrscheinlichkeit(smaß), wenn gilt:

1. $P(A) \geq 0 \ \forall A \in \mathfrak{F}$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B) \ \forall A, B \in \mathfrak{F}, A \cap B = \emptyset$

Anwendungsbeispiel:

- „LaPlace"-Raum: Ω ist endlich, d.h. $|\Omega| = N$, $N \in \mathbb{N}$, daher $\mathfrak{F}$ ebenfalls. Es gilt $\forall\, \omega \in \Omega$: $P(\{\omega\}) = N^{-1}$.
- $\mathfrak{F}$ ist bzgl. aller mengentheoretischen Operationen (Differenz, Durchschnitt, etc.) abgeschlossen (Beweis z.B. mit DeMorgan). Dies gilt ebenfalls, wenn $\mathfrak{F}$ σ-Algebra ist.

Satz: Sei $(\Omega, \mathfrak{F}, P)$ W-Raum. Dann gilt: $A \subseteq B \Longrightarrow P(A) \leq (B) \ \forall\ A, B, \in \mathfrak{F}$ (Monotonie).

Satz: Sei $(\Omega, \mathfrak{F}, P)$ W-Raum, $B \in \mathfrak{F},\ P(B) > 0$. Dann ist auch $P(\star|B) :\ \mathfrak{F} \longrightarrow \mathbb{R},\ A \longrightarrow P(A|B)$ ein W-Maß.

Definition: Bedingte Wkt.: $P(A|B) := \frac{P(A \cap B)}{P(B)}$

Satz: Formel der totalen Wkt:
Sei $(A_i)_{i \in I} \subseteq \mathfrak{F},\ A_i$ paarweise disjunkt, mit $I \subseteq \mathbb{N},\ I \neq \emptyset,\ P(A_i) > 0\ \forall i \in I$ und $\Omega = \bigcup A_i$. Dann gilt:
$P(F) = \sum_{i \in I} P(F|A_i) \cdot P(A_i)\ \forall F \in \mathfrak{F}$.

Satz: Bayes-Formel:
$P(A_i|F) = \frac{P(F|A_i) \cdot P(A_i)}{\sum_{j \in I} P(F|A_j) \cdot P(A_j)}$

Satz: $P(\bigcap_{i=1}^{n} A_i) = P(A_1) \cdot P(A_2|A_1) \cdot \ldots \cdot P(A_n|A_1 \cap \ldots \cap A_{n-1})$

Satz: Sei $(\Omega, \mathfrak{F}, P)$ W-Raum, $I \neq \emptyset$ beliebige Indexmenge. $\mathfrak{A} := (A_i)_{i \in I} \subseteq \mathfrak{F}$. Die Familie $\mathfrak{A}$ (bzw. die Ereignisse $A_i,\ i \in I$) heißen (vollständig) unabhängig, wenn für jede endliche Teilfamilie $(A_i)_{i \in J}$, mit $J \subseteq I$ endlich, gilt:
$P(\bigcap_{i \in J} A_i) = \prod_{i \in J} P(A_i)$

Bemerkungen:

- Aus vollst. Unabhängigkeit folgt paarweise Unabhängigkeit (direkt aus der Definition), umgekehrt gilt das jedoch nicht.
- $\perp\!\!\!\perp$ ist keine transitive Relation.
- $(A_i)_{i \in I}$ vollst. $\perp\!\!\!\perp \Longrightarrow (\tilde{A}_i)_{i \in I}$ vollst. $\perp\!\!\!\perp$, mit
  $\tilde{A}_i = \begin{cases} A_i & \text{für gewisse } i \\ \overline{A_i} & \text{für die übrigen } i \end{cases}$

Stochastik für Informatiker, Prof. Dr. Hans M. Dietz, Zusammenfassung von Florian Schoppmann

Definition: Eine σ-Algebra $\mathfrak{F}$ heißt diskret, wenn es eine Familie $(A_i)_{i\in I}$ von Teilmenge von Ω gibt mit $\emptyset \neq I \subseteq \mathbb{N}$, $A_i \cap A_j = \emptyset\ \forall i \neq j$ und $\Omega = \bigcup_{i\in I} A_i$, so dass sich jedes Element $A \in \mathfrak{F}$ als Vereinigung gewisser der A_i darstellen lässt.

Satz: Sei $\emptyset \neq \Omega$, $\mathfrak{F}$ diskrete σ-Algebra in Ω mit erzeugender Zerlegung $(A_i)_{i\in I}$.

i) Ist ein W-Maß P auf $(\Omega, \mathfrak{F})$ gegeben, so bildet $(P(A_i))_{i\in I}$ einen stochastischen Vektor ($\widehat{=}$ eine stochastische Folge).
ii) Gibt man einen stochastischen Vektor $(q_i)_{i\in I}$ vor, so existiert dazu genau ein W-Maß Q mit $Q(A_i) = q_i\ \forall i \in I$. Es gilt zudem: $Q(A) = \sum_{i\in I, A_i\subseteq A} q_i$

Beispiele für diskrete Verteilungen:

- 2-Punkt-Verteilung
- Geometrische Verteilung
- Binomialverteilung, Approximation unter gewissen Vor. durch Poissonverteilung – siehe folgenden Satz
- Hypergeometrische Verteilung („Lottoverteilung"), wird „für große Zahlen" durch die Binomialverteilung approximiert
- Multinomialverteilung
- Multi-Hypergeometrische Verteilung
- Poissonverteilung

Satz: Sei $(p_N)_{N\in\mathbb{N}} \subset [0,1]$ eine Folge derart, dass $\lim_{N\to\infty} N \cdot P_N =: \lambda > 0$ existiert. Dann gilt für alle $k \in \mathbb{N}_0$:
$\lim_{N\to\infty} \mathrm{Bi}(N, p_N)_k = \mathrm{Pois}(\lambda)_k = \frac{\lambda^k}{k!} e^{-\lambda}$

Satz: Der Durchschnitt beliebig vieler σ-Algebren in einer Menge Ω ist wiederum σ-Algebra in Ω.

Definition: Der Durchschnitt aller σ-Algebren in $\mathbb{R}$, die alle halboffenen Intervalle $(a,b],\ a,b \in \mathbb{R}$ enthalten, heißt borelsche σ-Algebra $\mathfrak{B}$. Die zu $\mathfrak{B}$ gehörigen Mengen heißen borelsche Mengen.
Bemerkungen: $\mathfrak{B}$ enthält alle Einpunktmengen, alle Intervalle und alle offenen, abgeschlossenen, kompakten Mengen.

Definition: Eine Funktion $F : \mathbb{R} \longrightarrow \mathbb{R}$, F monoton wachsend, rechtsseitig stetig und $\lim_{x\to-\infty} = 0$, $\lim_{x\to\infty} = 1$, heißt Verteilungsfunktion.
Bemerkungen: Durch die Verteilungsfunktion F wird eindeutig ein W-Maß Q auf $\mathfrak{B}$ gegeben: $Q((a,b]) = F(b) - F(a)\ \forall a,b \in \mathbb{R},\ a < b$.

Satz: Jede Verteilungsfunktion hat höchstens abzählbar viele Sprungstellen.
Beweis: Die Sprünge sind nach absteigende Höhe nummerierbar: Es gibt max. 1 der Höhe $> \frac{1}{2}$, max. 3 mit Höhe $\in (\frac{1}{4}, \frac{1}{2}]$, etc.

Definition: Die Funktionen $F^d, F^c : \mathbb{R} \longrightarrow \mathbb{R}$, $F^d(x) := \sum_{s\leq x} \Delta F(s)$, $F^c := F - F^d$ heißen diskreter Anteil bzw. stetiger Anteil von F.
Eine Verteilungsfunktion heißt diskret, wenn gilt: $F^d = F$.
Sie heißt absolutstetig, wenn $F = F^c$ gilt und zusätzlich eine Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ existiert mit $f \geq 0$ und $F(x) = \int_{-\infty}^{x} f(u)\mathrm{d}u,\ x \in \mathbb{R}$. In diesem Fall heißt jede derartige Funktion f eine Dichte von F.

§3 Zufallsgrößen

Definition: Sei $(\Omega, \mathfrak{F}, P)$ W-Raum und $X : \Omega \longrightarrow \mathbb{R}$ eine (messbare) Abbildung. Dann heißt X Zufallsgröße.

Definition: Sei $X : \Omega \longrightarrow \mathbb{R}$ eine Zufallsgröße mit Verteilungsfunktion F_X. Setze $S_X := \sum_{r\in X(\Omega)} r \cdot P(X = r)$ im diskreten Fall bzw. $S_X := \int_{r\in X(\Omega)} r \cdot F_X'(r)\mathrm{d}r$ im stetigen Fall. Falls $S_X < \infty$ heißt $\mathrm{E}(X) := S_X$ Erwartungswert von X, andernfalls „existiert der Erwartungswert nicht".

Satz: Sei $c \in \mathbb{R}$ bel. aber fest, X, Y Zufallsgrößen. Dann gelten folgende Rechenregeln:

i) $\mathrm{E}(c) = c$
ii) $\mathrm{E}(c \cdot X) = c \cdot \mathrm{E}(X)$
iii) $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$
iv) $\mathrm{E}(X \cdot Y) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$, falls $X \perp\!\!\!\perp Y$
v) $\mathrm{E}((X - \mathrm{E}(X))^2) = \mathrm{E}(X^2 - 2X \cdot \mathrm{E}(X) + \mathrm{E}(X)^2) = \mathrm{E}(X^2) - 2 \cdot \mathrm{E}(x) \cdot \mathrm{E}(X) + \mathrm{E}(X)^2 = \mathrm{E}(X^2) - \mathrm{E}(X)^2$

Anwendungsbeispiel: Erwartungswert der geom. Verteilung (p ist Einzelerfolgswahrscheinlichkeit):

$$\mathrm{E}(X) = \sum_{k=0}^{\infty} k(1-p)^{k-1}p = p \sum_{k=0}^{\infty} \underbrace{k(1-p)^{k-1}}_{=-\frac{\mathrm{d}}{\mathrm{d}p}(1-p)^k}$$

$$\begin{aligned} &= -p\frac{\mathrm{d}}{\mathrm{d}p}\sum_{k=0}^{\infty}(1-p)^k = -p\frac{\mathrm{d}}{\mathrm{d}p}\frac{1}{1-(1-p)} \\ &= -p(-\frac{1}{p^2}) = \frac{1}{p} \end{aligned}$$

Definition: Der Modalwert einer Zufallsgröße ist definiert als die Stelle des größten Sprunges der Verteilungsfunktion im diskreten Fall bzw. des Maximums der Dichte im stetigen Fall.

Definition: Sei X Zufalsgröße mit Verteilungsfunktion F_X und $\alpha \in (0,1)$. Dann heißt $Q_\alpha(X) := \{x \in \mathbb{R} \mid F_X(x-) \leq \alpha \leq F_X(x+)\}$ Menge der α-Quantile von X bzw. F_X.

Definition: Sofern $\mathrm{E}(X)$ existiert, heißt $\mathrm{Var}(X) := \mathrm{D}^2(X) = \mathrm{E}(X-\mathrm{E}(X))^2$ Streuung oder Varianz von X.

Satz: Zu den Zufallsgrößen X, Y existieren die Streuungen. Dann gelten folgende Rechenregeln:

i) $\mathrm{D}^2(c \cdot X) = c^2 \cdot \mathrm{D}^2(X)$
ii) $\mathrm{D}^2(X) = 0 \Longleftrightarrow P(X = \mathrm{E}(X)) = 1$
iii) $\mathrm{D}^2(X+a) = \mathrm{D}^2(X) \quad \forall a \in \mathbb{R}$
iv) $\mathrm{D}^2(X+Y) = \mathrm{D}^2(X) + \mathrm{D}^2(Y) + 2 \cdot \mathrm{E}[(X - \mathrm{E}(X))(Y-\mathrm{E}(Y))]$

Definition: Seien X, Y Zufallsgrößen mit $\mathrm{E}(X), \mathrm{E}(Y) < \infty$. Dann ist die Kovarianz von X, Y definiert als:
$\mathrm{CoV}(X,Y) := \mathrm{E}[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))]$
$= \mathrm{E}(XY) - \mathrm{E}(X)\cdot\mathrm{E}(Y)$
Es ist offensichtlich, dass $\mathrm{CoV}(X,Y) = 0$, falls $X \perp\!\!\!\perp Y$.

Definition: n Zufallsgrößen $X_1, \ldots, X_n$ heißen (vollständig) unabhängig, wenn für jede beliebige Auswahl von Borel-Mengen $B_1, \ldots, B_n$ die Ereignisse $[X_1 \in B_1], \ldots, [X_n \in B_n]$ (vollständig) unabhängig sind.

Satz: Sind $X, Y \perp\!\!\!\perp$, so gilt: $\mathrm{D}^2(X+Y) = \mathrm{D}^2(X) + \mathrm{D}^2(Y)$.

Satz: Sind X, Y stetig verteilt, so gilt: $X \perp\!\!\!\perp Y \Longleftrightarrow$ Man kann $f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y)$ wählen.

Definition: Zwei Zufallsgrößen X, Y heißen identisch verteilt, wenn $F_X = F_Y$ gilt.
Bsp.: $X \mathrel{\widehat{=}}$ Anz. Wappen, $Y \mathrel{\widehat{=}}$ Anz. Zahlen bei 5 Würfen. Es gilt: $X \sim \mathrm{Bi}(5, \frac{1}{2})$, $Y \sim \mathrm{Bi}(5,\frac{1}{2}) \implies X \stackrel{\mathrm{d}}{=} Y$. Dennoch natürlich $X \neq Y$, da $P(X \neq Y) = 1$.

§4 Zentraler Grenzwertsatz

Definition: Sei $X_1, X_2, \ldots$ eine Familie unabhängiger, identisch verteilte Zufallsgrößen mit $0 < \mathrm{D}^2 X_1 < \infty$. Für $S_n = \sum_{k=1}^n X_k$ gilt:
$\lim_{n\to\infty} P(\frac{S_n - n\cdot\mathrm{E}(X_1)}{\sqrt{n\cdot\mathrm{D}^2(X_1)}} \leq x) = \Phi_{0,1}(x)$.
(Levy-Lindeberg)

§5 Schätz- und verwandte Probleme

Satz: Ungleichung von Čebyšev:
$P(|X - \mathrm{E}(X)| \geq \epsilon) \leq \frac{\mathrm{D}^2(X)}{\epsilon^2} \; \forall \epsilon > 0, \; \mathrm{E}(X) < \infty$

Satz: Das schwache Gesetz der großen Zahl:
Seien $X_1, \ldots, X_n$ unabhängige, identisch verteile Zufallsgrößen mit $0 < \mathrm{D}^2 X_1 < \infty$. Sei $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ die „Mittelwert"-Variable. Für jedes $\epsilon > 0$ gilt:
$P(|\bar{X}_n - \mathrm{E}(X_1)| \geq \epsilon) \leq \frac{\mathrm{D}^2(X_1)}{\epsilon^2 n} \; (\stackrel{n\to\infty}{\longrightarrow} 0)$. Letzteres ist Aussage des „Kolmogorov'schen" starken GdgZ.

Anwendungsbeispiel: Für eine unfaire Münze soll $p := P(\text{„Kopf"})$ durch empirische Messung bestimmt werden. Bei Beibehaltung obiger Notation bezeichnet $\bar{X}_n$ den Quotienten $\frac{\text{Anzahl „Kopf"}}{\text{Anzahl der Versuche}}$.

Das GdgZ liefert für $\epsilon > 0$:

$$P(|\bar{X}_n - p| < \epsilon) \geq 1 - \frac{p(1-p)}{\epsilon^2 n} \stackrel{p(1-p)\leq\frac{1}{4}}{\geq} 1 - \frac{1}{4n\epsilon^2}$$

Sei nun bei 1000 Würfen 600 mal „Kopf" eingetreten, also das Ereignis $\bar{X}_n = 0{,}6$. Wähle bspw. $\epsilon = \frac{1}{10}$. Es folgt: $P(|p - 0{,}6| < 0{,}1) \geq 1 - \frac{1}{4\cdot 1000\cdot(1/10)^2} = 0{,}975$.
Also kann nach dem Experiment mit mindestens 97,5 %-iger Sicherheit gesagt werden, dass p im Intervall (0,5; 0,7) liegt.

§6 Zufällige Vektoren und Folgen

Definition: Seien $X_1, X_2, \ldots$ Zufallsgrößen auf $(\Omega, \mathfrak{F}, P)$. Dann nennt man $(X_1, \ldots, X_n)$, $(n \in \mathbb{N})$ zufälligen Vektor und $(X_n)_{n\in\mathbb{N}}$ zufällige Folge oder zufälligen stochastischen Prozess.

Satz: Seien X_1, X_2 unabhängige Zufallsgrößen auf $(\Omega, \mathfrak{F}, P)$. Für die Randverteilungen eines zufälligen Vektors (X_1, X_2) gilt dann:
Im diskreten Fall:
$P(X_1 = a) = \sum_{k \in X_2(\Omega)} P(X_1 = a, X_2 = k)$ (für X_2 analog).

Im stetigen Fall:
$P(X_1 \leq a) = \lim_{b \to \infty} F_{X_1,X_2}((a,b))$
$= \int_{-\infty}^{a} \int_{-\infty}^{\infty} f_{X_1,X_2}(t_1,t_2) \mathrm{d}t_2 \mathrm{d}t_1$
$= \int_{-\infty}^{a} f_{X_1}(t_1) \mathrm{d}t_1 = F_{X_1}(a)$

Definition: Seien X, Y Zufallsgrößen mit $0 < \mathrm{D}^2(X), \mathrm{D}^2(Y) < \infty$. Dann definiert man den Korrelationskoeffizient:
$\rho(X,Y) := \frac{\mathrm{CoV}(X,Y)}{\sqrt{\mathrm{D}^2(X) \cdot \mathrm{D}^2(Y)}}$.

Satz: Es gelten folgende Eigenschaften des Korrelationskoeffizienten:
i) $X \perp\!\!\!\perp Y \implies \rho(X,Y) = 0$
ii) $|\rho(X,Y)| \leq 1$
iii) $|\rho(X,Y)| = 1 \iff X$ und Y sind affin-linear abhängig, d.h. $Y = aX + b$ für geeignete $a, b \in \mathbb{R}$.