Zusammenfassung: Einführung in die Stochastik für Informatiker Prof. Dr. Udo Kamps - SS04

February 15, 2005

Inhaltsverzeichnis

Vorwort

Mit diesem Dokument stelle ich eine grobe Zusammenfassung der Vorlesung "Einführung in die Stochastik für Informatiker" im Sommerster 2004, gehalten von Prof. Dr. Udo Kamps, zur Verfügung. An einigen Stellen habe ich Beispiele aus verschiedenen Quellen eingefügt (z.B. selbige Vorlesung in früheren Semester).

Natürlich gilt auch hier die übliche Haftungsgeschichte (ich übernehme keine Verantwortung für die Fehler und Ähnliches). Falls jemand einen Fehler findet, sollte er mir dies doch bitte mitteilen - christian@steffens.cn

1 Diskrete Wahrscheinlichkeitsräume und deren Erweiterung

- 1. Aspekt: Die möglichen Ereignisse - Merkmalraum Ω
- 2. Aspekt: Die möglichen Fragestellungen - Ereignissystem $\mathcal{A}$
- 3. Aspekt: Die zugehörigen Wahrscheinlichkeiten - Wahrscheinlichtkeit P

$(\Omega, \mathcal{A}$,P) ist ein **diskreter Wahrscheinlichkeitsraum**
$(\Omega, \mathcal{A})$ ist ein **diskreter messbarer Raum**.

1.1 Merkmalraum Ω

1.1.1 Kartesisches Produkt

Das Kartesische Produkt $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ der Mengen $\Omega_1, \Omega_2, \cdots, \Omega_n$ ist die Menge $\Omega = \{(w_1, \cdots, w_n) | w_i \in \Omega_i$ für $i \in (1, \cdots, n)\}$. Statt $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ schreibt man auch $\bigotimes_{i=1}^n \Omega_i$, falls $\Omega_1 = \cdots = \Omega_n$ auch Ω_1^n.

1.2 Regeln von deMorgan

$$\begin{aligned}
(A \cup B)^c &= A^c \cap B^c \\
(A \cap B)^c &= A^c \cup B^c \\
(\bigcup_{i \in I} A_i)^c &= \bigcap_{i \in I} A_i^c \\
(\bigcap_{i \in I} A_i)^c &= \bigcup_{i \in I} A_i^c
\end{aligned}$$

einige Umformungen

- $B \backslash A = B \cap A^c$
- $P(A \backslash B) = P(A) - P(A \cap B)$
- $P(A \cup B) = P(A \cap B^c) + P(B)$
- $A \cap C = A - A \cap C^c$
- $A \cap C = (A \cap C \cap B) \cup (A \cap C \cap B^c)$

1.3 Ereignissystem $\mathcal{A}$

Es gelten folgende Prinzipien nach denen ein Ereignissystem $\mathcal{A}$ ausgewählt wird.

- Für Modelle mit endlichen oder abzählbaren Merkmalraum Ω wählt man für $\mathcal{A}$ die Potenzmenge $\Im(\mathcal{A})$, d.h. alle Teilmengen von Ω sind Ereignisse.
- Für Modelle mit $\Omega = \mathbb{R}$ (oder $\Omega = \mathbb{R}^n$) ist $\mathcal{A} = \Im(\Omega)$ nicht geeignet, weil es soviele Mengen sind. Man wählt für $\mathcal{A}$ stattdessen das System $\mathcal{B}$ der Borel-Mengen.

1.3.1 σ-Algebren

Das **erste Prinzip** verlangt, das man bei der Verknüpfung von Ereignissen wieder Ereignisse erhält. Es muss also die Vereinigung, Durchschnitt und das Komplement von Ereignissen stets wieder Ereignisse ergeben.

Definition: Seien $\Omega \neq \emptyset$ und $\mathcal{A} \in \Im(\Omega)$. $\mathcal{A}$ heisst **σ-Algebra** (von Ereignissen) über Ω falls gilt:

1. $\Omega \in \mathcal{A}$
2. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}\ \forall A \in \mathcal{A}$
3. $(A_n)_{n\in\mathbb{N}} \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_n \in \mathcal{A}$

Wir haben zwar nur die Abgeschossenheit unter der Vereinigung und Komplement gefordert, aber mit den Regeln von deMorgan (1.2, Seite 7) folgt auch die Abgeschlossenheit für den Durschnitt.

1.3.2 Erzeugte σ-Algebra $\mathcal{A}(\mathcal{E})$

Das **zweite Prinzip** verlangt, das man zunächst festlegt, welche Ereignisse mindestens in $\mathcal{A}$ enthalten sein sollen. Dann bestimmt man $\mathcal{A}$ als die kleinste σ-Algebra, die diese vorgegebenen Mengen enthält.

Definition: Ist $\mathcal{E}$ ein System von Teilmengen von Ω, dann heisst die kleinste σ-Algebra $\mathcal{A}$ (über Ω), die $\mathcal{E}$ enthält, die **von $\mathcal{E}$ erzeugte σ-Algebra**. $\mathcal{E}$ heisst Erzeuger von $\mathcal{A}$.

1.3.3 Borel-Mengen

Für das Ereignissystem $\mathcal{A}$ über $\Omega = \mathbb{R}$ wählt man als Erzeuger die Mengen aller Intervalle.

Definition: Es sei $\mathcal{G} := \{(a,b] | a, b \in \mathbb{R}, a \leq b\}$, die Menge der halb-offenen Intervalle in $\mathbb{R}$. Dann heisst die von $\mathcal{G}$ über $\mathbb{R}$ erzeugte σ-Algebra $\mathcal{B}$ die **Borel-σ-Algebra** über $\mathbb{R}$. Die einzelnen Ereignisse aus $\mathcal{B}$ heißen Borel-Mengen.

1.3.4 Spur-σ-Algebra

Seien $\emptyset \neq B \subset \Omega$ und $\mathcal{A}$ eine σ-Algebra über Ω. Dann ist

$$B \cap \mathcal{A} := \{B \cap A | A \in \mathcal{A}\}$$

σ-Algebra über B (nicht über Ω). $B \cap \mathcal{A}$ heisst **Spur-σ-Algebra**.

Beweis:

$B(\cong \Omega) \in B \cap \mathcal{A}$: $B \in B \cap \mathcal{A}$, da $B = B \cap \Omega$ und $\Omega \in \mathcal{A}$ wegen $\mathcal{A}$ σ-Algebra.

$C \in B \cap \mathcal{A} \to C^c \in B \cap \mathcal{A}$: Sei $C \in B \cap \mathcal{A}$. Dann gibt es ein $A \in \mathcal{A}$ so dass $C = B \cap A$. Das Komplement von C bzgl. B (auch hier NICHT bzgl. Ω) ist dann:

$$\begin{aligned} C^c &= B \cap (C)^c \\ &= B \cap (B \cap A)^c \\ &= B \cap (B^c \cup A^c) \\ &= (B \cap B^c) \cup (B \cap A^c) \\ &= \emptyset \cup (B \cap A^c) \\ &= B \cap A^c \in B \cap \mathcal{A} \text{ , da ja } A^c \in \mathcal{A}. \end{aligned}$$

$\{C_n\}_{n\in\mathbb{N}} \in B \cap \mathcal{A} \to \bigcup_{n\in\mathbb{N}} C_n \in B \cap \mathcal{A}$: D.h. es existieren $\{A_n\}_{n\in\mathbb{N}} \in \mathcal{A}$ mit $C_n = B \cap A_n$ $\forall n \in \mathbb{N}$. Dann gilt:

$$\begin{aligned} \bigcup_{n\in\mathbb{N}} C_n &= \bigcup_{n\in\mathbb{N}} (B \cap A_n) \\ &= B \cap (\bigcup_{n\in\mathbb{N}} A_n) \in B \cap \mathcal{A} \text{ , da } \bigcup_{n\in\mathbb{N}} \in \mathcal{A} \end{aligned}$$

1.3.5 $\mathcal{A}_{\bigcap}$ und $\mathcal{A}_{\bigcup}$

Seien $\Omega \neq \emptyset$, $\mathcal{I}$ eine Indexmenge und $\mathcal{A}_i$ für jedes $i \in \mathcal{I}$ eine σ-Algebra über Ω.

$$\begin{aligned} \mathcal{A}_{\bigcap} &:= \bigcap_{i\in\mathcal{I}} \mathcal{A}_i \quad \text{ist } \textbf{eine } \sigma\text{-Algebra} \\ \mathcal{A}_{\bigcup} &:= \bigcup_{i\in\mathcal{I}} \mathcal{A}_i \quad \text{ist } \textbf{keine } \sigma\text{-Algebra} \end{aligned}$$

1.4 Wahrscheinlichkeit P

Definition: Sei $\mathcal{A}$ eine σ-Algebra über $\Omega \neq \emptyset$. Eine Abbildung $P : \mathcal{A} \mapsto [0,1]$ mit

- $P(A) \geq 0 \ \forall A \in \mathcal{A}$
- $P(\Omega) = 1$
- $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ für alle *paarweise disjunkten* $A_i \in \mathcal{A}$ (σ-**Additivität**)

heißt **Wahrscheinlichkeitsverteilung** auf Ω.

1.4.1 Zähldichte

Es sei Ω ein abzählbarer Merkmalraum. Das Ereignissystem sei $\mathcal{A} := \Im(\Omega)$. Ist P ein W-Maß bzw. W-Verteilung über $(\Omega, \mathcal{A})$ und definiert man $f(\omega) := P(\{\omega\})$, mit $\omega \in \Omega$, dann gilt:

- $f(\omega) \geq 0 \quad \omega \in \Omega$
- $\sum_{\omega \in \,\Omega} f(\omega) = 1$
- $P(A) = \sum_{\omega \in \mathcal{A}} f(\omega) \quad A \in \mathcal{A}$

f nennt man **Zähldichte** zu P.

Beispiel: Seien $p \in (0,1)$ und $f : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{R}$ definiert durch

$$f((i,j)) := \begin{cases} 2^{-(j+1)} \cdot (1-p)^{i-j} \cdot p & \text{, für } i \geq j \\ 0 & \text{, sonst} \end{cases}$$

Zu zeigen: f ist Zähldichte einer diskreten Wahrscheinlichkeitsverteilung auf $(\mathbb{N}_0 \times \mathbb{N}_0, \Im(\mathbb{N}_0 \times \mathbb{N}_0))$.

i) $f(i,j) \geq 0 \; \forall i,j \in \mathbb{N}_0 \; \surd$

ii) $\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} f(i,j) = 1$
Beweis:

$$\begin{aligned}
\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} f(i,j) &= \sum_{j=0}^{\infty} \sum_{i=1}^{\infty} f(i,j) \quad \text{(so ist sicher dass } i \geq j) \\
&= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} 2^{-(j+1)} \cdot (1-p)^{i-j} \cdot p \\
&= p \cdot \sum_{j=0}^{\infty} \frac{1}{2^{-(j+1)}} \sum_{i=0}^{\infty} (1-p)^{i-j} \\
&= p \cdot \sum_{j=0}^{\infty} \frac{1}{2^{-(j+1)}} \sum_{i=j}^{\infty} (1-p)^{i}
\end{aligned}$$

$$\begin{aligned}
&= p \cdot \sum_{j=0}^{\infty} \frac{1}{2^{-(j+1)}} \cdot \frac{1}{1-(1-p)} \\
&= p \cdot \sum_{j=0}^{\infty} \frac{1}{2^{-(j+1)}} \cdot \frac{1}{p} \\
&= \frac{1}{2} \cdot \sum_{j=0}^{\infty} \frac{1}{2^j} \\
&= \frac{1}{2} \cdot 2 = 1 \ \surd
\end{aligned}$$

$\Rightarrow f$ ist Zähldichte

1.4.2 Einpunktverteilung

Einfachste diskrete Wahrscheinlichkeitsverteilung:
Sei $\Omega \neq \emptyset$ abzählbar, $\omega \in \Omega$ festgesetzt.

$$\mathcal{E}_\omega(A) = \begin{cases} 1 & \text{, für } \omega \in \text{A} \\ 0 & \text{, sonst} \end{cases}$$

$\mathcal{E}_\omega : \Im(\Omega) \mapsto \mathbb{R}$ heißt Dirac-Verteilung oder **Einpunktverteilung im Punkt** ω.

1.4.3 Träger

Sei $(\Omega, \mathcal{A}, P)$ ein diskreter Wahrscheinlichkeitsraum.

$$T := \text{supp}(P) := \{\omega \in \Omega | P(\omega) > 0\}$$

heißt Träger von P.

1.4.4 Lemma 1.12 - gewichtete Summe von Einpunktverteilungen

Sei $(\Omega, \mathcal{A}, P)$ ein diskreter Wahrscheinlichkeitsraum. Dann gilt: P ist darstellbar als gewichtete Summe von Einpunktverteilungen.

$$P(A) = \sum_{\omega \in \Omega} P(\omega) \cdot \mathcal{E}_\omega(A) \quad \forall A \in \mathcal{A}$$

1.5 Bernoulli-Experiment

Ein Zufallsexperiment mit zwei möglichen Ausgängen heißt **Bernoulli-Experiment**. Der entsprechende Wahrscheinlichkeitsraum: $(\Omega, \mathcal{A}, P)$

- $\Omega = \{0, 1\}$
- $\mathcal{A} = \Im(\Omega)$
- W-Maß P: $P(\{1\}) = p$ und $P(\{0\}) = 1 - p$

1.6 Laplace-Experiment

Ein Zufallsexperiment mit endlich vielen und gleichwertigen Ausgängen heißt Laplace-Experiment.

- $\Omega = \{1, 2, \cdots, N\}$
- $\mathcal{A} = \Im(\Omega)$
- W-Maß P: ergibt sich aus $P(\{1\}) = P(\{2\}) = \cdots = P(\{N\})$ als $P(\{\omega\}) = \frac{1}{N}$

Für die Wahrscheinlichkeit $P(A)$ eines beliebigen Ereignisses A gilt dann:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der günstigen Fälle}}{\text{Anzahl der möglichen Fälle}}$$

Das W-Maß P heisst **Laplace-Verteilung** oder **diskrete Gleichverteilung**.

1.7 Maß über $(\Omega, \mathcal{A})$

Ein Maß auf $\mathcal{A}$ ist eine Abbildung $\mu : \mathcal{A} \mapsto \mathbb{R} \cup \{\infty\}$ mit

- $\mu(A) \geq 0$
- $\mu(\emptyset) = 0$
- $\mu(A_1 + A_2 + \cdots) = \mu(A_1) + \mu(A_2) + \cdots$

Ein Maß kann auch den Wert ∞ annehmen.

2 Grundformen der Kombinatorik

2.1 Ziehen mit Zurücklegen in Reihenfolge

$$\begin{aligned}\Omega_1 &= \{\omega = (\omega_1, \cdots, \omega_k) | \omega_i \in \{1, \cdots, n\}, i \in \{1, \cdots, k\}\} \\ &= \{1, \cdots, n\}^k \\ |\Omega_1| &= n^k\end{aligned}$$

2.2 Ziehen ohne Zurücklegen in Reihenfolge

$$\begin{aligned}\Omega_2 &= \{\omega \in \{1, \cdots, n\}^k | \omega_i \neq \omega_j \text{ für } 1 \leq i \neq j \leq k\} \\ |\Omega_2| &= n \cdot (n-1) \cdots (n-k+1) \\ &= \frac{n!}{(n-k)!} \\ &=: n_k \\ & \quad \text{Falls n=k} \Rightarrow |\Omega_2| = n!\end{aligned}$$

2.3 Ziehen ohne Zurücklegen ohne Reihenfolge

Herleitung: entspricht dem Ziehen ohne Zurücklegen in Reihenfolge, bis auf die Tatsache dass eben hier die Reihenfolge keine Rolle spielt. Es existieren k!-Möglichkeiten die einzlenen Tupel zu permutieren, deshalb teilen wir durch diese Anzahl.

$$\begin{aligned}\Omega_3 &= \{\omega \in \{1, \cdots, n\}^k | \omega_1 < \omega_2 < \cdots < \omega_k\} \\ |\Omega_3| &= \binom{n}{k} \\ &= \frac{n!}{k! \cdot (n-k)!}\end{aligned}$$

Bemerkung: Die Anzahl der möglichen k-elementigen Teilmengen einer n-elementigen Menge ist gleich $\binom{n}{k}$.

2.4 Ziehen mit Zurücklegen ohne Reihenfolge

$$\begin{aligned}\Omega_4 &= \{\omega \in \{1, \cdots, n\}^k | \omega_1 \leq \omega_2 \leq \cdots \leq \omega_k\} \\ |\Omega_4| &= \binom{n+k-1}{k}\end{aligned}$$

2.5 Hypergeometrische-Verteilung

Definition

$$h(s; k, n, S) = \frac{\binom{S}{s} \cdot \binom{n-S}{k-s}}{\binom{n}{k}}$$

s gewünschte Anzahl in unserer Stichprobe

k Umfang der Stichprobe

n Umfang des Raumes

S Anzahl gewünschter Objekte im ganzen Raum

2.5.1 Beispiel

Skat: Wahrscheinlichkeit das jeder Spieler genau ein As hat.

$$\begin{aligned} A_i &\cong \text{Spieler i besitzt ein As} \\ \Rightarrow P(A_1 \cap A_2 \cap A_3) &= P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \end{aligned}$$

Modell 10 Karten an Spieler1, Spieler2 und Spieler3. Der Rest (2 Karten) geht auf den Skat. Ermittlung mit der hypergeometrischen Verteilung.

$$\begin{aligned} P(A_1) &= h(1;10,32,4) = \frac{\binom{4}{1} \cdot \binom{32-4}{10-1}}{\binom{32}{10}} = \frac{\binom{4}{1} \cdot \binom{28}{9}}{\binom{32}{10}} \\ P(A_2|A_1) &= h(1;10,22,3) = \frac{\binom{3}{1} \cdot \binom{22-3}{10-1}}{\binom{22}{10}} = \frac{\binom{3}{1} \cdot \binom{19}{9}}{\binom{22}{10}} \\ P(A_3|A_1 \cap A_2) &= h(1;10,12,2) = \frac{\binom{2}{1} \cdot \binom{12-2}{10-1}}{\binom{12}{10}} = \frac{\binom{2}{1} \cdot \binom{10}{9}}{\binom{12}{10}} \end{aligned}$$

2.5.2 Maximum-Likelihood-Verfahren

Messung der Anzahl Fische im Teich - über eine Art Abschätzng:

- fange S Fische, markiere sie und setze sie wieder aus
- warten, d.h. mischen des Raumes
- fange k Fische, darunter werden sich s makierte befinden

$$\begin{aligned} \Rightarrow \frac{S}{n} &\approx \frac{s}{k} \\ \Rightarrow n &\approx k \cdot \frac{S}{s} \end{aligned}$$

2.6 Binomial-Verteilung

Herleitung Die Binomialverteilung ist eine der wichtigsten diskreten Wahrscheinlichkeitsverteilungen. Sie beschreibt Ergebnisse von Bernoulli-Prozessen.

Ein Bernoulli-Prozess (auch: Bernoulli-Kette) besteht aus einer Abfolge mehrerer, unter gleichbleibenden Bedingungen durchgeführter Bernoulli-Versuche.

Ein Bernoulli-Versuch (auch: Bernoulli-Experiment) ist ein Zufallsversuch mit genau zwei möglichen Ergebnissen. Die Wahrscheinlichkeit dieser Ergebnisse wird durch die Bernoulli-Verteilung beschrieben.

Definition

$$b(s; k, \frac{s}{n}) = \binom{k}{s} \cdot p^s \cdot (1-p)^{k-s} \quad \text{mit } p = \frac{s}{n}$$

2.6.1 Beispiel

Softwarepaket per WWW verteilt. Möglich Übertragungen:

- 8 Datenpakete - keines darf defekt sein
- 16 Datenpakete - max. ein Paket darf defekt sein

Ist die zweite Installationsmethode effizienter als die erste, unter der Betrachtung, dass die Fehlerrate für ein Datenpaket exakt $\frac{1}{12}$ beträgt?

8 Datenpakete:

$$\begin{aligned} b(s;k,p) &= b(0;8,\frac{1}{12}) \\ &= \binom{8}{0} \cdot \frac{1}{12}^0 \cdot (1-\frac{1}{12})^{8-0} \\ &= 1 \cdot 1 \cdot \frac{11}{12}^8 \\ &\approx 0,49853 \end{aligned}$$

16 Datenpakete:

$$\begin{aligned} b(s;k,p) &= b(0;16,\frac{1}{12}) + b(1;16,\frac{1}{12}) \\ &= \binom{16}{0} \cdot \frac{1}{12}^0 \cdot (1-\frac{1}{12})^{16-0} + \binom{16}{1} \cdot \frac{1}{12}^1 \cdot (1-\frac{1}{12})^{16-1} \\ &= \frac{11}{12}^{16} + \frac{16!}{1! \cdot (16-1)!} \cdot \frac{1}{12}^1 \cdot \frac{11}{12}^{15} \\ &\approx 0,2485 + 0,3615 \\ &\approx 0,61 \end{aligned}$$

2.7 Poisson-Verteilung

Herleitung Die Poisson-Verteilung wird häufig dazu benutzt, zeitliche Ereignisse zu beschreiben. Gegeben sind ein zufälliges Ereignis, das durchschnittlich in einem zeitlichen Abstand t_1 stattfindet und ein Zeitraum t_2.

Die Poissonverteilung mit $\lambda = \frac{t2}{t1}$ gibt die Wahrscheinlichkeit dafür an, dass im Zeitraum t_2 genau $0, 1, 2, 3, \cdots, 1000, \cdots$ Ereignisse stattfinden.

Definition Die Wahrscheinlichkeitsverteilung auf $\mathbb{N}$ definiert durch die Zähldichte

$$p(s) = \frac{\lambda^s}{s!} \cdot e^{-\lambda} \qquad s \in \mathbb{N}, \lambda < 0$$

heißt **Poisson-Verteilung mit Parameter** λ.

Bezeichnung: $p_o(s, \lambda)$ bzw. $p_o(\lambda)$

2.7.1 Beispiel

Eine Kaufhaus wird an einem Samstag zwischen 14 und 15 Uhr durchschnittlich alle 10 Sekunden von einem Kunden betreten. Wie hoch ist die Wahrscheinlichkeit das in der nächsten Minute **genau** 9 Leute das Kaufhaus betreten?

$$\begin{aligned} \lambda &= \frac{t_2}{t_1} = \frac{60}{10} = 6 \\ p(9) &= \frac{6^9}{9!} \cdot e^{-6} \\ &= 0,0688348 \approx 6,88\% \end{aligned}$$

2.8 geometrische Verteilung

Definition:

$$geom(p) = p \cdot (1-p)^{k-1} \quad k \in \mathbb{N}$$

Bezeichnung: $geom(p)$

2.8.1 Beispiel

Paketorientiertes Netzwerk, ein einzelnes Datenpaket kommt mit Wahrscheinlichkeit p fehlerfrei am Ziel an. Wenn ein Fehler auftritt wird die Übertragung wiederholt, solange der Empfang bestätigt wurde.

X beschreibe die Anzahl Sendungen eines Paketes. Sei $Z \sim geom(p)$, so gilt $X = 1 + Z$. Für die Zähldichte von X gilt dann:

$$P(X = k) = p \cdot (1-p)^{k-1} \ \ \forall k \in \mathbb{N}$$

Die Übertragung des Paketes wird:

$$\begin{aligned}
&\textit{keinmal wiederholt} && \\
&\qquad Z = 0 &\Rightarrow\ & X = 1 \\
& &\Rightarrow\ & P(X = 1) = p \\
&\textit{zweimal wiederholt} && \\
&\qquad Z = 2 &\Rightarrow\ & X = 3 \\
& &\Rightarrow\ & P(X = 3) = p \cdot (1 - p)^2
\end{aligned}$$

Mit Wahrscheinlichkeit $0,99$ sollen höchstens drei erneute Übertragungen pro Pakte nötig sein, wie groß muss hierfür p sein.

$$\begin{aligned}
Z \leq 3 &\Rightarrow X \leq 4 \\
&\Rightarrow P(X \leq 4) = 0,99
\end{aligned}$$

$$\begin{aligned}
0,99 &= \sum_{k=1}^{4} P(X = k) = \sum_{k=1}^{4} p \cdot (1 - p)^{k-1} \\
&= p \cdot \sum_{k=1}^{4} (1 - p)^{k-1} \\
&= p \cdot \frac{1 - (1 - p)^4}{1 - (1 - p)} = 1 - (1 - p)^4 \\
&\Rightarrow p = 1 - \sqrt[4]{0,01} \\
&\approx 0,68
\end{aligned}$$

3 Eigenschaften von Wahrscheinlichkeitsräumen

3.1 Grenzwerte von Ereignisfolgen

Sei $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$, $\mathcal{A}$ sei σ-Algebra über $\Omega \neq \emptyset$.

$$\lim_{n\to\infty} A_n = \begin{cases} \bigcup_{n=1}^{\infty} A_n & \text{, falls } (A_n)_{n\in\mathbb{N}} \text{ wachsend, } A_n \subseteq A_{n+1} \\ \bigcap_{n=1}^{\infty} A_n & \text{, falls } (A_n)_{n\in\mathbb{N}} \text{ fallend, } A_n \supseteq A_{n+1} \end{cases}$$

3.1.1 limes superior

Für eine beliebige Ereignisfolge $(A_n)_{n\in\mathbb{N}}$ heißt

$$\begin{aligned} \limsup_{n\to\infty} A_n &= \lim_{n\to\infty} (\bigcup_{k=1}^{\infty} A_k) \\ &= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \\ &= \{\omega \in \Omega | \omega \text{ liegt in unendlich vielen der } A_n\} \\ &\cong \text{unendlich viele der } A_i \text{ treten ein} \end{aligned}$$

der **limes superior**.

3.1.2 limes inferior

Für eine beliebige Ereignisfolge $(A_n)_{n\in\mathbb{N}}$ heißt

$$\begin{aligned} \liminf_{n\to\infty} A_n &= \lim_{n\to\infty} (\bigcap_{k=1}^{\infty} A_k) \\ &= \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \\ &= \{\omega \in \Omega | \omega \text{ liegt in allen } A_n \text{ bis auf endlich viele}\} \\ &\cong \text{alle bis auf endlich viele der } A_n \text{ treten ein} \end{aligned}$$

der **limes inferior**.

3.2 Lemma 3.4 - limes und P

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $(A_n)_n \subset \mathcal{A}$. Dann gelten:

- $P(\bigcup_{n=1}^{\infty} A_n) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$, falls $(A_n)_n \nearrow$
- $P(\bigcap_{n=1}^{\infty} A_n) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$, falls $(A_n)_n \searrow$
- $P(\limsup_{n\to\infty} A_n) = \lim_{n\to\infty} P(\bigcup_{k=n}^{\infty} A_k)$

- $P(\liminf_{n\to\infty} A_n) = \lim_{n\to\infty} P(\bigcap_{k=n}^{\infty} A_k)$
- $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$ sub-σ-Additivität

3.3 Siebformel von Sylvester-Pointcarré

Für Ereignisse $(A_n)_{n\in\mathbb{N}}$ in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ gilt:

n=2: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

n=3: $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$

4 Bedingte Wahrscheinlichkeit

4.1 Definition

Seien A,B Ereignisse in Ω und sei $P(B) > 0$. Dann heißt

$$P(A|B) := \frac{P(A \cap B)}{P(B)}$$

die **bedingte Wahrscheinlichkeit von A unter** (der Bedingung) **B**, und es gilt

$$P(A \cap B) = P(B) \cdot P(A|B)$$

4.1.1 Beispiel

	WEIBLICH	MÄNNLICH	$\sum$
PRODUKTE α	0,1	0,2	0,3
PRODUKTE β	0,5	0,2	0,7
$\sum$	0,6	0,4	1

Gesucht: relative Häufigkeit für Verwendung von Produkte β in der Gruppe der Frauen.

$$\begin{aligned} A &\cong \text{Produkt } \beta \\ B &\cong \text{Gruppe der Frauen} \\ P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{0,5}{0,6} = \frac{5}{6} \end{aligned}$$

4.2 Lemma 4.4 - Rechenregeln

Seien $A, B, A_1, A_2, \dots \in \mathcal{A}$ sowie
$P(A) > 0;\ P(B) > 0;\ P(\bigcap_{i=1}^{n} A_i) > 0$

- $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$
- $P(\bigcap_{i=1}^{n} A_i) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdots P(A_n| \bigcap_{i=1}^{n-1} A_i)$

4.3 Lemma 4.6 - Formel für totale Wahrscheinlichkeit

Ist $(B_i, i \in \mathcal{I})$ eine abzählbare Zerlegung von Ω, d.h. es gilt $\Omega = \sum_{i \in \mathcal{I}} B_i$ und kennt man $P(B_i)$ und die bedingten Wahrscheinlichkeiten $P(A|B_i)$ für alle $i \in \mathcal{I}$, dann gilt:

$$P(A) = \sum_{i \in \mathcal{I}} P(A \cap B_i) = \sum_{i \in \mathcal{I}} P(B_i) \cdot P(A|B_i)$$

4.3.1 Beispiel - fehlerhafte Produktionsanlage

Serienartikel wird parallel auf Fertigungsanlagen produziert. Mengenanteile der 3 Anlagen werden als Wahrscheinlichkeiten aufgesetzt:

$$\begin{aligned}
A_i &\cong \text{Artikel wurde auf Anlage i gefertigt}\\
B &\cong \text{Stück ist fehlerhaft}\\
P(A_1) &= 0,3\\
P(A_2) &= 0,2\\
P(A_3) &= 0,5\\
P(B|A_1) &= 0,05\\
P(B|A_2) &= 0,03\\
P(B|A_3) &= 0,09
\end{aligned}$$

Gesucht Wahrscheinlichkeit, dass ein zufällig gegriffenes Stück fehlerhaft ist?

$$\Rightarrow P(B) = \sum_{i=1}^{3} P(B|A_i) \cdot P(A_i) = 0,066$$

4.3.2 Beispiel - faire/unfaire Münze

Eine unter 1.000.000 Münzen hat Zahl auf beiden Seiten. Eine zufällig ausgewählte Münze wird 20-mal geworfen.

Ergebnis: 20 mal Zahl

Frage: Wahrscheinlichkeit dass Münze trotzdem fair.

$$\begin{aligned}
A &\cong \text{faire Münze}\\
B &\cong \text{unfaire Münze}\\
P(A) &= \frac{(10^6)-1}{10^6} = \frac{10^6}{10^6} - \frac{1}{10^6} = 1 - 10^{-6}\\
P(B) &= 10^{-6}\\
Z_{20} &\cong \text{es fällt 20-mal Zahl}
\end{aligned}$$

$$\begin{aligned}
P(Z_{20}) &= P(Z_{20}|A) \cdot P(A) + P(Z_{20}|B) \cdot P(B)\\
&= (\frac{1}{2})^{20} \cdot (1 - 10^{-6}) + 1 \cdot 10^{-6}\\
\Rightarrow P(A|Z_{20}) &= P(Z_{20}|A) \cdot \frac{P(A)}{P(Z_{20})} = 0,4881
\end{aligned}$$

4.3.3 Beispiel - medizinisches Diagnoseverfahren für Krankheit

In 90 % der Fälle wird ein Kranker auch als krank erkannt.
In 5 % der Fälle wird ein Gesunder fälschlicherweise als krank eingestuft.

Modell: 1 % der Bevölkerung leidet an der gesuchten Krankheit.

Gesucht: Wahrscheinlichkeit das zufällig gewählte Person gesund ist, obschon die Diagnose einen positiven Befund liefert.

Lösung:

$$\begin{aligned}
A &\cong \text{Person gesund} \\
B &\cong \text{positiver Befund} \\
P(B|A^c) &= 0,9 \\
P(B|A) &= 0,05 \\
P(A^c) &= 0,01 \\
P(A) &= 0,99
\end{aligned}$$

$$\begin{aligned}
P(B) &= P(B|A^c) \cdot P(A^c) + P(B|A) \cdot P(A) \\
&= 0,9 \cdot 0,1 + 0,05 \cdot 0.99 = 0,0585
\end{aligned}$$

$$\Rightarrow P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)} = 0,846$$

4.4 Bayes´sche Formel

Ist $(B_i, i \in \mathcal{I})$ eine abzählbare Zerlegung von Ω, d.h. es gilt $\Omega = \sum_{i\in\mathcal{I}} B_i$ und kennt man $P(B_i)$ und die bedingten Wahrscheinlichkeiten $P(A|B_i)$ für alle $i \in \mathcal{I}$, dann gilt:

$$P(B_k|A) = \frac{P(B_k) \cdot P(A|B_k)}{\sum_{i\in\mathcal{I}} P(B_i) \cdot P(A|B_i)}$$

Hier wird von der Wirkung A auf die Ursachen B_k zurückgeschlossen.

4.4.1 Beispiel - allgemeiner Aufbau

Arzt stellt Sympton A fest, welches von Krankheiten $B_1, \cdots, B_k$ herrühren kann.

Gegeben

$$\begin{aligned}
P(B_i) &\cong \text{relative Häufigkeit einer Krankheit } B_i \\
P(A|B_k) &\cong \text{relative Häufigkeit des Symptons } A \text{ unter der Krankheit } B_k
\end{aligned}$$

Gesucht Wenn Sympton A auftritt, wie groß ist die Wahrscheinlichkeit für Krankheit B_k? Sprich gesucht ist $P(B_k|A)$

5 Stochastische Unabhängigkeit

5.1 Motivation

Beispiel Urne mit 2 weißen und 3 schwarzen Kugeln. Ziehen von 2 Kugeln mit Zurücklegen.
$\Omega = \{(i,j)|1 \leq i < j \leq 5\}$, $|\Omega| = 5^2 = 25$

weiße Kugeln: Nr.1,2
schwarze Kugeln: Nr.3,4,5

$$\begin{aligned}
A &\cong \text{2.Kugel schwarz} \\
&\Leftrightarrow A = \{(i,j) \in \Omega | j \in \{3,4,5\}\} \\
|A| &= 5 \cdot 3 = 15 \\
\\
B &\cong \text{1.Kugel weiß} \\
&\Leftrightarrow B = \{(i,j) \in \Omega | i \in \{1,2\}\} \\
|B| &= 2 \cdot 5 = 10 \\
\\
A \cap B &\cong \text{1. Kugel weiß und 2.Kugel schwarz} \\
|A \cap B| &= 6 \\
\\
P(A) &= \frac{15}{25} = \frac{3}{5} \\
P(B) &= \frac{10}{25} = \frac{2}{5} \\
P(A \cap B) &= \frac{6}{25} \\
&= P(A) \cdot P(B)
\end{aligned}$$

5.2 Definition

5.2.1 für zwei Ereignisse

Gegeben: Wahscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$
$A, B \in \mathcal{A}$ heißen **stochastisch unabhängig**, falls

$$P(A \cap B) = P(A) \cdot P(B)$$

5.2.2 paarweise stoch. unabh. Familie A_i

Eine Familie $(A_i)_{i \in \mathcal{I}} \subset \mathcal{A}$, $\mathcal{I}$ beliebige Indexmenge, heißt **paarweise stochastisch unabhängig**, falls

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j) \quad \forall i,j \in \mathcal{I}, i \neq j$$

5.2.3 stoch. unabh. Familie A_i

Eine Familie $(A_i)_{i \in \mathcal{I}} \subset \mathcal{A}$, $\mathcal{I}$ beliebige Indexmenge, heißt **(allgemein) stochastisch unabhängig**, falls für jede endliche Auswahl von Ereignissen gilt:

$$P(\bigcap_{j \in \mathcal{J}} A_j) = \prod_{j \in \mathcal{J}} P(A_j) \quad \forall\, \emptyset \neq \mathcal{J} \subset \mathcal{I}, |\mathcal{J}| < \infty$$

5.2.4 Bemerkung: paarweise $\neq$ allgemein stoch. unabh.

Die paarweise stochastische Unabhängigkeit von mehr als zwei Ereignissen hat nicht notwendig die allgemeine stochastische Unabhägnigkeit zur Folge.

Beispiel Werfen von zwei unverfälschten Würfeln. D.h. Laplaceverteilung.

$$\begin{aligned}
\Omega &= \{(i,j)|i,j \in \{1,2,3,4,5,6\}\} \\
A_i &\cong \text{Würfel i zeigt gerade Zahl, } i = 1,2 \\
A_1 \cap A_2 &= \{(2,2),(2,4),(2,6),(4,2),(4,4),(4,6),(6,2),(6,4),(6,6)\} \\
|A_1 \cap A_2| &= 9 \\
P(A_1 \cap A_2) &= \frac{9}{36} = \frac{1}{4} \\
&= P(A_1) \cdot P(A_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \\
&\Rightarrow A_1 \text{ und } A_2 \text{ stoch. unabh.}
\end{aligned}$$

$$A_3 \cong \text{Summe der Augenzahlen gerade}$$

$$\begin{aligned}
A_3 \cap A_1 &= \{(2,2),(2,4),(2,6),(4,2),(4,4),(4,6),(6,2),(6,4),(6,6)\} \\
|A_3 \cap A_1| &= 9 \\
P(A_3 \cap A_1) &= \frac{9}{36} = \frac{1}{4} \\
&= \frac{1}{2} \cdot \frac{1}{2} \\
&= P(A_1) \cdot P(A_3) \\
&\Rightarrow A_1 \text{ und } A_3 \text{ sind stoch. unabh.}
\end{aligned}$$

$$\begin{aligned}
A_3 \cap A_2 &= \{(2,2),(4,2),(6,2),(2,4),(4,4),(6,4),(2,6),(4,6),(6,6)\} \\
|A_3 \cap A_2| &= 9 \\
P(A_3 \cap A_2) &= \frac{9}{36} = \frac{1}{4} \\
&= \frac{1}{2} \cdot \frac{1}{2} \\
&= P(A_2) \cdot P(A_3) \\
&\Rightarrow A_2 \text{ und } A_3 \text{ sind stoch. unabh.}
\end{aligned}$$

D.h. $\{A_1, A_2, A_3\}$ **sind paarweise stochastisch unabhängig, aber nicht allgemein stoch. unabh. denn:**

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap A_2) = \frac{1}{4} \neq P(A_1) \cdot P(A_2) \cdot P(A_3) = \frac{1}{8}$$

5.2.5 Bemerkung: Bedeutung von Wahrscheinlichkeitsverteilung

Die stochastische Unabhängigkeit ist abhängig von der gewählten Wahrscheinlichkeitsverteilung.

Beispiel: Sei $\Omega = \{1,2,3\}$; $A = \{1\}$, $B = \{1,2\}$; e_1 = Einpunktverteilung und Q = Laplaceverteilung.

$$\begin{aligned}
A \cap B &= \{1\} \\
|A \cap B| &= 1 \\
e_1(A \cap B) &= e_1(1) \\
&= 1 \\
&= e_1(A) \cdot e_1(B)
\end{aligned}$$

$$\begin{aligned}
Q(A \cap B) &= \frac{1}{3} \\
&\neq \frac{2}{9} \\
&= \frac{1}{3} \cdot \frac{2}{3} \\
&= Q(A) \cdot Q(B)
\end{aligned}$$

5.2.6 Bemerkung: Relation "stoch. unabh." ist nicht transitiv

Relation "stochastisch unabhängig" ist nicht transitiv, d.h. aus der stoch. unabh. von A_1, A_2 und A_2, A_3 folgt nicht notwendig A_1, A_3 stochastisch unabhängig.

Beispiel: Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$
$\Omega = \{(0,0),(0,1),(1,0),(1,1)\}$
P = Laplace-Verteilung

$$\begin{aligned}
A_1 &= \{(0,0),(0,1)\} \\
A_2 &= \{(0,1),(1,0)\} \\
A_3 &= \{(1,0),(1,1)\} \\
P(A_i) &= \frac{1}{2} \quad \forall i \in \{1,2,3\}
\end{aligned}$$

weiter:

$$\begin{aligned}
P(A_1 \cap A_2) &= P((0,1)) = \frac{1}{4} = P(A_1) \cdot P(A_2) \\
P(A_2 \cap A_3) &= P((1,0)) = \frac{1}{4} = P(A_2) \cdot P(A_3)
\end{aligned}$$

aber:

$$P(A_1 \cap A_3) = P(\emptyset) = 0 \neq P(A_1) \cdot P(A_3)$$

5.3 Satz 5.7

- Sei $(A_i)_{i\in\mathcal{I}}$ eine Familie stoch. unabh. Ereignisse, $k \notin \mathcal{I}$ und $P(A_k) \in \{0,1\}$
  $\Rightarrow (A_i)_{i\in\,\mathcal{I}\cup\{k\}}$ stoch. unabh.
- $(A_i)_{i\in\mathcal{I}}$ sei stoch. unabh. $B_i \in \{A_i, A_i^c, \emptyset, \Omega\}$ $\forall i \in \mathcal{I}$
  $\Rightarrow (B_i)_{i\in\,\mathcal{I}}$ stoch. unabh.
- Sei $\mathcal{I} = \{1,\cdots,n\}$; $n \in \mathbb{N}$ fest gewählt.
  $(A_i)_{i\in\,\mathcal{I}}$ stoch. unabh. $\Leftrightarrow P(\bigcap_{i=1}^n B_i) = \prod_{i=1}^n P(B_i)$ $\forall B_i \in \{A_i, A_i^c\}$ $\forall i \in \mathbb{N}$

5.3.1 Folgerung

Sei die Familie $(A_i)_{i\in\,\{1,\cdots,n\}}$ stoch. unabh., oder kurz $A_1, A_2, \cdots, A_n$ stoch. unabh., dann gilt:

$$\begin{aligned} P(\bigcup_{i=1}^n A_i) &= 1 - P(\bigcup_{i=1}^n A_i)^c \\ &= 1 - P(\bigcap_{i=1}^n A_i^c) \text{ (nach deMorgan, 1.2 - Seite 7)} \\ &= 1 - \prod_{i=1}^n P(A_i^c) \quad \text{(nach 5.3 - Seite 28)} \\ &= 1 - \prod_{i=1}^n (1 - P(A_i)) \end{aligned}$$

5.4 Bemerkung: Konvergenz von $(A_i)_{i\in\,\mathcal{I}}$

Sei $(A_n)_n \subset \mathcal{A}$ Folge von Ereignissen.

$$(A_n)_n \text{ heißt } \textbf{konvergent} \Leftrightarrow \limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$$

Es gilt stets: $\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$
Denn:

$$\begin{aligned}
\omega \in \liminf_{n\to\infty} A_n &= \omega \in \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k \\
&\Rightarrow \exists n_o : \omega \in \bigcap_{k=n_0}^{\infty} A_k \\
&\Rightarrow \omega \in A_k \ \forall k \geq n_0 \\
&\Rightarrow \omega \in \bigcup_{k=n}^{\infty} A_k \ \forall n \geq 1 \\
&\Rightarrow \omega \in \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k \\
&= \omega \in \limsup_{n\to\infty} A_n
\end{aligned}$$

Weiterhin:

$$\begin{aligned}
(\limsup_{n\to\infty} A_n)^c &= (\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k)^c \\
&= \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k^c \\
&= \liminf_{n\to\infty} A_n^c
\end{aligned}$$

5.5 Borel-Cantelli - Lemma 5.14

Seien $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $(A_n)_{n\in\mathbb{N}}$. Dann gilt:

- $\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(\limsup_{n\to\infty} A_n) = 0$
- Ist zusätzlich $(A_n)_n$ stochastisch unabhängig:
  $\sum_{n=1}^{\infty} P(A_n) = \infty \Rightarrow P(\limsup_{n\to\infty} A_n) = 1$

5.5.1 Beweis

i) Wegen $\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k \subset \bigcup_{k=n}^{\infty} A_k \ \forall n \in \mathbb{N}$ ist

$$0 \leq P(\limsup_{n\to\infty} A_n) \leq P(\bigcup_{k=n}^{\infty} A_k) \leq_{\text{sub-}\sigma\text{-Additivität}} \sum_{k=n}^{\infty} P(A_k) \to_{n\to\infty} 0$$

ii) Es ist

$$\begin{aligned}
P(\limsup_{n=1} A_n) &= 1 - P(\limsup_{n=1} A_n)^c \\
&= 1 - P(\liminf_{n=1} A_n^c) \quad \text{(siehe 5.4 Seite 29)} \\
&= 1 - lim_{n\to\infty} P(\bigcap_{k=n}^{\infty} A_k^c) \text{ (siehe 3.1.2 Seite 18)} \\
&= 1 - \lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) \quad \text{(nach 5.3 - Seite 28)} \\
&\geq 1 - \lim_{n\to\infty} exp(-\sum_{k=n}^{\infty} P(A_k)) \\
&= 1
\end{aligned}$$

5.5.2 Beispiel

i) Unendliche Folge von Urnen, je eine Kugel wird gezogen. In der n-ten Urne befinden sich eine weiße und n-1 schwarze Kugeln.

$$\begin{aligned}
A_n &= \text{gezogene Kugel aus n-ter Urne ist weiß} \\
&\Rightarrow P(A_n) = \frac{1}{n} \quad \text{(Ziehungen geschehen unabh. voneinander)} \\
&\Rightarrow \sum_{n=1}^{\infty} P(A_n) = \infty \\
&\Rightarrow P(\limsup_{n\to\infty} A_n) = 1 \quad \text{(siehe oben)}
\end{aligned}$$

D.h. mit Wahrscheinlichkeit 1 werden unendlich viele weiße Kugeln gezogen.

ii) Wie i), aber n-te Urne enthält diesmal $n^2 - 1$ schwarze und eine weiße Kugel.

$$\begin{aligned}
&\Rightarrow \sum_{n=1}^{\infty} P(A_n) < \infty \\
&\Rightarrow P(\limsup_{n\to\infty} A_n) = 0
\end{aligned}$$

D.h. mit Wahrscheinlichkeit 1 werden nur endlich viele weiße Kugeln gezogen.

5.6 Definition: Produkt der Wahrscheinlichkeitsräume

Für diskrete Wahrscheinlichkeitsräume $(\Omega_i, \mathcal{A}_i, P_i)$, $1 \leq i \leq n$, heißt $(\Omega, \mathcal{A}, P)$ mit

- $\Omega = \prod_{i=1}^n \Omega_i = \{\omega = (\omega_1, \omega_2, \cdots, \omega_n) | \omega_i \in \Omega_i, 1 \leq i \leq n\}$
- P definiert durch $P(\omega) = \prod_{i=1}^n P(\omega_i) := \bigotimes_{i=1}^n P_i$

Produkt der Wahrscheinlichkeitsräume $(\Omega_i, \mathcal{A}_i, P_i)$, $1 \leq i \leq n$, Bezeichnung

$$(\Omega, \mathcal{A}, P) = \bigotimes_{i=1}^{n} (\Omega_i, \mathcal{A}_i, P_i)$$

6 Zufallsvariablen auf diskreten Wahrscheinlichkeitsräumen

6.1 $\{X \in \mathcal{A}'\}$ durch X beschreibbar

Ist X eine Abbildung $\Omega \to \Omega'$ und $\mathcal{A}' \subset \Omega'$, dann definiert man

$$\{X \in \mathcal{A}'\} := \{\omega \in \Omega | X(\omega) \in \mathcal{A}'\}$$

Ein Ergebnis der Form $\{X \in \mathcal{A}'\}$ heißt **durch X beschreibbar**.

6.2 Definition: Zufallsvariable

Urbildfunktion Sei $T : \Omega \to \Omega'$ eine Abbildung.

$$T^{-1} = \begin{cases} \mathfrak{P}(\Omega') \to \mathfrak{P}(\Omega) \\ A' \to T^{-1}(A') := \{\omega \in \Omega | T(\omega) \in A'\} \end{cases}$$

heißt die zu T gehörende Urbildfunktion.

Messbare Abbildung $(\Omega, \mathcal{A})$ heißt Messraum. Sind $(\Omega, \mathcal{A})$ und $(\Omega', \mathcal{A}')$ Messräume und $X : \Omega \to \Omega'$ mit

$$X^{-1}(A') = \{X \in A'\} \in \mathcal{A} \quad \forall A' \in \mathcal{A}'$$

so ist X messbar.

Herleitung Eine messbare Funktion von einem Wahrscheinlichkeitsraum in einen anderen heißt Zufallsvariable, bzw. Zufallsvektor falls $\Omega' = \mathbb{R}^n$.

Zufallsvariable ist eine Abbildung vom Merkmalraum Ω in eine Bildmenge Ω'. Gilt $\mathcal{A} \neq \mathfrak{P}(\Omega)$ und ist $\mathcal{A}'$ das Ereignissystem von Ω', dann wird für eine Zufallsvariable $X : \Omega \to \Omega'$ gefordert:

$$\{\omega \in \Omega | X(\omega) \in A'\} = \{X^{-1}(A')\} = \{X \in A'\} \in \mathcal{A} \quad \forall A' \in \mathcal{A}'$$

6.2.1 Beispiel 2.9 - Aktienkurse

Zufallsvariablen:

$$\begin{aligned} X_1 &\cong \text{"gestriger Aktienkurs"} \\ X_2 &\cong \text{"heutiger Aktienkurs"} \\ A &\cong \text{"der Kurs lag gestern über 500"} \\ B &\cong \text{"der Kurs ist von gestern auf heute gestiegen"} \\ A &= \{X_1 > 500\} \\ B &= \{X_1 < X_2\} \end{aligned}$$

6.3 Definition 6.4 - P^X Verteilung

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \Omega'$ eine Zufallsvariable. Dann ist

$$A' \to P^X(A') := P(X^{-1}(A')) = P(X \in A') \text{ mit } A' \in \mathcal{A}'$$

ein Wahrscheinlichkeitsmaß über $(\Omega', \mathcal{A}')$. P^X **heißt Verteilung von** X **bzgl.** P.

6.4 Zähldichte f^X

$X : \Omega \to \Omega'$ Zufallsvariable und Ω' abzählbar, dann hat P^X die Zähldichte (siehe 1.4.1 - Seite 10) f^X:

$$f^X(\omega') = P(X = \omega') \text{ mit } \omega' \in \Omega'$$

6.5 Verteilungsfunktion F^X

Sei X eine reellwertige ZV, dann hat P^X die Verteilungsfunktion

$$F^X(t) = P(X \leq t) \; t \in \mathbb{R}$$

6.5.1 Beispiel:

Es seien X, Y zwei stoch. unabh. ZV mit Verteilungsfunktion F^X und F^Y.

- Berechne die Verteilungsfunktion von $min(X, Y)$ und $max(X, Y)$ in Abhängigkeit von F^X und F^Y.
- Berechne obere Verteilungsfunktion für $X, Y \cong geom(p)$.

i)

$$\begin{aligned} F^{max(X,Y)}(t) &= P(max(X,Y) \leq t) \\ &= P(X \leq t, Y \leq t) \\ &= P(X \leq t) \cdot P(Y \leq t) \text{ da, } X, Y \text{ stoch. unabh.} \\ &= F^X(t) \cdot F^Y(t) \text{ für } t \in \mathbb{R} \end{aligned}$$

$$F^{min(X,Y)}(t) = 1 - [(1 - F^X(t)) \cdot (1 - F^Y(t))]$$

ii)

$$
\begin{aligned}
F^X(n) &= P(X \leq n) \\
&= \sum_{k=1}^{n} p \cdot (1-p)^{k-1} \\
&= p \cdot \sum_{k=1}^{n} (1-p)^{k-1} \\
&= p \cdot \sum_{k=1}^{n-1} (1-p)^{k} \\
&= p \cdot \sum_{k=0}^{n-1} (1-p)^{k} \quad \text{(für } k=0 \text{ ist die } \sum \text{ gleich null)} \\
&= p \cdot \frac{(1-p)^n - 1}{(1-p) - 1} \\
&= 1 - (1-p)^n
\end{aligned}
$$

6.6 Definition 6.9 - Randverteilung

Sei $\Omega = \Omega_1 \times \cdots \times \Omega_n = \bigotimes_{i=1}^{n} \Omega_i$ dann heißt die ZV

$$X_i : \Omega \to \Omega_i \ ; \ \omega \mapsto \omega_i$$

die i-te Projektion. Die Verteilung P^{X_i} heißt die **i-te Randverteilung**.

6.7 Definition 5.12 - stoch. unabh. ZV

$$\text{Die ZV } X_i; i \in \mathcal{I} \text{ sind stoch. unabh. } \Leftrightarrow$$

$$P(\bigcap_{i \in \mathcal{J}} \{X_i \in A_i\}) = \prod_{i \in \mathcal{J}} P(X_i \in A_i) \quad \forall \mathcal{J} \subset \mathcal{I} : |\mathcal{J}| < \infty \quad \text{und } \forall A_i \in \mathcal{A} \ , i \in \mathcal{J}$$

6.8 Faltung von ZV

Die Verteilung der Summe von stochastisch unabhängigen ZV heißt die **Faltung der Einzelwahrscheinlicheiten**: X, Y seien stoch. unabh. ZV auf $\mathbb{Z}$ mit den Zähldichten f bzw. g. D.h.

$$
\begin{aligned}
P(X = n) &=: f(n) \\
P(Y = n) &=: g(n)
\end{aligned}
$$

Dann hat $X+Y$ die Zähldichte h gegeben durch

$$\begin{aligned} h(k) &= \sum_{j\in\mathbb{Z}} f(j)\cdot g(k-j) \\ &= \sum_{j\in\mathbb{Z}} f(k-j)\cdot g(j) \\ &= P(X+Y=k) \end{aligned}$$

h ist die Faltung der Dichten f und g.

6.9 gedächtnislose Zufallsvariablen

Die Verteilung einer ZV X, mit $X \geq 0$, heißt **gedächtnislos**, wenn

$$P(X > x+y | X > x) = P(X > y) \quad \forall x, y \geq 0$$

7 Erwartungswerte

7.1 Motivation

Bei einem Examen seien die Noten 1 bis 5 vergeben. Mit relativer Häufigkeit 0,11 / 0,23 / 0,31 / 0,27 / 0,8. Der Mittelwert wäre dann

$$1 \cdot 0,11 + 2 \cdot 0,23 + 3 \cdot 0,31 + 4 \cdot 0,27 + 5 \cdot 0,8 = 2.98$$

7.2 Erwartungswert EX

Es sei $X : \Omega \to \Omega' \subset \mathbb{R}$ eine **diskrete** Zufallsvariable mit $X \geq 0$ oder Ω' endlich. Dann heißt

$$EX := \sum_{k \in \Omega'} k \cdot P(X = k) = \sum_{k \in \Omega'} k \cdot P^X(k) = \sum_{k \in \Omega'} k \cdot f^X(k)$$

der **Erwartungswert von X** (oder P^X).

7.2.1 Beispiel Erwartungswerte

Mit dieser Definition kann man für bereits alle bisherigen **diskreten** Verteilungen den Erwartungswert berechnen.

Name	EW
Laplace-Verteilung	$EX = \frac{N+1}{2}$
Einpunktverteilung	$EX = a$
Bernoulli-Versuch	$EX = p$
Binomial-Verteilung	$EX = n \cdot p$
Hypergeometrische-Verteilung	$EX = n \cdot \frac{S}{s}$
Poisson-Verteilung	$EX = \lambda$
Geometrische-Verteilung	$EX = \frac{1}{p}$

7.3 Eigenschaften des Erwartungswertes

- Der Erwartungswert ist **monoton**: Aus $X \leq Y$ folgt $EX \leq EY$, falls EX und EY existieren. Weiter folgt aus $a \leq X \leq b$ auch $a \leq EX \leq b$.
- Der Erwartungswert ist **linear**: $\exists\ EX$, dann $\exists\ E(a + bX)$ für $a, b \in \mathbb{R}$. Es gilt: $E(a + bX) = a + b \cdot EX$
- Existiert EX und EY, und ist $EX + EY$ definiert, dann existiert auch $E(X + Y)$, und es gilt $E(X + Y) = EX + EY$.
- Falls alle EX_i existieren und $EX_i \neq \pm\infty$, dann gilt: $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} EX_i$
- Sind X, Y **stoch. unabh.**, existieren EX und EY und endlich, dann exisiert EXY und es gilt: $EXY = EX \cdot EY$. Der sogenannte Multiplikationssatz.

7.4 k-tes Moment/Streuung/Varianz

Seien X, Y ZV´en und $c \in \mathbb{R}$, $k \in \mathbb{N}$.

- $E[(X - c)^k]$ k-tes Moment von X um c
- $E[(X - EX)^2] =: VarX$ **Varianz/Streuung** von X
- $E[(X - EX) \cdot (Y - EY)]$ **Kovarianz** von X und Y

7.4.1 Eigenschaften der Varianz: $E[(X - EX)^2]$

Sei $VarX < \infty$, es gilt:

- $Var(a \cdot X + b) = a^2 \cdot VarX \quad a, b \in \mathbb{R}$ - Eine Verschiebung hat keinen Einfluss auf die Varianz, ein Faktor verändert die Varianz quadratisch.
- $VarX = EX^2 - (EX)^2$
- $VarX = 0 \Leftrightarrow P(X \neq EX) = 0$
- $VarX = min\ E[(X - a)^2]\ \forall a \in \mathbb{R}$

7.4.2 Eigenschaften der Kovarianz: $E[(X - EX) \cdot (Y - EY)]$

- $Cov(X, Y) = E(X \cdot Y) - EX \cdot EY$
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, X) = VarX$
- $Cov(a \cdot X + b \cdot Y) = a \cdot Cov(X, Y)$
- $Cov^2(X, Y) \leq VarX \cdot VarY$
- X, Y unabh. $\Rightarrow Cov(X, Y) = 0$

7.4.3 Additionssatz

Seien X, Y ZV´en mit $VarX < \infty$ und $VarY < \infty$, dann gilt:

$$Var(X + Y) = Var(X) + Var(Y) + 2 \cdot Cov(X, Y)$$

Korollar $X_1, \cdots, X_n$ ZV´en mit $EX_i^2 < \infty$ für $i = 1, \cdots, n$, dann gilt:

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} Cov(X_i, Y_j)$$

7.4.4 Korrelationskoeffizient

$$Kor(X,Y) = \frac{Cov(X,Y)}{\sqrt{VarX \cdot VarY}}$$

7.5 Ungleichung von Lyapunoff

X sei reelle ZV, $E|X|^r < \infty$ für ein $r \in (0,\infty)$. Dann existiert auch $E(|X|^s)$ für alle $0 \leq s \leq r$ und es gilt:

$$(E(|X|^r))^{\frac{1}{r}} \geq (E(|X|)^s)^{\frac{1}{s}}$$

8 Das schwache Gesetz großer Zahlen

8.1 Motivation

Das Gesetz der großen Zahlen besagt, dass sich die relative Häufigkeit eines Zufallsergebnisses immer weiter an die theoretische Wahrscheinlichkeit für dieses Ergebnis (Erwartungswert) annähert, je häufiger das Zufallsexperiment durchgeführt wird.

Beispiel Die Wahrscheinlichkeit, dass eine Münze beim Werfen Kopf zeigt, beträgt $\frac{1}{2}$. Je häufiger die Münze geworfen wird, desto näher wird der Anteil der Würfe, bei denen Kopf erscheint, beim theoretischen Wert $\frac{1}{2}$ liegen. Trotzdem kann der absolute Abstand zwischen dem theoretischen und dem tatsächlich beobachteten Ergebnis immer weiter anwachsen.

#	K_{erw}	K_{beo}	Rhi_{erw}	Rhi_{beo}	ABSOLUTER ABSTAND
100	50	48	0,5	0,48	2
1000	500	491	0,5	0,491	9
10000	5000	4970	0,5	0,497	30

Man kann also aus dem Gesetz der großen Zahlen nicht die Schlussfolgerung ziehen, wenn ein Ereignis bislang nicht so häufig eintrat wie erwartet, muss es diesen Rückstand ausgleichen und folglich in Zukunft häufiger vorkommen.

8.2 stochastisch konvergierend

gegen 0: Eine Folge $(X_n)_n$ von ZV´en über $(\Omega, \mathcal{A}, P)$ heißt stochastisch konvergierend gegen 0, falls

$$\begin{aligned} \lim_{n\to\infty} P(|X_n| > \epsilon) &= 0 \;\; \forall \epsilon > 0 \\ \textit{Bezeichnung}: & \quad \text{P-lim}_{n\to\infty} X_n = 0 \end{aligned}$$

gegen c: Eine Folge $(X_n)_n$ von ZV´en heißt stochastisch konvergierend gegen $c \in \mathbb{R}$, falls

$$\begin{aligned} \text{P-lim}_{n\to\infty}(X_n - c) &= 0 \;\; \text{bzw.} \\ \text{P-lim}_{n\to\infty}(X_n - X) &= 0 \\ \textit{Bezeichnung}: & \\ \text{P-lim}_{n\to\infty} X_n &= c \;\; \text{bzw.} \\ \text{P-lim}_{n\to\infty} X_n &= X \end{aligned}$$

8.3 Markov´sche Ungleichung

Sei ZV X gegeben, $g : \mathbb{R}^+ \to \mathbb{R}^+$ monoton wachsenden, dann gilt:

$$P(|X| > \epsilon) \le P(|X| \ge \epsilon) \le \frac{1}{g(\epsilon)} \cdot E(g(|X|)) \;\; \forall \epsilon > 0 \text{ mit } g(\epsilon) > 0$$

8.4 Eine Version des schwachen Gesetz großer Zahlen

Das schwache Gesetz der großen Zahlen besagt, dass für eine unendliche Folge von Zufallsvariablen $X1, X2, X3, \cdots$, die alle den selben Erwartungswert μ und die selbe endliche Varianz σ^2 haben sowie unkorreliert sind (d.h. die Korrelation zwischen zwei beliebigen ZV beträgt Null), dann konvergiert die repräsentative Stichprobe

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

wahrscheinlich gegen μ.

Genauer: Für jede positive Zahl ϵ (beliebig klein) gilt:

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| < \epsilon) = 1$$

8.4.1 Beispiel

Wir haben eine evtl. gefälscht Münze(Qualitätskontrolle $\rightarrow$ Gut/Schlecht-Prüfung). Wie oft muss man werfen damit die Wahrscheinlichkeit für "Zahl" mit einer Sicherheitswahrscheinlichkeit von $\geq 0,95$ auf $0,01$ genau bestimmt werden kann.

$$\begin{aligned} X_i &= \mathcal{I}_{\text{Zahl im i-ten Versuch}} \\ EX_i &= P \quad \text{,wobei P unbekannt} \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad & (X_i)_{i\in\ \mathbb{N}} \text{ ZV´en iid} \cong b(1,p) \\ \Rightarrow \quad & P(|\frac{1}{n} \cdot \sum_{i=1}^{n} X_i - p| \geq 0,01) \\ \leq \quad & \frac{p\cdot(1-p)}{n\cdot 0,01^2} \\ \leq \quad & \frac{\frac{1}{4}}{n\cdot 0,01^2} \leq 0,05 \\ \Rightarrow \quad & n \geq \frac{1}{4\cdot 0,01^2 \cdot 0,05} \\ = \quad & \frac{20\cdot 1000}{4} = 5000 \end{aligned}$$

9 Borelmengen und Maße

Sei $f : \mathbb{R} \to \mathbb{R}$ eine Abbildung mit $f(x) \geq 0 \quad x \in \mathbb{R}$.
Sei $\mathcal{I} = (a, b) \subset \mathbb{R}$, $-\infty \leq a < b \leq \infty$, wobei $f(x) = 0 \quad \forall x \in \mathcal{I}^c$
f stetig auf $\mathcal{I}$ und

$$\int_{-\infty}^{\infty} f(x)dx = \int_a^b f(x)dx = 1 \quad \text{"Riemann-Integral"}$$

Definiere: **Verteilungsfunktion** $F : \mathbb{R} \to [0, 1]$ durch $F(y) = \int_{-\infty}^{y} f(x)dx$.

Eigenschaften: $F(y) \in [0, 1] \; \forall y \in \mathbb{R}$, $F \nearrow$, stetig.

$$\begin{aligned}
\lim_{y\to-\infty} F(y) &= \lim_{y\to-\infty} \int_{-\infty}^{y} f(x)dx \\
&= \int_{-\infty}^{-\infty} f(x)dx = 0 \\
\lim_{y\to\infty} F(y) &= \lim_{y\to\infty} \int_{-\infty}^{y} f(x)dx \\
&= \int_{-\infty}^{\infty} f(x)dx = 1
\end{aligned}$$

Bemerkung Nach dem Eindeutigkeits-/Existenzsatz existiert genau ein Wahrscheinlichkeitsmaß P mit der Verteilungsfunktion F, gegeben durch $F(y) = \int_{-\infty}^{y} f(x)dx$.

Dichtefunktion Die Funktion f heißt **(Riemann-) Dichte (-Funktion)** von P:

- $P(\{x\}) = 0 \; \forall x \in \mathbb{R}$
- $P(A) = 0$ für jede abzählbare Menge A
- $P((a, b]) = P((a, b)) = P([a, b)) = P([a, b]) = \int_a^b f(x)dx \quad \forall a < b$

$$f(x) = \begin{cases} 0 & , x \leq a \\ F'(x) & , a < x < b \\ 0 & , x \geq b \end{cases}$$

10 Wahrscheinlichkeitsmaße mit Riemann-Dichten über $\mathbb{R}$

10.1 Motivation

Es gibt Situationen in denen man Messwerte mit der Menge $\mathbb{R}$ der reellen Zahlen modelliert ($\rightarrow$ stetig), statt mit einer endlichen oder abzählbar unendlichen Menge($\rightarrow$ diskret). Mögliche Gründe hierfür:

- Feinheit der Diskretisierung um Vorfeld nicht festlegen
- mathematische Strukturen nutzen: Stetigkeit oder Ableitungen von Funktionen

Als Ereignissystem wählen wir die Borel-σ-Algebra.

Die wesentliche Idee zur Angabe der Wahrscheinlichkeit eines Ereignisses, z.b. eines Intervalls $(a, b]$, besteht darin, analog zum Aufsummieren von Zähldichten im diskreten Fall $[P(A) = \sum_{\omega \in A} f(\omega)]$, über eine gewisse Dichtefunktion f zu integrieren. Man erhält die Darstellung $P(A) = \int_a^b f(x)dx$.

10.2 Riemann-Dichte ("stetige Zähldichte")

Eine Riemann-integrierbare Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R} \text{ mit } f(x) \geq 0 \ (x \in \mathbb{R}) \text{ und } \int_{-\infty}^{\infty} f(x)\ dx = 1$$

heißt **Riemann-Dichte** über $\mathbb{R}$.

Jede Riemann-Dichte definiert ein eindeutige **Wahrscheinlichkeitsmaß** P über $(\mathbb{R}, \mathcal{B})$ mit der Eigenschaft

$$P((a, b]) = P([a, b]) = \int_a^b f(x)\ dx \ (a \leq b)\ ,\ P(a) = 0$$

10.3 Verteilungsfunktion $F(x)$

Motivation In den bisherigen Darstellungen von Wahrscheinlichkeitsmaßen durch Riemann-Dichten besteht ein gewisser Nachteil darin, dass für jedes Ereignis ein Integral ausgewertet werden muss. Man kann dies vermeiden, indem man das unbestimmte Integral, die sogenannte Stammfunktion, heranzieht.

Definition Ist P ein beliebiges W´maß über $(\mathbb{R}, \mathcal{B})$, dann heißt die Abbildung $F : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$F(x) := P((-\infty, x]) \quad \text{mit } x \in \mathbb{R}$$

die **Verteilungsfunktion von P**.

Folgerung Für ein Wahrscheinlichkeitsmaß P über $\mathbb{R}$ mit Riemann-Dichte f gilt nach dieser Definition:

$$F(x) = \int_{-\infty}^{x} f(t)\ dt$$

$$P((a,b]) = \int_{a}^{b} f(t)\ dt = F(b) - F(a)$$

Wir betrachten nun die wichtigsten Riemann-Dichte:

10.4 Rechteck-Verteilung

Der einfachste Fall: Wahrscheinlichkeit über einem Stück der reellen Geraden (eines Intervalls) gleichmäßig verteilt. Die Dichtefunktion ist dann auf diesem Stück konstant.

Dichte:

$$\begin{aligned} f(x) &= \frac{1}{b-a} \cdot \mathrm{I}_{[a,b]}(x) \\ &= \begin{cases} \frac{1}{b-a} & , x \in [a,b] \\ 0 & , \text{sonst} \end{cases} \end{aligned}$$

Verteilungsfunktion:

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

Bezeichnung: $\mathcal{R}(a,b)$ bzw. $\mathcal{R}_{[a,b]}$.

10.4.1 Beispiel

Zwei Freunde A und B verabreden sich zu treffen. A kommt zu einem zufälligen Zeitpunkt zwischen 21h und 22h und wartet 15 Minuten. B kommt zwischen 20h30 und 22h30 und wartet 30 Minuten. Beiden Ankunftszeiten können als iid angesehen werden. Mit welcher Wahrscheinlichkeit begegnen sich beide?

Lösung: X sei die Ankunftszeit von A: $X \sim \mathcal{R}[21;22]$
Y sei die Ankunftszeit von B: $Y \sim \mathcal{R}[20,5;22,5]$

$$\begin{aligned} f^{(X,Y)}(x,y) = f^X(x) \cdot f^Y(y) &= \frac{1}{22-21} \cdot \mathrm{I}_{[21;22]}(x) \cdot \frac{1}{22,5-20,5} \cdot \mathrm{I}_{[20,5;22,5]}(y) \\ &= \frac{1}{2} \cdot \mathrm{I}_{[21;22]}(x) \cdot \mathrm{I}_{[20,5;22,5]}(y) \end{aligned}$$

Die beiden treffen sich, wenn

$$(Y \leq X + \frac{1}{4}) \wedge (X \leq Y + \frac{1}{2}) \quad \Leftrightarrow \quad X - \frac{1}{2} \leq Y \leq X + \frac{1}{4}$$

$$\begin{aligned}
P(X - \frac{1}{2} \leq Y \leq X + \frac{1}{4}) &= \int_{21}^{22} \int_{x-\frac{1}{2}}^{x+\frac{1}{4}} \frac{1}{2} \cdot \mathrm{I}_{[21;22]}(x) \cdot \mathrm{I}_{[20,5;22,5]}(y)\ dx\ dy \\
&= \int_{21}^{22} \int_{x-\frac{1}{2}}^{x+\frac{1}{4}} \frac{1}{2}\ dy\ dx \\
&= \frac{1}{2} \int_{21}^{22} [y]_{x-\frac{1}{2}}^{x+\frac{1}{4}}\ dx \\
&= \frac{1}{2} \int_{21}^{22} [(x + \frac{1}{4}) - (x - \frac{1}{2})]\ dx \\
&= \frac{1}{2} \int_{21}^{22} \frac{3}{4}\ dx \\
&= \frac{3}{8} \cdot \int_{21}^{22}\ dx \\
&= \frac{3}{8} \cdot [x]_{21}^{22} \\
&= \frac{3}{8}
\end{aligned}$$

10.5 Exponentialverteilung

Sie verhält sich ähnlich der (diskreten) geometrischen Verteilung. Anstatt einer geometrisch abnehmenden Zähldichte, besitzt diese eine exponentiell abnehmende Riemann-Dichte. Wie schnell diese abnimmt ergibt sich aus dem Parameter λ. Exponentialverteilungen sind typisch für Wartezeiten bei gleichförmigen Ankunftsprozessen.

Dichte:

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

wobei $x \in \mathbb{R}$ und $\lambda > 0$.

Verteilungsfunktion:

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

Bezeichnung: $\mathrm{Exp}(\lambda)$; Exponentialverteilung mit Parameter λ.

10.5.1 Beispiel

Die Exponentialverteilung ist eine typische Lebensdauerverteilung. So ist beispielsweise die Lebensdauer von elektronischen Geräten häufig annähernd exponentialverteilt.

Beispiel: In einer Elektronikfirma werden Funkwecker produziert. Im Rahmen der Qualitätssicherung wird anhand von Reklamationen die Funktionsdauer der Wecker untersucht. Es ist definiert $X \cong$ Zeitdauer der Funktionsfähigkeit eines Funkweckers [Tage].

Bemerkung: Der Erwartungswerte der Exponentialverteilung ist $\frac{1}{\lambda}$ und die Varianz beträgt $\frac{1}{\lambda^2}$.

X ist exponentialverteilt mit dem Parameter $\lambda = 0,005$. In diesem Zusammenhang wird λ als Ausfallrate bezeichnet; es fallen also durchschnittlich pro Tag 5 Promille der Wecker aus. Entsprechend ist die Intensität $\frac{1}{\lambda}$ die durchschnittliche Zeitdauer, bis ein Wecker ausfällt. Es fällt also im Mittel alle 200 Tage ein Wecker aus.

$$
\begin{aligned}
X &\sim exp(\lambda) \ \text{ mit } \lambda = 0,005 \\
f(x) &= \lambda \cdot e^{-\lambda x} \cdot \mathrm{I}_{(0,\infty)}(x) \\
F^X(x) &= P(X \leq x) = (1 - e^{-\lambda x}) \cdot \mathrm{I}_{(0,\infty)}(x)
\end{aligned}
$$

Die Wahrscheinlichtkeit, dass ein Wecker **höchstens 20 Tage** hält, ist

$$P(X < 20) = 1 - e^{-\lambda x} = 1 - e^{-0,005 \cdot 20} = 0,0952$$

d.h. nach 20 Tagen sind ca. 10 % aller Wecker ausgefallen.

Entsprechend der Anteil der Wecker, die **mindestens 180 Tage** aushalten,

$$P(X > 180) = 1 - P(X \leq 180) = 1 - (1 - e^{-0,005 \cdot 180}) = 1 - 0,5934 = 0,4066$$

also halten ca. 40 % der Wecker länger als 180 Tage.

Der Anteil der Wecker mit einer Lebensdauer **zwischen 20 und 180 Tagen** beträgt dann,

$$
\begin{aligned}
P(20 \leq X \leq 180) &= P(X \leq 180) - P(X < 20) \\
&= P(X \leq 180) - P(X \leq 20) \\
&= F^X(180) - F^X(20) \\
&= (1 - e^{-0,005 \cdot 180}) - (1 - e^{-0,005 \cdot 20}) \\
&= 0,5934 - 0,0952 = 0,492
\end{aligned}
$$

Median der Lebensdauer: Gesucht ist ein x mit $P(X \leq x) = p = 0,5$.

$$\begin{aligned}
p = P(X \leq x) &= F^X(x) = 1 - e^{-\lambda x} \\
p &= 1 - e^{-\lambda x} \\
1 - p &= e^{-\lambda x} \\
\ln(1-p) &= -\lambda x \\
x &= -\frac{\ln(1-p)}{\lambda} \\
x &= -\frac{\ln(1-p)}{0,005} \\
x &= -200 \cdot \ln(1-p) = -200 \cdot \ln(0,5) \\
x &\approx 138,629
\end{aligned}$$

Schlußfolgerung: Man sieht, dass sehr viele Wecker schon nach kurzer Zeit defekt sind, ein Phänomen, das man bei elektronischen Geräten häufig beobachtet. Wenn sie kaputt gehen, dann meistens sofort. Ab dann halten sie quasi ewig.

10.6 Weibull-Verteilung

Dichte:

$$f(x) = \begin{cases} \alpha \cdot \beta \cdot x^{\beta-1} \cdot e^{-\alpha x^\beta} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

wobei $\alpha, \beta > 0$ und für $\alpha = \lambda$ und $\beta = 1$ gilt $Wei(\alpha, \beta) = Exp(\lambda)$.

Verteilungsfunktion:

$$F(x) = \begin{cases} 1 - e^{-\alpha x^\beta} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

Bezeichnung: $Wei(\alpha, \beta)$

10.6.1 Beispiel

Ein System S bestehe aus $n \in \mathbb{N}$ Komponenten $K_1, \cdots, K_n$ in Reihenschaltung. Ein Ausfall einer jeden Komponenten führt zum Ausfall von S. Die Lebensdauern $X_1, \cdots, X_n$ der einzelnen Komponenten seien stoch. unabh., und es gelte

$$X_i \sim Wei(\alpha_i, \beta) \quad , i \in \{1, \cdots, n\}$$

Zu zeigen: Die Lebensdauer des System S ist $Wei(\alpha, \beta)$-verteilt mit $\alpha = \alpha_1 + \cdots + \alpha_n$.

Beweis: U bezeichne die Lebensdauer des Systemes S, es gilt

$$U = \min\{X_1, \cdots, X_n\}$$

Es gilt für alle $x > 0$:

$$\begin{aligned}
P(U > x) &= P(X_1 > x, \cdots, X_n > x) \\
&= \prod_{i=1}^{n} P(X_i > x) \quad (X_i \text{ stoch. unabh.}) \\
&= \prod_{i=1}^{n} (1 - P(X_i \leq x)) \\
&= \prod_{i=1}^{n} (1 - F^{X_i}(x)) \\
&= \prod_{i=1}^{n} (1 - (1 - e^{-\alpha_i x^\beta})) \\
&= \prod_{i=1}^{n} e^{-\alpha_i x^\beta} \\
&= e^{-\alpha_1 x^\beta} \cdots e^{-\alpha_n x^\beta} \\
&= e^{-(\alpha_1 + \cdots + \alpha_n) x^\beta}
\end{aligned}$$

$$\begin{aligned}
F^U(x) &= P(U \leq x) = 1 - P(U > x) \\
&= 1 - (e^{-(\alpha_1 + \cdots + \alpha_n) x^\beta}) \\
&= 1 - e^{-\alpha x^\beta} \quad \forall x > 0
\end{aligned}$$

Es gilt für alle $x \leq 0 : F^U(x) = 0$

$$\Rightarrow U \sim Wei(\alpha, \beta)$$

10.7 Gamma-Verteilung

Dichte:

$$f_{\alpha,\beta}(x) = \frac{\alpha^\beta}{\Gamma(\beta)} \cdot x^{\beta-1} \cdot e^{-\alpha \cdot x} \cdot \mathcal{I}_{(0,\infty)}(x), \; x \in \mathrm{R}$$

wobei $\Gamma(\beta) = \int_0^\infty t^{\beta-1} \cdot e^{-t} \, dt$

Verteilungsfunktion: Da kein geschlossenes Integral existiert, können wir auch keine Verteilungsfunktion F angeben.

Bezeichnung: $\Gamma(\alpha, \beta)$, $\alpha, \beta > 0$

10.7.1 Beispiel: Faltung

X_1, X_2 stoch. unabh. mit $X_i \sim \Gamma(\alpha, \beta_i),\ i = 1, 2;\ \alpha, \beta_1, \beta_2 > 0$, dann gilt

$$X_1 + X_2 \sim \Gamma(\alpha, \beta_1 + \beta_2)$$

Hinweis: Es gilt

$$\int_0^z (z-x)^{a-1} \cdot x^{b-1}\, dx = \frac{\Gamma(a) \cdot \Gamma(b)}{\Gamma(a+b)} \cdot z^{a+b-1}, \quad a, b, > 0$$

Beweis: Nach 10.12, Seite 50, gilt:

$$\begin{aligned} f^{X_1+X_2}(y) &= \int_{-\infty}^{\infty} f^{X_1}(t) \cdot f^{X_2}(y-t)\, dt \; y \in \mathbb{R} \\ &= \int_{-\infty}^{\infty} \frac{\alpha^{\beta_1}}{\Gamma(\beta_1)} \cdot t^{\beta_1 - 1} \cdot e^{-\alpha t} \cdot \mathrm{I}_{(0,\infty)} \cdot \frac{\alpha^{\beta_2}}{\Gamma(\beta_2)} \cdot (y-t)^{\beta_2 - 1} \cdot e^{-\alpha(y-t)} \cdot \mathrm{I}_{(0,\infty)}\, dt \end{aligned}$$

Für $y \leq 0$ gilt sicherlich:

$$f^{X_1+X_2} = 0$$

Für $y > 0$ gilt:

$$\begin{aligned} f^{X_1+X_2} &= \frac{\alpha^{\beta_1+\beta_2}}{\Gamma(\beta_1) \cdot \Gamma(\beta_2)} \cdot \int_0^y e^{-\alpha t} \cdot e^{-\alpha(y-t)} \cdot t^{\beta_1 - 1} \cdot (y-t)^{\beta_2 - 1}\, dt \\ &= \frac{\alpha^{\beta_1+\beta_2}}{\Gamma(\beta_1) \cdot \Gamma(\beta_2)} \cdot e^{-\alpha y} \cdot \int_0^y (y-t)^{\beta_2 - 1} \cdot t^{\beta_1 - 1}\, dt \\ &= \frac{\alpha^{\beta_1+\beta_2}}{\Gamma(\beta_1) \cdot \Gamma(\beta_2)} \cdot e^{-\alpha y} \cdot [\frac{\Gamma(\beta_1) \cdot \Gamma(\beta_2)}{\Gamma(\beta_1 + \beta_2)} \cdot y^{\beta_1+\beta_2-1}] \quad \text{(siehe Hinweis oben)} \\ &= \frac{\alpha^{\beta_1+\beta_2}}{\Gamma(\beta_1 + \beta_2)} \cdot e^{-\alpha y} \cdot y^{\beta_1+\beta_2-1} \end{aligned}$$

Also gilt insgesamt:

$$\begin{aligned} f^{X_1+X_2}(y) &= \frac{\alpha^{\beta_1+\beta_2}}{\Gamma(\beta_1 + \beta_2)} \cdot e^{-\alpha y} \cdot y^{\beta_1+\beta_2-1} \cdot \mathrm{I}_{(0,\infty)}(y) \\ &\Rightarrow X_1 + X_2 \sim \Gamma(\alpha, \beta_1 + \beta_2) \end{aligned}$$

10.8 Normalverteilung

Die am weitesten verbreitet R-Dichte. Sie kommt dort vor (zumindest nährungsweise), wo sich viele kleine voneinander unabhängige Einflüsse (z.b. Messfehler) überlagern. Jede Normalverteilung ist durch zwei Paramater charakterisiert: Mittelwert μ und Streuung σ^2.

Dichte:

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad ; x \in \mathbb{R},\ \mu \in \mathbb{R},\ \sigma > 0$$

Bezeichnung: $\mathcal{N}(\mu, \sigma^2)$

Standard-Normalverteilung $\cong \mathcal{N}(0,1)$

10.9 Riemann-Dichten über $\mathbb{R}^n$

Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt Riemann-Dichte über $\mathbb{R}^n$, falls

- $f(x) \geq 0 \ \forall x \in \mathbb{R}^n$
- f ist Riemann-Integrierbar über $\mathbb{R}^n$ mit $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n)\ dx_1 \cdots dx_n = 1$

10.10 Verteilungsfunktion einer Riemann-Dichte über $\mathbb{R}^n$

Ist f eine Dichte über $\mathbb{R}^n$, so definiert

$$F(x_1, \cdots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(y_1, \cdots, y_n)\ dy_1 \cdots dy_n \quad , (y_1, \cdots, y_n) \in \mathbb{R}^n$$

eine stetige Verteilungsfunktion über $\mathbb{R}^n$.

10.11 Erwartungswert einer Riemann-Dichte

Sei f eine Riemann-Dichte von X. Dann heißt

$$EX = \int_{-\infty}^{\infty} x \cdot f(x)\ dx$$

der Erwartungswert von X.

Bemerkung: Es gelten alle bereits bewiesenen Eigenschaften und Ungleichungen, ersetzte hierfür überall die Summen durch entsprechende Integrale (ohne Beweis).

10.11.1 Beispiel: Rechteck-Verteilung

Sei ZV $X \cong \mathcal{R}(a,b)$ eine Rechteck-Verteilung.

$$\begin{aligned} EX = \int_{-\infty}^{\infty} x \cdot f(x)\ dx &= \int_a^b x \cdot \frac{1}{b-a}\ dx \\ &= \frac{1}{b-a} \cdot [\frac{x^2}{2}]_a^b \\ &= \frac{1}{b-a} \cdot [\frac{b^2}{2} - \frac{a^2}{2}] \\ &= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} \\ &= \frac{1}{b-a} \cdot \frac{(b-a) \cdot (b+a)}{2} \\ &= \frac{b+a}{2} \end{aligned}$$

10.12 Faltung von R-Dichten

Sind X und Y reellwertig mit gemeinsame R-Dichte $f^{(X,Y)}(x,y)$, dann hat $X+Y$ entsprechend die R-Dichte

$$f^{X+Y}(z) = \int_{-\infty}^{\infty} f^{(X,Y)}(x, z-x)\ dx \quad , z \in \mathbb{R}$$

Klar: wie oben beschrieben ersetzen wir in den bekannter Formel für Faltungen im diskreten Fall (siehe 6.8, Seite 34), die Summen durch entsprechende Integrale.

Sind X und Y stoch. unabh; so gilt:

$$f^{(X,Y)}(z_1, z_2) = f^X(z_1) \cdot f^Y(z_2)$$

und damit folgt

$$f^{X+Y}(z) = \int_{-\infty}^{\infty} f^X(x) \cdot f^Y(z-x)\ dx \quad , z \in \mathbb{R}$$

10.13 großes Beispiel

Die Dichtefunktion des Zufallvektors (X,Y) im $\mathbb{R}^2$ sei definiert durch

$$f^{(X,Y)}(x,y) = \begin{cases} c \cdot (x + 2xy) & , x, y \in [0,2] \\ 0 & , \text{sonst} \end{cases}$$

i) Bestimmen Sie die Konstante c: Da $f^{(X,Y)}$ R-Dichte muss gelten

$$\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^{(X,Y)}(x,y)\ dy\ dx = 1 \quad &\Leftrightarrow \quad \int_0^2\int_0^2 c \cdot (x+2xy)\ dy\ dx \\ &= \quad c \cdot \int_0^2\int_0^2 (x+2xy)\ dy\ dx \\ &= \quad c \cdot \int_0^2 x \cdot \int_0^2 (1+2y)\ dy\ dx \\ &= \quad c \cdot \int_0^2 x \cdot [y + \frac{2y^2}{2}]_0^2 \text{ dx} \\ &= \quad c \cdot 6 \cdot [\frac{x^2}{2}]_0^2 \\ c \quad &= \quad \frac{1}{12} \end{aligned}$$

ii) Ermitteln Sie die marginalen Dichten der ZV X und Y: Es gilt allgemein für Randdichten im stetigen Fall:

$$f^{X_i}(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_{i-1}, x, x_{i+1}, \cdots, x_n)\ dx_n \cdots dx_{i+1}\ dx_{i-1} \cdots dx_1 \quad \text{(i-te Stelle also fest)}$$

$$\begin{aligned}
f^{(X,Y)}(x) &= \int_0^2 f^{(x,y)}(x,y)\ dy \\
&= \int_0^2 c \cdot (x + 2xy)\ dy \\
&= c \cdot \int_0^2 x \cdot (1 + 2y)\ dy \\
&= c \cdot x \int_0^2 (1 + 2y)\ dy \\
&= c \cdot x \cdot [y + \frac{2y^2}{2}]_0^2 \\
&= \frac{1}{12} \cdot x \cdot 6 \\
&= \begin{cases} \frac{x}{2} & \text{für } x \in [0,2] \\ 0 & \text{für } x \notin [0,2] \end{cases}
\end{aligned}$$

$$\begin{aligned}
f^{(X,Y)}(y) &= \int_0^2 f^{(X,Y)}(x,y)\ dx \\
&= c \cdot \int_0^2 (x + 2xy)\ dx \\
&= c \cdot (1 + 2y) \int_0^2 x\ dx \\
&= \frac{1 + 2y}{12} \cdot [\frac{x^2}{2}]_0^2 \\
&= \begin{cases} \frac{1+2y}{6} & \text{für } y \in [0,2] \\ 0 & \text{für } y \notin [0,2] \end{cases}
\end{aligned}$$

iii) Sind X und Y stoch. unabh:

$$\begin{aligned}
X, Y \text{ stoch. unabh.} &\Leftrightarrow F^{(X,Y)}(x,y) = F^X(x) \cdot F^Y(y) \\
&\Leftrightarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^{(X,Y)}(x,y)\ dy\ dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^X(x) \cdot f^Y(y)\ dy\ dx
\end{aligned}$$

Beweis:

$$\begin{aligned}
\int_0^2 \int_0^2 f^{(X,Y)}(x,y)\ dy\ dx &= 1 \text{ (klar, siehe Definition der R-Dichte)} \\
&\Leftrightarrow \int_0^2 \int_0^2 f^X(x) \cdot f^Y(y)\ dy\ dx \\
&\quad \text{(Verwende die Ergebnisse aus Teil ii)} \\
&= \int_0^2 \int_0^2 \frac{x}{2} \cdot \frac{(1+2y)}{6}\ dy\ dx \\
&= \frac{1}{12} \cdot \int_0^2 x \cdot [y + \frac{2y^2}{2}]_0^2\ dx \\
&= \frac{1}{12} \cdot \int_0^2 x \cdot [(2+4) - (0+0)]\ dx \\
&= \frac{1}{2} \cdot [\frac{x^2}{2}]_0^2 \\
&= \frac{1}{2} \cdot 2 \\
&= 1 \checkmark
\end{aligned}$$

iv) Berechnen Sie EY und Var(Y):

$$\begin{aligned}
EY &= \int_0^2 y \cdot f^Y(y)\ dy \\
&= \int_0^2 y \cdot \frac{(1+2y)}{6}\ dy \\
&= \frac{1}{6} \cdot \int_0^2 (y + 2y^2)\ dy \\
&= \frac{1}{6} \cdot [\frac{y^2}{2} + \frac{2y^3}{3}]_0^2 \\
&= \frac{1}{6} \cdot [\frac{4}{2} + \frac{2 \cdot 8}{3}] \\
&= \frac{1}{3} + \frac{16}{18} \\
&= \frac{11}{9}
\end{aligned}$$

$$
\begin{aligned}
EY^2 &= \int_0^2 y^2 \cdot f^Y(y)\; dy \\
&= \int_0^2 y^2 \cdot \frac{(1+2y)}{6}\; dy \\
&= \frac{1}{6} \cdot \int_0^2 (y^2 + 2y^3)\; dy \\
&= \frac{1}{6} \cdot [\frac{y^3}{3} + \frac{2y^4}{4}]_0^2 \\
&= \frac{1}{6} \cdot [(\frac{8}{3} + \frac{16}{2}) - (0+0)] \\
&= \frac{8}{18} + \frac{8}{6} \\
&= \frac{16}{9}
\end{aligned}
$$

$$
\begin{aligned}
Var\ Y &= EY^2 - E^2Y \\
&= \frac{16}{9} - (\frac{11}{9})^2 \\
&= \frac{23}{81}
\end{aligned}
$$

11 Grundlagen der Simulation

Ziel: deterministische und reproduzierbare $(x_i)_{i\in\mathbb{N}} \subset [0,1]$, welche stetig gleichverteilt sind.

Gegeben: Grundraum $\Omega_m = \{\frac{0}{m}, \frac{1}{m}, \cdots, \frac{m-1}{m}\}$ mit diskreter Gleichverteilung P_m $(= \frac{1}{m}\ \forall)$.

Beschreibung der n-maligen unabh. zufälligen Auswahl einer Zahl aus Ω_m durch neuen Grundraum

$$\Omega_m^n := \{\omega = (\omega_1, \cdots, \omega_n) | \omega_j \in \Omega_m \text{ für } j \in \{1, \cdots, n\}\}$$

und der Gleichverteilung

$$P_m^n(\{\omega\}) := \prod_{j=1}^{n} P_m(\{\omega_j\}) = (\frac{1}{m})^n$$

11.1 Lineare Kongruenzgenerator

Seien
$m \in \mathbb{N}_0$ (Modul)
$a \in \mathbb{N}_0$ (Faktor)
$b \in \mathbb{N}_0$ (Inkrement)
$z_0 \in \mathbb{N}_0$, $z_0 \leq m-1$ (Anfangsglied)

Kongruenzschema:

$$z_{j+1} \equiv a \cdot z_j + b\ (mod\ m)\ ,\ j \in \mathbb{N}_0$$

Wir möchten aber Zahlen im Intervall $[0,1]$, daher eine Normierung nötig:

$$x_j := \frac{z_j}{m}\ ,\ j \in \mathbb{N}_0$$

liefert uns die gewünschte Folge $(x_n)_{n\in\mathbb{N}_0} \subset [0,1]$

11.1.1 Beispiel

Sei $m = 100,\ a = 18,\ b = 11,\ z_0 = 40$:

$$
\begin{aligned}
z_0 &= 40 \\
x_0 &= \frac{40}{100} = 0,4 \\
z_1 &= 18 \cdot 40 + 11\ (mod\ 100) \equiv 731 \equiv 31 \\
x_1 &= \frac{31}{100} = 0,31 \\
z_2 &= 18 \cdot 31 + 11\ (mod\ 100) \equiv 569 \equiv 69 \\
x_2 &= \frac{69}{100} = 0,69 \\
z_3 &= 18 \cdot 69 + 11\ (mod\ 100) \equiv 1253 \equiv 53 \\
x_3 &= \frac{53}{100} = 0,53 \\
\\
z_4 &= 18 \cdot 53 + 11\ (mod\ 100) \equiv 965 \equiv 65 \\
x_4 &= \frac{65}{100} = 0,65 \\
z_5 &= 18 \cdot 65 + 11\ (mod\ 100) \equiv 1181 \equiv 81 \\
x_5 &= \frac{81}{100} = 0,81 \\
z_6 &= 18 \cdot 81 + 11\ (mod\ 100) \equiv 1469 \equiv 69 \\
x_6 &= \frac{69}{100} = 0,69 \\
z_7 &= z_2 \\
x_7 &= x_2
\end{aligned}
$$

Es entstehen nur 6 verschiedene Zahlen, der Generator hat also eine Periode der Länge 4. Es muss also durch Wahl der Parameter die maximal mögliche Periodenlänge bestimmt werden.

12 Überblick Verteilungen

12.1 diskrete Verteilungen

Verteilung	Dichte	EX	$Var\ X$	Beispiel
Einpunkt'	$e_{\{a\}}(k)$	a	0	
Laplace'	$\frac{1}{N}$	$\frac{N+1}{2}$	$\frac{N^2-1}{12}$	Würfeln
Bernoulli'	$b(p) = p^k \cdot (1-p)^{1-k}$	p	$p \cdot (1-p)$	Münzwurf
Binomial'	$b(s;k,p) = \binom{k}{s} \cdot p^s \cdot (1-p)^{k-s}$	$n \cdot p$	$n \cdot p \cdot (1-p)$	16 Datenpakete: $b(0;16,\frac{1}{12}) + b(1;16,\frac{1}{12})$
Hypergeom'	$h(s;k,n,S) = \frac{\binom{S}{s} \cdot \binom{n-S}{k-s}}{\binom{n}{k}}$	$k \cdot \frac{S}{n}$	$k \cdot \frac{S}{n} \cdot \frac{n-S}{n} \cdot \frac{n-k}{n-1}$	1 As an jeden Spieler $P(A_3 \mid A_1 \cap A_2) = h(1;10,12,2)$
Poisson'	$p(s) = \frac{\lambda^s}{s!} \cdot e^{-\lambda}$	λ	λ	Kaufhaus: 10 sec $\approx$ 1 Besucher $p(9) = \frac{6^9}{9!} \cdot e^{-6}$
Geometrische'	$geom(p) = p \cdot (1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	IP-based Netzwerk $X = Z+1,\ Z \sim geom(p)$ $f^X(k) = p \cdot (1-p)^{k-1}$

12.2 stetige Verteilungen

VERTEILUNGEN	R-DICHTE:	VERTEILUNGSFUNKTION:
$\mathcal{R}_{[a,b]}$	$f^X(x) = \frac{1}{b-a} \cdot \mathrm{I}_{[a,b]}(x)$	$F^X(x) = \begin{cases} 0 & , x < a \\ \frac{x-a}{b-a} & , a \leq x \leq b \\ 1 & , x > b \end{cases}$
$Exp(\lambda)$	$f^X(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$	$F^X(x) = \begin{cases} 1 - e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$
$Wei(\alpha, \beta)$	$f^X(x) = \begin{cases} \alpha\beta \cdot x^{\beta-1} \cdot e^{-\alpha x^\beta} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$	$F^X(x) = \begin{cases} 1 - e^{-\alpha x^\beta} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$
$\Gamma(\alpha, \beta)$	$f^X(x) = \frac{\alpha^\beta}{\Gamma(\beta)} \cdot x^{\beta-1} \cdot e^{-\alpha \cdot x} \cdot \mathcal{I}_{(0,\infty)}(x)$ $\Gamma(\beta) = \int_0^\infty t^{\beta-1} \cdot e^{-t}\, dt$	
$\mathcal{N}(\mu, \sigma^2)$	$f^X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	
$\mathcal{N}(0,1)$	$f^X(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-(x)^2}{2}}$	

A Appendix

A.1 partielle Integration

Läßt sich die Funktion $f(x)$ darstellen als $f(x) = u(x) \cdot v'(x)$, dann gilt

$$\int u(x) \cdot v'(x)\ dx = u(x) \cdot v(x) - \int u'(x) \cdot v(x)\ dx$$

Beispiel:

$$\begin{aligned}
\int \ln x\ dx &= \int 1 \cdot \ln x\ dx \\
&= x \cdot \ln x - \int x \cdot \frac{1}{x}\ dx \\
&= x \cdot \ln x - \int\ dx \\
&= x \cdot \ln x - x + C
\end{aligned}$$

$$\begin{aligned}
\int x \cdot e^x\ dx &= x \cdot e^x - \int 1 \cdot e^x\ dx \\
&= x \cdot e^x - e^x + C
\end{aligned}$$

A.2 Substitution

Durch Substitution $x = \varphi(t)$ der unabhängigen Variable einer Funktion $y = f(x)$, ergibt sich das unbestimmte Integral

$$\int f(x)\ dx = \int f(\varphi(t)) \cdot \varphi'(t)\ dt$$

Beispiel:

$$\int \frac{1}{(2+3x)^2}\ dx \qquad t = 2 + 3x$$

$$x = \varphi(t) = \frac{t-2}{3}$$

$$\varphi'(t) = \frac{dx}{dt} = \frac{(t-2)' \cdot 3 - (t-2) \cdot 3'}{3^2} = \frac{1}{3}$$

$$dx = \frac{dt}{3}$$

$$\begin{aligned}
\int \frac{dx}{(2+3x)^2} &= \int \frac{1}{t^2} \cdot \frac{dt}{3} \\
&= \frac{1}{2-1} \cdot \frac{1}{t^{2-1}} \frac{1}{3} + C \\
&= \frac{1}{3t} + C = \frac{1}{3 \cdot (2+3x)} + C
\end{aligned}$$

$$\begin{aligned}
\int (1+x)^n \, dx \qquad & t = 1 + x \\
& x = \varphi(t) = t - 1 \\
& \varphi'(t) = \frac{dx}{dt} = 1 \\
& dx = dt \\
\int (1+x)^n \, dx &= \int t^n \, dt \\
&= \frac{t^{n+1}}{n+1} + C \\
&= \frac{(1+x)^{n+1}}{n+1} + C
\end{aligned}$$

A.3 geometrische Reihe

$$\sum_{k=1}^{n} p^{k-1} = \frac{1-p^n}{1-p}$$

Beispiel:

$$\sum_{k=1}^{5} 2^{k-1} = \frac{1-2^5}{1-2}$$

$$\sum_{k=1}^{10} 3^k = 3 \cdot \frac{1-3^{10}}{1-3}$$

A.4 Aufgabe 14

Gegeben seien $N+1$ Urnen, wobei die i-te Urne $N-i$ weiße und i schwarze Kugeln enthält. Eine Urne wird zufällig ausgewählt und man entnimmt m Kugeln. Dabei wird jede Kugel mit derselben Wahrscheinlichkeit gezogen.

Zu zeigen: Wahrscheinlichkeit im m+1-ten Zug aus der Urne eine schwarze Kugel ziehen, unter der Bedingung, auch in den ersten m Ziehungen eine schwarze Kugel vorzufinden, beträgt:

$$p_N^* = \frac{1}{N} \cdot \frac{\sum_{i=1}^N i^{m+1}}{\sum_{i=1}^N i^m}$$

Hinweis: Betrachte Ereignisse $H_i \cong$ Wahl der i-ten Urne und verwenden Sie die Formel der totalen Wahrscheinlichkeit (siehe 4.3, Seite 20)

Beweis: WR $(\Omega, \mathcal{A}, P)$, wobei

$$\begin{aligned}
\Omega &= \{(u, \omega_1, \cdots, \omega_{m+1}) | u \in \{0, \cdots, N\}, \omega_i \in \{0,1\}, i \in \{1, \cdots, m+1\}\} \\
\omega_i &= \begin{cases} 0 & \text{, weiß} \\ 1 & \text{, schwarz} \end{cases} \\
P(\{(u, \omega_1, \cdots, \omega_{m+1})\}) &= \frac{1}{N+1} \cdot \prod_{j=1}^{m+1} (\frac{u}{N})^{\omega_j} \cdot (1-\frac{u}{N})^{1-\omega_j} \quad \forall \omega \in \Omega \\
A &= \{(u, \omega_1, \cdots, \omega_{m+1}) \in \Omega | \omega_1 = \cdots = \omega_m = 1\} \\
&\cong \text{m-mal schwarz bei den ersten m Ziehungen} \\
B &= \{(u, \omega_1, \cdots, \omega_{m+1}) \in \Omega | \omega_{m+1} = 1\} \\
&\cong \text{schwarz im (m+1)-ten Zug} \\
H_i &= \{(u, \omega_1, \cdots, \omega_{m+1}) \in \Omega | u = i\} \\
&\cong \text{Wahl der Urne i}
\end{aligned}$$

$$\begin{aligned}
\textbf{\textit{Gesucht}}: \quad & P(B|A) \\
&= \frac{P(B \cap A)}{P(A)} \\
P(H_i) &= \frac{1}{N+1} \\
P(A \cap B | H_i) &= (\frac{i}{N})^{m+1}
\end{aligned}$$

Mit der Formel der totalen Wahrscheinlichkeit folgt:

$$\begin{aligned}
P(A \cap B) &= \sum_{i=0}^{N} P(A \cap B|H_i) \cdot P(H_i) \\
&= \sum_{i=0}^{N} (\frac{i}{N})^{m+1} \cdot \frac{1}{N+1} \\
&= (\frac{1}{N})^{m+1} \cdot \sum_{i=0}^{N} i^{m+1} \cdot \frac{1}{N+1}
\end{aligned}$$

$$\begin{aligned}
P(A) &= \sum_{i=0}^{N} P(A|H_i) \cdot P(H_i) \\
&= \sum_{i=0}^{N} \frac{i}{N}^{m} \cdot \frac{1}{N+1} \\
&= (\frac{1}{N})^{m} \cdot \sum_{i=0}^{N} i^{m} \cdot \frac{1}{N+1}
\end{aligned}$$

$$\begin{aligned}
P(B|A) &= \frac{P(B \cap A)}{P(A)} \\
&= \frac{(\frac{1}{N})^{m+1} \cdot \sum_{i=0}^{N} i^{m+1} \cdot \frac{1}{N+1}}{(\frac{1}{N})^{m} \cdot \sum_{i=0}^{N} i^{m} \cdot \frac{1}{N+1}} \\
&= \frac{1}{N} \cdot \frac{\sum_{i=1}^{N} i^{m+1}}{\sum_{i=1}^{N} i^{m}} \quad \text{(für i=0 ist die } \sum \text{ gleich null)} \\
&= p_N^*
\end{aligned}$$

A.5 Aufgabe 15

Sei $(\Omega, \mathcal{A}, P)$ ein WR und $A, B, B_1, \cdots, B_n \in \mathcal{A}$ mit

- $\bigcup_{i=1}^{n} B_i = \Omega$, wobei $B_1, \cdots, B_n$ paarweise disjunkt
- $P(A \cap B_i) > 0$ für alle $1 \leq i \leq n$

Zu zeigen:

$$P(B|A) = \sum_{i=1}^{n} P(B_i|A) \cdot P(B|A \cap B_i)$$

Beweis: Weiterhin gilt: $A \cap B = \sum_{i=1}^{n}(A \cap B \cap B_i)$, da ja die $B_i s$ eine disjkunkte Zerlegung von Ω.

$$\begin{aligned}
P(B|A) &= \frac{P(B \cap A)}{P(A)} = \frac{1}{P(A)} \cdot P(\sum_{i=1}^{n}(A \cap B \cap B_i)) \\
&= \frac{1}{P(A)} \cdot \sum_{i=1}^{n} P(A \cap B \cap B_i) \\
&= \frac{1}{P(A)} \cdot \sum_{i=1}^{n} P(B|A \cap B_i) \cdot P(A \cap B_i) \\
&= \sum_{i=1}^{n} P(B|A \cap B_i) \cdot \frac{P(A \cap B_i)}{P(A)} \\
&= \sum_{i=1}^{n} P(B|A \cap B_i) \cdot P(A|B_i)
\end{aligned}$$

A.6 Der Zonk

Ohne Beschränkung der Allgemeinheit (wegen der Symmetrie) sei folgende Situation gegeben:

- Kandidat wählt Tür 1
- Moderator öffnet Tür 3, d.h. Auto befindet sich hinter Tür 1 oder Tür 2.
- Frage: soll Kandidat bei Wahl der Tür 1 bleiben oder nun Tür 2 wählen.

Hierzu, ohne speziellen Wahrscheinlichkeitsraum:

$$\begin{aligned}
A_i &\cong \text{Auto befindet sich hinter Tür i} \\
&\ \text{Voraussetzung: } P(A_i) = \frac{1}{3}\ \forall i \in \{1,2,3\} \\
K_i &\cong \text{Kandidat wählt Tür i} \\
&\ \text{Voraussetzung: } P(K_i) = \frac{1}{3}\ \forall i \in \{1,2,3\} \\
Q_i &\cong \text{Moderator öffnet Tür i} \\
&\ \text{Nicht unabhängig von } A_i, K_i
\end{aligned}$$

Rechnung

$$\begin{aligned}
P(A_1|K_1 \cap Q_3) &= \frac{P(A_1 \cap K_1 \cap Q_3)}{P(K_1 \cap Q_3)} \\
&= \frac{P(Q_3|A_1 \cap K_1)}{P(K_1 \cap Q_3)} \cdot P(A_1 \cap K_1) \\
& \quad \text{(Benutze Formel der totalen Wahrscheinlichkeit)} \\
P(K_1 \cap Q_3) &= P(K_1 \cap Q_3 \cap A_1) + P(K_1 \cap Q_3 \cap A_2) + P(K_1 \cap Q_3 \cap A_3) \\
&= P(Q_3|A_1 \cap K_1) \cdot P(A_1 \cap K_1) + P(Q_3|A_2 \cap K_1) \cdot P(A_2 \cap K_1) \\
&= \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{1}{2} \\
\Rightarrow P(A_1|K_1 \cap Q_3) &= \frac{P(A_1 \cap K_1 \cap Q_3)}{P(K_1 \cap Q_3)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3} \\
\Rightarrow P(A_2|K_1 \cap Q_3) &= \frac{P(A_2 \cap K_1 \cap Q_3)}{P(K_1 \cap Q_3)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}
\end{aligned}$$

D.h. Änderung der Entscheidung verdoppelt die Gewinnwahrscheinlichkeit.

A.7 Aufgabe 23

Notebooks haben $0, \cdots, 5$ Fehler im Display. Es werden immer zufällig zwei Displays aus der Lieferung rausgegriffen, dass mit den meisten Fehler gelangt in Modell GR, dass andere in Modell GLX.

- Modellieren Sie die Anzahl Fehler mit einer Zufallsvariable X und geben Sie EX an.
- Modellieren Sie die Anzahl Fehler im Modell GR und GLX mit Hilfe der ZV X_1 und X_2 und berechnen die jeweiligen Erwartungswerte.

i) Es sei $(\Omega, \mathcal{A}, P)$ der WR und

$$X : \Omega \to \{0, \cdots, 5\} \text{ mit } P^{X(i)} = \frac{1}{6} \text{ , } i \in \{0, \cdots, 5\}$$

Dann gilt:

$$EX = \sum_{i=0}^{5} i \cdot P(X = i) = \frac{1}{6} \cdot \sum_{i=0}^{5} i = 2,5$$

ii)

$$(x, \tilde{x}) : \Omega \to \{0, \cdots 5\}^2 \text{ mit } P^X(i,j) = \frac{1}{36} \text{ ,} i,j \in \{0, \cdots, 5\}$$

Dann ist $X_2 = min(x, \tilde{x})$ (Modell GLX) und $X_1 = max(x, \tilde{x})$ (Modell GR). Für die Verteilung P^{X_1} und P^{X_2} gelten dann folgende Werte:

I	0	1	2	3	4	5
$P(X_2 = i)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$
$P(X_1 = i)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Wir erhalten:

$$\begin{aligned}
EX_2 &= (0 \cdot \frac{11}{36}) + (1 \cdot \frac{9}{36}) + (2 \cdot \frac{7}{36}) + (3 \cdot \frac{5}{36}) + (4 \cdot \frac{3}{36}) + (5 \cdot \frac{1}{36}) = \frac{55}{36} \cong 1,528 \\
EX_1 &= (0 \cdot \frac{1}{36}) + (1 \cdot \frac{3}{36}) + (2 \cdot \frac{5}{36}) + (3 \cdot \frac{7}{36}) + (4 \cdot \frac{9}{36}) + (5 \cdot \frac{11}{36}) = \frac{125}{36} \cong 3,472
\end{aligned}$$

A.8 Aufgabe 6 (K)

Firma besitzt Internet-Anbindung über drei Provider.
Zuverlässigkeit: Provider A: 95%, Provider B 92 % und Provider C 92%.
Anteile: Provider A: 45%, Provider B 25% und Provider C 25%.

a) Mit welcher Wahrscheinlichkeit erreicht ein Paket sein Ziel?

b) Falls ein Paket sein Ziel nicht erreicht, mit welcher Wahrscheinlichkeit kam es von Provider B.

c) Man vermutet falsche Angabe bzgl. der Zuverlässigkeit von Provider A. 40% der empfangenen Pakete stammen von Provider A. Welcher Anteil der über A gesendeten Pakete erreicht somit sein Ziel, unter der Annahme das die Angaben von B und C korrekt sind?

Lösung:

$$\begin{aligned}
Z &\sim \text{Paket erreicht sein Ziel} \\
A, B, C &\sim \text{Paket über Provider A,B,C gesendet} \\
P(Z|A) &= 0,95 \\
P(Z|B) &= 0,90 \\
P(Z|C) &= 0,92 \\
P(A) &= 0,45 \\
P(B) &= 0,30 \\
P(C) &= 0,25
\end{aligned}$$

a)
$$\begin{aligned} P(Z) &= P(Z|A) \cdot P(A) + P(Z|B) \cdot P(B) + P(Z|C) \cdot P(C) \\ &= 0,95 \cdot 0,45 + 0,9 \cdot 0,3 + 0,92 \cdot 0,25 = 0,9275 \end{aligned}$$

b)
$$\begin{aligned} P(B|Z^C) &= P(Z^C|B) \cdot \frac{P(B)}{P(Z^C)} \\ &= 0,10 \cdot \frac{0,3}{1 - 0,9275} = 0,414 \end{aligned}$$

c)
$$\begin{aligned} P(A|Z) &= 0,4 \\ P(Z|A) &= x \\ P(Z|A) &= P(A|Z) \cdot \frac{P(Z)}{P(A)} \\ &= 0,4 \cdot \frac{x \cdot 0,45 + 0,9 \cdot 0,3 + 0,92 \cdot 0,25}{0,45} \\ &= 0,4 \cdot \frac{x \cdot 0,45 + 0,5}{0,45} \\ x &= 0,4 \cdot (\frac{x \cdot 0,45}{0,45} + \frac{0,5}{0,45}) \\ x &= 0,4 \cdot x + \frac{0,4 \cdot 0,5}{0,45} \\ x - 0,4 \cdot x &= \frac{0,2}{0,45} \\ x \cdot (1 - 0,4) &= \frac{0,2}{0,45} \\ x &= \frac{0,2}{0,45 \cdot 0,6} \\ x &\approx 0,74 \end{aligned}$$

A.9 Aufgabe 2 (K)

Ein Mann besitzt fünf Münzen: zwei haben auf beiden Seiten Kopf, eine hat auf beiden Seiten Zahl, und die beiden letzten sind "fair".

a) Mann zieht zufällig eine Münze und wirft sie. Mit welcher Wahrscheinlichkeit liegt "Kopf" oben?

b) Mann öffnet beiden Augen und sieht das "Kopf" oben liegt. Mit welcher Wahrscheinlichkeit liegt auch unten "Kopf"?

c) Mann schließt die Augen wieder und wirft die selbe Münze nochmals. Wie groß ist die Wahrscheinlichkeit dass nun auf der unteren Seite "Kopf" liegt?

d) Er öffnet die Augen und sieht das "Kopf" oben liegt. Wie groß ist die Wahrscheinlichkeit für "Kopf" unten jetzt?

Vorbereitung: $K_1^o, K_1^u, K_2^o, K_2^u$ besagt das im ersten bzw. zweiten Wurf "Kopf" oben bzw. unten liegt. $M_i \sim$ Münze i gewählt für $i \in \{1, \cdots, 5\}$.

zu a)

$$P(K_1^o) = \sum_{i=1}^{5} P(K_1^o|M_i) \cdot P(M_i) = 1 \cdot \frac{1}{5} + 1 \cdot \frac{1}{5} + 0 \cdot \frac{1}{5} + \frac{1}{2} \cdot \frac{1}{5} + \frac{1}{2} \cdot \frac{1}{5} = \frac{3}{5}$$

zu b)

$$P(K_1^u|K_1^o) = \frac{P(K_1^u \cap K_1^o)}{P(K_1^o)} = \frac{\sum_{i=1}^{5} P(K_1^u \cap K_1^o|M_i) \cdot P(M_i)}{\frac{3}{5}} = \frac{5}{3} \cdot (1 \cdot \frac{2}{5}) = \frac{2}{3}$$

zu c)

$$\begin{aligned}
P(K_2^u|K_1^o) &= \frac{P(K_2^u \cap K_1^o)}{P(K_1^o)} \\
&= \frac{5}{3} \cdot [\sum_{i=1}^{5} P(K_2^u \cap K_1^0|M_i) \cdot P(M_i)] \\
&= \frac{5}{3} \cdot [1 \cdot \frac{2}{5} + 0 \cdot \frac{1}{5} + \frac{1}{4} \cdot \frac{1}{5} + \frac{1}{4} \cdot \frac{1}{5}] \\
&= \frac{5}{3} \cdot \frac{5}{10} = \frac{5}{6}
\end{aligned}$$

zu d)

$$\begin{aligned}
P(K_2^u|K_1^o \cap K_2^o) &= \frac{P(K_2^u \cap K_2^o \cap K_1^o)}{P(K_1^o \cap K_2^o)} \\
&= \frac{P(K_2^u \cap K_2^o)}{\sum_{i=1}^{5} P(K_1^o \cap K_2^o|M_i) \cdot P(M_i)} \\
&= \frac{P(M_1 \cap M_2)}{[1 \cdot \frac{2}{5} + 0 \cdot \frac{1}{5} + \frac{1}{4} \cdot \frac{2}{5}]} \\
&= \frac{2}{5} \cdot \frac{10}{5} = \frac{4}{5}
\end{aligned}$$