Stochastik Wintersemester 2009
Leibniz Universität Hannover
Vorlesungsmitschrift

Dozent: Prof. Dr. L. Baringhaus

Mitschrift von
Ronald Becher

28. Januar 2010

Zusammenfassung

Diese Mitschrift wird erstellt im Zuge der Vorlesung Stochastik A im Wintersemester 2009 an der Leibniz Universität Hannover. Obwohl mit großer Sorgfalt geschrieben, werden sich sicherlich Fehler einschleichen. Diese bitte ich zu melden, damit sie korrigiert werden können. Ihr könnt euch dazu an jeden der genannten Autoren (mit Ausnahme des Dozenten) wenden.

Dieses Skript wird primär über Github verteilt und „gepflegt“. Dort kann man es auch forken, verbessern und dann (idealerweise) ein „Pull Request“ losschicken. Siehe auch Github Guides. Auch wenn ihr gute Grafiken zur Verdeutlichung beitragen könnt, dürft ihr diese gerne schicken oder selbst einfügen (am besten als LaTeX-geeignete Datei (Tikz, SVG, PNG, ...)).

Hinweis: Es wird o.B.d.A. davon abgeraten, die Vorlesung zu versäumen, nur weil eine Mitschrift angefertigt wird! Selbiges gilt für den Übungsbetrieb! Stochastik besteht ihr nur, wenn ihr auch zu der Veranstaltung geht!

Inhaltsverzeichnis

Teil I
Einführung / Organisatorisches

Die Klausur wird am **01.02.2010 von 9:15 - 10:45** stattfinden, es wird wöchentliche Haus- und Stundenübungen geben.
Die abgegebenen Hausübungen werden korrigiert und in den Übungsgruppen zurückgegeben werden, es gibt allerdings keine Bonuspunkteregelung. Übungsgruppen wird es folgende geben:

- Mo, 12:15 - 13:00 Uhr (F309)
- Mo, 14:15 - 15:00 Uhr (G123) [max. 20 Teilnehmer]
- Di, 10:15 - 11:00 Uhr (F428)
- Di, 12:15 - 13:00 (F442)
- Mi, 12:15 - 13:00 (B305)

1 Literatur

Der Dozent empfiehlt das Buch "Stochastik für Einsteiger" von Herrn N. Henze (Vieweg Verlag)

Leider richtet sich das oben genannte Buch primär an Mathematiker. Im selben Verlag ist ein Buch für Informatiker erschienen.

1.1 Referenzen

Im Folgenden werden ich, wenn möglich, beide Quellen referenzieren.

Teil II
Grundlegende Wahrscheinlichkeitstheorie

„Was ist das richtige Modell, um ein Experiment zu beschreiben?“ ist die wichtigste Frage, die die mathematische Statistik zu beantworten hat, und diese Modelle an die Wirklichkeit anzupassen, ist die hauptsächliche Aufgabe der Stochastik. Der letzte Teil wird vor allem in der Vorlesung **Stochastik B** behandelt.

2 Wahrscheinlichkeitsräume

Zufallsexperimente werden beschrieben durch **Wahrscheinlichkeitsäume**. Ein Wahrscheinlichkeitsraum ist ein Tripel $(\Omega, \mathfrak{A}, P)$ und setzt sich wie folgt zusammen.

2.1 Die Ergebnismenge Ω

Die Menge der möglichen Ergebnisse heißt Ω ([Hüb09, S. 12]).
Im **Würfelwurf** sei beispielsweise $\Omega := \{1, 2, 3, 4, 5, 6\}$

2.2 Der Ereignisraum $\mathfrak{A}$

die Menge $\mathfrak{A} \subset \mathfrak{P}(\Omega)$[1] $:= \{A; A \subset \Omega\}$ als Menge der interessierenden Ergebnisse ([Hüb09, S. 14]).
$\mathfrak{A}$ ist eine so genannte σ-Algebra auf Ω, d.h.

1. $$\Omega \in \mathfrak{A} \tag{1}$$
   (Die Grundmenge Ω ist in $\mathfrak{A}$ enthalten)

2. $$A \in \mathfrak{A} \iff A^c \in \mathfrak{A} \tag{2}$$
   (Wenn $\mathfrak{A}$ eine Teilmenge $A \in \Omega$ als mögliches Ergebnis enthält, dann auch deren Komplement $A^c = \Omega \setminus A$)

3. $$A_1, A_2, \dots \in \mathfrak{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{A} \tag{3}$$
   (Wenn für jede natürliche Zahl n die Menge $A_n \in \mathfrak{A}$ ist, so ist auch die abzählbare Vereinigung aller $A_n \in \mathfrak{A}$)

[1]Potenzmenge

Im **Würfelwurf** seien beispielsweise gültige Ereignisse:

- $A := \{2, 4, 6\}$ das Ereignis „Es fällt eine gerade Zahl“, und
- $A := \{5, 6\}$ das Ereignis „Es fällt eine Zahl größer gleich 5“.

2.2.1 Rechnen mit Ereignissen

Wenn nun $A, B \in \mathfrak{A}$, dann gilt ([Hüb09, S. 15])

- $$A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \cdots \in \mathfrak{A}$$
  (Die Vereinigung von A, B ist auch ein interessierendes Ereignis und steht für “das Ereignis A oder B treten ein”)
- $$A \cap B = A \cap B \cap \Omega \cap \Omega \cap \cdots \in \mathfrak{A}$$
  (Die Schnittmenge von A, B ist auch ein interessierendes Ereignis und steht für “das Ereignis A **und** B treten ein”)
- $$A \Delta B = (A \cap B^c) \cup (B \cap A^c)$$
  (Die “symmetrische Differenz” beschreibt das Ereignis “Entweder A oder B”)
- $A \subseteq B$ heißt: “Das Ereignis A hat das Ereignis B zur Folge”.
  (Kurz erklärt: Wenn A eintritt, folgt $\omega \in A \ \wedge \ A \subseteq B \Rightarrow \omega \in B$)

2.3 Das Wahrscheinlichkeitsmaß P

durch ein Wahrscheinlichkeitsmaß P (lat: „**P**robabilitas“=Glaubwürdigkeit) mit $P : \mathfrak{A} \to [0, 1]$, also eine Mengenfunktion auf $\mathfrak{A}$ mit den Eigenschaften

1. $$P(\Omega) = 1 \tag{4}$$
   (Die so genannte “Normiertheit”).
2. Für jede Folge von **paarweise disjunkten** $A_1, A_2, \dots \in \mathfrak{A}$ gilt:
   $$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \tag{5}$$
   σ-Additivität von P

   Für $A \in \mathfrak{A}$ ist $P(A)$ die Wahrscheinlichkeitsfunktion (für das Eintreten) des Ereignisses A.

2.3.1 Eigenschaften von Wahrscheinlichkeits-Maßen

- Aus $P(\emptyset) = P(\emptyset \cup \emptyset \cup \dots) = \sum_{n=1}^{\infty} P(\emptyset)$ folgt $P(\emptyset) = 0$.

- Für $A, B \in \mathfrak{A}$ mit $A \subseteq B$ ist

$$P(B) = P(A \cup B \cap A^c) = P(A \cup B \cap A^c \cup \emptyset \cup \dots) = P(A) + P(B \cap A^c) \geq P(A) \tag{6}$$

  Also gelten folgende Eigenschaften / Gesetzmäßigkeiten:

  –
$$P(A) \leq P(B) \tag{7}$$
  (Isotonie)

  –
$$P(B \cap A^c) = P(B) - P(A) \tag{8}$$
  (Subtraktivität)

2.4 Ergebnisse von Zufallsexperimenten

$\omega \in \Omega$ nennen wir das Ergebnis des Zufallsexperimentes. Das Ereignis $A \in \mathfrak{A}$ tritt genau dann ein, wenn $\omega \in A$.

$$\omega \in \bigcup_{n=1}^{\infty} A_n \iff \exists n \in \mathbb{N} \wedge \omega \in A_n \tag{9}$$

Das durch $\bigcup_{n=1}^{\infty} A_n$ beschriebene Ereignis tritt genau dann ein, wenn mindestens eines der Ereignisse $A_n,\ n \in \mathbb{N}$ eintritt.
Auf die Frage, ob auch $\emptyset$ eine gültige Menge ist, sei gesagt, dass

$$\emptyset = \Omega^c \in \mathfrak{A}$$

das „unmögliche Ereignis“ genannt wird (Die Antwort ist also “ja”).

Seien $A_1, A_2, \dots \in \mathfrak{A}$. Dann

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c \in \mathfrak{A} \tag{10}$$

Anmerkung: $\bigcap_{n=1}^{\infty} A_n$ tritt genau dann ein, wenn alle $A_n, n \in \mathbb{N}$ eintreten.

2.5 Beispiel

$\emptyset \neq \Omega$ sei endlich, $|\Omega|$ Anzahl der Elemente von Ω und $\mathfrak{A} = \mathfrak{P}(\Omega)$.
Sei P definiert durch

$$P(A) = \frac{|A|}{|\Omega|},\ A \subset \Omega \tag{11}$$

P heißt „diskretes Laplace'sches W-Maß„ auf Ω.
Es ist

$$P(\{\omega\}) = \frac{1}{|\Omega|},\ \omega \in \Omega \tag{12}$$

also folgert man: „Alle möglichen Ergebnisse sind **gleich wahrscheinlich**."

3 Grundformeln der Kombinatorik

$n, r \in \mathbb{N}$, M_n ist eine n-elementige Menge, o.B.d.A. $M_n = \{1, \ldots, n\}$.

- $$P_n^r = \{(x_1, \ldots, x_r);\ x_i \in M_n, 1 \leq i \leq r\} \tag{13}$$
  das sind die r-Permutationen aus M_n mit Wiederholung.

- $$P_n^{(r)} = \{(x_1, \ldots, x_r);\ x_i \in M_n, 1 \leq i \leq r, \text{paarweise verschieden}\}\ (r \leq n) \tag{14}$$
  das sind die r-Permutationen aus M_n ohne Wiederholung.

- $$K_n^{(r)} = \{(x_1, \ldots, x_r);\ x_i \in M_n, 1 \leq i \leq r,\ x_1 < x_2 < \cdots < x_n\}\ (r \leq n) \tag{15}$$
  das sind die r-Kombinationen aus M_n ohne Wiederholung.

- $$K_n^r = \{(x_1, \ldots, x_r);\ x_i \in M_n, 1 \leq i \leq r,\ x_1 \leq x_2 \leq \cdots \leq x_n\}\ (r \leq n) \tag{16}$$
  das sind die r-Kombinationen aus M_n mit Wiederholung.

3.1 Zur Erklärung

1. Tupel in Permutationen sind nicht angeordnet, hier ist die Reihenfolge wichtig!
2. Tupel in Kombinationen sind o.B.d.A. angeordnet, wir können sie also "vergleichen".
3. Bei Varianten **mit Wiederholung**, dürfen also zwei Elemente "nebeneinander" auch gleich sein, bei Varianten **ohne Wiederholung** ist das nicht der Fall, hier müssen die Elemente also echt unterschiedlich sein.
4. Bei einer r-Kombination haben wir eine r-elementige Teilmenge (die also kleiner oder gleich groß der "großen" Menge M_n ist).
5. "mit Wiederholung" nennt man auch "mit Zurücklegen"

3.2 Anschauliche Beispiele

- Bei Kombinationen interessiert uns ausschließlich, welche Elemente wir aus M_n bekommen, ein gutes Beispiel für eine Kombination

  ohne Wiederholung ist also das übliche Lotto-Spiel "6 aus 49". Hier ist es uns egal, ob die Reihenfolge $(1, 2, 3, 4, 5, 6)$ ist oder $(6, 5, 4, 3, 2, 1)$.

mit Wiederholung ist die Frage, ob wir mindestens zwei Mal in fünf Versuchen eine Zahl ≥ 5 würfeln

- Bei Permutationen interessiert uns neben der Auswahl auch die Anordnung (man sagt: die Elemente sind “unterscheidbar”). Eine Permutation

ohne Wiederholung ist z.B. ein Kinobesuch. Ob Tim zwischen Tina und Claudia sitzt, ist für ihn nicht gleichwertig mit der Variante, zwischen Klaus und Jochen zu sitzen.
Anders ausgedrückt gilt in Permutationen

$$(Claudia, Tim, Tina, Klaus, Jochen) \neq (Claudia, Tina, Klaus, Tim, Jochen)$$

mit Wiederholung ist die Frage, ob ich beim Roulette einen aufsteigenden Run habe, also erst eine 0 fällt, dann eine 1, dann eine 2, usw. Das Standardbeispiel hierbei ist eine Urne, wo ich Kugeln oder Zettel heraushole und wieder hereinlege.
Eine Permutation mit Wiederholung ist also eine “Auswahl mit Wiederholungen unter Beachtung der Reihenfolge”.

3.3 Rechenbeispiele

Es gilt

- $$|P_n^r| = n^r \tag{17}$$
  (Erklärung: Man greift r Mal in die Urne, schaut sich einen der n Zettel an und legt ihn wieder herein, es gibt also danach genauso viele Möglichkeiten zu ziehen (also n))

- $$|P_n^{(r)}| = \begin{cases} r \leq n & n \cdot (n-1) \cdot \ldots \cdot (n-r+1) \\ r = n & r! \end{cases} \tag{18}$$
  (Erklärung: Lege ich meine Zettel nicht wieder in die Urne zurück, habe ich beim nächsten Versuch natürlich eine Möglichkeit weniger, welche Zettel ich ziehen könnte)

- $$r! \cdot |K_n^{(r)}| = |P_n^{(r)}| \tag{19}$$
$$\Updownarrow$$
$$|K_n^{(r)}| = \frac{n(n-1)\dots(n-r+1)}{r!} = \binom{n}{r} \tag{20}$$
$$\Downarrow$$
$$|\{A; A \subset M_n, |A| = r\}| = \binom{n}{r} \tag{21}$$
$$\Downarrow$$
$$|\mathfrak{P}(M_n)| = \sum_{r=0}^{r} \binom{n}{r} 1^r 1^{n-r} = (1+1)^n = 2^n \tag{22}$$

  (21) Gilt auch noch für $r = 0$.
  Erklärungen:
  (19) und äquivalent (20) kann man sich so vorstellen, dass man sich seine r-elementige Teilmenge (ohne Zurücklegen) zusammensuchen kann, aber die verschiedenen Möglichkeiten, sie zu permutieren ("die Kinobesucher umzusetzen"), interessieren den Betrachter hier nicht, also müssen wir die aus der Rechnung rausnehmen.
  (21) ist die Anzahl aller r-elementigen Teilmengen von M_n hoch r, und wegen (20) ebenfalls $\binom{n}{r}$.
  (22) ist die Länge der Potenzmenge, also aller möglichen Teilmengen von Elementen aus M_n

- $K_n^r \to (x_1, \dots, x_r) \underbrace{\longrightarrow}_{\textit{ist bijektiv}} (x_1, x_1 + 1, \dots, x_r + r - 1) \in K_{n+r-1}^{(r)}$
  $\Rightarrow |K_n^r| = |K_{n+r-1}^{(r)}| = \binom{n+r-1}{r}$
  Identifiziere $(x_1, \dots, x_r) \in K_n^r$ mit dem Besetzungszahlvektor $(k_1, \dots, k_n)$, wobei $k_i = |\{j \in \{1, \dots, r\}; x_j = i\}|,\ 1 \le i \le n$. Dann ist $k_1 + \ldots + k_n = r$.
  Dann ist $|\{(k_1, \dots, k_n) \in \mathbb{N}_0^n; k_1 + \cdots + k_n = r\}| = \binom{n+r-1}{r}$
  Folgerung: $|\{(k_1, \dots, k_n) \in \mathbb{N}^n; k_1 + \cdots + k_n = r\}| = \binom{n+(r-n)-1}{r-n} = \binom{r-1}{r-n} = \binom{r-1}{n-1}, r \ge n$

Aus der Menge der Permutationen betrachten wir eine spezifische Teilmenge.
Sei $k_1, \dots, k_n \in \mathbb{N}_0$ mit $k_1 + \dots + k_n = r$. Einen so genannter "Besetzungszahlvektor" drückt man beispielsweise so aus:

$$\begin{aligned}&\{(x_1,\dots,x_n) \in \mathbb{P}_n^r;\ |\{j; j \in \{1,\dots,r\}, x_j = i\}| = k_i, 1 \le i \le n\}| \\ =&\{(x_1,\dots,x_n) \in \mathbb{P}_n^r;\ \text{Genau } k_i \text{ der } r \text{ Komponenten } (x_1,\dots,x_r) \text{ sind gleich } i, 1 \le i \le n\}\end{aligned} \tag{23}$$

Die Länge oder Mächtigkeit dieser Menge ist wie folgt:

$$\begin{aligned}|23| &= \binom{r}{k_1}\binom{r-k_1}{k_2}\binom{r-k_1-k_2}{k_3}\cdots\binom{r-k1-\dots-k_{n-1}}{k_n} \\ &= \frac{r!}{k_1!\cancel{(r-k_1)!}} \cdot \frac{\cancel{(r-k_1)!}}{k_2!\cancel{(r-k_1-k_2)!}} \cdots \frac{\cancel{(r-k_1-\dots-k_n)!}}{k_n!} \\ &= \frac{r!}{k_1!\cdots k_n!}\end{aligned}$$

3.4 Multinomialer Lehrsatz

Es folgt aus dem (bekannten) binomischen Lersatz durch Verallgemeinerung der multinomiale Lehrsatz.

$$n^r = \sum_{\substack{(k_1,\dots,k_n)\in\mathbb{N}_0^n \\ k_1+\dots+k_n=r}} \frac{r!}{k_1!\cdots k_n!} \tag{24}$$

$$\text{Anzahl der Summanden aus } 24 = \binom{n+r-1}{r} \tag{25}$$

3.5 Beispiel

O.B.d.A. sind r Personen im Raum, $n = 365$ seien die Anzahle der Tage im Jahr und $r \le n$.
Gesucht: Wahrscheinlichkeit, dass mindestens 2 der r Personen am gleichen Tag Geburtstag haben.

- $\Omega = \{(x_1,\dots,x_r); x_i = \{1,\dots,n\}\}$
- $\mathfrak{A} = \mathfrak{P}(\Omega)$
- $P(A) = \frac{|A|}{|\Omega|}, A \subseteq \Omega$
- Alle möglichen r-Tupel von Geburtstagen sind gleich wahrscheinlich, also
- $A = \{(x_1,\dots,x_r) \in \Omega; i,j \in \{1,\dots,r\}, i \ne j, x_i = x_j\}$

- $P(A) = 1 - P(A^c)$
  In vielen Fällen ist es ein valider "Trick", statt der Ergebnismenge das Gegenergebnis zu berechnen
- $A^c = \{(x_1, \ldots, x_r) \in \Omega; x_1, \ldots, x_r \text{ paarweise verschieden}\}$
- $|A^c| = n(n-1)\cdots(n-r+1)$
- $P(A) = 1 - \frac{n(n-1)\cdots(n-r+1)}{n^r}$

$$P(A) = \begin{cases} > 1/2 & \text{wenn } r > 23, \\ 0,706 & \text{wenn } r = 30, \\ 0,994 & \text{wenn } r = 60. \end{cases}$$

3.6 Die Siebformel

Seien A_1, A_n Ereignisse in W-Raum $(\Omega, \mathfrak{A}, P)$

$$\begin{aligned} P(A_1 \cup A_2) &= P(A_1 \cup (A_2 \cap (A_1 \cap A_2)^c)) \\ &= P(A_1) + P(A_2 \cap (A_1 \cap A_2)^c) \\ &= P(A_1) + P(A_2) - P(A_1 \cap A_2) \end{aligned} \tag{26}$$

Dies führt uns zu folgendem: Seien $A_1, \ldots, A_n$ Ereignisse. Dann

$$P(\bigcup_{k=1}^{n} A_k) = \sum_{k-1}^{n} (-1)^{k-1} \sum_{1 \le j_1 < \ldots < j_k \le n} P(A_{j_1} \cap \ldots \cap A_{j_k}) \tag{27}$$

Wie leicht überprüfbar ist, ist (26) eine Anwendung von (27).

3.7 Beispiel

Sei

- $\Omega = \{\pi; \pi : \{1, \ldots, n\} \to \{1, \ldots, n\} \text{ bijektiv}$
- $|\Omega| = n!$
- $\mathfrak{A} = \mathfrak{P}(\Omega)$
- P das diskrete Laplace'sche W-Maß auf Ω
- $A_i = \{\pi \in \Omega; \pi(i) = i\},\ 1 \le i \le n$
- $\bigcup_{i=1}^{n} A_i = \{\pi \in \Omega; \text{Es gibt ein } i \in \{1, \ldots, n\} \text{ mit } \pi(i) = i\}$

Dann

$$
\begin{aligned}
P(\bigcup_{i=1}^{n} A_i) &= \sum_{k-1}^{n} (-1)^{k-1} \sum_{1\le j_1<\ldots<j_k\le n} \frac{|A_{j_1}\cap\ldots\cap A_{j_k}|}{|\Omega|} \\
&= \sum_{k=1}^{n} (-1)^{k-1} \sum_{\ldots} \frac{(n-k)!}{n!} \\
&= \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \cdot \frac{(n-k)!}{n!} \\
&= \sum_{k=1}^{n} (-1)^{k-1} \frac{1}{k!}
\end{aligned}
\tag{28}
$$

und wegen

$$
\begin{aligned}
\bigcap_{i=1}^{n} A_i^c &= \left(\bigcup_{i=1}^{n} A_i\right)^c \\
&= \{\pi\in\Omega; \forall i\in\{1,\ldots,n\} \text{ ist } \pi(i)\neq i\}
\end{aligned}
\tag{29}
$$

ist

$$
\begin{aligned}
P\left(\bigcap_{i=1}^{n} A_i^c\right) &= 1 - \sum_{k=1}^{n} (-1)^{k-1}\frac{1}{k!} \\
&= \sum_{k=0}^{n} (-1)^{k}\frac{1}{k!} \\
&\xrightarrow[n\to\infty]{} \sum_{k=0}^{\infty} (-1)^{k}\frac{1}{k!} \\
&= e^{-1} \\
&\approx 0.37
\end{aligned}
\tag{30}
$$

3.8 Weitere Beispiele für W-Maße

Sei $\emptyset \neq \Omega$, $\mathfrak{A}$ σ-Algebra mit $\{\omega\}\in\mathfrak{A}$ für jedes $\omega\in\Omega$.
Sei P W-Maß auf Ω mit der Eigenschaft, dass eine abzählbare Menge $\Omega_o \subseteq \Omega$ existiert mit $P(\Omega_0) = 1$
Dann

$$P\left(\Omega_0^c\right) = 1 - P\left(\Omega_0^c\right) = 0$$

Für $A \in \mathfrak{A}$ gilt

$$\begin{aligned} P(A) &= P\left((A \cap \Omega_0) + (A \cap \Omega_0^c)\right) \\ &= P(A \cap \Omega_0) + P(A \cap \Omega_0^c) \\ &= P\left(\sum_{\omega \in A \cap \Omega_0} \{\omega\}\right) \\ &= \sum_{\omega \in A \cap \Omega_o} P\left(\{\omega\}\right) \\ &= \sum_{\omega \in \Omega} P\left(\{\omega\}\right) \delta_\omega(A) \end{aligned} \tag{31}$$

Für $\omega \in \Omega$ heißt $\delta_\omega : \mathfrak{A} \to [0,1]$, definiert durch

$$\delta_\omega(A) = \begin{cases} 1 & \omega \in A, \\ 0 & \omega \notin A. \end{cases}$$

das Einpunktmaß oder **Dirac-Maß in** ω.

Es gilt also

$$P = \sum_{\omega \in \Omega_0} P(\{\omega\}) \cdot \delta_\omega.$$

Solche W-Maße heißen **diskrete W-Maße**.

3.9 Die Borelsche σ-Algebra

Betrachte den Fall $\Omega = \mathbb{R}$

$$\mathfrak{B} = \mathfrak{A} = \bigcap_{\substack{\mathfrak{A}^* \text{ } \sigma\text{-Algebra auf } \mathbb{R} \\ \mathfrak{A}^* \supseteq \{(-\infty, x]; x \in \mathbb{R}\}}} \mathfrak{A}^*$$

ist die Borelsche σ-Algebra auf $\mathbb{R}$.
Beachte

$$\mathfrak{B} \neq \mathfrak{P}(\mathbb{R})$$

$\mathfrak{B}$ enthält alle offenen, abgeschlossenen und ohne Intervalle Teilmengen von R.

3.10 Verteilungsfunktionen

Sei $F : \mathbb{R} \to \mathbb{R}$ Funktion mit folgenden Eigenschaften:

1. F ist monoton wachsend

2. F ist rechtsseitig stetig

3. $$\lim_{x \to -\infty} F(x) = 0$$

4. $$\lim_{x \to +\infty} F(x) = 1$$

Es existiert genau ein W-Maß P auf $\mathfrak{B}$ (auf $\mathbb{R}$) mit der Eigenschaft, dass

$$P\left((-\infty, x]\right)\right) = F(x), \ \forall x \in \mathbb{R} \tag{32}$$

F heißt die zu P gehörige **Verteilungsfunktion**.
Für $(a, b]$ mit $-\infty < a < b < +\infty$ ist

$$P((a, b]) = P((-\infty, b]) - P(-\infty, a]) = F(b) - F(a) \tag{33}$$

3.11 Verteilungsdichten

Sei $f : \mathbb{R} \to \mathbb{R}_+$ (uneigentlich) Riemann-integrierbar mit

$$\int_{-\infty}^{+\infty} f(t)dt = 1 \tag{34}$$

Setze
$$F(x) = \int_{-\infty}^{+\infty} f(t)dt, \ -\infty < x < +\infty \tag{35}$$

F ist Verteilungsfunktion eines W-Maßes P auf $\mathbb{R}$.
f heißt **Wahrscheinlichkeitsdichte** (oder einfach „Dichte") von P.

3.11.1 Spezialfälle

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

- $$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{1}{2}\frac{(t-a)^2}{\sigma^2}), -\infty < t < +\infty, (a \in R, \sigma^2 > 0)$$

(Anmerkung: Diese Verteilungsdichte wird auch Standardnormalverteilung genannt. Die Verteilungsfunktion dafür produziert die so genannte Gauß'sche Glockenkurve)
Dann ist

$$\int_{-\infty}^{+\infty} f(t)dt = 1$$

und somit

$$\begin{aligned} F(x) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{1}{2}\frac{(t-a)^2}{\sigma^2})dt \\ &= \phi(\frac{x-a}{\sigma}) \\ \phi(x) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{t^2}{2})dt, x \in \mathbb{R} \end{aligned}$$

Sei $(\Omega, \mathfrak{B}, P)$ der W-Raum unserer Wahl. P habe Dichte f, also eine Teilfunktion

$$f : R \to \int_{-\infty}^{+\infty} f(t)dt) = 1$$

$$F(x) = P((-\infty, x]) = \int_{-\infty}^{+\infty} f(t)dt), \ x \in \mathbb{R}$$

Allgemein gilt $P(E) = \int_E f(t)dt), \ E \in \mathfrak{B}(geeignet)$.

Beispiel:

$$f(t) = \begin{cases} 0 & t < 0 \\ \lambda \cdot exp(-\lambda t) & t \geq 0 \end{cases}$$

Aus

$$\int_0^x \lambda \cdot exp(-\lambda t)dt = [-exp(-\lambda t)]_o^x = 1 - exp(-\lambda x)$$

folgt

$$F(x) = \int_{-\infty}^{+\infty} f(t)dt = \begin{cases} 0 & x < 0 \\ 1 - exp(-\lambda x) & x \geq 0 \end{cases}$$

4 Weitere Eigenschaften von Wahrscheinlichkeitsmaßen

Sei $(\Omega, \mathfrak{A}, P)$ W-Raum. Seien $A_1, A_2, \ldots \in \mathfrak{A}$. Dann gilt

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n) \tag{36}$$

Diese Eigenschaft wird als **Sub-σ-Additivität** bezeichnet.

$$\bigcup_{n=1}^{\infty} A_n = A_1 + \sum_{n=2}^{\infty} (A_n \cap A_1^c \cap \ldots \cap A_{n-1}^c) \tag{37}$$

Wir können die Vereinigung auch als Summe von disjunkten Ereignissen auffassen und gelangen zu dieser Schreibweise in (36). Deshalb gilt auch (37)

$$\begin{aligned}
&P\left(\bigcup_{n=1}^{\infty} A_n\right) \\
&= \sum P(A_1) + \sum_{n=2}^{\infty} P(A_n \cap A_1^c \cap \ldots \cap A_{n-1}^c) \\
&\leq \sum_{n=1}^{\infty} P(A_n)
\end{aligned} \tag{38}$$

Speziell folgt hieraus die so genannte "Sub-Additivität von P":

$$P\left(\bigcup_{j=1}^{n} A_j\right) \leq \sum_{j=1}^{n} P(A_j) \tag{39}$$

Seien $A_n \in \mathfrak{A}, n = 1, 2, \ldots, A \in \mathfrak{A}$. Wir schreiben

- $A_n \uparrow A$ für $A_1 \subset A_2 \subset \ldots, A = P\left(\bigcup_{j=1}^{n} A_j\right)$
  Man sagt "A_n konvergiert **isoton** gegen A")

- $A_n \downarrow A$ für $A_1 \subset A_2 \subset \ldots, A = P\left(\bigcap_{j=1}^{n} A_j\right)$
  Man sagt "A_n konvergiert **antiton** gegen A")

5 Stetigkeitseigenschaften von W-Maßen

1. $A_n \uparrow A \Longrightarrow \lim_{n\to\infty} P(A_n) = P(A)$

2. $A_n \downarrow A \Longrightarrow \lim_{n\to\infty} P(A_n) = P(A)$

5.1 Beweis zu (1)

$$A = \bigcup_{n=1}^{\infty} A_n = A_1 + \sum_{n=2}^{\infty} (A_n \cap A_{n-1}^c) \tag{40}$$

Also

$$\begin{aligned}
P(A) &= P(A) + \sum_{n=2}^{\infty} P(A_n \cap A_{n-1}^c) \\
&= P(A) + \sum_{n=2}^{\infty} P(A_n - A_{n-1}^c) \\
&= P(A) + \lim_{m\to\infty} \sum_{n=2}^{m} P(A_n - A_{n-1}) \\
&= P(A_1) + \lim_{m\to\infty} (P(A_m - A_1)) \\
&= \lim_{m\to\infty} P(A_m)
\end{aligned} \tag{41}$$

5.2 Beweis zu (2)

Mit Gegenmenge ...

5.3 Beispiel

$\Omega = R, x \in R, (x_n) \subset R$ Folge, die monoton fallend gegen x konvergiert.

$$(-\infty, x] = \bigcap_{n=1}^{\infty} (-\infty, x_n]$$

P ist W-Maß auf R

$$F(x) = P((-\infty, x]) = \lim_{n\to\infty} P((-\infty, x_n]) = \lim_{n\to} F(x_n)$$

Hieraus folgt allgemein:
Ist $(x_n) \subset R$ eine Folge, die von rechts gegen x konvergiert (d.h. $x \leq x_n$ für jedes $n \leftarrow N$ und $\lim_{n\to} x_n = x$), so gilt

$$\lim_{n\to\infty} F(x_n) = F(x)$$

(“F ist rechtsseitig stetig”)

Teil III
Zufallsvariablen und ihre Verteilung

Wir gehen von einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ aus, R sei unsere gesuchte Bild-Menge, $\mathfrak{S}$ σ-Algebra auf R.

$$X : \Omega \to R$$

sei Wahrscheinlichkeitsfunktion mit der Eigenschaft, dass für jedes $B \in \mathfrak{S}$ gilt

$$X^{-1}(B) = \{\omega \in \Omega; X(\omega) \in B\} \in \mathfrak{A} \tag{42}$$

heißt Zufallsvariable.

- $$\{X \in B\} = \{\omega \in \Omega; X(\omega) \in B\} = X^{-1}(B) \tag{43}$$
- $$P(X \in B) \text{ für } P(\{X \in B\}) = P(X^{-1}(B)) \tag{44}$$
- Durch
  $$P^X : \mathfrak{S} \to [0, 1]$$
  definiert durch
  $$P(B) = P(X^{-1}(B)) = P(X \in B),\ B \in \mathfrak{S}$$
  wird ein W-Maß P^X auf $\mathfrak{S}$ definiert.
- $$X^{-1}\left(\sum_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} X^{-1}(B_n)$$

P^X heißt die **Verteilung von** X.

6 Spezialfälle

1. $R = \mathbb{R}, \mathfrak{S} = \mathfrak{B}$. Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ heißt **reelle Zufallsvariable**.
   $F(x) = P^X((-\infty, x]))P(X \in ((-\infty, x])) = P(X \leq x), x \in \mathbb{R}$ ist die Verteilungsfunktion von P^X, kurz: Die Verteilungsfunktion von X.
   Hat P^X die Dichte f, so ist
   $$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt, x \in \mathbb{R}$$
   Sprechweise: X hat die Dichte f.

2. $R \neq \emptyset$, $\mathfrak{S}$ enthalte die Mengen $\{x\}$ für $x \in R$.
   Es gibt eine abzählbare Menge $R_0 \subset R$ mit $P(X \in R_0) = 1$.
   Dann ist $P(X \notin R_0) = 1 - P(X \in R_0) = 0$
   Es gilt daher
   $$P(X \in R_0^c \cap B) = 0,\ \forall B \in \mathfrak{S}$$
   und
   $$\begin{aligned} P(X \in B) &= P(X \in R_0 \cap B) + \cancel{P(X \in R_0^c \cap B)} \\ &= \sum_{x \in R_0 \cap B} P(X = x) \\ &= \sum_{x \in R_0} P(X = x)\delta_x(B) \end{aligned}$$
   Also
   $$P^X = \sum_{x \in R_0} P(X = x)\delta_x$$
   X heißt dann diskret verteilt.
   P^X ist festgelegt durch die Werte $P(X = x), x \in R_0$

7 Beispiel: Würfelwurf von 2 echten Würfeln

- $\Omega = \{(\omega_1, \omega_2); \omega_1, \omega_2 \in \{1, \ldots, 6\}\}$
- $|\Omega| = 6^2 = 36$
- $\mathfrak{A} = \mathfrak{P}(\Omega)$
- $P(A) = \frac{|A|}{|\Omega|}, A \subset \Omega$
- $X : \Omega \to \mathbb{R}$, $X(\omega_1, \omega_2) = \omega_1 + \omega_2, (\omega_1, \omega_2) \in \Omega$
- $P(X = k)$

Also

$$\begin{aligned} P(X = k) &= P(\{(\omega_1, \omega_2) \in \Omega; \omega_1 + \omega_2 = k\}) \\ &= \frac{|\{(\omega_1, \omega_2) \in \Omega; \omega_1 + \omega_2 = k\}|}{36} \\ &= \begin{cases} \frac{1}{36} & , k = 2 \\ \frac{2}{36} & , k = 3 \\ \frac{3}{36} & , k = 4 \\ \vdots & \\ \frac{1}{36} & , k = 12 \end{cases} \end{aligned} \tag{45}$$

Natürlich kann man das auch allgemeiner schreiben:
Sei $R = \mathbb{R}^d = \mathfrak{B}^d$. X heißt dann d-dimensionaler Zufallsvektor $X : \Omega \to \mathbb{R}^d$.
$X = (X_1, \ldots, X_d)$ mit reellen Zufallsvariablen $X_1, \ldots, X_n$ die Komponenten von X. X heißt nun **Zufallsvektor**.
Sei $f : \mathbb{R}^d \to \mathbb{R}_+$ mit

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(t_1, \ldots, t_d) dt_1 \cdots dt_d = 1$$

f heißt Dichte des Zufallsvektors X, wenn gilt

$$\begin{aligned} F(x_1, \ldots, x_d) &= P(X_1 \leq x_1, \ldots, X_d \leq x_d) \\ &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(t_1, \ldots, t_d) dt_1 \cdots dt_d, \ (x_1, \ldots, x_d) \in \mathbb{R}^d \end{aligned}$$

F heißt Verteilungsfunktion des Zufallsvektors.

8 Beispiele für Zufallsvariablen und ihre Verteilungen

Sei $(\Omega, \mathfrak{A}, P)$ W-Raum und $X : \Omega \to \mathbb{R}$ reelle Zufallsvariable. X hat dann eine Dichtefunktion (oder kurz „Dichte") $f(\geq 0)$, wenn gilt

$$F(x) = P(X \leq x) = \int_{-\infty}^{+\infty} f(t)dt, \ -\infty < x < +\infty \tag{46}$$

....

8.1 Die Rechteck- oder Gleichverteilung

X heißt gleichverteilt oder Rechteck-verteilt auf $[a, b]$. Die Schreibweise hierfür sieht aus wie folgt

$$X \sim \mathfrak{R}(a, b)$$

Soviel zu Rechteck- und Gleich-verteilung.

8.2 Die Normalverteilung

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot exp(-\frac{1}{2} \cdot \frac{(t-\mu)^2}{\sigma^2}), \ -\infty < t < +\infty (\mu \in \mathbb{R}, \sigma^2 > 0) \tag{47}$$

Die Dichte sieht aus wie folgt

$$F(x) = P(X \leq x) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot exp(-\frac{1}{2} \cdot \frac{(t-\mu)^2}{\sigma^2})dt \tag{48}$$

Für

- $s := \frac{t-\mu}{\sigma}$
- $t = \mu + \sigma s$
- $dt = \sigma ds$

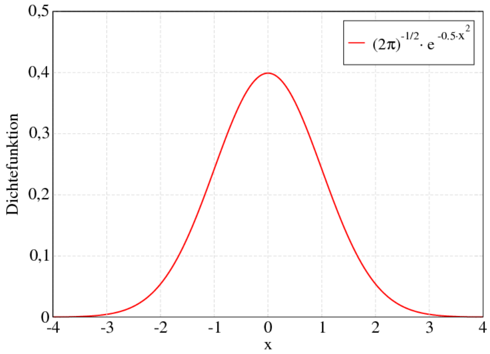
ist (oben)

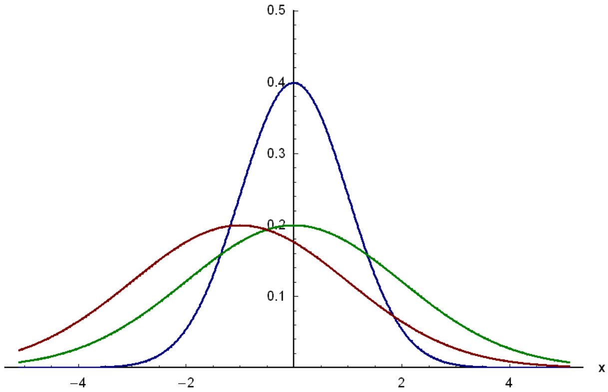
$$\begin{aligned} () &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot exp(-\frac{1}{2} \cdot s^2)dt \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right) \\ \Phi(x) &= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} exp(-\frac{1}{2} \cdot s^2)dt, \ -\infty < x < +\infty \end{aligned} \tag{49}$$

X heißt Normal-verteilt mit dem Parameter $\mu \in \mathbb{R}$ und $\sigma^2 > 0$. Schreibweise:

$$X \sim N(\mu, \sigma^2) \tag{50}$$

8.2.1 Beispiele für Normalverteilungen

- 

  Das [2] ist die so genannte Standard-Normal-Verteilung, auch bekannt als Gauß 'sche Glockenkurve
- 

  Zum Vergleich Dichtefunktionen der Normalverteilungen $N(0,1)$ (blau), $N(0,2)$ (grün) und $N(-1,2)$ (rot)

8.3 Die Exponential-Verteilung

$$f(t) = \begin{cases} 0 & , \ t < 0 \\ \lambda \cdot exp(-\lambda t) & , \ t \geq 0 \end{cases} \tag{51}$$

Hierbei sei $\lambda > 0$.
Dann

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \ x < 0 \\ 1 - exp(-\lambda t) & , \ x \geq 0 \end{cases} \tag{52}$$

X heißt Exponential-verteilt mit dem Parameter $\lambda > 0$. Schreibweise:

$$X \sim Exp(\lambda) \tag{53}$$

[2] Quelle Wikipedia

$X : \Omega \to \mathbb{R}^2$ sei ein 2-dimensoinaler Zufalsvektor, $X = (X_1, X_2)$. X habe die Dichte $f(\geq 0)$, d.h.

$$\begin{aligned} P(X_1 \leq x_1, X_2 \leq x_2) &= \int_{-\infty}^{x_2} \left(\int_{-\infty}^{x_1} f(t_1, t_2) dt_2 \right) dt_1 \\ &= \int_{-\infty}^{x_1} \left(\int_{-\infty}^{x_2} f(t_1, t_2) dt_1 \right) dt_2 \end{aligned} \tag{54}$$

Dann ist

$$\begin{aligned} P(X_1 \leq x_1) &= \int_{-\infty}^{x_1} \left(\int_{-\infty}^{+\infty} f(t_1, t_2) dt_2 \right) dt_1 \\ &= \int_{-\infty}^{x_1} f_{X_1}(t_1) dt_1 \end{aligned} \tag{55}$$

Also ist

$$f_{X_1}(t_1) = \int_{-\infty}^{+\infty} f(t_1, t_2) dt_2, \ -\infty < t_1 < +\infty \tag{56}$$

Dichte von X_1

Analog errechnet sich die Dichte von X_2

Teil IV
Bedingte Wahrscheinlichkeiten

9 Bedingte Wahrscheinlichkeiten

$(\Omega, \mathfrak{A}, P)$ ist W-Raum. Zunächst $|\Omega| < \infty$, $\mathfrak{A} = \mathfrak{P}(\Omega)$, $P(A) = \frac{|A|}{|\Omega|}$, $A \subset \Omega$, $A \subset \Omega, B \subset \Omega, |B| \geq 1$. Dann

$$\frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{P(A \cap B)}{P(B)} \tag{57}$$

„Wahrscheinlichkeit für das Eintreten von A unter Bedingung, dass B eintritt“, lese als „Wahrscheinlichket von A unter B“

9.1 Definition

$(\Omega, \mathfrak{A}, P)$ W-Raum (bel), $A, b \in \mathfrak{A}, P(B) > 0$. Dann heißt

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{58}$$

die bedingte Wahrscheinlichkeit von A unter B („A gegeben B“).

9.2 Satz

$(\Omega, \mathfrak{A}, P)$ W-Raum, $\emptyset \neq I$ abzählbare Indexmenge. Wir unterstellen $A_i \in \mathfrak{A}, i \in I$ seien paarweise disjunkt, $P(A_i) > 0$ für jede $i \in I$. Dann halten wir fest

1. Für $A \in \mathfrak{A}$ mit der Eigenschaft $A \in \sum_{i \in I} A_1$ ist $P(A) = \sum_{i \in I} P(A|A_i)P(A_i)$ (bekannt als „Satz von der totalen Wahrscheinlichkeit“)
2. Sei $A \in \mathfrak{A}$ mit $P(A) > 0, i_0 \in I$. Dann
$$P(A_{i_0}|A) = \frac{P(A|A_{i_0})P(A_{i_0})}{\sum_{i \in I} P(A|A_i)P(A_i)} \tag{59}$$
(bekannt als „Formel von Bayes“)
3. Sei $I = \{0, 1, \ldots, n\}$, $(n \in \mathbb{N}\text{fest})$. Dann gilt
$$P(A_0 \cap A_1 \cap \ldots \cap A_n) = P(A_0)P(A_1|A_0)P(A_2|A_0 \cap A_1) \cdots P(A_n|A_0 \cap \ldots \cap A_{n-1}) \tag{60}$$
(bekannt als „Multiplikationsformel“)

9.2.1 Beweis

1. Aus
$$A = \sum_{i \in I} (A \cap A_i)$$
folge sofort
$$P(A) \sum_{i \in I} P(A \cap A_i) = \sum_{i \in I} P(A|A_i)P(A_i)$$
2. folgt sofort aus 1.
3. $$\begin{aligned}&P(A_0)P(A_1|A_0)P(A_2|A_0 \cap A_1) \cdots P(A_n|A_0 \cap \ldots \cap A_{n-1})\\ =&\cancel{P(A_0)} \cdot \frac{\cancel{P(A_1 \cap A_0)}}{\cancel{P(A_0)}} \cdot \frac{P(A_0 \cap A_1 \cap A_2)}{\cancel{P(A_0 \cap A_1)}} \cdots\\ =&P(A_0 \cap A_1 \cap \ldots \cap A_n)\end{aligned} \tag{61}$$

9.3 Beispiel: Lostrommel

Eine Lostrommel enthält a Gewinnlose und b Nieten, $A \geq 1, b \geq 2$. Es gibt 2 Spielmöglichkeiten:

1. Ein Los wird gezogen. Das gezogene Los ist ein Gewinnlos oder eine Niete. Das Spiel ist beendet.

2. Ein Los wird gezogen und unbesehen weggeworfen. Daraufhin entfernt der Losverkäufer eine Niete aus der Trommel. Ein weiteres Los wird gezogen. Dieses Los ist ein Gewinnlos oder eine Niete.

Sei A das Ereignis, dass bei der ersten Ziehung ein Gewinnlos gezogen wird. Sei B das Ereignis, dass bei der zweiten Ziehung ein Gewinnlos gezogen wird.

Dann ergibt sich bei diesen Fällen

1.
$$P(A) = \frac{a}{a+b} \tag{62}$$

2.
$$\begin{aligned} P(B) &= P(B|A) \cdot P(A) + P(B|A^c)P(A^c) \\ &= \frac{a-1}{a+b-2} \cdot \frac{a}{a+b} + \frac{a}{a+b-2} \cdot \frac{b}{a+b} \\ &= \frac{a}{a+b} \cdot \left(\underbrace{\frac{a+b-1}{a+b-2}}_{\geq 1} \right) > \frac{a}{a+b} > P(A) \end{aligned} \tag{63}$$

Sei nun $a = 1,\ b = 2$. Dann folgt sofort

$$\begin{aligned} P(A) &= \frac{1}{3} \\ P(B) &= \frac{1}{3} \cdot 2 = \frac{2}{3} \end{aligned}$$

Sei $A, B \in \mathfrak{A}, P(B) > 0$. Dann gilt offenbar

$$P(A|B) = P(A) \iff P(A \cap B) = P(A)P(B) \tag{64}$$

10 Stochastische Unabhängigkeit von Ereignissen und Zufallsvariablen

10.1 Definition

$(\Omega, \mathfrak{A}, P)$ W-Raum, $A, B \in \mathfrak{A}$. A, B heißen (stochastisch) unabhängig, wenn

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

gilt.

10.2 Einsicht / Beweis

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)P(B), \text{ falls } A, B \text{ unabhängig} \\ &(1 - P(B))P(A) \\ &= P(A)P(B^c) \end{aligned} \tag{65}$$

10.3 Definition

- $\emptyset \neq I$ endliche Indexmenge, $A_i \in \mathfrak{A}, i \in I$ Ereignisse.
  A_i heißen (stochastisch) unabhängig, wenn für jede $\emptyset \neq J \subset I$ gilt

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j) \tag{66}$$

- $\emptyset \neq I$ beliebige Indexmenge, $A_i \in \mathfrak{A}, i \in I$ Ereignisse.
  $\mathfrak{A}_i, i \in I$ heißen unabhängig, wenn für jede Teilmenge $\emptyset \neq J \subset I,\ J$endlich die Ereignisse $A_j, j \in J$ unabhängig sind.

10.3.1 Bemerkung

$A_i \in \mathfrak{A}, i \in I$ seien Ereignisse. Dann gilt

$$A_i, i \in I \text{ unabhängig} \iff B_i, i \in I \text{ unabhängig },\ B_i \in \{A_i, A_i^c\} \text{ beliebig} \tag{67}$$

10.4 Definition

$(\Omega, \mathfrak{A}, P)$ W-Raum, $\emptyset \neq I$ Indexmenge, $(R_i, \mathfrak{S})$ Messräume, $i \in I$, $X_i : \Omega \to R_i$ Zufallsvariable, $i \in I$.
$X_i, i \in I$ heißen (stochastisch) unabhängig, wenn für jede Wahl von $B_i \in \mathfrak{S}_i, i \in I$ die Ereignisse $A_i = X_i^{-1}(B_i), i \in I$ unabhängig sind.

$(\Omega, \mathfrak{A}, P)$ sei W-Raum. Dann $X_i : \Omega \Rightarrow R_i,\ i = 1, \ldots, n$Zufallsvariable $X_1, \ldots, X_n$ sind unabhängig, wenn gilt

$$P(X_1 \in B_1, \ldots, X_n \in B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n)$$

für jede Auswahl von Ereignissen $B_i \in \mathfrak{A}, i = 1, \ldots, n$

...

10.5 Spezialfälle

10.5.1 Diskrete Verteilung

Die $X_1, \ldots, X_n$ haben je eine diskrete Verteilung. Dann sind

$$X_1, \ldots, X_n \text{ unabhängig} \iff P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n)$$

für jede Auswahl von Elementen $x_i \in R_i,\ i = 1, \ldots, n$.
(Beachte hier: $\{x\} \in R_i,\ \forall x \in R,\ \forall i \in I$)

10.5.2 Stetige Verteilung

Sei $X = (X_1, \ldots, X_n)$ mit den reellen Zufallsvariablen $X_1, \ldots, X_n$. Es habe X eine Dichte f, das heißt

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f(t_1, \ldots, t_n) dt_1, \ldots, dt_n,\ (x_1, \ldots, x_n \in R^n)$$

Dann hat X_i die Dichte

$$f_i(x_i) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(t_1, \ldots, t_{i-1}, x_i, t_{i+1}, \ldots, t_n) dt_1, \ldots, dt_n$$

Gilt

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i),\ (x_1, \ldots, x_n) \in \mathbb{R}^n \tag{68}$$

so sind die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig.

10.6 Beispiel

Ein Einzelexperiment, bei dem ein interessierendes Ereignis A mit Wahrscheinlichkeit $p \in [0, 1]$ eintritt, wird n-mal unter identischen Versuchsbedingungen ohne gegenseitige Beeinflussung (unabhängige Versuchswiederholungen) wiederholt.
Beschreibe das Einzelexperiment durch den W-Raum $(\Omega_0, \mathfrak{A}_0, P_0) = (\{0, 1\}, \mathfrak{P}(\{0, 1\}), P_0)$ mit

- $P_0(\{0\}) = 1 - p$

- $P_0(\{1\}) = p$

- $0 \cong A$ tritt nicht ein

- $1 \cong A$ tritt ein

Beschreibe das Gesamtexperiment durch

$$(\Omega, \mathfrak{A}, P) \text{ mit } \Omega = \{0,1\}^n = \{(\omega_1, \ldots, \omega_n); \omega_i \in \{0,1\}, i = 1, \ldots, n\}, \mathfrak{A} = \mathfrak{P}(\Omega)$$

$$\begin{aligned} P(\{\omega_1, \ldots, \omega_n\}) &= P_0(\{\omega_1\}) \cdot P_0(\{\omega_n\}) \\ &= p^{\omega_1} \cdot (1-p)^{1-\omega_1} \cdots p^{w_n} \cdot (1-p)^{1-\omega_n} \\ &= p^{\omega_1 + \ldots + \omega_n} \cdot (1-p)^{n-(\omega_1+\ldots+\omega_n)}, (\omega_1, \ldots, \omega_n) \in \Omega \end{aligned} \tag{69}$$

$X_i : \Omega \to \{0,1\}$ sei definiert durch

$$X_i(\omega_1, \ldots, \omega_n) = \omega_i, (\omega_1, \ldots, \omega_n) \in \Omega, i = 1, \ldots, n$$

Die $X_1, \ldots, X_n$ sind unabhängig, denn

$$\begin{aligned} &P(X_1 = \epsilon_1, \ldots, X_n = \epsilon_n) \\ =&P(\omega = \{1, \ldots, n\} \in \Omega;\ X_1(\omega_1, \ldots, \omega_n) = \epsilon_1, \ldots, X_n(\omega_1, \ldots, \omega_n) = \epsilon_n) \\ =&P(\{(\epsilon_1, \ldots, \epsilon_n)\} \\ =&p^{\epsilon_1 + \ldots + \epsilon_n} \cdot (1-p)^{n-(\epsilon_1+\ldots+\epsilon_n)} \\ =&\prod_{i=1}^{n} p_i^{\epsilon_i} \cdot (1-p)^{1-\epsilon_i} \end{aligned} \tag{70}$$

$$\begin{aligned} P(X_1 = \epsilon_1) &= P(\{(\omega_1, \ldots, \omega_n) \in \Omega; X_1(\omega_1, \ldots, \omega_n) = \epsilon_1\}) \\ &= P(\{(\epsilon_1, \omega_2, \ldots, \omega_n); \omega_i \in \{0,1\}, 2 \leq i \leq n\}) \\ &= \sum_{nnn} P \ldots \end{aligned} \tag{71}$$

$$P(X = k) = \ldots \tag{72}$$

11 Binomialverteilung

11.1 Definition

Eine reelle (oder $\{0,1\}$-wertige) Zufallsvariable X heißt Binomialverteilt mit den Parametern $n \in \mathbb{N}$ und $p \in [0,1]$, wenn gilt

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}, k \in \{0, 1, \dots, n\} \tag{73}$$

Schreibweise:

$$X \sim \mathfrak{B}(n, p)$$

11.2 Anwendung

Sei $(\Omega, \mathfrak{A}, P)$ W-Raum, $A \in \mathfrak{A}$ Ereignis. Dann heißt $I_A : \Omega \to \{0,1\}$ **Indikator**-funktion (oder Indikatorvariable) von A, definiert durch

$$I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A. \end{cases}$$

$I_A \sim \mathfrak{B}(1, p)$ mit $p = P(A)$.
Seien $A_1, \dots, A_n \in \mathfrak{A}$ unabhängige Ereignisse mit $p = P(A_i), i = 1, \dots, n$..
Dann gilt $\underbrace{X_1}_{=I_{A_1}}, \dots, \underbrace{X_n}_{=I_{A_n}}$ unabhängig und $X_i \sim \mathfrak{B}(1,p)$. $X = X_1 + \ldots + X_n \sim \mathfrak{B}(n, p)$.

11.3 Beispiel Urnenmodell

Eine Urne enthält r rote und s schwarze Kugeln. Uns können verschiedene Modi interessieren.

11.3.1 Mit Zurücklegen

Es wird n-mal mit Zurücklegen je eine Kugel gezogen.
Sei X Zufallsvariable und bezeichne die Anzahl der gezogenen roten Kugeln in n Ziehungen.
Dann $X \sim \mathfrak{B}\left(n, \frac{r}{r+s}\right)$

11.3.2 Ohne Zurücklegen

Es wird n-mal ohne Zurücklegen je eine Kugel gezogen mit $n \leq r + s =: a$.
Sei X Zufallsvariable und bezeichne die Anzahl der gezogenen roten Kugeln in n Ziehungen.

Nummeriere die Kugeln in folgender Weise:

$$\underbrace{1,\dots,r}_{\text{rot}},\underbrace{r+1,\dots,\overbrace{r+s}^{=a}}_{\text{schwarz}}$$

Dann

- $\Omega = \{(\omega_1,\dots,\omega_n);\ \omega_i \in \{1,\dots,a\}\text{paarweise verschieden}, 1 \leq i \leq n\}$
- $\mathfrak{A} = \mathfrak{P}(\Omega)$
- $P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{a\cdot(a-1)\cdot(a-n+1)},\ A \subset \Omega$
-

$$\begin{aligned}
P(X=k) &= \frac{|\{(\omega_1,\dots,\omega_n)\in\Omega; |j\in\{1,\dots,n\},\omega_j\in\{1,\dots,r\}|=k\}|}{|\Omega|}\\
&= \frac{\binom{n}{k}\cdot r\cdot(r-1)\cdots(r-k+1)\cdot s\cdot(s-1)\cdots(s-(n-k)+1)}{a\cdot(a-1)\cdots(a-n+1)} \qquad (74)\\
&= \frac{\binom{r}{k}\cdot\binom{a-r}{n-k}}{\binom{a}{n}}, k\in\mathbb{N}, max(0,a-r-n)\leq k\leq min(n,r)
\end{aligned}$$

12 Hypergeometrische Verteilung

12.1 Definition

Eine reelle Zufallsvariable mit der Verteilung

$$P(X=k) = \frac{\binom{r}{k}\cdot\binom{a-r}{n-k}}{\binom{a}{n}}$$

heißt hypergeometrisch verteilt mit den Parametern $a, r, n \in \mathbb{N}, r \leq a, n \leq a$ und der Schreibweise

$$X \sim \mathfrak{H}(a,r,n). \qquad (75)$$

Das bekannteste Beispiel dürfte das Zahlenlotto “6 aus 49” sein.

Teil V
Erwartungswerte

13 Erwartungswerte von Zufallsvariablen

Wir definieren Eerwartungswerte von Zufallsvariablen schrittweise, beginnend mit Indikatorvariablen. Es sei X eine eine reelle Zufallsvariable auf dem Wahrscheinlichkeitsraum ().

(1) Für eine Indikatorvariable$X = I_A$ mit $A \in \mathfrak{A}$ sei $E(X) = E_p(X) = P(A)$

(2) Für eine nicht negative einfache Zufallsvariable $X = \sum_{k=1}^{n} \alpha_k I_{A_k}$ mit nicht negativen reellen Zahlen $\alpha_1, \dots, \alpha_n$ und Ereignissen $A_1, \dots, A_n \in \mathfrak{A}$ sei

$$E(X) = E_p(X) = \sum \alpha_k \cdot P(A_k) \tag{76}$$

Der so definierte Wert $E(X)$ hängt nicht von der Darstellung von X ab, d.h. ist $X = \dots$

(3) Für eine nicht negative Zufallsvariable $X \geq 0$ gibt es eine Folge von nicht negativen einfachen Zufallsvariablen $X_n,\ n \in \mathbb{N}$, mit $0 \leq X_1 \leq \dots$ mit $X = \sup_{n \in \mathbb{N}} X_n$. Es sei $E(X) = E - p(X) = \sup_{n \in \mathbb{N}} E(X)$

(4) Jede reelle Zufallsvariable X lässt sich in der Form $X = X^+ - X^-$ mit den nicht negativen Zufallsvariablen $X^+ = \max(X, 0)$

14 Eigenschaften von Erwartungswerten

Es seien X, Y reelle Zufallsvariablen auf dem Wahrscheinlichkeitsraum (), α, β reelle Zahlen.

(1) Sind $X, Y \geq 0$ und $\alpha, \beta \geq 0$, so ist

$$E(\alpha \cdot X + \beta \cdot Y) = \alpha \cdot E(X) + \beta \cdot E(Y)$$

(2) Sind X, Y integrierbar, so ist $\alpha \cdot X + \beta \cdot Y$ integrierbar und es ist

$$E(\alpha \cdot X + \beta \cdot Y) = \alpha \cdot E(X) + \beta \cdot E(Y)$$

(3) ist $0 \leq X \leq Y$, so ist $0 \leq E(X) \leq E(Y)$.

(4) Existiert für die Zufallsvariable X der Erwartungswert, so ist $|E(X)| \leq E(|X|)$.

(5) Sind X, Y integrierbar und ist $X \leq Y$, so ist $E(X) \leq E(Y)$.

(6) X ist genau dann integrierbar, wen $|X|$ integrierbar ist.

(7) Ist $|X| \leq |Y|$ und Y integrierbar, so ist auch X integrierbar.

(8) Sind $X, \dots, X_n,\ n \in \mathbb{N}$ nicht negative Zufallsvariablen mit $0 \leq X_1 \leq X_2 \leq \dots$ und $X = \sup_{n\in\mathbb{N}} X_n$, so ist $E(X) = \sup_{n\in\mathbb{N}} E(X_n)$

(9) Sind $X_n,\ n \in \mathbb{N}$ nicht negative Zufallsvariablen, so ist

$$E(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} E(X_n)$$

(10) Sind $X, Y, X_n,\ n \in \mathbb{N}$ reelle Zufallsvariablen mit $X = \lim_{n\to\infty} X_n$ und $|X_n| \leq |Y|$ für jedes $n \in \mathbb{N}$, und es ist Y integrierbar, so sind auch X, X_n integrierbar und es gilt $\lim_{n\to\infty} E(|X_n - X|) = 0$. Insbesondere folgt hieraus auch $\lim_{n\to\infty} E(X) = E(X)$.

(11) Ist $X \geq 0$, so ist $E(X) = 0$ genau dann, wenn $P(X > 0) = 0$ ist.

Aus der Eigenschaft (8) ergibt sich für eine Folge von nicht negativen, reellen Zufallsvariablen X_n mit der Eigenschaft, dass die Reihe $\sum_{k=1}^{n} X_k$ gegen die reelle Zufallsvariable X konvergiert, die Identität

$$E\left(\sum_{n=1}^{\infty} X_n\right) = \dots \tag{77}$$

...folgt hieraus

$$|X| = \sum_{n=1} |X| \cdot I(n-1 < X \leq n), \tag{78}$$

dass

$$\begin{aligned} E(|X|) &= \sum_{n=1}^{\infty} E(|X| \cdot I(n-1 < X \leq n)) \\ &\leq \sum_{n=1}^{\infty} P(n-1 < |X| \leq n) \\ &= \dots \\ &= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} P(n-1 < |X| \leq n) \\ &= \sum_{k=1}^{\infty} P(|X| > n-1) \\ &= \sum_{k=0}^{\infty} P(|X| > n) \end{aligned} \tag{79}$$

ist. Analgo sieht man

$$E() \dots \tag{80}$$

...

15 Rechnen mit Erwartungswerten

() W-Raum, $\emptyset \neq R$ Menge, $\mathfrak{L}\sigma$-Algebra auf R, $f : R \to \mathbb{R}$, $f^{-1}(B) \in \mathfrak{L}$ für jedes $B \in \mathfrak{B}$. Dann P^X Verteilung von X, P^X ist W-Maß auf R und $(R, \mathfrak{L}, P^X)$ ist W-Raum für reelle Zufallsvarieblen auf $(R, \mathfrak{L}, P^X)$.

Der Erwartungswert von $F \cdot X$ existiert genau da„ wenn der Erwartungswert von f, aufgefasst als Erwartungswert auf $(R, \mathfrak{L}, P^X)$, existiert. Es gilt dann die **Transformationsformel**

$$E(f \cdot X) = E_P(f \cdot X) = E_{P^X}(f) \tag{81}$$

16 Spezialfälle

16.1 Stetige Verteilung

$(R, \mathfrak{L}) = (\mathbb{R}, \mathfrak{B})$, X habe die Dichte g.

(1) $f = I_{(-\infty,x]}$

$$\begin{aligned} F(x) = P(X \leq x) &= \underbrace{E(I_{(-\infty,x]} \cdot X)}_{=E(f \cdot X)} \\ &= E_{P^X}(I_{(-\infty,x]}) \\ &= P^X((-\infty, x]) = \int_{-\infty}^{x} \\ &= \int_{-\infty}^{x} f(t)g(t)dt \end{aligned} \tag{82}$$

(2) Die Identität

$$E(f \cdot X) = \int_{-\infty}^{x} f(t)g(t)dt$$

gilt für “beliebige” $f : \mathbb{R} \to \mathbb{R}$ (mit $f^{-1}(B) \in \mathfrak{B}$ für jedes $B \in \mathfrak{B}$, mit existierendem $E(f \cdot X)$).

16.2 Diskrete Verteilung

X habe eine diskrete Verteilung, d.h. es gibt eine abzählbare Menge $A \subset R$ mit $P(X \in A) = 1$, $(\{x\} \in \mathfrak{L} \forall x \in R)$.

$$P^X = \sum_{a \in A} P(X = a) \cdot \delta_a$$

Für $f \geq 0$ ist

$$\begin{aligned} E(f \cdot X) &= E\left(\sum_{a \in A} f(a) \cdot I(X = a) + \underbrace{E\left(X \cdot I(X \neq A)\right)} = 0\right) \\ &= \sum_{a \in A} f(a) \cdot E(I(X = a)) \\ &= \sum_{a \in A} f(a) \cdot P(X = a) \end{aligned} \tag{83}$$

Diese Gleichung ist besonders wichtig, da sie für beliebige f gilt mit existierendem $E(f \cdot X)$

17 Anwendungen

17.1 Erwartungswert der Rechteck-Verteilung

Mit $X \sim \mathfrak{R}(a, b)$ gilt

$$g(t) = \begin{cases} \frac{1}{b-a} & , \ a \leq t \leq b \\ b & , \ \text{sonst} \end{cases}$$

Dann ist der Erwartungswert

$$E(X) = \frac{1}{b-a} \cdot \int_a^b t \cdot dt = \frac{a+b}{2} \tag{84}$$

17.2 Erwartungswert der Normalverteilung

Sei $X \sim \mathfrak{N}(a, \sigma)$. Dann ist der Erwartungswert.

...

17.3 Beispiel

17.3.1 Binomialverteilt

$X \sim \mathfrak{B}(n,p)$
Ohne Einschränkung $X = \sum_{j=1}^n I_{A_j}$
$p = P(A_j),\ j = 1, \ldots, n$

$$E(X) = \sum_{j=1}^n \underbrace{E(I_{A_j})}_{=P(A)=p} = n \cdot p$$

17.3.2 Hypergeometrisch verteilt

$X \sim \mathfrak{H}(a,r,n)$
$X = \sum_{j=1}^n I_{A_j}$
A_j ist das Ereignis "Beim n-fachen Ziehen ohne Zurücklegen je einer Kugel aus einer Urne mit r roten und $s = a - r$ schwarzen Kugeln und bei der j-ten Ziehung wird eine rote Kugel gezogen".
$E(X) = \sum_{j=1}^n P(A_j)$

$$P(A_j) = \frac{r \cdot \cancel{(a-1)} \cdot \cancel{(a-2)} \cdots \cancel{(a-n+1)}}{a \cdot \cancel{(a-1)} \cdot \cancel{(a-2)} \cdots \cancel{(a-n+1)}} = \frac{r}{a} = \frac{r}{r+s},\ j = 1, \ldots, n$$

Also: $E(X) = n \cdot \frac{r}{a}$
Beim n-fachen Ziehen mit Zurücklegen je eine Kugel aus einer Urne mit r roten und $s = a - r$ schwarzen Kugeln ist Anzahl X der gezogenen roten Kugeln $\mathfrak{B}(n, \frac{r}{a})$-verteilt, also

$$E(X) = n \cdot \frac{r}{a}$$

18 Die Varianz

Sei $r \in \mathbb{N}$, X reelle Zufallsvariable mit $E(|X|^r) < \infty$.
Für $1 \leq s \leq r, s \in \mathbb{N}$ gilt dann wegen $|X|^s \leq |X|^r + 1$, dass $E(|X|^s) < \infty$ ist. Sei X reeelle Zufallsvariable mit $E(X^2) < \infty$. Dann ist $E(|X|) < \infty$.

Für $a \in \mathbb{R}$ ist $|X - a|^2 \leq X^2 + 2 \cdot |a| \cdot |X| + a^2$, also $E(|X - a|^2) < \infty$. Es ist

$$\begin{aligned} E([X-a]^2) &= E\left(([X - E(X)] + [E(X) - a])^2\right) \\ &= E\left([X - E(X)]^2\right) + [E(X) - a]^2 + \underbrace{E\left(2 \cdot [E(X) - a] \cdot [X - E(X)]\right)}_{=0} \\ &\geq E\left([X - E(X)]^2\right) \end{aligned}$$

18.1 Definition

Sei X reelle Zufallsvariable mit $E(X^2) < \infty$. Dann heißt

$$Var(X) = E\left([X - E(X)]^2\right) \tag{85}$$

die **Varianz** von X.

Es ist

$$Var(X) = \min_{a \in \mathbb{R}} E\left([X - a]^2\right) \tag{86}$$

18.2 Bemerkung

Es ist

$$\begin{aligned} Var(X) = E\left([X - a]^2\right) &= E\left(X^2 - 2 \cdot X \cdot E(X) + (E(X))^2\right) \\ &= E(X^2) - 2 \cdot \left(E(X)^2 + (E(X))^2\right) \\ &= E(X^2) - (E(X))^2 \geq 0 \end{aligned} \tag{87}$$

18.3 Spezialfälle

18.3.1 Zufallsvariablen mit diskreter Verteilung

$$\begin{aligned} Var(X) &= \sum_{\substack{x \in \mathbb{R} \\ P(X0x)<0}} (x - E(X))^2 \cdot P(X = x) \\ &= \sum_{\substack{x \in \mathbb{R} \\ P(X0x)<0}} x^2 \cdot P(X = x) - \left(\sum_{\substack{x \in \mathbb{R} \\ P(X0x)<0}} x \cdot P(X = x) \right)^2 \end{aligned} \tag{88}$$

18.3.2 Zufallsvariablen mit stetiger Verteilung

Hat X die Dichte f, so gilt

$$\begin{aligned} Var(X) &= \int_{-\infty}^{+\infty} (x - E(X))^2 \cdot f(x) \cdot dx \\ &= \int_{-\infty}^{+\infty} x^2 \cdot f(x) \cdot dx - \left(\int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx \right)^2 \end{aligned} \tag{89}$$

18.4 Beispiele

18.4.1 Rechteckverteilt

$X \sim \mathfrak{R}(a, b)$

18.4.2 Normalverteilt

$X \sim \mathfrak{N}(a, \sigma^2)$

18.4.3 Exponentialverteilt

$X \sim Exp(\lambda)$. X hat Dichte

$$f(x) = \begin{cases} \lambda \cdot \exp(-\lambda \cdot x) & , \ x \geq 0 \\ 0 & , \ sonst \end{cases}$$

Dann ist

$$E(X) = \frac{1}{\lambda} \tag{90}$$

$$\begin{aligned} Var(X) &= \lambda \cdot \int_0^{+\infty} \left[x^2 \cdot \exp(-\lambda \cdot x) \cdot dx\right] - \left(\frac{1}{\lambda}\right)^2 \\ &= \frac{1}{\lambda^2} \end{aligned} \tag{91}$$

19 Die Kovarianz

Seien X, Y reelle Zufallsvariablen mit $E(X^2) < \infty$ und $E(Y^2) < \infty$.

$$|X| \cdot |Y| \leq \frac{1}{2} \cdot \left(|X|^2 + |Y|^2\right) \iff 0 \leq (|X| - |Y|)^2 \Rightarrow E(|X \cdot Y|) < \infty$$

und es gilt

$$E\left((X+Y)^2\right) = E(X^2) + E(Y^2) + 2 \cdot E(X \cdot Y)$$

Also ist

$$\begin{aligned} Var(X+Y) &= E\left([X + Y - (E(X) + E(Y))]^2\right) \\ &= E\left([(X - E(X)) + (Y - E(Y))]^2\right) \\ &= Var(X) + Var(Y) + 2 \cdot E([X - E(X)] \cdot [Y - E(Y)]) \end{aligned} \tag{92}$$

19.1 Definition

X, Y sind reelle Zufallsvariablen mit $E(X^2) < \infty,\ E(Y^2) < \infty$. Dann heißt

$$Cov(X,Y) = E([X - E(X)] \cdot [Y - E(Y)]) \tag{93}$$

die Kovarianz von X und Y.

19.2 Bemerkung

Es ist

$$Cov(X,Y) = E(X \cdot Y) - E(X)E(Y)$$

Also gilt

$$Var(X+Y) = Var(X) + Var(Y) + 2 \cdot Cov(X,Y) \tag{94}$$

Mittels vollständiger Induktion folgt für reelle Zufallsvariablen $X_1, \ldots, X_n$ mit $E(X_j^2) < \infty, j = 1, \ldots, n$.

$$\begin{aligned} Var(X_1 + \ldots, +X_n) &= \sum_{j=1}^{n} Var(X_j) + 2 \cdot \sum_{1 \le i < j \le n} Cov(X_i, X_j) \\ &= \sum_{j=1}^{n} Var(X_j) + 2 \cdot \sum_{i,j=1, i \neq j}^{n} Cov(X_i, X_j) \end{aligned}$$

19.3 Spezialfall

$X = I_a, Y = I_B$. Dann

$$Cov(X,Y) = E(\underbrace{I_A \cdot I_B}_{I_{A \cap B}} - \underbrace{E(I_A)}_{P(A)} \cdot \underbrace{E(I_B)}_{P(B)}$$

$$Cov(I_A, I_B) = P(A \cap B) - P(A) \cdot P(B)$$

Also insbesondere $Cov(I_A, I_B) = 0$ falls A, B unabhängige Ereignisse. Ferner gilt für die Ereignisse $A_1, \ldots, A_n$:

$$\begin{aligned} Var(I_{A_1} + \ldots + I_{A_n}) &= \sum_{j=1}^{n} \underbrace{Var(I_{A_i})}_{P(A-i) \cdot (1-P(A_i))} + 2 \cdot \sum_{1 \le i < j \le n} [P(A_i \cap A_j) - P(A_i) \cdot P(A_j)] \\ &= \sum_{j=1}^{n} P(A-i) \cdot (1 - P(A_i)) + 2 \cdot \sum_{1 \le i < j \le n} [P(A_i \cap A_j) - P(A_i) \cdot P(A_j)] \end{aligned}$$

19.4 Beispiele

19.4.1 Binomialverteilt

Sei $X \sim \mathfrak{B}(n,p)$. Dann wäre eigentlich

$$Var(X) = \sum_{k=0}^{n} (k - n \cdot p)^2 \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Wir vereinfachen das im Folgenden: O.B.d.A. sei $X = \sum_{j=1}^{n} I_{A_i}$ mit **unabhängigen** Ereignissen $A_i, i = 1, \dots, n$. Dann

$$Var(X) = \sum_{j=1}^{n} Var(I_{A_i}) = n \cdot p \cdot (1-p) \tag{95}$$

19.4.2 Hypergeometrisch verteilt

Sei $X \sim \mathfrak{H}(a, r, n)$. Dann wäre eigentlich

$$Var(X) = \sum_{k=\max(0,n+r-a)}^{\min(r,n)} \left(k - n \cdot \frac{r}{a}\right)^2 \cdot \frac{\binom{r}{k} \cdot \binom{a-r}{n-k}}{\binom{a}{n}}$$

Wir vereinfachen das im Folgenden: O.B.d.A. fassen wir X auf als die Anzahl der gezogenen roten Kugeln in n Ziehungen bei n-fachem Ziehen ohne Zurücklegen je eine Kugel aus einer Urne mit r roten und $s = a - r$ schwarzen Kugeln.
Dann $X = \sum_{j=1}^{n} I_{A_i}$ und mit $P(A_i) = \frac{r}{a}$ sei nun

$$\begin{aligned} Var(X) &= n \cdot \frac{r}{a} \cdot \left(1 - \frac{r}{a}\right) + n \cdot (n-1) \cdot \left[\frac{r \cdot (r-1)}{a \cdot (a-1)} - \left(\frac{r}{a}\right)^2\right] \\ &= \underbrace{n \cdot \frac{r}{a} \cdot \left(1 - \frac{r}{a}\right)}_{\text{Varianz beim Ziehen mit Zurücklegen}} \cdot \frac{a-n}{n-1} \end{aligned} \tag{96}$$

da

$$P(A_i \cap A_j = \frac{r \cdot (r-1)(a-2) \cdots (a-n+1)}{a(a-1)(a-2) \cdots (a-n+1)} = \frac{r \cdot (r-1)}{a(a-1)}$$

Seien X, Y **unabhängige** Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$.

1 $X = I_A, Y = I_B$. Dann sind A und B unabhängig.

$$\begin{aligned} E(XY) &= E(I_A I_B) = E(I_{A\cap B}) \\ &= P(A\cap B) = P(A)P(B) \\ &= E(I_A)E(I_B) \\ &= E(X)E(Y) \end{aligned}$$

2

$$0 \leq X = \sum_{i=1}^{m} \alpha_i I_{A_i}$$

mit paarweise verschiedenen $\alpha_1, ..., \alpha_m$, paarweise disjunkten $A_1, ..., A_m, A_1 \cap ... \cap A_m = \Omega$.

$$0 \leq Y = \sum_{j=1}^{n} \beta_j I_{B_j}$$

mit paarweise verschiedenen $\beta_1, ..., \beta_n$, paarweise disjunkten $B_1, ..., B_n, B_1 \cap ... \cap B_n = \Omega$.
Dann ist

$$A_i = \{X = \alpha_i\},\ i = 1, ..., m$$

$$B_j = \{Y = \beta_j\},\ j = 1, ..., n$$

und die Ereignisse $A_i, 1 \leq i \leq m$, sind unabhängig von den Ereignissen $B_1, ..., B_n$.

$$\begin{aligned} E(XY) &= E(\sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\beta_j \underbrace{I_{A_i} I_{B_j}}_{= I_{A_i\cap B_j}}) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\beta_j \underbrace{P(A_i \cap B_j}_{= P(A_i)P(B_j)} \\ &= (\sum_{i=1}^{m} \alpha_i P(A_i))(\sum_{j=1}^{n} \beta_j P(B_i)) \\ &= E(X)E(Y) \end{aligned}$$

3 Sei $0 \leq X, (X_n)_{n=1}^{\infty}$ Folge von nicht negativen einfachen Zufallsvariablen $X_n =$

$f_n(X)$ der Form (ii) und sei $0 \leq Y, (Y_n)_{n=1}^{\infty}$ Folge von nicht negativen einfachen Zufallsvariablen $Y_n = g_n(Y)$ der Form (ii) mit $x_n \leq X_{n+1}, Y_n \leq Y_{n+1}, n \in \mathbb{N}$

$$\begin{aligned} X &= \sup_{n \leftarrow \infty} X_n = \lim_{n\to\infty} X_n \\ Y &= \sup_{n \leftarrow \infty} Y_n = \lim_{n\to\infty} Y_n, \end{aligned}$$

X_n, Y_n sind unabhängig.
Dann ist $XY = \lim n \to \infty X_n Y_n, X_n Y_n \leq X_{n+1} Y_{n+1}$.

$$\begin{aligned} E(XY) &= \lim n \to \infty E(X_n Y_n) \\ &\overset{(ii)}{=} \lim n \to \infty E(X_n) E(Y_n) \\ &= \lim n \to \infty E(X_n) \lim n \to \infty E(Y_n) \\ &= E(X)E(Y) \end{aligned}$$

4 $X = X^+ - X^-, Y = Y^+ - Y^-$
$E(X^+) < \infty, E(X^-) < \infty, E(Y^+) < \infty, E(Y^-) < \infty$
$XY = X^+Y^+ - X^+Y^- - X^-Y^+ + X^-Y^-$
X^+, X^-, Y^+, Y^- sind unabhängig.

$$\begin{aligned} E(XY) &= E(X^+)E(Y^+) - E(X^+)E(Y^-) - E(X^-)E(Y^+) + E(X^-)E(Y^-) \\ &= (E(X^+) - E(X^-))(E(Y^-) - E(Y^))) \\ &= E(X)E(Y) \end{aligned}$$

Also: Seien X, Y reelle undabhängige Zufallsvariablen.

1 $X, Y \geq 0 \Rightarrow E(XY) = E(X)E(Y)$

2 $E(|X|) < \infty$
$E(|Y|) < \infty \Rightarrow E(|XY|) < \infty$ und $E(XY) = E(X)E(Y)$

Folgerung: Sind X, Y **unabhängig** mit $E(X^2) < \infty, E(Y^2) < \infty$, so ist

$$Cov(X, Y) = E(XY) - E(X)E(Y) = 0$$

Bemerkung : Aus $Cov(X, Y) = 0$ folgt im Allgemeinen **nicht** X, Y unabhängig.

Beispiel dazu:

$$X_1, X_2 \sim \mathfrak{B}(1, \frac{1}{2})$$

$$\begin{aligned} X &= X_1 + X_2 \sim \mathfrak{B}(2, \frac{1}{2}) \\ Y &= X_1 - X_2 \end{aligned}$$

$$\begin{aligned} Cov(X,Y) &= Cov(X_1 + X_2, X_1 - X_2) \\ &= E((X_1 + X_2)(X_1 - X_2)) - (E(X_1) + E(X_2)) \underbrace{(E(X_1) - E(X_2))}_{=0} \\ &\overset{3.Bin.Formel}{=} E(X_1^2) - E(X_1X_2) + E(X_1X_2) - E(X_2^2) \\ &= E(X_1^2) - E(X_2^2) \\ &= 0 \end{aligned}$$

$P(X = 1, Y = 0) = 0 \neq \frac{1}{4} = P(X = 1)P(Y = 0)$
$P(X = 1) = frac12$
$P(Y = 0) = frac12$

19.5 Ungleichungen

19.5.1 Satz

Seien X, Y reelle Zufallsvariablen mit $E(X^2) < \infty$, $E(Y^2) < \infty$. Dann gilt:

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

(Cauchy-Schwarz-Ungleichung)

19.5.2 Beweis

$$\begin{aligned} 0 &\leq (|X| - a|Y|)^2 \\ 0 &\leq E((|X| - a|Y|)^2) \\ &= E(X^2) - 2aE(|X||Y|) + a^2E(Y^2) \\ &:= f(a) \end{aligned}$$

ohne Einschränkung: $E(Y^2) > 0$ (Sonst ist $P(Y=0) = 1$, also auch $E(|X||Y|) = 0$

$$
\begin{aligned}
f'(a) = 2aE(Y^2) - 2E(|X||Y|) \overset{!}{=} 0 \\
\Leftrightarrow a = a^* = fracE(|X||Y|)E(Y^2) \\
0 \leq f(a^*) &= E(X^2) - 2\frac{E(|XY|)^2}{E(Y^2)} + \frac{E(|XY|)^2}{E(Y^2)^2}E(Y^2) \\
&= E(X^2) - \frac{E(|XY|)^2}{E(Y^2)} \\
\Rightarrow E(|XY|)^2 \leq E(X^2)E(Y^2) \\
\Rightarrow E(|XY|) \leq \sqrt{E(X^2)E(Y^2)}
\end{aligned}
$$

19.5.3 Folgerung

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

19.5.4 Folgerung

$$
\begin{aligned}
\underbrace{|([X - E(X)][Y - E(Y)])|} &\leq \sqrt{E([X - E(X)]^2)E([Y - E(Y)]^2)} \\
|Cov(X,Y)| &\leq \sqrt{Var(X)Var(Y)}
\end{aligned}
$$

19.5.5 Definition

Seien X, Y relle Zufallsvariablen mit $E(X^2) < \infty$, $E(Y^2) < \infty$. Dann heiÃ?t

$$\rho(X,Y) = \frac{Cov(X,Y}{\sqrt{Var(X)Var(Y)}}$$

Korrelationskoeffizient von X und Y.
$\rho(X,Y) := 0$, falls $Var(X) = 0$ oder $Var(Y) = 0$, also nur wenn eine Zufallsvariable mit der Wahrscheinlichkeit 1 eine Konstante ist.
Einsicht: $-1 \leq \rho(X,Y) \leq 1$.

19.5.6 Neuer Sinnabschnitt

Seien X, Y reelle Zufallsvariablen mit $E(X) = 0 = E(Y)$ und $E(X^2) = 1 = E(Y^2)$.

$$\begin{aligned}
\inf_{a,b\in\mathbb{R}} E((X - (\underbrace{aY+b}_{\text{affin lineare Funktion von Y}}))^2) &= \inf_{a,b\in\mathbb{R}}[E(X^2 + a^2Y^2 + 2abY + b^2 - 2X(aY + \\
&= \inf_{a,b\in\mathbb{R}}[1 + a^2 - 1 + 0 + b^2 - 2aE(XY)] \\
&\overset{b=0}{=} \inf_{a\in\mathbb{R}}(1 + a^2 - 2aE(XY)) \\
&\overset{a=a^*=E(XY)}{=} 1 + (E(XY))^2 - 2(E(XY))^2 \\
&= 1 - (E(XY))^2 \\
&= 1 - \rho(XY)^2
\end{aligned}$$

Allgemein: X, Y reelle Zufallsvariablen. $E(X^2) < \infty, Var(X) > 0, Var(Y) > 0$.
Standardisieren:

$$\begin{aligned}
X^* &= \frac{X - E(X)}{\sqrt{Var(X)}} \\
Y^* &= \frac{Y - E(Y)}{\sqrt{Var(Y)}}
\end{aligned}$$

Dann: $E(X^*) = E(Y^*) = 0, E((X^*)^2) = E((Y^*)^2) = 1$

$$\inf_{a,b\in\mathbb{R}} E((X - (aY + b))^2) = ?$$

$$\begin{aligned}
E((X - (aY + b))^2) &= E((\sqrt{Var(X)}X^* + E(X) - (a\sqrt{Var(Y)}Y^* + aE(Y) \\
&= Var(X)E(X^* - (\frac{a\sqrt{Var(Y)}Y^* + aE(Y) + b - E(X}{\sqrt{Var(X)}}
\end{aligned}$$

$$\frac{a\sqrt{Var(Y)}}{\sqrt{Var(X)}} = a^*, \frac{aE(Y) + b - E(X)}{\sqrt{Var(X)}} = b^*$$

$$= Var(X)E(X^* - (a^*Y^* + b^*))$$

$$\begin{aligned}
\Rightarrow \inf_{a,b\in\mathbb{R}} E((X - (aY + b))^2) &= Var(X)(1 - \underbrace{\rho(X^*, Y^*)^2}_{= E(X^*Y^*)^2}) \\
&= Var(X)(1 - \rho(X, Y)^2)
\end{aligned}$$

Die bestmögliche Approximation von X durch eine affin lineare Funktion Y im quadratischen Mittel ist bestensfalls dieses Ergebnis.
Der Korrelationskeoffizient ist also ein Maß für den affin linearen Zusammenhang zwischen X und Y.
Eine perfekte Vorhersage für X ist möglich, falls Y bekannt (die Werte von X, Y müssen auf einer Gerade liegen).
Nur dieser Fall wird durch den Korrelationskoeffizienten abgedeckt, es ist aber auch so etwas möglich:
$\Rightarrow$ Der Korrelationskoeffizient hat nicht so viel Bedeutung, wie man oft glaubt.
Beispiel für eine Ünsinnskorrelation": Anzahl der Störche, Zahl der Geburten :)
In der Presse sind "Korrelationen"meist (empirische) Schätzwerte.

19.5.7 Satz:

Sei X eine reelle Zufallsvariable mit $E(X^2) < \infty$. Für jedes $\varepsilon > 0$ gilt:

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} Var(X)$$

(Chebyshevsche Ungleichung)

19.5.8 Beweis:

$$\begin{aligned} Var(X) &= E([X - E(X)]^2) \\ &\geq E([X - E(X)]^2 I(|X - E(X)| \geq \varepsilon)) \\ &\geq E(\varepsilon^2 I(|X - E(X)| \geq \varepsilon) \\ &= \varepsilon P(|X - E(X)| \geq \varepsilon) \end{aligned}$$

19.5.9 Satz:

Sei X eine reelle Zufallsvariable, $f : [0, \infty) \to [0, \infty)$ monoton wachsend, $\varepsilon > 0, f(\varepsilon) > 0$. Dann gilt

$$P(|X| \geq \varepsilon) \leq \frac{1}{f(\varepsilon)} E(f(|X|))$$

(Markovsche Ungleichung).

19.5.10 Beweis:

$$\begin{aligned} E(f(|X|)) &\geq E(\underbrace{f(|X|)}_{\geq \varepsilon} I(|X| \geq \varepsilon) \\ &\geq f(\varepsilon) P(|X| \geq \varepsilon) \end{aligned}$$

Auf diese Weise lassen sich noch unzählige weitere Ungleichungen beweisen.
Mit Chebyshev kann man das Gesetz der groÃ?en Zahlen beweisen, das Grundlage für so gut wie alle Simulationsverfahren ist.

19.5.11 Bemerkung zur letzten Vorlesung

Fehlerkorrektur. $X \sim \mathfrak{R}(a,b)$ Rechteckverteilung.

$$Var(X) = \frac{1}{b-a}\int\limits_a^1 (x - \frac{a+b}{2})^2 dx$$

(Der Bruch vor dem Integral wurde vergessen).

Teil VI

Gesetze der großen Zahlen

20 Einführung

Es handelt sich um Grenzwertsätze der Stochastik. Wir müssen uns hierfür erst einmal Konvergenzbegriffe überlegen.

Seien $X, X_n,\ n \in \mathbb{N}$ neue, d-dim. Zufallsvariablen auf einem W-Raum $(\Omega, \mathfrak{A}, P)$. Die Folge der X_n konvergiert "p-fast-sicher" gegen X, wenn gilt

$$P\left(\lim_{n\to\infty} X_n = X\right) = P\left(\bigcap_{l\in\mathbb{N}>0} \bigcup_{m\in\mathbb{N}} \bigcap_{n=m}^{\infty} \{|X_n - X| \leq \frac{1}{n}\}\right) = 1 \tag{97}$$

Hierbei sei $|\cdot|$ die euklidische Norm)

Schreibweise:

$$X_n \to X,\ P\text{-f.s.}$$

Die Folge der X_n konvergiert stochastisch gegen X, wenn gilt

$$\forall \epsilon > 0 : \lim_{n\to\infty} P\left(|X_n - X| \geq \epsilon\right) = 0 \tag{98}$$

Schreibweise

$$X_n \xrightarrow{P} X$$

21 Bemerkungen

21.1 a

$$X_n \to X P\text{-f.s.}$$

$= \iff$ Es gibt eine P-Nullmenge $N \in \mathfrak{A}$, (d.h. $P(N) = 0$) mit der Eig., dass

$$\begin{aligned}
\lim_{n\to\infty} X_n(\omega) = X(\omega) \forall\omega \in N^c \\
` \Rightarrow \text{“Wähle } N \\
= \left(\bigcap_{l\in\mathbb{N}>0} \bigcup_{m\in\mathbb{N}} \bigcap_{n=m}^{\infty} \{|X_n - X| \leq \frac{1}{n}\} \right)^c \\
= \bigcup_{l\in\mathbb{N}>0} \bigcap_{m\in\mathbb{N}} \bigcup_{n=m}^{\infty} \{|X_n - X| \leq \frac{1}{n}\} \\
` \Leftarrow \text{“} N^c \\
\subset \bigcap_{l\in\mathbb{N}>0} \bigcup_{m\in\mathbb{N}} \bigcap_{n=m}^{\infty} \{|X_n - X| \leq \frac{1}{n}\}
\end{aligned} \tag{99}$$

21.2 b

$X_n \to X$ P-f.s. $\Rightarrow X_n \xrightarrow{P} X$.
Für $l \in \mathbb{N}$ ist

$$\begin{aligned}
\bigcap_{m\in\mathbb{N}} \bigcup_{n=m}^{\infty} \{|X_n - X| \leq \frac{1}{n}\} \\
\subset \bigcup_{k\in\mathbb{N}>0} \bigcap_{m\in\mathbb{N}} \bigcup_{n=m}^{\infty} \{|X_n - X| \leq \frac{1}{k}\} \\
\Rightarrow = P\left(\bigcap_{m\in\mathbb{N}} \bigcup_{n=m}^{\infty} \{|X_n - X| \leq \frac{1}{n}\} \right) \\
= 0
\end{aligned} \tag{100}$$

Also

$$\begin{aligned} 0 &= P\left(\bigcap_{m\in\mathbb{N}}\bigcup_{n=m}^{\infty}\{|X_n - X| \leq \frac{1}{n}\}\right) \\ &= \lim_{m\to\infty} P\left(P\left(\bigcup_{n=m}^{\infty}\{|X_n - X| \leq \frac{1}{n}\}\right)\right) \\ &\geq \limsup_{m\to\infty} P\left(|X_n - X| \leq \frac{1}{n}\right) \end{aligned} \tag{101}$$

22 Chebyshevsches schwaches Gesetz der großen Zahlen

Seien X_n unabhängige, reelle zufallsvariablen, je mit derselben Verteilung. Es sei $E(|X_1|^2) < \infty$. Sei $X_n = \frac{1}{n} \cdot \sum_{j=1}^{n} X_j,\ n \in \mathbb{N}, \mu = E(X_1)$. Dann gilt

$$X_n \xrightarrow{P} \mu \tag{102}$$

22.1 Beweis

Mit der Chebyshevschen Ungleichung und $\mu = E(X_n)$.

$$P(|X_n - \mu|) \leq \frac{1}{\epsilon^2} Var(X_n) = \frac{1}{\epsilon^2 \cdot n^2}\sum_{j=1}^{n} Var(X_j) = \frac{1}{\epsilon \cdot n} Var(X_1)\ (n, \overset{\rightarrow}{\ldots}, \infty)\ 0 \tag{103}$$

23 Starkes Gesetz der großen Zahlen

von Kolmogorov, 1933.

Seien X_n unabhängige reelle Zufallsvariablen, ke mit derselben Verteilung mit endlichen Erwartungswert μ. Dann gilt

$$X_n \to \mu, P\text{-f.s.}$$

23.1 Anwendung

Berechnung von Integralen mittels Monte-Carlo-Simulation

$$g : R^k \to \mathbb{R}, Q \leq g \leq 1$$

Es sei

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} g(x_1, \ldots, x_n) \cdot dx_1 \cdots dx_k = \mu < \infty$$

Es ser $f > 0$ die Dichte eines k-dim. Zufallsvektors X. Dann ist

$$\mu = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \underbrace{\frac{g(x_1, \dots, x_k}{f(x_1, \dots, x_k)}}_{=h(x_1,\dots,x_k)} \cdot f(x_1, \dots, x_k) \cdot dx_1 \cdots dx_k = E(h(X))$$

Seien $X_1, X_2, \dots$ unabh., je mit derselben Vtl. wie X.
$Y_n = h(X_n), n \in \mathbb{N}$ die X_n sind unabh. je mit derselben Vtl. mit $\mu = E(Y_1)$.

$$\frac{1}{1} \cdot Y_n = \frac{1}{n} \cdot \sum_{j=1}^{n} h(X_j \to \mu, P - f.s.$$

Verschaffe Beobachtung $x1, \dots, x_n$ von $X_1, \dots, X_n$ und fasse
$\frac{1}{n} \cdot \sum_{j=1}^{n} h(x_j)$ als Approximation für $\mu = \int \int g(x_1, \dots, x_k) dx_1 \cdots dx_k$ auf

23.2 Anwendung

Einzelexperiment, bei dem ein bestimmtes Ereignis A mit der W-keit $p \in [0,1]$ eintritt, und unbeschränkt of unabhängig wiederholt. Es sei B_n das Ereignis, dass A bei der n-ten Wiederholung eintritt, $n \in \mathbb{N}$. Die B_n sind unabhängig, es ist $P(B_n) = p$ für jedes $n \in \mathbb{N}$. Sei $X_n = I_{B_n}$. Die X_n sind unabhängig, je mit derselben $Suett - B(1,p)$-Verteilung.

Dan gilt

$$\frac{1}{n}\sum_{j=1}^{n} X_j = \underbrace{\frac{1}{n}\sum_{j=1}^{n} I_{B_j}}_{\substack{\text{relative Häufigkeit}\\ \text{des Eintretens von } A\\ \text{bei den ersten } n \text{ Wdh.}}} \longrightarrow \mu = E(X_1) = p \ P\text{-fast sicher} \tag{104}$$

23.3 Die empirische Vertelungsfunktion

Seien $X_1, X_2 \dots$ unabhängige, reelle Zufallsvariablen, je mit derselben Verteilungs(funktion)

$$F(x) = P(X_1 \leq x), x \in \mathbb{R}$$
$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} I(X_j \leq x) \longrightarrow E(I(X_1 \leq x)) = P(X_1 \leq x) = F(x), P\text{-fast sicher } \forall x \in \mathbb{R}$$

Die Funktion $F_n(x)$ heißt empirische Verteilungsfunktion der $X_1, \dots, X_n$.
F_n ist die Verteilungsfunktion der empirischen Verteilung $P_n = \frac{1}{n}\sum_{j=1}^{n} \delta_{X_j}$, denn:

$$\begin{aligned} P_n((-\infty, x]) &= \frac{1}{n}\sum_{j=1}^{n} \underbrace{\delta_{X_j} \cdot ((-\infty, x])}_{=\begin{cases} 1, & X_j \leq x \\ 0, & X_j > x \end{cases}} \\ &= \frac{1}{n}\sum_{j=1}^{n} I(X_j \leq x) = F_n(x), x \in \mathbb{R} \end{aligned} \tag{105}$$

Für beobachtete Werte $x_1, \dots, x_n$ von $X_1, \dots, X_n$ heißt

$$\frac{1}{n}\sum_{j=1}^{n} I(x_j \leq x), \text{die beobachtete empirische Verteilungsfkt. der } X_1, \dots, X_n$$

die beobachtete empirische Verteilung der $X_1, \dots, X_n$.

24 Satz von Glivenko-Cantelli

$$P\left(\lim_{n\to\infty} \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| = 0\right) = 1 \tag{106}$$

Teil VII

Verteilungen von positiven ganzzahligen Zufallsvariablen

Sei für $n \in$ $X_n \sim \mathfrak{B}(n, p)$. Es gelte

$$\lim_{n \to n} n \cdot p_n = \lambda \in (0, \infty)$$

Für jedes $k \in \mathbb{N}$ ist im Falle $n \geq k$

$$\begin{aligned} P(X_n = k) &= \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} \\ &= \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \cdot p_n^k \cdot (1 - p_n)^{n-k} \\ &= \frac{1}{k!} \cdot (n \cdot p_n) \cdot (n-1) \cdot p_n \cdots (n-k+1) \cdot p_n \cdot (1 - \frac{n \cdot p_n}{n})^{n-k} \\ &\longrightarrow \frac{\lambda^k}{k!} \cdot \exp -\lambda \end{aligned} \tag{107}$$

25 Definition

Eine $\mathbb{N}_0$-wertige Zufallsvariable X, deren Verteilung festgesetzt ist durch

$$P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

heißt Poisson-verteilt mit dem Parameter $\lambda > 0$.

Schreibweise:

$$X \sim \mathfrak{P}(\lambda)$$

26 Das Gesetz der seltenen Ereignisse

Es gilt daher bei $X_n \sim \mathfrak{B}(n, p_n)$ mit $\lim_{n\to\infty} n \cdot p_n = \lambda \in (0, \infty)$ und $X \sim \mathfrak{P}(\lambda)$

$$\lim_{n \to \infty} P(X_n = k) = P(X = k), k \in \mathbb{N}_0 \tag{108}$$

27 Definition Erzeugende Funktionen

Sei X $\mathbb{N}_0$-verteilt Zufallsvariable. Dann heißt

$$f_x(t) = E(t^X) = \sum_{k=0}^{\infty} t^k \cdot P(X = k), \ |t| \leq 1 \tag{109}$$

erzeugende Funktion von X (genauer: der Verteilung von X).

Teil VIII
Das Gesetz der großen Zahlen

28 Verteilungskonvrgenz, Fourier-Transformierte, der zentrale Grenzwertsatz

Seien $X, X^{(n)}$ d-dimensionale Zufallsvektoren, $X = (X_1, \ldots, X_d)^T$, $X^{(n)} = (X_1^{(n)}, \ldots, X_d^{(n)})^T$.
Die Folge $(X^{(n)})_{n\in\mathbb{N}}$ konvergiert in Verteilung gegen X, in jedem $X^{(n)} \overset{v}{\to} X$, wenn gilt:
Für jede beschränkte, stetige Funktion $f : \mathbb{R}^d \to \mathbb{R}$,gilt

$$\lim_{n\to\infty} E(f(X^{(n)})) = E(f(X))$$

28.1 Satz

Seien F, F_n die Verteilungsfunktionen von $X, X^{(n)}, n \in \mathbb{N}$. Dann gilt $X^{(n)} \overset{v}{\to} X$ genau dann, wenn $\lim_{n\to\infty} F_n(x) = F(x)$ für jede $x \in \mathbb{R}^d$, x Stetigkeitsstelle von F.

28.1.1 Konkreter Anwendungsfall

Seien X, X_n relle Zufallsvariablen mit $X_n \overset{v}{\to} X$. Dann

$$\begin{aligned} F_n(x) &= P(X_n \leq x) \\ (n \to \infty) &\downarrow \\ F(x) &= P(X \leq x), \forall x \in \mathbb{R} \end{aligned} \tag{110}$$

Dann gilt auch für jede $a, b \in \mathbb{R}, a < b$

$$\begin{aligned} P(a < X_n \leq b) &= P(X_n \leq b) - P(X_n \leq a) \\ &\downarrow \\ &\overset{(n\to\infty)}{\to} P(X \leq b) - P(X \leq a) = P(a < X \leq b) \end{aligned} \tag{111}$$

Also:

Dann gilt auch

$$\begin{aligned} P(a \leq X_n \leq b) &\geq P(a < X_n \leq b) \\ \liminf_{n\to\infty} P(a \leq X_n \leq b &\geq P(a \leq X \leq b) \\ P(a \leq X_n \leq b) &\leq P(a - \epsilon < X_n \leq b),\ \forall \epsilon > 0 \\ \limsup_{n\to\infty} P(a \leq X_n \leq b) &\leq P(a - \epsilon < X \leq b),\ \forall \epsilon > 0 \end{aligned} \tag{113}$$

Also

$$P(a \leq X \leq b) \leq \liminf_{n\to\infty} P(a \leq X_n \leq b) \leq \limsup_{n\to\infty} P(a \leq X_n \leq b) \leq P(a \leq X \leq b)$$

29 Fourier-Transformierte

Siehe Datei im Stud.Ip „Eigenschaften von Fourier-Transformierten“

30 Der zentrale Grenzwertsatz von Lindeberg-Levy

Seien $X_1, X_2, \ldots$ unabhängig je mit derselben Verteilung mit $E(X_1) = 0, Var(X_1) = 1$. Dann gilt:

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j \xrightarrow{v} X, X \sim N(0,1)$$

30.1 Beweis

Die Fourier-Transformierte $N(0,1)$-Verteilung ist

$$\begin{aligned}\phi_X(z) &= \int_{-\infty}^{+\infty} \omega(z \cdot x) \frac{1}{\sqrt{2 \cdot \pi}} \cdot \exp -\frac{x^2}{2} \cdot dx + \underbrace{\int_{-\infty}^{+\infty} \sin{(z \cdot x)} \frac{1}{\sqrt{2 \cdot \pi}} \cdot \exp -\frac{x^2}{2} \cdot dx}_{=0} \\ &= \exp -\frac{z^2}{2}\end{aligned} \tag{114}$$

30.2 Anwendung

Seien $X_1, X_2, \ldots$ unabhängig, je mit derselben Verteilung mit $E(X_1^2) < \infty$. Sei $\mu = E(X_1)$, sei $\sigma^2 = Var(X_1) > 0$. Dann gilt

$$\frac{1}{\sqrt{n}} \cdot \sum_{j=1}^{n} \frac{X_i - \mu}{\sqrt{\sigma^2}} = \underbrace{\frac{\sum_{j=1}^{n} X_j - n \cdot \mu}{\sqrt{\mu \cdot \sigma^2}}}_{= \frac{\sqrt{n}\left(\sum_{j=1}^{n} X_j - \mu\right)}{\sqrt{\sigma^2}}} \xrightarrow{v} N(0,1) \tag{115}$$

Für $-\infty < a < b < +\infty$ gilt:

$$\begin{aligned}P(a \leq \sum_{j=1}^{n} X_j \leq b) &= P\left(\frac{a - n \cdot \mu}{\sqrt{n \cdot \sigma^2}} \leq \frac{\sum_{j=1}^{n} X_j - n \cdot \mu}{\sqrt{n \cdot \sigma^2}} \leq \frac{b - n \cdot \mu}{\sqrt{n \cdot \sigma^2}}\right) \\ &\approx \Phi(\frac{b - n \cdot \mu}{\sqrt{n \cdot \sigma^2}}) - \Phi(\frac{a - n \cdot \mu}{\sqrt{n \cdot \sigma^2}})\end{aligned} \tag{116}$$

30.3 Beispiel

$X_n \sim \mathfrak{B}(n, p$ sei eine binomialverteilte Zufallsvariable. Dann

$$\lim_{n\to\infty} P\left(a \leq \frac{X_n - n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right) = \Phi(b) - \Phi(a)$$

Angenähert gilt

$$P\left(a \leq \frac{X_n - n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}} \leq b\right) \approx \Phi(b) - \Phi(a)$$

$$\begin{aligned} P(a \leq X_n \leq b) &= P\left(\frac{a - n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}} \leq \frac{X_n - n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}} \leq \frac{b - n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right) \\ &\approx \Phi\left(\frac{b - n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right) - \Phi\left(\frac{a - n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right) \end{aligned}$$

30.4 Beispiel

$X_1, X_2, \ldots$ sind unabhängige Zufallsvariablen, je $\mathfrak{P}(\lambda)$ verteilt. $\lambda = E(X_j) = Var(X_j)$. Dann

$$\lim_{n\to\infty} P\left(a \leq \frac{\sum_{j=1}^{n} X_j - n\cdot\lambda}{\sqrt{n\cdot\lambda}} \leq b\right) = \Phi(b) - \Phi(a)$$

Hier $\sum_{j=1}^{n} X_j \sim \mathfrak{P}(n\cdot\lambda)$. Also

$$P\left(a \leq \frac{\sum_{j=1}^{n} X_j - n\cdot\lambda}{\sqrt{n\cdot\lambda}} \leq b\right) \approx \Phi\left(\frac{b - n\cdot\lambda}{\sqrt{n\cdot\lambda}}\right) - \Phi\left(\frac{a - n\cdot\lambda}{\sqrt{n\cdot\lambda}}\right)$$

$$P(a \leq S_n \leq b) = P\left(\frac{a - n\cdot\lambda}{\sqrt{n\cdot\lambda}} \leq \frac{S_n - n\cdot\lambda}{\sqrt{n\cdot\lambda}} \leq \frac{b - n\cdot\lambda}{\sqrt{n\cdot\lambda}}\right) \approx \Phi\left(\frac{b - n\cdot\lambda}{\sqrt{n\cdot\lambda}}\right) - \Phi\left(\frac{a - n\cdot\lambda}{\sqrt{n\cdot\lambda}}\right)$$

31 Grenzwertsätze für ausgewählte Verteilungen

31.1 Lokaler zentraler Grenzwertsatz für die Poisson-Verteilung

Tatsächlich lässt sich zeigen bei $X_\lambda \sim \mathfrak{P}(\lambda)$:

$$P(X_\lambda = k) \sim \frac{1}{\sqrt{2\cdot\pi\cdot\lambda}} \cdot \exp\left(-\frac{1}{2}\cdot\frac{(k-\lambda)^2}{\lambda}\right) \tag{117}$$

für $\lambda \to n, |k-\lambda| = \text{const} \cdot \sqrt{\lambda}$ das heißt

$$\sup_{\substack{k\in\mathbb{N} \\ |k-|\leq \text{const}\cdot\sqrt{\lambda}}} \left| \frac{P(X_k = k)}{\frac{1}{\sqrt{2\cdot\pi\cdot\lambda}} \cdot \exp\left(-0.5 \cdot \frac{(k-\lambda)^2}{\lambda}\right)} - 1 \right| \to 0 \tag{118}$$

31.2 Lokaler zentraler Grenzwertsatz für die Binomial-Verteilung

Analaog für $X_n \sim \mathfrak{P}(n,p)$

$$\binom{n}{k} p^k (1-p)^{n-k} = P(X_n = k) \sim \frac{1}{\sqrt{2\cdot\pi\cdot n\cdot p\cdot(1-p)}} \cdot \exp\left(-\frac{1}{2}\right) \cdot \frac{(k-np)^2}{np(1-p)}, \; n \to \infty \tag{119}$$

32 Satz von der stetigen Abbildung

Auch genannt „Continuous Mapping Theorem“.
Seien $X_1, \ldots, X_n, \; n \in \mathbb{N}$ d-dim. Zufallsvariablen mit $X_n \xrightarrow{v} X$. $h : \mathbb{R}^d \to \mathbb{R}^p$ stetig. Dann gilt

$$h(X_n) \xrightarrow{v} h(X) \tag{120}$$

32.1 Beweis

Sei $f : \mathbb{R}^d \to \mathbb{R}$ stetig und beschränkt. Dann

$$E(f(h(X_n))) = E(\underbrace{f \cdot h}_{\text{stetig, beschränkt}}(X_n)) \longrightarrow E(f \cdot h(X)) = E(f(h(X_n)))$$

32.2 Satz

Seien X, X_n, Y_n reelle Zufallsvarieblen, $c \in \mathbb{R}$ konstant. Es gelte $X_n \xrightarrow{v} X, Y_n \xrightarrow{P} c$ (d.h.: Für jedes $\epsilon > 0$ gilt $\lim_{n\to\infty} P(|Y_n - c| > \epsilon) = 0$) . Dann

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{v} \begin{pmatrix} X \\ c \end{pmatrix} \tag{121}$$

32.3 Anwendung

$X_n \xrightarrow{v} X, Y_n \xrightarrow{P} c$. Dann

1 $X_n + Y_n \xrightarrow{v} X + c$

2 $X_n \cdot Y_n \xrightarrow{v} X \cdot c$

3 $\frac{X_n}{Y_n} \xrightarrow{v} \frac{X}{c}, \; c \neq 0$

33 Satz: Fehlerfortpflanzungsgesetz

X_n reelle Zufallsvariablen, $a_n > 0, n \in \mathbb{N}$ reelle Zahlen mit $a_n \to \infty$. Es gelte

$$a_n \cdot (X_n - \mu) \xrightarrow{v} N(0, \sigma^2)$$

Es sei $f : \mathbb{R} \to \mathbb{R}$ differenzierbar in μ mit $f'(\mu) \neq 0$. Dann gilt

$$a_n \cdot (f(X_n) - f(\mu)) \xrightarrow{v} N(0, \sigma^2 \cdot f'(\mu)^2) \tag{122}$$

Allgemein $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty, -\infty\}, f|_U$ reellwertig für eine Umgebung U von μ. Dann gilt

$$\lim_{n\to\infty} P(a_n(f(X_n) - f(a)) \leq x = P(Z \leq x), Z \sim (N(0, \sigma^2 f'(\mu)^2)$$

Schreibweise: $a_n(f(X_n) - f(\mu)) \xrightarrow{v} N(0, \sigma^2 f'(\mu)^2)$

$X_\lambda \sim \mathfrak{P}(\lambda)$

$2(\sqrt{X_\lambda} - \sqrt{\lambda}) \xrightarrow{v} N(0,1)$ „Wurzeltransformation bei Poisson-Verteilung“

33.1 Anwendung

$X_1, X_2, \ldots$ unabhängig, jemit derselben Poisson-Verteilung $X_i \sim \mathfrak{P}(\lambda),\ \lambda > 0$.

$$\sqrt{n} \cdot \left(\frac{1}{n} \cdot \sum_{j=1}^{n} X_j - \lambda \right) \xrightarrow{\nu} N(0, \lambda) \tag{123}$$

Sei $f(x) = \sqrt{x}$ für $x \geq 0$ für differenzierbar in $\lambda > 0$ und $f'(\lambda) = \frac{1}{2}\lambda^{-1/2}$.
Damit gilt

$$\sqrt{n} \cdot \left(\sqrt{\frac{1}{n} \cdot \sum_{j=1}^{n} X_j} - \sqrt{\lambda} \right) \xrightarrow{\nu} N(0, \lambda \cdot \frac{1}{4} \cdot \lambda^{-1}) = N(0, \frac{1}{4}) \tag{124}$$

Allgemein gilt

$$2 \cdot \sqrt{n} \cdot \left(\frac{1}{n} \cdot \sum_{j=1}^{n} X_j - \lambda \right) \xrightarrow{\nu} N(0, 1) \tag{125}$$

33.2 Spezialfall

Sei $\lambda = 1$. Dann ist

$$S_N := \sum_{j=1}^{n} X_j \sim \mathfrak{P}(n)$$

Es ist dann

$$2 \cdot \left(\sqrt{S_n} - \sqrt{n} \right) \xrightarrow{\nu} N(0, 1)$$

Allgemein soll

$$2 \cdot \left(\sqrt{X_j} - \sqrt{\lambda} \right) \xrightarrow{\nu} N(0, 1),\ \forall \lambda \to \infty$$

gezeigt werden. Zeige zunächst

$$\frac{X_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow{\nu} N(0, 1)$$

Beweis mit Fourier-Transformierte (ausgelassen).

33.3 Wurzeltransformation für die Poisson-Verteilung

Es gilt also

$$\sqrt{\lambda} \cdot \left(\frac{X_\lambda}{\lambda} - 1 \right) \xrightarrow{\nu} N(0, 1) \tag{126}$$

Mit dem Fehlerfortpflanzungsgesetz folgt

$$\sqrt{\lambda} \cdot \left(\sqrt{\frac{X_\lambda}{\lambda}} - \sqrt{1} \right) \xrightarrow{\nu} N(0, \frac{1}{4}) \tag{127}$$

so dass gilt

$$\left(\sqrt{X_\lambda} - \sqrt{\lambda}\right) \xrightarrow{\nu} N(0, \frac{1}{4}) \tag{128}$$

oder alternativ

$$2 \cdot \left(\sqrt{X_\lambda} - \sqrt{\lambda}\right) \xrightarrow{\nu} N(0, 1) \tag{129}$$

Diese Wurzeltransformation ist eine varianzstabilisierende Transformation für die Poissonverteilung.

33.4 Bemerkung

Als Ergänzung zu den lokalen zentralen Grenzwertsätzen für die Binomial-, bzw. Poisson-Verteilung lässt sich auch ein lokaler zentraler Grenzwertsatz für Hypergeometrische Verteilung formulieren:
Für $n \in \mathbb{N}$ sei $X_n \sim \mathfrak{h}(a_n, r_n, n)$ mit $Var(X_n) = n \cdot \frac{r_n}{a_n} \cdot \frac{a_n - r_n}{a_n} \cdot a_n - na_n - 1 \to \infty,\ (n \to \infty)$. Dann gilt der „Zentrale Grenzwertsatz für die Hypergeometrsiche Verteilung in lokaler Form“:

$$P(X_n = k) \sim \frac{1}{\sqrt{2 \cdot \pi \cdot Var(X_n)}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(k - E(X_n))^2}{Var(X_n)}\right) \tag{130}$$

für $|k - E(X_n)| \leq const\ \sqrt{Var(X_n)},\ n \to \infty$.

Damit gilt für $-\infty < a \leq b < +\infty$ der „Zentrale Grenzwertsatzfür die Hypergeometrische Verteilung in kumulativer Form“:

$$P\left(a \leq \frac{X_n - E(X_n)}{\sqrt{Var(X_n)}} \leq b\right) \to \frac{1}{\sqrt{2 \cdot \pi}} \int_a^b \exp\left(-\frac{1}{2} \cdot x^2\right) \cdot dx \tag{131}$$

Teil IX
Anhang

Literaturangaben

[Hüb09] Gerhard Hübner. *Stochastik: Eine anwendungsorientierte Einführung für Informatiker, Ingenieure und Mathematiker*. Vieweg+Teubner5. auflage. , 2009. ISBN 9783834807175.
URL http://amazon.de/o/ASIN/3834807176/

[Hen09] Norbert Henze. *Stochastik für Einsteiger: Eine Einführung in die faszinierende Welt des Zufalls. Mit über 220 Übungsaufgaben und Lösungen*. Vieweg+Teubner8., erweiterte auflage. , 2009. ISBN 9783834808158.
URL http://amazon.de/o/ASIN/3834808156/