Einführung in die Stochastik für Mathematiker - SS 03
Prof. Dr. M. Schaefer, RWTH Aachen
Definitionen und Sätze

Erstellt von Lars Otten
lars.otten@kullen.rwth-aachen.de

5. September 2003

Diese Aufzeichnungen stammen **nicht** vom Lehrstuhl und wurden auch **nicht** von diesem kontrolliert. Dementsprechend können noch Fehler enthalten sein, für die ich keine Verantwortung übernehme!

Unter `http://bart.kullen.rwth-aachen.de/~lotten/stochastik/` ist immer die aktuellste Version dieses Dokuments erhältlich (zumindest in der näheren Zukunft).

Erstellt mit LaTeX.

§1 Diskrete Wahrscheinlichkeitsräume

Definition 1.1

Sei $\Omega \neq \emptyset$ abzählbar:

(a) Eine Abbildung $p : \Omega \to \mathbb{R}$ heißt Zähldichte, falls

 (i) $p(\omega) \geq 0 \quad \forall \omega \in \Omega$

 (ii) $\sum\limits_{\omega \in \Omega} p(\omega) = 1$

(b) Ist $p : \Omega \to \mathbb{R}$ eine Zähldichte, so heißt das Paar (Ω, p) ein diskreter Wahrscheinlichkeitsraum.

(c) Ist (Ω, p) ein diskreter Wahrscheinlichkeitsraum, so heißt die Abbildung

$$P : \begin{cases} \mathfrak{P}(\Omega) \to \mathbb{R} \\ A \mapsto \sum_{\omega \in A} p(\omega) \end{cases}$$

die zu p gehörige Wahrscheinlichkeitsverteilung.

Satz 1.1

Sei (Ω, p) ein diskreter Wahrscheinlichkeitsraum und P die zu p gehörige Wahrscheinlichkeitsverteilung. Dann gilt:

(a) $P(A) \geq 0 \quad \forall\, A \in \mathfrak{P}(\Omega)$ („P ist nicht-negativ“)

(b) $P(\Omega) = 1$ („P ist normiert“)

(c) Sind $A_n, n \in \mathbb{N}$, paarweise disjunkt, so ist

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \qquad (\text{„P } \sigma\text{-additiv“})$$

Satz 1.2

Sei $\Omega \neq \emptyset$ abzählbar und $P : \mathfrak{P}(\Omega) \to \mathbb{R}$ mit (a)–(c) aus Satz 1.1; ist dann

$$p : \begin{cases} \Omega \to \mathbb{R} \\ \omega \mapsto P(\{\omega\}) \end{cases},$$

so ist p Zähldichte und P die zu p gehörige Wahrscheinlichkeitsverteilung.

Bemerkung 1.1

(a) Ist $P : \mathfrak{P}(\Omega) \to \mathbb{R}$ nicht-negativ, normiert und σ-additiv, so gilt:

  (i) Ist I höchstens abzählbar unendlich und sind $A_i, i \in I$ paarweise disjunkt, so ist $P\left(\sum_{i\in I} A_i\right) = \sum_{i\in I} (P(A_i))$.

  (ii) Satz 1.1 und 1.2 zeigen, dass ein diskreter Wahrscheinlichkeitsraum auch als Paar (Ω, P) definiert werden kann, wobei P nicht-negativ, normiert und σ-additiv ist.

Satz 1.3

Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum, dann gilt:

(a) $P(A^C) = 1 - P(A) \quad \forall\, A \in \mathfrak{P}(\Omega)$

(b) $P(A) \leq P(B) \quad \forall\, A, B \in \mathfrak{P}(\Omega), A \subset B$

(c) $P(A) = P(AB) + P(AB^C) \quad \forall\, A, B \in \mathfrak{P}(\Omega)$

(d) $P(A \cup B) = P(A) + P(B) - P(AB) \quad \forall A, B \in \mathfrak{P}(\Omega)$

(e) $P\left(\bigcup\limits_{n=1}^{\infty} A_n\right) \leq \sum\limits_{n=1}^{\infty} P(A_n) \quad \forall\, A_n \in \mathfrak{P}(\Omega), n \in \mathbb{N}$

(f) $P\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = \lim\limits_{n\to\infty} P(A_n) \quad \forall A_n \in \mathfrak{P}(\Omega), n \in \mathbb{N}, A_n \subset A_{n+1} \,\forall\, n \in \mathbb{N}$

(g) $P\left(\bigcap\limits_{n=1}^{\infty} A_n\right) = \lim\limits_{n\to\infty} P(A_n) \quad \forall A_n \in \mathfrak{P}(\Omega), n \in \mathbb{N}, A_n \supset A_{n+1} \,\forall\, n \in \mathbb{N}$

§2 Grundbegriffe der Kombinatorik

Zwei wichtige Prinzipien der Kombinatorik

(1) $M = \sum\limits_{i=1}^{r} M_i \quad \Rightarrow \quad |M| = \sum\limits_{i=1}^{r} |M_i|$

(2) $M_1, M_2 : \quad |M_1| = |M_2| \quad \Leftrightarrow \quad \exists\, f : M_1 \to M_2$ bijektiv

Hilfsmittel

(1) Fakultätsfunktion (erklärt für $n \in \mathbb{N} \cup \{0\} =: N_0$)

$$n! := \begin{cases} 1 & , n = 0 \\ \prod_{j=1}^{n} j & , n \geq 1 \end{cases}$$

(2) Binomialkoeffizient (erklärt für $k \in \mathbb{Z}$ und $n \in \mathbb{N}$)

$$\binom{n}{k} := \begin{cases} \frac{n!}{k! \cdot (n-k)!} & , 0 \leq k \leq n \\ 0 & , \text{sonst} \end{cases}$$

Abkürzungen

(1) $M = \{1, \ldots, n\}$, $n, r \in \mathbb{N}$

 (i) $\mathcal{G}_r(M) := \{(x_1, \ldots, x_r) \mid x_i \in M,\ 1 \leq i \leq r\}$
 („Menge der geordneten r–Tupel mit Komponenten aus M **mit** Wiederholung“)

 (ii) $\mathcal{G}_r^o(M) := \{(x_1, \ldots, x_r) \mid x_1, \ldots, x_r \text{ paarw. verschieden}\}$
 („Menge der geordneten r–Tupel mit Komponenten aus M **ohne** Wiederholung“)

 (iii) $\mathcal{U}_r(M) := \{(x_1, \ldots, x_r) \in \mathcal{G}_r(M) \mid x_1 \leq x_2 \leq \ldots \leq x_r\}$
 („Menge der ungeordneten r–Tupel mit Komponenten aus M **mit** Wiederholung“)

 (iv) $\mathcal{U}_r^o(M) := \{(x_1, \ldots, x_r) \in \mathcal{G}_r(M) \mid x_1 < x_2 < \ldots < x_r\}$
 („Menge der ungeordneten r–Tupel mit Komponenten aus M **ohne** Wiederholung“)

(2) $M_1, M_2 \neq \emptyset, \quad \mathcal{F}(M_1, M_2) := \{f : M_1 \to M_2 \mid \text{f injektiv}\}$

(3) $M \neq \emptyset$ beliebig, $k \in \mathbb{Z}$

 (i) $Per(M) := \mathcal{F}(M, M)$

 (ii) $\mathfrak{P}_k(M) := \{A \in \mathfrak{P}(M) \mid |A| = k\}$

Satz 2.1

(a) $n \in \mathbb{N}, M_1, M_2$ beliebige Mengen: $|M_1| = |M_2| = n$
$\Rightarrow \quad |\mathcal{F}(M_1, M_2)| = n!$

(b) $k \in \mathbb{Z}, n \in \mathbb{N}$, M beliebige Menge: $|M| = n$. Dann gilt:

 (i) $|Per(M)| = n!$

 (ii) $|\mathfrak{P}_k(M)| = \binom{n}{k}$

 (iii) $|\mathfrak{P}(M)| = 2^n$

Satz 2.2

(a) $r \in \mathbb{N}, M_1, \ldots, M_r \neq \emptyset, |M_i| < \infty, 1 \leq i \leq r$
$\Rightarrow |M_1 \times \ldots \times M_r| = \prod_{i=1}^{r} |M_i|$

(b) $r, n \in \mathbb{N}, M = \{1, \ldots, n\}$ Dann gilt:

(i) $|\mathcal{G}_r(M)| = n^r$

(ii) $|\mathcal{G}_r^o(M)| = r!\binom{n}{r}$

(iii) $|\mathcal{U}_r(M)| = \binom{n+r-1}{r}$

(iv) $|\mathcal{U}_r^o(M)| = \binom{n}{r}$

Satz 2.3 („Hypergeometrische Verteilung“)

Seien $r, s \in \mathbb{N}_0,\ r+s > 0,\ N := r+s,\ n \in \mathbb{N},\ M_1 := \{1, \ldots, r\},\ M_2 := \{r+1, \ldots, N\},\ M = M_1 + M_2$,
$\Omega = \mathcal{U}_n^o(M) = \mathfrak{P}_n(M)$, P Laplace-Verteilung auf Ω, $k \in \{1, \ldots, n\}$, $A_k \hat{=}$ interessantes Ereignis, $A_k := \{\omega \in \Omega \mid |\omega \cap M_1| = k\}$. Dann gilt:

$$P(A_k) = \frac{\binom{r}{k} \cdot \binom{N-r}{n-k}}{\binom{N}{n}}$$

Beachte: $P(A_k) = 0$, falls $k > r$ oder $n - k > s = N - r$.

Satz 2.4 („Siebformel“)

Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum, $n \in \mathbb{N},\ A_1, \ldots, A_n \subset \Omega,\ M = \{1, \ldots, n\}$

$$\Rightarrow P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \cdot \sum_{I \in \mathfrak{P}_k(M)} P\left(\bigcap_{i \in I} A_i\right)$$

§3 Bedingte Wahrscheinlichkeit, stochastische Unabhängigkeit

Definition 3.1

Sei (Ω, P) diskreter Wahrscheinlichkeitsraum, $A, B \in \mathfrak{P}(\Omega)$ mit $P(B) > 0$.
Dann heißt

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

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

Satz 3.1

Sei $n \in \mathbb{N},\ n \geq 2,\ (\Omega, P)$ diskreter Wahrscheinlichkeitsraum, $A_1, \ldots, A_n \in \mathfrak{P}(\Omega), P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$.
Dann gilt:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P\left(A_i \middle| \bigcap_{\rho=1}^{i-1} A_\rho\right)$$

Satz 3.2

Sei (Ω, P) diskreter Wahrscheinlichkeitsraum, dann gilt:

(a) Ist $B \in \mathfrak{P}(\Omega)$ mit $P(B) > 0$, so ist

$$P_B := \begin{cases} \mathfrak{P}(\Omega) \to \mathbb{R} \\ A \mapsto P(A|B) \end{cases}$$

eine diskrete Wahrscheinlichkeitsverteilung.

(b) („Satz von der totalen Wahrscheinlichkeit“)
Sind $n \in \mathbb{N},\ A, B_1, \ldots, B_n \in \mathfrak{P}(\Omega)$ mit $P(B_i) > 0$, $1 \leq i \leq n$ und $\Omega = \sum_{i=1}^{n} B_i$ $(\Rightarrow B_i \cap B_j = \emptyset \,\forall\, i, j)$, so folgt:

(i) $P(A) = \sum_{i=1}^{n} P(B_i) \cdot P(A|B_i)$ („Satz von der totalen Zerlegung“)

(ii) („Formel von Bayes“) Ist auch $P(A) > 0$, so gilt

$$P(B_i|A) = \frac{P(B_i) \cdot P(A|B_i)}{\sum_{j=1}^{n} P(A|B_j) \cdot P(B_j)}$$

Definition 3.2

Sei (Ω, P) diskreter Wahrscheinlichkeitsraum, $I \neq \emptyset$ (*beliebige* Index-Menge), $A_i \in \mathfrak{P}(\Omega)$, $i \in I$. Dann heißt $\{A_i \mid i \in I\}$ stochastisch unabhängig (bzgl. P), falls für *jede endliche* Teilmenge $J \subset I$ gilt:

$$P\Bigl(\bigcap_{j \in J} A_j\Bigr) = \prod_{j \in J} P(A_j)$$

Satz 3.3

Sei (Ω, P) diskreter Wahrscheinlichkeitsraum, $I \neq \emptyset,\ A_i \in \mathfrak{P}(\Omega)$, $i \in I,\ \{A_i \mid i \in I\}$ stochastisch unabhängig. Dann gilt:

(a) Ist $A^* \in \mathfrak{P}(\Omega)$ mit $P(A^*) \in \{0, 1\}$, so ist auch $\{A_i \mid i \in I\} \cup \{A^*\}$ stochastisch unabhängig.

(b) Ist $J \subset I$, $J \neq \emptyset$, so ist auch $\{A_j \mid j \in J\}$ stochastisch unabhängig.

(c) Sind $B_i \in \mathfrak{P}(\Omega)$ mit $B_i \in \{A_i, A_i^C\}$, $i \in I$, so ist (auch) $\{B_i \mid i \in I\}$ stochastisch unabhängig.

Satz 3.4 („Modellierung von unabhängigen Zufalls-Experimenten“)

Sei $n \in \mathbb{N}$ (O.B.d.A. sei $n \geq 2$). Weiter seien (Ω_i, P_i) diskrete Wahrscheinlichkeitsräume mit zugehöriger Zähldichte p, $1 \leq i \leq n$, $\Omega := \Omega_1 \times \ldots \times \Omega_n$. Außerdem sei die Abb. $\pi_i : \Omega \to \Omega_i$ definiert durch $(\omega_1, \ldots, \omega_n) \mapsto \omega_i$ („Projektion von Ω auf die i-te Komponente“). Dann sei

$$\pi_i^{-1} : \begin{cases} \mathfrak{P}(\Omega_i) \to \mathfrak{P}(\Omega) \\ B \mapsto \{(\omega_1, \ldots, \omega_n) \in \Omega \mid \pi_i(\omega_1, \ldots, \omega_n) \in B\} \end{cases}$$

Dann gilt:

(a) $p : \Omega \to \mathbb{R},\ (\omega_1, \ldots, \omega_n) \mapsto \prod\limits_{i=1}^{n} p_i(\omega_i)$ ist Zähldichte auf Ω.

(b) $p \leftrightarrow P,\ B_i \in \mathfrak{P}(\Omega_i),\ 1 \le i \le n,\ A_i := \pi_i^{-1}(B_i),\ 1 \le i \le n.$

   (i) $P(B_1 \times \ldots \times B_n) = \prod\limits_{i=1}^{n} P_i(B_i)$

   (ii) $\{A_1, \ldots, A_n\}$ stochastisch unabhängig bzgl. P mit $P(A_i) = P_i(B_i)$.

(c) Ist Q diskrete Wahrscheinlichkeitsverteilung auf Ω, so dass $\pi_1^{-1}(B_1), \ldots, \pi_n^{-1}(B_n)$ bzgl. Q stochastisch unabhängig sind mit $Q\left(\pi_i^{-1}(B_i)\right) = P_i(B_i),\ 1 \le i \le n$, so folgt $Q = P$ (aus (a)). („P ist eindeutig bestimmt“)

Beispiel 3.6(a) (Bernoulli-Experiment, Binomialveteilung)

Geg. Zufalls-Experiment mit 2 möglichen Ergebnissen: <u>0 oder 1</u> $\quad 1 \to \vartheta \in [0,1]$, $0 \to 1-\vartheta$

$n \in \mathbb{N}$ (Experiment wird n-mal unabhängig durchgeführt)

$\Omega_i = \{0,1\},\quad p_i(1) = \vartheta = 1 - p_i(0),\quad 1 \le i \le n,\quad \Omega = \Omega_1 \times \ldots \times \Omega_n = \{0,1\}^n$ ($\to$ Produktraum)

$p(\omega) = p(\omega_1, \ldots, \omega_n) = \prod\limits_{i=1}^{n} p_i(\omega_i) = \vartheta^{t(\omega)} \cdot (1-\vartheta)^{n-t(\omega)}$ mit $t(\omega) = \sum\limits_{i=1}^{n} \omega_i$

$$\boxed{\text{„}(\Omega, P)\text{ Bernoulli-Experiment“}}$$

$k \in \{0, 1, \ldots, n\},\ p \leftrightarrow P,\ A_k = \{\omega \in \Omega \mid t(\omega) = k\}$ (bei diesen n Durchführungen tritt „1“ genau k-mal ein). $P(A_k) = \sum\limits_{\omega \in A_k} p(\omega) = \binom{n}{k} \vartheta^k (1-\vartheta)^{n-k}$

Sei nun $\Omega' = \{0, 1, \ldots, n\},\quad p'(k) := \binom{n}{k} \vartheta^k (1-\vartheta)^{n-k},\ k \in \Omega'$

$\Rightarrow p'(k) \ge 0$ <u>und</u> $\sum\limits_{k \in \Omega'} p'(k) = (\vartheta + 1 - \vartheta)^n = 1^n = 1$

<u>Also:</u> (Ω', p') diskreter Wahrscheinlichkeitsraum: $p' \leftrightarrow P'$,
P' heißt <u>Binomialverteilung</u> mit Parametern n und ϑ.

Beispiel 3.6(b) (Negative Binomialverteilung mit Parametern $r \in \mathbb{N}$, $\vartheta \in (0,1]$)

Seien $r, n \in \mathbb{N},\ n \ge r,\ \vartheta \in (0,1]$. Betrachte „Bernoulli-Experiment“ mit n und ϑ. $\Omega = \{0,1\}^n$,
$p(\omega_1, \ldots, \omega_n) = \vartheta^{t(\omega)} (1-\vartheta)^{n-t(\omega)}$, wobei $t(\omega) := t(\omega_1, \ldots, \omega_n) = \sum\limits_{i=1}^{n} \omega_i$.

$A_r^{(n)} := \{(\omega_1, \ldots, \omega_n) \in \Omega \mid t(\omega) = r,\ \omega_n = 1\}$

$p \leftrightarrow P$: $\quad P(A_r^{(n)}) = \binom{n-1}{r-1} \vartheta^r (1-\vartheta)^{n-r}$

$k := n - r$, dann $n \to k + r$

$P(A_r^{(k+r)}) = \binom{k+r-1}{r-1} \vartheta^r (1-\vartheta)^k = \binom{k+r-1}{k} \vartheta^r (1-\vartheta)^k$

$k \in \mathbb{N}_0 =: \Omega',\quad p'(k) := \vartheta^r \binom{k+r-1}{k} (1-\vartheta)^k$

Dann gilt: (1) $p'(k) \ge 0\ \forall\, k \in \Omega'$ und (2) $\sum\limits_{k=0}^{\infty} p'(k) = 1$

$p' \leftrightarrow P'$: P' heißt <u>negative Binomialverteilung</u> mit Parametern r und ϑ.

Interpretation: Ein 0-1-wertiges Experiment werde so lange unabhängig durchgeführt, bis genau r „Einsen“ eintreten. Die Wahrscheinlichkeit von 1 sei ϑ. Dann „ist“ die Wahrscheinlichkeit, dass genau $k+r$ Ausführungen nötig sind, gleich $\vartheta^r\binom{k+r-1}{k}(1-\vartheta)^k$.

Spezialfall: Für $r=1$ erhält man die geometrische Verteilung.

Satz 3.5 (Poisson'scher Grenzwertsatz)

Sei $\lambda > 0$, $\vartheta_n := \frac{\lambda}{n}$, $n \in \mathbb{N}$, dann gilt:

$$\lim_{n\to\infty} \binom{n}{k} \vartheta_n^k (1-\vartheta_n)^{n-k} = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad \forall\, k \in \mathbb{N}_0$$

Bemerkung 3.2

Seien $\Omega := \mathbb{N}_0$, $p(\omega) = e^{-\lambda} \cdot \frac{\lambda^\omega}{\omega!}$, $\omega \in \Omega$

Dann gilt: (1) $p(\omega) \geq 0 \;\forall\, \omega \in \Omega$ und (2) $\sum\limits_{k=0}^{\infty} p(k) = 1$

Also: p Zähldichte auf Ω, $p \leftrightarrow P$ heißt Poissonverteilung mit Parameter λ.

§4 Zufallsvariablen

Definition 4.1

Seien $\Omega, \mathfrak{X} \neq \emptyset$, $f : \Omega \to \mathfrak{X}$, dann heißt

$$f^{-1} : \mathfrak{P}(\mathfrak{X}) \to \mathfrak{P}(\Omega),\; B \mapsto \{\omega \in \Omega \mid f(\omega) \in B\}$$

die „Mengen-Inverse“ von f.

Andere Schreibweisen: $\omega \in \Omega : \{f(\omega) \in B\} = \{f \in B\} := f^{-1}(B)$
$x \in \mathfrak{X} : \{f = x\} := f^{-1}(\{x\})$

Lemma 4.1

Seien $\Omega, \mathfrak{X} \neq \emptyset$, $f\colon \Omega \to \mathfrak{X}$, $I \neq \emptyset$ (beliebige Index-Menge), $B_i \in \mathfrak{P}(\mathfrak{X})$, $i \in I$, dann gilt:

(a) $f^{-1}\Big(\bigcup\limits_{i\in I} B_i\Big) = \bigcup\limits_{i\in I} f^{-1}(B_i)$

(b) $f^{-1}\Big(\bigcap\limits_{i\in I} B_i\Big) = \bigcap\limits_{i\in I} f^{-1}(B_i)$ („f^{-1} ist operationstreu“)

(c) $f^{-1}(B^c) = \left(f^{-1}(B)\right)^c$, $B \in \mathfrak{P}(\mathfrak{X})$

Definition 4.2

Seien (Ω, P) diskreter Wahrscheinlichkeitsraum; $\mathfrak{X} \neq \emptyset$ und $X : \Omega \to \mathfrak{X}$

(a) X heißt Zufallsvariable auf (Ω, P) mit Werten in $\mathfrak{X}$.

(b) $P^X\colon \mathfrak{P}(\mathfrak{X}) \to \mathbb{R}$, $B \mapsto P\big(X^{-1}(B)\big)$ heißt die Verteilung von X (bzgl. P).

Lemma 4.2

Seien (Ω, P) diskreter Wahrscheinlichkeitsraum, $\mathfrak{X} \neq \emptyset$, $X : \Omega \to \mathfrak{X}$ Dann gilt:

(a) P^X ist nicht-negativ (d.h. $P^X(B) \geq 0 \; \forall \, B$)

(b) P^X ist normiert (d.h. $P^X(\mathfrak{X}) = 1$)

(c) P^X ist σ-additiv (d.h. sind $B_n \in \mathfrak{P}(\mathfrak{X}), n \in \mathbb{N}$, paarweise disjunkt, so ist $P^X\Big(\sum\limits_{n=1}^{\infty} B_n\Big) = \sum\limits_{n=1}^{\infty} P^X(B_n)$)

(d) Es gibt abzählbare, nicht-leere Teilmenge $X_0 \subset \mathfrak{X}$ mit

(i) $P^X(X_0) = 1$

(ii) $P^X(B) = \sum_{x \in X_0} q(x) \mathbb{1}_B(x) \quad \forall \, B \in \mathfrak{P}(\mathfrak{X})$,
dabei ist $q(x) := P^X\big(\{x\}\big) = P\big(\{X = x\}\big)$, $x \in \mathfrak{X}$
(Sprechweise: P^X ist eine quasi-diskrete Wahrscheinlichkeitsverteilung auf $\mathfrak{X}$)

Bemerkung 4.1

Ist X eine Zufallsvariable aus (Ω, P) mit Werten in $\mathbb{R}^1$ bzw. $\mathbb{R}^n$, so nennt man X auch eine reelle bzw. n-dim. Zufallsvariable.

Lemma 4.3

Seien (Ω, P) diskreter Wahrscheinlichkeitsraum, $\mathfrak{X}_1, \mathfrak{X}_2 \neq \emptyset$ (beliebig), X_1, X_2 Zufallsvariablen auf (Ω, P) mit Werten in $\mathfrak{X}_1, \mathfrak{X}_2$.
$\tilde{X} : \Omega \to \mathfrak{X}_1 \times \mathfrak{X}_2,\ \omega \mapsto \big(X_1(\omega), X_2(\omega)\big), \quad \mathfrak{X}_1^* := X_1(\Omega),\ \mathfrak{X}_2^* := X_2(\Omega)$ *(! $\mathfrak{X}_1^*$ und $\mathfrak{X}_2^*$ abzählbar !)*
Dann gilt:

(a) (i) $P^{X_1}(B_1) = \sum\limits_{x_2 \in \mathfrak{X}_2^*} P^{\tilde{X}}(B_1 \times \{x_2\}), \quad B_1 \in \mathfrak{P}(\mathfrak{X}_1)$

(ii) $P^{X_2}(B_2) = \sum\limits_{x_1 \in \mathfrak{X}_1^*} P^{\tilde{X}}(\{x_1\} \times B_2), \quad B_2 \in \mathfrak{P}(\mathfrak{X}_2)$

(b) $\mathfrak{X}_1 = \mathfrak{X}_2 = \mathbb{R},\ Z = X_1 + X_2 \quad \Rightarrow \quad P^Z\big(\{z\}\big) = \sum\limits_{x_1 \in \mathfrak{X}_1^*} P^{\tilde{X}}\big(\{x_1, z - x_1\}\big)$

$$\Big(= \sum_{x_1 \in \mathfrak{X}_1^*} P\big(\{X_1 = x_1\} \cap \{X_2 = z - x_1\}\big)\Big) = \sum_{x_2 \in \mathfrak{X}_2^*} P\big(\{X_1 = z - x_2\} \cap \{X_2 = x_2\}\big)$$

Definition 4.3

Sei (Ω, P) diskreter Wahrscheinlichkeitsraum; $I \neq \emptyset$ (beliebig), $\mathfrak{X}_i \neq \emptyset$, X_i Zufallsvariable auf (Ω, P) mit Werten in $\mathfrak{X}_i$, $i \in I$.
Dann heißt $(X_i)_{i \in I}$ stochastisch unabhängig, wenn für jede Wahl von $B_i \in \mathfrak{P}(\mathfrak{X}_i)$, $i \in I$, die Ereignisse $A_i := \{X_i \in B_i\}$, $i \in I$, stochastisch unabhängig sind.

Bemerkung 4.2

(a) Unabhängigkeit von Ereignissen ist äquivalent mit der Unabhängigkeit jeder endlichen Auswahl der gegebenen Ereignisse. Daher ist die Unabhängigkeit von $(X_i)_{i \in I}$ äquivalent mit der Unabhängigkeit von $(X_j)_{j \in J}$ für jede endliche Menge $J \subset I$.

(b) Im folgenden betrachten wir daher zunächst den Fall $I = \{1, \dots, n\}$, $n \in \mathbb{N}$, $n \geq 2$.

Satz 4.1

(Ω, P) diskreter Wahrscheinlichkeitsraum; $n \in \mathbb{N}$, $\mathfrak{X}_1, \ldots, \mathfrak{X}_n \neq \emptyset$, $X_1, \ldots, X_n$ Zufallsvariablen auf (Ω, P) mit Werten in $\mathfrak{X}_1, \ldots$ bzw. $\mathfrak{X}_n$. Dann sind die folgenden Aussagen äquivalent:

(a) $X_1, \ldots, X_n$ sind stochastisch unabhängig.

(b) $P\Big(\bigcap\limits_{i=1}^{n} \{X_i = x_i\}\Big) = \prod\limits_{i=1}^{n} P\big(\{X_i = x_i\}\big) \quad \forall\, \tilde{x} = (x_1, \ldots, x_n) \in \mathfrak{X}_1 \times \ldots \times \mathfrak{X}_n$

(c) $P\Big(\bigcap\limits_{i=1}^{n} \{X_i \in B_i\}\Big) = \prod\limits_{i=1}^{n} P\big(\{X_i \in B_i\}\big) \quad \forall\, B_i = (B_1, \ldots, B_n) \in \mathfrak{P}(\mathfrak{X}_1) \times \ldots \times \mathfrak{P}(\mathfrak{X}_n)$

Lemma 4.4

$X_1, X_2, \ldots, X_n$ seien stochastisch unabhängige Zufallsvariablen auf (Ω, P) mit Werten in $\mathfrak{X}_1, \mathfrak{X}_2, \ldots$ bzw. $\mathfrak{X}_n$. Ferner seien $\mathfrak{X}_i^* := X_i(\Omega), 1 \leq i \leq n$. Außerdem seien $\mathfrak{X}^* := \mathfrak{X}_1^* \times \ldots \times \mathfrak{X}_n^*$ (abzählbar) und $\mathfrak{X} := \mathfrak{X}_1 \times \ldots \times \mathfrak{X}_n$. Dann gilt:

(a) Ist $B \in \mathfrak{P}(\mathfrak{X})$, $B^* \in B \cap \mathfrak{X}^*$, so folgt:
$P\big(\,\{(X_1, \ldots, X_n) \in B\}\,\big) = P\big(\,\{(X_1, \ldots, X_n) \in B^*\}\,\big) = \sum\limits_{(x_1^*, \ldots, x_n^*) \in B^*} \prod_{i=1}^{n} P\big(\{X_i = x_i^*\}\big)$

(b) Ist $J \subset \{1, \ldots, n\}, J \neq \emptyset$, so sind auch $(X_j)_{j \in J}$ stochastisch unabhängig.

Satz 4.2

Mit den Bezeichnungen und Voraussetzungen aus Lemma 4.4 seien ferner $k \in \{1, \ldots, n-1\}$, $\mathfrak{I}_1 := \mathfrak{X}_1 \times \ldots \times \mathfrak{X}_k$, $\mathfrak{I}_2 := \mathfrak{X}_{k+1} \times \ldots \times \mathfrak{X}_n$, $\mathfrak{Z}_1, \mathfrak{Z}_2 \neq \emptyset$ und $g_1 := \mathfrak{I}_1 \to \mathfrak{Z}_1$, $g_2 := \mathfrak{I}_2 \to \mathfrak{Z}_2$; definiert man dann $Z_1 \colon \Omega \to \mathfrak{Z}_1$, $\omega \mapsto g_1(X_1(\omega), \ldots, X_k(\omega))$ und $Z_2 \colon \Omega \to \mathfrak{Z}_2$, $\omega \mapsto g_2(X_{k+1}(\omega), \ldots, X_n(\omega))$, so sind Z_1, Z_2 stochastisch unabhängig.

§5 Erwartungswert, Varianz, Kovarianz, Korrelation

Definition 5.1

Ist X *reelle* Zufallsvariable auf (Ω, P) mit

$$\sum_{\omega \in \Omega} |X(\omega)| \cdot P\left(\{\omega\}\right) < \infty \qquad \text{(absolute Konvergenz)}$$

so heißt

$$E(X) = EX := \sum_{\omega \in \Omega} X(\omega) \cdot P\left(\{\omega\}\right)$$

der Erwartungswert von X (bzgl. P).

Bemerkung 5.1

Ist $|\Omega| < \infty$, so ist die Bedingung $\sum\limits_{\omega \in \Omega} |X(\omega)| \cdot P\left(\{\omega\}\right) < \infty$ überflüssig.
Ist dagegen $|\Omega| = \infty$, so wird durch diese Bedingung sichergestellt, dass $\sum\limits_{\omega \in \Omega} X(\omega) \cdot P\left(\{\omega\}\right)$ unabhängig von der Abzählung von Ω ist.

Satz 5.1 („Transformationssatz“)

X Zufallsvariable auf (Ω, P) mit Werten in $\mathfrak{X} \neq \emptyset$. Weiter seien $g\colon \mathfrak{X} \to \mathbb{R}$; $Y : \Omega \to \mathbb{R}$, $\omega \mapsto g(X(\omega))$ (also $Y = g \circ X$). Ist dann $\mathfrak{X}^*$ abzählbar mit $\mathfrak{X}^* \supset X(\Omega)$, so gilt:

(a) $\sum\limits_{\omega\in\Omega} |Y(\omega)| \cdot P(\{\omega\}) < \infty \quad \Leftrightarrow \quad \sum\limits_{x\in\mathfrak{X}^*} |g(x)| \cdot P(\{X = x\}) < \infty$

(b) $\sum\limits_{\omega\in\Omega} |Y(\omega)| \cdot P(\{\omega\}) < \infty \quad \Rightarrow \quad EY = \sum\limits_{x\in\mathfrak{X}^*} g(x) \cdot P(\{X = x\})$

Bemerkung 5.2

(a) Man beachte, dass X nicht notwendig reellwertig ist.

(b) Ist X reelle Zufallsvariable mit existierendem Erwartungswert, so folgt aus Satz 5.2(b):

$$EX = \sum_{x\in\mathfrak{X}^*} x \cdot P(\{X = x\}) \qquad \text{(Wähle } \mathfrak{X}^* := X(\Omega),\ g\colon \mathbb{R} \to \mathbb{R},\ x \mapsto x\text{)}$$

Das heißt der Erwartungswert von X *hängt nur ab von der Verteilung von* X.

Beispiel 5.2

(a) X sei $b(n, \vartheta)$-verteilt ($n \in \mathbb{N},\ \vartheta \in [0,1]$)

$$\Rightarrow \quad EX = n\vartheta$$

(b) X sei hypergeometrisch verteilt mit Parametern $r, s, n \in \mathbb{N}_0,\ n \geq 1,\ r + s \geq 1$

$$\Rightarrow \quad EX = n\,\frac{r}{r+s}$$

(c) X sei negativ binomialverteilt mit Parameter $r \in \mathbb{N},\ \vartheta \in (0,1)$

$$\Rightarrow \quad EX = r\,\frac{1-\vartheta}{\vartheta}$$

(d) X sei Poisson-verteilt mit Parameter $\lambda > 0$

$$\Rightarrow \quad EX = \lambda$$

Satz 5.2

X, Y seien reelle Zufallsvariablen auf (Ω, P) mit existierenden Erwartungswerten. Dann gilt:

(a) Ist $\alpha \in \mathbb{R}$, so existiert der Erwartungswert von αX und es gilt $E(\alpha X) = \alpha EX$.

(b) Der Erwartungswert von $X + Y$ exisitiert und es gilt $E(X + Y) = EX + EY$.

(c) Sind X, Y stochastisch unabhängig, so existiert der Erwartungswert von $X \cdot Y$ und es gilt $E(X \cdot Y) = EX \cdot EY$

Lemma 5.1

X, Y reelle Zufallsvariablen auf (Ω, P), so dass der Erwartungswert von X^2 und Y^2 existiert. Dann gilt:

(a) Es existieren Erwartungswerte von X und Y.

(b) Bezeichnet a bzw. b den Erwartungswert von X bzw. Y, so existieren Erwartungswerte von $(X-a)^2, (Y-b)^2$ und $(X-a)(Y-b)$.

(c) Es existiert der Erwartungswert von $(X+Y)^2$.

Definition 5.2

X, Y seien reelle Zufallsvariablen auf (Ω, P) mit $EX^2 < \infty$, $EY^2 < \infty$, ferner sei $a := EX, b := EY$

(a) $VarX := E(X-a)^2$ heißt die Varianz von X

(b) $Cov(X, Y) := E(X-a)(Y-b)$ heißt die Kovarianz von X und Y

(c) Ist $VarX > 0$ und $VarY > 0$, so heißt

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

die Korrelation zwischen X und Y.
Gilt $\rho(X, Y) = 0$, also $Cov(X, Y) = 0$, so nennt man X und Y unkorreliert.

Satz 5.3

X, Y reelle Zufallsvariablen auf (Ω, P) mit $EX^2 < \infty$, $EY^2 < \infty$; $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. Dann gilt:

(a) $VarX = EX^2 - (EX)^2$

(b) $Cov(X, Y) = E(XY) - EX \cdot EY$

(c) $Var(\alpha X + \beta) = \alpha^2 VarX$

(d) $Cov(\alpha X + \beta, \gamma Y + \delta) = \alpha\gamma\, Cov(X, Y)$

(e) $Var(X+Y) = VarX + VarY + 2\, Cov(X, Y)$

(f) X, Y stochastisch unabhängig $\Rightarrow$ $Var(X+Y) = VarX + VarY$

Satz 5.4

X, Y seien reelle Zufallsvariablen auf (Ω, P) mit $EX^2 < \infty$ und $EY^2 < \infty$, dann gilt:

(a) („Cauchy-Schwarz-Ungleichung“)

$$\big(E(X \cdot Y)\big)^2 \leq EX^2 \cdot EY^2$$

(b) $|Cov(X, Y)|^2 \leq VarX \cdot VarY$

(c) Ist die Korrelation zwischen X und Y erklärt, so ist $|\rho(X, Y)| \leq 1$

Beispiel 5.4

(a) $x_1, \ldots, x_n \in \mathbb{R}$, X Zufallsvariable mit $P(X = x_i) = \frac{1}{n}$. Dann ist $EX = \frac{1}{n} \sum\limits_{i=1}^{n} x_i =: \bar{x}$ und es gilt:

$$VarX = \frac{1}{n} \left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right)$$

(b) $n \in \mathbb{N}, \vartheta \in [0,1]$; X sei $b(n, \vartheta)$-verteilt. Dann ist $EX = n\vartheta$ und es gilt:

$$VarX = n\vartheta(1 - \vartheta)$$

(c) X sei hypergeometrisch verteilt mit $r, s, n;\ N := r + s$. Dann ist $EX = n\frac{r}{N}$ und es gilt:

$$VarX = n \frac{r}{N} \frac{s}{N} \left(\frac{N - n}{N - 1} \right)$$

(d) $r \in \mathbb{N}$, $\vartheta \in (0,1)$; X sei negativ binomialverteilt mit r und ϑ. Dann ist $EX = \frac{r(1-\vartheta)}{\vartheta}$ und es gilt:

$$VarX = \frac{r(1 - \vartheta)}{\vartheta^2}$$

(e) Sei $\lambda > 0$, X Poisson-verteilt mit Parameter λ. Dann ist $EX = \lambda$ und es gilt:

$$VarX = \lambda$$

Satz 5.5 („Tschebyscheff-Ungleichgung“)

Sei X reelle Zufallsvariable auf (Ω, P) mit $EX^2 < \infty$

$$\Rightarrow \quad P\big(\{|X - EX| \geq \varepsilon\} \big) \leq \frac{VarX}{\varepsilon^2} \quad \forall \varepsilon > 0$$

Satz 5.6 („Markow-Ungleichung“)

Sei Z reelle Zufallsvariable auf (Ω, P); $g\colon [0, +\infty) \to [0, +\infty)$ mit
(1) g isoton (d.h. monoton wachsend)
(2) $Eg(|Z|) < \infty$

$$\Rightarrow \quad g(\varepsilon) \cdot P\big(\{|Z| \geq \varepsilon\}\big) \leq Eg(|Z|)$$

Bemerkung 5.3

(a) Die Ungleichung in (5.5) ist relativ grob, aber von theoretischem Nutzen.

(b) Der Nutzen wird deutlich im nächsten Satz:

Satz 5.7 („Schwaches Gesetz der großen Zahlen“)

Seien $n \in \mathbb{N}$, $X_1, \ldots, X_n$ reelle stochastisch unabhängige Zufallsvariablen auf (Ω, P). Alle $X_1, \ldots, X_n$ mögen dieselbe Verteilung besitzen ($X_1, \ldots, X_n$ seien stochastisch unabhängige „identisch verteilte“ Zufallsvariablen). Ist dann $EX^2 < \infty$, $a := EX_1$, $\sigma^2 := VarX_1$, $\bar{X}_{(n)} := \frac{1}{n}\sum\limits_{i=1}^{n} X_i$, so gilt:

$$P\big(\{|\bar{X}_{(n)} - a| \geq \varepsilon\}\big) \leq \frac{\sigma^2}{\varepsilon^2 n} \quad \forall \varepsilon > 0$$

§6 Diskrete Zufallsvariablen auf allgemeinen Räumen

Definition 6.1

Seien $\Omega \neq \emptyset$, $\mathfrak{A} \subset \mathfrak{P}(\Omega)$, dann heißt $\mathfrak{A}$ $\underline{\sigma\text{-Algebra in } \Omega}$ falls folgende Bedingungen erfüllt sind:

(1) $\Omega \in \mathfrak{A}$

(2) Für alle $A \in \mathfrak{A}$ gilt auch $A^c \in \mathfrak{A}$

(3) Für jede Folge $A_n \in \mathfrak{A},\ n \in \mathbb{N}$, gilt: $\left(\bigcup\limits_{i=1}^{\infty} A_n\right) \in \mathfrak{A}$

Satz 6.1

$\Omega \neq \emptyset$, $\mathfrak{A}$ σ-Algebra in Ω, dann gilt:

(a) $\emptyset \in \mathfrak{A}$

(b) Ist $I \neq \emptyset$ abzählbar und sind $A_i \in \mathfrak{A}$, $i \in I$, dann folgt:

  (i) $\bigcup\limits_{i \in I} A_i \in \mathfrak{A}$

  (ii) $\bigcap\limits_{i \in I} A_i \in \mathfrak{A}$

Satz 6.2

$\Omega \neq \emptyset$, $\mathfrak{A}$ σ-Algebra in Ω; $A_n \in \mathfrak{A}$, $n \in \mathbb{N}$

$$\Rightarrow \quad B := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m \in \mathfrak{A} \qquad (\text{„lim inf } A_n\text{“})$$

$$C := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m \in \mathfrak{A} \qquad (\text{„lim sup } A_n\text{“})$$

Definition 6.2

$\Omega \neq \emptyset$, $\mathfrak{A} \subset \mathfrak{P}(\Omega)$, $P\colon \mathfrak{A} \to \mathbb{R}$

(a) $(\Omega, \mathfrak{A})$ heißt Messraum, falls $\mathfrak{A}$ σ-Algebra in Ω ist.

(b) Ist $(\Omega, \mathfrak{A})$ ein Messraum, so heißt P ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathfrak{A})$, falls P nicht-negativ, normiert und σ-additiv ist.

(c) $(\Omega, \mathfrak{A}, P)$ heißt ein (allgemeiner) Wahrscheinlichkeitsraum, falls $(\Omega, \mathfrak{A})$ ein Messraum und P ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathfrak{A})$ ist.

Bemerkung 6.1

(a) Die Sätze 1.3 und 2.4 gelten auch, mit folgenden Modifikationen:

  (i) Ersetze „(Ω, P) diskreter Wahrscheinlichkeitsraum“ durch „$(\Omega, \mathfrak{A}, P)$ Wahrscheinlichkeitsraum“.

  (ii) Von allen auftretenden Teilmengen von Ω wird zusätzlich verlangt, dass sie Element von $\mathfrak{A}$ sind.

(b) Mit den selben Modifikationen werden die Definition 3.1 und 3.2 auf allgemeine Wahrscheinlichkeitsräume übertragen.

(c) Mit diesen Modifikationen bleiben die Sätze 3.1 und 3.3 gültig.

(d) Es gilt folgende Verallgemeinerung von Satz 3.4:

Satz 6.3 („Produktmaßsatz“)

Seien $(\Omega_1, \mathfrak{A}_1, P_1), (\Omega_2, \mathfrak{A}_2, P_2)$ Wahrscheinlichkeitsräume und $\Omega := \Omega_1 \times \Omega_2$. Dann gilt:

(a) Es gibt σ-Algebra $\mathfrak{A}$ in Ω und ein Wahrscheinlichkeitsmaß P auf $(\Omega, \mathfrak{A})$ mit

  (i) $A_1 \times A_2 \in \mathfrak{A} \qquad \forall\, A_1 \in \mathfrak{A}_1, A_2 \in \mathfrak{A}_2$

  (ii) $P(A_1 \times A_2) = P(A_1) \cdot P(A_2) \qquad \forall\, A_1 \in \mathfrak{A}_1, A_2 \in \mathfrak{A}_2$

(b) Wählt man $\mathfrak{A}$ als kleinste σ-Algebra in Ω, die (a)(i) erfüllt, so ist P eindeutig bestimmt.

Zu Zufallsvariablen: $(\Omega, \mathfrak{A}, P)$, $X\colon \Omega \to \mathfrak{X}$. Ω ist im Allgemeinen *nicht* abzählbar und damit ist $X(\Omega)$ im Allgemeinen nicht abzählbar.

Definition 6.3 (vgl. Def. 4.2)

Seien $(\Omega, \mathfrak{A}, P)$ Wahrscheinlichkeitsraum, $\mathfrak{X} \neq \emptyset$ (beliebig), $X\colon \Omega \to \mathfrak{X}$

(a) X heißt diskrete Zufallsvariable auf $(\Omega, \mathfrak{A}, P)$ mit Werten in $\mathfrak{X}$, falls gilt:

  (i) $X(\Omega)$ abzählbar

  (ii) $X^{-1}(B) \in \mathfrak{A} \qquad \forall\, B \in \mathfrak{P}(\mathfrak{X})$

(b) Ist X diskrete Zufallsvariable auf $(\Omega, \mathfrak{A}, P)$ mit Werten in $\mathfrak{X}$, so heißt

$$P^X\colon \begin{cases} \mathfrak{P}(\mathfrak{X}) \to \mathbb{R} \\ B \mapsto P\left(X^{-1}(B)\right) \end{cases} \quad \text{die Verteilung von } X.$$

Satz 6.4

Sei $(\Omega, \mathfrak{A})$ Messraum, $\mathfrak{X} \neq \emptyset$ beliebig, $X\colon \Omega \to \mathfrak{X}$ mit $X(\Omega)$ abzählbar. Dann sind folgende Aussagen äquivalent:

(a) $X^{-1}(B) \in \mathfrak{A} \qquad \forall\, B \in \mathfrak{P}(\mathfrak{X})$

(b) $X^{-1}(\{x\}) \in \mathfrak{A} \qquad \forall\, x \in \mathfrak{X}$

Satz 6.5

Seien $(\Omega, \mathfrak{A}, P)$ Wahrscheinlichkeitsraum, $\mathfrak{X}, \mathfrak{Z} \neq \emptyset$, $X\colon \Omega \to \mathfrak{X}$, $g\colon \mathfrak{X} \to \mathfrak{Z}$, so dass X diskrete Zufallsvariable auf $(\Omega, \mathfrak{A}, P)$ ist. Dann ist

$$Y\colon \begin{cases} \Omega \to \mathfrak{Z} \\ \omega \mapsto g\big(X(\omega)\big) \end{cases} \quad \text{eine diskrete Zufallsvariable auf } (\Omega, \mathfrak{A}, P) \text{ mit Werten in } \mathfrak{Z}.$$

Satz 6.6

Seien $(\Omega, \mathfrak{A}, P)$ Wahrscheinlichkeitsraum, $\mathfrak{X}_1, \mathfrak{X}_2 \neq \emptyset$, $\mathfrak{X} := \mathfrak{X}_1 \times \mathfrak{X}_2$, $X_i\colon \Omega \to \mathfrak{X}_i, i = 1,2$ und $\tilde{X}\colon \Omega \to \mathfrak{X}$, $\omega \mapsto \big(X_1(\omega), X_2(\omega)\big)$. Dann gilt:

$\tilde{X}$ ist diskrete Zufallsvariable auf $(\Omega, \mathfrak{A}, P)$ mit Werten in $\mathfrak{X}$ genau dann, wenn X_i diskrete Zufallsvariable auf $(\Omega, \mathfrak{A}, P)$ mit Werten in $\mathfrak{X}_i$ ist $(i = 1, 2)$.

Bemerkung 6.2

(a) Lemma 4.3 bleibt gültig mit folgenden Modifikationen:

 (i) Ersetze „(Ω, P) diskreter Wahrscheinlichkeitsraum“ durch „$(\Omega, \mathfrak{A}, P)$ allgemeiner Wahrscheinlichkeitsraum“.

 (ii) Ersetze „X_1, X_2 Zufallsvariablen auf (Ω, P)“ durch „X_1, X_2 diskrete Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$“.

(b) Mit den entsprechenden Modifikationen wird Definition 4.3 auf diskrete Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$ übertragen.

(c) Sätze 4.1 und 4.2 gelten mit den entsprechenden Modifikationen auch für diskrete Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$.

(d) $(\Omega, \mathfrak{A}, P)$ Wahrscheinlichkeitsraum, X diskrete Zufallsvariable auf $(\Omega, \mathfrak{A}, P)$ mit Werten in $\mathbb{R}$; $\mathfrak{X}^* := X(\Omega)$ *(! $\mathfrak{X}^*$ abzählbar !)*. Ist die Reihe $\sum\limits_{x \in \mathfrak{X}^*} |x| \cdot P\big(\{X = x\}\big) < \infty$, so heißt

$$EX := \sum_{x \in \mathfrak{X}^*} x \cdot P\big(\{X = x\}\big) \text{ der } \underline{\text{Erwartungswert von } X}.$$

(e) Die Begriffe „Varianz“, „Kovarianz“ und „Korrelation“werden analog wie in §5 eingeführt.

(f) Alle Aussagen in §5 bleiben gültig.

(g) Alle Voraussetzungen der im folgenden Paragraphen angegebenen Sätze sind nicht leer. Insbesondere gilt der folgende Satz:

Satz 6.7

Für $n \in \mathbb{N}$ sei (Ω_n, P_n) ein diskreter Wahrscheinlichkeitsraum, dann gibt es einen allgemeinen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ und darauf eine Folge von stochastisch unabhängigen Zufallsvariablen $X_n, n \in \mathbb{N}$, mit Werten in $\Omega_n, n \in \mathbb{N}$, so dass $P^{X_n} = P_n \quad \forall n \in \mathbb{N}$.

§7 Gesetze der großen Zahlen

Satz 7.1 („Schwaches Gesetz der großen Zahlen")

Seien $X_n, n \in \mathbb{N}$, diskrete stochastisch unabhängige Zufallsvariablen auf einem allgemeinen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit Werten in $\mathbb{R}$; alle X_n mögen dieselbe Verteilung besitzen; ist dann $EX_1^2 < \infty$, $a := EX_1$, so gilt:

$$\lim_{n\to\infty} P\left(\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| \geq \varepsilon\right\}\right) = 0 \quad \forall\, \varepsilon > 0$$

Satz 7.2

Seien $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum; $Y, Y_n, n \in \mathbb{N}$, diskrete Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$ mit Werten jeweils in $\mathbb{R}$; ist dann $C_0 := \big\{\omega \in \Omega \mid \lim\limits_{n\to\infty} Y_n(\omega) = Y(\omega)\big\}$, so gilt $C_0 \in \mathfrak{A}$.

Satz 7.3

$(\Omega, \mathfrak{A}, P)$ sei Wahrscheinlichkeitsraum, $Y, Y_n, n \in \mathbb{N}$, diskrete Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$ mit Werten jeweils in $\mathbb{R}$. Dann sind folgende Aussagen äquivalent:

(a) $P\big(\{\omega \in \Omega \mid \lim\limits_{n\to\infty} Y_n(\omega) = Y(\omega)\}\big) = 1$

(b) $\exists\, N \in \mathfrak{A}$ mit $P(N) = 0$ und $\lim\limits_{n\to\infty} Y_n(\omega) = Y(\omega) \quad \forall\, \omega \in N^c$

(c) $P\left(\bigcup\limits_{n=1}^{\infty} \bigcap\limits_{m=n}^{\infty} \left\{|Y_m - Y| < \frac{1}{k}\right\}\right) = 1 \quad \forall\, k \in \mathbb{N}$

(d) $P\left(\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{m=n}^{\infty} \left\{|Y_m - Y| \geq \frac{1}{k}\right\}\right) = 0 \quad \forall\, k \in \mathbb{N}$

(e) $\lim\limits_{n\to\infty} P\left(\bigcup\limits_{m=n}^{\infty} \left\{|Y_m - Y| \geq \frac{1}{k}\right\}\right) = 0 \quad \forall\, k \in \mathbb{N}$

Satz 7.4 („Kolmogoroff-Ungleichung")

Seien $(\Omega, \mathfrak{A}, P)$ Wahrscheinlichkeitsraum, $n \in \mathbb{N}$, $X_1, \cdots, X_n$ stochastisch unabhängige diskrete Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$ mit Werten jeweils in $\mathbb{R}$.
Ferner seien $EX_i^2 < \infty$, $a_i := EX_i$, $\sigma_i^2 := VarX_i$, $1 \leq i \leq n$, $\tau^2 := \sum\limits_{i=1}^{n} \sigma_i^2 = Var\Big(\sum\limits_{i=1}^{n} X_i\Big)$. Ist dann $\tau^2 > 0$, so gilt

$$P\left(\bigcup_{j=1}^{n}\left\{\left|\sum_{\rho=1}^{j}(X_\rho - a_\rho)\right| \geq k \cdot \tau\right\}\right) \leq \frac{1}{k^2} \quad \forall\, k > 0$$

Definition 7.1

Seien $Y, Y_n, n \in \mathbb{N}$ diskrete Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit Werten jeweils in $\mathbb{R}$.

(a) $(Y_n)_{n\in\mathbb{N}}$ heißt stochastisch konvergent gegen Y, falls

$$\lim_{n\to\infty} P\big(\{|Y_n - Y| \geq \varepsilon\}\big) = 0 \quad \forall\, \varepsilon > 0$$

(b) $(Y_n)_{n\in\mathbb{N}}$ konvergiert (P-) fast sicher gegen Y, falls

$$P\left(\left\{\lim_{n\to\infty} Y_n = Y\right\}\right) = 1$$

Satz 7.5 („Starkes Gesetz der großen Zahlen")

$X_n, n \in \mathbb{N}$, seien diskrete, stochastisch unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit Werten jeweils in $\mathbb{R}$. Es gelte $EX_n^2 < \infty\ \forall\, n \in \mathbb{N}$. Ist dann $a_n := EX_n$, $\sigma_n^2 := Var X_n$, $Y_n := \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - a_i), n \in \mathbb{N}$, und gilt die „Kolmogoroff-Bedingung" $\sum\limits_{n=1}^{\infty}\frac{\sigma_n^2}{n^2} < \infty$, so folgt:

$(Y_n)_{n\in\mathbb{N}}$ konvergiert fast sicher gegen 0

Satz 7.6

Mit den Bezeichnungen und Voraussetzungen aus Satz 7.5 gelte ferner:
Es gibt $a \in \mathbb{R}$ mit $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{i=1}^{n} a_i = a$. Dann folgt:

$\left(\frac{1}{n}\sum\limits_{i=1}^{n} X_i\right)_{n\in\mathbb{N}}$ konvergiert fast sicher gegen a

Satz 7.7

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von stochastisch unabhängigen diskreten Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit Werten jeweils in $\mathbb{R}$. Alle $X_n, n \in \mathbb{N}$, mögen dieselbe Verteilung besitzen (also $P(\{X_n = x\}) = P(\{X_1 = x\})\ \forall\, x \in \mathbb{R}\ \forall\, n \in \mathbb{N}$); gilt dann $EX_1^2 < \infty$, so folgt:

$\left(\frac{1}{n}\sum\limits_{i=1}^{n} X_i\right)_{n\in\mathbb{N}}$ konvergiert fast sicher gegen $a_1 := EX_1$

Satz 7.8

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von stochastisch unabhängigen diskreten Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit Werten jeweils in einer (beliebigen) Menge $\mathfrak{X} \neq \emptyset$. Alle $X_n, n \in \mathbb{N}$, mögen dieselbe Verteilung besitzen.
Ist dann $g\colon \mathfrak{X} \to \mathbb{R}$, $Y_n := g(X_n), n \in \mathbb{N}$, und gilt $EY_1^2 < \infty$, so folgt:

$\left(\frac{1}{n}\sum\limits_{i=1}^{n} Y_i\right)_{n\in\mathbb{N}}$ konvergiert fast sicher gegen $b_1 := EY_1$