Bew: $P[H_i \mid A] = \frac{P[A \cap H_i]}{P[A]} = \frac{P[A \mid H_i]\, P[H_i]}{\sum_{\substack{k \in I \\ P[H_k] \neq 0}} P[A \mid H_k]\, P[H_k]}$

Bsp 2.7: · Eine Krankheit trete mit Häufigkeit 10^{-4} auf

· Ein Test sei positiv (+) bei 96% der Kranken und 0,1% der Gesunden

$\Rightarrow$ A-priori: $P[K] = 10^{-4}$, $P[K^c] = 1 - 10^{-4}$

Bed. W-keiten: $P[+ \mid K] = 0{,}96$

$P[+ \mid K^c] = 0{,}001$

A posteriori: $P[K \mid +]$

$$= \frac{P[+ \mid K]\, P[K]}{P[+ \mid K]\, P[K] + P[+ \mid K^c]\, P[K^c]} \approx \frac{1}{11}$$

$\Rightarrow$ Die W'keit, dass man krank ist, wenn man positiv getestet wird ist nur $\approx$ 9%.

2.3 Mehrstufige Modelle

Wir betrachten ein n-stufiges Zufallsexperiment mit diskreten Stichprobenräumen $\Omega_1, \dots, \Omega_n$.

Setze $\Omega := \Omega_1 \times \dots \times \Omega_n = \{\omega = (\omega_1, \dots, \omega_n) : \omega_i \in \Omega_i, 1 \leq i \leq n\}$

und $X_k(\omega) = \omega_k$.

Wir interpretieren den Index k als Zeit.

Bsp: a) ~~Akt~~ Aktienkurse

b) Wasserstand des Rheins, montags 8^{00}

Satz 2.8:

Seien $p_1(\cdot)$ und $p_k(\cdot \mid x_1,\dots,x_{k-1})$ für jedes $2 \le k \le n$ und $x_1 \in \Omega_1, \dots, x_{k-1} \in \Omega_{k-1}$ Massenf'kt einer W'vtlg auf Ω_k. Dann $\exists!$ W'vtlg $\mathbb{P}$ auf $(\Omega, \mathcal{P}(\Omega))$

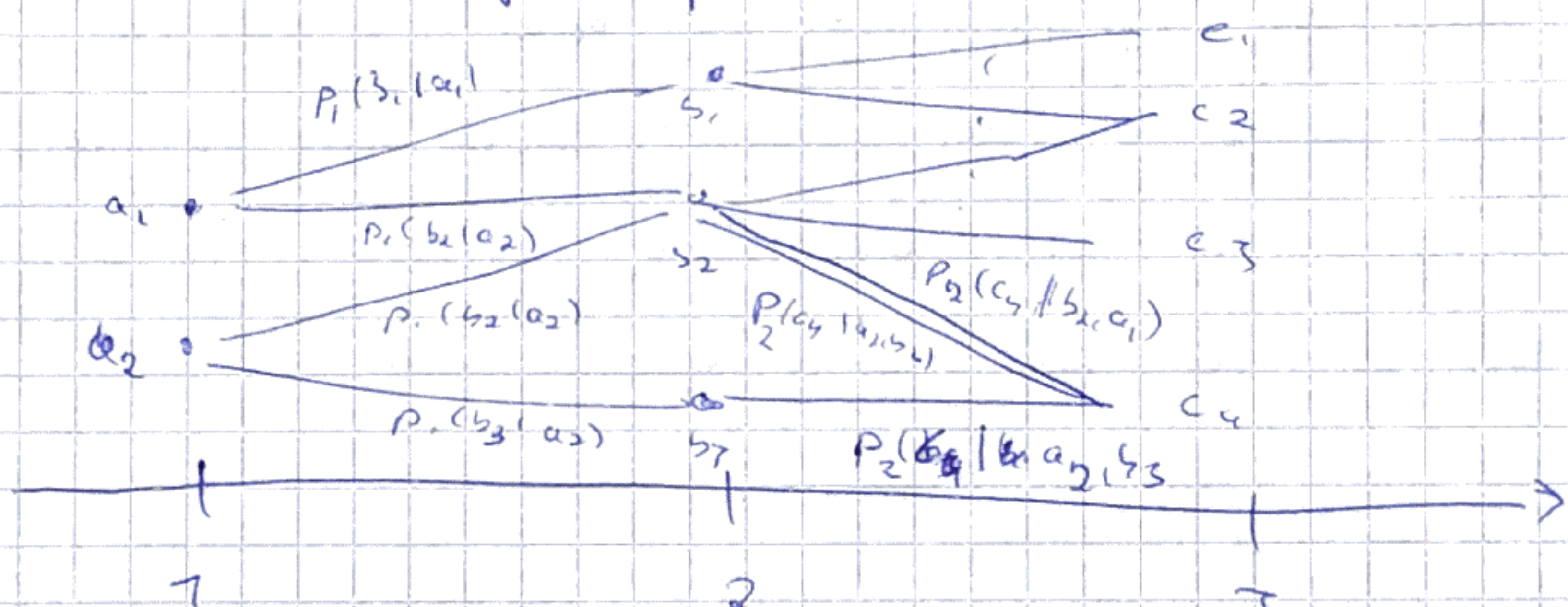
s.d. a) $\mathbb{P}[X_1 = x_1] = p_1(x_1)$

b) $\mathbb{P}[X_k = x_k \mid X_1 = x_1, \dots, X_{k-1} = x_{k-1}] = p_k(x_k \mid x_1, \dots, x_{k-1})$

für $2 \le k \le n$, $x_i \in \Omega_i$ mit $\mathbb{P}[X_1 = x_1, \dots, X_{k-1} = x_{k-1}] \neq 0$.

$\mathbb{P}$ hat die Massenf'kt $p(x_1,\dots,x_n) = p_1(x_1)\, p_2(x_2 \mid x_1) \cdots p_n(x_n \mid x_1,\dots,x_{n-1})$

Bsp:



Beweis:

Eindeutigkeit: Zu zeigen für $1 \le k \le n$

(*) $\mathbb{P}[X_1 = x_1, \dots, X_k = x_k] = p_1(x_1)\, p_2(x_2 \mid x_1) \cdots p_k(x_k \mid x_1, \dots, x_{k-1})$

Beweis durch Induktion:

IA: $k = 1$, stimmt per Annahme.

IV: (*) gelte für $k-1$ $(k \ge 2)$.

IS: Falls $\mathbb{P}[X_1 = x_1, \dots, X_{k-1} = x_{k-1}] = 0$ gilt $0 = 0 \Leftarrow$ (*).

Sonst:

$\mathbb{P}[X_1 = x_1, \dots, X_k = x_k] = \mathbb{P}[X_1 = x_1, \dots, X_{k-1} = x_{k-1}] \cdot$

$\mathbb{P}[X_k = x_k \mid X_1 = \dots, X_{\dots} = \dots]$

Existenz: $p \geq 0$. Zu zeigen noch: Normierung. Da $p_1, \dots, p_n(\cdot|\cdot)$ W'vtlg folgt:

$$\sum_{x\in\Omega} p(x) = \sum_{x_1\in\Omega_1} \dots \sum_{x_n\in\Omega_n} p(x_1,\dots,x_n)$$

$$= \sum_{x_1\in\Omega_1} p_1(x_1) \sum_{x_2\in\Omega_2} \dots \underbrace{\sum_{x_n\in\Omega_n} p_n(x_n|x_1,\dots,x_{n-1})}_{=1 \text{ (nach Ann.)}} = \dots = 1$$

□

Falls $p_k(x_k|x_1,\dots,x_{k-1}) = p_k(x_k|x_{k-m},\dots,x_{k-1})$, d.h. das k-te Teilexperiment wird nur von den Experimenten $k-1,\dots,k-m$ beeinflusst, so sagen wir unser Modell hat ein Gedächtnis von "$m-1$ Schritten".

↳ 2 wichtige Klassen:

i) $m=0$: Produktmodelle (⇝ keine Abhängigkeit)

ii) $m=1$: Markovkette (⇝ nur kein "Gedächtnis")

2.3.1 Produktmodelle

Annahme: $p_k(x_k|x_1,\dots,x_{k-1}) = p_k(x_k) \overset{\text{Satz 2.8}}{\rightsquigarrow} p(x_1,\dots,x_n) = \prod_{k=1}^{n} p_k(x_k)$

Def 2.9: Seien P_i W'vtlg auf Ω_i mit Massenfkt p_i. Die W'vtlg P auf $\Omega = \Omega_1 \times \dots \times \Omega_n$ mit Massenfkt $p(x_1,\dots,x_n) = \prod_{k=1}^{n} p_k(x_k)$ heißt Produkt von $P_1,\dots,P_n$.
Wir schreiben $P = P_1 \otimes \dots \otimes P_n$.

Bsp 2.10 (n-dimensionale Bernoulli-Vtlg zum Parameter p)

$\Omega_i = \{0,1\}$, $p_i(1) = p = (1-p)\left(\frac{p}{1-p}\right)^1$

$p_i(0) = (1-p) = (1-p)\left(\frac{p}{1-p}\right)^0$ $\quad 1 \leq i \leq n$

Produktmodell: $p(x_1,\dots,x_n) = (1-p)^n \prod_{k=1}^{n} \left(\frac{p}{1-p}\right)^{x_k} = p^{\sum x_k}(1-p)^{n-\sum x_k}$

Satz 2.11

Sei $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ mit $\Omega = \Omega_1 \times \dots \times \Omega_n$, $\mathbb{P} = \mathbb{P}_1 \otimes \dots \otimes \mathbb{P}_n$ ein Produktmodell. Dann gilt für beliebig $A_i \subseteq \Omega_i$ $1 \le i \le n$:

$$\mathbb{P}[A_1 \times \dots \times A_n] = \prod_{i=1}^{n} \mathbb{P}_i[A_i]$$

Beweis: $\mathbb{P}[A_1 \times \dots \times A_n] = \sum_{(x_1,\dots,x_n) \in A_1 \times \dots \times A_n} p(x_1, \dots, x_n)$

$$= \sum_{x_1 \in A_1} \dots \sum_{x_n \in A_n} \prod_{k=1}^{n} p_k(x_k) = \prod_{k=1}^{n} \sum_{x_k \in A_k} p_k(x_k)$$

$$= \prod_{k=1}^{n} \mathbb{P}_k[A_k] \qquad \square$$

Bemerkung 2.12: Satz 2.11 besagt, dass in Produktmodell die ZV $X_i(x_1, \dots, x_n) = x_i$ unabhängig sind (vgl. Def 1.28, Bem 1.42 c) und Abschnitt 2.4)

2.3.2 Markovketten (MK)

Es ist Konvention den "Zeitindex" bei 0 anstarten.

Fixiere abzählbares S und betrachte mit $\Omega_i = S$, $0 \le i \le n$

$$\Omega = S^{n+1} = \{x = (x_0, \dots, x_n) : x_i \in S\}$$

Def 2.13: i) Eine MK ist ein mehrstufiges Modell mit

$$p_k(x_k \mid x_1 \dots x_{k-1}) = p_k(x_k \mid x_{k-1}) \quad 1 \le k \le n.$$

ii) Eine Matrix $(P(x,y))_{x,y \in S}$ mit

a) $P(x,y) \ge 0$, $x, y \in S$

b) $\sum_{y \in S} P(x,y) = 1$, $\forall x \in S$

heißt stochastische Matrix auf S.

Warum MK und stoch. Matrize in einer Def?

Lemma 2.14: $P_k(x,y) := p_k(y \mid x)$ definiert eine stoch. Matrix

Beweis: Konsequenz aus Ann., dass $p_k(\cdot \mid x)$ W'keit ist!

(38) Warum nützlich?

Dies erlaubt häufig folgende Größen zu berechnen.

$P[X_n = x]$, $\lim_{n\to\infty} P[X_n = x]$ (abhängig vom Startwert?)

Grund: Sind P, Q stoch. Matrizen auf S so ist auch $P\cdot Q$ stoch. Matrix, denn (nicht-neg. Einträge sind klar)

$$\sum_{y\in S} P\cdot Q(x,y) = \sum_{y\in S}\sum_{z\in S} P(x,z)\,Q(z,y)$$

$$= \sum_{z\in S} P(x,z) \underbrace{\sum_{y\in S} Q(z,y)}_{=1} = 1$$

Damit:

Satz 2.15: Sei μ_0^{tr} der Zeilenvektor mit Einträgen $p_0(x)$, $x\in S$, und $P_1, \dots, P_n$ die „Übergangsmatrizen" einer MK geg. durch $P_j(x,y) = p_j(y\,|\,x)$. Dann gilt

$$\mu_n^{tr}(x) := P[X_n = x] = (\mu_0^{tr} P_1 \cdots P_n)(x), \quad x\in S.$$

Beweis: $p(x_0,\dots,x_n) \overset{\text{Satz 2.8}}{\underset{\text{Def 2.13}}{=}} p_0(x_0)\, \underbrace{p_1(x_1\,|\,x_0)}_{P_1(x_0,x_1)} \cdots \underbrace{p_n(x_n\,|\,x_{n-1})}_{P_n(x_{n-1},x_n)}$

$$P[X_n = x] = \sum_{x_0\in S}\cdots\sum_{x_{n-1}\in S} p(x_0,\dots,x_{n-1},x) = (\mu_0^{tr} P_1 \cdots P_n)(x) \quad \square$$

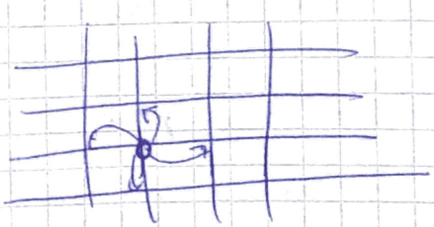
VS

Bsp 2.16: a) Produktmodelle sind MK

b) Irrfahrt auf $S=\mathbb{Z}^d$ (random walk), $d\in\mathbb{N}$.

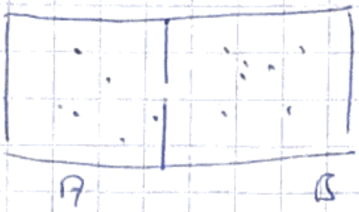
symmetrische

$$p_n(x,y) = \begin{cases} \frac{1}{2d}, & \text{falls } \|x-y\| = 1 \\ 0, & \text{sonst} \end{cases}$$



c) Urnenmodell von (Paul und Tatiana) Ehrenfest

Modell zum Austausch von n Gasmolekülen zwischen zwei Behältern



4×10^{23}

n Teilchen in A und B,

zu jeder Zeit $t \in \mathbb{N}$, wechselt ein zufällig gewähltes Teilchen die Urne.

Zwei verschiedene Beschreibungen durch (MK):

1. Makroskopisch:

$n_A :=$ # Teilchen in Urne A

makroskopische Größe:

$$\varrho_A := \frac{n_A}{n} \in \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\right\} =: S$$

Übergangsmatrix / W'keit

$$p(x,y) = \begin{cases} x, & \text{falls } y = x - \frac{1}{n} \\ 1-x, & \text{falls } y = x + \frac{1}{n} \\ 0, & \text{sonst} \end{cases}$$

beschreibt diskrete (zeitlich homogene) Entwicklung ~~des Systems~~ von ϱ_A.

2. Mikroskopisch:

$$S := \{0,1\}^n = \{\sigma = (\sigma_1, \dots, \sigma_n) : \sigma_i \in \{0,1\}\}, \text{ wobei}$$

$$\sigma_i = \begin{cases} 1 & \text{falls Teilchen } i \text{ in Urne A} \\ 0 & \text{———— B} \end{cases}$$

(S ist sehr groß! 2^n Elemente)

Übergangsw'keit:

$$p(\sigma, \tilde{\sigma}) = \begin{cases} \frac{1}{n}, & \text{falls } \sum |\sigma_i - \tilde{\sigma}_i| = 1 \\ 0, & \text{sonst} \end{cases}$$

Dies ist eine symmetrische Irrfahrt auf dem Hyperwürfel $\mathbb{R}^n$

(40) Um Übergangswahrscheinlichkeiten zwischen zwei beliebigen Zeiten $0 \le k < l \le n$ zu berechnen, bietet sich folgende Verallg. von Satz 2.15 an:

Satz 2.17:

In der Situation von Satz 2.15 gilt für alle $0 \le k < l \le n$ und $x_0, \dots, x_k \in S$ mit $\mathbb{P}[X_0 = x_0, \dots, X_k = x_k] \neq 0$

$$\mathbb{P}[X_l = x_l \mid X_0 = x_0, \dots, X_k = x_k] = \mathbb{P}[X_l = x_l \mid X_k = x_k]$$

$$= P_{k+1} \cdots P_l (x_k, x_l)$$

Beweis: $\mathbb{P}[X_l = x_l \mid X_0 = x_0, \dots, X_k = x_k]$

$$= \frac{\mathbb{P}[X_0 = x_0, \dots, X_k = x_k, X_l = x_l]}{\mathbb{P}[X_0 = x_0, \dots, X_k = x_k]}$$

$$= \frac{\sum_{x_{k+1}, \dots, x_{l-1} \in S} p_0(x_0) p_1(x_0, x_1) \cdots p_k(x_{k-1}, x_k) \, p_{k+1}(x_k, x_{k+1}) \cdots p_l(x_{l-1}, x_l)}{p_0(x_0) p_1(x_0, x_1) \cdots p_k(x_{k-1}, x_k)}$$

$$= \sum_{x_{k+1}, \dots, x_{l-1} \in S} p_{k+1}(x_k, x_{k+1}) \cdots p_l(x_{l-1}, x_l) = (P_{k+1} \cdots P_l)(x_k, x_l)$$

Mit Satz 2.15 folgt

$$\mathbb{P}[X_l = x_l \mid X_k = x_k] = \frac{\mathbb{P}[X_k = x_k, X_l = x_l]}{\mathbb{P}[X_k = x_k]}$$

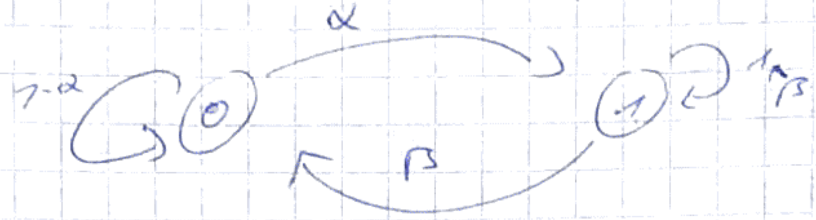
$$= \sum_{x_{k+1}, \dots, x_{l-1} \in S} \frac{(\mu_0^T P_1 \cdots P_k)(x_k) \, p_{k+1}(x_k, x_{k+1}) \cdots p_l(x_{l-1}, x_l)}{(\mu_0^T P_1 \cdots P_k)(x_k)}$$

$$= (P_{k+1} \cdots P_l)(x_k, x_l)$$

Bem 2.18: Falls $P_k = P$ für alle k heißt die MK homogen.

Bsp 2.19: $S=\{0,1\}$, $\alpha,\beta\in(0,1)$

Graphisch:



Ü-matrix: $P=\begin{pmatrix}1-\alpha & \alpha\\ \beta & 1-\beta\end{pmatrix}$

Methoden P^n zu berechnen: i) wenn möglich: diagonalisieren

ii) Formel finden und dann induktiv beweisen.

Bedingen auf Zeitpunkt $n-1$ liefert:

$$
\begin{aligned}
P^n(0,0) &= P^{n-1}(0,0)\,P(0,0) + P^{n-1}(0,1)\,P(1,0)\\
&= P^{n-1}(0,0)\,(1-\alpha) + \underbrace{P^{n-1}(0,1)}_{\overset{P^{n-1}\text{ stoch. Matrix}}{=}(1-P^{n-1}(0,0))}\,\beta\\
&= \beta + (1-\alpha-\beta)\,P^{n-1}(0,0)\\
&= \beta + (1-\alpha-\beta)\cdot\beta + (1-\alpha-\beta)^2 P^{n-2}(0,0)\\
&= \dots = \underbrace{P^0(0,0)}_{=1}(1-\alpha-\beta)^n + \sum_{k=0}^{n-1}\beta(1-\alpha-\beta)^k\\
&= (1-\alpha-\beta)^n + \beta\,\frac{1-(1-\alpha-\beta)^n}{1-(1-\alpha-\beta)}\\
&= \frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^n\,\frac{\alpha}{\alpha+\beta} = 1-P^n(0,1)
\end{aligned}
$$

Symmetrisches Argument für $P^n(1,1)$ liefert

$$
P^n = \begin{pmatrix}\frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta}\\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta}\end{pmatrix} + \underbrace{(1-\alpha-\beta)^n\begin{pmatrix}\frac{\alpha}{\alpha+\beta} & \frac{-\alpha}{\alpha+\beta}\\ \frac{-\beta}{\alpha+\beta} & \frac{\beta}{\alpha+\beta}\end{pmatrix}}_{\xrightarrow{n\to\infty} 0 \text{ exponentiell schnell!}}
$$

(42) Insbesondere folgt

$$\lim_n P^n(0) = (p_0(0), p_1(1)) \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix} = \frac{\beta}{\alpha+\beta}$$

unabhängig von der Startverteilung (p_0, p_1)
("exponentieller Gedächtnisverlust") und Konvergenz ins Gleichgewicht)

2.4. Unabhängigkeit von Ereignissen.

Sei $(\Omega, \mathcal{P}(\Omega), P)$ diskreter W'raum, $A, B \in \mathcal{P}(\Omega)$. A und B sind unabh., falls „A tritt ein" keine Info über B enthält und vice versa.

$\hookrightarrow \quad P[B \mid A] = P[B]$, falls $P[A] \neq 0$

$P[A \mid B] = P[A]$, — $P[B] \neq 0$

$\Rightarrow \quad P[A \cap B] = P[A] \cdot P[B]$

Def 2.20: Sei $(\Omega, \mathcal{P}(\Omega), P)$ ein diskreter W'raum. Eine Familie von Ereignissen $(A_k)_{k \in I}$ heißt unabhängig (bzgl P), falls

$$P[A_{i_1} \cap \dots \cap A_{i_k}] = \prod_{j=1}^{k} P[A_{i_j}]$$

$\forall\, k \leq |I|$ und paarweise verschiedene $i_1, \dots, i_k \in I$.

Bsp 2.21: i) $P[A] \in \{0,1\} \Rightarrow A$ unabhängig von B für alle $B \in \mathcal{P}(\Omega)$.

$P[A] = 0 \Rightarrow 0 \leq P[A \cap B] \leq P[A] = 0 = P[B] \cdot P[A]$.

$P[A] = 1 \Rightarrow P[B] \cdot P[A] = P[B] = P[B \cap A] + \underbrace{P[B \cap A^c]}_{=0,\text{ da } P[A^c]=0.}$

(ii) 3-facher fairer Münzwurf:

$\Omega = \{0,1\}^3$, $P = \text{Unif}(\Omega)$, $X_k((\omega_1, \omega_2, \omega_3)) = \omega_k$, $1 \leq k \leq 3$

$A_1 = \{X_1 = X_2\}$, $A_2 = \{X_2 = X_3\}$, $A_3 = \{X_1 = X_3\}$

$\mathbb{P}[A_j] = \frac{4}{8} = \frac{1}{2} \qquad 1 \le j \le 3$

$\mathbb{P}[A_i \cap A_j] = \frac{2}{8} = \frac{1}{4} = \mathbb{P}[A_i]\mathbb{P}[A_j] \quad i \neq j$

$\Rightarrow A_1, A_2, A_3$ sind paarweise unabhängig.

Aber: $\mathbb{P}[A_1 \cap A_2 \cap A_3] = \frac{1}{4} \neq \mathbb{P}[A_1]\,\mathbb{P}[A_2]\,\mathbb{P}[A_3] = \frac{1}{8}$.

Bem / Lemma 2.22:

Sind die Ereignisse $A_1, \dots, A_n \in \mathcal{P}(\Omega)$ unabhängig und $B_k = A_k$ oder A_k^c für $1 \le k \le n$ so sind auch die Ereignisse $B_1, \dots, B_n$ unabhängig.

Beweis: ÜA.

Seien $A_1, A_2, \dots \in \mathcal{P}(\Omega)$ unabhängige Ereignisse mit $\mathbb{P}[A_i] = p \in [0,1]$ für alle i. Setze

$T(\omega) = \inf\{n \in \mathbb{N} : \omega \in A_n\}$

mit ~~inf~~ $\inf(\emptyset) = \infty$. Dann folgt:

$$\mathbb{P}[T = n] = \mathbb{P}[A_1^c \cap A_2^c \cap \dots \cap A_{n-1}^c \cap A_n] = (1-p)^{n-1} p = \left(\prod_{i=1}^{n-1} \mathbb{P}[A_i^c]\right) \mathbb{P}[A_n]$$

Def 2.23) Sei $p \in [0,1]$. Die W'vtlg μ auf $\mathbb{N} \cup \{\infty\}$ mit Massenfkt $\mu(n) = p(1-p)^{n-1}$, $n \ge 1$ heißt geometrische Vtlg zum Parameter p, Bezeichnung Geom(p).

Bem 2.24: $\mathbb{P}[T > n] = \mathbb{P}[A_1^c \cap \dots \cap A_n^c] = (1-p)^n \xrightarrow[n \to \infty]{} 0$, falls $p \neq 0$ $\Rightarrow \mathbb{P}[T = \infty] = 0$.

Falls $p = 0$ folgt $\mathbb{P}[T = \infty] = 1$.

Bsp:

(44)

Seien nun wieder $A_1, \dots, A_n \in \mathcal{P}(\Omega)$ unabhängig und setze

$S_n(\omega) = |\{1 \leq i \leq n : \omega \in A_i\}| = \sum_{i=1}^{n} 1_{A_i}(\omega)$

Dann folgt für $0 \leq k \leq n$

$$\mathbb{P}[S_n = k] = \sum_{\substack{I \subseteq \{1,\dots,n\} \\ |I| = k}} \mathbb{P}\Big[\bigcap_{i \in I} A_i \bigcap_{j \in \{1,\dots,n\} \setminus I} A_j^c\Big]$$

$$\overset{\text{Lemma 2.22}}{=} \sum_{\substack{I \subseteq \{1,\dots,n\} \\ |I| = k}} \underbrace{\prod_{i \in I} \mathbb{P}[A_i]}_{p^k} \underbrace{\prod_{j \in \{1,\dots,n\} \setminus I} \mathbb{P}[A_j^c]}_{(1-p)^{n-k}}$$

$$= \binom{n}{k}$$

$\Rightarrow \quad S_n \sim Bin(n, p)$

Wir wissen $\mathbb{E}\, S_n = np$, s.d. $\mathbb{E}\left[\frac{S_n}{n}\right] = p \quad \forall n.$

Wie groß ist die Abweichung von $\frac{S_n}{n}$ vom erwarteten Wert p?

Satz 2.25: (Bernstein Ungleichung)

$\forall \varepsilon > 0$ und $n \in \mathbb{N}$ gilt:

$$\mathbb{P}\left[\frac{S_n}{n} \geq p + \varepsilon\right] \leq e^{-2\varepsilon^2 n}$$

$$\text{und } \mathbb{P}\left[\frac{S_n}{n} \leq p - \varepsilon\right] \leq e^{-2\varepsilon^2 n}$$

Insbesondere gilt:

$$\mathbb{P}\left[\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right] \leq 2 \cdot e^{-2\varepsilon^2 n}$$

Beweis: Schreibe $q = (1-p)$. Dann gilt für alle $\lambda \geq 0$

$$\mathbb{P}[S_n \geq n(p+\varepsilon)] = \sum_{k \geq n(p+\varepsilon)} \binom{n}{k} p^k q^{n-k}$$

$$\leq \sum_{k \geq n(p+\varepsilon)} \binom{n}{k} p^k q^{n-k} e^{\lambda(k - n(p+\varepsilon))}$$

$$\leq \sum_{k=0}^{n} \binom{n}{k} (pe^{\lambda})^k q^{n-k} e^{-n(p+\varepsilon)\lambda}$$

$$\leq (pe^{\lambda} + q)^n (e^{-\lambda p})^n e^{-n\varepsilon\lambda}$$

$$= (pe^{\lambda q} + qe^{-\lambda p})^n e^{-n\varepsilon\lambda}$$

Fakt: $pe^{\lambda q} + qe^{-\lambda p} \leq e^{\lambda^2/8}$ $(*)$ (z.B. Skript Eberle)

Mit $(*)$ folgt: $\mathbb{P}[S_n \geq n(p+\varepsilon)] \leq e^{n(\frac{\lambda^2}{8} - \lambda\varepsilon)}$

Wählen wir nun $\lambda = 4\varepsilon$. Dann

$$\mathbb{P}[S_n \geq n(p+\varepsilon)] \leq e^{-2n\varepsilon^2}$$

Für die zweite Ungleichung sei $\tilde{S}_n = n - S_n \Rightarrow \tilde{S}_n \sim Bin(n,q)$

$$\mathbb{P}[\tilde{S}_n \geq n(q+\varepsilon)] = \mathbb{P}[S_n \leq n(p-\varepsilon)]$$
$$\leq e^{-2\varepsilon^2 n} \qquad \square$$

2.5 Unabhängige ZV und Infofkt.

Def 2.26: Sei $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ ein diskreter W'raum. Eine Familie $X_i : \Omega \to S_i$, $i \in I$ von ZV heißt unabhängig, falls für alle $a_i \in S_i$ die Ereignisse $\{X_i = a_i\}$ unabhängig sind.

Bem 2.27: Sind die ZV $(X_i)_{i\in I}$ unabhängig, so auch $(X_j)_{j\in J}$ für $J \subseteq I$.

↳ Im folgenden meist $I = \{1, \dots, n\}$. Sind $X_i : \Omega \to S_i$ ZV für $1 \leq i \leq n$ so ist auch $(X_1, \dots, X_n) : \Omega \to S_1 \times \dots \times S_n$ ZV.

Def 2.28: Die Vtlg $P_{X_1,\dots,X_n}$ des Zufallsvektors $(X_1, \dots, X_n)$ unter $\mathbb{P}$ heißt gemeinsame Vtlg der ZV $X_1, \dots, X_n$ und hat Massenfkt $p_{X_1,\dots,X_n}(a_1, \dots, a_n) = \mathbb{P}[X_1 = a_1, \dots, X_n = a_n]$

Satz 2.29: Die folgenden Aussagen sind äquivalent:

a) $X_1,\dots,X_n$ sind unabhängig

b) $p_{X_1,\dots,X_n}(a_1,\dots,a_n) = \prod_{k=1}^{n} p_k(a_k)$, $\forall a_k \in S_k$, $k=1,\dots,n$

c) $P_{X_1,\dots,X_n} = \bigotimes_{k=1}^{n} P_{X_k}$

d) Die Ereignisse $\{X_1 \in A_1\},\dots,\{X_n \in A_n\}$ sind unabhängig $\forall A_k \subseteq S_k$, $k=1,\dots,n$

e) Die Ereignisse $\{X_1 = a_1\},\dots,\{X_n = a_n\}$ sind unabhängig $\forall a_k \in S_k$, $k=1,\dots,n$.

Beweis: a) $\Leftrightarrow$ b) $p_{X_1,\dots,X_n}(a_1,\dots,a_n) = P[X_1=a_1,\dots,X_n=a_n] = \prod_{i=1}^{n} P[X_i=a_i] = \prod_{i=1}^{n} p_{X_i}(a_i)$

b) $\Leftrightarrow$ c) folgt aus Def 2.9

c) $\Rightarrow$ d) Sei $1 \le i_1 < i_2 < \dots < i_k \le n$ und $A_{i_j} \subseteq S_{i_j}$ für $1 \le j \le k$. Setze $A_i = S_i$ für $i \notin \{i_1,\dots,i_k\}$.

Dann folgt aus Satz 2.11

$$P[X_{i_1} \in A_{i_1},\dots,X_{i_k} \in A_{i_k}] = P[X_1 \in A_1,\dots,X_n \in A_n]$$

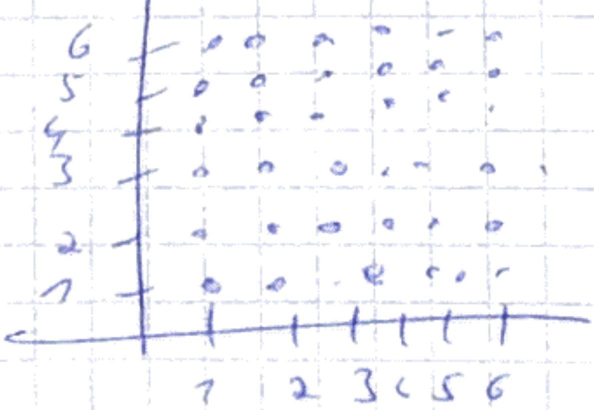
$$= \prod_{i=1}^{n} P[X_i \in A_i] = \prod_{j=1}^{k} P[X_{i_j} \in A_{i_j}]$$

d) $\Rightarrow$ e) $\Rightarrow$ a) folgt direkt □

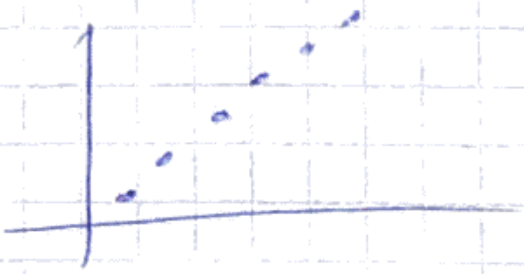
2.30: $X, Y : \Omega \to \{1,\dots,6\}$ gleichverteilte ZV.

Die Punkte illustrieren Ereignisse mit positiver W'keit

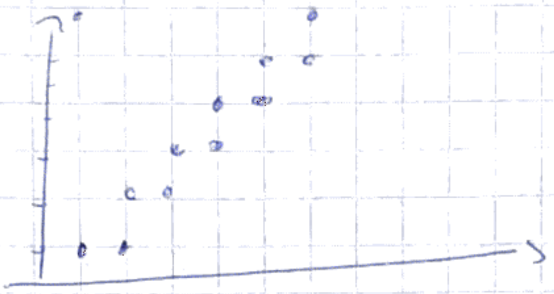
a) X, Y unabhängig



b) $X = Y$



c) $X = (Y + Z) \bmod 6 \quad Z = \mathrm{Ber}(\frac{1}{2})$



2.5.1. Die Irrfahrt auf $\mathbb{Z}$.

Seien $X_1, X_2, X_3,$ unabhängige und identisch verteilte („iid": independent and identically distributed; „uiv") ZV auf einem diskreten W'raum $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ mit

$$\mathbb{P}[X_k = 1] = p$$
$$\mathbb{P}[X_k = -1] = (1-p) \quad , \quad p \in (0,1)$$

Wir fixieren $a \in \mathbb{Z}$ und betrachten die MK auf $\mathbb{Z}$ mit

$$\begin{cases} S_0 = a & , a \in \mathbb{Z}. \text{ (fest)} \\ S_{n+1} := S_n + X_{n+1} & , n \in \mathbb{N}_0. \end{cases} ,$$

d.h. $S_n = \sum_{i=1}^{n} X_i + a$. Die Folge von ZV (MK) $(S_0, S_1, S_2, \dots)$ heißt Irrfahrt auf $\mathbb{Z}$ mit S_n = Position zum Zeitpunkt n.

Fragen: (1) Rückkehrzeit:

$\mathbb{P}[S_n = a] = ?$

(2) Treffzeit o' Hg:

Sei $b \in \mathbb{Z}$ und

$T_b(\omega) = \inf\{n \geq 1 : S_n(\omega) = b\}$, $\inf \emptyset = \infty$

$\Rightarrow \mathbb{P}[T_b \leq n] = ?$

Für $b = a$ ist T_b die erste Rückkehrzeit.

Für $b \neq a$ ist T_b die erste Treffzeit von b.

48

(3) Vtlg von Maximum

Sei $M_n := \max_{0 \le k \le n} S_k \Rightarrow \mathbb{P}[M_n \le \lambda] = ?$

Zuerst berechnen wir die Vtlg von S_n

Prop 2.31: Für $k \in \mathbb{Z}$,

$$\mathbb{P}[S_n = a+k] = \begin{cases} 0 & , |k| > n \text{ oder } n+k \text{ ungerade}, \\ \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}} & , \text{sonst}. \end{cases}$$

Beweis: $S_n = a+k \Leftrightarrow X_1 + \dots + X_n = k$

$$\Leftrightarrow \begin{cases} X_i = 1 & \text{genau } \frac{n+k}{2}\text{-mal} \\ X_i = -1 & \text{— } \frac{n-k}{2}\text{-mal}. \end{cases} \quad \square$$

Damit:

(1) Rückkehrw'keit:

$\mathbb{P}[S_{2n-1} = a] = 0$ für $n \in \mathbb{N}$

$\mathbb{P}[S_{2n} = a] = \binom{2n}{n} p^n (1-p)^n$

und mit der Stirling Formel

$$n! = \sqrt{2\pi n}\, n^n e^{-n} \left(1 + \mathcal{O}\left(\tfrac{1}{n}\right)\right)$$

$$\approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (\text{für } n \to \infty)$$

folgt für große n:

$$\mathbb{P}[S_{2n} = a] = \frac{(2n)!}{(n!)^2} p^n (1-p)^n$$

$$\sim \frac{\sqrt{4\pi n}}{2\pi n} \frac{(2n/e)^{2n}}{(n/e)^{n2}} p^n (1-p)^n = \frac{1}{\sqrt{\pi n}} \left(2p(1-p)\right)^n$$

wobei zwei Folgen $(a_n)_n$ und $(b_n)_n$ asymptotisch äquivalent sind, $a_n \sim b_n$, falls $\lim_{n \to \infty} \frac{a_n}{b_n} = 1$.

$\Rightarrow$ Falls $p \neq \frac{1}{2}$ gilt $(4p(1-p)) < 1$ und daher
$\mathbb{P}[S_{2n} = a] \xrightarrow[n\to\infty]{} 0$ exponentiell schnell.

Falls $p = \frac{1}{2}$ $\Rightarrow$ $\mathbb{P}[S_{2n} = a] \approx \frac{1}{\sqrt{\pi n}} \xrightarrow[n\to\infty]{} 0$ (langsam)

(2) Treffzeit Vtlg

Wir betrachten nur den Fall der einfachen Irrfahrt (manchmal auch symmetrische Irrfahrt genannt) mit $p = \frac{1}{2}$.

Da $\{T_b \leq n\} = \bigcup_{k=0}^{n} \{S_k = b\}$ müssen wir für die Vtlg von T_b die gemeinsame Vtlg μ_a der ZV $S(\omega) = (S_0(\omega), S_1(\omega), \dots, S_n(\omega))$:

Lemma 2.52:

μ_a ist die Gleichvtlg auf dem Pfadraum

$$\Omega_{a,n} := \{(s_0, \dots, s_n) : s_0 = a,\ s_k \in \mathbb{Z},\ |s_j - s_{j-1}| = 1,\ 1 \leq j \leq n\}.$$

Beweis: $\mu_a(\{(s_0, \dots, s_n)\}) = \mathbb{P}[S_0 = s_0, S_1 = s_1, \dots, S_n = s_n]$

$= \mathbb{P}[S_0 = s_0, X_1 = s_1 - s_0, X_2 = s_2 - s_1, \dots, X_n = s_n - s_{n-1}]$

X_i i.i.d. ~~…~~ $= \delta_a(s_0) \prod_{k=1}^{n} \mathbb{P}[X_k = s_k - s_{k-1}]$

$$= \begin{cases} 0, & \text{falls } s_0 \neq a \text{ oder } |s_i - s_{i-1}| \neq 1 \text{ für ein } i \in \{1, \dots, n\} \\ 2^{-n} & \text{sonst für } (s_0, \dots, s_n) \in \Omega_{a,n} \end{cases}$$

$\square$

Damit können wir W'keiten für die Irrfahrt durch Abzählen berechnen. Kombiniert mit einer schönen geometrischen Idee, erlaubt dies Treffzeitw'keiten zu berechnen.

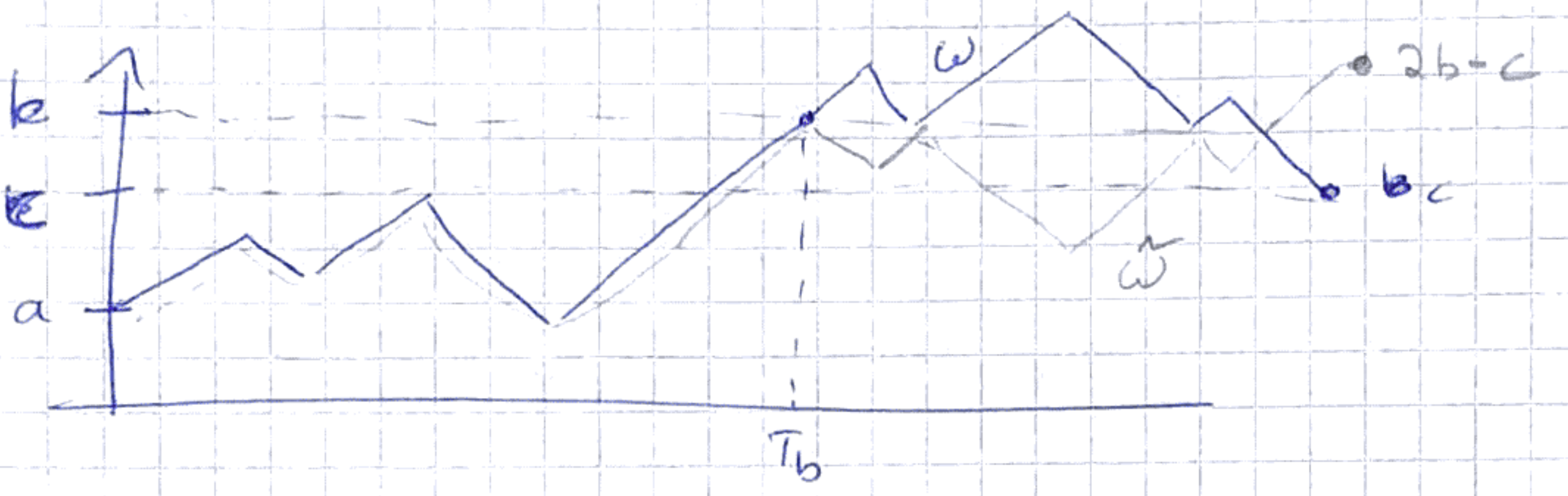
Satz 2.33.

Seien $b, c \in \mathbb{Z}$ mit $b \neq a$.

i) Falls ($a < b$ und $c \geq b$) oder ($a > b$ und $c \leq b$), gilt $P[T_b \leq n, S_n = c] = P[S_n = c]$

ii) Falls ($a < b$ und $c < b$) oder ($a > b$ und $c > b$)

$\Rightarrow P[T_b \leq n, S_n = c] = P[S_n = 2b - c]$



Beweis. i) Folgt aus $\{S_n = c\} \subseteq \{T_b \leq n\}$

ii) Wir nehmen an $a < b$ und $c < b$ (andere Fall analog)

$$(*) \quad P[T_b \leq n, S_n = c] = \frac{\#\text{ Pfade mit } T_b \leq n,\ S_n = c}{2^{-n}}$$

$$(**) \quad P[S_n = 2b - c] = \frac{\#\text{ Pfade mit } S_n = 2b - c}{2^{-n}}$$

Dann $(*) = (**)$, falls eine Bijektion zwischen den Mengen (bzw Pfaden) in $(*)$ und $(**)$ existiert.

Sei $\omega \in (*)$, d.h. $T_b(\omega) \leq n$, $S_n(\omega) = c$. Definiere

$\phi : \omega \to \phi(\omega)$ durch

$$\begin{cases} (\phi(\omega))_k = \omega_k, & k \leq T_b(\omega) \\ (\phi(\omega))_k = \underbrace{2b - \omega_k}_{= b + b - \omega_k}, & k > T_b(\omega) \end{cases}$$

$\Rightarrow \tilde{\omega} = \phi(\omega)$ ist in $(**)$, d.h. $S_n(\tilde{\omega}) = 2b - c$

und $\phi(\phi(\omega)) = \omega \Rightarrow \phi$ ist eine Bijektion. $\square$

Aus der bekannten Vtlg von S_n können wir nun $\mathbb{P}[T_b \leq n]$ berechnen.

Kor 2.34:

i) $$\mathbb{P}[T_b \leq n] = \begin{cases} \mathbb{P}[S_n \geq b] + \mathbb{P}[S_n > b], & b > a \\ \mathbb{P}[S_n \leq b] + \mathbb{P}[S_n < b], & b < a \end{cases}$$

ii) $$\mathbb{P}[T_b = n] = \begin{cases} \frac{1}{2}\mathbb{P}[S_{n-1} = b-1] - \frac{1}{2}\mathbb{P}[S_{n-1} = b+1], & b > a \\ \frac{1}{2}\mathbb{P}[S_{n-1} = b+1] - \frac{1}{2}\mathbb{P}[S_{n-1} = b-1], & b < a \end{cases}$$

Beweis: ObdA $b > a$

i) $$\mathbb{P}[T_b \leq n] = \sum_{c \in \mathbb{Z}} \mathbb{P}[T_b \leq n, S_n = c]$$

$$\overset{\text{Satz 2.33}}{=} \sum_{c \geq b} \mathbb{P}[S_n = c] + \underbrace{\sum_{c < b} \mathbb{P}[S_n = 2b - c]}_{= \sum_{c > b} \mathbb{P}[S_n = c]}$$

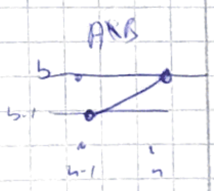
$$\overset{\sigma\text{-Add.}}{=} \mathbb{P}[S_n \geq b] + \mathbb{P}[S_n > b]$$

ii) $$\mathbb{P}[T_b = n] = \mathbb{P}[T_b \leq n] - \mathbb{P}[T_b \leq n-1]$$

$$\overset{i)}{=} \mathbb{P}[S_n \geq b] - \mathbb{P}[S_{n-1} \geq b] + \mathbb{P}[S_n \geq b+1] - \mathbb{P}[S_{n-1} \geq b]$$

Wegen $\mathbb{P}[A] - \mathbb{P}[B] = \mathbb{P}[A \setminus B] - \mathbb{P}[B \setminus A]$ folgt

$$\mathbb{P}[\underbrace{S_n \geq b}_{A}] - \mathbb{P}[\underbrace{S_{n-1} \geq b}_{B}]$$



$$= \mathbb{P}[S_n \geq b, S_{n-1} < b] - \mathbb{P}[S_{n-1} \geq b, S_n < b]$$

$$\overset{(MK)}{=} \frac{1}{2}\mathbb{P}[S_{n-1} = b-1] - \frac{1}{2}\mathbb{P}[S_{n-1} = b]$$

$$\Rightarrow \mathbb{P}[T_b = n] = \frac{1}{2}\Big(\mathbb{P}[S_{n-1} = b-1] - \mathbb{P}[S_{n-1} = b] + \mathbb{P}[S_{n-1} = b] - \mathbb{P}[S_{n-1} = b+1]\Big) \quad \square$$

(52) Vtlg vom Max. mean:

Kor 2.35: Sei $M_n = \max_{0 \le k \le n} S_k$. Für $b > a$,

$$\mathbb{P}[M_n \geq b] = \mathbb{P}[T_b \leq n] = \mathbb{P}[S_n \geq b] + \mathbb{P}[S_n > b]$$

Beweis: $\{M_n \geq b\} = \{T_b \leq n\}$ □.

3. Kovarianz, Korrelation und Gesetz der großen Zahlen

Def 3.1.: Sei $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ ein diskreter W'raum. Dann setzen wir

$$\mathcal{L}^2(\Omega, \mathcal{P}(\Omega), \mathbb{P}) := \{ X : \Omega \to \mathbb{R} : \mathbb{E}(X^2) < \infty \}.$$

(Cauchy-Schwarz Ungleichung)

Lemma 3.2: i) Falls $X, Y \in \mathcal{L}^2(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ gilt $\mathbb{E}[X \cdot Y] \leq \sqrt{\mathbb{E}[X^2]}\sqrt{\mathbb{E}[Y^2]}$

ii) $\mathcal{L}^2(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ ist ein Vektorraum mit Skalarprodukt

$$(X, Y)_{\mathcal{L}^2} = \mathbb{E}(X \cdot Y)$$

(streng genommen ist $(X,Y)_{\mathcal{L}^2}$ nur positiv semi-definit).

Beweis: (i) Mit der Cauchy-Schwarz Ungleichung für Summen folgt:

$$(\mathbb{E}[|X \cdot Y|])^2 = \Big(\sum_{\substack{a \in X(\Omega) \\ b \in Y(\Omega)}} |a \cdot b| \, \mathbb{P}[X=a, Y=b] \Big)^2$$

$$= \Big(\sum_{\substack{a \in X(\Omega) \\ b \in Y(\Omega)}} |a| \sqrt{\mathbb{P}[X=a, Y=b]} \; |b| \sqrt{\mathbb{P}[X=a, Y=b]} \Big)^2$$

$$\leq \sum_{a,b} a^2 \, \mathbb{P}[X=a, Y=b] \sum_{a,b} b^2 \, \mathbb{P}[X=a, Y=b]$$

$$= \sum_a a^2 \, \mathbb{P}[X=a] \sum_b b^2 \, \mathbb{P}[Y=b]$$

$$= \mathbb{E}[X^2] \; \mathbb{E}[Y^2]$$

(ii) $X, Y \in \mathcal{L}^2(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, $a \in \mathbb{R} \Rightarrow aX + Y$ ist ZV mit

$$\mathbb{E}[(aX+Y)^2] = \mathbb{E}[a^2X^2 + Y^2 + 2aXY] \leq 2a^2\mathbb{E}X^2 + \mathbb{E}Y^2 < \infty$$

da $u \cdot v \leq \frac{u^2}{2} + \frac{v^2}{2}$ (Young Ungleichung).

$\Rightarrow \alpha X + Y \in \mathcal{L}^2(\Omega, \mathcal{P}(\Omega), P)$, Bilinearität und Symmetrie ist klar und $(X, X)_{\mathcal{L}^2} = E[X^2] \geq 0 \quad \forall X \in \mathcal{L}^2$. □

Def 3.3: Sei $X, Y \in \mathcal{L}^2(\Omega, \mathcal{P}(\Omega), P)$.

a) Die Kovarianz von X und Y ist definiert als

$$\mathrm{Cov}(X, Y) = E[(X - EX)(Y - EY)]$$
$$= E[XY] - E[X]E[Y]$$

b) Die Standardabweichung von X ist definiert als

$$\sigma(X) = \sqrt{\mathrm{Var}(X)}$$

c) Falls $\sigma(X), \sigma(Y) \neq 0$ ist der Korrelationskoeffizient von X und Y definiert als

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

d) X, Y heißen unkorreliert, falls $\mathrm{Cov}(X, Y) = 0$.
Falls $\mathrm{Cov}(X, Y) > 0$ (< 0) heißen X, Y positiv korreliert (negativ korreliert).

Bem 3.4:

i) Cauchy Schwarz liefert Wohldefiniertheit von $E\, X \cdot Y$.

ii) ——— " ——— $|\rho(X, Y)| \leq 1$

iii) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

iv) Unabhängigkeit $\Rightarrow$ Unkorreliertheit:
$\not\Leftarrow$

denn:

$$E[XY] = \sum_{\substack{a \in X(\Omega) \\ b \in Y(\Omega)}} ab\, P[X = a, Y = b]$$

$$\overset{\text{unab.}}{=} \sum_a a\, P[X = a] \sum_b b\, P[Y = b]$$

$$= EX \quad EY$$

Andererseits: Sei $P[X = 1] = P[X = -1] = P[X = 0] = \frac{1}{3}$
$Y = X^2$

54

$\Rightarrow \mathbb{E}[X \cdot Y] = \mathbb{E}[X^3] = 0$ (antisymmetri)

$\mathbb{E}[X]\mathbb{E}[Y] = 0$ $(\mathbb{E}X = 0)$.

aber $\mathbb{P}[X \in \{-1,1\}, Y = 0] = 0 \neq \mathbb{P}[X \in \{-1,1\}]\,\mathbb{P}[Y=0]$
$= \frac{2}{9}$

Satz 3.5:

Seien $X: \Omega \to S$, $Y: \Omega \to T$ diskrete reelle ZV auf $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.

Dann sind die folgenden Aussagen äquivalent:

i) X und Y sind unabhängig

ii) $f(X)$ und $g(Y)$ sind unkorreliert für alle Funktionen $f: S \to \mathbb{R}$, $g: T \to \mathbb{R}$ mit $f(X), g(Y) \in \mathcal{L}^2(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.

Beweis: i) $\Rightarrow$ ii) wie in Bem 3.4. iv):

$$\mathbb{E}[f(X)g(Y)] = \sum_{\substack{a \in S \\ b \in T}} f(a)g(b)\,\mathbb{P}[X=a, Y=b]$$

$$= \sum_{a \in S} f(a)\,\mathbb{P}[X=a] \sum_{b \in T} g(b)\,\mathbb{P}[Y=b]$$

ii) $\Rightarrow$ i) $\forall a \in S, b \in T$ gilt

$$\mathbb{P}[X=a, Y=b] = \mathbb{E}[1_{\{a\}}(X)\,1_{\{b\}}(Y)]$$

$$\overset{\text{unkorreliert}}{=} \mathbb{E}[1_{\{a\}}(X)]\,\mathbb{E}[1_{\{b\}}(Y)] = \mathbb{P}[X=a]\,\mathbb{P}[Y=b] \quad \square$$

Bem: Für Unabhängigkeit müssen wir also Unkorreliertheit von ~~gege~~ $f(X)$ und $g(Y)$ für "alle" Fkt f, g testen. Für Unkorreliertheit nur für die Fkt $f(X) = X$ und $g(Y) = Y$.

Satz 3.6: Seien $X_1, \dots, X_n \in \mathcal{L}^2(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Dann

$$\mathrm{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \mathrm{Var}(X_i) + 2 \sum_{1 \le k < \ell \le n} \mathrm{Cov}(X_k, X_\ell)$$

Beweis: Kovarianz ist bilinear und $\mathrm{Var}(X) = \mathrm{Cov}(X, X)$. Damit

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i\right) = \sum_{k,l=1}^{n} \mathrm{Cov}(X_k, X_l)$$

$$= \sum_{k=1}^{n} \mathrm{Var}(X_k) + 2 \sum_{1 \le k < l \le n} \mathrm{Cov}(X_k, X_l) \qquad \square$$

3.1 Gesetze der großen Zahlen.