27. April

2) Bedingte Wahrscheinlichkeiten und Unabhängigkeit.

2.1) Bedingte Wahrscheinlichkeit.

Beispiel: Aus Statistiken (hier sind fiktiven Daten) über Kleinkinder, sind 2 Merkmale "Krabbeln" und "Laufen" zeitlich gemessen.

- Seien $A = \{\text{Läuft von 10 Monaten}\}$ und $B = \{\text{Krabbelt vor 6 Monaten}\}$ zwei Ereignisse, mit

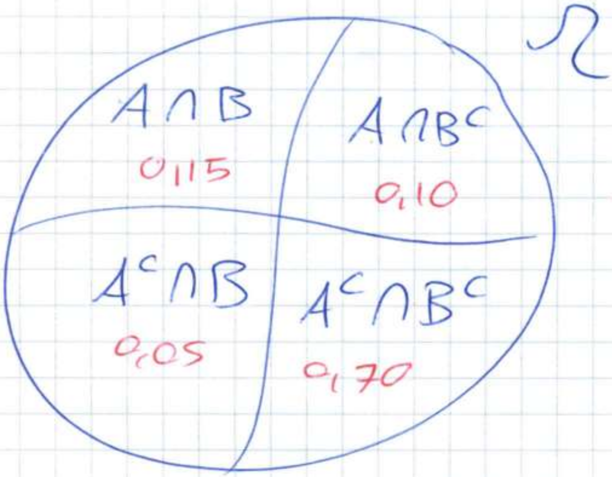
- $\mathbb{P}(\text{Läuft vor 10 Monaten}) = 0{,}25$ und $\mathbb{P}(\text{Krabbelt vor 6 Monaten}) = 0{,}20$.

Frage: Sei nun ein Kind 6 Monaten alt und krabbelt (d.h., B tritt ein). Wie wahrscheinlich ist es, dass in 4 Monaten das Kind schon läuft?

- Um die Frage eine Antwort zu geben braucht man mehr Informationen, z.B.

W-keit	A	A^c
B	0,15	0,05
B^c	0,10	0,70

· Man kann dann die Frage antworten, unter die Annahme, dass das Kind "ein "Zufällige" Kind der Statistik ist".



· W-Keiten, p

· Wir wissen, dass B tritt ein.

$\Rightarrow$ Teilen wir $\Omega = \Omega_B \cup \Omega_B^c$

wobei $\Omega_B = \{\omega \in \Omega \mid \omega \in B\}$.

· Um die Frage zu beantworten, definieren wir eine neue W-Keit auf Ω, mit Gewichte $p(\cdot|B)$, die die Information $\omega \in B$ berücksichtig, d.h.,

$$p(\omega|B) := 0 \quad \text{für alle } \omega \in \Omega_B^c.$$

· Da die Information $\omega \in B$ ist die gleiche $\forall \omega \in \Omega_B$, setzen wir

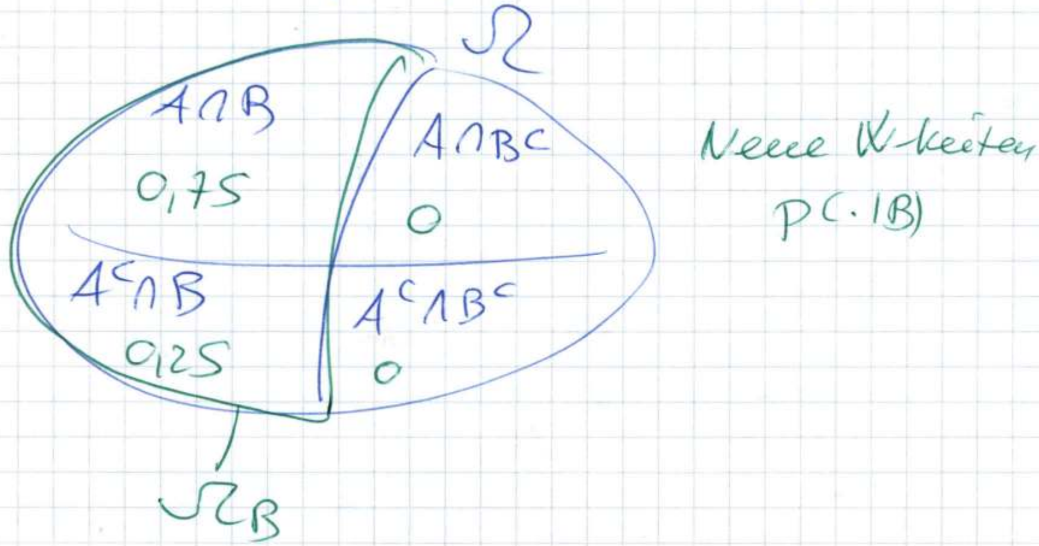
$$p(\omega|B) := \text{const} \cdot p(\omega) \quad \text{für } \omega \in \Omega_B.$$

Die Normierung $\sum_{\omega \in \Omega} p(\omega|B) = \sum_{\omega \in \Omega_B} p(\omega|B) = 1$

liefert dann $const = \frac{1}{\mathbb{P}(B)}$, d.h.,

$$p(\omega|B) = \frac{p(\omega)}{\mathbb{P}(B)}.$$

In unsere Beispiel, $\mathbb{P}(B) = 0{,}2$, deshalb



$\Rightarrow$ Antwort: Mit 75% wird das Kind mit 10 Monaten laufen.

Definition 1) Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein W-raum, $A, B \in \mathcal{F}$ Ereignisse mit $\mathbb{P}(B) \neq 0$. Dann,

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

heisst die bedingte W-keit von A gegeben B.

<u>Bemerkung:</u> ⓐ $\mathbb{P}(\cdot|B): A \longmapsto \mathbb{P}(A|B)$
ist eine W-Verteilung auf $(\Omega, \mathcal{F})$,
die "<u>bedingte Verteilung gegeben B</u>".

ⓑ $\mathbb{P}(\cdot|B)$ hat Gewichte

$$p(\omega|B) = \begin{cases} \frac{p(\omega)}{\mathbb{P}(B)} & , \omega \in B, \\ 0 & , \omega \notin B. \end{cases}$$

ⓒ Sei $X: \Omega \to S' \subset \mathbb{R}$ eine (diskrete) Zufallsvariable mit Verteilung $\mathbb{P}(\cdot|B)$. Dann X hat Erwartungswert

$$\mathbb{E}(X|B) = \sum_{s \in S'} s \cdot \mathbb{P}(X = s|B),$$

die "<u>bedingte Erwartung von X gegeben B</u>".

<u>Beispiel:</u> . N fairen Münzwürfen ergeben n mal die Zahl.
. Was ist die W-keit, dass bei der ersten m Würfe immer die Zahl fällt?

. $\Omega = \{\omega = (x_1, \ldots, x_N) \mid x_k \in \{0,1\}\}$
wobei $x_k = \begin{cases} 1 & , \text{Ausgang } k\text{-te Wurf} = \text{Zahl} \\ 0 & , \quad " \quad = \text{Kopf}. \end{cases}$
mit (wegen "fairen") Gleichverteilung $\mathbb{P}$.

Sei $X_k(\omega) = x_k$, $k = 1, \dots, N$.

$$\Rightarrow \mathbb{P}\left(X_1 = \dots = X_m = 1 \,\Big|\, \sum_{k=1}^{N} x_k = n\right) =$$

$$= \frac{\mathbb{P}\left(X_1 = \dots = X_m = 1, \sum_{k=m+1}^{N} x_k = n-m\right)}{\mathbb{P}\left(\sum_{k=1}^{N} x_k = n\right)}$$

$$= \frac{2^{-N} \cdot \binom{N-m}{n-m}}{2^{-N} \cdot \binom{N}{n}} = \frac{(N-m)!}{N!} \cdot \frac{n!}{(n-m)!}.$$

Bislang: $\mathbb{P} \longrightarrow \mathbb{P}(\cdot \mid B)$.
Es gilt aber auch eine Identität in die andere Richtung.

Satz 2) Sei $\Omega = \bigcup_{k \in I} H_k$ eine disjunkte Zerlegung von Ω in (abzählbar) viele Hypothesen H_k, $k \in I$.

Dann, $\forall A \in \mathcal{F}$,

$$\mathbb{P}(A) = \sum_{\substack{k \in I, \\ \mathbb{P}(H_k) \neq 0}} \mathbb{P}(A \mid H_k)\, \mathbb{P}(H_k).$$

Beweis: $\forall A \in \mathcal{F}$, $A = A \cap \Omega = A \cap \left(\bigcup_{k \in I} H_k\right) \overset{\text{disjunkte}}{=} \bigcup_{k \in I} (A \cap H_k)$

$$\Rightarrow \mathbb{P}(A) \overset{\sigma\text{-add.}}{=} \sum_{k \in I} \mathbb{P}(A \cap H_k) = \sum_{\substack{k \in I \\ \mathbb{P}(H_k) \neq 0}} \mathbb{P}(A \cap H_k)$$ und mit $\mathbb{P}(A \cap H_k) = \mathbb{P}(A \mid H_k)\mathbb{P}(H_k)$

$0 \leq \mathbb{P}(A \cap H_k) \leq \mathbb{P}(H_k)$

#

Beispiel: Urne A enthält 2 rote und 3 blaue Kugeln, Urne B dagegen 3 rote und 2 blaue.

Wir ziehen eine Kugel (K_1) aus A und legen wir in B. Danach ziehen wir eine Kugel (K_2) aus B.

$\mathbb{P}(K_2 \text{ ist rot}) = ?$

A

$B \cup K_1$

ⓐ Falls K_1 rot $\Rightarrow$ 

$B \cup K_1$

ⓑ Falls K_2 blau $\Rightarrow$ 

$$\Rightarrow \mathbb{P}(K_2 \text{ rot}) = \mathbb{P}(K_2 \text{ rot} \mid K_1 \text{ rot})\,\mathbb{P}(K_1 \text{ rot}) + \mathbb{P}(K_2 \text{ rot} \mid K_1 \text{ blau})\,\mathbb{P}(K_1 \text{ blau})$$

$$= \frac{4}{6} \cdot \frac{2}{5} + \frac{3}{6} \cdot \frac{2}{5} = \frac{17}{30}.$$

2.2) Bayessche Regel

- In Bayessche Statistik $\mathbb{P}(H_k)$ ist die "a priori" Einschätzung der W-keit von Hypothese H_k
  (Bsp: H_k = Unfallkosten pro Jahr $\in [k \cdot 100€, (k+1) \cdot 100€)$).

- Aus statistische Daten weiss man auch, dass ein Ereignis $A \in \mathcal{F}$ mit $\mathbb{P}(A) \neq 0$ eintritt
  (Bsp: ein Auffahrunfall)

und dazu kennt man (die empirische)
$\mathbb{P}(A|H_k)$, die bedingte W-keit für das Eintreten von A unter Hypothese H_k.
(Bsp: W-keit, dass beim Auffahrunfall die Unfallkosten in $[k \cdot 100€, (k+1)100€)$ liegen).

. Falls A passiert, wird man (die Versicherungskosten neu berechnen auf der Basis von) $\mathbb{P}(H_k|A)$: die "a posteriori"-Verteilung.

Korollar 3) Für $A \in \mathcal{F}$ mit $\mathbb{P}(A) \neq 0$, gilt es:

$$\mathbb{P}(H_k|A) = \frac{\mathbb{P}(A|H_k)\mathbb{P}(H_k)}{\sum\limits_{\substack{e \in I \\ \mathbb{P}(H_e) \neq 0}} \mathbb{P}(A|H_e) \cdot \mathbb{P}(H_e)}$$

$\forall k \in I$ mit $\mathbb{P}(H_k) \neq 0$.

Beweis: $$\mathbb{P}(H_k|A) \overset{\text{def}}{=} \frac{\mathbb{P}(A \cap H_k)}{\mathbb{P}(A)} \underset{\text{Satz 2}}{\overset{\uparrow}{=}} \frac{\mathbb{P}(A|H_k)\mathbb{P}(H_k)}{\sum\limits_{\substack{e \in I, \\ \mathbb{P}(H_e) \neq 0}} \mathbb{P}(A|H_e) \cdot \mathbb{P}(H_e)} \quad \#$$

Beispiel: Eine Krankheit K tritt mit 10^{-4} Häufigkeit ein.

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

$\Rightarrow$ Apriori: $\mathbb{P}(K) = 10^{-4}$; $\mathbb{P}(K^c) = 1-10^{-4}$

Bed. W-Keiten: $\mathbb{P}(+|K) = 0{,}96$,

$\mathbb{P}(+|K^c) = 0{,}001$

Aposteriori: $\mathbb{P}(K|+) =$

$$= \frac{\mathbb{P}(+|K)\,\mathbb{P}(K)}{\mathbb{P}(+|K)\,\mathbb{P}(K) + \mathbb{P}(+|K^c)\,\mathbb{P}(K^c)}$$

$\cong 1/11$.

$\Rightarrow$ Die W-Keit, dass man krank ist, wenn man positiv getestet ist nur $\approx$ 9%

2.3) Mehrstufige Modelle

. Wir betrachten eine Folge von n (diskrete) Zufallsexperimenten, mit $\Omega_1, \dots, \Omega_n$ die Stichproberäume.

. Damit definiert man ein <u>n-stufiges Zufallsexp.</u>

$$\Omega = \Omega_1 \times \dots \times \Omega_n = \{\omega = (\omega_1, \dots, \omega_n) \mid \omega_k \in \Omega_k, \ k = 1, \dots, n\}$$

und setzen wir

$$\mathcal{F} = \mathcal{P}(\Omega).$$

Dazu setzen wir die Zufallsvariablen

$$X_k(\omega) = \omega_k, \quad k = 1, \dots, n.$$

. Der Index "k" kann als "Zeit" gesehen.

. Falls wir die <u>Anfangsverteilung</u>

$$\mathbb{P}(X_1 = x_1) =: p_1(x_1) \quad , \forall x_1 \in \Omega_1$$

und die bedingten Verteilungen

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

$\forall \ x_\ell \in \Omega_\ell, \ \ell = 1, \dots, k \qquad \mathbb{P}(X_1 = x_1, \dots, X_{k-1} = x_{k-1}) \neq 0,$

dann kann man <u>die</u> Verteilung $\mathbb{P}$ auf Ω konstruieren.

<u>Beispiel:</u> Lotto Ziehung.

Satz 4) Seien $P_1(\cdot)$ und $P_k(\cdot | x_1, \ldots, x_{k-1})$ $\forall k = 2, \ldots, n$, $x_1 \in \Omega_1, \ldots, x_{k-1} \in \Omega_{k-1}$ Massenfunktionen einer W-Verteilung auf Ω_k.

Dann, $\exists!$ W-Verteilung $\mathbb{P}$ auf $(\Omega, \mathcal{F})$ s.d.

(a) $\mathbb{P}(X_1 = x_1) = P_1(x_1)$

(b) $\mathbb{P}(X_k = x_k | X_1 = x_1, \ldots, X_{k-1} = x_{k-1}) = P_k(x_k | x_1, \ldots, x_{k-1})$ $k = 2, \ldots, n$.

. $\mathbb{P}$ hat die Massenfunktion

$$p(x_1, \ldots, x_n) = P_1(x_1) P_2(x_2|x_1) \cdots P_n(x_n | x_1, \ldots, x_{n-1}).$$

Beweis: Eindeutigkeit:

Zu zeigen: $\forall k = 1, \ldots, n$, $\mathbb{P}(X_1 = x_1, \ldots, X_k = x_k) =$
$= p(x_1, \ldots, x_k)$. $(*)$

IA: Für $k = 1$ stimmt.

IV: Sei $(*)$ gültig für $k-1$. Dann:

(a) $p(x_1, \ldots, x_{k-1}) = 0 \Rightarrow 0 = 0$ ✓.

(b) Falls $\mathbb{P}(X_1 = x_1, \ldots, X_{k-1} = x_{k-1}) \neq 0$, dann

$$\mathbb{P}(X_1 = x_1, \ldots, X_k = x_k) = \mathbb{P}(X_1 = x_1, \ldots, X_{k-1} = x_{k-1}) \cdot \mathbb{P}(X_k = x_k | X_1 = x_1, \ldots, X_{k-1} = x_{k-1})$$

$$\stackrel{I.V.}{=} \left(P_1(x_1) \cdots P_{k-1}(x_{k-1} | x_1, \ldots, x_{k-2}) \right) \cdot P_k(x_k | x_1, \ldots, x_{k-1})$$ ✓.

Existenz: Normierung:

$\forall x \in \Omega \Rightarrow x = (x_1, \ldots, x_n)$ mit $x_k \in \Omega_k$, d.h.,

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

$$= \sum_{x_1 \in \Omega_1} P_1(x_1) \cdots \underbrace{\sum_{x_n \in \Omega_n} P_n(x_n | x_1, \ldots, x_{n-1})}_{=1} = \cdots = 1.$$

Dazu, ⓑ gilt weil:

$$\mathbb{P}(X_1 = x_1, \ldots, X_k = x_k) = \sum_{x_{k+1} \in \Omega_{k+1}} \cdots \sum_{x_n \in \Omega_n} p(x_1, \ldots, x_n)$$

$$= \overbrace{P_1(x_1) \cdots P_{k-1}(x_{k-1} | x_1, \ldots, x_{k-2})}^{= \mathbb{P}(X_1 = x_1, \ldots, X_{k-1} = x_{k-1})}$$

$$\cdot P_k(x_k | x_1, \ldots, x_{k-1})$$

$$\Rightarrow \mathbb{P}(X_k = x_k | X_1 = x_1, \ldots, X_{k-1} = x_{k-1}) = P_k(x_k | x_1, \ldots, x_{k-1}).$$

#

. Falls $P_k(x | x_1, \ldots, x_{k-1})$ nur von $x_{k-1}, \ldots, x_{k-m+1}$ hängt, dann sagen wir, dass unsere Modell hat ein "Gedächtnis" von "$m-1$" Schritten.

. 2 Wichtige Klassen von Modellen:

ⓐ $m = -1$: Produktmodelle

ⓑ $m = 0$: Markovketten (MK).

2.3.1) Produktmodelle

. Falls $P_k(x_k | x_1, \ldots, x_{k-1}) = P_k(x_k)$, dann aus Satz 4 folgt:

$$P(x_1, \ldots, x_n) = \prod_{k=1}^{n} P_k(x_k), \quad x = (x_1, \ldots, x_n) \in \Omega.$$

Def. 5) Die W-Verteilung $\mathbb{P}$ auf $\Omega = \Omega_1 \times \dots \times \Omega_n$ mit Massenfunktion

$$p(x_1, \dots, x_n) = \prod_{k=1}^{n} p_k(x_k)$$

heisst Produkt von $\mathbb{P}_1, \dots, \mathbb{P}_n$.

Notation: $\mathbb{P} = \mathbb{P}_1 \otimes \dots \otimes \mathbb{P}_n$.

Beispiel: Seien n unabhängige 0-1-Experimente mit Erfolgswahrscheinlichkeit p,

d.h., $\Omega_1 = \dots = \Omega_n = \{0, 1\}$

mit $p_k(1) = p = 1 - p_k(0)$, $k = 1, \dots, n$.

Dann, $p_k(1) = (1-p) \cdot \left(\frac{p}{1-p}\right)^1$ und

$$p_k(0) = (1-p) \left(\frac{p}{1-p}\right)^0$$

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

heisst die n-dimensionale Bernoulli Verteilung (zum Parameter p).

Bem.: Produktmodelle

$\Updownarrow$

Unabhängige Z.V.

(siehe Kapitel 2.4).

Satz 6) Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Produktmodell. Dann, für beliebige Ereignisse $A_k \subseteq \Omega_k$, $k=1,\dots,n$ gilt:

$$\mathbb{P}(A_1 \times \dots \times A_n) = \prod_{k=1}^{n} \mathbb{P}_k(A_k)$$

(und $\mathbb{P}(A_k) = \mathbb{P}_k(A_k)$).

Beweis: $\mathbb{P}(A_1 \times \dots \times A_n) = \mathbb{P}((X_1,\dots,X_n) \in A_1 \times \dots \times A_n)$

$$= \sum_{(x_1,\dots,x_n) \in A_1 \times \dots \times A_n} p(x_1,\dots,x_n) \overset{\text{Prod. Modell}}{=} \sum_{x_1 \in A_1} \dots \sum_{x_n \in A_n} \prod_{k=1}^{n} p_k(x_k)$$

$$= \prod_{k=1}^{n} \sum_{x_k \in A_k} p_k(x_k) = \prod_{k=1}^{n} \mathbb{P}_k(A_k).$$

Ausserdem, $\mathbb{P}(X_k \in A_k) =$

$$= \mathbb{P}(X_\ell \in \Omega_\ell, \ell \neq k, X_k \in A_k)$$

$$= \underbrace{\prod_{\ell \neq k} \mathbb{P}_\ell(X_\ell \in \Omega_\ell)}_{=1} \cdot \mathbb{P}_k(X_k \in A_k) \quad \#$$

2.3.2) Markovketten (MK)

- Kein Gedächtnis: $m=0$.
- Es ist üblich statt mit 1, mit 0 anzufangen.

Betrachten $\Omega_k = S$ für ein festes (abzählbares) S und

$$\Omega = S^{n+1} = \{(x_0,\dots,x_n) \mid x_k \in S, k=0,\dots,n\}.$$

Da $m=0$: $P_k(x_k|x_1,\dots,x_{k-1}) = P_k(x_k|x_{k-1})$.

Def. 7) Eine Markovkette (MK) ist ein zweistufiges Modell mit

$$P_k(x_k|x_1,\dots,x_{k-1}) = P_k(x_k|x_{k-1}).$$

Def. 8) Eine Matrix $\mathbb{P}(x,y)$, $x,y\in S'$ mit

ⓐ $\mathbb{P}(x,y)\geq 0$, $\forall x,y\in S'$

und

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

heisst eine stochastische Matrix auf S'.

Was ist den Link zwischen MK und stochastische Matrizen?

Lemma 9) Sei $\mathbb{P}_k(x,y) := p_k(y|x)$, ist eine stochastische Matrix.

Beweis: ⓐ V. klar.

ⓑ $$\sum_{y\in S'}\mathbb{P}(x,y) = \sum_{y\in S'}\mathbb{P}(X_k=y|X_{k-1}=x)$$

$$\overset{\sigma\text{-add.}}{=} \mathbb{P}\Big(\underbrace{\bigcup_{y\in S'}\{X_k=y\}}_{=\Omega_k=S'}\Big|X_{k-1}=x\Big) \overset{\text{Normierung}}{=} 1.$$

#

Fragen:
- $\mathbb{P}(X_k = x) = ?$
- $\lim_{k\to\infty} \mathbb{P}(X_k = x) = ?$
- Ist ↗ von x_0 abhängig?

---

4. Mai

Bemerkung: Seien P und Q zwei stochastische Matrizen auf S', dann ist $P \cdot Q$ auch eine stochastische Matrix auf S'.

In der Tat, positivität ist klar und
$$\sum_{y\in S'} (P\cdot Q)(x,y) = \sum_{y\in S'} \sum_{z\in S'} P(x,z)P(z,y)$$
$$= \sum_{z\in S'} P(x,z) \cdot \underbrace{\sum_{y\in S'} P(z,y)}_{=1} = 1.$$

F: Wie berechnet man die Verteilung zur Zeitpunkt n?

Satz 10) Sei μ_0^t den (Zeilen) Vektor mit $p_0(x),\ x\in S'$ als Einträge und $P_1, \dots, P_n$ die "Übergangmatrizen" einer MK auf S'. Dann die Verteilung zur Zeit n ist durch $\mu_n^t(x) := \mathbb{P}(X_n = x) = (\mu_0 P_1 \cdots P_n)(x)$, $x\in S'$ gegeben.

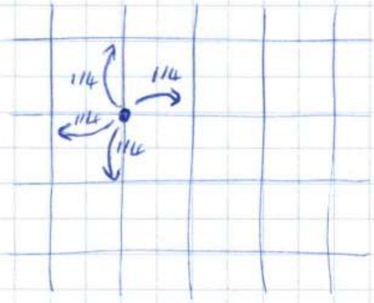
Beweis: Aus Satz 4 und Def 7 gilt
$$P(x_0, \dots, x_n) = p_0(x_0) \cdot \underbrace{P_1(x_1|x_0)}_{=P_1(x_0,x_1)} P_2(x_2|x_1) \cdots \underbrace{P_n(x_n|x_{n-1})}_{=P_n(x_{n-1},x_n)}$$
$$\mu_n^t(x) := \sum_{x_0\in S'} \cdots \sum_{x_{n-1}\in S'} P(x_0, \dots, x_n) = (\mu_0^t P_1 \cdots P_n)(x) \quad \#$$

Beispiele: (a) Produktmodelle sind MK.

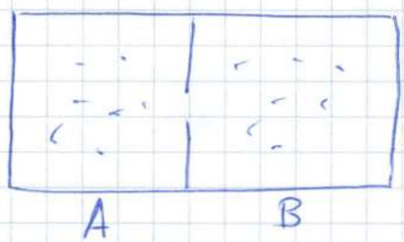
(b) (symmetrische) Random walk auf $S = \mathbb{Z}^d$, $d \in \mathbb{N}$.

(auf Deutsch: "Irrfahrt").

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



(c) Urnenmodell von Ehrenfest



A B

- N Teilchen verteilt in A und B.

Zu jeder Zeit $t \in \mathbb{N}$ wechselt eine zufällig ausgewählte Kugel die Urne.

- Makroskopische Grösse:

Sei n_A = # Kugeln in Urne A, dann eine Makroskopische Grösse ist

$$S_A := \frac{n_A}{N} \in \{0, \tfrac{1}{N}, \tfrac{2}{N}, \dots, \tfrac{N-1}{N}, 1\} = S'$$

Die Markovkette mit Übergangsmatrix

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

beschreibt die (diskrete) Zeitentwicklung von S_A.

. <u>Mikroskopisches Modell:</u>

. Sei $S' = \{0,1\}^N = \{(\sigma_1, \dots, \sigma_N) = \sigma \mid \sigma_k \in \{0,1\}\}$

wobei $\sigma_k = \begin{cases} 1, & \text{falls Teilchen } k \text{ ist in Urne } A, \\ 0, & \text{sonst}. \end{cases}$

Dann, die MK die die Zeitentwicklung von $\sigma \in S'$ beschreibt ist gegeben durch

$$\mathbb{P}(\sigma, \tilde{\sigma}) = \begin{cases} \frac{1}{N}, & \text{falls } \underbrace{\sum_{k=1}^{N} |\sigma_k - \tilde{\sigma}_k| = 1}_{\text{d.h., } \|\sigma - \tilde{\sigma}\| = 1.}, \\ 0, & \text{sonst}. \end{cases}$$

$\Rightarrow$ Das ist ein <u>Random Walk auf dem Hyperwürfel</u> $\{0,1\}^N$.

. <u>Frage:</u> Wie berechnet man die Übergangs W-keit zwischen zwei Zeiten $0 \le k < \ell \le n$?

<u>Satz 11)</u> <u>(Markov Eigenschaft)</u>

Für alle $0 \le k < \ell \le n$ und $x_0, \dots, x_\ell \in S'$ mit $\mathbb{P}(X_0 = x_0, \dots, X_k = x_k) \neq 0$ gilt:

$$\mathbb{P}(X_\ell = x_\ell \mid X_0 = x_0, \dots, X_k = x_k) = \mathbb{P}(X_\ell = x_\ell \mid X_k = x_k)$$
$$= (P_{k+1} \cdots P_\ell)(x_k, x_\ell).$$

Beweis: $\mathbb{P}(X_\ell = x_\ell \mid X_0 = x_0, \ldots, X_k = x_k) =$

$$= \frac{\mathbb{P}(X_0 = x_0, \ldots, X_k = x_k, X_\ell = x_\ell)}{\mathbb{P}(X_0 = x_0, \ldots, X_k = x_k)}$$

$$= \frac{\sum\limits_{x_{k+1}, \ldots, x_{\ell-1} \in S} \cancel{P_0(x_0) P_1(x_0, x_1) \cdots P_k(x_{k-1}, x_k)}\, P_{k+1}(x_k, x_{k+1}) \cdots P_\ell(x_{\ell-1}, x_\ell)}{\cancel{P_0(x_0) P_1(x_0, x_1) \cdots P_k(x_{k-1}, x_k)}}$$

$$= (P_{k+1} \cdots P_\ell)(x_k, x_\ell).$$

und $\mathbb{P}(X_\ell = x_\ell \mid X_k = x_k) = \dfrac{\mathbb{P}(X_\ell = x_\ell, X_k = x_k)}{\mathbb{P}(X_k = x_k)}$

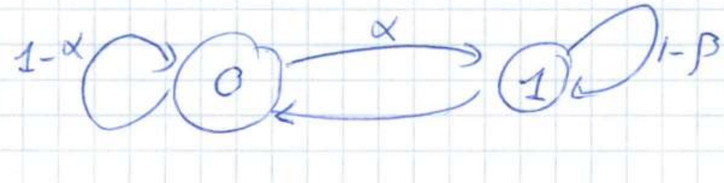
$$\underset{\text{Satz 10}}{=} \sum_{x_{k+1}, \ldots, x_{\ell-1} \in S} \frac{\cancel{(\mu_0^t P_1 \cdots P_k)(x_k)} \cdot P_{k+1}(x_k, x_{k+1}) \cdots P_\ell(x_{\ell-1}, x_\ell)}{\cancel{(\mu_0^t P_1 \cdots P_k)(x_k)}}$$

$$= (P_{k+1} \cdots P_\ell)(x_k, x_\ell). \quad \#$$

Bem.: Falls $P_k = P \ \forall k$, dann die MK ist <u>homogen</u>.

Beispiel 1: $S = \{0, 1\}$ und $\alpha, \beta \in (0, 1)$.

Graphische Darstellung



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

Um P^n zu berechnen kann man:

ⓐ → (falls möglich) P diagonalisieren

ⓑ → eine Formel finden und durch Induktion beweisen.

58

. Fangen wir von (b):

. <u>Man weiss:</u> $\left.\begin{array}{l} P^n(0,0) + P^n(0,1) = 1 \\ P^n(1,0) + P^n(\phi,1) = 1 \end{array}\right\} (*)$

weil P^n ist eine stochastische Matrix.

. Dazu: $P^n(0,0) = P^{n-1}(0,0)\,P(0,0) + P^{n-1}(0,1)\,P(1,0)$

$$= P^{n-1}(0,0)\cdot(1-\alpha) + \underbrace{P^{n-1}(0,1)}_{=1-P^{n-1}(0,0)}\cdot\beta$$

$$= P^{n-1}(0,0).(1-\alpha-\beta) + \beta$$

$$\Rightarrow \boxed{P^n(0,0) =} P^{n-2}(0,0).(1-\alpha-\beta)^2 + \beta(1-\alpha-\beta) + \beta$$

$$= \dots = P^0(0,0)(1-\alpha-\beta)^n + \sum_{k=0}^{n-1} \beta\,(1-\alpha-\beta)^k$$

$$= (1-\alpha-\beta)^n + \beta\cdot\frac{1-(1-\alpha-\beta)^n}{1-(1-\alpha-\beta)}$$

$$= \frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^n\cdot\left(1-\frac{\beta}{\alpha+\beta}\right)$$

$$= \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}\,(1-\alpha-\beta)^n.$$

Zusammen mit (*) und, dass die Formel für $P^n(1,1)$ ist die Gleiche aber mit α,β vertauscht, findet man: Exponentiell schnell Verfall!

$$P^n = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix} + (1-\alpha-\beta)^n\cdot\begin{pmatrix} \frac{\alpha}{\alpha+\beta} & -\frac{\alpha}{\alpha+\beta} \\ -\frac{\beta}{\alpha+\beta} & \frac{\beta}{\alpha+\beta} \end{pmatrix}$$

<u>Insbesondere:</u> $\lim_{n\to\infty} P^n = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix}$ weil $|1-\alpha-\beta| < 1$.

Bem.: . Da $\lim_{n\to\infty} \mathbb{P}^n$ ist eine Matrix mit Gleiche Zeilen ist kein Zufall.

. Es folgt, dass

$$\lim_{n\to\infty} \mathbb{P}(X_n=0) = \left(P_0(0) \;\; P_0(1)\right)\begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix}$$

$$\underset{P_0(0)+P_0(1)=1}{=} \left(\frac{\beta}{\alpha+\beta} \;\; \frac{\alpha}{\alpha+\beta}\right)$$ ist unabhängig von der Anfangsbedienung $P_0(0), P_0(1)$.

. Methode ⓐ: (Skip erthl.) . Eigenwerte von $\mathbb{P}$:

$$\det(\lambda \mathbb{1} - \mathbb{P}) = \det\begin{pmatrix} \lambda-(1-\alpha) & -\alpha \\ -\beta & \lambda-(1-\beta) \end{pmatrix}$$

$$= \lambda^2 - (2-\alpha-\beta)\lambda - \alpha^2 + (1-\alpha)(1-\beta) = 0$$

$$\Leftrightarrow \begin{cases} \lambda_1 = 1 & \longrightarrow \vec{v}_1 = (1,1)^t \\ \lambda_2 = 1-\alpha-\beta & \longrightarrow \vec{v}_2 = (\alpha, -\beta)^t \end{cases}$$

Bem.: . $\lambda_1 = 1$ ist immer dabei, weil $\mathbb{P}$ eine Stochastische Matrix.
In der Tat, $\vec{v} = (1,1,\dots,1)^t$ ist ein Eigenvektor von $\mathbb{P}$ mit Eigenwert 1:

$$.(\mathbb{P}\cdot\vec{v})(x) = \sum_{y\in S} \mathbb{P}(x,y)\underbrace{\vec{v}(y)}_{=1} = 1 = \vec{v}(x).$$

. Da die Eigenwerte (alle) verschiedene sind, kann man $\mathbb{P}$ diagonalisieren:

Sei $U = (v_1 \mid v_2) = \begin{pmatrix} 1 & \alpha \\ 1 & -\beta \end{pmatrix} \Rightarrow \mathbb{P} = U\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}U^{-1}$

$$\Rightarrow P^n = U \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} U^{-1}$$

$$= \begin{pmatrix} 1 & \alpha \\ 1 & -\beta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{pmatrix} \cdot \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{1}{\alpha+\beta} & -\frac{1}{\alpha+\beta} \end{pmatrix}.$$

2.4) Unabhängigkeit von Ereignisse

- Seien A und B zwei Ereignisse und $(\Omega, \mathcal{F}, \mathbb{P})$ den W-raum.

- A und B sind unabhängig, falls die Information dass A eintritt hat keinen Einfluss an die W-keit dass B eintritt, und viceversa.

$\Rightarrow \quad \mathbb{P}(A|B) = \mathbb{P}(A) \quad$ falls $\mathbb{P}(B) \neq 0$
und $\mathbb{P}(B|A) = \mathbb{P}(B) \quad$ falls $\mathbb{P}(A) \neq 0$.

$\Rightarrow \quad \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

Def. 12) Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein W-raum.

(a) $A, B \in \mathcal{F}$ sind <u>unabhängig</u> (bzgl. $\mathbb{P}$) falls
$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

(b) Eine Familie A_k, $k \in I$, von Ereignissen heisst unabhängig, falls
$$\mathbb{P}(A_{i_1} \cap \ldots \cap A_{i_n}) = \prod_{k=1}^{n} \mathbb{P}(A_{i_k})$$
$\forall\, n \leq |I|$ und paarweise verschiedenen $i_1, \ldots, i_n \in I$.

Beispiele: (a) Falls $\mathbb{P}(A) \in \{0,1\}$, dann ist A unabhängig von B, $\forall B \in \mathcal{F}$.

In der Tat: • **Falls** $\mathbb{P}(A)=0$, dann

$$0 \leq \mathbb{P}(A \cap B) \leq \mathbb{P}(A) = 0$$

$$\Rightarrow \mathbb{P}(A \cap B) = 0 = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

• **Falls** $\mathbb{P}(A) = 1$, dann

$$\mathbb{P}(A \cap B) + \underbrace{\mathbb{P}(A^c \cap B)}_{=0} = \mathbb{P}(B) = \mathbb{P}(B) \cdot \mathbb{P}(A).$$

(b) Wir werfen 3 mal eine faire Münze.

. $\Omega = \{\underset{\substack{\uparrow \\ \text{Kopf}}}{0}, \underset{\substack{\uparrow \\ \text{Zahl}}}{1}\}^3$, $\mathbb{P}$ = Gleichverteilung.

. $X_k(\omega) = \omega_k$: k-te Würf.

. Betrachten wir die Ereignisse:

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

Dann, $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = \frac{4}{8} = \frac{1}{2}$.

Dazu, $\mathbb{P}(A_1 \cap A_2) = \frac{2}{8} = \frac{1}{4} = \mathbb{P}(A_1)\mathbb{P}(A_2)$

(ähnlich mit $(1,2) \to (2,3)$; $(1,2) \to (3,1)$).

$\Rightarrow$ Man sagt, dass A_1, A_2, A_3 sind paarweise unabhängig.

Aber: $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \frac{2}{8} = \frac{1}{4} \neq \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3) = \frac{1}{8}$

9. May

Lemma 13) Seien $A_1,\dots,A_n \in \mathcal{F}$ unabhängig und $B_k = A_k$ oder $B_k = A_k^c$, $\forall k=1,\dots,n$.

Dann sind $B_1,\dots,B_n$ unabhängig.

Beweis: O.b.d.A., $B_1 = A_1,\dots,B_\ell = A_\ell$, $B_{\ell+1} = A_{\ell+1}^c,\dots,B_n = A_n^c$.

$$\Rightarrow \mathbb{P}(B_1\cap\dots\cap B_n) = \mathbb{P}\left(\left(\bigcap_{k=1}^{\ell} A_k\right)\cap\left(\bigcap_{m=\ell+1}^{n} A_m^c\right)\right)$$

$$= \mathbb{E}\Big(\prod_{k=1}^{\ell} \mathbb{1}_{A_k}\cdot\underbrace{\prod_{m=\ell+1}^{n}(1-\mathbb{1}_{A_m})}_{=\sum_{J\subseteq\{\ell+1,\dots,n\}}(-1)^{|J|}\prod_{j\in J}\mathbb{1}_{A_j}}\Big)$$

$$\overset{\text{Linearität}}{=} \sum_{J\subseteq\{\ell+1,\dots,n\}}(-1)^{|J|}\underbrace{\mathbb{E}\Big(\prod_{k=1}^{\ell}\mathbb{1}_{A_k}\prod_{j\in J}\mathbb{1}_{A_j}\Big)}_{\overset{\text{unab. von } A\text{'s}}{=}\prod_{k=1}^{\ell}\mathbb{P}(A_k)\prod_{j\in J}\mathbb{P}(A_j)}$$

$$= \prod_{k=1}^{\ell}\underbrace{\mathbb{P}(A_k)}_{=\mathbb{P}(B_k)}\prod_{m=\ell+1}^{n}\underbrace{(1-\mathbb{P}(A_m))}_{=\mathbb{P}(A_m^c)=\mathbb{P}(B_m)}$$

$$= \prod_{k=1}^{n}\mathbb{P}(B_k). \quad \#$$

2.4.1) Verteilung der Wartezeit

- Seien $A_1, A_2,\dots,A_L \in \mathcal{F}$ unabhängige Ereignisse bzgl. $\mathbb{P}$ mit $\mathbb{P}(A_k) = p \in [0,1]$.
- Betrachten wir die Zufallsvariable

$$T_L(\omega) = \inf\{n\in\{1,\dots,L\} \mid \omega\in A_n\}$$
mit $\inf\emptyset := \infty$

. $T_L(\omega)$ ist die Wartezeit auf das erste Eintreten der Ereignisse $A_1, \ldots, A_L$.

Was ist die Verteilung von T_L?

Für $n \leq L$: $\mathbb{P}(T_L = n) = \mathbb{P}(A_1^c \cap \cdots \cap A_{n-1}^c \cap A_n)$

$$\overset{\text{Lem 13}}{=} \mathbb{P}(A_n) \cdot \prod_{k=1}^{n-1} \mathbb{P}(A_k^c)$$

$$= p \cdot (1-p)^{n-1}$$

und für $n = \infty$: $\mathbb{P}(T_L = \infty) = \mathbb{P}(A_1^c \cap \cdots \cap A_L^c) = (1-p)^L$.

. Im Limes $L \to \infty$, sei $T = \lim_{L \to \infty} T_L$.

Dann

$$\mathbb{P}(T = n) = p(1-p)^{n-1}, \quad n \geq 1$$

Bemerkung: . Die Verteilung mit Massenfunktion

$$p(n) = p(1-p)^{n-1}, \quad n \geq 1$$

ist eine geometrische Verteilung zum Parameter p.

. Die Verteilung mit Massenfunktion

$$p(n) = (1-p)p^n, \quad n \geq 0$$

heisst ebenfalls geometrische Verteilung zum Parameter p.

$\Rightarrow$ Es gibt zwei "standards".

. $\mathbb{P}(T \geq n) = \mathbb{P}(A_1^c \cap \cdots \cap A_{n-1}^c) = (1-p)^{n-1}$

• Was ist den Erwartungswert von T?

$$\mathbb{E}(T) \underset{\substack{\uparrow \\ \text{(Siehe Übungen)}}}{=} \sum_{n\geq 1} \mathbb{P}(T \geq n) = \sum_{n \geq 1} (1-p)^{n-1} \underset{\substack{\uparrow \\ \text{geom. Reihe}}}{=} \frac{1}{1-(1-p)} = \frac{1}{p}.$$

2.4.2) Verteilung der eingetretene Ereignisse

• Seien $A_1, \dots, A_n \in \mathcal{F}$ unabhängig mit $\mathbb{P}(A_k) = p$.
und die Zufallsvariable

$$S(\omega) = |\{1 \leq k \leq n \mid \omega \in A_k\}| = \sum_{k=1}^{n} \mathbb{1}_{A_k}(\omega),$$

die die Anzahl der eingetretene Ereignisse zählt.

Dann, $\mathbb{P}(S = k) = \binom{n}{k} \cdot p^k (1-p)^{n-k}$.

[Das haben wir schon nach Def. 12, Kap 1 gezeigt.

$$\mathbb{P}(S_n = k) = \sum_{\substack{I \subseteq \{1,\dots,n\} \\ |I| = k}} \mathbb{P}\left(\left(\bigcap_{j \in I} A_j \right) \cap \left(\bigcap_{\ell \notin I} A_\ell^c \right) \right)$$

$$\overset{\text{Lem. 13}}{=} \sum_{\substack{I \subseteq \{1,\dots,n\} \\ |I| = k}} \prod_{j \in I} \underbrace{\mathbb{P}(A_j)}_{= p} \prod_{\ell \notin I} \underbrace{\mathbb{P}(A_\ell^c)}_{= (1-p)}$$

$$= \sum_{\substack{I \subseteq \{1,\dots,n\} \\ |I| = k}} p^k (1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}.]$$

• Wir haben schon gesehen, dass $\mathbb{E}(S_n) = n \cdot p$.

$\Rightarrow \mathbb{E}\left(\frac{S_n}{n}\right) = p \quad \forall n.$

Das folgende Satz, gibt uns Informationen über die W-keit, dass $\frac{S_n}{n}$ aus p abweicht.

Satz 14) $\forall \varepsilon > 0$ und $n \in \mathbb{N}$ gilt:

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

und $$\mathbb{P}\left(\frac{S_n}{n} \leq p - \varepsilon\right) \leq e^{-2\varepsilon^2 n}.$$

Insbesondere gilt:

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

Bem.: ⓐ Die relative Häufigkeit bei n unabhängige Stichproben ist $\frac{S_n}{n} \approx p$ für grosse n.

Beweis: Setze $q = 1 - p$.

$\forall \lambda > 0$:
$$\mathbb{P}(S_n \geq n(p+\varepsilon)) = \sum_{k \geq np + n\varepsilon} \binom{n}{k} p^k q^{n-k}$$
$$\leq \sum_{k \geq n(p+\varepsilon)} \binom{n}{k} p^k q^{n-k} \cdot e^{\lambda(k - n(p+\varepsilon))}$$
$$\leq \sum_{k=0}^{n} \binom{n}{k} (pe^{\lambda})^k \cdot q^{n-k} e^{-n(p+\varepsilon)\lambda}$$
$$\leq (pe^{\lambda} + q)^n (e^{-\lambda p})^n \cdot e^{-n\lambda\varepsilon}$$
$$= (pe^{\lambda q} + qe^{-\lambda p})^n \cdot e^{-\lambda n \varepsilon}$$

Behauptung: $pe^{\lambda q} + qe^{-\lambda p} \leq e^{\frac{\lambda^2}{8}}$ $(*)$

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

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 $\tilde{S}_n = n - S_n \Rightarrow \tilde{S}_n \sim \mathrm{Bin}(n,q)$.

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

Noch zu zeigen, ⊛.

- Sei $f(\lambda) = \ln(pe^{\lambda q} + q e^{-\lambda p})$

$$\Rightarrow f(\lambda) = -\lambda p + \ln(pe^{\lambda} + q).$$

- $f(0) = 0$ weil $p+q=1$.
- $f'(\lambda) = -p + \dfrac{pe^{\lambda}}{pe^{\lambda}+q} = -p + \dfrac{p}{p+qe^{-\lambda}}$ und $f'(0)=0$
- $f''(\lambda) = \dfrac{p \cdot q \cdot e^{-\lambda}}{(p+qe^{-\lambda})^2}$

Wegen

$$\left.\begin{array}{l}(a+b)^2 = a^2+b^2+2ab \\ (a-b)^2 = a^2+b^2-2ab \geq 0\end{array}\right\} \Rightarrow (a+b)^2 \geq 4ab,$$

$$f''(\lambda) \leq \frac{p \cdot q \cdot e^{-\lambda}}{4 p \cdot q \cdot e^{-\lambda}} = \frac{1}{4}.$$

Aber: $f(\lambda) = f(0) + \int_0^{\lambda} f'(x)\,dx$

$$= \int_0^{\lambda} dx \int_0^{x} dy\, f''(y) \leq \int_0^{\lambda} dx\, \frac{x}{4} = \frac{\lambda^2}{8},\ \forall \lambda \geq 0.$$

#

Do: Beispiel/"Illustration" mit Mathematica.

2.5) Unabhängige ZV und Random Walk.

Def. 15) Seien $X_k : \Omega \to S'_k$, $k = 1, \dots, n$ diskrete ZV auf $(\Omega, \mathcal{F}, \mathbb{P})$.

Die Verteilung $\mu_{X_1, \dots, X_n}$ des Zufallsvektor $(X_1, \dots, X_n)$ heisst gemeinsame Verteilung der ZV. $X_1, \dots, X_n$ und hat Massenfunktion

$$p_{X_1, \dots, X_n}(a_1, \dots, a_n) = \mathbb{P}(X_1 = a_1, \dots, X_n = a_n).$$

Def. 16) Die diskreten ZV. $X_1, \dots, X_n$ heissen unabhängig, falls gilt:

$$\mathbb{P}(X_1 = a_1, \dots, X_n = a_n) = \prod_{k=1}^{n} \mathbb{P}(X_k = a_k),$$

$\forall a_k \in S_k$, $k = 1, \dots, n$.

Satz 17) Die folgende Aussagen sind äquivalent:

(a) $X_1, \dots, X_n$ sind unabhängig,

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

(c) $\mu_{X_1 \dots X_n} = \bigotimes_{k=1}^{n} \mu_{X_k}$.

(d) Die Ereignisse $\{X_1 \in A_1\}, \dots, \{X_n \in A_n\}$ sind unabhängig $\forall A_k \subset S'_k$, $k = 1, \dots, n$.

(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): Folgt aus der Def. 15.

(b) $\Leftrightarrow$ (c): Folgt aus der Def. von $\bigotimes_{k=1}^{n} \mu_{X_k}$.

(c) $\Rightarrow$ (d): Seien $1 \leq i_1 < i_2 < \dots < i_m \leq n$ und $A_{i_k} \subseteq S'_{i_k}$, $k = 1, \dots, m$. Setzen wir $A_i := \Omega$ für $i \notin \{i_1, \dots, i_m\}$. Dann aus (c) folgt (zusammen mit Satz 4):

$$\mathbb{P}(X_{i_1} \in A_{i_1}, \dots, X_{i_m} \in A_{i_m}) = \mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n)$$

$$= \mu_{X_1, \dots, X_n}(A_1 \times \dots \times A_n)$$

$$= \prod_{k=1}^{n} \mu_{X_k}(A_k) = \prod_{k=1}^{n} \mathbb{P}(X_k \in A_k)$$

$$= \prod_{k=1}^{m} \mathbb{P}(X_{i_k} \in A_{i_k})$$

(d) $\Rightarrow$ (e) $\Rightarrow$ (a): klar. #

Def. 18) Eine (beliebige) Kollektion $X_k : \Omega \to S_k$, $k \in I$, von diskreten ZV heisst unabhängig, falls die Ereignisse

$\{X_k = a_k\}, k \in I, \forall a_k \in S_k,$

unabhängig sind.

2.5.1) Der Random Walk auf $\mathbb{Z}$.

. Seien $X_1, X_2, \dots$ iid Z.V. auf $(\Omega, \mathcal{F}, \mathbb{P})$, mit

$$\begin{cases} \mathbb{P}(X_k = 1) = p \\ \mathbb{P}(X_k = -1) = 1-p \end{cases}, \quad p \in (0,1).$$

. Betrachten wir die MK auf $\mathbb{Z}$ mit

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

Dann, $\quad S_n = \sum_{k=1}^{n} X_k + a.$

. $(S_0, S_1, \dots)$ ist ein <u>Random Walk</u> auf $\mathbb{Z}$

mit S_n = Position zur Zeit n.

<u>Fragen:</u> (1) <u>Rückkehrwahrscheinlichkeit</u>.

. $\mathbb{P}(S_n = a) = ?$

(2) <u>Trittzeitverteilung</u>.

. Sei $\lambda \in \mathbb{Z}$ und

$$T_\lambda(\omega) := \inf\{n \geq 1 \mid S_n(\omega) = \lambda\}$$

mit $\inf \emptyset := \infty$.

$\Rightarrow \mathbb{P}(T_\lambda \leq n) = ?$

. Für $a = \lambda$, T_λ ist die erste Rückkehrzeit.

. Für $a \neq \lambda$, T_λ ist die erste Treffzeit von λ.

(3) Verteilung vom Maximum.

. Sei $M_n := \max_{0 \le k \le n} S_n$.

$\Rightarrow \mathbb{P}(M_n \le \lambda) = ?$

Zuerst berechnen wir die Verteilung von S_n.

<u>Prop. (9)</u> Für $k \in \mathbb{Z}$,

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

<u>Beweis:</u> $S_n = a + k \Longleftrightarrow X_1 + \dots + X_n = k$

$\Longleftrightarrow \frac{n+k}{2}$-mal $X_e = +1$

und $\frac{n-k}{2}$-mal $X_e = -1$. #

(1) RückkehrW-keit.

. $\mathbb{P}(S_{2n-1} = a) = 0$ für $n \in \mathbb{N}$

. $\mathbb{P}(S_{2n} = a) = \frac{(4p(1-p))^n}{\sqrt{\pi n}} \left(1 + O\left(\frac{1}{n}\right)\right)$

für $n \in \mathbb{N}$.

. In der Tat, aus Stirling Formula,

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

folgt:

$$\mathbb{P}(S_{2n}=0) = \binom{2n}{n} p^n (1-p)^n$$

$$= \frac{(2n)!}{(n!)^2} \cdot p^n (1-p)^n$$

$$= \frac{\sqrt{4\pi n}\, 2^{2n} n^{2n} e^{-2n}}{\left(\sqrt{2\pi n}\, n^n e^{-n}\right)^2} \cdot p^n (1-p)^n \left(1 + O\left(\tfrac{1}{n}\right)\right)$$

$$= \frac{(4p(1-p))^n}{\sqrt{\pi n}} \cdot \left(1 + O\left(\tfrac{1}{n}\right)\right).$$

Bemerkung: . Falls $p \neq \frac{1}{2} \Rightarrow \mathbb{P}(S_{2n}=0) \xrightarrow[n\to\infty]{} 0$

exponentiell schell, d.h.,

$$c := \lim_{n\to\infty} -\frac{\ln(\mathbb{P}(S_{2n}=0))}{n} \in (0,\infty)$$

$$\left(= -\ln(4p(1-p))\right).$$

. Falls $p = \frac{1}{2} \Rightarrow \mathbb{P}(S_{2n}=0) \simeq \frac{1}{\sqrt{\pi n}} \xrightarrow{n\to\infty} 0$

sehr langsam.

2.5.2) Symmetrischer R.W.

. $P = \frac{1}{2}$.

. Wir wollen T_λ berechnen.

Es gilt: $\mathbb{P}(T_\lambda \leq n) = \mathbb{P}\left(\bigcup_{k=1}^{n} \{S_k = \lambda\}\right)$

$\Rightarrow$ Brauchen wir die gemeinsame Verteilung, μ_a, von $S^!(\omega) := (S_0(\omega), S_1(\omega), \dots, S_n(\omega))$.

Lemma 20) μ_a ist die Gleichverteilung auf dem Pfadraum

$$\tilde{\Omega}_{a,n} := \{(s_0, \dots, s_n) \mid s_0 = a, s_k \in \mathbb{Z} \text{ mit } |s_k - s_{k-1}| = 1, \ k = 1, \dots, n\}.$$

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

$$= \mathbb{P}(S_0 = s_0, X_1 = s_1 - s_0, \dots, X_n = s_n - s_{n-1})$$

$$\overset{\text{Uk}}{=} \delta_{s_0, a} \cdot \prod_{k=1}^{n} \mathbb{P}(X_k = s_k - s_{k-1})$$

$$= \begin{cases} 0, & \text{für } (s_0, \dots, s_n) \notin \tilde{\Omega}_{a,n}, \\ \frac{1}{2^n}, & \text{für } (s_0, \dots, s_n) \in \tilde{\Omega}_{a,n}. \end{cases} \quad \#$$

Die Treffzeit wird mit Hilfe vom Reflexionsprinzip berechnet:

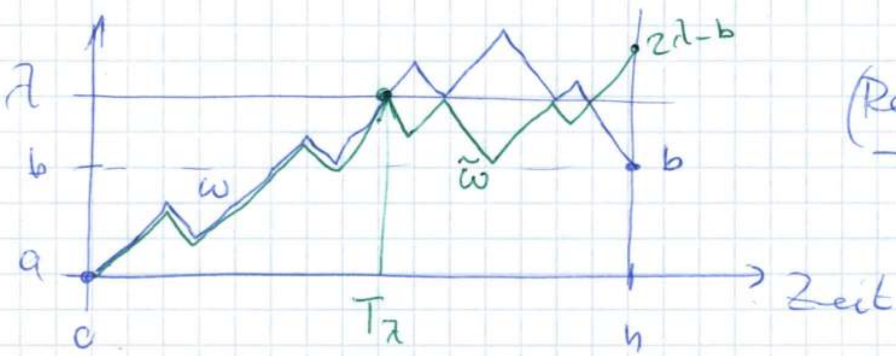
Satz 21) Seien $\lambda, b \in \mathbb{Z}$ mit $\lambda \neq a$.

(a) Falls ($a < \lambda$ und $b \geq \lambda$) oder ($a > \lambda$ und $b \leq \lambda$)

$$\Rightarrow \mathbb{P}(T_\lambda \leq n, S_n = b) = \mathbb{P}(S_n = b)$$

(b) Falls ($a < \lambda$ und $b \leq \lambda$) oder ($a > \lambda$ und $b \geq \lambda$)

$$\Rightarrow \mathbb{P}(T_\lambda \leq n, S_n = b) = \mathbb{P}(S_n = 2\lambda - b).$$



(Reflexionsprinzip)

Beweis: (a) Klar, weil $\{T_\lambda \le n\} \subseteq \{S_n = b\}$.

(b) Fall $a < \lambda$, $b \le \lambda$; der andere ist analog.

$(*)$ $\mathbb{P}(T_\lambda \le n, S_n = b) = \dfrac{\#\text{ Pfade mit } T_\lambda \le n, S_n = b}{2^n}$.

$(**)$ $\mathbb{P}(S_n = 2\lambda - b) = \dfrac{\#\text{ Pfade mit } S_n = 2\lambda - b}{2^n}$.

. $(*) = (**)$ falls $\exists$ Bijektion zwischen Pfaden in $(*)$ und in $(**)$.

. Sei ω in $(*)$. Definiere

$$\Phi : \omega \to \Phi(\omega)$$

mit $\begin{cases} (\Phi(\omega))_k = \omega_k, & k \le T_\lambda \\ (\Phi(\omega))_k = 2\lambda - \omega_k, & k > T_\lambda . \end{cases}$

$\Rightarrow \tilde{\omega} = \Phi(\omega)$ ist in $(**)$ und $\Phi(\Phi(\omega)) = \omega$. #

$\Rightarrow \Phi$ ist eine Bijektion.

- Aus Satz 21, wenn wir über $b \in \mathbb{Z}$ summieren erhalten wir $\mathbb{P}(T_\lambda \leq n)$.

Kor. 22) ⓐ $\mathbb{P}(T_\lambda \leq n) = \begin{cases} \mathbb{P}(S_n \geq \lambda) + \mathbb{P}(S_n > \lambda), & \lambda > a, \\ \mathbb{P}(S_n \leq \lambda) + \mathbb{P}(S_n < \lambda), & \lambda < a, \end{cases}$

ⓑ $\mathbb{P}(T_\lambda = n) = \begin{cases} \frac{1}{2}\mathbb{P}(S_{n-1} = \lambda - 1) - \frac{1}{2}\mathbb{P}(S_{n-1} = \lambda + 1), & \lambda > a, \\ \frac{1}{2}\mathbb{P}(S_{n-1} = \lambda + 1) - \frac{1}{2}\mathbb{P}(S_{n-1} = \lambda - 1), & \lambda < a. \end{cases}$

Beweis: (für $\lambda > a$).

ⓐ $$\mathbb{P}(T_\lambda \leq n) = \sum_{b \in \mathbb{Z}} \mathbb{P}(T_\lambda \leq n, S_n = b)$$

$$\overset{\text{Satz 21}}{=} \sum_{b \geq \lambda} \mathbb{P}(S_n = b) + \underbrace{\sum_{b < \lambda} \mathbb{P}(S_n = 2\lambda - b)}_{= \sum_{b > \lambda} \mathbb{P}(S_n = b)}$$

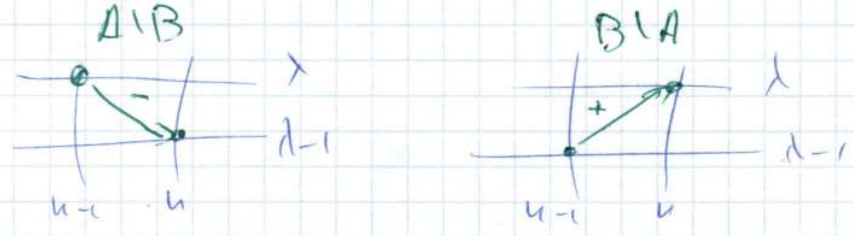
$$\overset{\sigma\text{-add.}}{=} \mathbb{P}(S_n \geq \lambda) + \mathbb{P}(S_n > \lambda).$$

ⓑ $$\mathbb{P}(T_\lambda = n) = \mathbb{P}(T_\lambda \leq n) - \mathbb{P}(T_\lambda \leq n-1)$$

$$\overset{ⓐ}{=} \mathbb{P}(S_n \geq \lambda) - \mathbb{P}(S_{n-1} \geq \lambda)$$
$$+ \mathbb{P}(S_n \geq \lambda + 1) - \mathbb{P}(S_{n-1} \geq \lambda + 1)$$

- Aus $\mathbb{P}(A) - \mathbb{P}(B) = \mathbb{P}(A \setminus B) - \mathbb{P}(B \setminus A)$ folgt:

$$\mathbb{P}(\underbrace{S_n \geq \lambda}_{A}) - \mathbb{P}(\underbrace{S_{n-1} \geq \lambda}_{B}) =$$



$$= \mathbb{P}(S_{n-1} = \lambda - 1, S_n = \lambda) - \mathbb{P}(S_{n-1} = \lambda, S_n = \lambda - 1)$$

$$\overset{uk}{=} \mathbb{P}(S_{n-1} = \lambda - 1) \cdot \frac{1}{2} - \mathbb{P}(S_{n-1} = \lambda) \cdot \frac{1}{2}.$$

$$\Rightarrow \mathbb{P}(T_\lambda = n) = \frac{1}{2}\Big(\mathbb{P}(S_{n-1} = \lambda - 1) - \cancel{\mathbb{P}(S_{n-1} = \lambda)} + \cancel{\mathbb{P}(S_{n-1} = \lambda)} - \mathbb{P}(S_{n-1} = \lambda + 1)\Big) \quad \#$$

(3) Maximum.

Kor. 23) Sei $M_n := \max_{0 \leq k \leq n} S_k$.

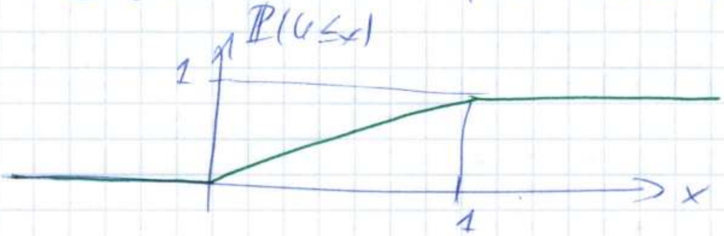
Für $\lambda > a$,
$$\mathbb{P}(M_n \geq \lambda) = \mathbb{P}(T_\lambda \leq n).$$

Beweis: klar. #

2.6) Simulationsverfahren.

Def. 24) (a) U ist eine <u>reellwertige Z.V.</u> falls
$$\{\omega \in \Omega \mid U(\omega) \leq x\} \in \mathcal{F}, \quad \forall x \in \mathbb{R}.$$

(b) $U : \Omega \to [0,1]$ ist <u>gleichverteilt auf $[0,1]$</u> falls
$$\mathbb{P}(U \leq x) = x, \quad \forall x \in [0,1].$$



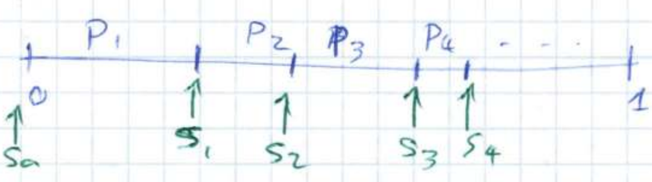
<u>Notation:</u> $U \sim \mathrm{Unif}([0,1])$ oder $U \sim \mathcal{U}([0,1])$

(c) $U_k : \Omega \to \mathbb{R}$, $k \in I$, heissen <u>unabhängig</u> falls $\{U_k \leq x_k\}, k \in I, \forall x_k \in \mathbb{R}$ unabhängig sind.

2.6.1) Direkte Verfahren.

- Sei $S = \{a_1, a_2, \dots\}$ diskret, μ eine W-Verteilung auf S mit Gewichte
$$p_k := \mathbb{P}(a_k)$$

Wegen $\sum_{k=1}^{\infty} p_k = 1$, setzen wir $s_n := \sum_{k=1}^{n} p_k$, $n \geq 1$. $s_0 = 0$



Lemma: Sei $U: \Omega \to [0,1)$, $U \sim \mathrm{Unif}([0,1))$, und $X(\omega) := a_n$, falls $s_{n-1} < U(\omega) \leq u_n$, $n \geq 1$.

Dann, $X \sim \mu$.

Beweis: $\forall n \geq 1$,

$$\mathbb{P}(X = a_n) = \mathbb{P}(s_{n-1} < U(\omega) \leq u_n)$$
$$= \mathbb{P}(U(\omega) \leq s_n) - \mathbb{P}(U(\omega) \leq s_{n-1})$$
$$= s_n - s_{n-1} = p_n. \quad \#$$

Mit Hilfe von Lemma 25:

Algo 26)

Input: $p_1, p_2, \ldots$

Output: X mit $X \sim \mu$.

```
n := 1;
s := p;
erzeuge u ~ Unif([0,1))
while u > s do
    n := n+1;
    s := s + p_n;
end while
return x := a_n
```

Bemerkung: $\mathbb{E}(\#\text{schritten}) = \sum_{n=1}^{\infty} n \cdot p_n$.

Problem: Falls $|S|$ gross ist, kann den Verfahren nicht so gut sein (Siehe Übungen).

2.6.2) Acceptance-Rejection-Verfahren.

- Sei μ eine W-Vert. mit Massenfunktion p,
  ν " " " q
  s.d. $\exists\, c \in [1,\infty)$ mit $\underline{p(x) \leq c \cdot q(x),\ \forall x \in S'}$

$$\Rightarrow 0 \leq \frac{p(x)}{c \cdot q(x)} \leq 1,\ \forall x \in S'.$$

Das wird die "Akzeptanz W-keit".

- Ziel: μ simulieren wenn ν leicht zu simulieren ist.

Algo 27) Input: $p(x), q(x), x \in S'$, c.
Output: x mit $x \sim \mu$.

```
repeat
        erzeuge x ~ ν;
        erzeuge u ~ Unif([0,1]);
until ( p(x)/(c·q(x)) ≥ u )
return x;
```

(d.h., Akzeptiere den Vorschlag mit W-keit $\frac{p(x)}{c \cdot q(x)}$).

- zu zeigen: Algo funktioniert und ist "schnell".

- Seien $X_1, X_2, \ldots \sim \nu$ die Vorschläge,
  $U_1, U_2, \ldots \sim \mathrm{Unif}([0,1])$;

Setze $T := \min\left\{ n \geq 1 \,\middle|\, \frac{p(X_n)}{c \cdot q(X_n)} \geq U_n \right\}$ und
$X_T(\omega) := X_{T(\omega)}(\omega)$.

Satz 28)

$$\left[\begin{array}{l} \text{(a)}\ X_T \sim \mu \\ \text{(b)}\ T \sim \text{geom}(1/c): \\ \quad \mathbb{P}(T=n) = \frac{1}{c-1}\left(1-\frac{1}{c}\right)^n, n \geq 1 \\ \quad \Rightarrow \mathbb{E}(T) = c. \end{array}\right.$$

Beweis: (b) Wegen unab. von $X_1, X_2, \dots$ und $U_1, U_2, \dots$, sind

$$A_n := \left\{ \frac{p(X_n)}{c \cdot q(X_n)} \geq U_n \right\}$$

unabhängig für verschiedene n.

$$\mathbb{P}(T=n) = \mathbb{P}(A_1^c \cap \dots \cap A_{n-1}^c \cap A_n).$$

$\Rightarrow$ Berechnen wir $\mathbb{P}(A_n)$:

$$\mathbb{P}(A_n) \overset{\sigma\text{-add.}}{=} \sum_{a \in S'} \mathbb{P}\left(\left\{U_n \leq \frac{p(a)}{c \cdot q(a)}\right\} \cap \{X_n = a\}\right)$$

$$\overset{\text{unab.}}{=} \sum_{a \in S'} \underbrace{\mathbb{P}\left(U_n \leq \frac{p(a)}{c \cdot q(a)}\right)}_{= \frac{p(a)}{c \cdot q(a)} \text{ weil } 0 \leq \frac{p(a)}{c \cdot q(a)} \leq 1} \cdot \underbrace{\mathbb{P}(X_n = a)}_{= q(a)}$$

$$= \sum_{a \in S'} \frac{p(a)}{c \cdot \cancel{q(a)}} \cdot \cancel{q(a)} = \frac{1}{c}$$

$$\Rightarrow \mathbb{P}(T=n) \overset{\text{unab.}}{=} \left(1-\frac{1}{c}\right)^{n-1} \cdot \frac{1}{c} = \frac{1}{c-1} \cdot \left(1-\frac{1}{c}\right)^n, n \geq 1.$$

(a)
$$\mathbb{P}(X_T = a) = \sum_{n \geq 1} \mathbb{P}(X_T = a, T = n)$$

$$= \sum_{n \geq 1} \mathbb{P}(A_1^c \cap \dots \cap A_{n-1}^c \cap (X_T = a \cap A_n))$$

$$= \sum_{n \geq 1} \left(1-\frac{1}{c}\right)^{n-1} \cdot q(a) \cdot \frac{p(a)}{c \cdot q(a)} = p(a). \qquad \#$$