Vorlesung 10a

Mehrstufige Zufallsexperimente

und

Markovketten

Teil 1

Spielregel (im diskreten Fall):

Für jedes $i = 1, \ldots, n-1$ hat man

Übergangswahrscheinlichkeiten

$$P(a_1 \ldots a_i, a_{i+1}) = \mathbf{P}_{a_1 \ldots a_i}(X_{i+1} = a_{i+1}) \ ,$$

die angeben,

mit welcher Wahrscheinlichkeit in der $(i+1)$-ten Stufe

das Ereignis $\{X_{i+1} = a_{i+1}\}$ eintritt,

gegeben das Eintreten von $\{X_1 = a_1, \ldots, X_i = a_i\}$.

Die gemeinsame Verteilung ist nun gegeben durch die

Multiplikationsregel

$$\mathbf{P}(X_1 = a_1, \ldots, X_n = a_n)$$
$$= \rho(a_1) P(a_1, a_2) \cdots P(a_1 \ldots a_{n-1}, a_n)\ .$$

Beispiel: Die Pólya-Urne.

In einer Urne befindet sich anfangs
eine weiße und eine blaue Kugel.
In jedem Schritt wird eine Kugel rein zufällig gezogen und gemeinsam mit einer zusätzlichen Kugel derselben Farbe zurückgelegt.
Die Zufallsvariable Z_i mit Werten in $\{0, 1\}$ bezeichne die im i-ten Zug vorgefundene Farbe (0 für blau, 1 für weiß).

Die Übergangswahrscheinlichkeiten sind

$$\mathbf{P}_{a_1 \dots a_i}(Z_{i+1} = a_{i+1}) = \frac{1+k}{2+i} \tag{1}$$

mit $a_1, \dots, a_i = 0, 1$

und

$$k = k(a_1, \dots, a_{i+1}) = \#\{j : 1 \le j \le i, a_j = a_{i+1}\} \ ,$$

$k+1$ ist also die Zahl der Kugeln in der Urne,

die nach i Zügen die Farbe a_{i+1} haben.

Z. B. ist für $(a_1, \dots, a_8) := (1, 1, 0, 1, 0, 0, 1, 1)$

$$\mathbf{P}\big((Z_1, \dots, Z_8) = (a_1, \dots, a_8)\big) = \frac{1}{2}\,\frac{2}{3}\,\frac{1}{4}\,\frac{3}{5}\,\frac{2}{6}\,\frac{3}{7}\,\frac{4}{8}\,\frac{5}{9} = \frac{5!\,3!}{9!}\,.$$

Für $0 \le k \le n$ hat jede 01-Zugfolge $(a_1, \dots, a_n)$

mit $a_1 + \dots + a_n = k$ dieselbe Wahrscheinlichkeit, nämlich

$$\mathbf{P}\big((Z_1, \dots, Z_n) = (a_1, \dots, a_n)\big) = \frac{k!(n-k)!}{(n+1)!} = \frac{1}{n+1}\binom{n}{k}^{-1}\,.$$

Für $0 \le k \le n$ hat jede 01-Zugfolge $(a_1, \dots, a_n)$

mit $a_1 + \cdots + a_n = k$ dieselbe Wahrscheinlichkeit, nämlich

$$\mathbf{P}\big((Z_1, \dots, Z_n) = (a_1, \dots, a_n)\big) = \frac{k!(n-k)!}{(n+1)!} = \frac{1}{n+1}\binom{n}{k}^{-1} .$$

Es gibt $\binom{n}{k}$ derartige Zugfolgen. Also ist

$$\mathbf{P}(Z_1 + \cdots + Z_n = k) = \frac{1}{n+1}\,, \quad k = 0, \dots, n\,.$$

Fazit:

Die Anzahl der weißen Kugeln nach n Zügen

ist uniform verteilt in $\{0, 1, \dots, n\}$.

Pólya-Urne mit r Farben:

Wieder wird in jedem Zug die gezogene Kugel zusammen mit einer gleichfarbigen Kugel zurückgelegt.

Die Anfangsbesetzung sei $(1, \dots, 1)$, also je eine Kugel von jeder Farbe.

$X_{jn} :=$ # Neuzugänge der Farbe j in n Schritten.

Sei $(k_1, \dots, k_r) \in S_{n,r}$,

d.h. $k_j \in \mathbb{N}_0$ mit $k_1 + \cdots + k_r = n$.

Man sieht wie im Fall $r = 2$:

Alle möglichen Zugfolgen

von $(1, \dots, 1)$ nach $(1 + k_1, \dots, 1 + k_r)$

haben dieselbe Wahrscheinlichkeit, nämlich

$$\frac{k_1! \cdots k_r!}{r \cdot (r+1) \cdots (n+r-1)} = \frac{k_1! \cdots k_r!}{(n+r-1)!}(r-1)!\ .$$

Alle möglichen Zugfolgen

von $(1, \ldots, 1)$ nach $(1 + k_1, \ldots, 1 + k_r)$

haben dieselbe Wahrscheinlichkeit, nämlich

$$k_1! \cdots k_r! \frac{(r-1)!}{(n+r-1)!} \ .$$

Es gibt $\binom{n}{k_1,\ldots,k_r}$ solche Zugfolgen. Also ist

$$\mathbf{P}(X_{1n} = k_1, \ldots, X_{rn} = k_r) = \frac{n!(r-1)!}{(n+r-1)!} = \frac{1}{\binom{n+r-1}{n}}$$

$$\mathbf{P}(X_{1n} = k_1, \ldots, X_{rn} = k_r) = \frac{n!(r-1)!}{(n+r-1)!} = \frac{1}{\binom{n+r-1}{n}},$$

d.h. $(X_{1n}, \ldots, X_{rn})$ ist uniform verteilt auf $S_{n,r}$.

Fazit:

Die Pólya-Urne mit Anfangsbesetzung $(1, \ldots, 1)$

liefert

uniform verteilte Besetzungen!

Veranschaulichung von mehrstufigen Experimenten

durch *Bäume*:

$k_i = a_1 \dots a_i$ ist ein *Knoten der Tiefe* i

Die *Nachfolger* von k_i sind von der Form $k_i \, a_{i+1}$

Die *Kanten* des Baums erhalten die Gewichte

$$g(*, k_1) := \rho(a_1) \;, \quad g(k_i, k_{i+1}) := P(a_1 \dots a_i, a_{i+1})$$

mit $k_1 = a_1$, $k_i = a_1 \dots a_i$, $k_{i+1} = a_1 \dots a_{i+1}$.

Die Wahrscheinlichkeit, in einem bestimmten Blatt zu enden,

ergibt sich als Produkt der Kantengewichte

entlang des Weges von der Wurzel zum Blatt.

Prominente Beispiele von mehrstufigen Zufallsexperimente

sind solche, bei denen

die Übergangswahrscheinlichkeiten der nächsten Stufe

nur von der aktuellen Stufe abhängen

(und nicht von den vorhergehenden):

$$P(a_1 \ldots a_{i-2}\, a_{i-1}, a_i) = P(a_{i-1}, a_i)$$

In dem Fall spricht man von einer Markovkette

(mit Übergangsmatrix P)

Markovketten

Es geht um spezielle mehrstufige Zufallsxperimente

beschrieben durch Zufallsvariable $X_0, X_1, \ldots$

mit ein-und demselben Zielbereich S

(genannt “Zustandsraum”)

und Übergangswahrscheinlichkeiten,

die nur vom aktuellen Zustand abhängen.

Eine (endliche oder unendliche) Folge $X_0, X_1, \ldots$

von Zufallsvariablen mit abzählbarem Zielbereich S

heißt *Markovkette* mit *Zustandsraum* S*,*

Startverteilung ρ und *Übergangsmatrix* P, falls

$$\mathbf{P}(X_0 = a_0, \ldots, X_n = a_n) = \rho(a_0) P(a_0, a_1) \cdots P(a_{n-1}, a_n)$$

mit $n = 0, 1, \ldots$ und $a_0, \ldots, a_n \in S$.

Meist denkt man die Übergangsmatrix als fest

und notiert die Startverteilung als Subskript:

$$\mathbf{P}_{\rho}(X_0 = a_0, \ldots, X_n = a_n) =$$

$$\rho(a_0)P(a_0, a_1) \cdots P(a_{n-1}, a_n)$$

Startet die Kette in a, dann ist $\rho = \delta_a$

(die auf a konzentrierte Verteilung).

Statt $\mathbf{P}_{\delta_a}$ schreibt man auch $\mathbf{P}_a$

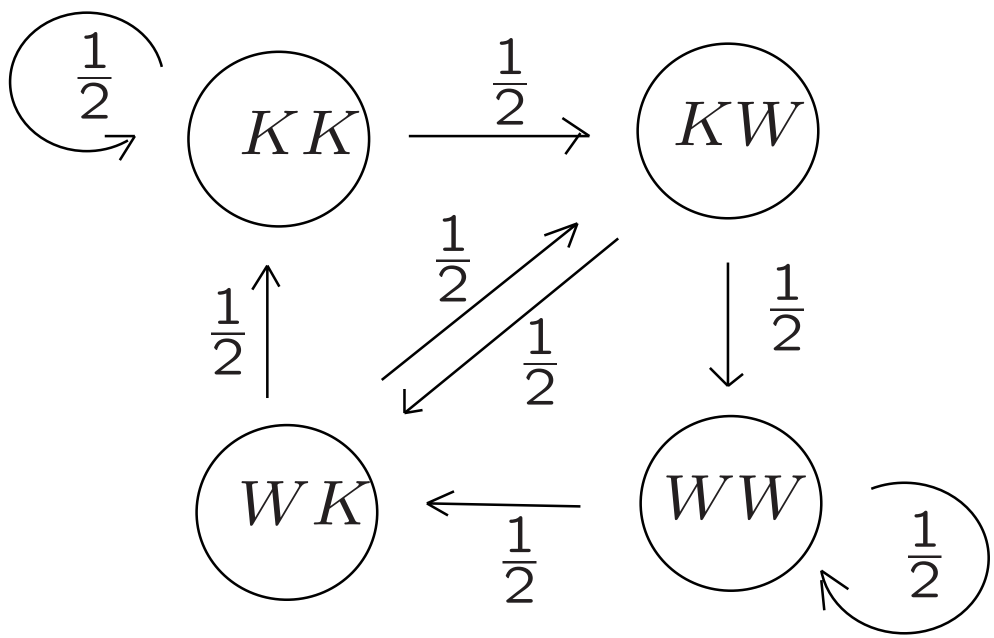
und erhält

$$\mathbf{P}_a(X_1 = a_1, \ldots, X_n = a_n) = P(a, a_1) \cdots P(a_{n-1}, a_n) \ .$$

Beispiel 1:

Muster der Länge 2 beim fairen Münzwurf

Graph der Übergangswahrscheinlichkeiten:



Beispiel 2:

Zufällige Wanderung durch einen Baum

von der Wurzel zur Krone (vgl Vorlesung 8b)

Die Zustände sind die Knoten;

die Übergangswahrscheinlichkeiten sind die Kantengewichte:

$$P(k,l) := \begin{cases} g(k,l), & \text{falls } l \text{ Nachfolger von } k, \\ 0 & \text{sonst.} \end{cases}$$

Die Blätter werden zu *absorbierenden Zuständen*.

Beispiel 3:

Irrfahrt auf $\mathbb{Z}^d$

Sind $Z_1, Z_2, \ldots$ unabhängige Kopien einer $\mathbb{Z}^d$-wertigen Zufallsvariablen Z und ist $a \in \mathbb{Z}^d$, dann ist

$$X_0 = a\ , \quad X_n := a + Z_1 + \cdots + Z_n\ , \quad n = 1, 2, \ldots\ ,$$

eine Markovkette mit Start in a und Übergangsmatrix

$$P(b,c) := \mathbf{P}(b + Z = c)\ , \quad b, c \in \mathbb{Z}^d\ .$$

Ist Z uniform verteilt auf den $2d$ Nachbarn des Ursprungs, dann sprechen wir von der *einfachen Irrfahrt* auf $\mathbb{Z}^d$.

Beispiel 4:

Pólya-Urne

Sei $(Z_1, Z_2, \ldots,)$ die Farbfolge der gezogenen Kugeln in einer Pólya$(1,1)$-Urne.

$$W_n := 1 + Z_1 + \cdots + Z_n\,, \quad B_n := 2 + n - W_n\,.$$

$X_n := (W_n, B_n)$ ist dann eine Markovkette mit Zustandsraum $S = \mathbb{N}^2$, Startzustand $(1,1)$ und Übergangswahrscheinlichkeiten

$$P((w,b),(w+1,b)) = \tfrac{w}{w+b}\,,$$
$$P((w,b),(w,b+1)) = \tfrac{b}{w+b}\,.$$