Vorlesung 9a

Kontinuierliche Übergangswahrscheinlichkeiten;

mehrstufige Zufallsexperimente

1. Addieren von unabhängigen ZV'en
– zweistufig aufgefasst:

Der diskrete Fall

(Buch S. 92)

Y und Z seien unabhängige $\mathbb{Z}$-wertige Zufallsvariable.

Wie ist $Y + Z$ verteit?

Wir können $Y + Z$ auffassen als

zweite Stufe eines Zufallsexperiments.

Die erste Stufe ist Y. Gegeben $\{Y = a\}$ ist

$$X_2 := Y + Z$$

so verteilt wie $a + Z$.

$$\mathbf{P}(Y = a, Y + Z = b) = \mathbf{P}(Y = a, a + Z = b)$$

$$= \mathbf{P}(Y = a)\,\mathbf{P}(a + Z = b)$$

$$= \mathbf{P}(Y = a)\,\mathbf{P}(Z = b - a)$$

Summation über a ergibt die “totale Wahrscheinlichkeit”:

$$\mathbf{P}(Y + Z = b) = \sum_a \mathbf{P}(Y = a)\,\mathbf{P}(Z = b - a)$$

(Zerlegung von $\mathbf{P}(Y + Z = b)$ nach den Ausgängen von Y,

“Zerlegung nach dem ersten Schritt”)

$$\mathbf{P}(Y+Z=b)=\sum_a \mathbf{P}(Y=a)\,\mathbf{P}(Z=b-a)$$

Beispiel:

Y, Z unabhängig und Geom(p)-verteilt

$$\mathbf{P}(Y+Z=b)=\sum_{a=1}^{b-1} pq^{a-1}pq^{b-a-1}$$
$$=(b-1)p^2q^{b-2}, \quad b=2,3,\ldots$$

Dies sind die Gewichte der sogenannten

negativen Binomialverteilung mit Parametern 2, p

Diese ist die Verteilung der Anzahl der Versuche

in einem p-Münzwurf bis einschließlich zum zweiten Erfolg.

2. Addieren von unabhängigen ZV'en
– zweistufig aufgefasst:

Der Fall mit Dichten

(Buch S. 92)

Im diskreten Fall hatten wir

$$\mathbf{P}(Y = a, Y + Z = b) = \mathbf{P}(Y = a, a + Z = b)$$

$$= \mathbf{P}(Y = a)\,\mathbf{P}(a + Z = b)$$

$$= \mathbf{P}(Y = a)\,\mathbf{P}(Z = b - a)$$

Haben Y und Z die Dichten $f(y)\,dy$ und $g(z)\,dz$,

so bekommt man analog die gemeinsame Dichte von (Y, Z):

$$\mathbf{P}(Y \in da, Y + Z \in db) = \mathbf{P}(Y \in da, a + Z \in db)$$

$$= \mathbf{P}(Y \in da)\,\mathbf{P}(a + Z \in db)$$

$$= f(a)da\,g(b - a)\,db$$

$$\boxed{\mathbf{P}(Y \in da, Y + Z \in db) = f(a)\, g(b-a)\; da\, db}$$

Integration über a gibt die Dichte von $Y + Z$:

$$\mathbf{P}(Y + Z \in db) = \Big(\int f(a)\, g(b-a)\; da\Big)\, db\ .$$

Beispiel: Für Y und Z unabhängig und Exp(1)-verteilt ist

$$\mathbf{P}(Y + Z \in db) = \Big(\int_0^b e^{-a} e^{-(b-a)}\; da\Big)\, db$$

$$= b e^{-b}\, db, \quad b \geq 0.$$

(Dichte der Gamma(2)-Verteilung)

Folglich ist für Y und Z unabhängig und Exp(1)-verteilt,

sowie für $\lambda > 0$

$$\mathbf{P}\left(\frac{Y+Z}{\lambda} \in db\right) = (\lambda b)e^{-(\lambda b)}\, d(\lambda b)$$
$$= \lambda^2 b\, e^{-\lambda b}\, db, \quad b \geq 0.$$

(Dichte der Gamma-Verteilung

mit Formparameter 2 und Skalenparameter λ)

3. Münzwurf mit zufälliger Erfolgswahrscheinlichkeit

(Buch S. 93)

Münzwurf mit zufälliger Erfolgsw'keit.

Erste Stufe: Zufällige Erfolgswahrscheinlichkeit U.

Zweite Stufe: Gegeben $U = p$ führe p-Münzwurf durch.

Dadurch entsteht eine zufällige 01-Folge $Z_1, Z_2, \ldots$

Ist U uniform verteilt auf dem Intervall $[0, 1]$,

dann ergibt sich für $X_n := Z_1 + \cdots + Z_n$

$$\mathbf{P}(U \in dp, X_n = k) = dp \binom{n}{k} p^k (1-p)^{n-k} \ , \qquad 0 \le p \le 1,\ k = 0, \ldots, n$$

$$\mathbf{P}(X_n = k) = \binom{n}{k} \int_0^1 p^k (1-p)^{n-k} \, dp \ , \quad k = 0, 1, \ldots, n \ .$$

Eine Darstellung des Münzwurfs mit zufälliger, uniform verteilter Erfolgswahrscheinlichkeit:

Seien $U, U_1, U_2, \ldots$ unabhängig und uniform auf $[0, 1]$.

U gibt die zufällige Erfolgswahrscheinlichkeit, und

$$Z_1 := I_{\{U_1 < U\}}\,, \quad Z_2 := I_{\{U_2 < U\}}\,, \quad \ldots$$

stellt einen Münzwurf mit zufälligem Parameter U dar.

$$\{X_n = k\} = \{Z_1 + \cdots + Z_n = k\}$$

$$= \Big\{\text{von den } U_1, \ldots, U_n \text{ sind genau } k \text{ Stück kleiner als } U\Big\}.$$

Aus Symmetriegründen hat jedes der Ereignisse

$$E_k := \{\text{genau } k \text{ der } U_1, \dots, U_n \text{ fallen kleiner aus als als } U\}$$

die Wahrscheinlichkeit $\frac{1}{k+1}$

(denn die Anordnung der $U, U_1, U_2, \dots$ ist rein zufällig).

Also folgt:

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

Fazit:

Die Anzahl der Erfolge im n-fachen Münzwurf mit zufälliger, auf $[0, 1]$ uniform verteilter Erfolgswahrscheinlichkeit ist uniform verteilt auf $\{0, 1, \dots, n\}$.

4. Mehrstufigkeit und Multiplikationsregel

(Buch S. 94)

bisher Zweistufigkeit: von “heute” zu “morgen”

jetzt: von “heute und morgen” zu “übermorgen”, etc.

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 von $X_1, \ldots, X_n$

ergibt sich rekursiv als

Multiplikationsregel

$$\mathbf{P}(X_1 = a_1, \ldots, X_n = a_n)$$

$$= \mathbf{P}(X_1 = a_1, \ldots, X_{n-1} = a_{n-1}) P(a_1 \ldots a_{n-1}, a_n)$$

$$= \cdots$$

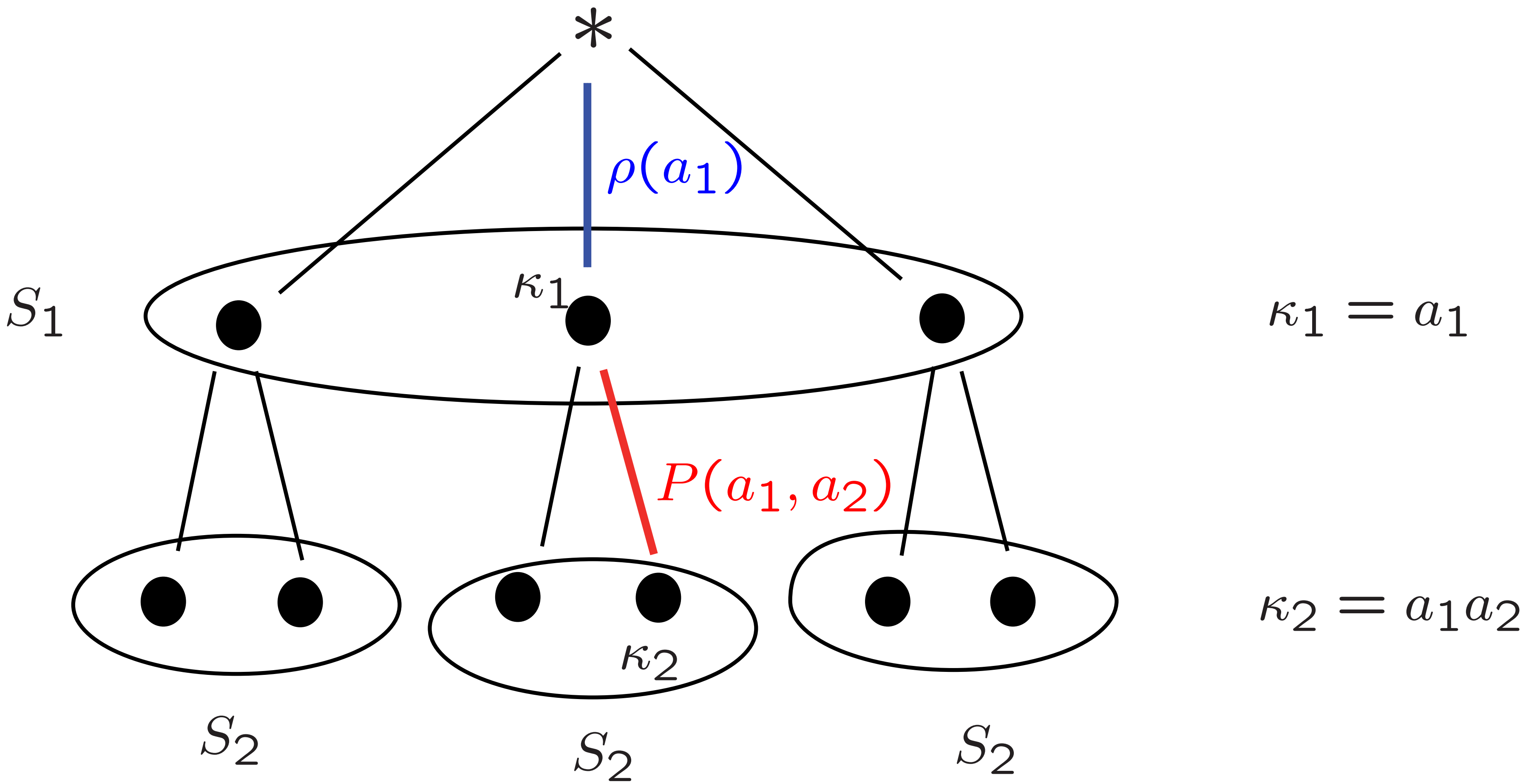
$$= \rho(a_1) P(a_1, a_2) \cdots P(a_1 \ldots a_{n-1}, a_n) \; .$$

5. Veranschaulichung von mehrstufigen Zufallsexperimenten durch Bäume

(Buch S. 95)

Wir erinnern uns an die

Veranschaulichung von zweistufigen Experimenten durch *Bäume*:



$$\mathbf{P}(X_1 = a_1, X_2 = a_2) = \rho(a_1) P(a_1, a_2)\,.$$

(Produkt der Kantengewichte von $*$ zum Knoten κ_2)

Was 2 recht ist, ist $i+1$ billig:

$\kappa_i = a_1 \ldots a_i$ ist ein *Knoten der Tiefe* i

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

Die *Kanten* des Baums erhalten die Gewichte

$$g(*, \kappa_1) := \rho(a_1) \;, \quad g(\kappa_i, \kappa_{i+1}) := P(a_1 \ldots a_i, a_{i+1})$$

mit $\kappa_1 = a_1, \;\; \kappa_i = a_1 \ldots a_i, \;\; \kappa_{i+1} = a_1 \ldots a_{i+1}$.

Die Wahrscheinlichkeit, in einem bestimmten Blatt zu enden,

ergibt sich als Produkt der Kantengewichte

entlang des Weges von der Wurzel zum Blatt.

6. Die Pólya-Urne

Oder

Wo Tauben sind, fliegen Tauben zu ....

(Buch S. 94)

In einer Urne befinden 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).

Wir nennen dann $(Z_1, \ldots, Z_n)$ auch

eine (Standard-) *Pólya-Folge* der Länge n.

$$\mathbf{P}(Z_1 = 0) = \mathbf{P}(Z_1 = 1) = \tfrac{1}{2}$$

$$\mathbf{P}_0(Z_2 = 0) = \tfrac{2}{3}$$

(das ist die W'keit, beim zweiten Zug blau zu ziehen,

wenn beim ersten Zug blau gezogen wurde)

$$\mathbf{P}_0(Z_2 = 1) = \tfrac{1}{3}$$

$$\mathbf{P}_{00}(Z_3 = 0) = \tfrac{3}{4}$$

$$\mathbf{P}_{01}(Z_3 = 0) = \tfrac{2}{4}$$

u. s. w.

Die Übergangswahrscheinlichkeiten sind

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

mit $a_1, \dots, a_i, a_{i+1} = 0, 1$,

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

$1 + \ell$ ist also die Zahl der Kugeln in der Urne nach i Zügen, deren Farbe mit der Farbe der $i+1$-ten gezogenen Kugel übereinstimmt.

Z. B. ist

$$\mathbf{P}\big((Z_1,\dots,Z_8)=(1,1,0,1,0,0,1,1)\big)=\frac{1}{2}\,\frac{2}{3}\,\frac{1}{4}\,\frac{3}{5}\,\frac{2}{6}\,\frac{3}{7}\,\frac{4}{8}\,\frac{5}{9}$$

$$=\frac{5!\;3!}{9!}\,.$$

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

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

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

nach n Zügen hinzugekommenen weißen Kugeln

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

7. Darstellung der Pólya-Urne als Münzwurf mit uniform verteilter Erfolgswahrscheinlichkeit

(Buch S. 113)

$U_0, U_1, U_2, \ldots$ seien unabhängig und Unif[0, 1]-verteilt,

$$Z_i := I_{\{U_i < U_0\}}, \quad i \geq 1.$$

Satz: Für jedes $n \in \mathbb{N}$ ist $(Z_1, \ldots, Z_n)$ so verteilt wie eine Standard-Pólyafolge der Länge n.

Beweis: Zum Aufwärmen betrachten wir $n = 2$.

$$\mathbf{P}(\{Z_1 = 1, Z_2 = 1\}) = \mathbf{P}(\{U_1 < U_0, U_2 < U_0\}) = \tfrac{2}{6},$$

$$\mathbf{P}(\{Z_1 = 1, Z_2 = 0\}) = \mathbf{P}(\{U_1 < U_0, U_2 \geq U_0\}) = \tfrac{1}{6}$$

(weil alle 3! = 6 “Rangfolgen” von (U_1, U_2, U_3) gleich wahrscheinlich sind)

Andererseits gilt für eine Standard-Pólyafolge (F_1, F_2) der Länge 2:

$$\mathbf{P}(F_1 = 1, F_2 = 1) = \tfrac{1}{2}\tfrac{2}{3} = \tfrac{1}{3},$$

$$\mathbf{P}(F_1 = 1, F_2 = 0) = \tfrac{1}{2}\tfrac{1}{3} = \tfrac{1}{6}.$$

Jetzt zu allgemeinem n:

Für gegebenes $(a_1, \ldots, a_n)$ berechnen wir die W'keit
des Ereignisses $E := \{Z_1 = a_1, \ldots, Z_n = a_n\}$
$= \{U_i < U_0$ genau dann, wenn $a_i = 1, i = 1, \ldots, n\}$.

Wir definieren die *Rangfolge* von $(U_0, \ldots, U_n)$ als
diejenige Permutation $(R(i))$ von $0, 1, \ldots, n$, für die gilt:
$R(i) = j$ falls U_0 das j-kleinste der $U_0, \ldots, U_n$ ist.
(Dabei wird von 0 bis n statt von 1 bis $n + 1$ gezählt.)

Merke: R ist mit W'keit 1 wohldefiniert,
und für jede feste Permutation σ von $0, 1, \ldots, n$ ist
$\mathbf{P}(R = \sigma) = \frac{1}{(n+1)!}$

Sei G die Menge der Permutation π von $0, \ldots, n$ mit der Eigenschaft $\{\pi(i) < \pi(0)$ genau dann, wenn $a_i = 1\}$.

Sei $k := a_1 + \cdots + a_n$.
Die Menge G hat $k!(n-k)!$ Elemente. Also ist
$$\mathbf{P}(E) = \mathbf{P}(R \in G) = \frac{1}{(n+1)!} k!(n-k)! = \frac{1}{n+1}\binom{n}{k}^{-1}.$$

Das ist dieselbe W'keit wie die, welche wir in Formel $(*)$ auf der vorletzten Folie von Abschnitt 5 berechnet hatten. $\square$

8. Pólya-Urne mit r Farben

(Buch S. 95)

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

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

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

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

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

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

Fazit:

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

liefert

uniform verteilte Besetzungen!