Vorlesung 8b

Zweistufige Zufallsexperimente

Teil 1

Stellen wir uns ein zufälliges Paar $X = (X_1, X_2)$ vor,

das auf zweistufige Weise zustande kommt:

es gibt eine Regel, die besagt, wie X_2 verteilt ist,

gegeben dass X_1 den Ausgang a_1 hat.

Beispiel 1:

In Stufe 1

wird eine (zufällige) reelle Zahl eingestellt.

In Stufe 2 wird

ein um diese Zahl standard-normalverteiltes Ergebnis

produziert:

Gegeben $X_1 = a_1$

hat X_2 die Verteilung N$(a_1, 1)$.

Beispiel 2:

In Stufe 1 wird (aus einer Population) ein Mensch ausgewählt.

Dieser hat eine bestimme Krankheit ($\{X_1 = 1\}$)

oder er hat sie nicht ($\{X_1 = 0\}$).

In Stufe 2 wird dieser Mensch medizinisch getestet:

Gegeben $\{X_1 = 1\}$ sei der Test positiv mit W'keit 1.

Gegeben $\{X_1 = 0\}$ sei der Test positiv mit W'keit 0.01.

Beispiel 3:

In Stufe 1 entscheiden wir uns

mit Wahrscheinlichkeit p für einen fairen Würfel

und mit W'keit $1-p$ für einen gezinkten: alle 6 Seiten mit 6.

X_2:= die dann geworfene Augenzahl.

$$\mathbf{P}_{\text{fair}}(X_2 = 6) = 1/6$$
$$\mathbf{P}_{\text{gezinkt}}(X_2 = 6) = 1.$$

$$\mathbf{P}(X_2 = 6) = p \cdot 1/6 + (1-p) \cdot 1.$$

Zweistufige Zufallsexperimente - der diskrete Fall:

Sei $X = (X_1, X_2)$ ein zufälliges Paar

mit diskretem Zielbereich $S = S_1 \times S_2$.

Für jedes $a_1 \in S_1$ sei $P(a_1, .)$ eine Verteilung auf S_2

Vorstellung: gegeben $X_1 = a_1$

hat die die Zufallsvariable X_2 die Verteilung $P(a_1, .)$.

Schreibweise:

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

Aus der Verteilung ρ von X_1

und den *Übergangsverteilungen* P

gewinnt man die gemeinsame Verteilung von X_1 und X_2:

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

oder anders geschrieben

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

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

Wir bemerken: Genau dann

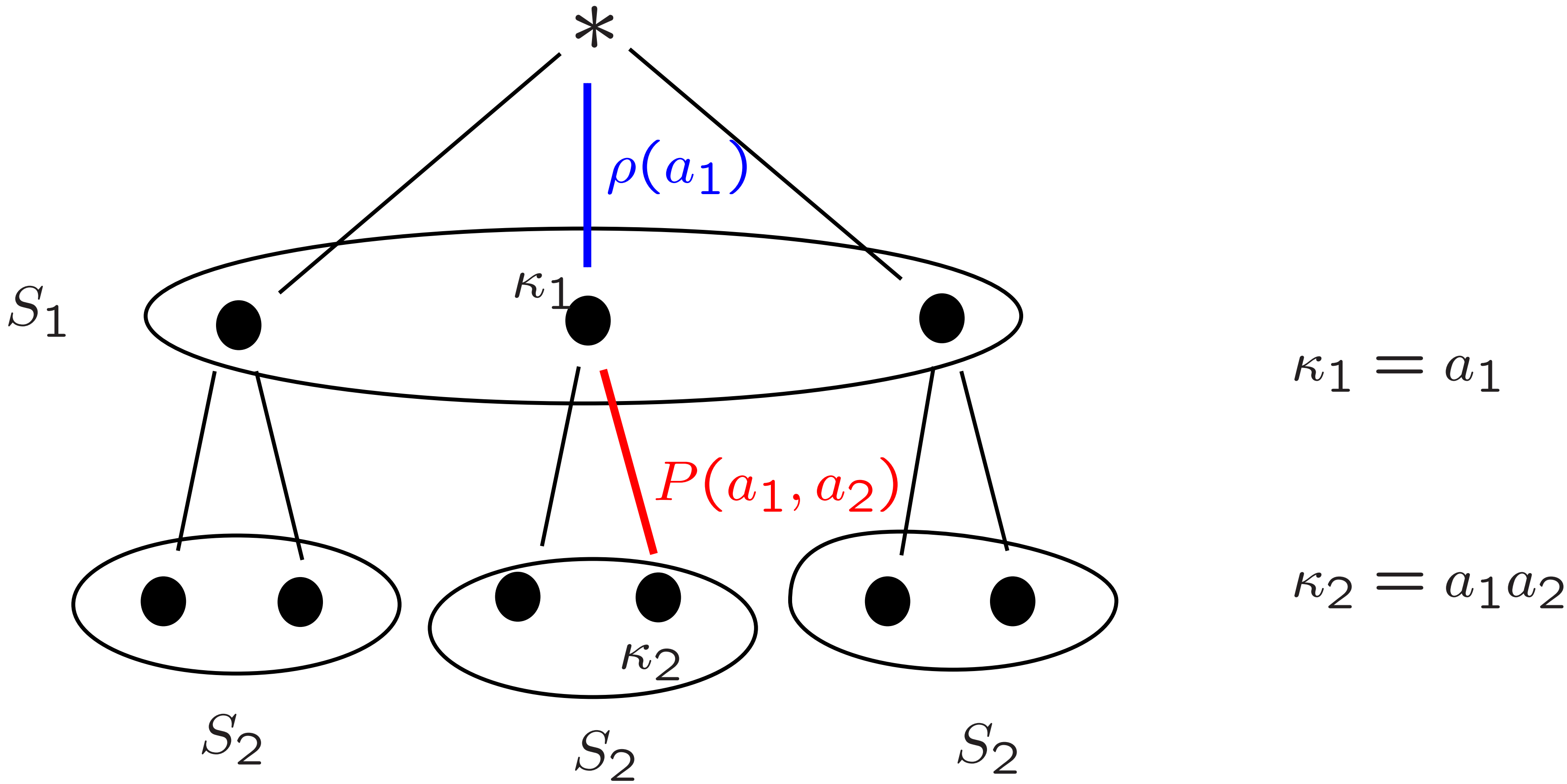
hängen die Verteilungen $\mathbf{P}_{a_1}(X_2 \in \cdot)$ nicht von a_1 ab,

wenn X_1 und X_2 unabhängig sind.

Ein zweistufiges Zufallsexperiment

kann in seiner Abfolge

durch einen *Baum der Tiefe* 2 veranschaulicht werden:



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

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

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

Summiert über $a_2 \in A_2$, mit $A_2 \subset S_2$, erhält man daraus:

$$\mathbf{P}(X_1 = a_1, X_2 \in A_2) = \mathbf{P}(X_1 = a_1)\, \mathbf{P}_{a_1}(X_2 \in A_2)\ .$$

$$\mathbf{P}(X_1 = a_1, X_2 \in A_2) = \mathbf{P}(X_1 = a_1)\,\mathbf{P}_{a_1}(X_2 \in A_2)\,.$$

Summation über $a_1 \in A_1$, mit $A_1 \subset S_1$:

$$\mathbf{P}(X_1 \in A_1, X_2 \in A_2)$$

$$= \sum_{a_1 \in A_1} \mathbf{P}(X_1 = a_1)\mathbf{P}_{a_1}(X_2 \in A_2).$$

$$\mathbf{P}(X_1 \in A_1, X_2 \in A_2)$$
$$= \sum_{a_1 \in A_1} \mathbf{P}(X_1 = a_1)\mathbf{P}_{a_1}(X_2 \in A_2).$$

Speziell mit $A_1 = S_1$ bekommt man die

Formel von der totalen Wahrscheinlichkeit:

$$\mathbf{P}(X_2 \in A_2) = \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1)\,\mathbf{P}_{a_1}(X_2 \in A_2)\,.$$

Diese zerlegt

die Wahrscheinlichkeit des Ereignisses $\{X_2 \in A_2\}$

nach den Ausgängen von X_1.

Ist $S_2 \subset \mathbb{R}$, dann setzen wir für jedes $a_1 \in S_1$

$$\mathbf{E}_{a_1}[X_2] := \sum_{a_2 \in S_2} a_2 \, \mathbf{P}_{a_1}(X_2 = a_2) \ ,$$

vorausgesetzt, die rechte Seite existiert,

und sprechen vom

Erwartungswert von X_2, gegeben $X_1 = a_1$.

$$\mathbf{E}[X_2] = \sum_{a_2 \in S_2} a_2 \, \mathbf{P}(X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} a_2 \, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} a_2 \, \mathbf{P}(X_1 = a_1) \mathbf{P}_{a_1}(X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} \mathbf{E}_{a_1}[X_2] \, \mathbf{P}(X_1 = a_1).$$

Also haben wir die einprägsame Formel

$$\mathbf{E}[\mathbf{E}_{X_1}[X_2]] = \mathbf{E}[X_2] \; .$$

(Zerlegung des Erwartungswertes von X_2 nach X_1.)

Beispiel: Suchen in Listen

(vgl. Buch S. 85-87)

n Namen werden in r Listen einsortiert. Dadurch ergibt sich eine Besetzung $k = (k_1, \ldots, k_r)$.

Jeder Name steht in seiner Liste Nr. j an einer der Stellen $i = 0, \ldots, k_j - 1$.

Stochastisches Modell für die erste Stufe:

Die zufällige Besetzung $Z = (Z_1, \ldots, Z_r)$ kommt durch n-maliges Würfeln mit den Gewichten $p_1, \ldots, p_r$ zustande.

Aus den n Namen wird rein zufällig einer herausgegriffen.

Es sei M die Stelle, die er in seiner Liste einnimmt.

Aufgabe: Berechne $\mathbf{E}[M]$.

(Dieser Erwartungswert beschreibt die mittlere Suchzeit
nach einem zufälligen Namen,
wenn man auf Grund einer mitgelieferten Kennzahl weiß,
in welcher Liste man zu suchen hat;
vgl. Buch S. 86)

Der Erwartungswert von M, gegeben $Z = k$, ist

$$\mathbf{E}_k[M] = \frac{1}{n} \sum_{j=1}^{r} \sum_{i=0}^{k_j-1} i = \frac{1}{n} \sum_{j=1}^{r} \frac{k_j(k_j-1)}{2} \,.$$

Mit der oben hergeleitetet Zerlegung des Erwartungswertes

$$\mathbf{E}[M] = \mathbf{E}\big[\mathbf{E}_Z[M]\big]$$

erhalten wir

$$\mathbf{E}[M] = \frac{1}{n} \sum_{j=1}^{r} \mathbf{E}\Big[\frac{Z_j(Z_j-1)}{2}\Big] \,.$$

Weil nach Annahme Z_j Binomial(n, p_j)-verteilt ist, folgt

$$\mathbf{E}[Z_j^2] = \mathbf{Var} Z_j + (\mathbf{E}[Z_j])^2$$
$$= np_j(1 - p_j) + (np_j)^2 = p_j^2 n(n-1) + np_j.$$

Weiter ergibt sich

$$\frac{1}{2n}(\mathbf{E}[Z_j^2] - \mathbf{E}[Z_j]) = \frac{1}{2} p_j^2 (n-1),$$

$$\mathbf{E}[M] = \frac{n-1}{2}(p_1^2 + \cdots + p_r^2)\ .$$

$$\mathbf{E}[M] = \frac{n-1}{2}(p_1^2 + \cdots + p_r^2) \ .$$

Im Fall uniformer Gewichte

$$p_1 = \cdots = p_r = 1/r$$

ergibt sich

$$\mathbf{E}[M] = \frac{n-1}{2r} \ .$$

Eine Modifikation des vorigen Beispiels:

Sei Z wieder multinomial $(n,\ p_1, \ldots, p_r)$ verteilt,

J sei unabhängig von K, mit $\mathbf{P}(J = j) = p_j, \quad j = 1, \ldots, r$.

Berechne den Erwartungswert von $X := Z_J$.

(Dieser Erwartungswert beschreibt die mittlere Suchzeit

nach einem in den Listen nicht vorhandenen Namen,

vgl. Buch S. 85)

Wir zerlegen $\mathbf{E}[X]$ nach den Ausgängen von Z:

$$\mathbf{E}[X] = \mathbf{E}[\mathbf{E}_Z[X]] = \sum_k \mathbf{P}(Z = k)\,\mathbf{E}_k[X]$$

$$= \sum_k \mathbf{P}(Z = k) \sum_{j=1}^{r} p_j k_j$$

$$= \sum_{j=1}^{r} p_j \sum_k \mathbf{P}(Z = k) k_j$$

$$= \sum_{j=1}^{r} p_j\,\mathbf{E} Z_j = n \sum_{j=1}^{r} p_j^2 .$$

Im Fall uniformer Gewichte

$$p_1 = \cdots = p_r = 1/r$$

ergibt sich

$$\mathbf{E}[X] = \frac{n}{r}.$$

Im Vergleich dazu war (siehe voriges Besipiel)

die mittlere “Suchtiefe” eines rein zufällig aus den n

herausgegriffenen Namens

$$\mathbf{E}[M] = \frac{n-1}{2r}.$$