Vorlesung 8a

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 entscheiden wir uns

mit Wahrscheinlichkeit $2/3$ für einen fairen Würfel

und mit W'keit $1/3$ für einen gezinkten:

drei Seiten mit 5, drei mit 6.

X_2:= die dann geworfene Augenzahl.

$$\mathbf{P}_{\text{fair}}(X_2 = 6) = \frac{1}{6}$$

$$\mathbf{P}_{\text{gezinkt}}(X_2 = 6) = \frac{1}{2}.$$

$$\mathbf{P}(X_2 = 6) = \frac{2}{3} \cdot \frac{1}{6} + \frac{1}{3} \cdot \frac{1}{2}.$$

Beispiel 2:

In Stufe 1

wählen wir eine auf $\{1, 2, 3\}$ uniform verteilte Zahl

In Stufe 2 verschieben wir das Ergebnis

mit W'keit 1/2 um eins nach rechts

und mit W'kt 1/2 um eins nach links.

Gegeben $X_1 = 3$, ist X_2 uniform verteilt auf $\{2, 4\}$.

$$\mathbf{P}(X_1 = 3, X_2 = 2) = \frac{1}{3} \cdot \frac{1}{2}\,.$$

$$\mathbf{P}(X_2 = 2) = \sum_{i=1}^{3} \mathbf{P}(X_1 = i, X_2 = 2) = \frac{1}{3}.$$

Beispiel 3:

In Stufe 1

stellt sich eine reelle Zahl X_1 ein.

In Stufe 2 wird dazu

eine unabhängige standard-normalverteilte ZV'e addiert:

Gegeben $X_1 = a_1$

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

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 von X_1

(der sogenannten *Startverteilung*, hier bezeichntet mit ρ)

und den Verteilungen $P(a_1, \,.)$

(den sogenannten *Übergangsverteilungen*)

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

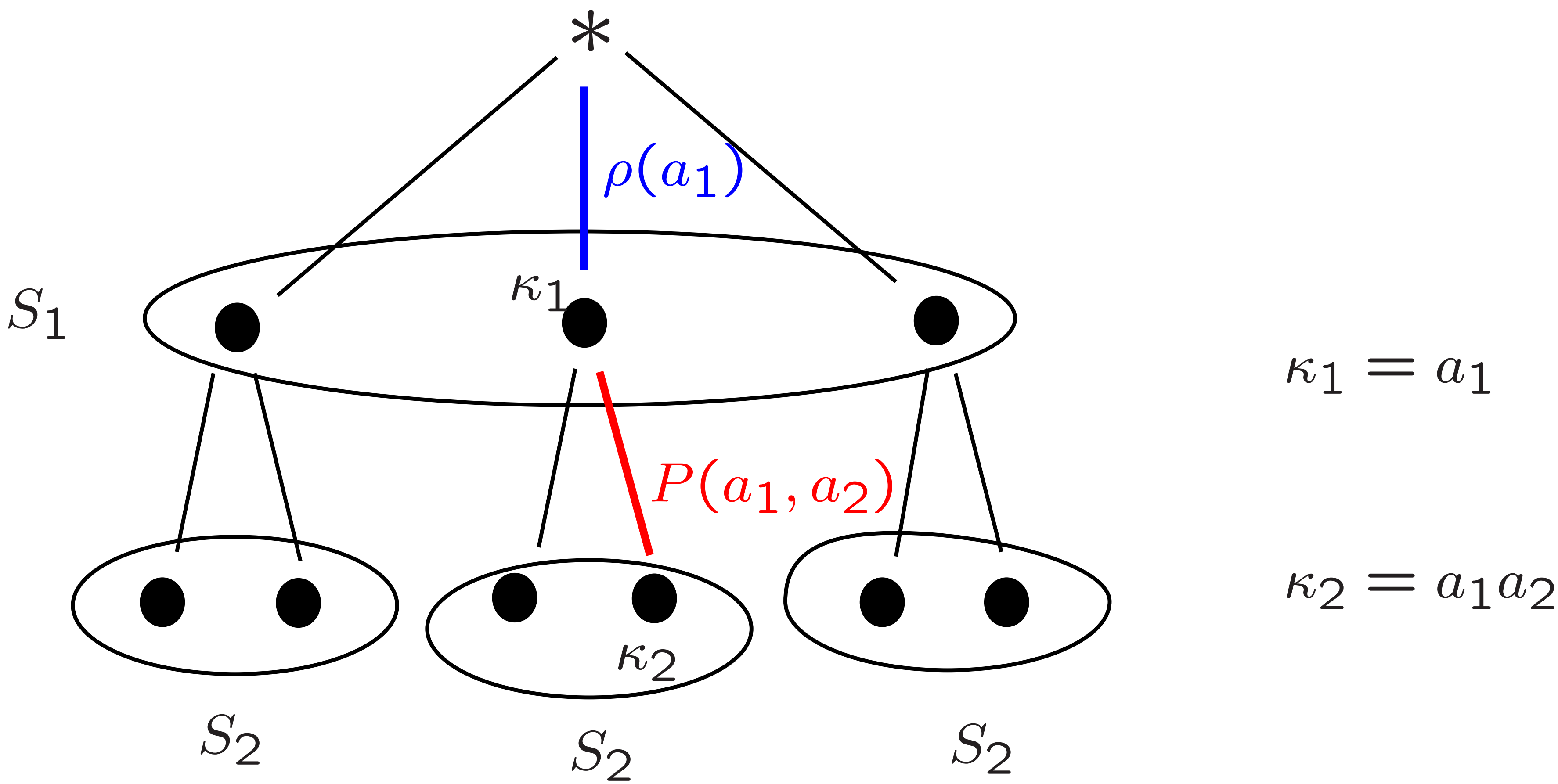
hängen die Verteilungen $\mathbf{P}_{a_1}(X_2 \in \cdot)$ nicht von a_1 ab,

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 $X_1 = a_1$.

Nicht nur die Verteilung von X_2 lässt sich
nach den Ausgängen von X_1 zerlegen,
sondern auch sein Erwartungswert.

Anders gesagt:

Es gibt auch die “Formel vom totalen Erwartungswert”:

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

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

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

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

Also haben wir die einprägsame Formel

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

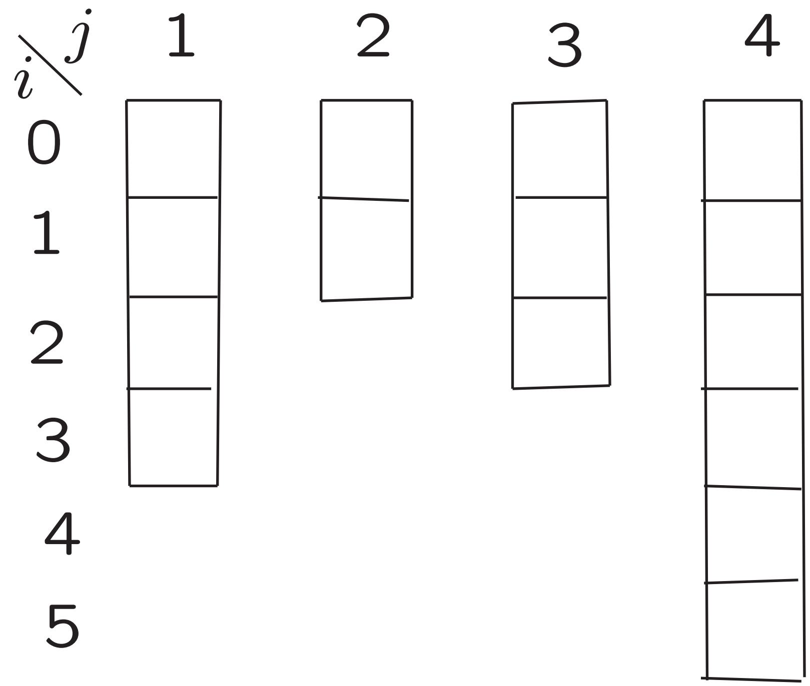
(Zerlegung des Erwartungswertes von X_2 nach X_1.)

Beispiel: Suchen in Listen.

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 $i = 0, \ldots, k_j - 1$.

Wir beginnen sachte: Was ist für gegebenes k der Erwartungswert der Stellennummer M (der “Suchtiefe”) eines rein zufällig aus den n herausgegriffenen Names?



Die Antwort ist

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

Wir betrachten jetzt ein

stochastisches Modell für die erste Stufe:

Annahme:

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

(Suchtiefe)

für einen aus den Listen zufällig gewählten Namen.)

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

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

Mit der oben hergeleitetet Zerlegung des Erwartungswertes

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

erhalten wir

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

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

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

(vgl. Übung 13 b)

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

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

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

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)

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

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

Zur Erinnerung: Wir denken uns das zufällige Paar

$$X = (X_1, X_2)$$

auf zweistufige Weise zustande gekommen:

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

$$\mathbf{E}_{a_1}\Big[h(X_1, X_2)\Big] := \sum_{a_2 \in S_2} h(a_1, a_2) \mathbf{P}_{a_1}(X_2 = a_2)$$

nennen wir den

bedingten Erwartungswert von $h(X_1, X_2)$*,*

gegeben $\{X_1 = a_1\}$.

"Zerlegung des Erwartungswertes nach der ersten Stufe":

$$\mathbf{E}\big[h(X_1, X_2)\big]$$

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

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

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

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

$$= \mathbf{E}\big[\mathbf{E}_{X_1}[h(X_1, X_2)]\big].$$

Merke: Der bedingte Erwartungswert

$$\mathbf{E}_{a_1}\big[h(X_1, X_2)\big]$$

ist eine Zahl.

$$\mathbf{E}_{X_1}\big[h(X_1, X_2)\big]$$

ist eine Zufallsvariable.

Wir sprechen von der

bedingten Erwartung von $h(X_1, X_2)$ gegeben X_1.

Die bedingte Erwartung als

beste Prognose im quadratischen Mittel:

Satz:

Sei X_2 reellwertige Zufallsvariable mit $\mathbf{E}[X_2^2] < \infty$.

Dann minimiert die bedingte Erwartung $\mathbf{E}_{X_1}[X_2]$

unter allen reellwertigen Zufallsvariablen der Form $h(X_1)$

den erwarteten quadratischen Abstand

$$\mathbf{E}\Big[(X_2 - h(X_1))^2\Big].$$

Beweis:

Wir zerlegen $\mathbf{E}\big[(X_2 - h(X_1))^2\big]$ nach X_1:

$$\mathbf{E}\big[(X_2 - h(X_1))^2\big] = \mathbf{E}\Big[\mathbf{E}_{X_1}\big[(X_2 - h(X_1))^2\big]\Big]$$
$$= \sum_{a_1} \mathbf{P}(X_1 = a_1)\mathbf{E}_{a_1}\big[(X_2 - h(a_1))^2\big]$$

Wir wissen schon (aus Vorlesung 7):

$\mathbf{E}_{a_1}\big[(X_2 - h(a_1))^2\big]$ wird minimal für

$h(a_1) := \mathbf{E}_{a_1}[X_2]$. $\square$