Vorlesung 5b

Unabhängigkeit

Teil 1

Zwei (diskrete) Zufallsvariable

(Buch S. 61)

Zwei binäre Zufallsvariable:

Ausgeglichene Verhältnisse?

Der Wertebereich von $X = (X_1, X_2)$ sei $\{o, u\} \times \{\ell, r\}$

	ℓ	r
o		
u		

Wie sollten die Verteilungsgewichte der 4 Ausgänge

$\{(o,\ell),(o,r),(u,\ell),(u,r)\}$ aussehen,

damit es gerechtfertigt ist zu sagen:

X_1 und X_2 sind unabhängig?

Beispiel:

	ℓ	r
o	$2/6$	$1/6$
u	$1/6$	$2/6$

Das Chancenverhältnis von ℓ zu r,

d.h. das Verhältnis $\mathbf{P}(X_2 = \ell) : \mathbf{P}(X_2 = r)$,

ist $(\frac{2}{6} + \frac{1}{6}) : (\frac{1}{6} + \frac{2}{6}) = 1 : 1.$

So wird man bezüglich X_2 wetten,

wenn man nichts über den Ausgang von X_1 weiß.

Beispiel:

	ℓ	r
o	$2/6$	$1/6$
u	$1/6$	$2/6$

Wenn man aber weiß, dass das Ereignis $\{X_1 = o\}$ eintritt,

wird man bezüglich des Ausgangs von X_2

anders wetten als $1 : 1$.

Das Chancenverhältnis von ℓ zu r ist dann $2 : 1$.

Die Verhältnisse in den Zeilen sind nicht ausgeglichen!

Beispiel: Ausgeglichene Verhältnisse

Der Wertebereich von $X = (X_1, X_2)$ sei $\{o, u\} \times \{\ell, r\}$

	ℓ	r
o	$2/9$	$4/9$
u	$1/9$	$2/9$

Hier sind die Chancenverhältnisse von ℓ zu r

in allen Zeilen gleich -

unabhängig vom Ausgang von X_1.

Definition

Zufallsvariable X_1, X_2 heißen *(stochastisch)* *unabhängig*,

wenn für alle Ereignisse $\{X_1 \in A_1\}$, $\{X_2 \in A_2\}$ gilt:

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

(“Produktformel für Wahrscheinlichkeiten”)

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

Zum Merken :

Für jedes Paar von Teilmengen $A_2,\ A_2' \subset S_2$ gilt:

Die Verhältnisse

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

hängen nicht ab von A_1 $(\subset S_1)$

(und sind in diesem Sinn “ausgeglichen”).

Beispiel: Zweimaliges (gewöhnliches) Würfeln (X_1, X_2):

X_1 und X_2 sind unabhängig. In der Tat:

Seien $A_1, A_2 \subset \{1, \dots, 6\}$ mit $\# A_1 =: m_1$, $\# A_2 =: m_2$.

Dann ist $\# A_1 \times A_2 = m_1 \cdot m_2$,

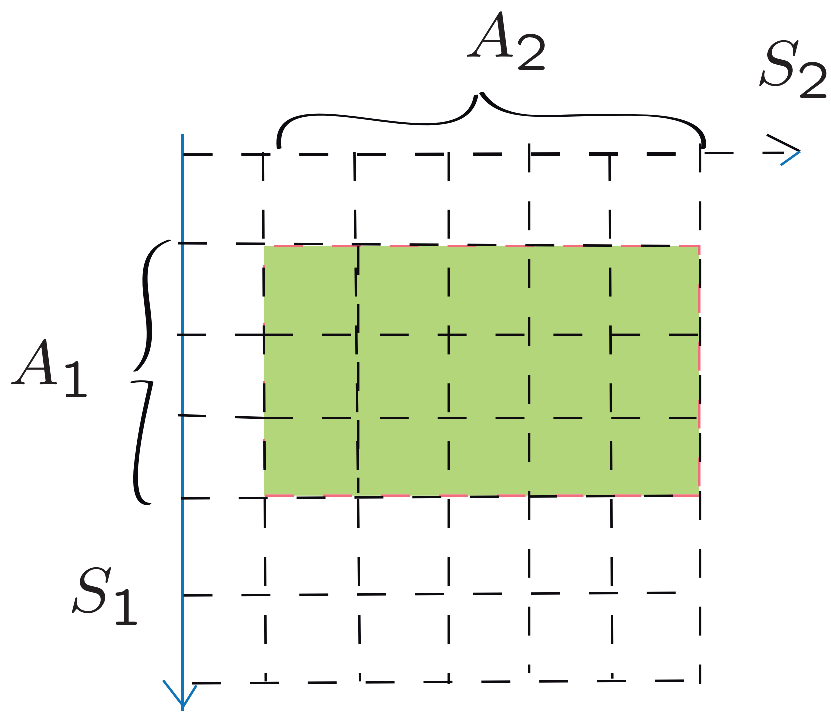
also

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

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

$$= \frac{m_1 m_2}{36} = \frac{m_1}{6} \cdot \frac{m_2}{6}$$

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



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

Für die Prüfung der Unabgängigkeit *diskreter* Zufallsvariabler

reicht es fürs Nachprüfen der Produktformel,

statt der *Teilmengen* $A_1 \subset S_1,\ A_2 \subset S_2$

die einzelnen *Ausgänge* $a_1 \in S_1$, $a_2 \in S_2$ zu betrachten:

Für diskrete Zufallsvariable gilt: X_1, X_2 sind unabhängig

genau dann,

wenn für alle $a_1 \in S_1$ und $a_2 \in S_2$ gilt:

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

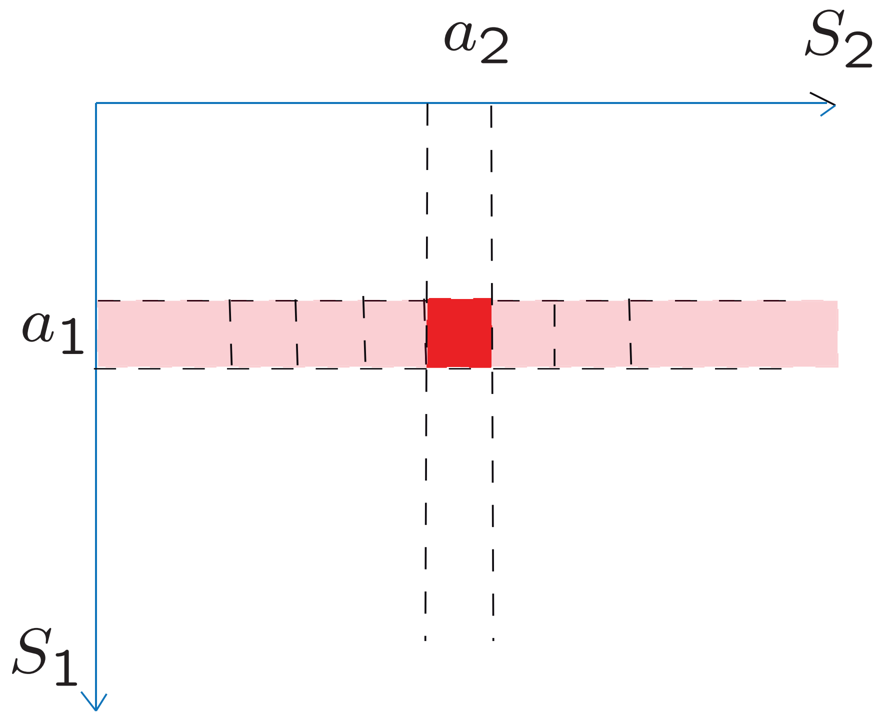
Denn:

"$\Longrightarrow$" ist klar (wähle $A_1 := \{a_1\},\ A_2 := \{a_2\}$).

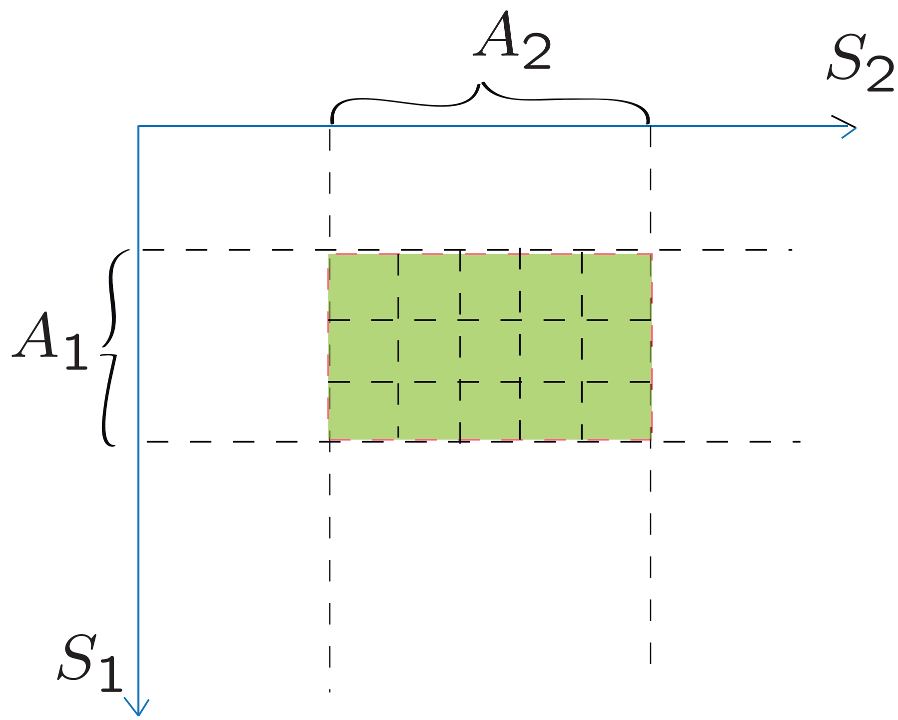
"$\Longleftarrow$": $\mathbf{P}(X_1 \in A_1, X_2 \in A_2) = \sum_{a_1 \in A_1,\, a_2 \in A_2} \mathbf{P}(X_1 = a_1, X_2 = a_2)$

$$= \sum_{a_1 \in A_1,\, a_2 \in A_2} \mathbf{P}(X_1 = a_1)\mathbf{P}(X_2 = a_2)$$

$$= \sum_{a_1 \in A_1} \mathbf{P}(X_1 = a_1) \sum_{a_2 \in A_2} \mathbf{P}(X_2 = a_2) = \mathbf{P}(X_1 \in A_1)\mathbf{P}(X_2 \in A_2).$$



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



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

Ein handliches Kriterium für die Unabhängigkeit zweier diskreter Zufallsvariabler ist die Proportionalität der Zeilen* der Matrix ihrer gemeinsamen Verteilungsgewichte $\rho(a_1, a_2)$.

In der Tat ist das hinreichend: dann ist nämlich $\rho(a_1, \cdot)$ *(das ist die Zeile namens a_1)* ja auch proportional zu $\rho_2(\cdot) := \sum_{a_1' \in S_1} \rho(a_1', \cdot)$ *(das ist die Summe der Zeilen)*, also gilt

$$(*) \qquad \rho(a_1, a_2) = k(a_1)\rho_2(a_2)$$

für irgendwelche nichtnegativen Zahlen $k(a_1)$.
Summiert man $(*)$ über $a_2 \in S_2$, so erhält man $\rho_1(a_1) = k(a_1)$.
Damit wird $(*)$ zur Produktformel von Folie 11.

Umgekehrt impliziert die Produktformel von Folie 11 die Proportionalität der Zeilen* der Matrix $(\rho(a_1, a_2))_{a_1 \in S_1,\, a_2 \in S_2}$ (Übung).

*oder auch der Spalten