

Vorlesung 5a

Unabhängigkeit

Teil 1

1. Unabhängigkeit von zwei Zufallsvariablen

Unabhängigkeit von zwei Zufallsvariablen

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”)

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

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

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

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

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

2. Produktformel für Erwartungswerte

Sind X_1, X_2 unabhängig mit Werten in S_1 bzw. S_2 ,
dann gilt für Mengen $A_1 \subset S_1, A_2 \subset S_2$:

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

Anders geschrieben:

$$\mathbf{E}[1_{A_1}(X_1)1_{A_2}(X_2)] = \mathbf{E}[1_{A_1}(X_1)] \mathbf{E}[1_{A_2}(X_2)]$$

In Worten:

Der Erwartungswert des Produktes von $1_{A_i}(X_i)$, $i = 1, 2$
ist das Produkt der Erwartungswerte.

Wir werden gleich sehen, dass allgemeiner gilt:

Sind X_1, X_2 unabhängig,
und h_1, h_2 reellwertige “Verarbeitungen”, dann gilt:

Der Erwartungswert des Produktes $h_1(X_1) \cdot h_2(X_2)$
ist gleich dem Produkt der Erwartungswerte.

Genauer:

Satz:

X_1, X_2 unabhängige ZV'e mit Zielbereichen S_1, S_2 ,
 h_1, h_2 Abbildungen von S_1 bzw. S_2 in die reellen Zahlen.

Haben $h_1(X_1)$ und $h_2(X_2)$ endlichen Erwartungswert,
so folgt

$$\mathbf{E}[h_1(X_1)h_2(X_2)] = \mathbf{E}[h_1(X_1)] \mathbf{E}[h_2(X_2)] .$$

(“Produktformel für Erwartungswerte”)

Beweis für diskrete ZV'e:

$$\mathbf{E}[h_1(X_1)h_2(X_2)]$$

$$= \sum_{a_1, a_2} h_1(a_1)h_2(a_2) \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= \sum_{a_1, a_2} h_1(a_1)\mathbf{P}(X_1 = a_1) h_2(a_2)\mathbf{P}(X_2 = a_2)$$

$$= \sum_{a_1} h_1(a_1)\mathbf{P}(X_1 = a_1) \sum_{a_2} h_2(a_2)\mathbf{P}(X_2 = a_2)$$

$$= \mathbf{E}[h_1(X_1)] \mathbf{E}[h_2(X_2)] \quad \square$$

3. Unabhängigkeit von mehreren Zufallsvariablen:

Zufallsvariable X_1, \dots, X_n mit Zielbereichen S_1, \dots, S_n
heißen

(stochastisch) *unabhängig*, falls für alle Ereignisse $\{X_i \in A_i\}$
folgende Produktformel gilt:

$$\mathbf{P}(X_1 \in A_1, \dots, X_n \in A_n) = \mathbf{P}(X_1 \in A_1) \cdots \mathbf{P}(X_n \in A_n) .$$

Unabhängigkeit von abzählbar unendlich vielen Zufallsvariablen:

Sei X_1, X_2, \dots eine Folge von Zufallsvariablen.

Definition:

Die Zufallsvariablen X_1, X_2, \dots sind unabhängig
: \iff für jedes n sind X_1, \dots, X_n unabhängig.

Beispiele:

Fortgesetzter Münzwurf, fortgesetztes Würfeln

Für diskrete Zufallsvariable X_1, \dots, X_n
ist die **Unabhängigkeit** gleichbedeutend mit der
Produktform der Verteilungsgewichte:

$$\mathbf{P}(X_1 = a_1, \dots, X_n = a_n) = \rho_1(a_1) \cdots \rho_n(a_n)$$

Die $\rho_i(a_i)$ sind dann die Verteilungsgewichte von X_i .

4. Unabhängigkeit von Ereignissen

Ereignisse E_1, \dots, E_n heißen **unabhängig**
: $\iff I_{E_1}, \dots, I_{E_n}$ sind **unabhängig**.

Satz:

Dafür reicht aus, dass

$$\mathbf{P}(E_{i_1} \cap \dots \cap E_{i_k}) = \mathbf{P}(E_{i_1}) \cdots \mathbf{P}(E_{i_k})$$

für alle $1 \leq i_1 < \dots < i_k \leq n$.

Einen eleganten Beweis führt man über eine Rechnung mit Indikatorvariablen (ähnlich wie bei der Einschluss-Ausschlussformel), vgl. Buch Seite 67.

Korollar zum vorigen Satz:

Die Unabhängigkeit zweier Ereignisse E_1, E_2
ist äquivalent zur Produktformel

$$\mathbf{P}(E_1 \cap E_2) = \mathbf{P}(E_1) \mathbf{P}(E_2)$$

Und die **Unabhängigkeit dreier Ereignisse** E_1, E_2, E_3 ist äquivalent dazu, dass **beide** der folgenden **Bedingungen a) und b)** erfüllt sind:

$$\text{a) } \mathbf{P}(E_1 \cap E_2) = \mathbf{P}(E_1)\mathbf{P}(E_2),$$

$$\mathbf{P}(E_1 \cap E_3) = \mathbf{P}(E_1)\mathbf{P}(E_3),$$

$$\mathbf{P}(E_2 \cap E_3) = \mathbf{P}(E_2)\mathbf{P}(E_3).$$

$$\text{b) } \mathbf{P}(E_1 \cap E_2 \cap E_3) = \mathbf{P}(E_1)\mathbf{P}(E_2)\mathbf{P}(E_3)$$

a) oder b) allein reichen i.a. nicht für die Unabhängigkeit:

Beispiel:

(Z_1, Z_2) sei ein zweifacher $\frac{1}{2}$ -Münzwurf,

$$E_1 := \{Z_1 = 1\}, E_2 := \{Z_2 = 1\},$$

$$E_3 := \{Z_1 = Z_2\}$$

E_1, E_2, E_3 sind paarweise unabhängig,
aber nicht unabhängig.

5-EURO-Preisaufgabe von vor 5 Jahren:

Finden Sie ein möglichst einfaches Beispiel mit 3 Ereignissen E_1, E_2, E_3 , die trotz der Gleichheit $\mathbf{P}(E_1 \cap E_2 \cap E_3) = \mathbf{P}(E_1)\mathbf{P}(E_2)\mathbf{P}(E_3)$ nicht unabhängig sind.

Lösung (nach dem Vorschlag von Lena Walter):

(X_1, X_2, X_3) sei uniform verteilt auf
 $S := \{(0, 0, 1), (0, 1, 0), (1, 0, 0), (1, 0, 1), (1, 1, 0), (1, 1, 1)\}$

$$E_1 := \{X_1 = 1\}, E_2 := \{X_2 = 1\}, E_3 := \{X_3 = 1\}$$

Dafür gilt

$$\mathbf{P}(E_1 \cap E_2 \cap E_3) = \frac{1}{6} = \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} = \mathbf{P}(E_1) \cdot \mathbf{P}(E_2) \cdot \mathbf{P}(E_3)$$

aber $\mathbf{P}(E_2 \cap E_3) = \frac{1}{6} \neq \frac{1}{4} = \mathbf{P}(E_2) \cdot \mathbf{P}(E_3)$.

5. Unabhängigkeit von Teilbeobachtungen

Sind X und Y unabhängig,

dann auch $g(X)$ und $h(Y)$.

Durch den Übergang zu “Teilbeobachtungen” $g(X)$ und $h(Y)$
können aber auch aus abhängigen Zufallsvariablen X, Y
voneinander unabhängige Zufallsvariable
 $g(X), h(Y)$ entstehen:

Gewisse **Teilaspekte** von **abhängigen Zufallsvariablen**
können **unabhängig** sein:

Beispiel: (X, Y) sei ein rein zufälliges Paar
verschiedener Zahlen aus $\{1, 2, \dots, 10\}$.

Offenbar sind X und Y nicht unabhängig.

Aber: die Ereignisse

$$E_1 := \{X \in \{1, 6\}\}, \quad E_2 := \{1 \leq Y \leq 5\}.$$

sind **unabhängig**.

Denn

$$\mathbf{P}(E_1) = \frac{1}{5}, \quad \mathbf{P}(E_2) = \frac{1}{2}, \quad \mathbf{P}(E_1 \cap E_2) = \frac{4+5}{10 \cdot 9} = \frac{1}{10}.$$

6. Positiv korrelierte Ereignisse

Was bedeutet die Beziehung

$$(*) \quad \mathbf{P}(E_1 \cap E_2) > \mathbf{P}(E_1) \cdot \mathbf{P}(E_2) ?$$

Offenbar sind dann E_1 und E_2 nicht unabhängig.

Schreiben wir $\mathbf{P}(E_1) =: p_1$, $\mathbf{P}(E_2) =: p_2$. Dann ist

$$\begin{aligned} (*) &\iff \mathbf{E}[I_{E_1}I_{E_2} - p_1p_2] > 0 \\ &\iff \mathbf{E}[I_{E_1}I_{E_2} - p_1I_{E_2} - p_2I_{E_1} + p_1p_2] > 0 \\ &\iff \mathbf{E}[(I_{E_1} - p_1)(I_{E_2} - p_2)] > 0. \end{aligned}$$

E_1 und E_2 haben dann also die Tendenz,
gemeinsam einzutreten oder gemeinsam nicht einzutreten.

Man nennt solche Ereignisse E_1 und E_2 **positiv korreliert**.

7. Indirekte Abhängigkeiten

Ein Beispiel für “indirekte Abhängigkeiten”:

Wir haben zwei gezinkte Münzen, mit $P(\text{Kopf} = 1) = 0.9$,
und eine faire Münze (mit $P(\text{Kopf} = 1) = 0.5$).

Jede der drei Münzen wird einmal geworfen; die Ausgänge
sind G , H für die gezinkten und F für die faire Münze.

Lernt man aus der Information, ob F so wie G ausfällt, etwas
für die Prognose, ob F so wie H ausfällt?

Vielleicht doch, denn:

Wenn F wie G ausfällt, fällt F eher als Kopf aus, und dann
fällt wohl H auch eher so aus wie F

Hier ist eine mathematische Analyse:

(G, H) sei ein p -Münzwurf, F sei uniform verteilt auf $\{0, 1\}$,

G, H, F seien unabhängig.

$$\mathbf{P}(G = F) = p \frac{1}{2} + q \frac{1}{2} = \frac{1}{2}, \quad \text{analog: } \mathbf{P}(H = F) = \frac{1}{2}$$

$$\mathbf{P}(G = F, H = F) = (p^2 + q^2) \frac{1}{2}$$

$$p^2 + q^2 = \frac{1}{2}((p + q)^2 + (p - q)^2) = \frac{1}{2}(1 + (p - q)^2) \geq \frac{1}{2},$$

mit “=” genau dann wenn $p = \frac{1}{2}$.

Für $p \neq \frac{1}{2}$ sind $\{G = F\}$ und $\{Y = F\}$ positiv korreliert!