Vorlesung 7a

Unabhängigkeit

Teil 1

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

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

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

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

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

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

anders geschrieben:

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

Der Erwartungswert des Produktes

ist das Produkt der Erwartungswerte.

Die Indikatorvariablen $I_{\{X_1 \in A_1\}}$ und $I_{\{X_2 \in A_2\}}$

sind unkorreliert.

Wir werden gleich sehen, dass allgemeiner gilt:

Sind X_1, X_2 unabhängig,
dann sind reellwertige “Verarbeitungen”
$h_1(X_1)$, $h_2(X_2)$ unkorreliert.

Anders gesagt:
Der Erwartungwert des Produktes $h_1(X_1) \cdot h_2(X_2)$
ist das 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

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

("Produktformel für Erwartungswerte")

Insbesondere sind (im Fall endlicher Varianzen)

$h_1(X_1)$ und $h_2(X_2)$ unkorreliert.

Beweis für diskrete ZV'e:

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

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

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

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

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

Unabhängigkeit von mehreren Zufallsvariablen:

Zufallsvariable $X_1, \ldots, X_n$ mit Zielbereichen $S_1, \ldots, S_n$

heißen

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

folgende Produktformel gilt:

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

Unabhängigkeit von abzählbar vielen Zufallsvariablen:

Sei $X_1, X_2, \ldots$ eine Folge von Zufallsvariaben.

Definition:

Die Zufallsvariablen $X_1, X_2, \ldots$ sind unabhängig

$:\Longleftrightarrow$ für jedes n sind $X_1, \ldots, X_n$ unabhängig.

Beispiele:

Fortgesetzter Münzwurf, fortgesetztes Würfeln

Für diskrete Zufallsvariable $X_1, \ldots, X_n$

ist die Unabhängigkeit geichbedeutend mit der

Produktform der Verteilungsgewichte:

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

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

Ereignisse $E_1, \ldots, E_n$ heißen unabhängig

$:\Longleftrightarrow I_{E_1}, \ldots, I_{E_n}$ sind unabhängig.

Satz:

Dafür reicht aus, dass

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

für alle $1 \le i_1 < \cdots < i_k \le n$.

Einen eleganten Beweis führt man mit einem Faktorisierungsargument für Indikatorvariable (ähnlich wie bei der Einschluss-Ausschlussformel), vgl. Buch Seite 67.

Korollar zum vorigen Satz:

Die Unahängigkeit 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:

a) $$\mathbf{P}(E_1 \cap E_2) = \mathbf{P}(E_1)\mathbf{P}(E_2),$$
$$\mathbf{P}(E_1 \cap E_3) = \mathbf{P}(E_1)\mathbf{P}(E_3),$$
$$\mathbf{P}(E_2 \cap E_3) = \mathbf{P}(E_2)\mathbf{P}(E_3).$$

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, Z_3 sei ein $\frac{1}{2}$-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 im SS 2011:

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

$$S := \{(0,0,1),(0,1,0),(1,0,0),(1,0,1),(1,1,0),(1,1,1)\}$$

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

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

Gewisse Teilaspekte von abhängigen Zufallsvariablen

können unabhängig sein:

Beispiel:

(X,Y) seien rein zufällige “Zwei aus $\{1,2,\ldots,32\}$”.

Offenbar sind X und Y nicht unabhängig.

Aber: die Ereignisse

$E_1 := \{X$ ist durch 8 teilbar $\}$, $\quad E_2 := \{17 \le Y \le 24\}$.

sind unabhängig.

Denn

$$\mathbf{P}(E_1) = \frac{1}{8}\,, \mathbf{P}(E_2) = \frac{1}{4}\,, \mathbf{P}(E_1 \cap E_2) = \frac{3 \cdot 8 + 7}{32 \cdot 31} = \frac{1}{32}.$$

Ein Beispiel für indirekte Abhängigkeiten:

G_1, G_2 seien gezinkte Münzen,

mit “W’keit für Kopf gleich 0.9”.

F sei eine faire Münze.

Jede der drei Münzen wird einmal geworfen.

Lernt man aus der Information, ob F so wie G_1 ausfällt,

etwas über die Prognose, ob F so wie G_2 ausfällt?

Wenn F wie G_1 ausfällt, fällt F eher als Kopf aus, und dann

fällt wohl G_2 auch eher so aus wie F ....

Hier ist eine mathematische Analyse:

Ein Beispiel für indirekte Abhängigkeiten:

(X, Y) sei ein p-Münzwurf, U sei uniform verteilt auf $\{0, 1\}$,

X, Y, U seien unabhängig.

$$\mathbf{P}(X = U) = p\frac{1}{2} + q\frac{1}{2} = \frac{1}{2}, \quad \text{analog: } \mathbf{P}(Y = U) = \frac{1}{2}$$

$$\mathbf{P}(X = U, Y = U) = (p^2 + q^2)\frac{1}{2}$$

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

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

Für $p \neq \frac{1}{2}$ sind $I_{\{X=U\}}$ und $I_{\{Y=U\}}$ positiv korreliert!