Vorlesung 6a

Varianz und Kovarianz

Teil 4

Die Varianz einer Summe von ZV’en und die Kovarianz von zwei ZV’en

(Buch S. 60)

Beim zweifachen p-Münzwurf Z_1, Z_2

ergab sich aus der Unabhängigkeit der Z_i:

$$\mathbf{Var}[Z_1 + Z_2] = \mathbf{Var}[Z_1] + \mathbf{Var}[Z_2].$$

Wie “streuen” Summen von

nicht unabhängigen Zufallsgrößen?

Wie steht’s mit der

Varianz einer Summe von Zufallsvariablen?

$$\mathbf{Var}[X+Y] = \mathbf{E}[((X-\mu_X)+(Y-\mu_Y))^2]$$

$$= \mathbf{E}[(X-\mu_X)^2] + \mathbf{E}[(Y-\mu_Y)^2] + 2\mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

Mit der Definition der Kovarianz

$$\mathbf{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

bekommen wir

$$\boxed{\mathbf{Var}[X+Y] = \mathbf{Var}\ X + \mathbf{Var}\ Y + 2\mathbf{Cov}[X,Y].}$$

Die Kovarianz

$$\mathbf{Cov}[X,Y] = \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

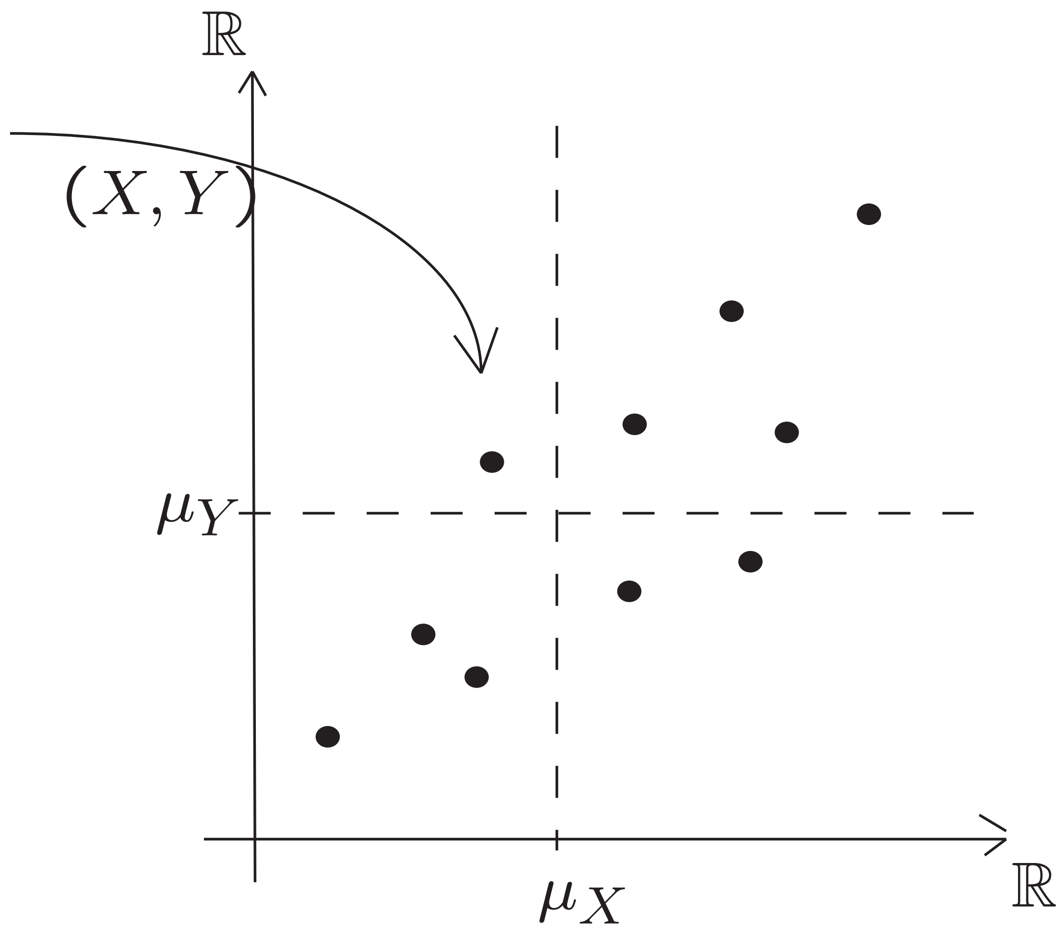
ist positiv,

wenn X und Y die Tendenz haben,

gemeinsam über bzw. gemeinsam unter

ihrem Erwartungswert auszufallen.

(Größere Abweichungen fallen dabei mehr ins Gewicht.)


$\mathbb{R}$
(X, Y)
μ_Y
μ_X
$\mathbb{R}$

Ist $\mathbf{Cov}[X, Y]$

$= 0$,	dann nennt man X, Y	unkorreliert
> 0,	$\ldots$	positiv korreliert
< 0,	$\ldots$	negativ korreliert.

Zwei Spezialfälle:

$$Y = X:$$

$$\mathbf{Cov}[X, Y] = \mathbf{E}[(X - \mu_X)(X - \mu_X)] = \mathbf{Var}[X]$$

$$Y = -X:$$

$$\mathbf{Cov}[X, Y] = \mathbf{E}[(X - \mu_X)(-X + \mu_X)] = -\mathbf{Var}[X]$$

$$\boxed{\mathbf{Var}[X+Y] = \mathbf{Var}X + \mathbf{Var}Y + 2\mathbf{Cov}[X,Y].}$$

Ganz analog ergibt sich:

$$\mathbf{Var}[Z_1 + \cdots + Z_n]$$

$$= \mathbf{Var}\, Z_1 + \cdots + \mathbf{Var}\, Z_n + 2 \sum_{i<j} \mathbf{Cov}[Z_i, Z_j]$$

Wir halten fest:

Sind $X_1, \dots, X_n$ reellwertige Zufallsvariable

mit endlicher Varianz und

$$\mathbf{Cov}[X_i, X_j] = 0 \quad \text{für } i \neq j$$

(man sagt dafür auch: die X_i sind *paarweise unkorreliert*)

dann gilt:

$$\mathbf{Var}[X_1 + \cdots + X_n] = \mathbf{Var}\, X_1 + \cdots + \mathbf{Var}\, X_n$$

Eine Umformung von $\mathbf{Cov}[X,Y]$:

$$\begin{aligned}\mathbf{E}[(X-\mu_X)(Y-\mu_Y)] &= \mathbf{E}[XY - \mu_X Y - X\mu_Y + \mu_X\mu_Y] \\ &= \mathbf{E}[XY] - \mu_X\mu_Y\end{aligned}$$

wegen der Linearität des Erwartungswertes.

So bekommen wir die (manchmal) hilfreiche Formel

$$\boxed{\mathbf{Cov}[X,Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]}$$

$$\boxed{\mathbf{Cov}[X,Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]}$$

Speziell für zwei Ereignisse E_1, E_2:

$$\mathbf{Cov}[I_{E_1}, I_{E_2}] = \mathbf{P}(E_1 \cap E_2) - \mathbf{P}(E_1)\,\mathbf{P}(E_2).$$

Vgl. die Definition

der positiven/negativen Korreliertheit von Ereignissen

in V5b6.

$$\boxed{\mathbf{Cov}[X,Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]}$$

Aus der Multiplikationsformel für den Erwartungswert

sehen wir

Unabhängige Zufallsvariable X und Y sind unkorreliert.

Speziell ist für unabhängige Zufallsvariable

mit endlichen Varianzen

die Varianz der Summe gleich der Summe der Varianzen.