

Kansrekening

8078254

June 12, 2025

Covariantie en correlatie

Definitie. Zij X, Y kansvariabelen, dan heet

$$\text{Cov}(X, Y) = \mathbb{E}([X - \mathbb{E}(X)][Y - \mathbb{E}(Y)])$$

heet de covariantie tussen X en Y . De correlatie tussen X en Y is

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \quad \text{als } \text{Var}(X), \text{Var}(Y) > 0$$

En 0 als een van de varianties = 0

Belangrijke eigenschap:

$$\text{Cor}(aX, bY) = \text{Cor}(X, Y)$$

voor $a, b > 0$

Voobeeld (lineair model). Zij X een kansvariabele, ε een kansvariabele. Zij $a, b \in \mathbb{R}$

$$Y = aX + b + \varepsilon$$

(ε een ruisterm die bijv. meetfouten kan modelleren)

Als X, Y kansvariabelen zijn, dan noemen we X en Y

- positief gecorreleerd als $\text{Cov}(X, Y) > 0$
- negatief gecorreleerd als $\text{Cov}(X, Y) < 0$
- ongecorreleerd als $\text{Cov}(X, Y) = 0$

Opmerking: Als X, Y onafhankelijk zijn, dan zijn X en Y niet gecorreleerd.

Waarom: Je kan laten zien

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

In het bijzonder geldt dus als X en Y onafhankelijk zijn, dan is

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

en dus $\text{Cov}(X, Y) = 0$

Belangrijk!! $\text{Cov}(X, Y) = 0 \not\Rightarrow X, Y$ onafhankelijk

Zij $(\vec{X}^{(i)})$ een rij kansvectoren, $\vec{X}$ een kansvector in $\mathbb{R}^d$. We zeggen dat $\vec{X}^{(i)} \stackrel{idd}{\sim} \vec{X}$ als

- $F_{\vec{X}^{(i)}} = F_{\vec{X}}$
- voor elke $I \subset \mathbb{N}$ zijn $\vec{X}^{(i)}, i \in I$ onafhankelijk:

$$F_{\vec{X}_{i \in I}^{(i)}} = \prod_{i \in I} F_{\vec{X}^{(i)}}$$

Definitie. Verwachtingswaarde van een kansvector

$\vec{X} = \begin{pmatrix} \vec{x}_1 \\ \vdots \\ \vec{x}_d \end{pmatrix}$ is gedefiniëerd door

$$\mathbb{E}(\vec{X}) = \begin{pmatrix} \mathbb{E}(\vec{x}_1) \\ \mathbb{E}(\vec{x}_2) \\ \vdots \\ \mathbb{E}(\vec{x}_d) \end{pmatrix}$$

Uitbreiding van het begrip “variantie” naar kansvectoren.

Definitie. De covariantiematrix van een d -dimensionale kansvector $\vec{X}$ is gedefiniëerd door

$$\text{Cov}(\vec{X}) = \text{Cov}(\vec{X}_i, \vec{X}_j)^d$$

Zij $(\vec{X}^{(i)})_{i \geq 1}$ kansvectoren, $\vec{X}$ een kansvector, $\vec{\mu} = \mathbb{E}(\vec{X})$ Definiëer

$$\vec{Y}^{(n)} = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\vec{X}^{(i)} - \vec{\mu})$$

Is er een kansvector $\vec{Y}$ zodat

$$\vec{Y}^{(n)} \xrightarrow{d} \vec{Y}$$

Dat wil zeggen

$$\lim_{n \rightarrow \infty} F_{\vec{Y}^{(n)}}(y) = F_Y(y)$$

voor alle $y \in \mathbb{R}^d$ zodat F_Y continu is in y

$$\vec{Y}_k^{(n)} = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\vec{X}_k^{(i)} - \mathbb{E}(\vec{X}_k^{(i)})) \xrightarrow{d} W_k W_k \sim N(0, \text{Var}(\vec{X}_k^{(i)}))$$

Stelling. Zij $(\vec{Z}^{(n)})_{n \geq 1}$ kansvectoren, $\vec{Z}$ een kansvector in $\mathbb{R}^d$ Als voor elke $\vec{v} \in \mathbb{R}^d$ geldt dat

$$\langle \vec{Z}^{(n)}, \vec{v} \rangle \xrightarrow{d} \langle \vec{Z}, \vec{v} \rangle$$

dan geldt

$$\vec{Z}^{(n)} \xrightarrow{d} \vec{Z}$$

$$\vec{Y}_k^{(n)} = \langle \vec{Y}^{(n)}, v \rangle, \text{ waarbij } v = e_k$$

Zij $v \in \mathbb{R}^d$

Definitie. Een kansvector $\vec{Z}$ met $\vec{\mu} = \mathbb{E}(\vec{X})$ en $\text{Cov}(\vec{Z}) = \Sigma$ heet normaal verdeeld, notatie

$$\vec{Z} \overset{d}{\sim} N(\vec{\mu}, \Sigma)$$

Definitie. Zij $\vec{Z}$ een d -dimensionale kansvector. Dan is de momentgenererendefunctie van $\vec{Z}$ gedefiniëerd door:

$$M_{\vec{Z}}(\vec{u}) = \mathbb{E}(e^{\langle \vec{u}, \vec{Z} \rangle})$$

test

$$\begin{array}{c} \vdots \\ X \cdot Y \end{array}$$

Deze was (*is at the time of writing*) niet grappig om te T_EX'en