Stochastik I

11. Große Übung

Martin Dattge und Leonardo Vela

09.12.2020

Multivariate Verteilungsfunktion

Definition

Ist $\mathbb{P}$ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(\mathbb{R}^d)$, so heißt

$$F_{\mathbb{P}}(t_1, \dots, t_d) = \mathbb{P}((-\infty, t_1] \times \dots \times (-\infty, t_d])$$

für $t_1, \dots, t_d \in \mathbb{R}$ multivariate Verteilungsfunktion von $\mathbb{P}$.

Multivariate Verteilungsfunktionen

Satz

Ist F die multivariate Verteilungsfunktion eines Wahrscheinlichkeitsmaßes auf $\mathcal{B}(\mathbb{R}^d)$, so gelten:

i) $F\colon \mathbb{R}^d \longrightarrow [0, 1]$

ii) $\lim\limits_{t_1 \to -\infty} F(t_1, \ldots, t_d) = \lim\limits_{t_2 \to -\infty} F(t_1, \ldots, t_d) = \ldots = \lim\limits_{t_d \to -\infty} F(t, \ldots, t_d) = 0,$

iii) $\lim\limits_{t_1, t_2, \ldots, t_d \to \infty} F(t_1, \ldots, t_d) = 1,$

iv) *F ist in jeder Koordinate rechtsstetig,*

v) *F ist rechtecksmonoton.*

- Mit dem Fortsetzungssatz von Carathéodory und dem Hauptsatz über Dynkin-Systeme existiert zu jeder multivariaten Verteilungsfunktion auch wieder ein eindeutig bestimmtes Wahrscheinlichkeitsmaß auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Rechtecksmonotonie

Definition

Eine Funktion $F\colon \mathbb{R}^d \to \mathbb{R}$ heißt rechtecksmonoton, falls für $t^1 \leq t^2 \in \mathbb{R}^d$ (d.h. $t_k^1 \leq t_k^2$, $\forall k \in \{1, \dots, d\}$)

$$\Delta_{t^1}^{t^2} F := \sum_{i_1,\dots,i_d \in \{1,2\}} (-1)^{i_1+\dots+i_d} F(t_1^{i_1}, \dots, t_d^{i_d}) \geq 0$$

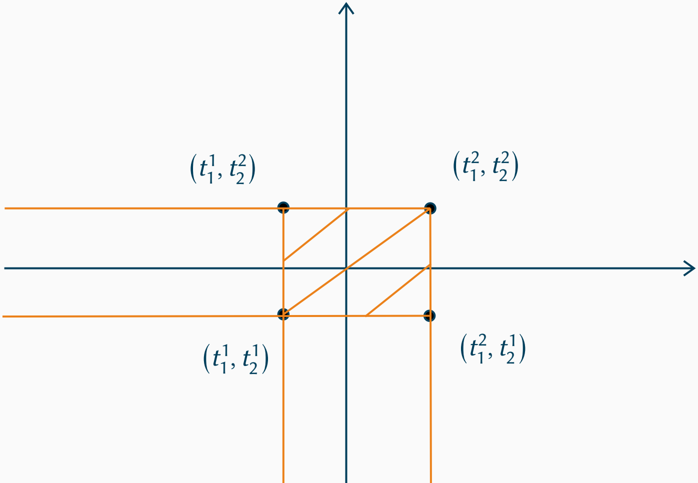
gilt.

- Die Rechtseckmonotonie ist eine Verallgemeinerung der Monotonie von Verteilungsfunktionen. Für $d = 1$ entspricht sie gerade der normalen Monotonie von Verteilungsfunktionen.
- Es gilt: $\Delta_{t_1}^{t^2} F = \mathbb{P}_F((t_1^1, t_1^2] \times \cdots \times (t_d^1, t_d^2])$. Deshalb darf dieser Wert auch nicht negativ werden!

Rechtseckmonotonie in $\mathbb{R}^2$

Für d=2 entspricht die Rechtseckmonotonie gerade

$$F(t_1^2, t_2^2) - F(t_1^1, t_2^2) - F(t^2, t_2^1) + F(t_1^1, t_1^1) \geq 0, \quad \forall t^1 \leq t^2 \in \mathbb{R}^2.$$



Nützliches zur Rechtseckmonotonie

Für den Differenzoperator $\Delta_{t^1}^{t^2}$ gilt zudem für eine Funktion $F\colon \mathbb{R}^d \longrightarrow \mathbb{R}$ und $t^1 \leq t^2 \in \mathbb{R}^d$:

$$\begin{aligned}
\Delta_{t^1}^{t^2} F &= \sum_{i_1,\dots,i_d \in \{1,2\}} (-1)^{i_1+\dots+i_d} F(t_1^{i_1}, \dots, t_d^{i_d}) \\
&= \sum_{i_1=1}^{2} \sum_{i_2=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\dots+i_d} F(t_1^{i_1}, \dots, t_d^{i_d}) \\
&= \sum_{i_1=1}^{2} (-1)^{i_1} \sum_{i_2=1}^{2} (-1)^{i_2} \cdots \sum_{i_d=1}^{2} (-1)^{i_d} F(t_1^{i_1}, \dots, t_d^{i_d}).
\end{aligned}$$

Aufgabe 1

Seien $F_1, F_2, \dots, F_d \colon \mathbb{R} \longrightarrow [0,1]$ mit $d \in \mathbb{N}$ Verteilungsfunktionen. Zeige, dass

$$F \colon \mathbb{R}^d \longrightarrow [0,1], \quad (t_1, t_2, \dots, t_d)^T \longrightarrow F(t_1, t_2, \dots, t_d) := F_1(t_1)F_2(t_2)\cdots F_d(t_d),$$

eine multivariate Verteilungsfunktion definiert.

Lösung Aufgabe 1

i) Es gilt $0 \leq F \leq 1$, da $0 \leq F_k \leq 1$ für alle $k \in \{1, \dots, d\}$.

ii) Für $k \in \{1, \dots, d\}$ gilt $\lim_{t_k \to -\infty} F_k(t_k) = 0$, da F_k eine Verteilungsfunktion ist. Also gilt

$$0 \leq \lim_{t_k \to -\infty} F(t_1, \dots, t_d) = \lim_{t_k \to -\infty} \prod_{i=1}^{d} F_i(t_i) \leq \lim_{t_k \to -\infty} F_k(t_k) = 0.$$

iii) Da die F_k eine Verteilungsfunktion ist für alle $k \in \{1, \dots, d\}$ gilt $\lim_{t_k \to \infty} F_k(t_k) = 1$ und damit

$$\lim_{t_1, \dots, t_d \to \infty} F(t_1, \dots, t_d) = \lim_{t_1, \dots, t_d \to \infty} \prod_{i=1}^{d} F_i(t_i) = \prod_{i=1}^{d} \lim_{t_i \to \infty} F_i(t_i) = 1$$

Lösung Aufgabe 1

iv) Seien $k \in \{1, \ldots, d\}$ und $t_i \in \mathbb{R}$, $i = 1, \ldots, d$ beliebig. Dann gilt für jede Folge $(t_k^n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ mit $t_k^n \downarrow t_k$ wegen der Rechtsstetigkeit von F_k

$$\lim_{n\to\infty} F(t_1, \ldots, t_k^n, \ldots, t_d) = \lim_{n\to\infty} \prod_{i\neq k}^{d} F_i(t_i) F_k(t_k^n) = \prod_{i\neq k}^{d} F_i(t_i) \lim_{n\to\infty} F_k(t_k^n) = \prod_{i=1}^{d} F_i(t_i).$$

iiv) Für $t^1 \leq t^2$ (d.h. $t_k^1 \leq t_k^2 \ \forall k \in \{1, \ldots, d\}$) gilt

$$\Delta_{t^1}^{t^2} = \sum_{i_1=1}^{2} (-1)^{i_1} \cdot \sum_{i_2=1}^{2} (-1)^{i_2} \cdot \ldots \cdot \sum_{i_d=1}^{2} (-1)^{i_d} F(t_1^{i_1}, \ldots, t_d^{i_d})$$

Lösung Aufgabe 1

Betrachten wir zunächst nur die letzte Summe des Ausdrucks.

$$\begin{aligned}\sum_{i_d=1}^{2}(-1)^{i_d}F(t_1^{i_1},...,t_d^{i_d}) &= F(t_1^{i_1},...,t_d^{2}) - F(t_1^{i_1},...,t_d^{1}) \\ &= F_1(t_1^{i_1})\cdot\ldots\cdot F_d(t_d^2) - F_1(t_1^{i_1})\cdot\ldots\cdot F_d(t_d^1) \\ &= F_1(t_1^{i_1})\cdot\ldots\cdot F_{d-1}(t_{d-1}^{i_{d-1}})\cdot(F_d(t_d^2) - F_d(t_d^1))\end{aligned}$$

Wir können den Ausdruck also schreiben als

$$\Delta_{t^1}^{t^2} = (F_d(t_d^2) - F_d(t_d^1))\sum_{i_1=1}^{2}(-1)^{i_1}\cdot\sum_{i_2=1}^{2}(-1)^{i_2}\cdot\ldots\cdot\sum_{i_{d-1}=1}^{d-1}(-1)^{i_{d-1}}F(t_1^{i_1})\cdot...\cdot F_{d-1}(t_{d-1}^{i_{d-1}})$$

Damit folgt induktiv $\Delta_{t^1}^{t^2} = \prod_{i=1}^{d}(F_i(t_i^2) - F(t_i^1))$ und damit $\Delta_{t_1}^{t^2} \geq 0$, da $F_i(t_i^2) - F(t_i^1) \geq 0 \forall i \in \{1,..,d\}$.

Zufallsvektoren

Definition

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Eine Abbildung

$$X\colon \Omega \longrightarrow \mathbb{R}^d, \quad \omega \mapsto X(\omega) = \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ \vdots \\ X_d(\omega) \end{pmatrix},$$

heißt Zufallsvektor, falls sie $\mathcal{A}$-$\mathcal{B}(\mathbb{R}^d)$-messbar ist.

- X ist genau dann ein Zufallsvektor, wenn alle $X_i\colon \Omega \longrightarrow \mathbb{R}$ Zufallsvariablen sind, also wenn sie $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-messbar sind.

Verteilungen von Zufallsvektoren

Definition

Das Maß

$$\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B)), \quad B \in \mathcal{B}(\mathbb{R}^d),$$

auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ heißt Verteilung des Zufallsvektors X.

- Um die Verteilung der Zufallsvariablen X_i, $i \in \{1, \ldots, d\}$, zu berechnen nutzen wir folgende Gleichung

$$\mathbb{P}(X_i \in B) = \mathbb{P}(X_1 \in \mathbb{R}, X_2 \in \mathbb{R}, \ldots, X_i \in B, \ldots, X_d \in \mathbb{R}) = \mathbb{P}_X(\mathbb{R} \times \ldots \times B \times \ldots \times \mathbb{R}), \quad B \in \mathcal{B}(\mathbb{R}),$$

man nennt $\mathbb{P}_{X_i}$ dann auch eindimensionale Randverteilung von X_i.

Identisch verteilte Zufallsvariablen

Definition

Seien X, Y Zufallsvariablen auf beliebigen Wahrscheinlichkeitsräumen. Dann heißen X und Y identisch verteilt, falls

$$F_X = F_Y$$

gilt.

Mit dem Hauptsatz über Dynkinsysteme gilt

$$F_X = F_Y \quad \Longleftrightarrow \quad \mathbb{P}_X = \mathbb{P}_Y,$$

weshalb die Definition über Verteilungsfunktionen hier ausreicht.

Unabhängige Zufallsvariablen

Definition

Seien $X_1, \ldots, X_d$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Die Zufallsvariablen heißen dann unabhängig, falls

$$F_{(X_1,\ldots,X_d)}(t_1, \ldots, t_d) = F_{X_1}(t_1) \cdot \ldots \cdot F_{X_d}(t_d)$$

für alle $t_1, \ldots, t_d \in \mathbb{R}$ gilt.

Es gilt zudem für unabhängige Zufallsvariablen und messbare Funktionen $f_1, \ldots, f_d \colon \mathbb{R} \to \mathbb{R}$:

$$\begin{aligned} \mathbb{P}(X_1 \in B_1, \ldots, X_d \in B_d) &= \mathbb{P}(X_1 \in B_1) \cdots \mathbb{P}(X_d \in B_d), \quad B_i \in \mathcal{B}(\mathbb{R}^d),\ i = 1, \ldots, d, \\ \mathbb{E}[f_1(X_1) \cdots f_d(X_d)] &= \mathbb{E}[f_1(X_1)] \cdots \mathbb{E}[f_d(X_d)], \end{aligned}$$

wobei die letzte Gleichung nur gilt, falls die Erwartungswerte existieren.

Kovarianz und Korrelation

Definition

Seien X, Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$, sodass die Erwartungswerte $\mathbb{E}[X], \mathbb{E}[Y], \mathbb{E}[XY]$ existieren. Dann heißt

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

Kovarianz von X und Y. Die Korrelation von X und Y ist dann, falls zudem $\mathbb{V}(X), \mathbb{V}(Y) > 0$ gilt, definiert durch

$$\mathrm{Cor}(X, Y) := \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}.$$

Insbesondere gilt

$$\mathrm{Cov}(X, X) = \mathbb{V}(X).$$

Interpretation der Kovarianz

- Eine positive Kovarianz wird so interpretiert, dass falls X einen großen (niedrigen) Wert annimmt, Y wahrscheinlich auch einen großen (niedrigen) Wert annimmt.
- Eine negative Kovarianz wird so interpretiert, dass falls X einen großen (niedrigen) Wert annimmt, Y wahrscheinlich einen niedrigen (großen) Wert annimmt.
- Sie sagt aber nichts über die Stärke dieser Beziehung aus, deshalb verwendet man oft die Korrelation, eine standardisierte Form der Kovarianz.

Aufgabe 2

a) Seien X, Y unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ für die die Kovarianz definiert ist. Zeige, dass

$$\operatorname{Cov}(X, Y) = 0$$

gilt.

b) Seien X, Y unabhängige standardnormalverteilte Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Berechne

$$\operatorname{Cov}(aX + bY, cX + dY), \quad \operatorname{Cor}(aX + bY, cX + dY), \quad a, b, c, d \in \mathbb{R}.$$

c) Zeige, dass im Allgemeinen für zwei Zufallsvariablen X, Y für die die Korrelation definiert ist,

$$|\operatorname{Cor}(X, Y)| \leq 1$$

gilt.

Hinweis für $c)$: Hölder-Ungleichung

Lösung Aufgabe 2

a) Definiere $f\colon \mathbb{R} \to \mathbb{R}, t \mapsto t - \mathbb{E}[X]$ und $g\colon \mathbb{R} \to \mathbb{R}, t \mapsto t - \mathbb{E}[Y]$, so sind beide Funktionen stetig und somit messbar.
Damit gilt für die Kovarianz, dass:

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] \\
&= \mathbb{E}[f(X)g(Y)] \\
&= \mathbb{E}[f(X)] \cdot \mathbb{E}[g(Y)] \\
&= \mathbb{E}[X - \mathbb{E}[X]] \cdot \mathbb{E}[Y - \mathbb{E}[Y]] \\
&= (\mathbb{E}[X] - \mathbb{E}[X]) \cdot (\mathbb{E}[Y] - \mathbb{E}[Y]) \\
&= 0
\end{aligned}$$

wobei wir die Unabhängigkeit von X und Y benutzt haben.

Lösung Aufgabe 2

b) Zunächst gilt mit der Linearität des Erwartungswerts, da X und Y standardnormalverteilt:

$$\mathbb{E}[aX + bY] = a \cdot \mathbb{E}[X] + b \cdot \mathbb{E}[Y] = 0 = \mathbb{E}[cX + dY]$$

Damit folgt für die Kovarianz, dass:

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathbb{E}[(aX + bY) \cdot (cX + dY)] \\
&= \mathbb{E}[acX^2 + adXY + bcXY + bdY^2] \\
&= ac \cdot \mathbb{E}[X^2] + (ad + bc) \cdot \mathbb{E}[XY] + bd \cdot \mathbb{E}[Y^2] \\
&= ac \cdot \mathbb{E}[X^2] + (ad + bc) \cdot \mathbb{E}[X] \cdot \mathbb{E}[Y] + bd \cdot \mathbb{E}[Y^2]
\end{aligned}$$

wobei wir die Unabhängigkeit von X und Y benutzt haben.
Wegen der Verschiebungsformel der Varianz und weil X und Y standardnormalverteilt folgt:

$$\mathrm{Cov}(X, Y) = ac + bd$$

Lösung Aufgabe 2

Für die Varianz gilt nach der Verschiebungsformel:

$$\begin{aligned}\mathbb{V}(aX + bY) &= \mathbb{E}[(aX + bY)^2] - (\mathbb{E}[aX + bY])^2 = \mathbb{E}[(aX + bY)^2] - 0 \\ &= \mathbb{E}[a^2 X^2 + 2abXY + b^2 Y^2] \\ &= a^2 \cdot \mathbb{E}[X^2] + 2ab \cdot \mathbb{E}[XY] + b^2 \cdot \mathbb{E}[Y^2] \\ &= a^2 + b^2\end{aligned}$$

Analog ergibt sich:

$$\mathbb{V}(cX + dY) = c^2 + d^2$$

Somit folgt:

$$\operatorname{Cor}(aX + bY, cX + dY) = \frac{ac + bd}{\sqrt{(a^2 + b^2) \cdot (c^2 + d^2)}}$$

Lösung Aufgabe 2

c) Mit der Dreiecksungleichung für Integrale und der Hölderungleichung für $p = q = 2$ gilt:

$$\begin{aligned} |\operatorname{Cov}(X,Y)| &= |\mathbb{E}[(X-\mathbb{E}[X])\cdot(Y-\mathbb{E}[Y])]| \\ &\le \mathbb{E}[|(X-\mathbb{E}[X])\cdot(Y-\mathbb{E}[Y])|] \\ &\le \sqrt{\mathbb{E}[|X-\mathbb{E}[X]|^2]\cdot\mathbb{E}[|Y-\mathbb{E}[Y]|^2]} \\ &= \sqrt{\mathbb{V}(X)\cdot\mathbb{V}(Y)} \end{aligned}$$

Daraus folgt für die Korrelation:

$$|\operatorname{Cor}(X,Y)| = \frac{|\operatorname{Cov}(X,Y)|}{\sqrt{\mathbb{V}(X)\cdot\mathbb{V}(Y)}} \le 1$$