Stochastik I

12. Große Übung

Martin Dattge und Leonardo Vela

16.12.2020

Faltung

Definition

Sind $\mu_1, \ldots, \mu_n$ Wahrscheinlichkeitsmaße auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ so heißt das Bildmaß des Produktmaßes $\mu_1 \otimes \cdots \otimes \mu_n$ unter der Abbildung

$$h_n \colon \mathbb{R}^n \longrightarrow \mathbb{R}, \quad (x_1, \ldots, x_n) \mapsto x_1 + \cdots + x_n,$$

Faltung von $\mu_1, \ldots, \mu_n$. Man schreibt

$$(\mu_1 * \cdots * \mu_n)(B) := (\mu_1 \otimes \cdots \otimes \mu_n)(h_n^{-1}(B)), \quad B \in \mathcal{B}(\mathbb{R}).$$

Summen von unabhängigen Zufallsvariablen

Seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen. Dann gilt für die Verteilung der Summe dieser Zufallsvariablen

$$\begin{aligned}
\mathbb{P}(X_1 + \cdots + X_n \in B) &= \mathbb{P}(h_n(X_1, \ldots, X_n) \in B) \\
&= \mathbb{P}((X_1, \ldots, X_n) \in h_n^{-1}(B)) \\
&= \mathbb{P}_{(X_1, \ldots, X_n)}(h_n^{-1}(B)) \\
&= (\mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n})(h_n^{-1}(B)) \\
&= (\mathbb{P}_{X_1} * \cdots * \mathbb{P}_{X_n})(B),
\end{aligned}$$

wobei B eine Menge aus $\mathcal{B}(\mathbb{R})$ sei. Die Faltung der Verteilungen von unabhängigen Zufallsvariablen ist also genau die Verteilung der Summe der entsprechenden Zufallsvariablen!

Faltung diskreter Verteilungen

Für diskrete Verteilungen braucht man die Faltung eigentlich überhaupt nicht! Seien X_1, X_2 unabhängige, diskrete Zufallsvariablen. Dann gilt

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 = a) &= \mathbb{P}(X_1 + X_2 = a, X_2 \in X_2(\Omega)) \\
&= \mathbb{P}\left(X_1 + X_2 = a, X_2 \in \dot{\bigcup_{b \in X_2(\Omega)}} \{b\}\right) \\
&= \sum_{b \in X_2(\Omega)} \mathbb{P}(X_1 + X_2 = a, X_2 = b) \\
&= \sum_{b \in X_2(\Omega)} \mathbb{P}(X_1 = a - b, X_2 = b) \\
&= \sum_{b \in X_2(\Omega)} \mathbb{P}(X_1 = a - b)\mathbb{P}(X_2 = b).
\end{aligned}$$

Beispiel: Summe diskreter Verteilungen

Seien X_1, X_2 unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_2 = 1) = \mathbb{P}(X_1 = 2) = \mathbb{P}(X_2 = 2) = \frac{1}{2}.$$

Dann gilt

$$X_2(\Omega) = \{1, 2\}$$

und somit

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 = 4) &= \sum_{b \in X_2(\Omega)} \mathbb{P}(X_1 = 4 - b)\mathbb{P}(X_2 = b) \\
&= \mathbb{P}(X_1 = 4 - 1)\mathbb{P}(X_2 = 1) + \mathbb{P}(X_1 = 4 - 2)\mathbb{P}(X_2 = 2) \\
&= \mathbb{P}(X_1 = 3)\mathbb{P}(X_2 = 1) + \mathbb{P}(X_1 = 2)\mathbb{P}(X_2 = 2) \\
&= \frac{1}{4}.
\end{aligned}$$

Faltung stetiger Verteilungen

Für Summen unabhängiger stetiger Zufallsvariablen braucht man hingegen die Faltung. Diese liefert dann eine Regel für die Dichte der Zufallsvariable $X_1 + X_2$, wenn X_1, X_2 unabhängige, stetige Zufallsvariablen, denn nach der Vorlesung gilt für diese

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(x-y) f_Y(y) \, \mathrm{d}y.$$

Aufgabe 1

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

$$\mathbb{P}_X = \mathbb{P}_Y = \mathcal{U}([0, 1]).$$

Berechne die Dichte der Zufallsvariable $Z := X + Y$ und den Erwartungswert von $\max\{1, Z\}$.

Lösung

Nach der Vorlesung gilt

$$\begin{aligned} f_Z(x) &= \int_{\mathbb{R}} f_X(x-y) f_Y(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}} \mathbb{1}_{[0,1]}(x-y)\mathbb{1}_{[0,1]}(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}} \mathbb{1}_{[x-1,x]}(y)\mathbb{1}_{[0,1]}(y)\,\mathrm{d}y, \end{aligned}$$

denn

$$0 \leq x-y \leq 1 \quad \Leftrightarrow \quad x \geq y \geq x-1.$$

Also folgt

$$f_Z(x) = \int_{\mathbb{R}} \mathbb{1}_{[x-1,x]\cap[0,1]}(y)\,\mathrm{d}y.$$

Lösung

Mit

$$f_Z(x) = \int_{\mathbb{R}} \mathbb{1}_{[x-1,x]\cap[0,1]}(y)\,\mathrm{d}y$$

folgt

$$f_Z(x) = \begin{cases} 0 & , x < 0, \\ x & , 0 \leq x \leq 1, \\ 1 - (x-1) & , 1 < x \leq 2, \\ 0 & , 2 \leq x. \end{cases}$$
$$= x\mathbb{1}_{[0,1]}(x) + (2-x)\mathbb{1}_{(1,2]}(x).$$

Also kennen wir die Dichte von Z und können somit auch $\mathbb{E}[\max\{1, Z\}]$ ausrechnen.

Nach der Vorlesung gilt, da Z eine Zufallsvariable mit bekannter Dichte ist

$$
\begin{aligned}
\mathbb{E}[\max\{1, Z\}] &= \int_{\mathbb{R}} \max\{1, z\} f_Z(z)\, \mathrm{d}z \\
&= \int_{\mathbb{R}} \max\{1, z\} \left(z \mathbb{1}_{[0,1]}(z) + (2 - z)\mathbb{1}_{(1,2]}(z)\right) \mathrm{d}z \\
&= \int_{\mathbb{R}} \max\{1, z\} z \mathbb{1}_{[0,1]}(z)\, \mathrm{d}z + \int_{\mathbb{R}} \max\{1, z\}(2 - z)\mathbb{1}_{(1,2]}(z)\, \mathrm{d}z \\
&= \int_0^1 \max\{1, z\} z\, \mathrm{d}z + \int_1^2 \max\{1, z\}(2 - z)\, \mathrm{d}z \\
&= \int_0^1 z\, \mathrm{d}z + \int_1^2 2z - z^2\, \mathrm{d}z \\
&= \left[\frac{1}{2} z^2\right]_0^1 + \left[z^2 - \frac{1}{3} z^3\right]_1^2 \\
&= \left(\frac{1}{2} - 0\right) + \left(4 - \frac{8}{3} - \left(1 - \frac{1}{3}\right)\right) = \frac{7}{6}.
\end{aligned}
$$

Bedingte Wahrscheinlichkeiten

Definition

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und A, B Mengen aus $\mathcal{A}$ (also Ereignisse) mit $\mathbb{P}(B) > 0$. Dann heißt

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

die bedingte Wahrscheinlichkeit des Ereignisses A gegeben B.

- $\mathbb{P}(A|B)$ gibt an, wie groß die Wahrscheinlichkeit ist, dass A eintritt, falls B schon eingetreten ist.
- Für unabhängige Ereignisse folgt insbesondere $\mathbb{P}(A|B) = \mathbb{P}(A)$

Aufgabe 2

Sei X eine $\mathbb{N}$-wertige Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Zeige die Äquivalenz der folgenden Aussagen:

1) $X \sim \mathrm{Geo}(p)$ für ein $p \in (0,1)$, d.h. es gilt
$$\mathbb{P}_X = \sum_{n=1}^{\infty} p(1-p)^{n-1}\delta_n.$$

2) Die Verteilung von X ist gedächtnislos, d.h. für alle $k, n \in \mathbb{N}$ gilt:
$$\mathbb{P}(X = n + k | X > k) = \mathbb{P}(X = n).$$

3) Für alle $n \in \mathbb{N}$ gilt $\mathbb{P}(X = n + 1 | X > 1) = \mathbb{P}(X = n)$.

Hinweis: Für $q \in (0,1)$ gilt
$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}.$$

Lösung

1) $\implies$ 2) Sei $X \sim \mathrm{Geo}(p)$. Dann gilt per Definition für $k, n \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{P}(X = n+k | X > k) &= \frac{\mathbb{P}(X = n+k, X > k)}{\mathbb{P}(X > k)} = \frac{\mathbb{P}(X = n+k)}{\mathbb{P}(X > k)} \\
&= \frac{p(1-p)^{n+k-1}}{1 - \mathbb{P}(X \leq k)} \\
&= \frac{p(1-p)^{n+k-1}}{1 - \sum_{i=1}^{k} p(1-p)^{i-1}} \\
&= \frac{p(1-p)^{n+k-1}}{1 - \sum_{i=0}^{k-1} p(1-p)^{i}} \\
&= \frac{p(1-p)^{n+k-1}}{1 - p(\frac{1-(1-p)^k}{1-(1-p)})} = \frac{p(1-p)^{n+k-1}}{(1-p)^k}
\end{aligned}$$

Lösung

Insgesamt folgt:

$$\mathbb{P}(X = n + k | X > k) = \frac{p(1-p)^{n+k-1}}{(1-p)^k} = p(1-p)^{n-1} = \mathbb{P}(X = n).$$

2) $\implies$ 3) folgt direkt mit der Wahl $k = 1$.

3) $\implies$ 1) Gelte dafür

$$\mathbb{P}(X = n + 1 | X > 1) = \mathbb{P}(X = n)$$

für alle $n \in \mathbb{N}$.

Lösung

Dann gilt wegen 3) auch

$$\begin{aligned}\mathbb{P}(X = n) &= \mathbb{P}(X = n+1 | X > 1) \\ &= \frac{\mathbb{P}(X = n+1, X > 1)}{\mathbb{P}(X > 1)} \\ &= \frac{\mathbb{P}(X = n+1)}{1 - \mathbb{P}(X = 1)}\end{aligned}$$

und dies ist äquivalent zu

$$\mathbb{P}(X = n)(1 - \mathbb{P}(X = 1)) = \mathbb{P}(X = n+1), \quad \forall n \in \mathbb{N}.$$

Das brauchen wir gleich um 1) zu zeigen.

Lösung

Wir zeigen nun per vollständiger Induktion mit der Wahl $p = \mathbb{P}(X = 1)$, dass

$$\mathbb{P}(X = n) = p(1 - p)^{n-1}, \quad \forall n \in \mathbb{N}$$

gilt.
Für $n = 1$ gilt per Definition $\mathbb{P}(X = 1) = p = p(1 - p)^0$.
Angenommen die Aussage gilt für ein bel. aber festes $n \in \mathbb{N}$, so folgt

$$\begin{aligned}\mathbb{P}(X = n + 1) &= \mathbb{P}(X = n)(1 - \mathbb{P}(X = 1)) \\ &= p(1 - p)^{n-1}(1 - p) \\ &= p(1 - p)^n,\end{aligned}$$

wobei wir im ersten Schritt das Resultat der vorherigen Folie genutzt haben.

Lösung

Da X nur Werte in $\mathbb{N}$ annimmt, folgt aus

$$\mathbb{P}(X = n) = p(1-p)^{n-1}, \quad \forall n \in \mathbb{N},$$

dass

$$\mathbb{P}_X = \sum_{n=1}^{\infty} p(1-p)^{n-1} \delta_n$$

gilt und somit folgt

$$X \sim \mathrm{Geo}(p).$$