

1.3 Kolmogorov-Axiome und σ -Algebren

1.3.1 σ -Algebren

Definition 88

Sei Ω eine Menge. Eine Menge $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ heißt σ -Algebra über Ω , wenn folgende Eigenschaften erfüllt sind:

(E1) $\Omega \in \mathcal{A}$.

(E2) Wenn $A \in \mathcal{A}$, dann folgt $\bar{A} \in \mathcal{A}$.

(E3) Für $n \in \mathbb{N}$ sei $A_n \in \mathcal{A}$. Dann gilt auch $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Für jede (endliche) Menge Ω stellt die Menge $\mathcal{P}(\Omega)$ eine σ -Algebra dar.

Für $\Omega = \mathbb{R}$ ist die Klasse der *Borel'schen Mengen*, die aus allen Mengen $A \subseteq \mathbb{R}$ besteht, welche sich durch (abzählbare) Vereinigungen und Schnitte von Intervallen (offen, halboffen oder geschlossen) darstellen lassen, eine σ -Algebra.

1.3.2 Kolmogorov-Axiome

Definition 89 (Wahrscheinlichkeitsraum, Kolmogorov-Axiome)

Sei Ω eine beliebige Menge und \mathcal{A} eine σ -Algebra über Ω . Eine Abbildung

$$\Pr[\cdot] : \mathcal{A} \rightarrow [0, 1]$$

heißt **Wahrscheinlichkeitsmaß** auf \mathcal{A} , wenn sie folgende Eigenschaften besitzt:

- 1 (W1) $\Pr[\Omega] = 1$.
- 2 (W2) A_1, A_2, \dots seien paarweise disjunkte Ereignisse. Dann gilt

$$\Pr \left[\bigcup_{i=1}^{\infty} A_i \right] = \sum_{i=1}^{\infty} \Pr[A_i].$$

Für ein Ereignis $A \in \mathcal{A}$ heißt $\Pr[A]$ **Wahrscheinlichkeit** von A . Ein **Wahrscheinlichkeitsraum** ist definiert durch das Tupel $(\Omega, \mathcal{A}, \Pr)$.

Die in obiger Definition aufgelisteten Eigenschaften eines Wahrscheinlichkeitsmaßes wurden von dem russischen Mathematiker **Andrei Nikolaevich Kolmogorov** (1903–1987) formuliert. Kolmogorov gilt als einer der Pioniere der modernen Wahrscheinlichkeitstheorie, leistete jedoch auch bedeutende Beiträge zu zahlreichen anderen Teilgebieten der Mathematik. Informatikern begegnet sein Name auch im Zusammenhang mit der so genannten Kolmogorov-Komplexität, einem relativ jungen Zweig der Komplexitätstheorie.

Die Eigenschaften in obiger Definition nennt man auch **Kolmogorov-Axiome**.

Lemma 90

Sei $(\Omega, \mathcal{A}, \Pr)$ ein Wahrscheinlichkeitsraum. Für Ereignisse A, B, A_1, A_2, \dots gilt

- 1 $\Pr[\emptyset] = 0, \Pr[\Omega] = 1.$
- 2 $0 \leq \Pr[A] \leq 1.$
- 3 $\Pr[\bar{A}] = 1 - \Pr[A].$
- 4 *Wenn $A \subseteq B$, so folgt $\Pr[A] \leq \Pr[B].$*

Lemma 90

- 5 (Additionssatz) Wenn die Ereignisse A_1, \dots, A_n paarweise disjunkt sind, so folgt

$$\Pr \left[\bigcup_{i=1}^n A_i \right] = \sum_{i=1}^n \Pr[A_i].$$

Für disjunkte Ereignisse A, B erhalten wir insbesondere

$$\Pr[A \cup B] = \Pr[A] + \Pr[B].$$

Für eine unendliche Menge von paarweise disjunkten Ereignissen A_1, A_2, \dots gilt analog

$$\Pr \left[\bigcup_{i=1}^{\infty} A_i \right] = \sum_{i=1}^{\infty} \Pr[A_i].$$

Beweis:

Wenn wir in Eigenschaft (W2) $A_1 = \Omega$ und $A_2, A_3, \dots = \emptyset$ setzen, so ergibt die Eigenschaft, dass $\Pr[\Omega] + \sum_{i=2}^{\infty} \Pr[\emptyset] = \Pr[\Omega]$.

Daraus folgt $\Pr[\emptyset] = 0$.

Regel 2 und Regel 5 gelten direkt nach Definition der Kolmogorov-Axiome und Regel 1.

Regel 3 erhalten wir mit Regel 5 wegen

$$1 = \Pr[\Omega] = \Pr[A] + \Pr[\bar{A}].$$

Für Regel 4 betrachten wir die disjunkten Ereignisse A und $C := B \setminus A$, für die gilt, dass $A \cup B = A \cup C$. Mit Regel 5 folgt die Behauptung. □

1.3.3 Lebesgue-Integrale

Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt **messbar**, falls das Urbild jeder Borel'schen Menge ebenfalls eine Borel'sche Menge ist.

Z.B. ist für jede Borel'sche Menge A die Indikatorfunktion

$$I_A : x \mapsto \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{sonst} \end{cases}$$

messbar. Jede stetige Funktion ist messbar. Auch Summen und Produkte von messbaren Funktionen sind wiederum messbar. Jeder messbaren Funktion kann man ein Integral, das so genannte **Lebesgue-Integral**, geschrieben $\int f \, d\lambda$, zuordnen.

Ist $f : \mathbb{R} \rightarrow \mathbb{R}_0^+$ eine messbare Funktion, so definiert

$$\Pr : A \mapsto \int f \cdot I_A \, d\lambda$$

eine Abbildung auf den Borel'schen Mengen, die die Eigenschaft (W2) der Kolmogorov-Axiome erfüllt. Gilt daher zusätzlich noch $\Pr[\mathbb{R}] = 1$, so definiert f auf natürliche Weise einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \Pr)$, wobei $\Omega = \mathbb{R}$ und \mathcal{A} die Menge der Borel'schen Mengen ist.

1.4 Rechnen mit kontinuierlichen Zufallsvariablen

1.4.1 Funktionen kontinuierlicher Zufallsvariablen

Sei $Y := g(X)$ mit einer Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$.

Die Verteilung von Y erhalten wir durch

$$F_Y(y) = \Pr[Y \leq y] = \Pr[g(X) \leq y] = \int_C f_X(t) \, dt.$$

Hierbei bezeichnet $C := \{t \in \mathbb{R} \mid g(t) \leq y\}$ alle reellen Zahlen $t \in \mathbb{R}$, für welche die Bedingung „ $Y \leq y$ “ zutrifft. Das Integral über C ist nur dann sinnvoll definiert, wenn C ein zulässiges Ereignis darstellt. Aus der Verteilung F_Y können wir durch Differenzieren die Dichte f_Y ermitteln.

Beispiel 91

Sei X gleichverteilt auf dem Intervall $]0, 1[$. Für eine Konstante $\lambda > 0$ definieren wir die Zufallsvariable $Y := -(1/\lambda) \ln X$.

$$\begin{aligned} F_Y(y) &= \Pr[-(1/\lambda) \ln X \leq y] = \Pr[\ln X \geq -\lambda y] \\ &= \Pr[X \geq e^{-\lambda y}] \\ &= 1 - F_X(e^{-\lambda y}) \\ &= \begin{cases} 1 - e^{-\lambda y} & \text{für } y \geq 0, \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Beispiel (Forts.)

Damit folgt mit $f_Y(y) = F'_Y(y)$ sofort

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y} & \text{für } y \geq 0, \\ 0 & \text{sonst.} \end{cases}$$

Eine Zufallsvariable mit einer solchen Dichte f_Y nennt man *exponentialverteilt*.

Beispiel 92

Sei X eine beliebige Zufallsvariable. Für $a, b \in \mathbb{R}$ mit $a > 0$ definieren wir die Zufallsvariable $Y := a \cdot X + b$.

Es gilt

$$F_Y(y) = \Pr[aX + b \leq y] = \Pr\left[X \leq \frac{y-b}{a}\right] = F_X\left(\frac{y-b}{a}\right),$$

und somit

$$f_Y(y) = \frac{d F_Y(y)}{d y} = \frac{d F_X((y-b)/a)}{d y} = f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}.$$

Simulation von Zufallsvariablen

Unter der **Simulation** einer Zufallsvariablen X mit Dichte f_X versteht man die algorithmische Erzeugung von Zufallswerten, deren Verteilung der Verteilung von X entspricht.

Dazu nehmen wir an, dass die zu simulierende Zufallsvariable X eine stetige, im **Bildbereich** $]0, 1[$ streng monoton wachsende Verteilungsfunktion F_X besitzt. Weiter nehmen wir an, dass U eine auf $]0, 1[$ gleichverteilte Zufallsvariable ist, die wir simulieren können.

Aus unserer Annahme über F_X folgt, dass es zu F_X eine (eindeutige) inverse Funktion F_X^{-1} gibt mit $F_X(F_X^{-1}(x)) = x$ für alle $x \in]0, 1[$.

Sei nun

$$\tilde{X} := F_X^{-1}(U),$$

dann gilt

$$\begin{aligned}\Pr[\tilde{X} \leq t] &= \Pr[F_X^{-1}(U) \leq t] \\ &= \Pr[U \leq F_X(t)] \\ &= F_U(F_X(t)) \\ &= F_X(t).\end{aligned}$$

Beispiel 93

Im obigen Beispiel der Exponentialverteilung gilt $F_X(t) = 1 - e^{-t}$ für $t \geq 0$, und wir erhalten auf $]0, 1[$ die Umkehrfunktion $F_X^{-1}(t) = -\ln(1 - t)$. Also gilt $\tilde{X} = F_X^{-1}(U) = -\ln(1 - U)$.

Statt \tilde{X} haben wir im Beispiel die Zufallsvariable $-\ln U$ betrachtet, die aber offensichtlich dieselbe Verteilung besitzt.

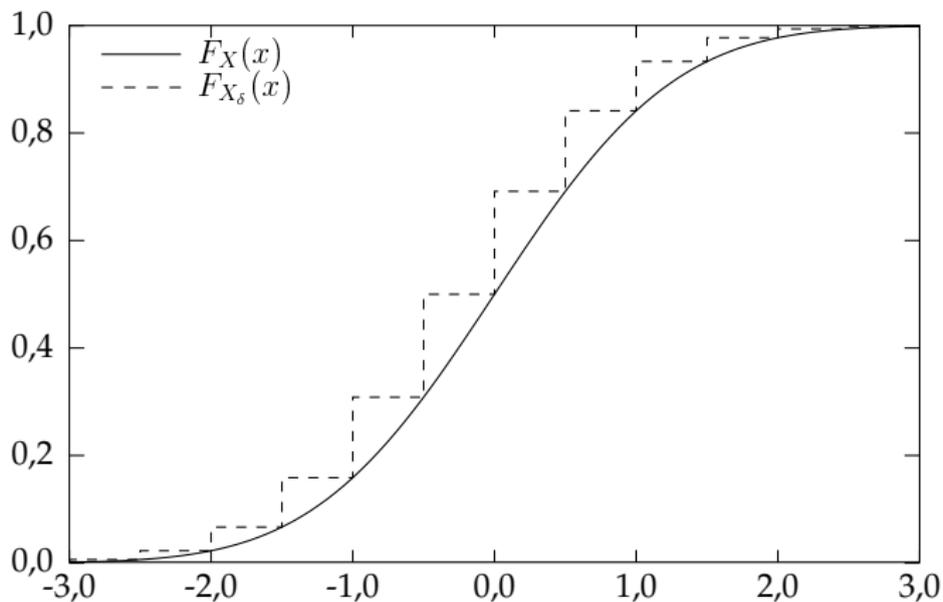
1.4.2 Kontinuierliche Zufallsvariablen als Grenzwerte diskreter Zufallsvariablen

Sei X eine kontinuierliche Zufallsvariable. Wir können aus X leicht eine diskrete Zufallsvariable konstruieren, indem wir für ein festes $\delta > 0$ definieren

$$X_\delta = n\delta \iff X \in [n\delta, (n+1)\delta[\text{ für } n \in \mathbb{Z}.$$

Für X_δ gilt

$$\Pr[X_\delta = n\delta] = F_X((n+1)\delta) - F_X(n\delta).$$



Für $\delta \rightarrow 0$ nähert sich die Verteilung von X_δ der Verteilung von X immer mehr an.

1.4.3 Erwartungswert und Varianz

Definition 94

Für eine kontinuierliche Zufallsvariable X ist der Erwartungswert definiert durch

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} t \cdot f_X(t) \, dt,$$

sofern das Integral $\int_{-\infty}^{\infty} |t| \cdot f_X(t) \, dt$ endlich ist.

Für die Varianz gilt entsprechend

$$\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \int_{-\infty}^{\infty} (t - \mathbb{E}[X])^2 \cdot f_X(t) \, dt,$$

wenn $\mathbb{E}[(X - \mathbb{E}[X])^2]$ existiert.

Lemma 95

Sei X eine kontinuierliche Zufallsvariable, und sei

$$Y := g(X).$$

Dann gilt

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} g(t) \cdot f_X(t) dt.$$

Beweis:

Wir zeigen die Behauptung nur für den einfachen Fall, dass g eine lineare Funktion ist, also $Y := a \cdot X + b$ für $a, b \in \mathbb{R}$ und $a > 0$.

Es gilt (siehe obiges Beispiel)

$$\mathbb{E}[a \cdot X + b] = \int_{-\infty}^{\infty} t \cdot f_Y(t) \, dt = \int_{-\infty}^{\infty} t \cdot f_X\left(\frac{t-b}{a}\right) \cdot \frac{1}{a} \, dt.$$

Durch die Substitution $u := (t-b)/a$ mit $du = (1/a) dt$ erhalten wir

$$\mathbb{E}[a \cdot X + b] = \int_{-\infty}^{\infty} (au + b) f_X(u) \, du.$$



Beispiel 96

Für Erwartungswert und Varianz der Gleichverteilung ergibt sich

$$\begin{aligned}\mathbb{E}[X] &= \int_a^b t \cdot \frac{1}{b-a} \, dt = \frac{1}{b-a} \cdot \int_a^b t \cdot dt \\ &= \frac{1}{2(b-a)} \cdot [t^2]_a^b \\ &= \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2},\end{aligned}$$

$$\mathbb{E}[X^2] = \frac{1}{b-a} \cdot \int_a^b t^2 \cdot dt = \frac{b^2 + ba + a^2}{3},$$

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \dots = \frac{(a-b)^2}{12}.$$