Repetition der diskreten Wahrscheinlichkeitstheorie

1.1 Zufallsexperiment, Ereignis, Wahrscheinlichkeit

Definition 1.1 *Ein* endliches (diskretes) Zufallsexperiment *ist beschrieben durch die endliche (abzählbar unendliche) Menge $\mathcal{E}$ der Elementarereignisse sowie durch ein Wahrscheinlichkeitsmass P. Dies ist eine Funktion von $\mathcal{E}$ auf die nicht-negativen reellen Zahlen $\mathbb{R}^+$, d.h. $P: \mathcal{E} \to \mathbb{R}^+$, für die gilt*

$$\sum_{e \in \mathcal{E}} P(e) = 1.$$

Definition 1.2 *Ein* Ereignis $\mathcal{A}$ *ist eine Teilmenge von $\mathcal{E}$ (die leere Menge $\emptyset$ und $\mathcal{E}$ selbst sind zulässig). Das Wahrscheinlichkeitsmass P kann auf natürliche Weise erweitert werden zu einer Funktion, die jedem Ereignis eine Wahrscheinlichkeit zuordnet:*

$$P(\mathcal{A}) = \sum_{e \in \mathcal{A}} P(e).$$

Es gilt

$$P(\mathcal{E}) = 1 \quad \text{und} \quad P(\emptyset) = 0$$

sowie

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B}).$$

Definition 1.3 *Zwei Ereignisse $\mathcal{A}$ und $\mathcal{B}$ heissen (statistisch) unabhängig falls $P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A}) \cdot P(\mathcal{B})$.*

Definition 1.4 *Die* bedingte Wahrscheinlichkeit *des Ereignisses $\mathcal{A}$, gegeben das Ereignis $\mathcal{B}$, ist definiert als*

$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})},$$

falls $P(\mathcal{B}) > 0$, und ist nicht definiert wenn $P(\mathcal{B}) = 0$.

1.2 Zufallsvariablen

Definition 1.5 *Eine* diskrete Zufallsvariable X *ist eine Abbildung von der Menge $\mathcal{E}$ der Elementarereignisse auf eine Wertemenge $\mathcal{X}$ mit endlich vielen oder abzählbar unendlich vielen Elementen. Die Wahrscheinlichkeitsverteilung von X ist die Funktion $P_X : \mathcal{X} \to \mathbb{R}^+$ definiert durch*

$$P_X(x) = \sum_{e \in \mathcal{E}: X(e) = x} P(e)$$

Man schreibt auch $P(X = x)$ statt $P_X(x)$, wobei $X = x$ das Ereignis $\{e \in \mathcal{E} : X(e) = x\}$ ist, d.h. das Ereignis „die Zufallsvariable X nimmt den Wert x an".

Mehrere Zufallsvariablen können als eine einzige Zufallsvariable (oder Zufallsvektor) betrachtet werden. Ein Beispiel ist das Paar $[X, Y]$, dessen Wahrscheinlichkeitsverteilung $P_{XY} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^+$ eine Funktion von zwei Variablen ist und durch

$$P_{XY}(x, y) \; = \; P(X = x, Y = y)$$

gegeben ist. $P(X = x, Y = y)$ steht für $P(\{e \in \mathcal{E} : X(e) = x \text{ und } Y(e) = y\})$. Die Wahrscheinlichkeitsverteilungen von X und Y sind durch P_{XY} eindeutig bestimmt:

$$P_X(x) = \sum_y P_{XY}(x, y)$$

(und analog $P_Y(y) = \sum_x P_{XY}(x, y)$). Dieses Prinzip der Elimination einer Zufallsvariablen in der Verteilung durch Summation über alle ihre Werte gilt allgemein.

Definition 1.6 *Die* bedingte *Wahrscheinlichkeitsverteilung der Zufallsvariablen X, gegeben das Ereignis $\mathcal{A}$ mit $P(\mathcal{A}) > 0$, ist definiert als*

$$P_{X|\mathcal{A}}(x) = P(X = x|\mathcal{A}).$$

Die bedingte Wahrscheinlichkeitsverteilung *von X, gegeben Y, ist wie folgt definiert:*

$$P_{X|Y} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^+ : (x, y) \mapsto P_{X|Y}(x, y),$$

wobei

$$P_{X|Y}(x, y) \;=\; P(X = x|Y = y) \;=\; \frac{P(X = x, Y = y)}{P(Y = y)}.$$

Für Werte y mit $P_Y(y) = 0$ ist $P_{X|Y}(x, y)$ nicht definiert.

Man beachte, dass $P_{X|Y}$ eine reellwertige Funktion von *zwei* Argumenten ist. Um dies zu verdeutlichen, verwenden wir nicht die in der Literatur ebenfalls übliche Notation $P_{X|Y}(x|y)$, welche die Bedingung in der Argumentenliste nochmals wiederholt. Man beachte ferner, dass für jedes y gilt:

$$\sum_x P_{X|Y}(x, y) = 1,$$

d.h. $P_{X|Y}(\cdot, y)$ ist, als ein-argumentige Funktion aufgefasst, selbst eine zulässige (normierte) Wahrscheinlichkeitsverteilung (falls $P_Y(y) > 0$).

Definition 1.7 *Zwei Zufallsvariablen X und Y heissen* statistisch unabhängig *genau dann wenn*

$$P_{XY}(x, y) = P_X(x) \cdot P_Y(y)$$

für alle $x \in \mathcal{X}$ und $y \in \mathcal{Y}$.

Es folgt aus dieser Definition, dass für statistisch unabhängige Zufallsvariablen X und Y gilt: $P_{X|Y}(x, y) = P_X(x)$ für alle $x \in \mathcal{X}$ und $y \in \mathcal{Y}$ mit $P_Y(y) \neq 0$.

1.3 Erwartungswert und Varianz

Von speziellem Interesse sind Zufallsvariablen X, die als Werte *reelle Zahlen* annehmen. Der *Erwartungswert* $E(X)$ und die *Varianz* $\mathrm{Var}(X)$ von X sind dann wie folgt definiert:

$$E(X) = \sum_x x P_X(x)$$

und

$$\mathrm{Var}(X) = \sum_x (x - E(X))^2 P_X(x).$$

Es ist einfach zu zeigen, dass

$$E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$$

gilt. Falls $X_1, \ldots, X_n$ (auch nur paarweise) statistisch unabhängig sind, gilt

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n).$$

Für eine reellwertige Funktion f, deren Definitionsbereich $\mathcal{X}$ einschliesst, gelten

$$E(f(X)) = \sum_x f(x) P_X(x)$$

und

$$\mathrm{Var}(f(X)) = \sum_x (f(x) - E(f(X)))^2 P_X(x).$$