Statistik I für Human- und SozialwissenschaftlerInnen

Ergänzung zu Zufallsvariablen

Prof. Dr. Michael Kohler
Fachbereich Mathematik
Technische Universität Darmstadt

kohler@mathematik.tu-darmstadt.de

1. Ein Beispiel

Beispiel: Student S. beschließt, seinen Lebensunterhalt durch Betreiben eines Glücksrads auf dem Münchner Oktoberfest aufzubessern.

Nach Drehen bleibt dieses rein zufällig auf einem von **64 Feldern** stehen. Bleibt es auf einem der **fünf braun gefärbten Felder** stehen, so erhält der Spieler einen Mohrenkopf (Wert **20 Cent**). Bleibt es auf einem der **beiden rot gefärbten Felder** stehen, so so erhält der Spieler eine rote Rose (Wert **3 Euro**). Und bleibt es auf dem **einzigen schwarzen Feld** stehen, so erhält der Spieler das Buch Statistik - Der Weg zur Datenanalyse von Fahrmeir, Künstler, Pigeot und Tutz, Springer 1998 (Wert ca. **25 Euro**). Auf den 56 übrigen weißen Feldern wird kein Gewinn ausgegeben.

Wie modelliert man den zufälligen Gewinn stochastisch ?

2. Zufallsvariablen

Wir beschreiben Zufallsexperimente mit zufälligem Ergebnis $X(\omega) \in \mathbb{R}$ durch Festlegung von

$$\mathbf{P}[X \in B] = \text{ Wk., dass zufälliger Ergebnis } X(\omega) \text{ in } B \text{ liegt}$$

für $B \subseteq \mathbb{R}$.

Im obigen Beispiel:

Hier nimmt X nur endlich viele Werte an, nämlich nur die Werte $0, 20, 300$ und 2500 (in Cent). Wir bestimmen

$$\mathbf{P}[X = x]$$

für jeden dieser Werte:

Für $x = 300$: X nimmt den Wert 300 an, wenn das Glücksrad auf einem der 2 roten Felder stehenbleibt. Dass genau eines der beiden roten Felder von den insgesamt 64 Feldern auftritt, kommt mit Wk. $2/64$ vor. Daher gilt:

$$\mathbf{P}[X = 300] = \frac{2}{64}.$$

Analog bestimmt man

$$\mathbf{P}[X=0]=\frac{56}{64}, \mathbf{P}[X=20]=\frac{5}{64} \text{ und } \mathbf{P}[X=2500]=\frac{1}{64}.$$

Für alle anderen Werte $x \in \mathbb{R}$ gilt $\mathbf{P}[X=x]=0$.Wir setzen dann z.B.

$$\mathbf{P}[X \in \{20, 2500\}] = \mathbf{P}[X=20]+\mathbf{P}[X=2500]=\frac{5}{64}+\frac{1}{64}=\frac{6}{64}.$$

Allgemein:

$$\mathbf{P}[X \in B] = \sum_{k \in \mathbb{N}_0 \cap B} \mathbf{P}[X=k] = \sum_{k \in \{0,20,300,2500\} \cap B} \mathbf{P}[X=k]$$

für $B \subseteq \mathbb{R}$.

Formal ist X eine sogenannte Zufallsvariable, die definiert werden kann wie folgt:

Wir beschreiben das Drehen am Glücksrad durch einen Laplaceschen W-Raum $(\Omega, \mathbf{P})$ mit

$$\Omega = \{1, 2, \dots, 64\} \quad \text{und} \quad \mathbf{P}(A) = \frac{|A|}{|\Omega|} \quad (A \subseteq \Omega).$$

Hierbei ist $\omega \in \Omega$ die Nummer des Feldes, auf dem das Glücksrad stehenbleibt. Die Felder 1 bis 5 seien braun, die Felder 6 und 7 seien rot, Feld 8 sei schwarz und die Felder 9 bis 64 seien weiß.

Der bei Auftreten von Feld ω ausgezahlte Gewinn ist gegeben durch

$$X(\omega) = \begin{cases} 20 & \text{für} \quad \omega \in \{1, \dots, 5\} \\ 300 & \text{für} \quad \omega \in \{6, 7\} \\ 2500 & \text{für} \quad \omega = 8 \\ 0 & \text{für} \quad \omega \in \{9, 10, \dots, 64\}. \end{cases}$$

Die Wahrscheinlichkeit, dass $X(\omega)$ in der Menge $\{20, 2500\}$ landet, wird dann festgelegt gemäß

$$\begin{aligned}\mathbf{P}[X \in \{20, 2500\}] &= \mathbf{P}(\{\omega \in \Omega \; : \; X(\omega) \in \{20, 2500\}\}) \\ &= \mathbf{P}(\{1, 2, 3, 4, 5, 8\}) = \frac{6}{64}.\end{aligned}$$

Allgemein: Die Wahrscheinlichkeit, dass $X(\omega)$ in einer Menge $B \subseteq \mathbb{R}$ landet, wird dann festgelegt gemäß

$$\mathbf{P}_X(B) := \mathbf{P}[X \in B] := \mathbf{P}(\{\omega \in \Omega \; : \; X(\omega) \in B\})$$

Bezeichnung: $\mathbf{P}_X$ heißt **Verteilung** der Zufallsvariablen X.

Oben gilt:

$$\mathbf{P}[X \in \{20, 2500\}] = \frac{6}{64} = \mathbf{P}[X = 20] + \mathbf{P}[X = 2500].$$

Allgemein kann man zeigen:

$$\mathbf{P}[X \in B] = \sum_{k \in \{0,20,300,2500\} \cap B} \mathbf{P}[X = k]$$

für $B \subseteq \mathbb{R}$.

Die Abbildung X oben ist ein Beispiel für eine sogenannte **diskrete Zufallsvariable.** Häufig auftretende weitere Beispiele dafür sind:

1. Seien $n \in \mathbb{N}$ und $p \in [0, 1]$. Eine $b(n, p)$**-verteilte Zufallsvariable** ist eine diskrete Zufallsvariable X, die mit Wahrscheinlichkeit Eins nur Werte aus $\{0, 1, \ldots, n\}$ annimmt und für die gilt:

$$\mathbf{P}[X = k] = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{für } k \in \{0, \ldots, n\}.$$

2. Sei $\lambda \in \mathbb{R}_+$. Eine $\pi(\lambda)$-verteilte Zufallsvariable ist eine diskrete Zufallsvariable X, die mit Wahrscheinlichkeit Eins nur Werte aus $\mathbb{N}_0$ annimmt und für die gilt:

$$\mathbf{P}[X = k] = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad \text{für } k \in \mathbb{N}_0.$$