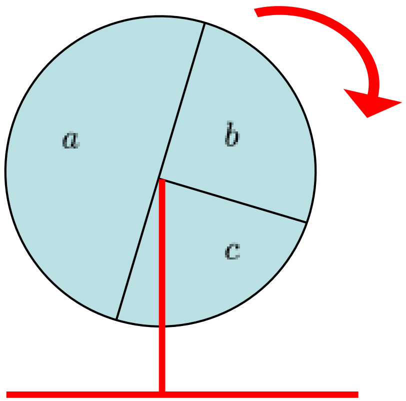
3. diskrete Zufallsvariablen

Diskrete Stochastik

Prof. Dr. Andreas Vogt

Fachhochschule
Nordwestschweiz

Zufallsvariablen



$\Omega = \{a, b, c\}$

P definiert durch $f_P(a) = \frac{1}{2}$, $f_P(b) = f_P(c) = \frac{1}{4}$.

Bei a gewinnt der Spieler CHF 10, sonst muss er 5 zahlen.

Jetzt interessiert man sich nur noch für seinen Gewinn.
Dies wird (oft) mit Hilfe von Zufallsvariablen modelliert.

Definition: Eine Zufallsvariable X über Ω ist eine Abbildung von Ω in eine Menge $\mathcal{X}$.

Im Folgenden wird stets $\mathcal{X} \subseteq \mathbb{R}$ sein. Dann sagt man auch reellwertige Zufallsvariable.

In obigem Beispiel: $\mathcal{X} = \{-5, 10\}$ und $X(a) = 10$, $X(b) = X(c) = -5$.

X induziert eine Wahrscheinlichkeitsverteilung P^X auf $\mathcal{X}$ durch
$P^X(A) = P\left(X^{-1}(A)\right)$, wobei $X^{-1}(A)$ das Urbild von A bezeichnet.

Für $A \subseteq \mathcal{X}$ ist $X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$.

P^X heisst Verteilung von X oder Bildmass von X unter P.

In obigem Beispiel: $P^X(\{-5\}) = P(X^{-1}\{-5\}) = P(\{b, c\}) = \frac{1}{2}$.

Andere Schreibweisen: $P(X \in A) = P^X(A)$, $A \subseteq \mathcal{X}$; $P(X = a) = P^X(\{a\})$, $a \in \mathcal{X}$.

In obigem Beipiel: $P(X = 10) \mathrel{\widehat{=}}$ Wahrscheinlichkeit dafür, dass der Gewinn 10 ist.

Falls $\mathcal{X}$ endlich oder abzählbar ist, dann heisst die Funktion $f : \mathcal{X} \to [0, 1]$, definiert durch $f(x) = P(X = x) = P^X(\{x\})$ Zähldichte von X.

Zufallsvariablen

Eine Zufallsvariable ist eine normale mathematische Funktion.

Da bei jeder Durchführung des Zufallsexperimentes ein zufälliges Ergebnis ω eintritt, ist auch der zugehörige Wert $X(\omega)$ nicht vorhersagbar.

Daher kommt der Name "Zufallsvariable", der am Anfang eher missverständlich ist. Bei praktischen Beispielen ist er passend.

Zufallsvariablen

Beispiel: Zweimaliges Würfeln: Man gewinnt CHF i bei Augensumme i.

Gesucht: Wahrscheinlichkeit für "Gewinn ist k"

Modellierung:

Zähldichte, d.h. Wahrscheinlichkeit für die Ergebnisse

Schlechte Variante:

mögliche Augensummen $\rightarrow$ $\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ $f_P(2) = \frac{1}{36}$, $f_P(3) = \frac{2}{36}$, ...

Bei dieser Modellierung erkennt man das "Konzept" nicht so gut.

Besser:

Modellierung als Kopplung zweier Laplace-Experimente

$\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$ $\quad f_P(i, j) = \frac{1}{6} \cdot \frac{1}{6}$

$X : \Omega \to \mathbb{R}$ def. durch $X(i, j) = i + j$

Gesucht ist die Verteilung von X

Es gilt etwa $P(X = 6) = P(\{(i, j) \mid i + j = 6\} =$

$$= P(\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1))\} = \frac{5}{36}.$$

Die Überlegungen sind letzlich dieselben wie in obiger Variante, aber die Annahmen (unabhängige Kopplung zweier Laplace-Experiment) wird sichtbar.

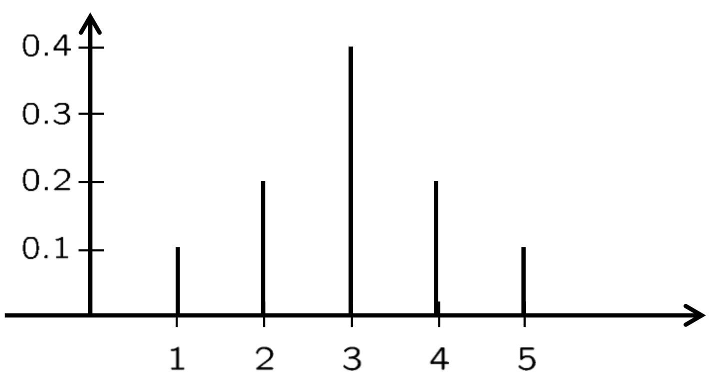
Zudem hat man nun das Modell gebaut, und kann auch andere Zufallsvariablen definieren, etwa: $Y(i, j) = i \cdot j$. ...

Zufallsvariablen

Die Zähldichte einer diskreten Zufallsvariablen bestimmt die Verteilung eindeutig, oft wird diese in einer Tabelle oder mit Hilfe eines Stabdiagramms dargestellt.

Beispiel:

x_i	1	2	3	4	5
$P(X = x_i)$	0.1	0.2	0.4	0.2	0.1



Verteilungsfunktion

Definition:

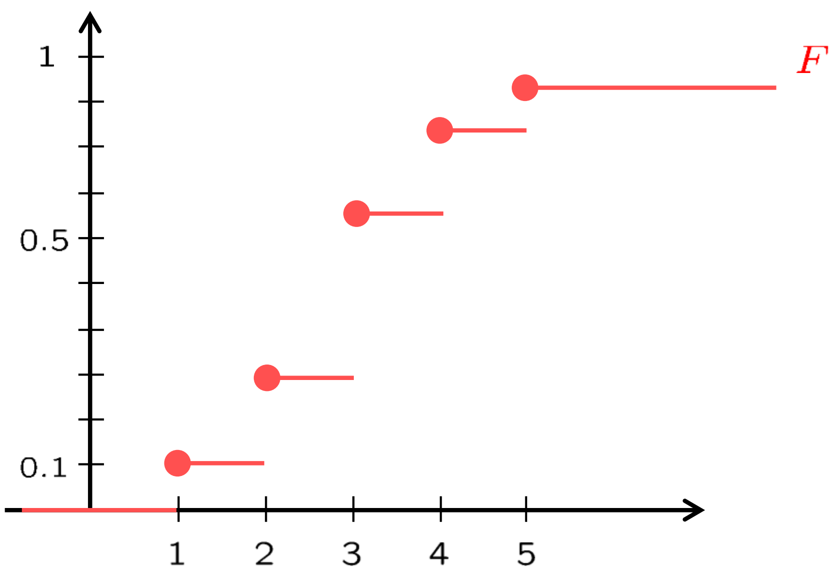
Es sei $X : \Omega \to \mathcal{X}$ eine Zufallsvariable, wobei $\mathcal{X} \subset \mathbb{R}$ eine endliche oder abzählbare Menge ist. Zudem sei f die Zähldichte von X.

Dann heisst

$$F(x) = P(X \leq x) = \sum_{t \in \mathcal{X} : t \leq x} f(t) \text{ Verteilungsfunktion von } X.$$

Beispiel:

x_i	1	2	3	4	5
$P(X = x_i)$	0.1	0.2	0.4	0.2	0.1



Auch die Verteilungsfunktion bestimmt die Verteilung eindeutig.

Lernkontrolle

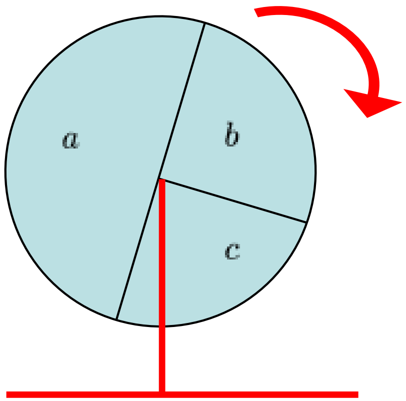
Wir werfen 4 mal mit einer ausgewogenen Münze. Die Zufallsvariable X zähle die Anzahl Köpfe.

(1) Welche Werte kann X annehmen?

(2) Geben Sie die Verteilung von X in Tabellenform an.

(3) Skizzieren Sie die Verteilungsfunktion.

Erwartungswert



$$\Omega = \{a, b, c\}$$

P definiert durch $f_p(a) = \frac{1}{2}$, $f_P(b) = f_P(c) = \frac{1}{4}$.

Bei a verliert der Betreiber CHF 10, bei b gewinnt er 5 und bei c 7.

Was ist auf lange Sicht hin der Gewinn/Verlust des Betreibers?

Bei 10.000 Spielen kommt a bei ca. 5.000 und b und c bei jeweils ca. 2.500 vor.

Dann ist der Gewinn/Verlust pro Spiel:

$$(-10CHF \cdot 5000 + 5CHF \cdot 2500 + 7CHF \cdot 2500)/10000 =$$

$$= -10CHF \cdot \frac{5000}{10000} + 5CHF \cdot \frac{2500}{10000} + 7CHF \cdot \frac{2500}{10000} =$$

$$= X(a) \cdot f_P(a) + X(b) \cdot f_P(b) + X(c) \cdot f_P(c) = -2CHF$$

Die Zufallsvariable X modelliere die Auszahlung des Betreibers.

Erwartungswert

Definition:

Es sei $X : \Omega \to \mathcal{X}$ eine Zufallsvariable, wobei $\mathcal{X} \subset \mathbb{R}$ eine endliche oder abzählbare Menge ist. Zudem sei f die Zähldichte von X.

Dann heisst $E(X) = \sum_{x \in \mathcal{X}} x \cdot f(x) = \sum_{x \in \mathcal{X}} x \cdot P(X = x)$ Erwartungswert von X.
(im Falle der Existenz)

Beispiel:

Beim zweimaligen Würfelwurf erhält man i CHF, wenn beide Würfel i zeigen, und sonst nichts. X beschreibe diese Auszahlung.

$\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$, P Gleichverteilung

$\mathcal{X} = \{0,1,2,3,4,5,6\}$, $X : \Omega \to \mathcal{X}$ def. durch $X(i,j) = \begin{cases} i, & i = j \\ 0, & \text{sonst} \end{cases}$

Dann ist

$$E(X) = \tfrac{21}{36}$$

$$E(X) = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2) + 3 \cdot P(X=3) + 4 \cdot P(X=4) + 5 \cdot P(X=5) + 6 \cdot P(X=6) = \tfrac{21}{36}$$

Varianz und Standardabweichung

Definition:

Es sei $X : \Omega \to \mathcal{X}$ eine Zufallsvariable, wobei $\mathcal{X} \subset \mathbb{R}$ eine endliche oder abzählbare Menge ist. Zudem sei f die Zähldichte von X und μ der Erwartungswert von X.

Dann heisst $V(X) = \sum_{x \in \mathcal{X}} (x-\mu)^2 \cdot f(x) \ = \sum_{x \in \mathcal{X}} (x-\mu)^2 \cdot P(X=x)$ Varianz von X.

Durch das Quadrieren heben sich Abweichungen nach unten und nach oben nicht auf, zudem werden grössere Abweichungen stärker gewichtet.

$\sigma(X) = \sqrt{V(X)}$ heisst Standardabweichung von X.

Die Varianz und die Standdardabweichung sind Masse für die Streuung der Zufallsvariable um den Erwartungswert.

Beispiel:

Zwei Hersteller produzieren Bauteile, wobei die Lebensdauern wie folgt verteilt sind:

Hersteller 1 (X_1):

x_i	1	2	3	4	5
$P(X_1 = x_i)$	0.1	0.2	0.4	0.2	0.1

Erwartungswert $E(X_1) = 3$

$$V(X_1) = (1-3)^2 \cdot 0.1 + (2-3)^2 \cdot 0.2 + (3-3)^2 \cdot 0.4 + (4-3)^2 \cdot 0.2 + (5-3)^2 \cdot 0.1 = 1.2$$

Hersteller 2 (X_2):

x_i	1	2	3	4	5
$P(X_2 = x_i)$	0.3	0.1	0.2	0.1	0.3

Erwartungswert $E(X_2) = 3$

$$V(X_2) = (1-3)^2 \cdot 0.3 + (2-3)^2 \cdot 0.1 + (3-3)^2 \cdot 0.2 + (4-3)^2 \cdot 0.1 + (5-3)^2 \cdot 0.3 = 2.6$$

Varianz und Standardabweichung

Zur Berechnung der Varianz ist es manchmal einfacher, folgende Formel zu verwenden:

Satz:

Es sei $X : \Omega \to \mathcal{X}$ eine Zufallsvariable, wobei $\mathcal{X} \subset \mathbb{R}$ eine endliche oder abzählbare Menge ist.

Dann gilt
(im Falle der Existenz)

$$V(X) = E(X^2) - E(X)^2.$$

Serie 3, komplett