Wahrscheinlichkeitstheorie und Statistik

Kapitel 2: Zufallsvariablen

2 Zufallsvariablen (Random Variables)

Wurf 2 Würfel: Summe $X = \omega_1 + \omega_2$

$X : \mathcal{S} \longrightarrow \mathbb{R}$

ist eine Zufallsvariable (ZV, RV)

Idee: Wir interessieren uns nicht für beliebige Ereignisse, sondern nur für solche, die durch X beschrieben werden können

Etwa: Gewicht ≥ 100 kg

Höhe $< 1{,}60$ m

Temp $> 30°$

etc.

Zurück zu den Würfeln: Mögliche Werte von X sind in $\{2, \dots, 12\}$.

Ereignisse, die beschrieben werden können, sind der Art

$$\underset{\text{mögliche Werte}}{X \in B}, \quad B \subseteq \{2, \dots, 12\}$$

Einfachste Mengen

$$B = \{k\}, \quad k \in \{2, \dots, 12\}$$

W.keiten

$P[X = 2] = \frac{1}{36}$

$P[X = 3] = \frac{2}{36}$

$P[X = 4] = \frac{3}{36}$

$P[X = 5] = \frac{4}{36}$

$P[X = 6] = \frac{5}{36}$

$P[X = 7] = \frac{6}{36}$

$P[X = 8] = \frac{5}{36}$

$P[X = 9] = \frac{4}{36}$

$P[X = 10] = \frac{3}{36}$

$P[X = 11] = \frac{2}{36}$

$P[X = 12] = \frac{1}{36}$

Ereignisse in Termini von X
(in terms of)

$$P[5 \le X \le 9] = \frac{24}{36} = \frac{2}{3}$$

$$\sum_{k=5}^{9} P[X = k]$$

$$P[X = 1] = 0$$

Prinzip:

$$P[\mathfrak{X} \in B] = \sum_{k=B} P[\mathfrak{X} = k]$$

Die ZV $\mathfrak{X}$ ist

- (discrete)
  diskret, wenn $P[\mathfrak{X} = x_i] > 0$ für endlich viele $x_1, \dots, x_n$
  oder abzählbar unendlich viele $x_1, \dots, x_n, \dots$

  Bsp für diskret und unendlich: Würfeln hintereinander bis zur ersten 6 (Münzwurf bis zum ersten Kopf)

  Mögliche Werte: $1, 2, \dots, n, \dots, \infty$

- stetig, kontinuierlich (continuous), wenn sie ein Kontinuum von Werten annehmen kann

  Etwa: Größe, Gewicht, Temperatur

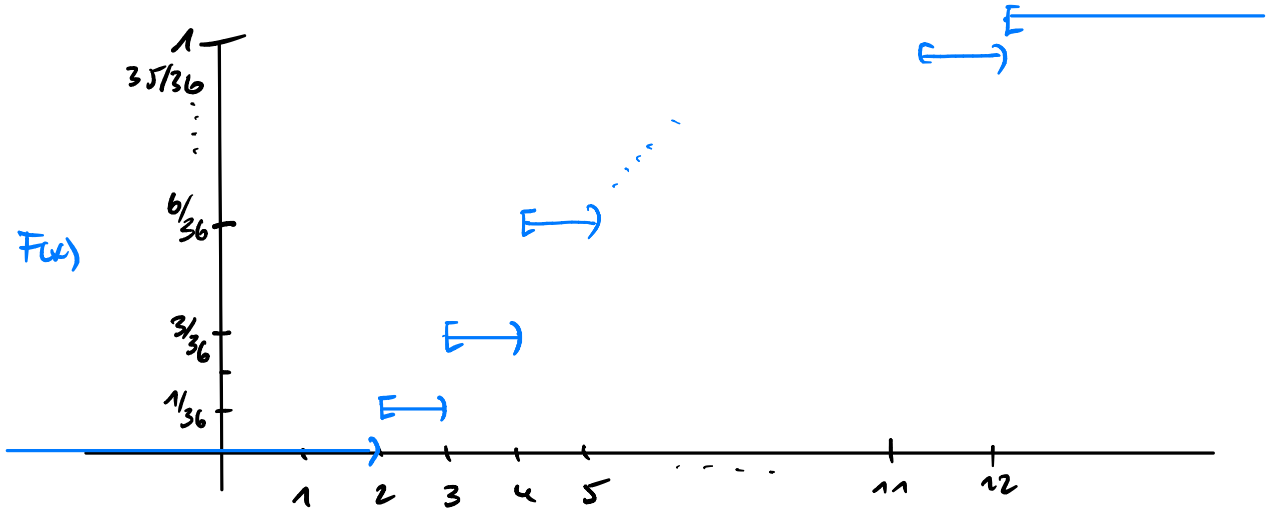
Definition 24: Die (kumulative) Verteilungsfunktion (cumulative distribution function, cdf) von $\mathfrak{X}$ ist

$$F : \underbrace{\mathbb{R}}_{\text{mögliche Werte von } \mathfrak{X}} \longrightarrow [0,1] \quad \text{mögliche W.keiten}$$

$$F(x) = P[\mathfrak{X} \leq x], \quad x \in \mathbb{R}$$

„$\mathfrak{X} \sim F$" heißt „F ist Verteilungsfkt. von $\mathfrak{X}$"

Verteilung von $\mathfrak{X} = \omega_1 + \omega_2$



$$P[\mathfrak{X} \leq x] = F(x)$$

$$P[\mathfrak{X} \leq 3] = P[\mathfrak{X} = 2] + P[\mathfrak{X} = 3]$$
$$= \frac{1}{36} + \frac{2}{36} = \frac{3}{36}$$

Beispiel: Wartezeit bei einem Prozess ohne Gedächtnis

mit Gedächtnis: Bus mit 1-stündigen Intervallen,
je länger wir warten, desto wahrscheinlicher

ohne Gedächtnis:

- ein Atom zerfällt — ideal
- ein Kunde ruft an — näherungsweise
- ein Computer ausfällt — " "

Sei X die Wartezeit:
Wenn wir s Minuten gewartet,
ist die W.keit noch t Min. zu warten,
gleich der W.keit zu Anfang t Min. zu warten

Formal: $P[X > s+t \mid X > s] = P[X > t]$

für alle $s, t \geq 0$

$$P[X > s+t \mid X > s] = P[X > t] \qquad \text{für alle } s,t \geq 0$$

bedeutet

$$\frac{P[X > s+t,\ X > s]}{P[X > s]} = P[X > t]$$

$$\Leftrightarrow \quad P[X > s+t] = P[X > s] \cdot P[X > t]$$

sei $G(s) := P[X > s]$. Dann gilt

- $G(0) = 1$ $\quad G(s) = 1$ f.a. $s < 0$
- $G(s+t) = G(s) \cdot G(t)$
- $G(s) = P[X > s] = 1 - P[X \leq s] = 1 - F_X(s)$
- $\lim_{s \to \infty} G(s) = 0$

Welche Fkt. G haben diese Eigenschaften

- $G(0) = 1$ $G(s) = 1$ f.a. $s < 0$
- $\boxed{G(s+t) = G(s) \cdot G(t)}$
- $G(s) = P[X > s] = 1 - P[X \leq s] = 1 - F_X(s)$
- $\lim_{s \to \infty} G(s) = 0$

Antwort:

$$G(s) = \begin{cases} 1, & s < 0 \\ A^s, & s \geq 0 \end{cases}$$ für ein A mit $0 < A < 1$

$$G(s) = \begin{cases} 1, & s < 0 \\ A^s, & s \geq 0 \end{cases}$$ für ein A mit $0 < A < 1$

Erinnerung: $0 < A < 1 \Rightarrow \log A < 0$

$\lambda := -\log A > 0$

Dann

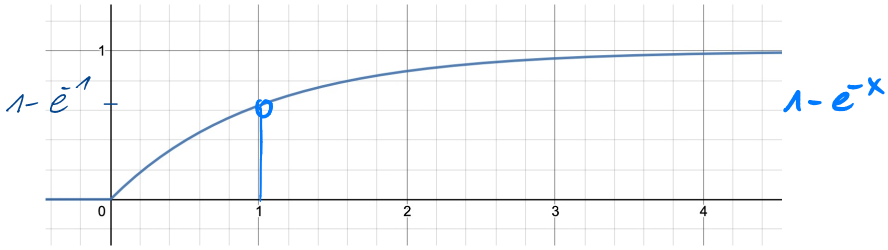
$$G(s) = A^s = e^{\log A \cdot s} = e^{-\lambda s}$$

$$F_X(s) = 1 - G(s) = 1 - e^{-\lambda s}, \quad \text{für ein } \lambda > 0$$

Verteilungsfkt. der W.keit höchstens s Min zu warten, bei einem Proz. ohne Gedächtnis

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-x} \end{cases} \qquad 1 - e^{-\lambda x}$$

$F(x) = P[X \leq x]$



$, \lambda = 1$

$e^{-x} = F'(x) = f(x)$

Dichte von F

$P[X \leq 1] = F(1) = 1 - e^{-1}$

$$\Rightarrow F(x) = \int_{-\infty}^{x} f(y)\, dy$$

Eine Verteilungsfkt F hat folgende Charakteristika

- $0 \leq F(x) \leq 1$
- F monoton wachsend (monotonically increasing)

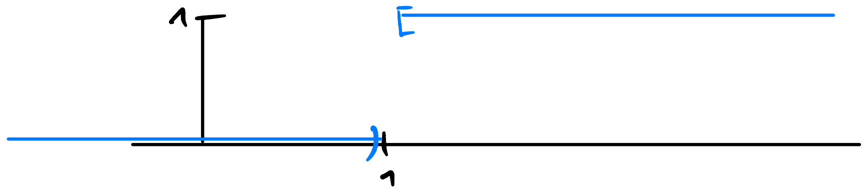
$x \leq y \Rightarrow [X \leq x] \subseteq [X \leq y]$

$\Rightarrow F(x) = P[X \leq x] \leq P[X \leq y] = F(y)$

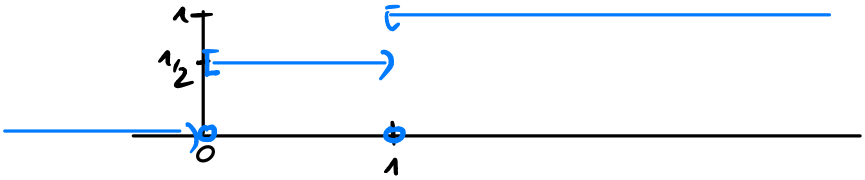
- $\lim_{x \to -\infty} F(x) = 0$ $\qquad \lim_{x \to -\infty} P[X \leq x] = P(\emptyset) = 0$
- $\lim_{x \to \infty} F(x) = 1$ $\qquad \lim_{x \to \infty} P[X \leq x] = P(\Omega) = 1$
- $\lim_{\substack{x \to +x_0 \\ x > x_0}} F(x) = F(x_0)$

Jede Funktion mit diesen Eigenschaften ist eine Verteilungsfkt.

Mögliche Formen Verteilungsfunktionen



X hat nur Wert 1,

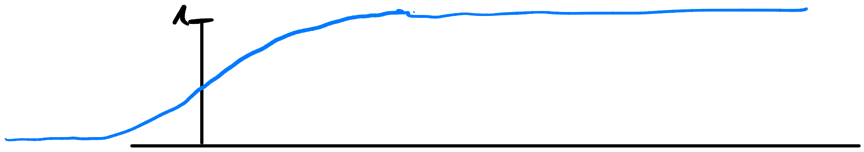


X hat Werte $\{0,1\}$

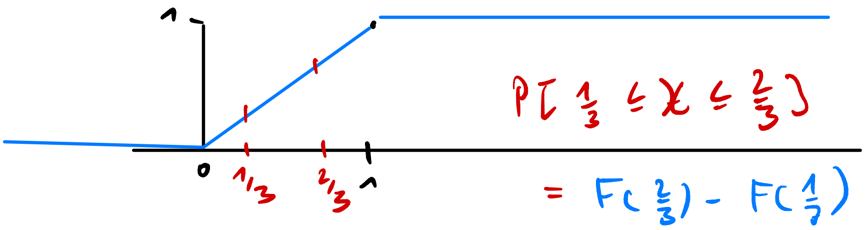
0 Zahl   faire Münze

1 Kopf

$P[X=0] = \frac{1}{2}$, $P[X=1]=\frac{1}{2}$



$P[X \leq 1] = P[X \leq y], \ y > 1$



$P[\frac{1}{3} \leq X \leq \frac{2}{3}]$

$= F(\frac{2}{3}) - F(\frac{1}{3})$

Mögliche Werte von X nur zw. 0 und 1

2.1 Arten von Zufallsvariablen

Sei X diskret. Dann ist

$$p : \mathbb{R} \longrightarrow [0,1] \qquad p(x) = P[X=x]$$

die W.keitsfunktion von X

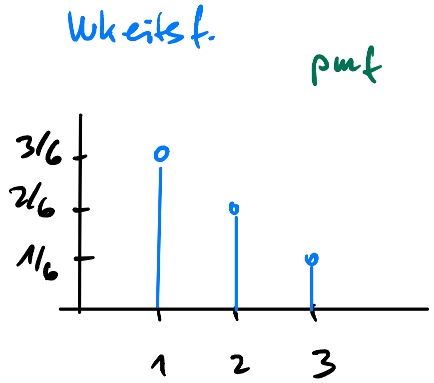
(probability mass function, pmf)

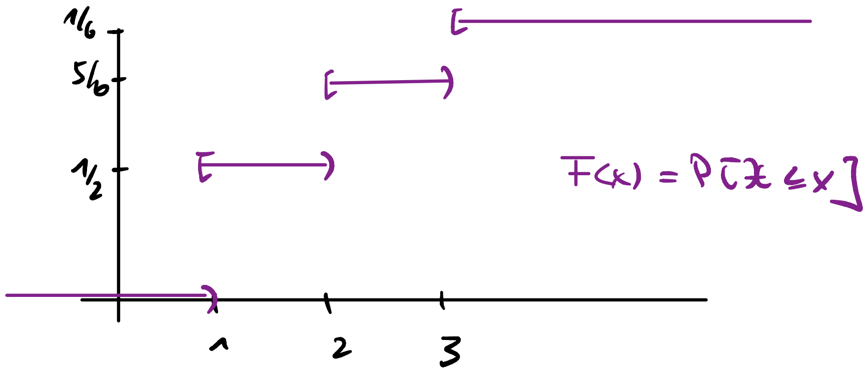
Seien $x_1, \dots, x_n, \dots$ die möglichen Werten. Dann

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

Beispiel 26: X habe die Werte $\{1, 2, 3\}$

$$p(2) = \frac{1}{3}, \quad p(3) = \frac{1}{6}, \quad p(1) = \frac{3}{6} = \frac{1}{2}$$





F konstant auf Intervallen $[x_i, x_{i+1})$

$\Rightarrow$ F Treppenfunktion (step function)

X ist stetig (continuous) wenn es eine Fkt f gibt

$$f: \mathbb{R} \to \mathbb{R}_0^+, \quad f \geq 0 \quad \text{so dass}$$

$$P[X \in B] = \int_B f(x)\, dx$$

für alle „vernünftigen" $B \in \mathbb{R}$

(im Wesentlichen, $B = \bigcup_{i=1}^{\infty} [a_i, b_i]$)

Wir nennen f eine W.keitsdichte oder Dichte

probability density fct.

X diskret, B diskret

$$P[X \in B] = \sum_{x \in B} P_X(x)$$

W.keitsfkt.
pmf