Grundlagen zur Anwendung von Zufallsvariablen in der Signalverarbeitung

Wahrscheinlichkeiten

Werfen eines Würfels

Zufallsexperiment: Reales oder hypothetisches Experiment mit unvorhersehbarem Ergebnis

Ergebnismenge: Menge Ω aller möglichen Ergebnisse

$$\Omega = \{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}$$

Ereignis: Beschreibung eines bestimmten Ausgangs des Zufallsexperimentes

Ereignis	**Ereignismenge** H
Gerade Augenzahl	$\{\epsdice{2}, \epsdice{4}, \epsdice{6}\}$
Augenzahl Drei	$\{\epsdice{3}\}$
Augenzahl größer als Vier	$\{\epsdice{5}, \epsdice{6}\}$

Wahrscheinlichkeit I

Definition der Wahrscheinlichkeit als Maß für das Auftreten eines Ereignisses:

$$P(\text{"Ereignis"}) \rightarrow [0,\ 1]$$

Berechnung der Wahrscheinlichkeit

$$P(\text{"Ereignis"}) = \frac{\text{Anzahl der Elemente in Ereignismenge}}{\text{Anzahl der Elemente in Ergebnismenge}} = \frac{|H|}{|\Omega|}$$

Beispiele:

$$P(\text{"Gerade Augenzahl"}) = \frac{3}{6} = 0.5$$

$$P(\text{"Augenzahl Drei"}) = \frac{1}{6}$$

$$P(\text{"Augenzahl größer als Vier"}) = \frac{2}{6} = \frac{1}{3}$$

Wahrscheinlichkeit II

Entspricht das Ereignis der Ergebnismenge gilt

$$P(\text{"Sicheres Ereignis"}) = 1$$

$$P(\text{"Werfen einer Zahl zwischen Eins und Sechs"}) = \frac{6}{6} = 1$$

Kein gemeinsames Element von Ereignismenge und Ergebnismenge

$$P(\text{"Unmögliches Ereignis"}) = 0$$

$$P(\text{"Werfen einer Sieben"}) = \frac{0}{6} = 0$$

Gegenteiliges Ereignis

$$P(\text{"Gegenteiliges Ereignis"}) = P(\overline{\text{"Ereignis"}}) = 1 - P(\text{"Ereignis"})$$

$$P(\text{"Augenzahl kleiner gleich Vier"}) = 1 - P(\text{"Augenzahl größer als Vier"}) = 1 - \frac{1}{3} = \frac{2}{3}$$

Verbundereignisse

Gegeben sind zwei Ereignisse A und B.

Wahrscheinlichkeit für das Auftreten von B unter der Voraussetzung das A eintritt:

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

Beispiel:

$$P(\text{"Augenzahl größer als Vier"}|\text{"Augenzahl gerade"}) =$$

$$= \frac{P\left(\{\text{⚄}, \text{⚅}\} \cap \{\text{⚁}, \text{⚃}, \text{⚅}\}\right)}{P\left(\{\text{⚁}, \text{⚃}, \text{⚅}\}\right)} = \frac{P\left(\{\text{⚅}\}\right)}{P\left(\{\text{⚁}, \text{⚃}, \text{⚅}\}\right)} = \frac{1/6}{3/6} = \frac{1}{3}$$

Satz von Bayes

Umrechnung der bedingten Wahrscheinlichkeiten

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Beispiel:

$A =$ "Augenzahl gerade"

$B =$ "Augenzahl größer als Vier"

$$P(B|A) = \frac{1}{3}$$

$P($"Augenzahl gerade" | "Augenzahl größer als Vier"$) = P(A|B) =$

$$= \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{1/3 \cdot 1/2}{1/3} = \frac{1}{2}$$

Statistische Unabhängigkeit

Zwei Ereignisse gelten als statistisch unabhängig falls

$$P(A \cap B) = P(A) \cdot P(B)$$

Beispiel: Werfen zweier Würfel W_1 und W_2

Ereignis A: "W_1 zeigt eine Drei"

$$P(A) = \frac{1}{6}$$

Ereignis B: "W_2 zeigt eine ungerade Zahl"

$$P(B) = \frac{1}{2}$$

Verbundwahrscheinlichkeit:

$$P(A \cap B) = \frac{1 \cdot 3}{6 \cdot 6} = \frac{1}{12} = P(A) \cdot P(B)$$

Ereignis A und Ereignis B sind voneinander stochastisch unabhängig.

Zufallsvariablen

Zufallsvariable

Zur Analyse von Zufallsexperimenten ist die Verwendung von Mengen unhandlich.

Verwendung einer Zufallsvariable X:

Jedem Element der Ergebnismenge Ω wird eine reelle Zahl $X \in \mathbb{R}$ zugeordnet.

Beispiel: Werfen eines Würfels

Anzahl der Augen des Wurfes entspricht der Zufallsvariablen X:

$$X \in \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$$

Wahrscheinlichkeit einer Drei:

$$P(X = 3) = \frac{1}{6}$$

Wahrscheinlichkeit einer Fünf oder Sechs:

$$P(X > 4) = \frac{2}{6} = \frac{1}{3}$$

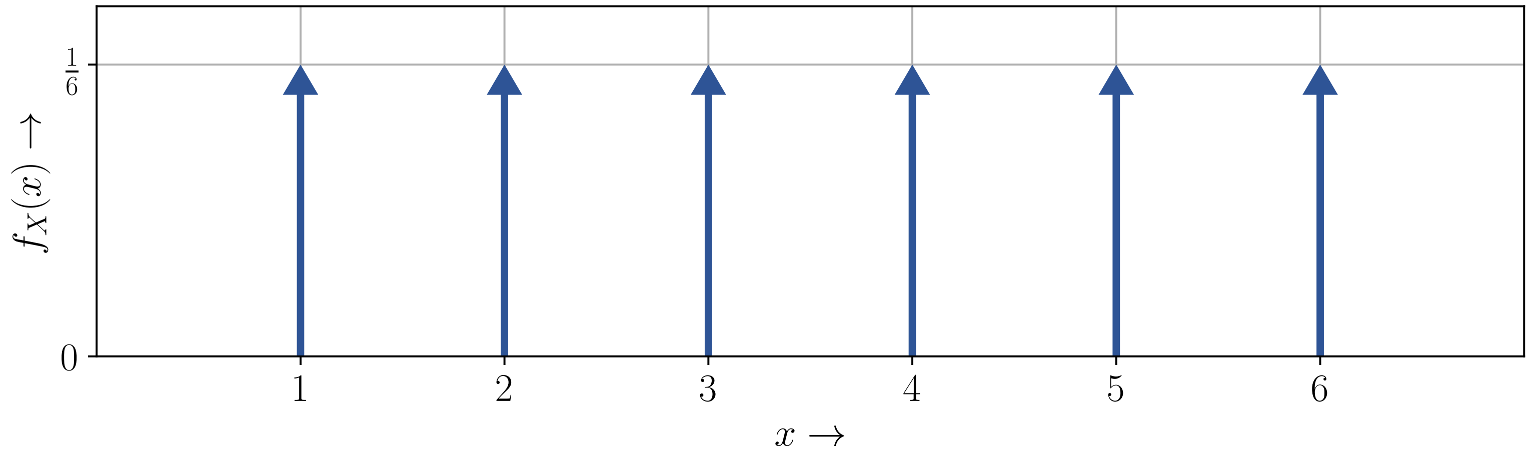
Wahrscheinlichkeitsdichte

Beschreibung aller Einzelwahrscheinlichkeiten über die *Wahrscheinlichkeitsdichtefunktion*

$$f_X(x) = \sum_i P(X = x_i)\delta(x - x_i)$$

Beispiel: Werfen eines Würfels

$$f_X(x) = \sum_{i=1}^{6} \frac{1}{6}\delta(x - i) = \frac{1}{6}\sum_{i=1}^{6}\delta(x - i)$$



Eigenschaften der Wahrscheinlichkeitsdichtefunktion

Für eine Wahrscheinlichkeitsdichtefunktion gilt immer

$$\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$$

Eigenschaften

- Summe der Wahrscheinlichkeiten sämtlicher Ereignisse muss Eins sein
- $f_X(x) \geq 0$

Wahrscheinlichkeitsverteilung

Definition der Wahrscheinlichkeitsverteilung

$$F_X(x) = \int_{-\infty}^{x} f_X(\xi)\mathrm{d}\xi$$

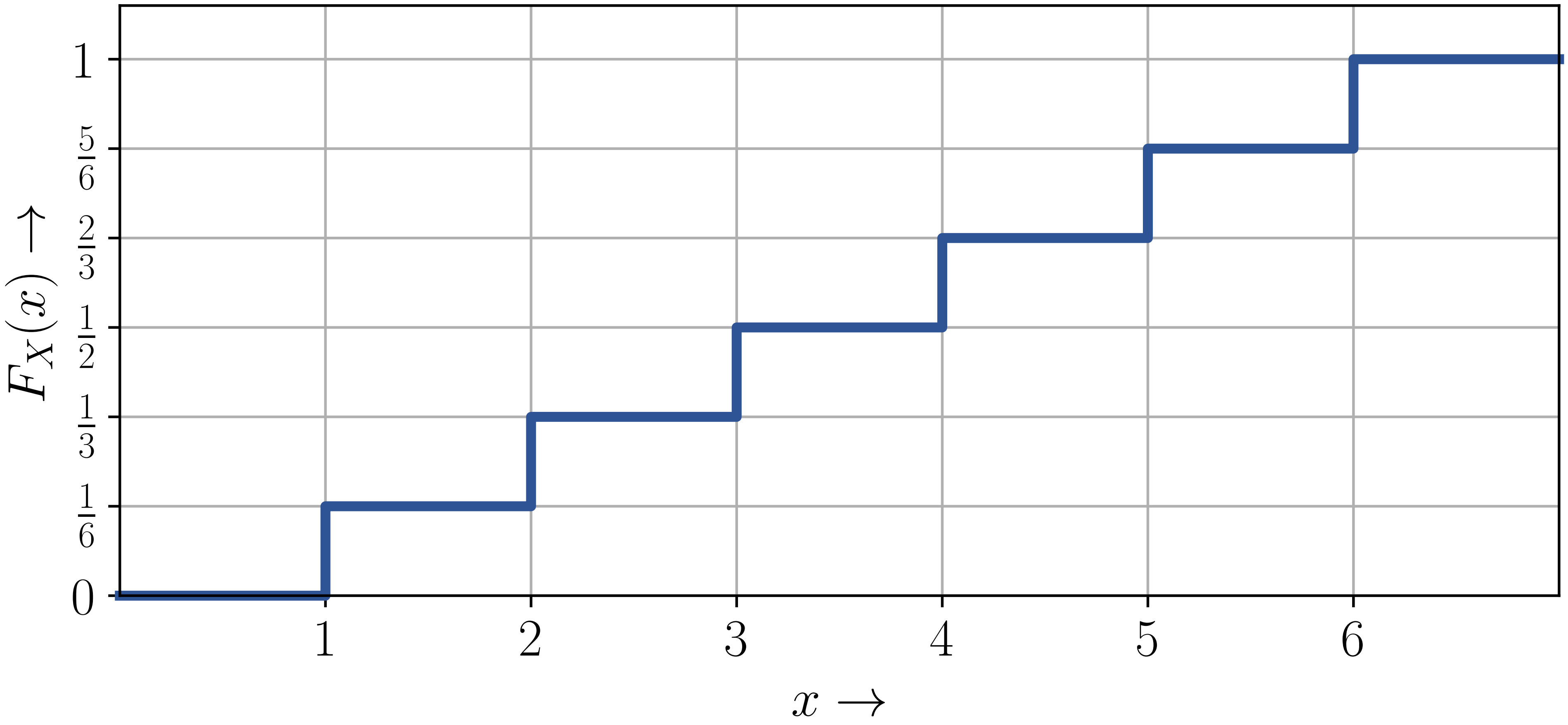
Eigenschaften der Wahrscheinlichkeitsverteilung:

$$F_X(x) \geq 0$$

$$\lim_{x \to -\infty} F_X(x) = 0$$

$$\lim_{x \to \infty} F_X(x) = 1$$

Wahrscheinlichkeitsverteilung eines Würfelwurfes



Berechnen von Wahrscheinlichkeiten mit der Wahrscheinlichkeitsverteilung

Gegeben Wahrscheinlichkeitsverteilung $F_X(x)$ einer Zufallsvariablen X

Berechnung der Wahrscheinlichkeit, dass Zufallsvariable X einen Wert x unterschreitet

$$P(X \leq x) = F_X(x)$$

Berechnung der Wahrscheinlichkeit, dass Zufallsvariable X einen Wert x überschreitet

$$P(X > x) = P(\overline{X \leq x}) = 1 - F_X(x)$$

Berechnung der Wahrscheinlichkeit, dass Zufallsvariable X innerhalb zweier Grenzen x_1 und x_2 liegt

$$P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$$

Kontinuierliche Zufallsvariable

Beispiel des Werfens eines Würfels ist ein diskretes Zufallsexperiment:

Jedem Ergebnis Zufallsvariable X kann eine konkrete Wahrscheinlichkeit zugeordnet werden.

$$X \in \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$$

Zufallsvariable X wird auch als *diskrete* Zufallsvariable bezeichnet.

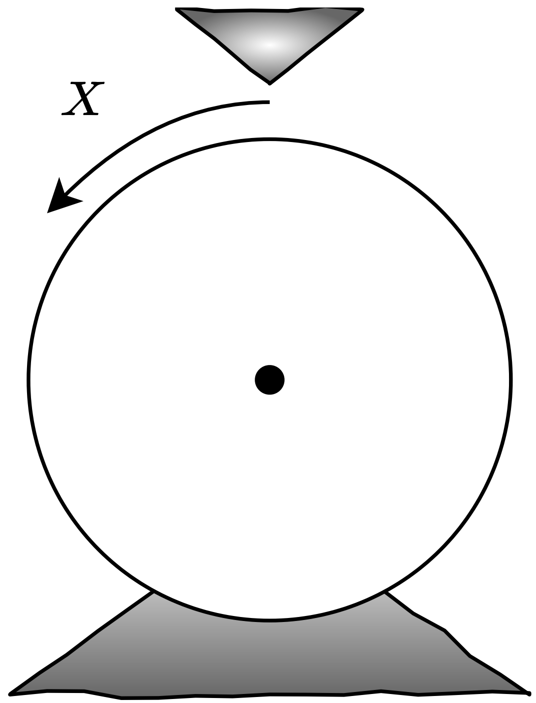
Beispiel: Drehen eines Glücksrades

Ergebnismenge:

$$X \in \{0 \ldots 2\pi\}$$

Angeben von konkreten Einzelwahrscheinlichkeiten nicht möglich

$$P(X = x) = 0 \qquad \forall x \in \{0 \ldots 2\pi\}$$

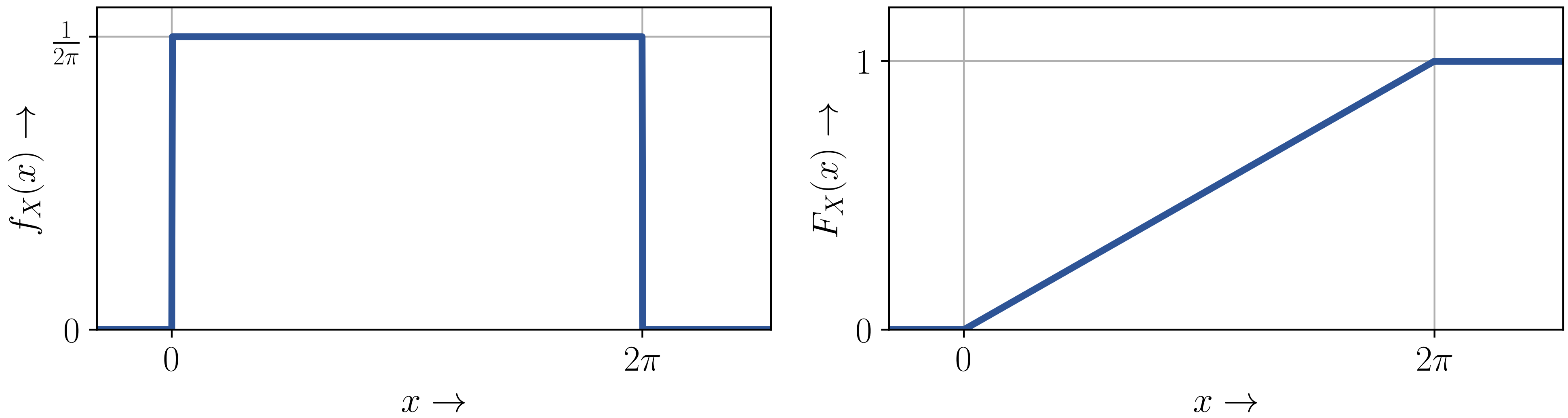


Wahrscheinlichkeitsdichte und -verteilung von kontinuierlichen Zufallsvariablen

Auch für kontinuierliche Zufallsvariable X lasse sich die

- Wahrscheinlichkeitsdichtefunktion $f_X(x)$ mit $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$
- Wahrscheinlichkeitsverteilungsfunktion $F_X(x) = \int_{-\infty}^{x} f_X(\xi)\mathrm{d}\xi$

Beispiel Glücksrad: *Gleichverteilung*



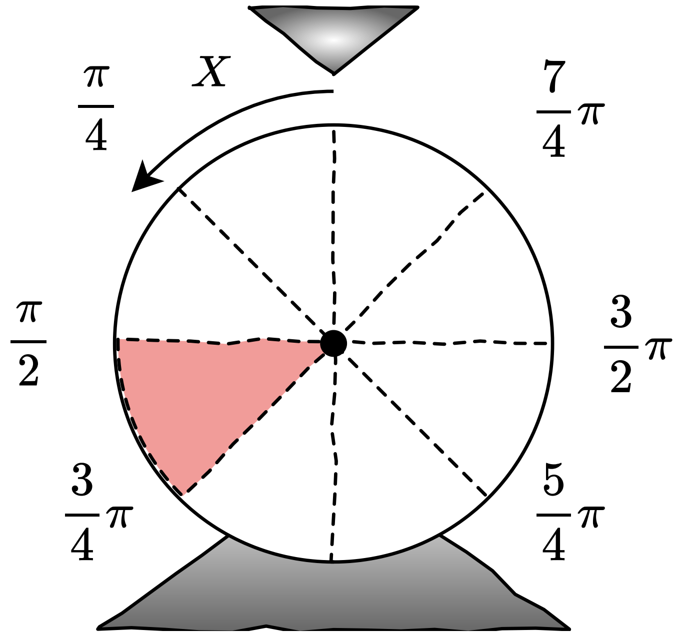
Berechnen von Wahrscheinlichkeiten kontinuierlicher Zufallsvariablen

Beispiel Glücksrad: Unterteilung des Rades in verschiedene Segmente

Berechnung der Wahrscheinlichkeit einen bestimmten Bereich zu treffen

Beispiel Glücksrad: Treffen des Bereiches zwischen $\pi/2$ und $3/4\pi$

$$P\left(\frac{\pi}{2} < X < \frac{3\pi}{4}\right) = \int_{\frac{\pi}{2}}^{\frac{3\pi}{4}} f_X(x)\mathrm{d}x = F_X\left(\frac{3\pi}{4}\right) - F_X\left(\frac{\pi}{2}\right)$$



Erwartungswert

Erwartungswert ist der im Mittel zu erwartende Wert der Zufallsvariable X

$$\mathrm{E}\{X\} = \mu_X = \int\limits_{-\infty}^{+\infty} x \cdot f_X(x)\mathrm{d}x$$

Vereinfachte Berechnung des Erwarungswertes bei diskreten Zufallsvariable (wie z.B. Würfelwurf)

$$\mathrm{E}\{X\} = \mu_X = \sum_i x_i \cdot P(X = x_i)$$

Würfel-Beispiel:

$$\mathrm{E}\{X\} = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3.5$$

Momente von Zufallsvariablen

Für eine Zufallsvariable X lässt sich allgemein das n-te Moment berechnen (n ganzzahlig und positiv)

$$m_X^{(n)} = \mathrm{E}\{X^n\} = \int\limits_{-\infty}^{+\infty} x^n \cdot f_X(x)\,\mathrm{d}x$$

Das 1. Moment entspricht dem Mittelwert

$$m_X^{(1)} = \mathrm{E}\{X\} = \int\limits_{-\infty}^{+\infty} x \cdot f_X(x)\,\mathrm{d}x$$

Weiterhin lässt sich das n-te *zentrale* Moment berechnen

$$\mu_X^{(n)} = \mathrm{E}\Big\{\Big(X - m_X^{(1)}\Big)^n\Big\} = \int\limits_{-\infty}^{+\infty} \Big(x - m_X^{(1)}\Big)^n \cdot f_X(x)\,\mathrm{d}x$$

Varianz und Standardabweichung

Varianz als Maß für die Streuung der Wahrscheinlichkeitsdichte (zweites zentrales Moment)

$$\mathrm{E}\big\{(X-\mu_X)^2\big\} = \sigma_X^2 = \int\limits_{-\infty}^{+\infty} (x-\mu_X)^2 \cdot f_X(x)\,\mathrm{d}x = \mu_X^{(2)}$$

Für eine diskrete Zufallsvariable (wie im Würfel-Beispiel) kann Varianz auch wie folgt berechnet werden:

$$\mathrm{E}\big\{(X-\mu_X)^2\big\} = \sigma_X^2 = \sum_i (x_i-\mu_X)^2 \cdot P(X=x_i)$$

Würfel-Beispiel:

$$\sigma_X^2 = \frac{(1-3.5)^2}{6} + \frac{(2-3.5)^2}{6} + \frac{(3-3.5)^2}{6} + \frac{(4-3.5)^2}{6} + \frac{(5-3.5)^2}{6} + \frac{(6-3.5)^2}{6} = \frac{17.5}{6} \approx 2.9$$

Standardabweichung (häufige Anwendung in der Statistik)

$$\sigma_X = \sqrt{\mathrm{E}\{(X-\mu_X)^2\}}$$

Erwartungswert und Varianz einer kontinuierlichen Zufallsvariable

Erwartungswert: $\mu_X = \mathrm{E}\{X\} = \int\limits_{-\infty}^{\infty} x \cdot f_X(x)\mathrm{d}x$

Varianz: $\sigma_X^2 = \mathrm{E}\big\{(X-\mu_X)^2\big\} = \int\limits_{-\infty}^{\infty} (x-\mu_X)^2 \cdot f_X(x)\mathrm{d}x$

Beispiel Glücksrad:

$$\mu_X = \int\limits_0^{2\pi} x \cdot \frac{1}{2\pi}\mathrm{d}x = \frac{1}{2\pi}\left[\frac{x^2}{2}\right]_0^{2\pi} = \frac{1}{2\pi}\left(\frac{(2\pi)^2}{2} - 0\right) = \pi$$

$$\sigma_X^2 = \int\limits_0^{2\pi} (x-\mu_X)^2 \cdot \frac{1}{2\pi}\mathrm{d}x = \frac{1}{2\pi}\left[\frac{(x-\pi)^3}{3}\right]_0^{2\pi} = \frac{1}{2\pi}\left(\frac{\pi^3}{3} - \frac{-\pi^3}{3}\right) = \frac{\pi^2}{3}$$

Beispiel: Gaußverteilung

Wahrscheinlichkeitsdichtefunktion:

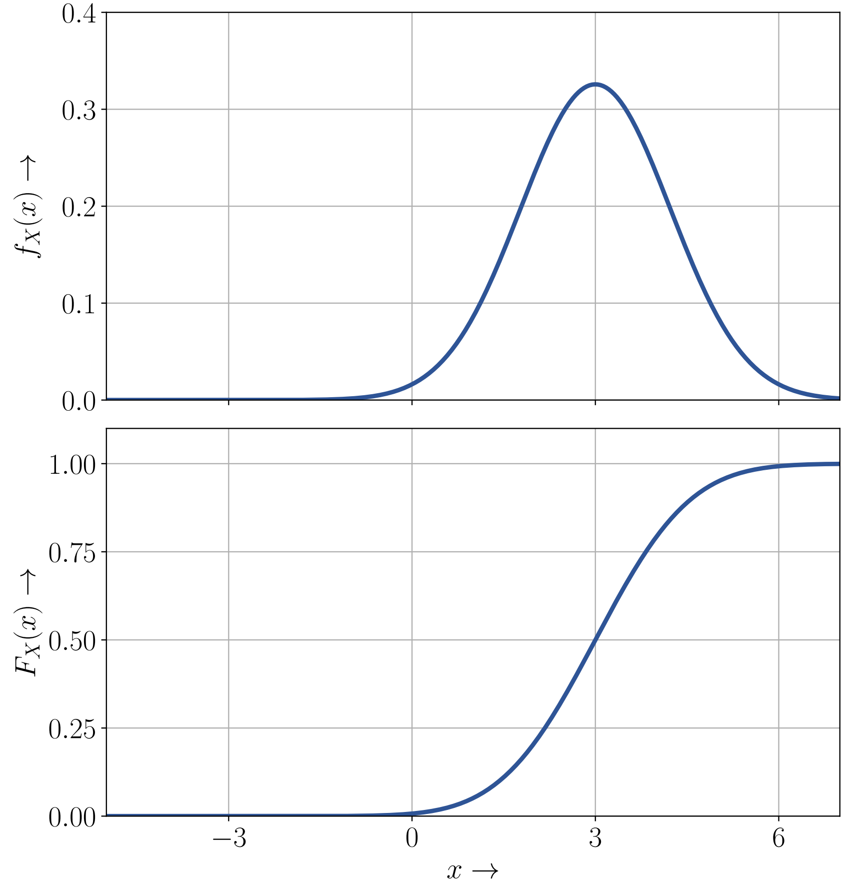
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \cdot \mathrm{e}^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}$$

Verteilungsfunktion:

$$F_X(x) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x-\mu_X}{\sqrt{2\sigma_X^2}}\right)\right)$$

Fehlerfunktion (*Error-Function*)

$$\operatorname{erf}(x) = \frac{2}{\pi}\int\limits_0^x \mathrm{e}^{-\xi^2}\mathrm{d}\xi$$



Verbundwahrscheinlichkeit

Beschreibung der gemeinsamen Wahrscheinlichkeitsdichte von zwei Zufallsvariablen X und Y

$$f_{XY}(x,y)$$

Berechnung der Verbundverteilungsfunktion

$$F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(\xi,\gamma)\mathrm{d}\xi\mathrm{d}\gamma$$

Berechnung der Randdichten allgemein für kontinuierliche Zufallsvariablen bei gegebener Dichte $f_{XY}(x,y)$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\,\mathrm{d}y$$

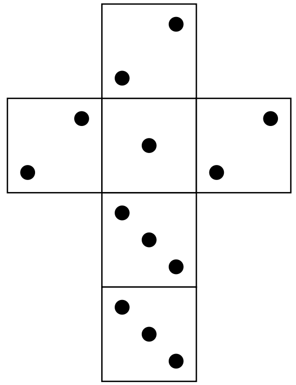
$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\,\mathrm{d}x$$

Verbundwahrscheinlichkeiten von zwei Würfeln I

Werfen zweier Würfel: Zufallsvariablen X und Y

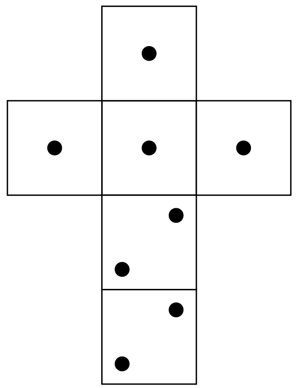
1. Würfel hat auf einer Seite 1 Auge, auf drei Seiten zwei Augen und auf zwei Seiten drei Augen

- $P(X = 1) = \frac{1}{6}$
- $P(X = 2) = \frac{1}{2}$
- $P(X = 3) = \frac{1}{3}$



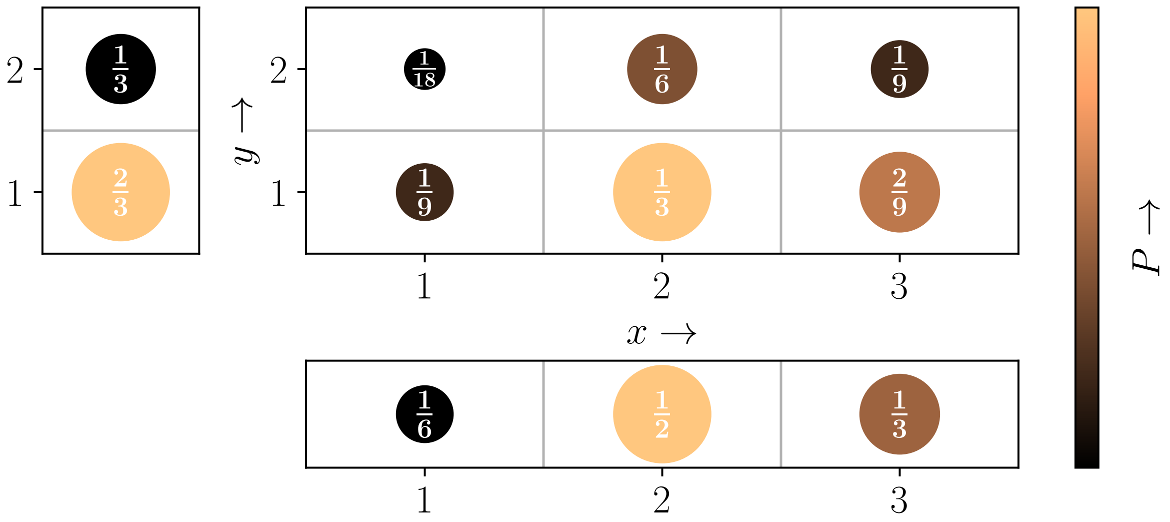
2. Würfel hat auf zwei Seiten 1 Auge und auf 4 Seiten zwei Augen

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



Verbundwahrscheinlichkeiten von zwei Würfeln II

Berechnen der Verbundwahrscheinlichkeiten $P(X = x \cap Y = y)$



Hier: X und Y statistisch unabhängig:

$$P(X = x \cap Y = y) = P(X = x) \cdot P(Y = y)$$

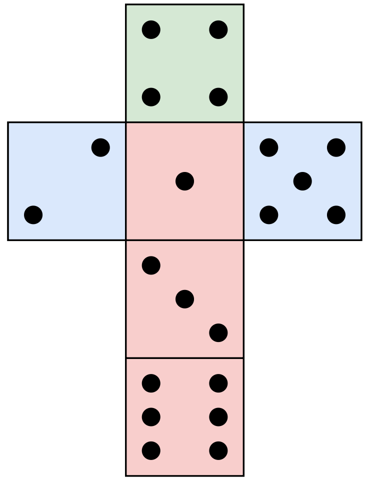
Allgemein für kontinuierliche Zufallsvariablen:

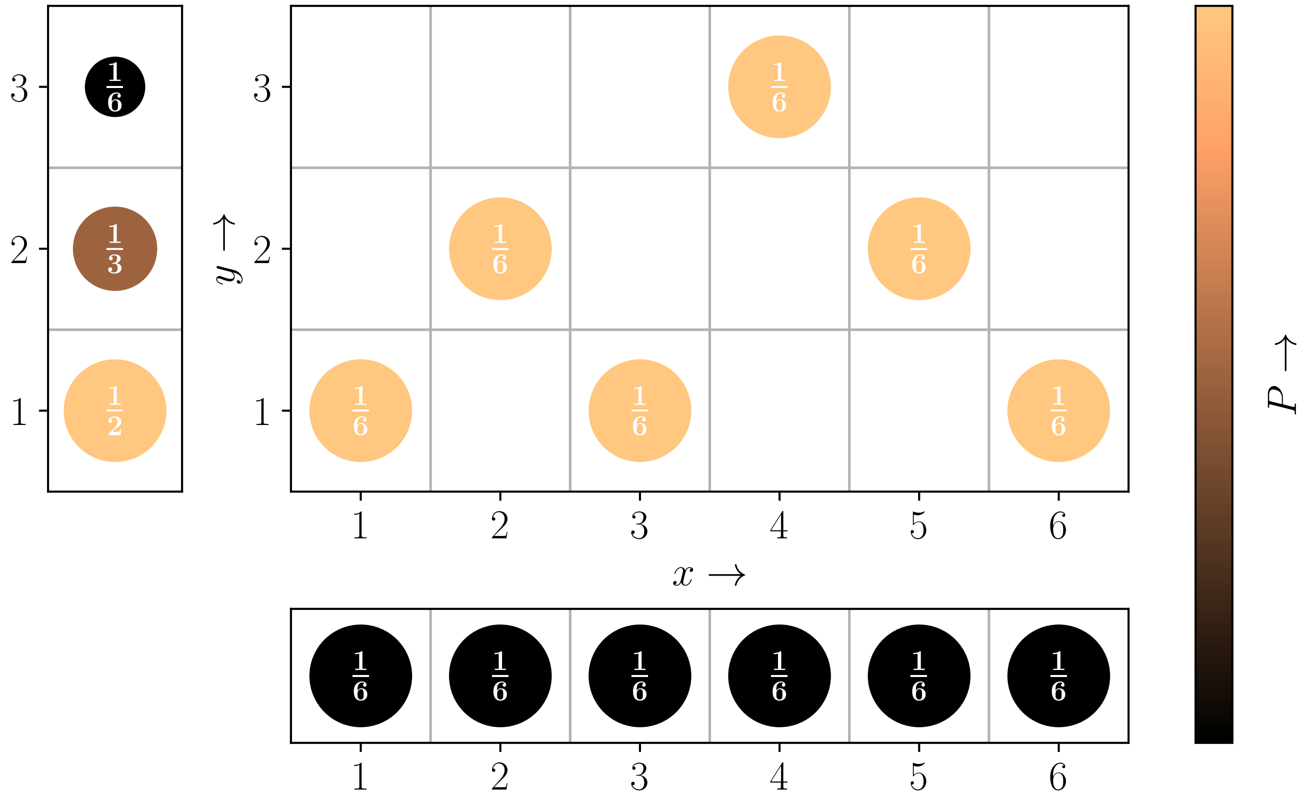
$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$$

Verbundwahrscheinlichkeiten von einem Würfel mit zwei Attributen

Zufallsvariable X: Anzahl der Augen des Würfels

Zufallsvariable Y: rot $\Rightarrow Y = 1$ blau $\Rightarrow Y = 2$ grün $\Rightarrow Y = 3$





Rechnen mit Erwartungswerten I

Skalierung einer Zufallsvariable mit konstantem Faktor a

$$\mathrm{E}\{a \cdot X\} = \int\limits_{-\infty}^{+\infty} a \cdot x \cdot f_X(x)\,\mathrm{d}x = a \cdot \int\limits_{-\infty}^{+\infty} x \cdot f_X(x)\,\mathrm{d}x = a \cdot \mathrm{E}\{X\}$$

Addition zweier beliebiger Zufallsvariablen X und Y

$$\mathrm{E}\{X + Y\} = \iint\limits_{-\infty}^{+\infty} (x + y) \cdot f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y = \iint\limits_{-\infty}^{+\infty} x \cdot f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y + \iint\limits_{-\infty}^{+\infty} y \cdot f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y =$$

$$= \int\limits_{-\infty}^{+\infty} x \int\limits_{-\infty}^{+\infty} f_{XY}(x, y)\,\mathrm{d}y\mathrm{d}x + \int\limits_{-\infty}^{+\infty} y \int\limits_{-\infty}^{+\infty} f_{XY}(x, y)\,\mathrm{d}x\mathrm{d}y =$$

$$= \int\limits_{-\infty}^{+\infty} x \cdot f_X(x)\,\mathrm{d}x + \int\limits_{-\infty}^{+\infty} y \cdot f_Y(y)\,\mathrm{d}y = \mathrm{E}\{X\} + \mathrm{E}\{Y\}$$

Rechnen mit Erwartungswerten II

Wichtige Eigenschaft des Erwartungswertoperators:

Damit ist der Erwartungswertoperator ein *linearer* Operator

$$\boxed{\begin{gathered}\mathrm{E}\{a \cdot X\} = a \cdot \mathrm{E}\{X\} \\[2em] \mathrm{E}\{X + Y\} = \mathrm{E}\{X\} + \mathrm{E}\{Y\}\end{gathered}}$$

Kovarianz und Korrelation

Ermittlung des linearen Zusammenhanges zweier Zufallsvariablen X und Y mittels *Kovarianz*

$$C_{XY} = \mathrm{E}\{(X - \mu_X) \cdot (Y - \mu_Y)\} = \mathrm{E}\{XY - Y \cdot \mu_X - X \cdot \mu_Y + \mu_X \mu_Y\} = \mathrm{E}\{XY\} - \mu_X \mu_Y$$

Kovarianz liegt im Bereich

$$-\sigma_X \sigma_Y \leq C_{XY} \leq \sigma_X \sigma_Y$$

Falls $C_{XY} = 0$ gilt werden X und Y als *unkorreliert* bezeichnet

Korrelationskoeffizient: Normierung der Kovarianz auf Standardabweichung

$$c_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y} \qquad \text{mit} \qquad -1 \leq c_{XY} \leq +1$$

Kovarianz und Korrelationskoeffizient geben ein Maß für den linearen Zusammenhang der Zufallsvariablen an. Je größer der Wert desto größer der Zusammenhang.

Statistisch unabhängige und orthogonale Zufallsvariablen

Falls X und Y statistisch unabhängig:

$$\mathrm{E}\{X \cdot Y\} = \mathrm{E}\{X\} \cdot \mathrm{E}\{Y\}$$

Statistisch unabhängige Zufallsvariablen sind unkorreliert, d.h. $C_{XY} = 0$

Vorsicht: Unkorrelierte Zufallsvariablen sind nicht automatisch statistisch unabhängig

Zwei Zufallsvariablen X und Y werden als *orthogonal* bezeichnet falls

- X und Y statistisch unabhängig sind
- zumindest eine der beiden Zufallsvariablen mittelwertfrei ist, d.h. $\mu_X = 0$ oder $\mu_Y = 0$

Für orthogonale Zufallsvariablen X und Y gilt

$$\mathrm{E}\{X \cdot Y\} = 0$$

Zufallsexperimente

Zufallsexperiment

- Reales (oder hypothetisches) Experiment mit unvorhersebarem Ergebnis
- Alle möglichen Ergebnisse sind vorab bekannt
- Ergenis jedes individuellen Experiments ist nicht verhersehbar

N-fache Wiederholung eines Experimentes ergibt N konkrete Zufallszahlen

$$\mathbf{X} = [X_1,\ X_2,\ X_3,\ \ldots,\ X_N] = [X_i] \qquad X_i \in \Omega$$

Beispiel: 10-faches Werfen eines Würfels

$$\mathbf{X} = [1,\ 2,\ 6,\ 1,\ 2,\ 4,\ 2,\ 2,\ 1,\ 4]$$

Schwache Gesetz der großen Zahlen

Eine direkte Messung von Wahrscheinlichkeiten ist nicht möglich.

Eine messbare Größe mit den gleichen Eigenschaften der Wahrscheinlichkeit ist die *relative Häufigkeit*.

Hierfür wird das Zufallsexperiment N-fach wiederholt und die Zahl n_X des Auftretens der Zufallsvariable $X = x$ bestimmt.

Die relative Häufigkeit ergibt sich aus

$$h_X(x) = \frac{n_X}{N}$$

Schwache Gesetz der Großen Zahlen: Bei hinreichend hoher Zahl N an Wiederholungen strebt relative Häufigkeit gegen die Wahrscheinlichkeit

$$\lim_{N \to \infty} h_X(x) = P(X = x)$$

Beispiel: Berechnung der Wahrscheinlichkeit beim Würfelwurf

Beispiel: 10-faches Werfen eines Würfels

$$h_X(2) = \tfrac{4}{10} = 0{,}4 \neq P(X=2) = \tfrac{1}{6} \approx 0{,}1666$$

$$h_X(4) = \tfrac{2}{10} = 0{,}2 \neq P(X=4) = \tfrac{1}{6} \approx 0{,}1666$$

Beispiel: 100-faches Werfen eines Würfels

$$h_X(2) = 0{,}17 \neq P(X=2) = \tfrac{1}{6} \approx 0{,}1666$$

$$h_X(4) = 0{,}19 \neq P(X=4) = \tfrac{1}{6} \approx 0{,}1666$$

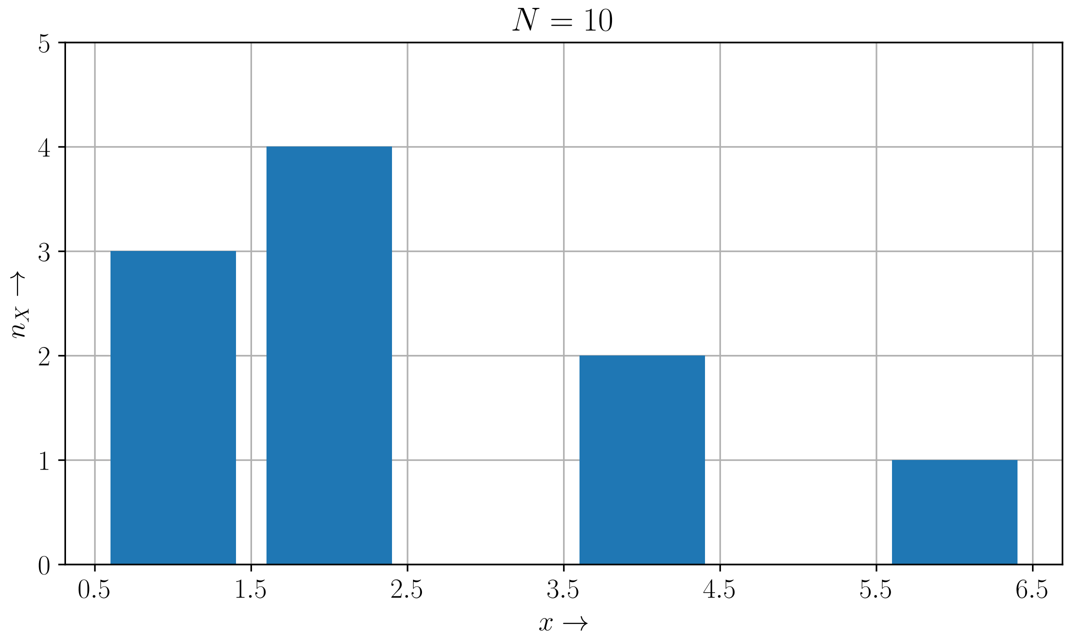
Je größer die Anzahl N an Wiederholungen des Zufallsexperimentes, desto näher liegt die relative Häufigkeit an der Wahrscheinlichkeit.

Histogramm

Gegeben: Ergebnisse $\mathbf{X} = [X_i]$ eines konkreten Zufallsexperimentes

Histogramm:

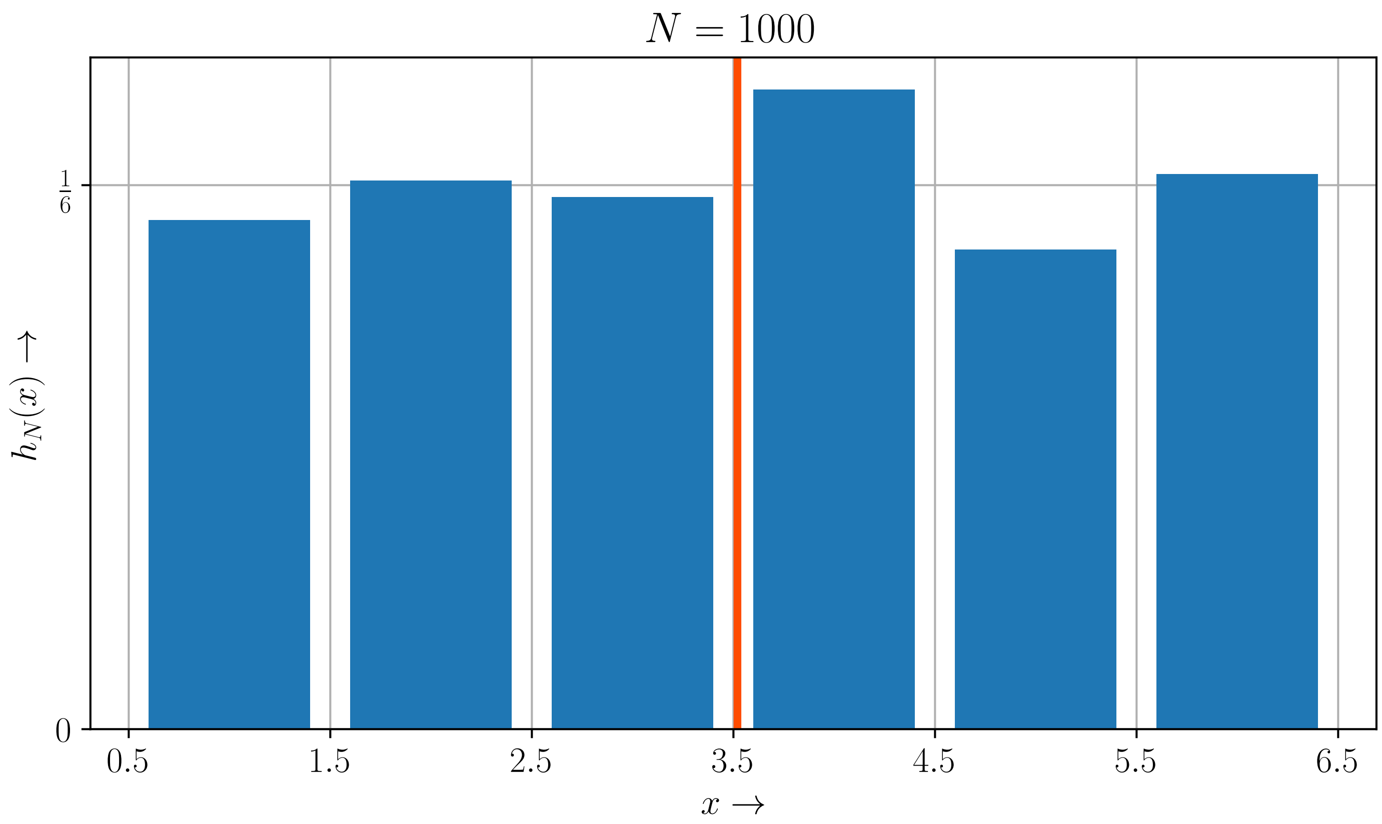
- Definition von Intervallen (*Bins*)
- Bestimmung der Anzahl der Ergebnisse X_i in den jeweiligen Intervallen



Histogramm vs. Wahrscheinlichkeitsdichtefunktion

Normierung des Histogramms auf relative Häufigkeiten

Für $N \to \infty$ nähert sich das Histogramm der Wahrscheinlichkeitsdichtefunktion



Berechnung von Mittelwert und Varianz von konkreten Zufallsexperimenten

Falls statistische Eigenschaften des Zufallsprozesses bekannt berechnen sich Mittelwert und Varianz durch

$$\mu_X = \int_{-\infty}^{+\infty} x \cdot f_X(x)\,\mathrm{d}x \qquad \sigma_X^2 = \int_{-\infty}^{+\infty} (x-\mu_X)^2 \cdot f_X(x)\,\mathrm{d}x$$

Bei konkreten Zufallsexperimenten lassen sich nur Schätzwerte ermitteln

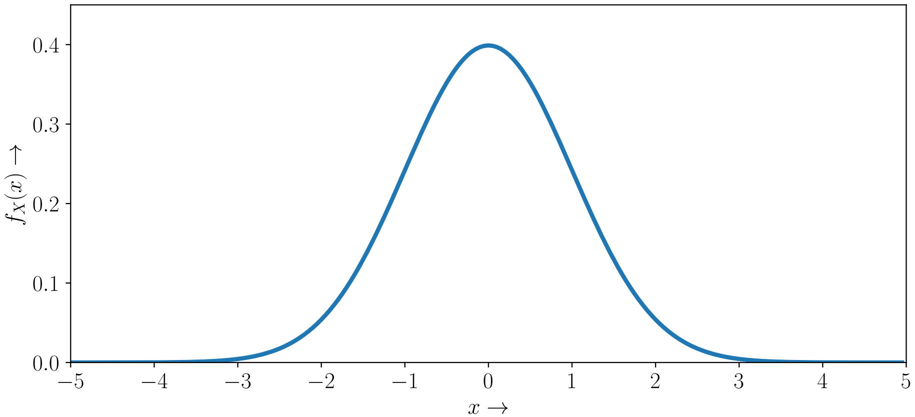
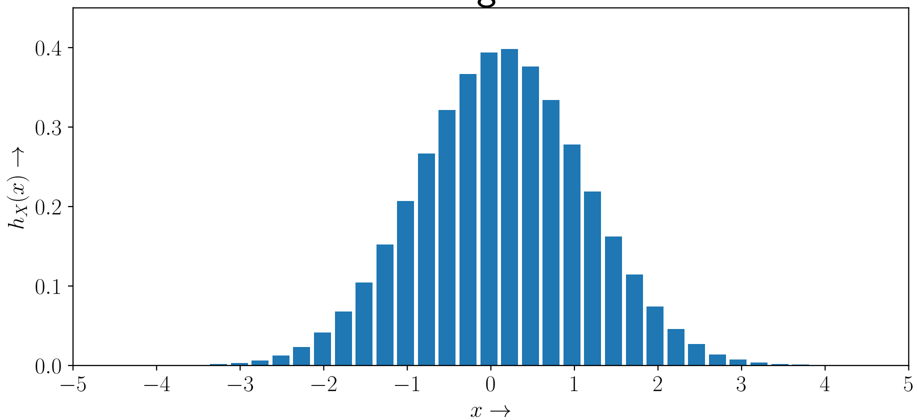
$$\hat{\mu}_X = \frac{1}{N}\sum_{i=1}^{N} X_i$$

$$\hat{\sigma}_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i-\hat{\mu}_X)^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \frac{1}{N}\sum_{k=1}^{N} X_k\right)^2$$

Falls Mittelwert μ_X bekannt (z.B. falls X sicher mittelwertfrei) ist ein besserer Schätzwert für die Varianz

$$\hat{\sigma}_X^{*2} = \frac{1}{N}\sum_{i=1}^{N}(X_i-\mu_X)^2$$

Vergleich zwischen Wahrscheinlichkeiten und Zufallsexperimenten

Rechnen mit Zufallsvariablen	Rechnen mit relativen Häufigkeiten
$P(X = x)$ bzw. $\int_{x_1}^{x_2} f_X(x)\,\mathrm{d}x = F(x_2) - F(x_1)$	$h_X(x) = \frac{n_X}{N}$
Wahrscheinlichkeitsdichtefunktion 	Histogramm 
$\mu_X = \int_{-\infty}^{+\infty} x \cdot f_X(x)\,\mathrm{d}x$	$\hat{\mu}_X = \frac{1}{N} \sum_{i=1}^{N} X_i$
$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 \cdot f_X(x)\,\mathrm{d}x$	$\hat{\sigma}_X^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \hat{\mu}_X)^2$

Beispiel: Berechnung von Mittelwert und Varianz beim Würfelwurf

Beispiel: 10-faches Werfen eines Würfels

$$\begin{aligned}
\hat{\mu}_X &= \tfrac{1+2+6+1+2+4+2+2+1+4}{10} &= 2.5 \quad &\neq \quad \mu_X &= 3.5 \\
\hat{\sigma}_X^2 &= \tfrac{(1-2.5)^2+(2-2.5)^2+(6-2.5)^2+\dots}{10-1} &= 2.72 \quad &\neq \quad \sigma_X^2 &= \tfrac{17.5}{6} \approx 2.92
\end{aligned}$$

Beispiel: 100-faches Werfen eines Würfels

$$\begin{aligned}
\hat{\mu}_X &= \tfrac{3+1+4+\dots}{100} &= 3.47 \quad &\rightarrow \quad \mu_X &= 3.5 \\
\hat{\sigma}_X^2 &= \tfrac{(3-3.47)^2+(1-3.47)^2+(4-3.47)^2\dots}{100-1} &= 2.96 \quad &\rightarrow \quad \sigma_X^2 &\approx 2.92
\end{aligned}$$

Je größer die Anzahl N an Wiederholungen des Zufallsexperimentes, desto besser sind die Schätzwerte für Mittelwert und Varianz.

Simulation von Zufallsexperimenten in Matlab

Gauß-verteilte Zufallsvariablen

```
>> help randn
```

```
randn Normally distributed pseudorandom numbers.
```

```
R = randn(N) returns an N-by-N matrix containing pseudorandom values drawn
from the standard normal distribution.
```

`randn()` liefert Gaussverteilte Zufallsvariablen mit Mittelwert $\mu_X = 0$ und Varianz $\sigma_X^2 = 1$

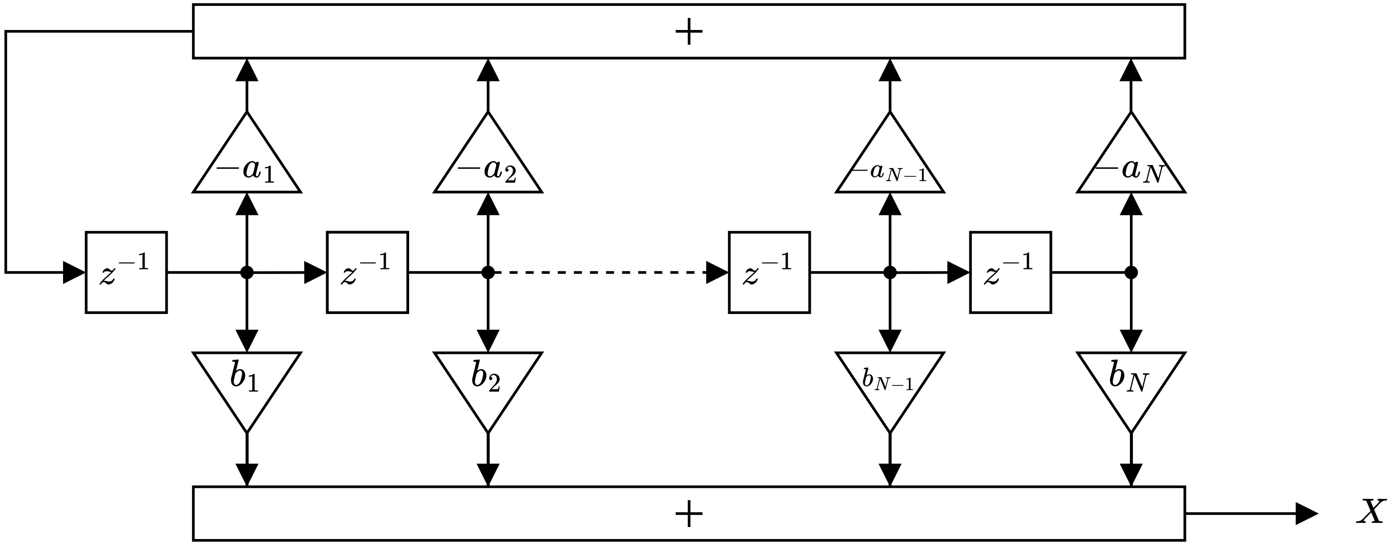
Transformation der Zufallsvariable auf beliebige Werte für Mittelwert und Varianz, z.B. $\mu_X = 1.5$ und $\sigma_X^2 = 2$:

```
>> sqrt(2) * randn(1, 1000) + 1.5
```


Pseudozufallszahlen

Berechnung von Zufallszahlen im Rechner (z.B. mit Matlab) in der Regel nicht zufällig sondern mittels Algorithmus

Eine Möglichkeit zur Berechnung von Pseudozufallszahlen: *Rückgekoppeltes Schieberegister*



Nach zentralem Grenzwertsatz ist Zufallsvariable X näherungsweise Gauß-verteilt

Zustände $z_1,\ z_2,\ \ldots,\ z_N$ definieren Sequenz an Zufallsvariablen

Verwendung von Pseudozufallszahlen in Matlab

- Befehl `rand()` liefert immer neue Zufallszahlen
- In manchen Situationen ist es notwendig immer die gleiche Sequenz an Zufallsvariablen zu verwenden

Beeinflussung der Sequenz an Zufallsvariablen mittels sogenannten *Seed*

Festlegung des Seeds in Matlab über `rng()`

Beispiel:

```
>> rng(32168)

>> randn(1,10)

ans =

   -0.9460 -0.6412 -0.6984 0.1844 1.7311 1.2937 1.7639 0.9946 -0.6087
-2.1481
```

Weitere Zufallsgeneratoren in Matlab

Gleichverteilung im Intervall $[0, 1]$

```
>> help rand
```

```
rand Uniformly distributed pseudorandom numbers.
```

Zufällige ganze Zahlen (Radom Integers)

```
>> help randi
```

```
randi Pseudorandom integers from a uniform discrete distribution.
```

Zufällige Permutationen einer Sequenz (*Mischen*)

```
>> help randperm
```

```
randperm Random permutation.
```


Erstellung von Histogrammen in Matlab

`histogram(X)` erstellt ein Figure mit dem Histogramm einer Zufallsvariablen `X`

`[N,EDGES] = histcounts(X)` gibt die Häufigkeit `N` an, mit der die Zufallsvariablen `X` innerhalb verschiedener Bereiche (Ränder `EDGES`) auftreten

Weitere Konfiguration möglich durch

- Definition der Bereiche (`BINS`)
- Normalisierung (z.B. Anzahl, Wahrscheinlichkeitsdichte, Wahrscheinlichkeitsverteilung, etc.)
- etc.

Gegeben: Vektor an Einzelwahrscheinlichkeiten `P`

Berechnung der Wahrscheinlichkeitsverteilung mittels

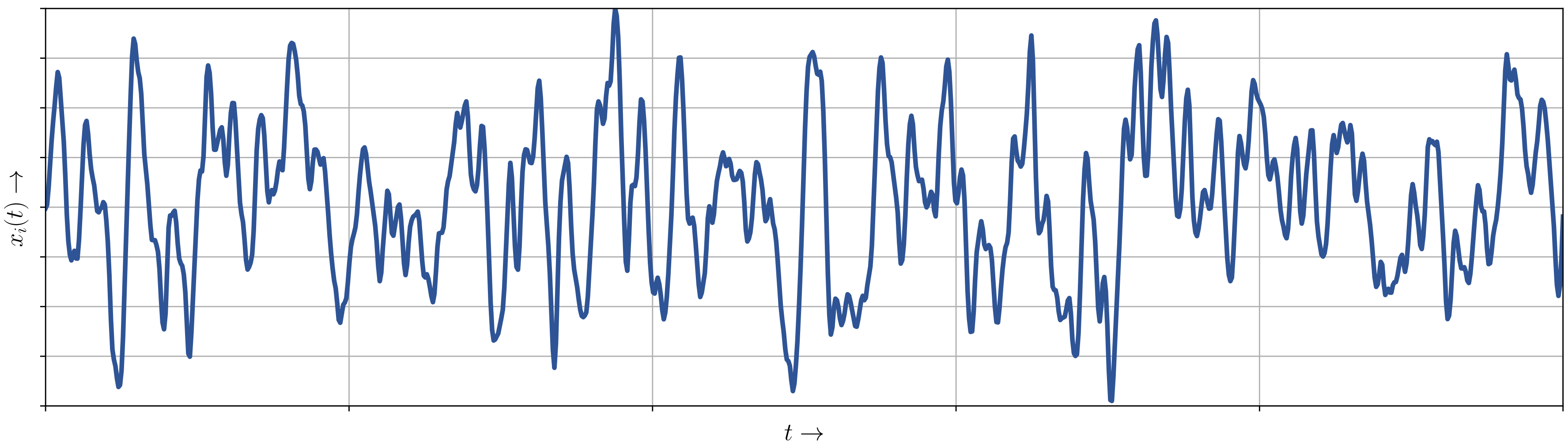
```
>> cumsum(P)
```


Zufallsprozesse

Zufallsprozess

Zufallsvariable: Jedem Ergebnis eines Zufallsexperimentes wird eine Zahl x_i zugeordnet.

Zufallsprozess: Jedem Ergebnis eines Zufallsexperimentes wird eine Zeitfunktion $x_i(t)$ zugeordnet.

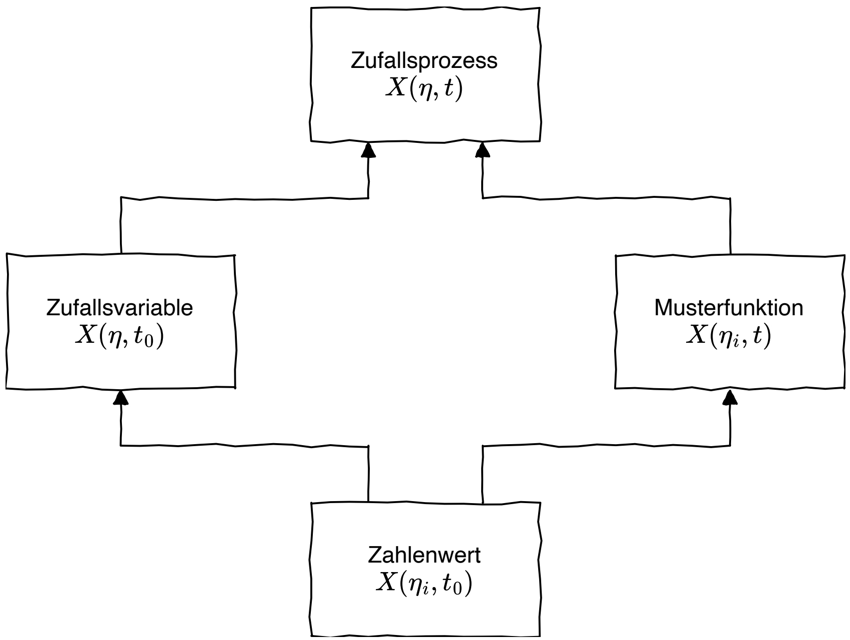


Beschreibung des Zufallsprozesses mittels der Funktion $X(\eta, t)$ mit Zufallsvariable η

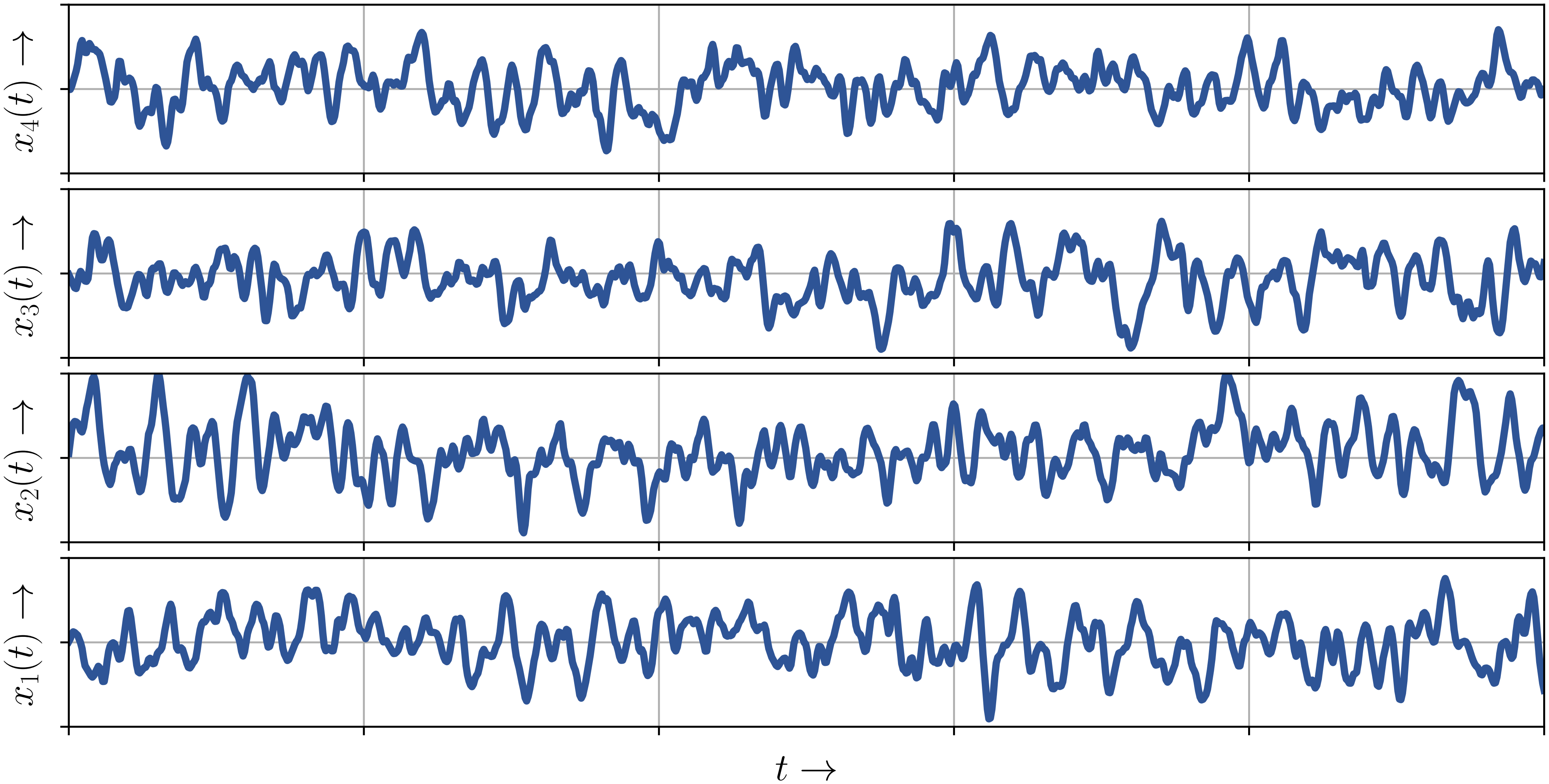
Für jedes konkrete Ergebnis η_i: Zuweisung einer konkreten Zeitfunktion $X(\eta_i, t) = X_i(t)$.

Zufallsprozess, Zufallsvariable, Musterfunktion und Variable

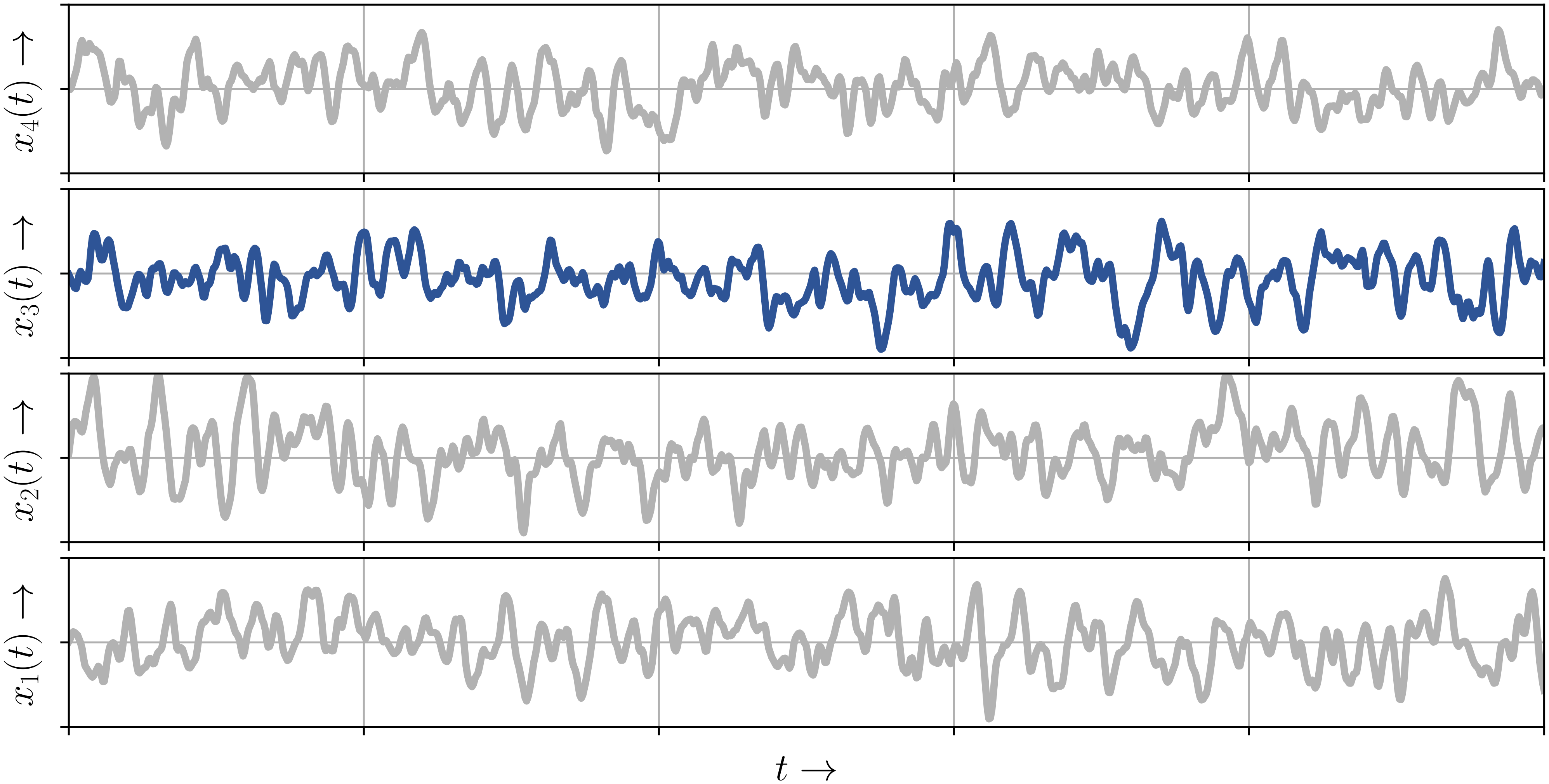
1. η und t variabel: $X(\eta, t)$ beschreibt *Zufallsprozess* bzw. Schar von Musterfunktionen
2. η variable, t fest: $X(\eta, t_0)$ ist *Zufallsvariable*
3. η fest, t variabel: $X(\eta_i, t) = x_i(t)$ ist *Musterfunktion* (deterministische Funktion)
4. η und t fest: $X(\eta_i, t_0) = x_i(t_0)$ ist gewöhnlicher Zahlenwert



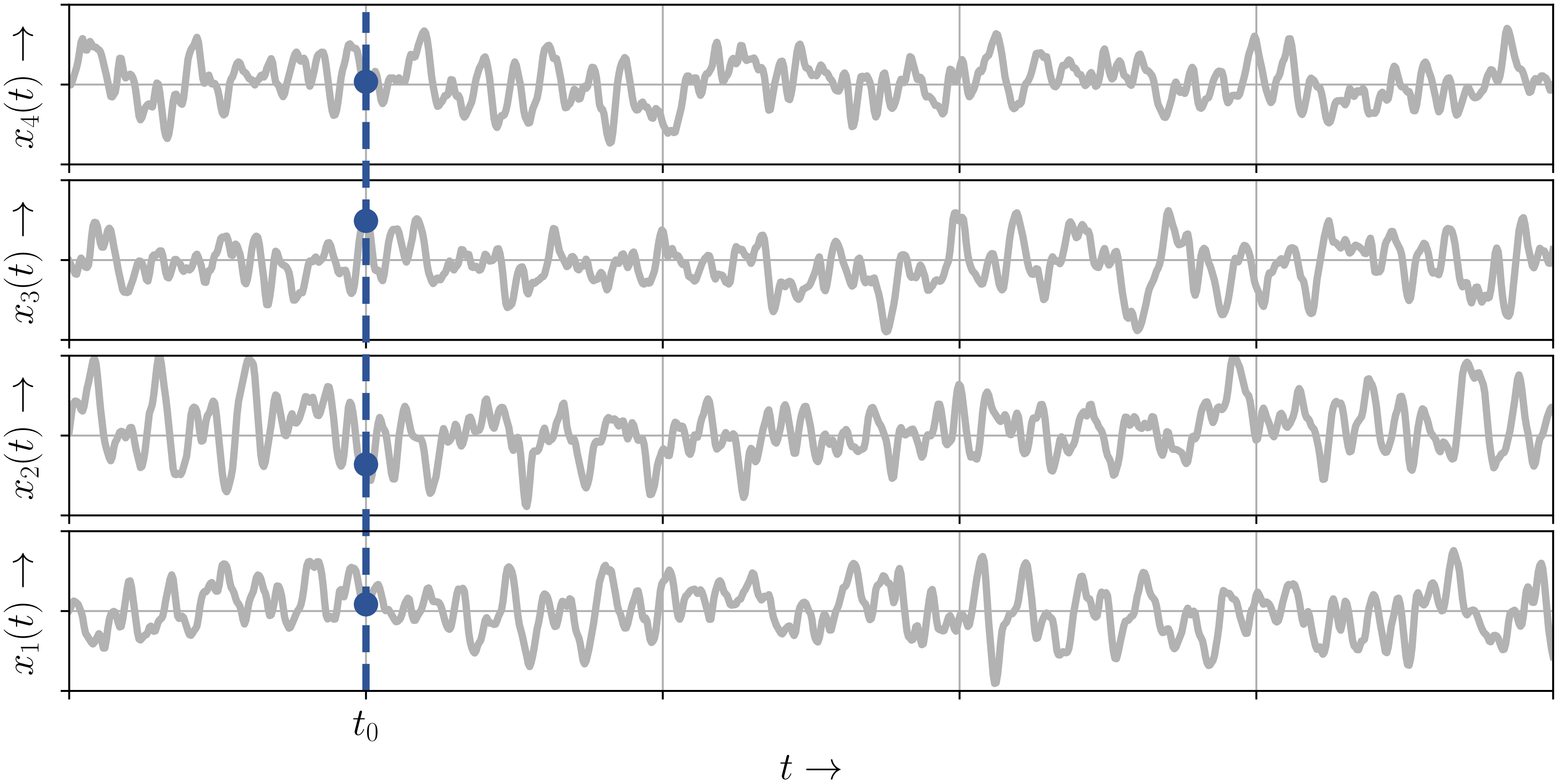
4 Realisierungen eines Zufallsprozesses



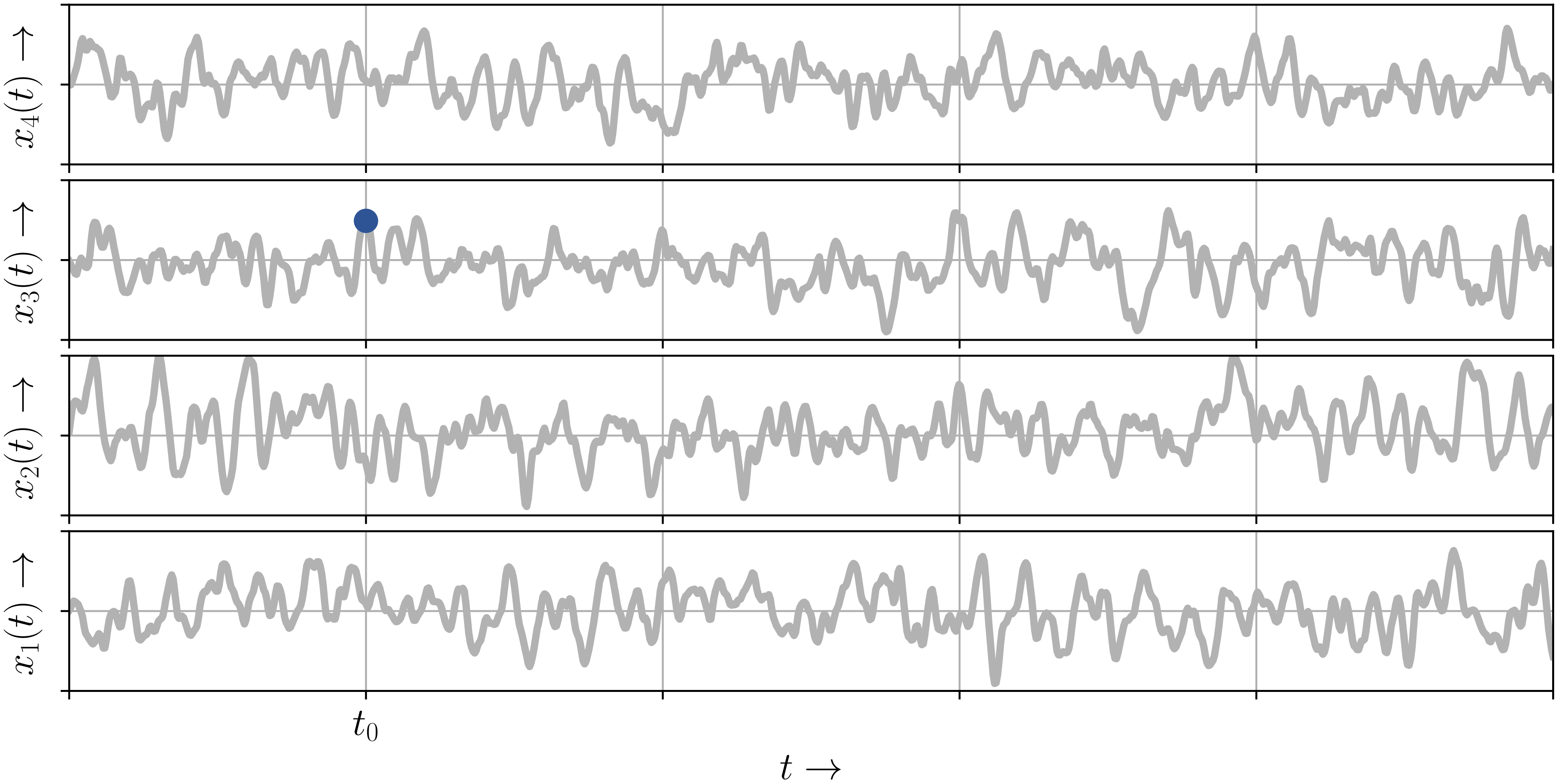
Musterfunktion



Zufallsvariable



Zahlenwert



$x_4(t) \rightarrow$
$x_3(t) \rightarrow$
$x_2(t) \rightarrow$
$x_1(t) \rightarrow$
t_0
$t \rightarrow$

Diskussion

- Die Betrachtung von Zufallsprozessen berücksichtigt zwei Dimensionen:
    1. Musterfunktionen
    2. Zeit
- Zeitliche Betrachtung ermöglicht zusätzlich die Analyse im Frequenzbereich mittels Fourier-Transformation
- Im Folgenden werden Anwendungen betrachtet bei denen die zeitliche Dimension keine Rolle spielt
- Diese Anwendungen lassen sich unverändert auf Zufallsprozesse anwenden

Summe von Zufallsvariablen

Summe von Zufallsvariablen

Gegeben: Zwei Zufallsvariablen X und Y

Abbildung auf $Z = X + Y$

Falls X und Y statistisch unabhängig, d.h. $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(x, z - y)\,\mathrm{d}x = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z - y)\,\mathrm{d}x$$

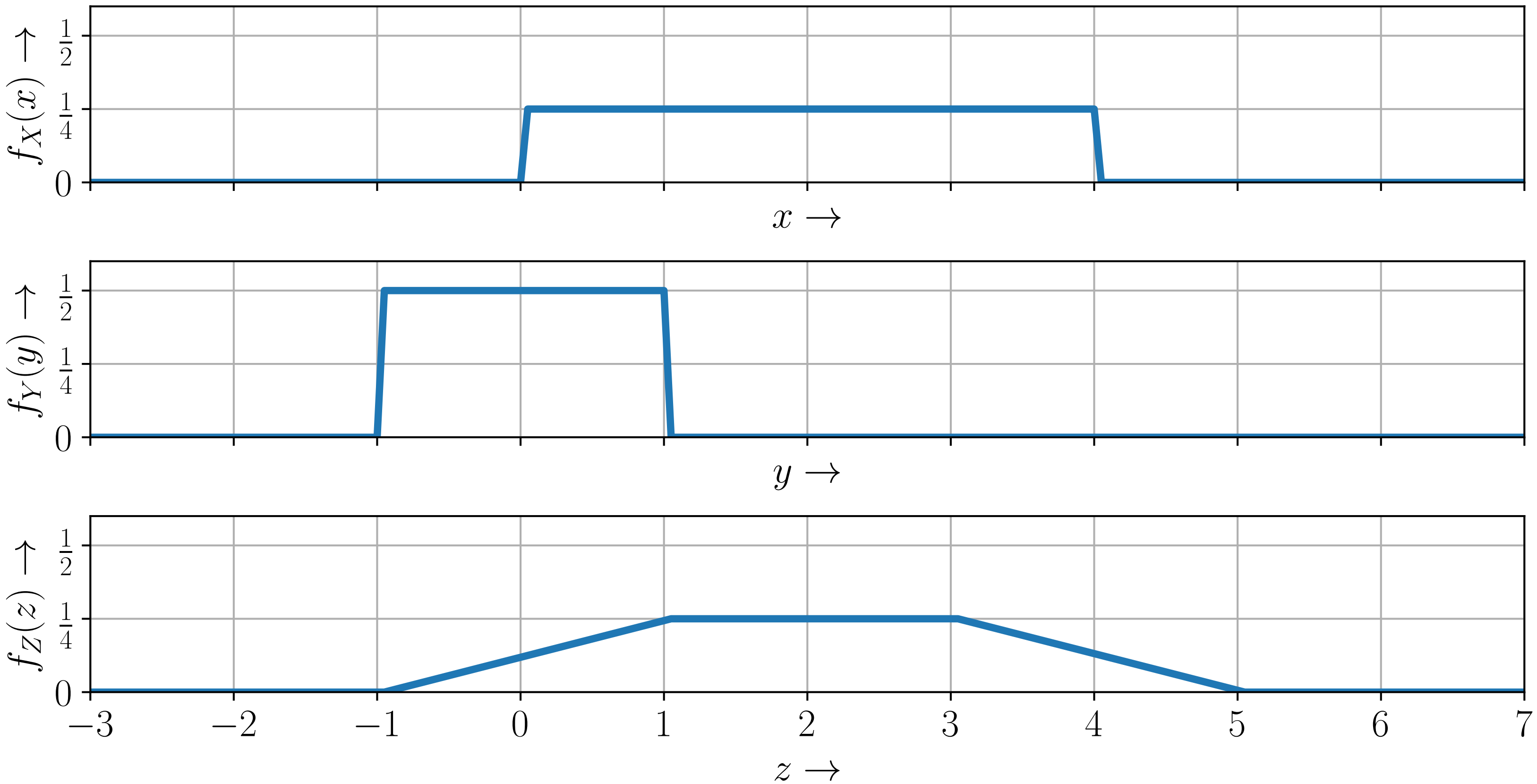
Wahrscheinlichkeitsdichtefunktion von Z entspricht der Faltung der jeweiligen Funktionen von X und Y:

$$f_Z(z) = f_X(z) * f_Y(z)$$

Beispiel: Addition zweier gleichverteitler Zufallsvariablen

Gegeben: zwei Zufallsvariablen X und Y mit $f_X(x) = \frac{\mathrm{rect}((x-2)/4)}{4}$ und $f_Y(y) = \frac{\mathrm{rect}(y/2)}{2}$

Berechnung der Summe Z mit $f_Z(z) = f_X(x) * f_Y(y)$



Zentraler Grenzwertsatz

N-fache Überlagerung von Zufallsvariablen

$$Y = X_1 + X_2 + X_3 + \ldots + X_N$$

Für $N \to \infty$ nähert sich Y einer normalverteilten Zufallsvariable.

Sonderfall: Normierte Summe von Zufallsvariablen mit $\mathrm{E}\{X_i\} = \mu_X$ und $\mathrm{E}\big\{(X_i - \mu_X)^2\big\} = \sigma_X^2$ für alle i

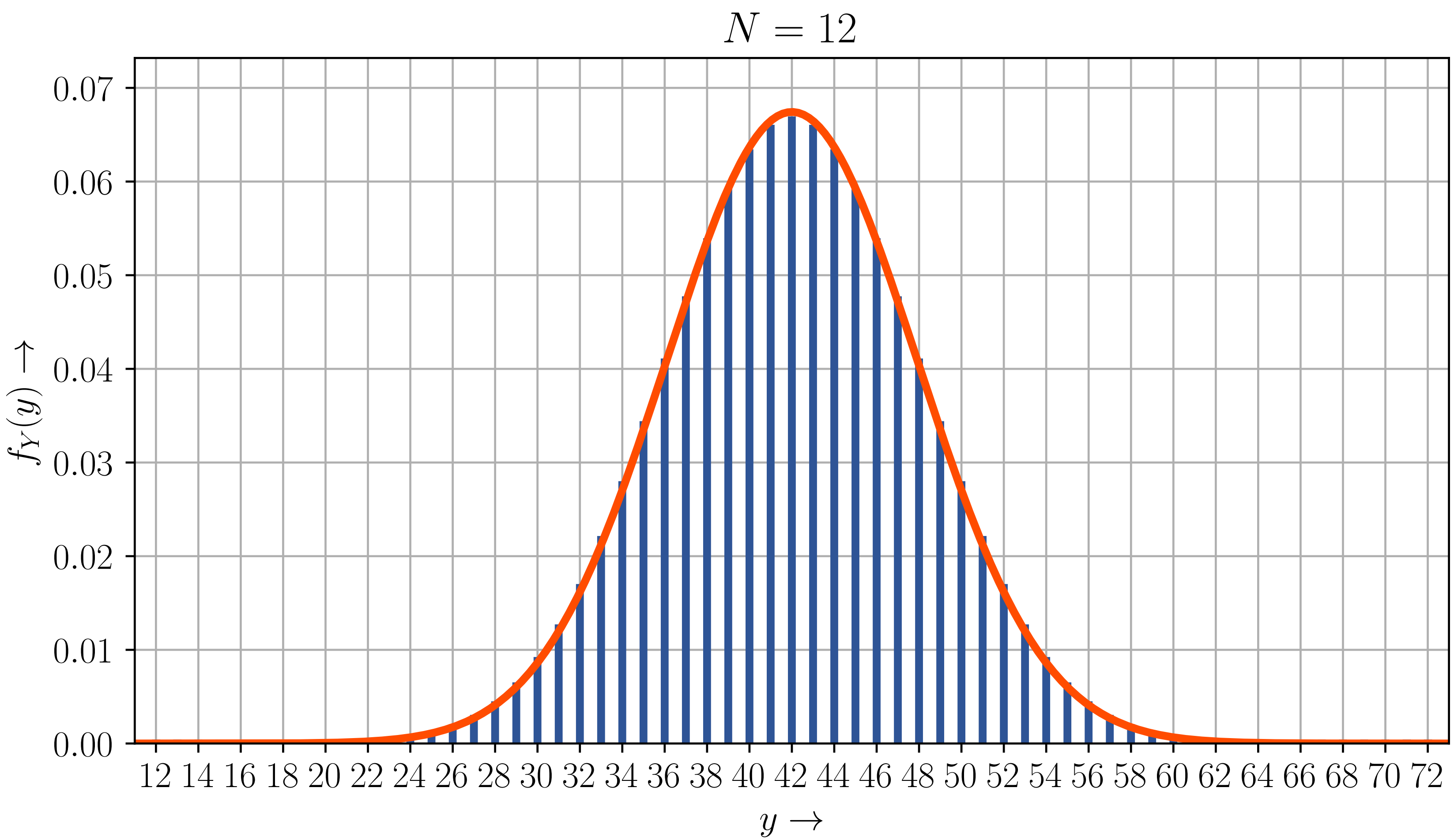
$$Y = \frac{X_1 + X_2 + X_3 + \ldots + X_N}{N}$$

Zufallsvariable Y ist Gauß-verteilt

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} \cdot \mathrm{e}^{-\frac{(x-\mu_Y)^2}{2\sigma_Y^2}}$$

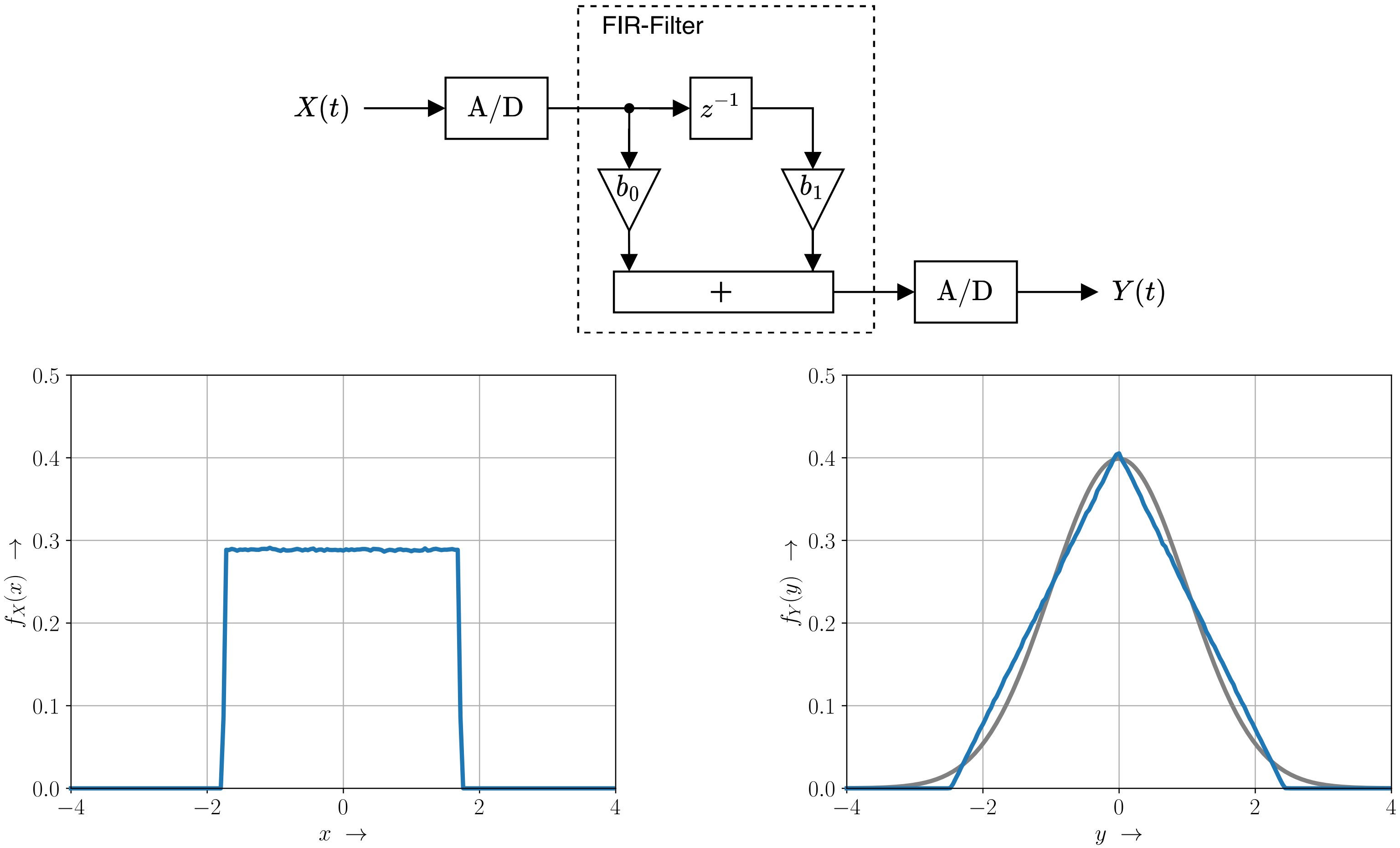
mit $\mu_Y = N \cdot \mu_X$ und $\sigma_Y^2 = N \cdot \sigma_X^2$

Zentraler Grenzwertsatz am Beispiel eines Würfels



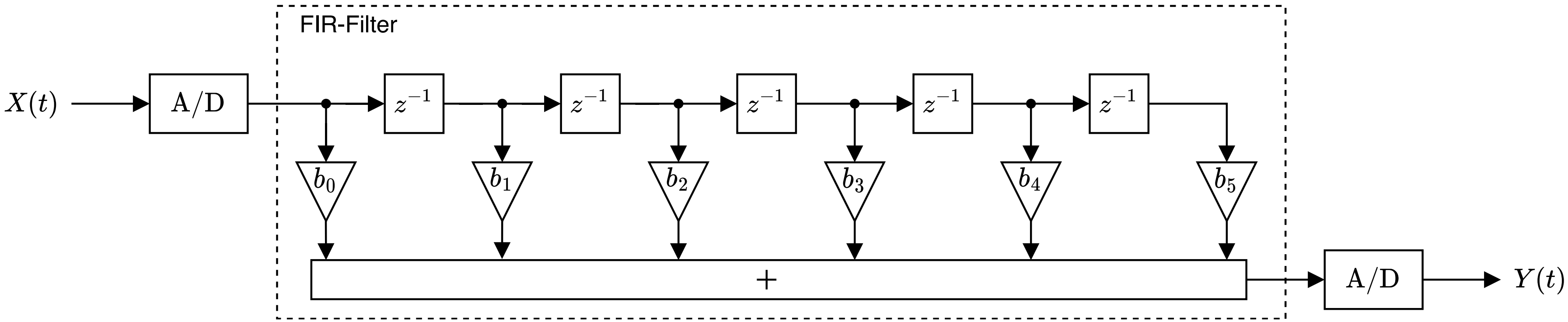
Anwendungsbeispiel: Tiefpassfilter 1. Ordnung

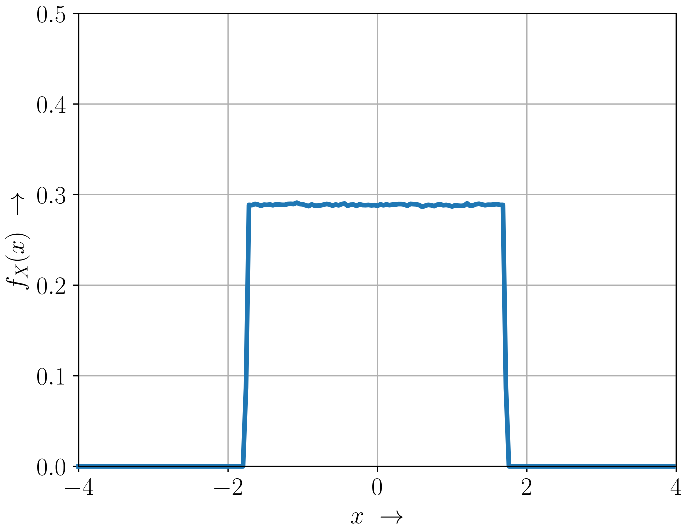
Tiefpassfilterung eines gleichverteilten Eingangsprozesses $X(t)$ mit zeitdiskreten FIR-Filter 1. Ordnung

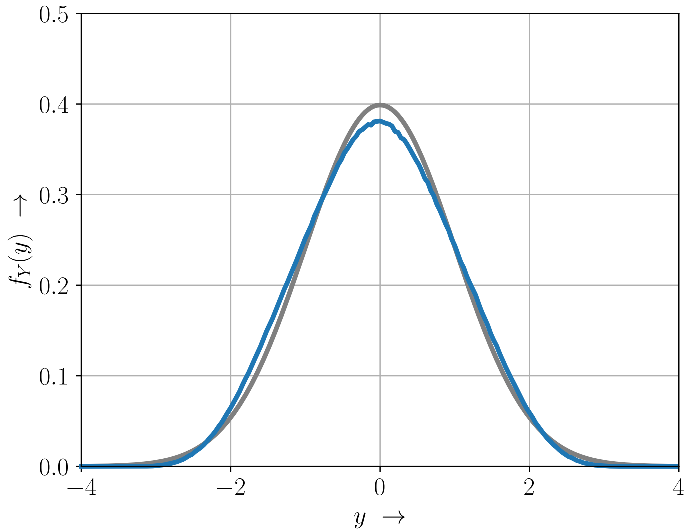


Anwendungsbeispiel: Tiefpassfilter 5. Ordnung

Tiefpassfilterung eines gleichverteilten Eingangsprozesses $X(t)$ mit zeitdiskreten FIR-Filter 5. Ordnung

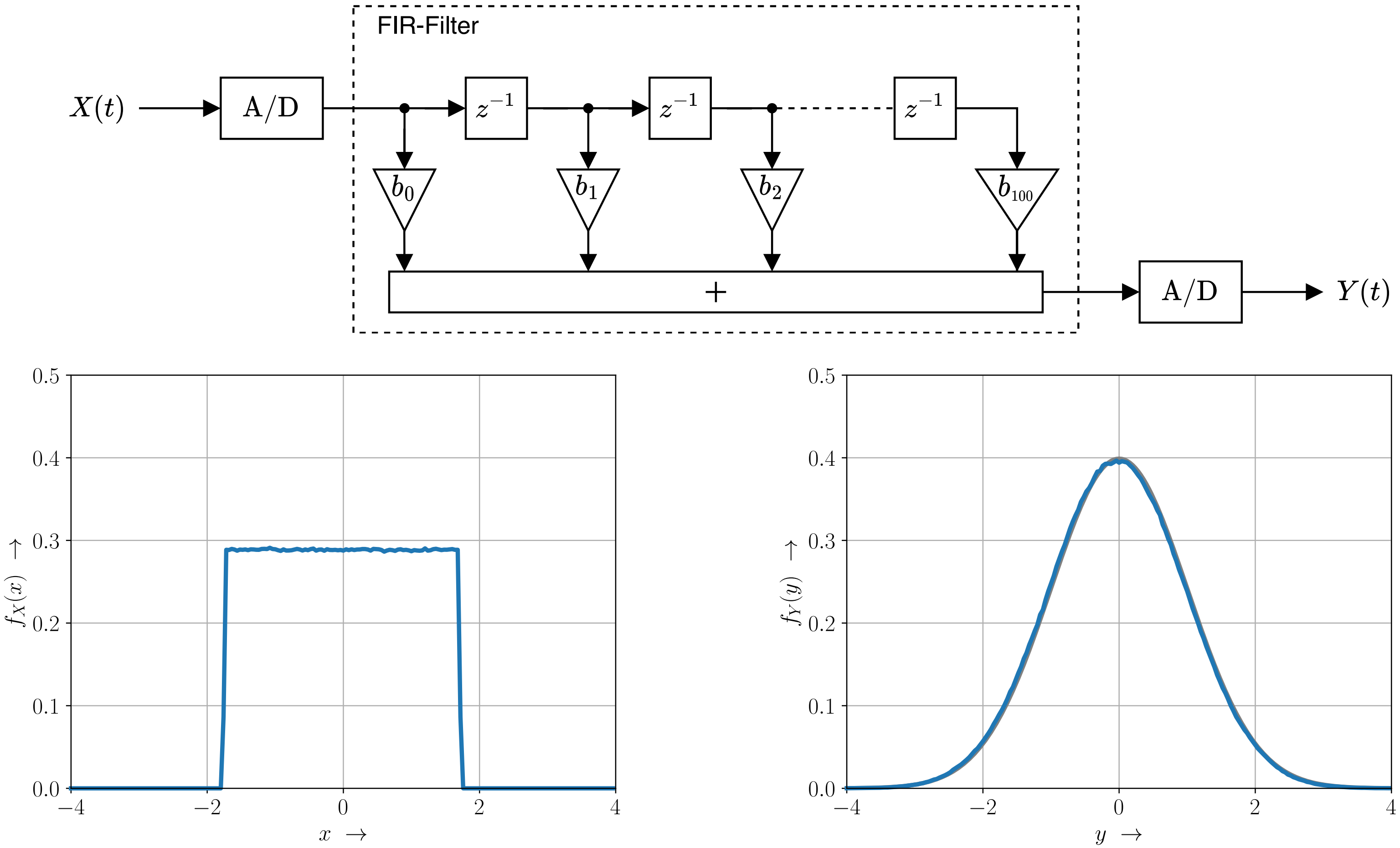






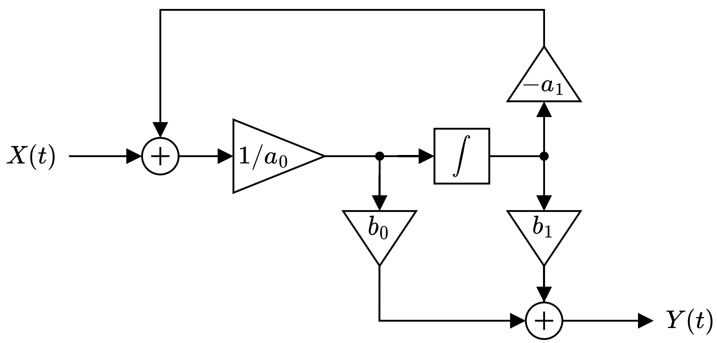
Anwendungsbeispiel: Tiefpassfilter 100. Ordnung

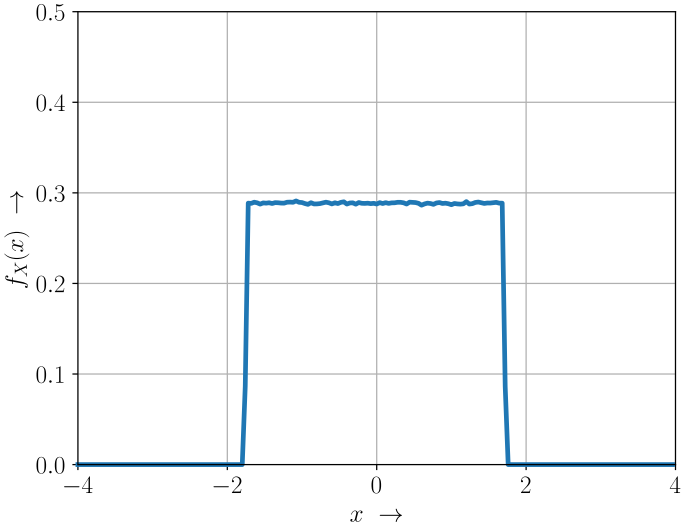
Tiefpassfilterung eines gleichverteilten Eingangsprozesses $X(t)$ mit zeitdiskreten FIR-Filter 100. Ordnung



Anwendungsbeispiel: Butterworth Filter 1. Ordnung

Tiefpassfilterung eines gleichverteilten Eingangsprozesses $X(t)$ mit Filter 1. Ordnung





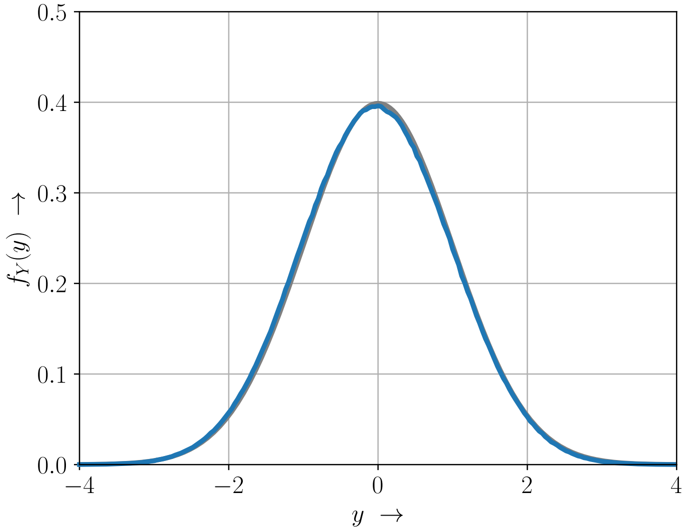


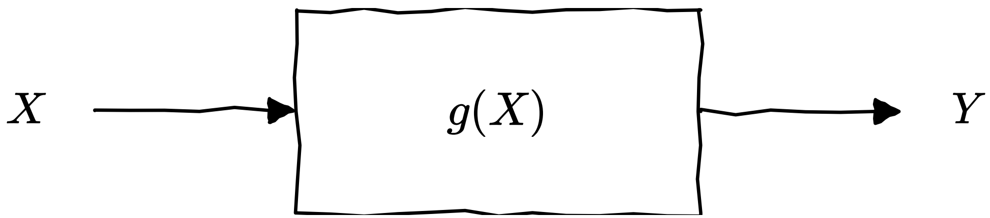
Abbildung von Zufallsvariablen

Abbildung von Zufallsvariablen

Gegeben eine Zufallsvariable X mit Wahrscheinlichkeitsdichte $f_X(x)$

Erzeugen einer neuen Zufallsvariablen mittels beliebiger Abbildung

$$Y = g(X)$$



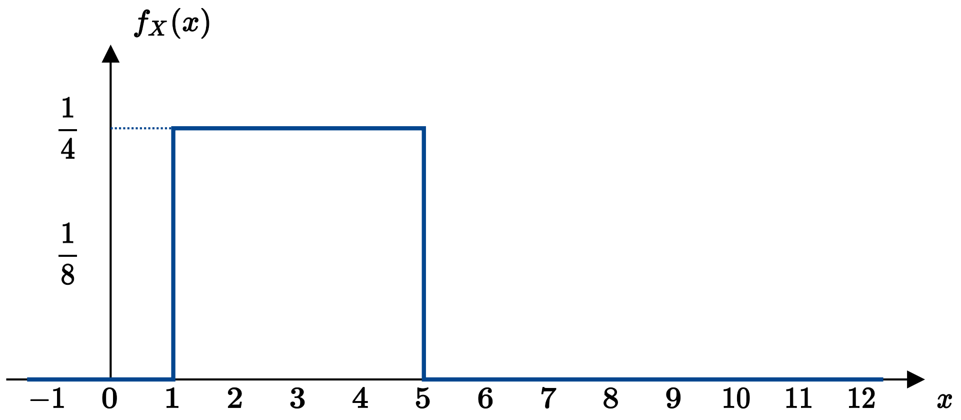
Aufgabe: Bestimmung der Wahrscheinlichkeitsdichtefunktion von $f_Y(y)$

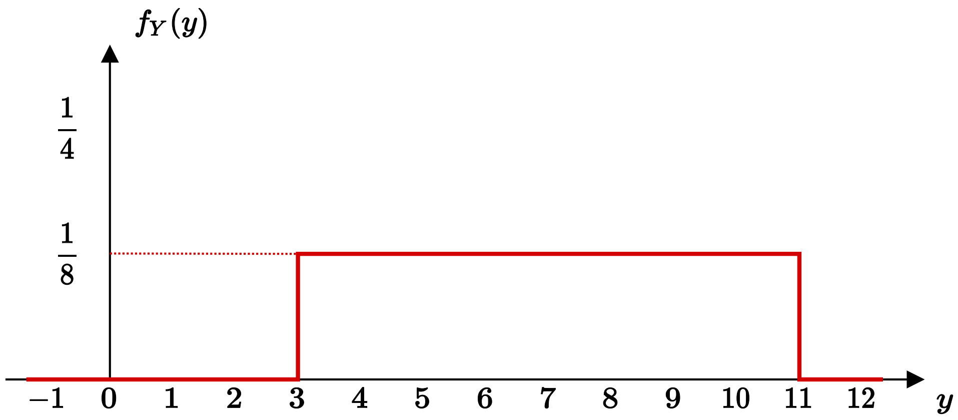
Beispiel: Abbildung einer Gleichverteilung mit linearer Funktion I

Gegeben: Gleichverteilte Zufallsvariable X im Bereich $x = 1, \ldots, 5$

Lineare Abbildung mittels Funktion

$$Y = g(X) = 2 \cdot x + 1$$





Beispiel: Abbildung einer Gleichverteilung mit linearer Funktion II

Wahrscheinlichkeit für X im Intervall der Breite Δx: $[x_0, x_0 + \Delta x]$

$$P(x_0 < X < x_0 + \Delta x) = \int_{x_0}^{x_0+\Delta x} f_X(x)\,\mathrm{d}x$$

Gleiche Wahrscheinlichkeit der Zufallsvariable Y

$$P(y_0 < Y < y_0 + \Delta y) = \int_{y_0}^{y_0+\Delta y} f_Y(y)\,\mathrm{d}y$$

mit

$$y_0 = g(x_0)$$

$$y_0 + \Delta y = g(x_0 + \Delta x)$$

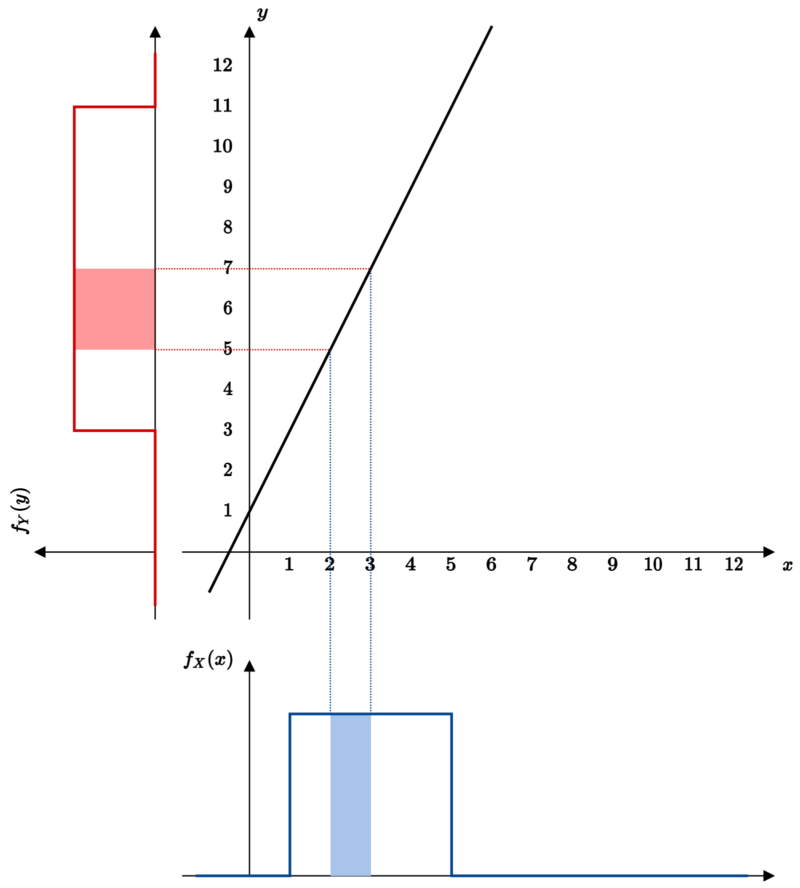


Abbildung einer beliebigen Verteilung mit linearer Funktion I

Gegeben: Zufallsvariable X mit Wahrscheinlichkeitsdichtefunktion $f_X(x)$

Abbildungsvorschrift: $Y = g(X) = a \cdot X + b$ mit $a > 0$

Übergang der Betrachtung zu einem infinitessimal kleinen Interval $\Delta x \to \mathrm{d}x$

$$P(x_0 < X < x_0 + \mathrm{d}x) = f_X(x_0) \cdot \mathrm{d}x \quad \overset{!}{=} \quad P(y_0 < Y < y_0 + \mathrm{d}y) = f_Y(y_0) \cdot \mathrm{d}y$$

Da Beginn des Intervalls x_0 bzw. y_0 beliebig gilt

$$f_X(x) \cdot \mathrm{d}x = f_Y(y) \cdot \mathrm{d}y$$

Damit gilt für die Wahrscheinlichkeitsdichtefunktion der transformierte Zufallsvariable Y

$$f_Y(y) = f_X(x) \cdot \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{f_X(x)}{\frac{\mathrm{d}y}{\mathrm{d}x}} = \frac{f_X(x)}{g'(x)}$$

Achtung: Für $a < 0$ wäre $f_Y(y) < 0$ und keine gültige Wahrscheinlichkeitsdichtefunktion!

Beispiel: Abbildung mit einer linearen Funktion mit negativer Steigung

Lineare Abbildung mittels Funktion

$$Y = g(X) = -2 \cdot x + 12$$

Vertauschen der Intervallgrenzen von $f_Y(y)$

Damit gilt für die Wahrscheinlichkeiten der beiden Intervalle

$$P(x_0 < X < x_0 + \Delta x) = P(y_0 + \Delta y < Y < y_0)$$

Durch die negative Steigung gilt für die Intervallbreite bei Y

$$\Delta y < 0$$

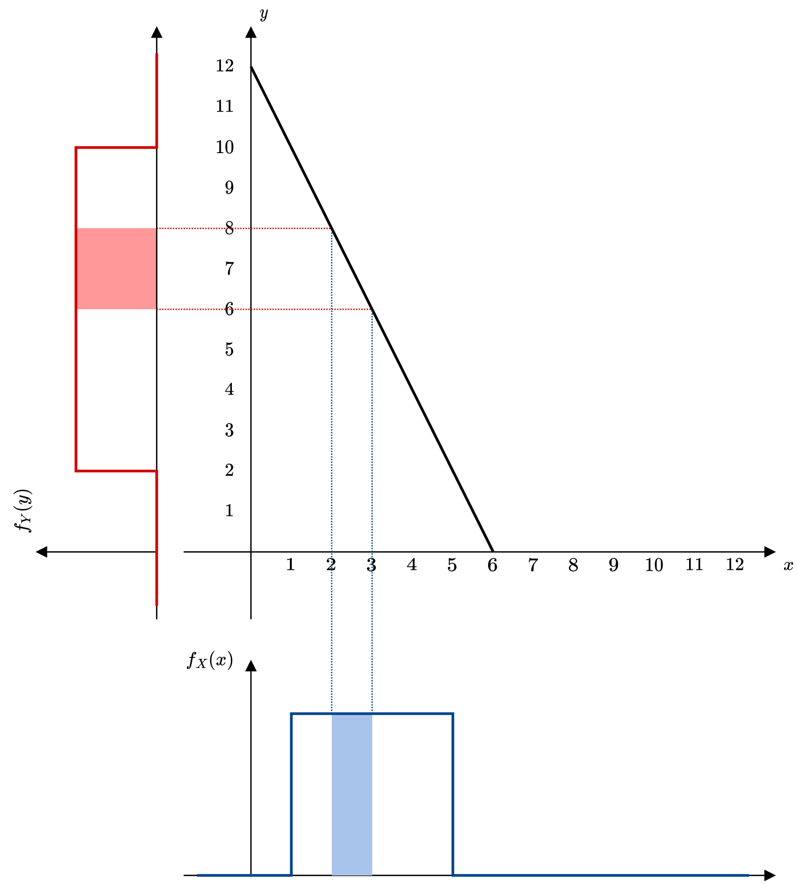


Abbildung einer beliebigen Verteilung mit linearer Funktion II

Gegeben: Zufallsvariable X mit Wahrscheinlichkeitsdichtefunktion $f_X(x)$

Abbildungsvorschrift: $Y = g(X) = a \cdot X + b$ mit $a < 0$

Übergang der Betrachtung zu einem infinitessimal kleinen Interval $\Delta x \rightarrow \mathrm{d}x$

$$P(x < X < x + \mathrm{d}x) = f_X(x) \cdot \mathrm{d}x \quad \stackrel{!}{=} \quad P(y + \mathrm{d}y < Y < y) = f_Y(y) \cdot (-\mathrm{d}y) = f_Y(y) \cdot |\mathrm{d}y|$$

Damit gilt für die Wahrscheinlichkeitsdichtefunktion der transformierte Zufallsvariable Y

$$f_Y(y) = f_X(x) \cdot \frac{\mathrm{d}x}{|\mathrm{d}y|} = \frac{f_X(x)}{\left|\frac{\mathrm{d}y}{\mathrm{d}x}\right|} = \frac{f_X(x)}{|g'(x)|}$$

Durch Anwendung des Betrages gilt dieser Zusammenhang allgemein $\forall a \in \mathbb{R}$

Allgemeiner Zusammenhang der Wahrscheinlichkeitsdichtefunktionen

Die Überlegungen lassen sich auf auf beliebige Funktionen $Y = g(X)$ erweitern

Voraussetzung: $g(x)$ ist streng monoton (injektiv) über dem Definitionsbereich

Berechnung der Wahrscheinlichkeitsdichtefunktion von Y

$$f_Y(y) = \frac{f_X(x)}{|g'(x)|}$$

Falls $g(x)$ nicht streng monoton

- Anwendung der gleichen Überlegungen ebenfalls möglich
- Berücksichtigung Mehrdeutigkeiten der Zufallsvariablen X bei Betrachtung eines Intervalls von Y

Einfluss einer Abbildung durch lineare Funktion auf Erwartungswert und Varianz I

Gegeben: Zufallsvariable X mit beliebiger Verteilung $f_X(x)$ und Mittelwert μ_X und Varianz σ_X^2

Abbildung auf neue Zufallsvariable Y mit linearer Funktion (z.B. Verstärkerkennlinie mit Offset)

$$Y = g(X) = a \cdot X + b \qquad \text{mit} \qquad a > 0$$

Mittelwert der neuen Zufallsvariablen

$$\mu_Y = \mathrm{E}\{Y\} = \int\limits_{-\infty}^{+\infty} y \cdot f_Y(y)\,\mathrm{d}y = \int\limits_{-\infty}^{+\infty} g(x) \cdot \frac{f_X(x)}{\frac{\mathrm{d}y}{\mathrm{d}x}}\,\mathrm{d}y = \int\limits_{-\infty}^{+\infty} g(x) \cdot f_X(x)\,\mathrm{d}x =$$

$$= \int\limits_{-\infty}^{+\infty} (a \cdot x + b) \cdot f_X(x)\,\mathrm{d}x = a \cdot \int\limits_{-\infty}^{+\infty} x \cdot f_X(x)\,\mathrm{d}x + b \cdot \int\limits_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x =$$

$$= a \cdot \mu_X + b$$

Einfluss einer Abbildung durch lineare Funktion auf Erwartungswert und Varianz II

Varianz der neuen Zufallsvariablen

$$\sigma_Y^2 = \mathrm{E}\{(Y-\mu_Y)^2\} = \int\limits_{-\infty}^{+\infty} (y-\mu_Y)^2 \cdot f_Y(y)\,\mathrm{d}y = \int\limits_{-\infty}^{+\infty} (g(x)-\mu_Y)^2 \cdot \frac{f_X(x)}{\frac{\mathrm{d}y}{\mathrm{d}x}}\,\mathrm{d}y =$$

$$= \int\limits_{-\infty}^{+\infty} (a\cdot x + b - a\cdot\mu_X - b)^2 \cdot f_X(x)\,\mathrm{d}x = a^2 \cdot \int\limits_{-\infty}^{+\infty} (x-\mu_X)^2 \cdot f_X(x)\,\mathrm{d}x =$$

$$= a^2 \cdot \sigma_X^2$$

Offset b hat keinen Einfluss auf die Varianz

Varianz skaliert mit quadratischem Verstärkungsfaktor

Umrechnung verschiedener Wahrscheinlichkeitsdichtefunktionen

Nicht für jede Wahrscheinlichkeitsverteilung existiert ein Zufallszahlengenerator (z.B. in Matlab).

Gegeben: Zwei Wahrscheinlichkeitsdichtefunktionen $f_X(x)$ und $f_Y(y)$

Hierfür lässt sich allgemein eine Abbildung $Y = g(X)$ bestimmen mit

$$F_Y(g(x)) = F_X(x)$$

Somit lassen sich ausgehend von einem Zufallsgenerator beliebig verteilte Zufallszahlen generieren.

Umrechnung einer gleichverteilten Zufallsvariablen in eine Gauß-verteilte

Beispiel: Umrechnung einer gleichverteilten Zufallsvariablen X in eine Gauß-verteilte Zufallsvariable Y

Gegeben: Wahrscheinlichkeitsverteilungen für X und Y

$$F_X(x) = \frac{x}{2} + \frac{1}{2} \qquad -1 \leq x \leq +1$$

$$F_Y(y) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{y}{\sqrt{2}}\right)\right)$$

Gesucht: Umrechnung für $Y = g(X)$

$$F_Y(y) = F_Y(g(x)) = F_X(x)$$

$$\frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{y}{\sqrt{2}}\right)\right) = \frac{x}{2} + \frac{1}{2}$$

$$\operatorname{erf}\left(\frac{y}{\sqrt{2}}\right) = x$$

$$y = g(x) = \sqrt{2} \cdot \operatorname{erf}^{-1}(x)$$

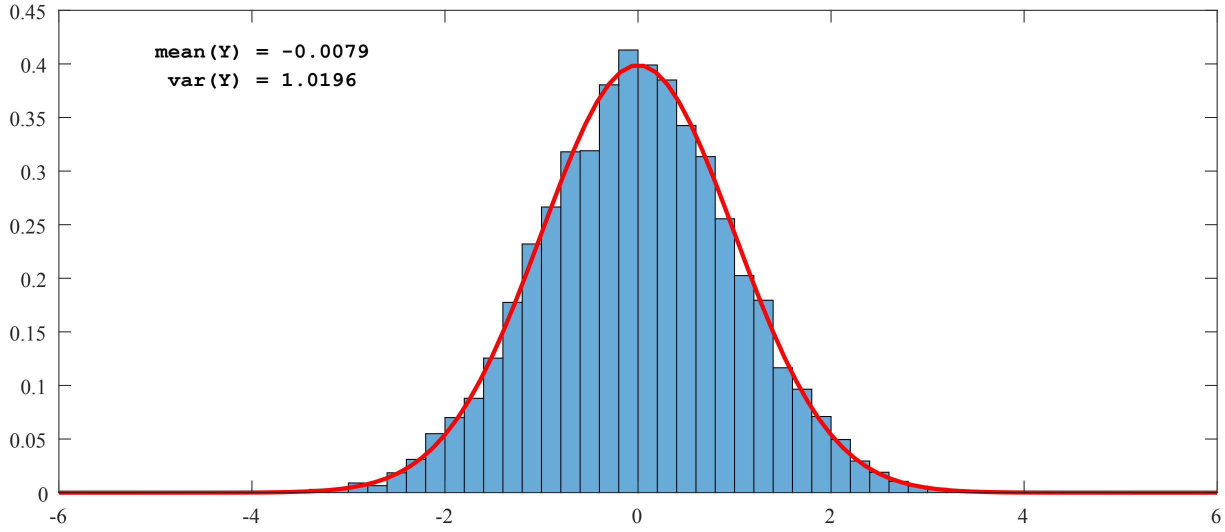
Matlab Beispiel: Umrechnung einer Gleichverteilung in eine Gauß-Verteilung

```
N = 10000;
X = 2*rand(1, N)-1;
Y = sqrt(2)*erfinv(X);

figure(1), clf
histogram(Y, 'Normalization', 'pdf')
hold on

y = linspace(-6, 6, 100);
fY = @(x) 1/sqrt(2*pi) * exp(-x.^2/2);
plot(y, fY(y), 'r', 'LineWidth', 2)

text(-5, 0.4, sprintf('mean(Y) = %6.4f\n var(Y) = %6.4f', mean(Y), var(Y)), 'FontName', 'Monospaced', 'FontWeight', 'bold')
```



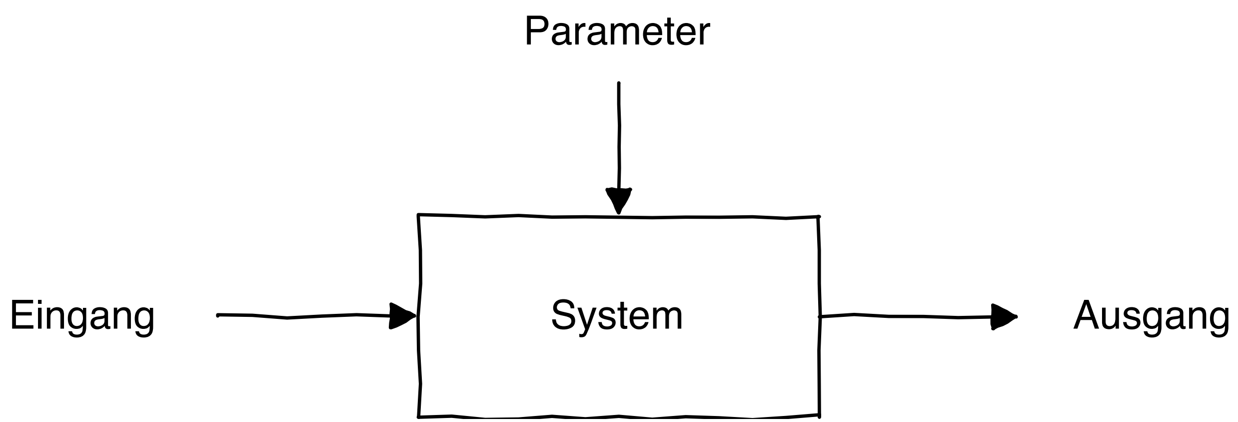
Schätztheorie

Aufgabenstellung der Schätztheorie

Gegeben ist ein System dessen Eigenschaften durch *Parameter* definiert werden.

Das System bildet eine oder mehrere Variablen am *Eingang* auf eine oder mehrere Variablen am *Ausgang* ab.

Die Variable am Ausgang wird auch als *Beobachtung* bezeichnet.

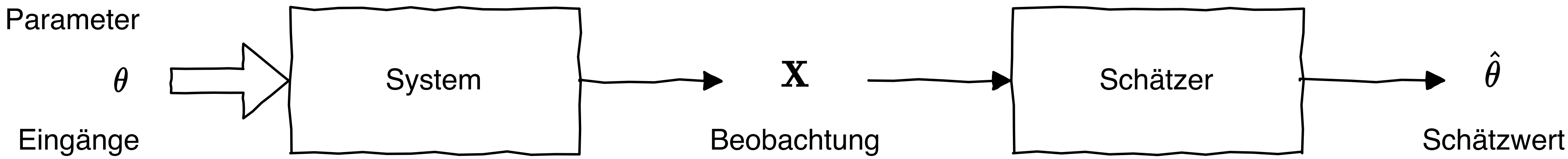


Im Folgenden werden die Eingangsgrößen und Parameter zu einem Vektor θ zusammengefasst.

Dieser Vektor θ ist deterministisch und keine Zufallsvariable.

Aufgabe: Finden eines Schätzwertes $\hat{\theta}$ durch Beobachtung der Variablen am Ausgang.

System und Schätzer



Vektor θ ist deterministische Größe

Beobachtungsvektor $\mathbf{X}$ ist eine Zufallsvariable

Schätzer beschreibt einen Algorithmus oder eine Rechenvorschrift, um aus $\mathbf{X}$ einen Schätzwert $\hat{\theta}$ zu erhalten.

Hinweis: Häufig wird nicht zwischen Eingängen und Parametern unterschieden und der Vektor θ wird nur als Parametervektor bezeichnet.

Entwurfskriterium für einen Schätzer

Aufgabe des Schätzers ist es einen möglichst guten Schätzwert θ zu finden.

Was bedeutet dabei ein *möglichst guter Schätzwert*?

Von Interesse ist vor allem der Fehler zwischen Parameter θ und Schätzwert $\hat{\theta}$

$$\mathbf{e} = \theta - \hat{\theta}$$

Häufig wird als Entwurfskriterium der *mittlere quadratische Fehler* (engl. mean squared error MSE) verwendet

$$\mathrm{MSE} = \mathrm{E}\big\{\mathbf{e}^{\mathrm{T}} \cdot \mathbf{e}\big\} = \mathrm{E}\Big\{(\theta - \hat{\theta})^{\mathrm{T}} \cdot (\theta - \hat{\theta})\Big\}$$

Ziel ist es natürlich den MSE zu minimieren.

Damit lässt sich der Entwurf des Schätzers als Optimierungskriterium formulieren

$$\mathrm{MSE} \to \min$$

Dieses Kriterium wird als als *minimum mean squared error* (MMSE) bezeichnet.

Bewertungskriterien für einen Schätzer

Die Aussagekraft der Schätzwerte eines Schätzers lässt sich anhand folgender Bewertungskriterien ermitteln

- Erwartungswert der Schätzwerte

$$\mathrm{E}\left\{\hat{\theta}\right\}$$

- Systematischer Fehler (engl. *bias*)

$$\mathrm{E}\left\{\hat{\theta}-\theta\right\}=\mathrm{E}\left\{\hat{\theta}\right\}-\theta$$

- Kovarianz der Schätzung

$$C_{\hat{\theta}\hat{\theta}}=\mathrm{E}\left\{(\hat{\theta}-\mathrm{E}\{\theta\})\cdot(\hat{\theta}-\mathrm{E}\{\theta\})^{\mathrm{T}}\right\}$$

Der Schätzer ist besser je kleiner der systematische Fehler und die Kovarianz (d.h. Streuung) der Schätzung

Eigenschaften von Schätzern

Mit Hilfe der Bewertungskriterien für Schätzer lassen sich auch gewisse Eigenschaften formulieren

- Erwartungstreuer Schätzer (engl. *unbiased estimator*)

$$\mathrm{E}\left\{\hat{\theta}\right\} = \theta$$

- Asymptotisch erwartungstreuer Schätzer (engl. *asymptotically unbiased estimator*)

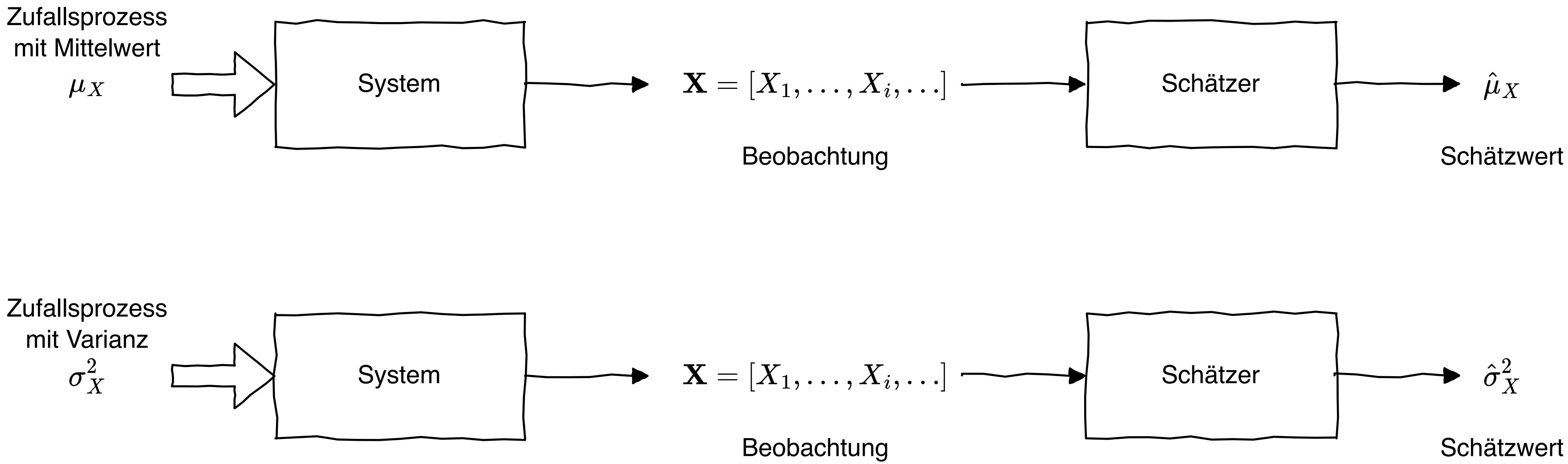
$$\lim_{N\to\infty} \mathrm{E}\left\{\hat{\theta}\right\} = \theta \qquad N \text{ Anzahl der Datenpunkte}$$

- Konsistenter Schätzer (engl. *consistent estimator*): $\lim\limits_{N\to\infty} P(|\hat{\theta} - \theta| > \epsilon) = 0$
- Effizienter Schätzer (engl. *efficient estimator*): $C_{\hat{\theta}^*\hat{\theta}^*} = \min\limits_{\hat{\theta}} C_{\hat{\theta}\hat{\theta}}$

Zwei Beispiele: Berechnung von Mittelwert und Varianz eines Zufallsexperimentes

Beispiel für einen Schätzer: Berechnung der Eigenschaften (Parameter) eines konkreten Zufallsexperimentes

- Mittelwert μ_X
- Varianz σ_X^2



Arithmetischer Mittelwert als Schätzwert für den Mittelwert einer Zufallsvariablen

Gegeben: N Ergebnisse X_i eines Zufallsexperimentes mit unbekannten statistischen Eigenschaften

Gesucht: Schätzwert für den Erwartungswert $\hat{\mu}_X$

Ansatz: Arithmetischer Mittelwert als Schätzer

$$\hat{\mu}_X = \frac{1}{N}\sum_{i=1}^{N} X_i$$

Analyse des mittleren Schätzwertes

$$\mathrm{E}\{\hat{\mu}_X\} = \mathrm{E}\left\{\frac{1}{N}\sum_{i=1}^{N} X_i\right\} = \frac{1}{N}\mathrm{E}\left\{\sum_{i=1}^{N} X_i\right\} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\{X_i\} = \frac{1}{N}\sum_{i=1}^{N}\mu_X = \mu_X$$

Damit ist der arithmetische Mittelwert ein *erwartungstreuer Schätzwert* für den Mittelwert einer Zufallsvariablen.

Schätzwert für die Varianz einer Zufallsvariablen I

Gegeben: N Ergebnisse X_i eines Zufallsexperimentes mit unbekannten statistischen Eigenschaften

Gesucht: Schätzwert für die Varianz $\hat{\sigma}_X^2$

Ansatz: Arithmetischer Mittelwert der quadratischen Abweichung vom Schätzwert $\hat{\mu}_X$ des Mittelwertes

$$\hat{\sigma}_X^{*2} = \frac{1}{N} \sum_{i=1}^{N} (X_i - \hat{\mu}_X)^2 = \frac{1}{N} \sum_{i=1}^{N} \left(X_i - \frac{1}{N} \sum_{k=1}^{N} X_k \right)^2$$

(*Hinweis*: Wir bezeichnen den Schätzer zunächst als $\hat{\sigma}_X^{*2}$, da wir später noch einen besseren Ansatz finden)

Diskussion:

- *Spoiler*: Dieser Ansatz ist leider nicht optimal im Sinne der Erwartungstreue
- Trotzdem wird der Ansatz häufig verwendet
- Falls Mittelwert μ_X bekannt (z.B. falls Zufallsvariable mittelwertfrei) ist der Ansatz wieder korrekt

Schätzwert für die Varianz einer Zufallsvariablen II

Betrachtung des mittleren Schätzwertes der Varianz

$$\begin{aligned}
\mathrm{E}\left\{\hat{\sigma}_X^{*2}\right\} &= \mathrm{E}\left\{\frac{1}{N}\sum_{i=1}^{N}\left(X_i - \frac{1}{N}\sum_{k=1}^{N}X_k\right)^2\right\} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left\{\left(X_i - \frac{1}{N}\sum_{k=1}^{N}X_k\right)^2\right\} \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left\{X_i^2 - \frac{2}{N}X_i\sum_{k=1}^{N}X_k + \left(\frac{1}{N}\sum_{k=1}^{N}X_k\right)^2\right\} \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{E}\{X_i^2\} - \frac{2}{N}\mathrm{E}\left\{X_i\sum_{k=1}^{N}X_k\right\} + \mathrm{E}\left\{\left(\frac{1}{N}\sum_{k=1}^{N}X_k\right)^2\right\}\right)
\end{aligned}$$

Im Folgenden werden die drei Summanden innerhalb der Summe $\sum$ einzeln betrachtet.

Zwischenrechnung 1: Auswertung der drei Erwartungswerte innerhalb der Summe

Für die Varianz der Zufallsvariablen gilt:

$$\sigma_X^2 = \mathrm{E}\big\{(X-\mu_X)^2\big\} = \mathrm{E}\big\{X^2 - 2X\mu_X + \mu_X^2\big\} = \mathrm{E}\big\{X^2\big\} - 2\mathrm{E}\{X\}\,\mu_x + \mu_X^2 =$$

$$= \mathrm{E}\big\{X^2\big\} - 2\mu_x^2 + \mu_X^2 = \mathrm{E}\big\{X^2\big\} - \mu_X^2$$

Damit ergibt sich für den Erwartungswert der quadrierten Zufallsvariablen:

$$\mathrm{E}\big\{X^2\big\} = \mathrm{E}\big\{X_i^2\big\} = \sigma_X^2 + \mu_X^2$$

Zwischenrechnung 2: Auswertung der drei Erwartungswerte innerhalb der Summe

$$\mathrm{E}\left\{X_i \sum_{k=1}^{N} X_k\right\} = \mathrm{E}\left\{\sum_{k=1}^{N} X_i \cdot X_k\right\} = \sum_{k=1}^{N} \mathrm{E}\{X_i \cdot X_k\}$$

Für $k = i$ ergibt sich die Varianz der quadrierten Zufallsvariable:

$$\mathrm{E}\{X_i \cdot X_k\} = \mathrm{E}\{X_i^2\} = \sigma_X^2 + \mu_X^2$$

Da X_i, X_k stochastisch unabhängig für $k \neq i$:

$$\mathrm{E}\{X_i \cdot X_k\} = \mathrm{E}\{X_i\} \cdot \mathrm{E}\{X_k\} = \mu_X^2$$

Somit ergibt sich für den Erwartungswert

$$\mathrm{E}\left\{X_i \sum_{k=1}^{N} X_k\right\} = \sigma_X^2 + \mu_X^2 + (N-1) \cdot \mu_X^2 = \sigma_X^2 + N \cdot \mu_X^2$$

Zwischenrechnung 3: Auswertung der drei Erwartungswerte innerhalb der Summe

$$\left(\frac{1}{N}\sum_{k=1}^{N} X_k\right)^2 = \frac{1}{N^2}\left(X_1 + X_2 + X_3 + \ldots + X_N\right)^2 =$$

$$= \frac{1}{N^2}\left(X_1^2 + X_1X_2 + X_1X_3 + \ldots + X_1X_N + X_2X_1 + X_2^2 + X_2X_3 + \ldots + X_N^2\right)$$

$$= \frac{1}{N^2}\left(X_1\sum_{j=1}^{N} X_j + X_2\sum_{j=1}^{N} X_j + \ldots + X_N\sum_{j=1}^{N} X_j\right) = \frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N} X_kX_j$$

Da X_k, X_j stochastisch unabhängig für $k \neq j$:

$$\mathrm{E}\left\{\frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N} X_kX_j\right\} = \frac{1}{N^2}\sum_{k=1}^{N}\sum_{j=1}^{N}\mathrm{E}\{X_kX_j\} = \frac{1}{N^2}\left(N\cdot(\sigma_X^2 + \mu_X^2) + N\cdot(N-1)\cdot\mu_X^2\right) =$$

$$= \frac{1}{N^2}\left(N\sigma_X^2 + N^2\mu_X^2\right) = \frac{\sigma_X^2}{N} + \mu_X^2$$

Schätzwert für die Varianz einer Zufallsvariablen III

Mit Hilfe der drei Zwischenergebnisse folgt für den Schätzwert der Varianz

$$
\begin{aligned}
\mathrm{E}\{\hat{\sigma}_X^{*2}\} &= \frac{1}{N}\sum_{i=1}^{N}\left(\sigma_X^2 + \mu_X^2 - \frac{2}{N}(\sigma_X^2 + N\cdot\mu_X^2) + \frac{\sigma_X^2}{N} + \mu_X^2\right) = \\
&= \frac{1}{N}\cdot N\cdot\left(\sigma_X^2 + \mu_X^2 - \frac{2\sigma_X^2}{N} - 2\mu_X^2 + \frac{\sigma_X^2}{N} + \mu_X^2\right) = \\
&= \sigma_X^2\cdot\left(1 - \frac{2}{N} + \frac{1}{N}\right) = \\
&= \sigma_X^2\cdot\frac{N-2+1}{N} = \\
&= \sigma_X^2\cdot\frac{N-1}{N}
\end{aligned}
$$

Schätzwert für die Varianz einer Zufallsvariablen IV

$$\mathrm{E}\left\{\hat{\sigma}_X^{*2}\right\} = \sigma_X^2 \cdot \frac{N-1}{N} = \sigma_X^2 \cdot \left(1 - \frac{1}{N}\right)$$

Diskussion:

- Der hier verwendete Schätzwert ist offensichtlich nicht erwartungstreu
- Es tritt ein systematischer Fehler auf

$$\mathrm{E}\left\{\hat{\sigma}_X^{*2}\right\} - \sigma_X^2 = -\frac{1}{N}\sigma_X^2$$

- Der Schätzwert ist aber asymptotisch erwartungstreu, da

$$\lim_{N\to\infty} \mathrm{E}\left\{\hat{\sigma}_X^{*2}\right\} = \lim_{N\to\infty} \sigma_X^2 \cdot \left(1 - \frac{1}{N}\right) = \sigma_X^2$$

Verbesserung des Schätzwertes für die Varianz einer Zufallsvariablen

Beseitigung es systematischen Fehlers durch Multiplikation mit $N/(N-1)$

Damit ergibt sich ein verbesserter Schätzwert

$$\hat{\sigma}_X^2 = \frac{N}{N-1} \cdot \hat{\sigma}_X^{*2} = \frac{N}{N-1} \cdot \frac{1}{N} \sum_{i=1}^{N} \left(X_i - \frac{1}{N} \sum_{k=1}^{N} X_k \right)^2$$

$$= \frac{1}{N-1} \sum_{i=1}^{N} \left(X_i - \frac{1}{N} \sum_{k=1}^{N} X_k \right)^2$$

Der Erwartungswert dieses Schätzers entspricht nun der Varianz

$$\mathrm{E}\{\hat{\sigma}_X^2\} = \mathrm{E}\left\{ \frac{N}{N-1} \cdot \hat{\sigma}_X^{*2} \right\} = \frac{N}{N-1} \cdot \mathrm{E}\{\hat{\sigma}_X^{*2}\} = \frac{N}{N-1} \cdot \sigma_X^2 \cdot \frac{N-1}{N} = \sigma_X^2$$

Damit ist dieser Schätzer erwartungstreu!

Referenzen

[1] B.-U. Köhler, *Konzpete der statistischen Signalverarbeitung,* Springer Verlag.

[2] E. Hänsler, *Statistische Signale,* Springer Verlag.

[3] E. Weitz, *Konkrete Mathematik (nicht nur) für Informatiker,* Springer Spektrum.

