4. Vorlesung Statistik I
Wahrscheinlichkeitstheorie I

Gesamtüberblick

- bislang: Deskriptive Statistik
- heute und nächste Vorlesungen: Wahrscheinlichkeitstheorie
- danach: Inferenzstatistik

Ausblick Inferenzstatistik

Deskriptive Statistik und Inferenzstatistik

- Die **Deskriptive Statistik** umfasst alle Methoden, die der Beschreibung von Daten dienen.

- Die **Inferenzstatistik** umfasst alle statistischen Verfahren, die es erlauben, trotz der Informationsunvollständigkeit der Stichprobendaten Aussagen über eine Population zu treffen.

- Zu klären ist:
  - Was ist eine Population (Synonym: Grundgesamtheit)?
  - Was ist eine Stichprobe?
  - Was meint Informationsunvollständigkeit?
  - Was hat die Wahrscheinlichkeitstheorie damit zu tun?

Definition: Population (Synonym: Grundgesamtheit)

- Definition **Population:** Gesamtheit aller Merkmalsträger*innen, auf die eine Untersuchungsfrage gerichtet ist.

- Die Definition der Population ist an eine exakte sachliche, räumliche und zeitliche Abgrenzung gebunden.
- Beispiel:
  - Population: Alle Personen im Alter von 12 bis 80 Jahren in Deutschland im Jahr 2018
  - Untersuchungsfrage: Wie viele dieser Personen leiden an einer sozialen Phobie?

- Wenn wir alle Personen der Population erheben können (= Vollerhebung), können wir die Untersuchungsfrage im Rahmen der Deskriptivstatistik beantworten.
- Im Beispiel oben würden wir zum Beispiel einfach die absolute oder relative Häufigkeit der sozialen Phobie in der Population berechnen.

Definition: Stichprobe

- Es leuchtet unmittelbar ein, dass ab einer gewissen Größe der Population nicht alle Personen erfasst werden können.

- Man erhebt in diesem Fall lediglich einen Teil der Population, eine sogenannte Stichprobe.

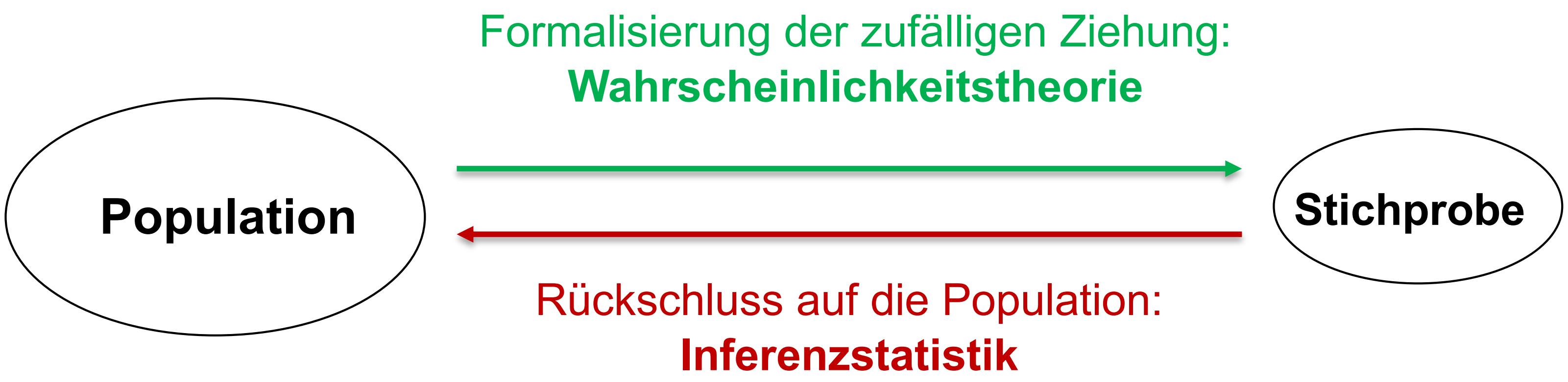
- Definition **Stichprobe:** Auswahl bestimmter Merkmalsträger*innen aus einer Population.

Inferenzstatistik I

- Problem:
  - Wenn nur ein Teil der Grundgesamtheit erfasst wird, z.B. 500 Personen, ist die Informationslage in Bezug auf die Untersuchungsfrage unvollständig. Wir können nicht einfach deskriptiv-statistische Methoden verwenden.
  - Wie kann man trotzdem Aussagen treffen, die sich auf alle Personen der Grundgesamtheit beziehen, obwohl nur die Daten einer Stichprobe vorliegen?

- Idee:
  - Wir ziehen die Personen **zufällig** aus der Population in die Stichprobe.
  - Wir greifen auf mathematische Methoden zur Formalisierung von Zufallsprozessen zurück → **Wahrscheinlichkeitstheorie**
  - Aus diesen ergeben sich Methoden, die Rückschlüsse von der Stichprobe auf die Population erlauben → **Inferenzstatistik**

Inferenzstatistik II



Deskriptive Statistik:
Beschreibung der Population
(nur bei Vollerhebung)

Deskriptive Statistik:
Beschreibung der Stichprobe

Überblick

- bislang:
    - Ausblick Inferenzstatistik

- jetzt: Zufallsexperimente

Zufallsexperimente

Definition: Zufallsexperiment

- Ein **Zufallsexperiment** ist ein Vorgang, der **mehr als einen möglichen Ausgang** haben kann und bei dem wir **vorher nicht wissen, welcher dieser Ausgänge eintreten wird**.

- Charakteristika des Zufallsexperiments:
  - Es existiert eine Menge möglicher Ausgänge des Zufallsexperiments.
  - Es ist vor der Durchführung unbekannt, welcher der möglichen Ausgänge tatsächlich eintreten wird (Zufallsabhängigkeit).

- Wichtig: Ein Zufallsexperiment muss kein „Experiment“ im klassischen Sinne sein.

Definition: Ergebnis/Ereignis/Ergebnismenge/Ereignismenge

- Die möglichen Ausgänge eines Zufallsexperimentes nennt man **Ergebnisse**. Ein beliebiges Ergebnis wird mit dem Symbol ω bezeichnet.

- Wenn man alle möglichen Ergebnisse eines Zufallsexperimentes in einer Menge zusammenfasst, erhält man die **Ergebnismenge**. Sie wird mit dem Symbol Ω bezeichnet.

- Jede Zusammenfassung von mehreren Ergebnissen eines Zufallsexperiments in einer Menge nennt man **Ereignis**. Ein Ereignis ist damit eine Teilmenge von Ω. Wir bezeichnen Ereignisse meistens mit Großbuchstaben: A, B, …

- Enthält ein Ereignis nur ein Ergebnis, spricht man von einem **Elementarereignis**.

- Wenn alle möglichen Ereignisse in einer Menge zusammengefasst werden, erhält man die **Ereignismenge**. Sie wird mit dem Symbol $\mathcal{A}$ bezeichnet. $\mathcal{A}$ ist somit die Menge aller möglichen Teilmengen von Ω.
  Hinweis: Per Definition ist in $\mathcal{A}$ auch Ω selbst, sowie das (eigentlich unmögliche) Ereignis { }, dass kein Ergebnis aus Ω eintritt, enthalten.

Beispiel Zufallsexperiment: Würfelwurf

- Zufallsexperiment: Wurf eines Würfels

- Mögliche **Ergebnisse**: ⚀⚁⚂⚃⚄⚅

- Ergebnismenge: $\Omega = \{$⚀ , ⚁ , ⚂ , ⚃ , ⚄ , ⚅$\}$

- Einige denkbare **Ereignisse** (z.B. in einem Würfelspiel):
  - „Beim Wurf einer ⚅ muss die nächste Spielerin eine Runde aussetzen.“
    → Elementarereignis {⚅}
  - „Bei einem Wurf mit **gerader** Anzahl der Würfelaugen, darf eine Spielerin gleich noch mal würfeln.“
    → Ereignis $A_1 = \{$⚁ , ⚃ , ⚅$\}$
  - „Bei einem Wurf mit einer Anzahl der **Würfelaugen** die **kleiner ist als 4**, werden die Plätze von zwei Spieler*innen getauscht.“
    → Ereignis $A_2 = \{$⚀ , ⚁ , ⚂$\}$

- Ereignismenge: Die Menge **aller Ereignisse** die bei einem Würfelwurf möglich sind. Lassen wir hier aus, da sehr groß.

Überblick

- bislang:
    - Ausblick Inferenzstatistik
    - Zufallsexperimente
- jetzt: Wahrscheinlichkeitsverteilungen

Wahrscheinlichkeitsverteilungen

Beispiel fairer Würfelwurf (W6)

- $\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$

- Eine **Wahrscheinlichkeitsverteilung** P ist eine Funktion, die jedem Ereignis A aus der Ereignismenge $\mathcal{A}$ eine Wahrscheinlichkeit in Form einer reellen Zahl $P(A)$ zwischen 0 und 1 zuweist.

- Die Wahrscheinlichkeitsverteilung würde also jetzt **allen Ereignissen** eine Wahrscheinlichkeit zuordnen.

- Die Menge $\mathcal{A}$ aller Ereignisse ist hier schon zu groß um das noch übersichtlich darzustellen.

- Praktischerweise genügt es, lediglich die Wahrscheinlichkeiten der **Elementarereignisse** eines fairen Würfels zu kennen:

$$P(\{⚀\}) = P(\{⚁\}) = P(\{⚂\}) = P(\{⚃\}) = P(\{⚄\}) = P(\{⚅\}) = \frac{1}{6}$$

- Alle Wahrscheinlichkeiten der weiteren Ereignisse können daraus berechnet werden.

Beispiel Zufallsexperiment: Würfelwurf

- Einige denkbare **Ereignisse** (z.B. in einem Würfelspiel) und deren Wahrscheinlichkeiten:
  - „Beim Wurf einer ⚅ muss die nächste Spielerin eine Runde aussetzen.“
    → Elementarereignis {⚅}
    → $P(\{⚅\}) = \frac{1}{6}$
  - „Bei einem Wurf mit **gerader** Anzahl der Würfelaugen, darf eine Spielerin gleich noch mal würfeln.“
    → Ereignis $A_1 = \{⚁, ⚃, ⚅\}$
    → $P(\{A_1\}) = P(\{⚁\}) + P(\{⚃\}) + P(\{⚅\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 0{,}5$
  - „Bei einem Wurf mit einer Anzahl der **Würfelaugen** die **kleiner ist als 4**, werden die Plätze von zwei Spieler*innen getauscht.“
    → Ereignis $A_2 = \{⚀, ⚁, ⚂\}$
    → $P(\{A_2\}) = P(\{⚀\}) + P(\{⚁\}) + P(\{⚂\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 0{,}5$

Was ist eigentlich eine Wahrscheinlichkeit?

- Eine Wahrscheinlichkeit muss die **Kolmogoroff-Axiome** erfüllen:
  - Eigenschaft 1: $P(A) \geq 0$
  - Eigenschaft 2: $P(\Omega) = 1$
  - Eigenschaft 3: Wenn $A \cap B = \{\}$, dann ist $P(A \cup B) = P(A) + P(B)$.
- Wahrscheinlichkeiten sind nur für Ereignisse definiert, nicht für Ergebnisse.
- **Aber**: Wie interpretieren wir eigentlich eine Aussage wie $P(\{\text{⚅}\}) = \frac{1}{6}$?
- Die wichtigsten Interpretationen sind
  - die intuitive alltagssprachliche Interpretation (erfüllt häufig nicht die Axiome)
  - die frequentistische Interpretation
  - die bayesianische Interpretation
  - die spieltheoretische Interpretation
- Wir werden im Rahmen dieser Vorlesung ausschließlich die frequentistische Interpretation besprechen und verwenden.

frequentistische Interpretation I

- Ausgangspunkt der frequentistischen Wahrscheinlichkeitsinterpretation ist die hypothetische Überlegung, dass das Zufallsexperiment n mal unter identischen Bedingungen durchgeführt werden könnte.

- Bei jedem dieser Durchgänge könnte man dann feststellen, ob ein Ereignis A aufgetreten ist oder nicht.

- Die Anzahl der Durchgänge, in denen das Ereignis A aufgetreten ist, wird mit n_A bezeichnet.

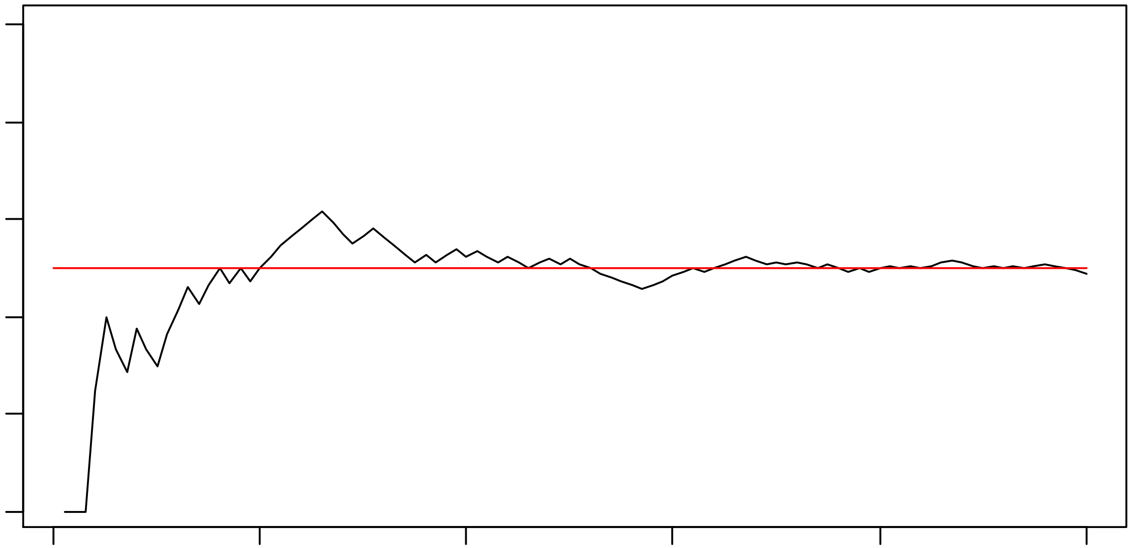
- Der Quotient $\frac{n_A}{n}$ ist die relative Häufigkeit des Auftretens des Ereignisses A bei n Durchgängen des Zufallsexperiments.

- Die Wahrscheinlichkeit eines Ereignisses A wird dann als Grenzwert dieser relativen Häufigkeit $\frac{n_A}{n}$ bei $n \to \infty$ interpretiert (vereinfacht gesagt).

frequentistische Interpretation II

- Illustration: Eine faire Münze wird insgesamt 100-mal geworfen. Es wird ermittelt, wie oft „Kopf" bei einer unterschiedlichen Anzahl von Durchgängen beobachtet wird.

$\boldsymbol{n}$	$\boldsymbol{n_A}$	$\boldsymbol{\frac{n_A}{n}}$
10	3	.30
20	10	.50
30	17	.57
40	21	.53
50	26	.52
60	29	.48
70	36	.51
80	40	.50
90	45	.50
100	49	.49



- bei wachsendem n: Stabilisierung des Quotienten $\frac{n_A}{n}$ um einen konstanten Wert.
- Dieser Grenzwert wird als Wahrscheinlichkeit interpretiert.

frequentistische Interpretation III

- Die Aussage $\mathrm{P}(\{Kopf\}) = 0.5$ bei einem Münzwurf hat also in der frequentistischen Wahrscheinlichkeitsinterpretation folgende Bedeutung:

- Wenn man die Münze unendlich oft werfen würde, würde bei 50% dieser Würfe Kopf oben landen.

Überblick

- bislang:
    - Ausblick Inferenzstatistik
    - Mathematische Grundbegriffe
    - Zufallsexperimente
    - Wahrscheinlichkeitsverteilungen

- jetzt: Zufallsvariablen

Zufallsvariablen

Zufallsvariablen

- Eine Zufallsvariable (ZV) ist eine Funktion X, die jedem Ergebnis $\omega \in \Omega$ eines Zufallsexperiments eine reelle Zahl $\mathrm{X}(\omega) \in \mathbb{R}$ zuordnet.
- In vielen Fällen repräsentieren Zufallsvariablen einen Messvorgang:
  Ω ist dann die Menge der Merkmalsausprägungen, denen durch X jeweils ein Variablenwert zugeordnet wird (siehe VL 1).
- Durch Zufallsvariablen können bestimmte Aspekte des Zufallsexperiments abgebildet werden.
- Man kann mehrere Zufallsvariablen für das gleiche Zufallsexperiment betrachten.
- Durch die „Übersetzung“ in Zahlen wird die Anwendung mathematischer Modelle möglich gemacht oder vereinfacht.
- Indem wir das Zufallsexperiment über die ZV formalisieren, können wir uns auch direkt die Eigenschaften der ZV ansehen, ohne dass wir das Zufallsexperiment wirklich durchführen müssen. Aussagen über die Zufallsvariablen können dann direkt auf das zugrunde liegende Experiment bezogen werden.

Zufallsvariablen - Beispiele

- Münzwurf – (Wie oft) liegt „Kopf“ oben?

ω	„Zahl“	„Kopf“
$X(\omega)$	0	1

- Würfelwurf – Wie viele Würfelaugen liegen oben?

ω	⚀	⚁	⚂	⚃	⚄	⚅
$X(\omega)$	1	2	3	4	5	6

- Schulabschluss – Hat eine Person Abitur?

ω	Kein Abitur	Abitur
$X(\omega)$	0	1

- Intelligenz – Wie intelligent ist eine Person im Vergleich zu einer Gruppe von Personen?

ω	eine s_{emp} weniger als der Durchschnitt	…	durchschnittlich	…	eine s_{emp} mehr als der Durchschnitt
$X(\omega)$	85	…	100	…	115

Zufallsvariablen und ihre Realisationen

- Zufallsvariablen werden im Allgemeinen mit Großbuchstaben bezeichnet: X, Y, etc.
- Nach der Durchführung des Zufallsexperiments tritt ein Ergebnis ω ein und die Zufallsvariable X nimmt einen Wert $\mathrm{x} = X(\omega)$ an.
- Dieser Wert x wird **Realisation** der ZV X genannt
- Realisationen von ZVs werden mit den der ZV entsprechenden Kleinbuchstaben bezeichnet.
- Falls im Beispiel auf der letzten Folie z.B. das Ergebnis $\omega =$ ⚅ auftreten würde, würde sich die Zufallsvariable *X* im Wert $\mathrm{x} = X($⚅$) = 6$ realisieren.

Neue Ergebnisse und Ereignisse

- Die Menge aller möglichen Realisationen einer ZV X nennt man **Träger** der Zufallsvariable X. Er wird mit dem Symbol T_X bezeichnet. T_X ist eine **neue Ergebnismenge**, diesmal bestehend aus Zahlen.

- Die Teilmengen des Trägers T_X sind **neue Ereignisse** A_X.

- Die Menge aller dieser Ereignisse ist eine **neue Ereignismenge** $\mathcal{A}_X$.

Beispiel: einfacher Würfelwurf

- Zufallsexperiment/-variable: (Wie oft) wird bei einem einmaligen Würfelwurf eine ⚄ oder ⚅ gewürfelt.

- $\Omega = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$

ω	⚀	⚁	⚂	⚃	⚄	⚅
$X(\omega)$	0	0	0	0	1	1

- Träger $T_X = \{0,1\}$

- Ereignismenge $\mathcal{A}_X = \{\{\}, \{0\}, \{1\}, \{0, 1\}\}$

  Hinweis: Per Definition sind in jeder Ereignismenge auch das sichere Ereignis Ω („irgendein Ergebnis aus dem Träger T_X tritt auf“) sowie das unmögliche Ereignis $\{\}$ („kein Ergebnis aus dem Träger T_X tritt auf“) enthalten.

Wahrscheinlichkeitsverteilung von Zufallsvariablen I

- Jede ZV X hat eine Wahrscheinlichkeitsverteilung P_X, die jedem Ereignis A_x aus $\mathcal{A}_X$ eine Wahrscheinlichkeit $P_X(A_x)$ zuweist.
- P_X kann aus der Wahrscheinlichkeitsverteilung P des zugrundeliegenden Zufallsexperiments berechnet werden.
- Da $\mathcal{A}$ und $\mathcal{A}_X$ ggf. sehr groß sein können, schauen wir uns mal nur die Elementarereignisse an:

ω	P	$X(\omega)$	P_X
⚀	$P(\{⚀\}) = 1/6$	0	$P_X(\{0\}) =$ $P(\{⚀\}) + P(\{⚁\}) + P(\{⚂\}) + P(\{⚃\}) =$ $1/6 + 1/6 + 1/6 + 1/6 = 2/3$
⚁	$P(\{⚁\}) = 1/6$	0	
⚂	$P(\{⚂\}) = 1/6$	0	
⚃	$P(\{⚃\}) = 1/6$	0	
⚄	$P(\{⚄\}) = 1/6$	1	$P_X(\{1\}) =$ $P(\{⚄\}) + P(\{⚅\}) =$ $1/6 + 1/6 = 1/3$
⚅	$P(\{⚅\}) = 1/6$	1	

Wahrscheinlichkeitsverteilung von Zufallsvariablen II

Bemerkung: In der Praxis geht es fast immer nur um P_X (und nicht mehr um P)

- In vielen Fällen interessiert man sich nur für die Wahrscheinlichkeitsverteilung P_X der Zufallsvariable und nicht für die zugrunde liegende Wahrscheinlichkeitsverteilung P.
- Man gibt P_X dann einfach direkt an (d.h. man *trifft eine Annahme* über die Verteilung der Zufallsvariable X), ohne wie im Beispiel auf der letzten Folie zuerst P anzugeben und dann P_X aus P zu berechnen.
- Die zugrunde liegende Wahrscheinlichkeitsverteilung P mit Ereignisraum $\mathcal{A}$ und Ergebnisraum Ω taucht dann explizit in der Praxis meistens gar nicht mehr auf.

- Vereinfachte Notation:
    - statt $P_X(\{x_j\})$ schreibt man häufig einfach $P(X = x_j)$
    - $P(X = x_j)$ steht dann für „die Wahrscheinlichkeit, dass sich die ZV X in einem Wert gleich x_j realisiert.
    - $P(X \leq c)$ steht für die Wahrscheinlichkeit, dass sich die ZV X in einem Wert kleiner oder gleich c realisiert.

diskrete und stetige Zufallsvariablen

- Wie bei „normalen“ Variablen (siehe VL 1) unterscheidet man zwischen diskreten und stetigen (oder kontinuierlichen) Zufallsvariablen:

- Von **diskreten Zufallsvariablen** sprechen wir, wenn der Träger eine **endliche Anzahl** an unterschiedlichen möglichen Realisationen enthält, d.h. $T_X = \{x_1, x_2, \dots, x_j, \dots, x_m\}$.

- Von **stetigen Zufallsvariablen** spricht man, wenn der Träger eine **unendliche Anzahl** an unterschiedlichen möglichen Realisationen (mit unendlich feinen Abstufungen) enthält, z.B. $T_X = \mathbb{R}$ oder $T_X = [0, 1]$.

<u>Bemerkung:</u> Streng genommen, können in der Wahrscheinlichkeitstheorie diskrete Zufallsvariablen auch unendlich viele Ausprägungen haben, z.B. $T_X = \{0, 1, 2, 3, \dots\} = \mathbb{N}_0^+$. Für uns ist jedoch die Definition mit endlich vielen Ausprägungen ausreichend.

Überblick

- bislang:
  - Ausblick Inferenzstatistik
  - Mathematische Grundbegriffe
  - Zufallsexperimente
  - Wahrscheinlichkeitsverteilungen
  - Zufallsvariablen

- jetzt: Wahrscheinlichkeitsfunktionen und Wahrscheinlichkeitsdichtefunktionen

Wahrscheinlichkeitsfunktionen und Wahrscheinlichkeitsdichtefunktionen

Motivation

- Wenn wir eine konkrete Wahrscheinlichkeitsverteilung P_X einer Zufallsvariablen *X* vollständig angeben wollen, müssen wir für jedes Ereignis A_X aus $\mathcal{A}_X$ eine Wahrscheinlichkeit $P_X(A_X)$ angeben.

- Problem:
  - Für diskrete Zufallsvariablen ist dies in den meisten Fällen aufgrund der Größe von $\mathcal{A}_X$ extrem aufwendig.
  - Für stetige Zufallsvariablen ist es sogar unmöglich, da man z.B. nicht einmal alle Elementarereignisse explizit angeben bzw. aufschreiben kann.

- Lösung:
  - Es lässt sich für diskrete ZV zeigen, dass eine Funktion, die die Wahrscheinlichkeit der Elementarereignisse abbilden kann, völlig ausreichend ist.
  - Ein ähnliches Prinzip lässt sich auch für stetige ZV ableiten.

Wahrscheinlichkeitsfunktion

- Allgemein kann man für diskrete ZVs die Wahrscheinlichkeit für jedes beliebige Ereignis A_X aus der Summe der Wahrscheinlichkeiten der Elementarereignisse berechnen, aus denen sich dieses Ereignis A_X zusammensetzt:

$$P_X(A_X) = \sum_{x_j \epsilon A_X} P_X(\{x_j\}) = \sum_{x_j \epsilon A_X} P(X = x_j)$$

Man summiert über alle Elemente x_j der Menge A_X

- Um die Wahrscheinlichkeitsverteilung einer diskreten Zufallsvariable anzugeben, reicht es also, eine Funktion anzugeben, die jedem Element x_j des Trägers T_X die Wahrscheinlichkeit $P(X = x_j)$ des entsprechenden Elementarereignisses zuweist.

- Eine solche Funktion wird **Wahrscheinlichkeitsfunktion** f genannt.

Eigenschaften der Wahrscheinlichkeitsfunktion

- Die Wahrscheinlichkeit für ein Ereignis A_X lässt sich additiv aus der Wahrscheinlichkeitsfunktion zusammensetzen:

$$P_X(A_X) = \sum_{x_j \epsilon A_X} f(x_j)$$

- Für alle Ausprägungen x_j nimmt die Wahrscheinlichkeitsfunktion Werte ≥ 0 an.

$f(x_j) \geq 0$ für alle $x_j \in T_X$

- Die Summe der Wahrscheinlichkeitsfunktion über **alle** Elemente x_j des Trägers ist immer $= 1$.

$$\sum_{x_j \epsilon T_X} f(x_j) = \sum_{j=1}^{m} f(x_j) = 1$$

Wahrscheinlichkeitsfunktion - Beispiel

- Beispiel: Zweimaliges Werfen einer Münze. Die ZV *X* gibt die Anzahl Kopf an.

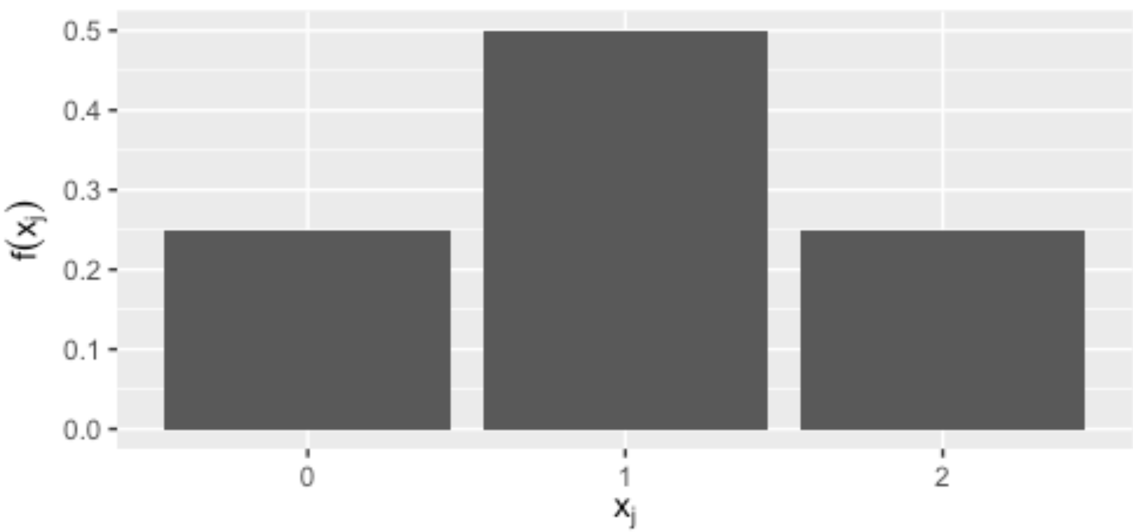
1. Wurf	2. Wurf	P	$X(\omega)$	P_X	$f(x_j)$
Zahl	Zahl	$P(\{ZZ\}) = {}^1/_4$	0	$P_X(\{0\}) = {}^1/_4$	$f(0) = 0.25$
Zahl	Kopf	$P(\{ZK\}) = {}^1/_4$	1	$P_X(\{1\}) = {}^2/_4$	$f(1) = 0.5$
Kopf	Zahl	$P(\{KZ\}) = {}^1/_4$	1		
Kopf	Kopf	$P(\{KK\}) = {}^1/_4$	2	$P_X(\{2\}) = {}^1/_4$	$f(2) = 0.25$

- $\Omega = \{ZZ, ZK, KZ, KK\}$
- $T_X = \{0, 1, 2\}$

Wahrscheinlichkeitsfunktion - Beispiel

- Tabellarische und graphische Darstellung dieser Wahrscheinlichkeitsfunktion:

x_j	$f(x_j)$
0	0,25
1	0,50
2	0,25



Wahrscheinlichkeitsfunktion - Beispiel

Beispiele für die Berechnung der Wahrscheinlichkeit von beliebigen Ereignissen mithilfe der Wahrscheinlichkeitsfunktion und der Formel

$$P_X(A_X) = \sum_{x_j \epsilon A_X} f(x_j)$$

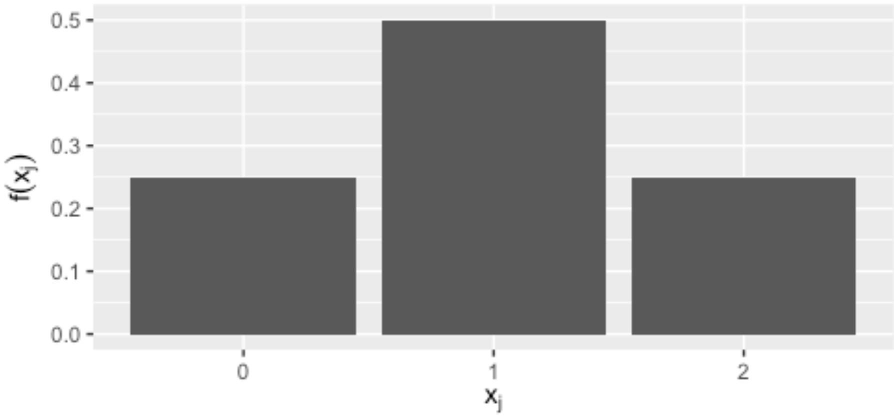
- Wahrscheinlichkeit des Ereignisses $A_X = \{1, 2\}$, dass mindestens einmal Kopf auftritt:

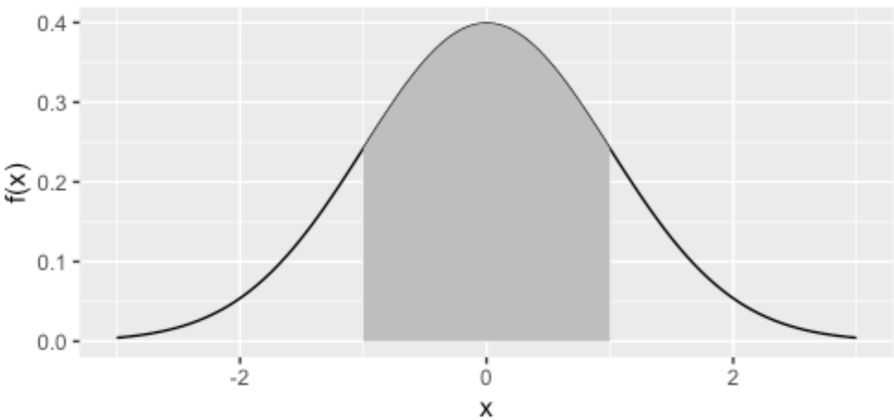
$$P_X(\{1, 2\}) = \sum_{x_j \epsilon \{1,2\}} f(x_j) = f(1) + f(2) = 0.5 + 0.25 = 0.75$$

- Wahrscheinlichkeit des Ereignisses $A_X = \{0, 2\}$, dass kein- oder zweimal Kopf auftritt:

$$P_X(\{0, 2\}) = \sum_{x_j \epsilon \{0,2\}} f(x_j) = f(0) + f(2) = 0.25 + 0.25 = 0.5$$

Diskrete ZV vs. stetige ZV





- Für eine diskrete ZV ist auch die Wahrscheinlichkeitsfunktion $f(x_j)$ diskret.
- Entsprechend kann die Wahrscheinlichkeit für jedes Ereignis A_X aus der Summe der $f(x_j)$ der $x_j \in A_X$ gebildet werden.
- Für eine stetige ZV wäre eine ähnliche Funktion entsprechend auch stetig.
- Statt einer Summe müssten wir dann die Wahrscheinlichkeit für ein Ereignis A_X aus der Fläche unter der stetigen Kurve $f(x)$ im entsprechenden Intervall berechnen.

Wahrscheinlichkeitsdichtefunktion I

- Für stetige Zufallsvariablen hilft uns die Idee der Wahrscheinlichkeitsfunktion nicht weiter.
- Grund: Ereignisse sind hier sehr häufig Intervalle. Intervalle können nicht als (endliche) Vereinigungsmengen von Elementarereignissen dargestellt werden.
- Die Berechnung der Wahrscheinlichkeit eines Ereignisses durch die Addition der Wahrscheinlichkeiten der Elementarereignisse klappt deshalb nicht mehr.
- Deshalb: Angabe einer Funktion, deren Integral für jedes Ereignis die Wahrscheinlichkeit dieses Ereignisses angibt:

$$P_X(A_X) = \int_{A_X} f(x)dx$$

- Eine solche Funktion wird **Wahrscheinlichkeitsdichtefunktion** f genannt.

Wahrscheinlichkeitsdichtefunktion II

- Im Gegensatz zur Wahrscheinlichkeitsfunktion entspricht der Funktionswert $f(x)$ einer Wahrscheinlichkeitsdichtefunktion **nicht** der Wahrscheinlichkeit des Elementarereignisses $\{x\}$.
- Der konkrete Wert $f(x)$ wird Wahrscheinlichkeitsdichte der Realisation x genannt. Er kann auch Werte größer 1 annehmen und besitzt keine intuitive Interpretation.
- Die Wahrscheinlichkeitsdichtefunktion wird oft auch einfach Dichtefunktion oder Dichte genannt.

Eigenschaften der Wahrscheinlichkeitsdichtefunktion

- Die Wahrscheinlichkeit eines Ereignisses A_X wird aus der Fläche unter der Wahrscheinlichkeitsdichte $f(x)$ in dem Intervall über x berechnet, aus dem sich das Ereignis A_X zusammensetzt.

$$P_X(A_X) = \int_{A_X} f(x)dx$$

- Für alle Ausprägungen x nimmt die Wahrscheinlichkeitsdichtefunktion Werte ≥ 0 an.

$$f(x) \geq 0$$

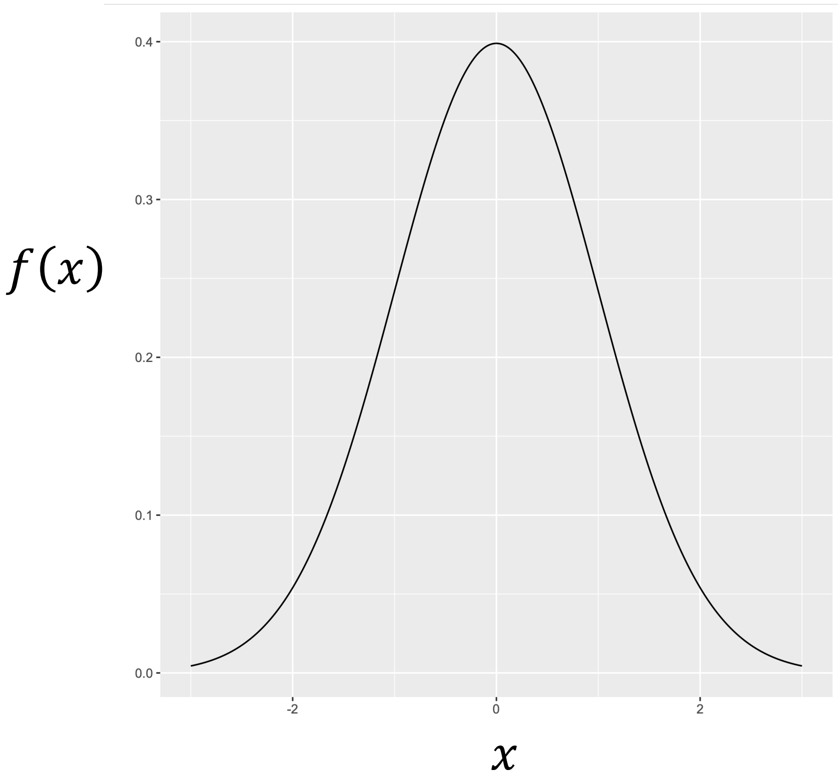
- Die Fläche unter der Wahrscheinlichkeitsdichtefunktion $f(x)$ für den gesamten Wertebereich von x (also im Bereich des Trägers T_X) ist immer 1.

$$\int_{T_X} f(x)dx = 1$$

Beispiel I

- Beispiel: ZV mit $T_X = \mathbb{R}$
- Beispiel für eine Wahrscheinlichkeitsdichtefunktion:

$$f(x) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right)$$



Beispiel II

- Beispiele für die Berechnung der Wahrscheinlichkeit von beliebigen Ereignissen mithilfe der Wahrscheinlichkeitsdichtefunktion und der Formel

$$P_X(A_X) = \int_{A_X} f(x)dx$$

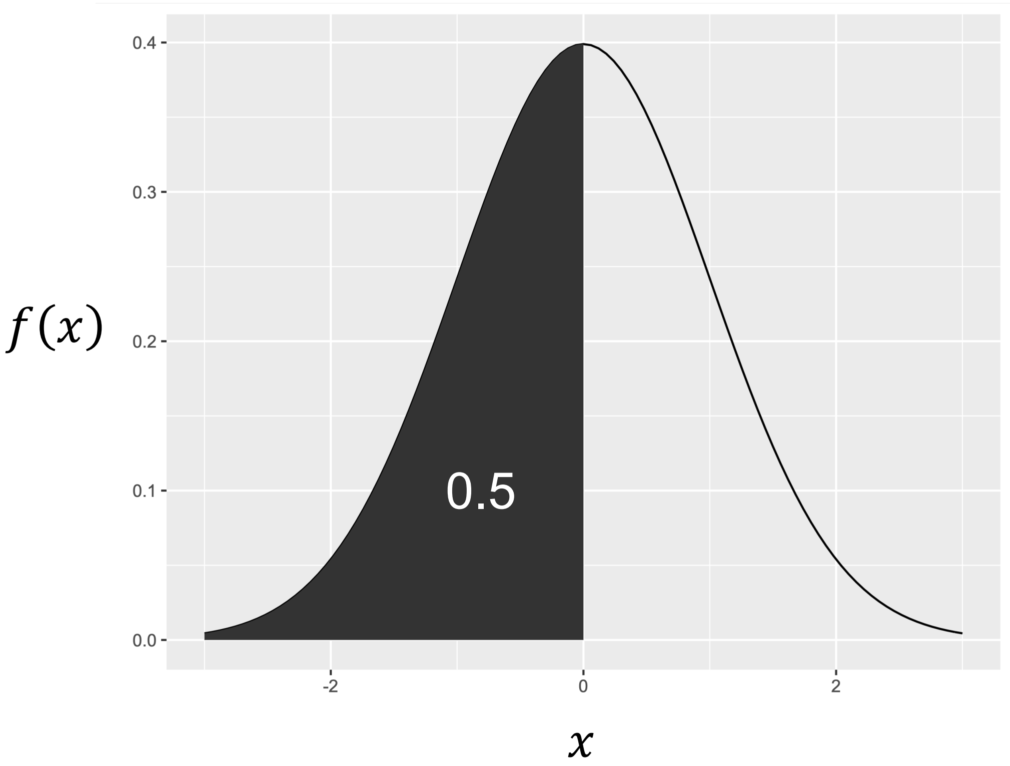
also durch Berechnung der **Fläche unter der Wahrscheinlichkeitsdichtefunktion im Bereich $\boldsymbol{A_X}$.**

- Keine Sorge: Sie müssen in Statistik I und II keine Intervalle „händisch berechnen“.

Beispiel III

- Wahrscheinlichkeit des Ereignisses $A_X =]-\infty, 0]$, dass sich die ZV in einem Wert kleiner oder gleich 0 realisiert:

$$P_X(]-\infty, 0]) = P(X \leq 0) = \int_{-\infty}^{0} f(x)dx = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right) dx = \cdots = 0.5$$

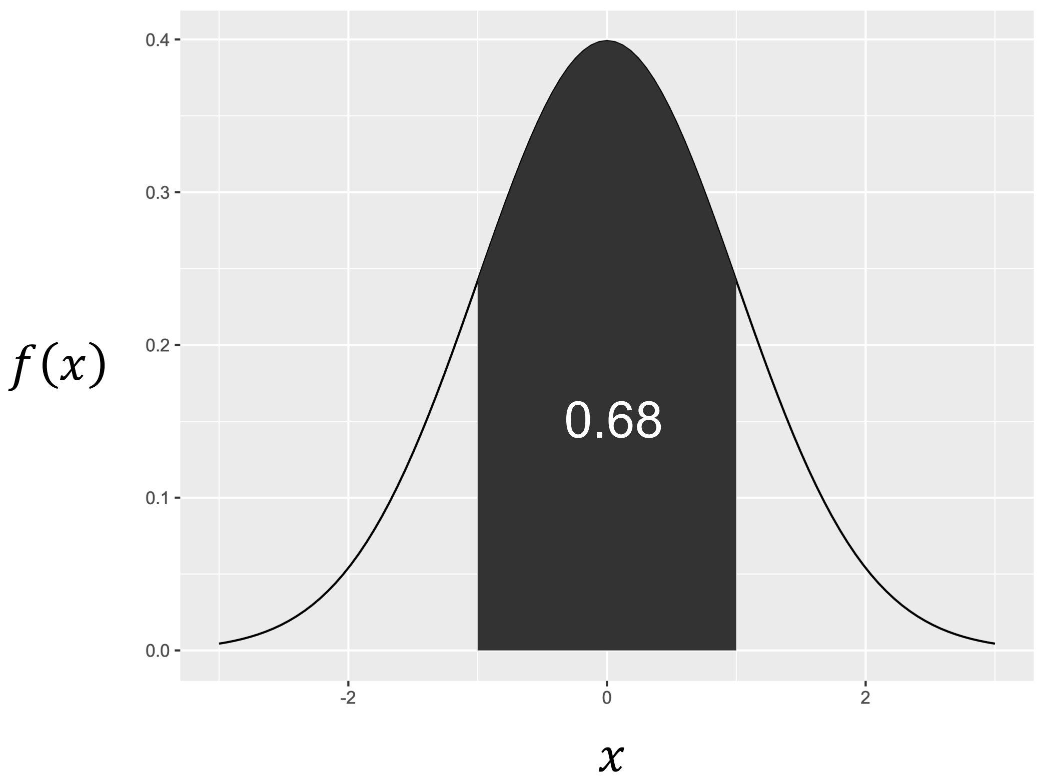


Hinweis:
Bei stetigen Zufallsvariablen gilt
$P(X \leq x) = P(X < x)$
bzw.
$P(X \geq x) = P(X > x)$

Beispiel IV

- Wahrscheinlichkeit des Ereignisses $A_X = [-1, 1]$, dass sich die ZV in einem Wert zwischen -1 und 1 realisiert:

$$P_X([-1,1]) = P(-1 \leq X \leq 1) = \int_{-1}^{1} f(x)dx = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right) dx = \cdots \approx 0.68$$



Überblick

- bislang:
    - Ausblick Inferenzstatistik
    - Mathematische Grundbegriffe
    - Zufallsexperimente
    - Wahrscheinlichkeitsverteilungen
    - Zufallsvariablen
    - Wahrscheinlichkeitsfunktionen und Wahrscheinlichkeitsdichtefunktionen

- jetzt: Verteilungsfunktionen

Verteilungsfunktionen

Verteilungsfunktion

- Eine weitere wichtige Funktion zur Charakterisierung der Wahrscheinlichkeitsverteilung P_X einer Zufallsvariable ist die **Verteilungsfunktion** *F* der Zufallsvariable.

- Sie ist sowohl für diskrete als auch für stetige Zufallsvariablen definiert.

- Sie weist jedem Element *x* des Trägers T_X der Zufallsvariablen die Wahrscheinlichkeit zu, dass sich die Zufallsvariable in diesem Wert *x* oder einem kleineren Wert realisiert.

- Die Verteilungsfunktion einer stetigen Variable gibt also die Fläche unter der Wahrscheinlichkeitsdichtefunktion im Intervall vom kleinsten Wert des Trägers bis zu einem beliebigen Endpunkt an.

- Liegt einem die Verteilungsfunktion vor, können Wahrscheinlichkeiten für Intervalle sehr viel einfacher berechnet werden, als jedes Mal ein Integral über eine beliebig komplexe Funktion $f(x)$ zu berechnen.

Verteilungsfunktion und Wahrscheinlichkeitsfunktion

- Im Fall einer diskreten Zufallsvariable *X* besteht folgender Zusammenhang zwischen der Verteilungsfunktion *F* und der Wahrscheinlichkeitsfunktion *f*:

$$F(x_k) = \sum_{j=1}^{k} f\left(x_j\right)$$

  wobei die x_1, x_2, …, x_m die aufsteigend der Größe nach geordneten Elemente des Trägers sind.

- Dieser Zusammenhang ergibt sich direkt aus der Definition der Verteilungsfunktion und den Eigenschaften der Wahrscheinlichkeitsfunktion.

- Bemerkung: Der Zusammenhang zwischen *f* und *F* bei diskreten Zufallsvariablen ist sehr ähnlich wie der Zusammenhang zwischen relativen Häufigkeiten und kumulierten relativen Häufigkeiten in der Deskriptivstatistik.

Beispiel: Verteilungsfunktion einer diskreten Zufallsvariable

- Beispiel: Zweimaliges Werfen einer Münze. Die ZV *X* gibt die Anzahl Kopf an.

- $T_X = \{0, 1, 2\}$ mit $x_1 = 0$, $x_2 = 1$ und $x_3 = 2$

- Wahrscheinlichkeitsfunktion von *X*:

$$f(x_1) = f(0) = 0.25$$

$$f(x_2) = f(1) = 0.50$$

$$f(x_3) = f(2) = 0.25$$

- Verteilungsfunktion von *X*:

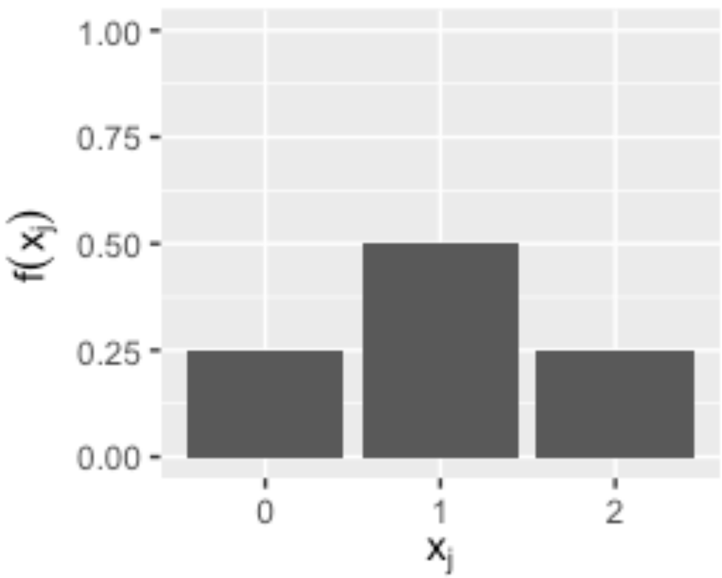
$$F(x_1) = \sum_{j=1}^{1} f(x_j) = f(x_1) = f(0) = 0.25$$

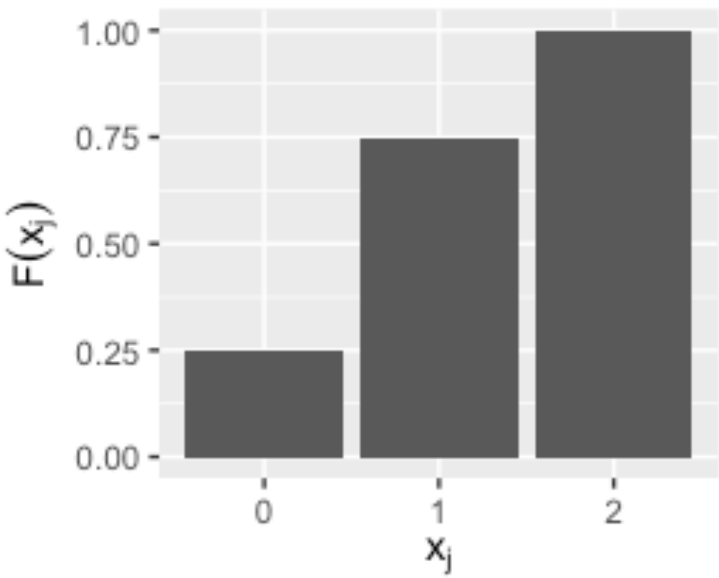
$$F(x_2) = \sum_{j=1}^{2} f(x_j) = f(x_1) + f(x_2) = f(0) + f(1) = 0.25 + 0.50 = 0.75$$

$$F(x_3) = \sum_{j=1}^{3} f(x_j) = f(x_1) + f(x_2) + f(x_3) = f(0) + f(1) + f(2) = 0.25 + 0.50 + 0.25 = 1$$

Beispiel: Verteilungsfunktion einer diskreten Zufallsvariable

x_j	$f(x_j)$	$F(x_j)$
0	0.25	0.25
1	0.50	0.75
2	0.25	1





Verteilungsfunktion und Wahrscheinlichkeitsdichtefunktion

- Im Fall einer stetigen Zufallsvariable X besteht folgender Zusammenhang zwischen der Verteilungsfunktion F und der Wahrscheinlichkeitsdichtefunktion f:

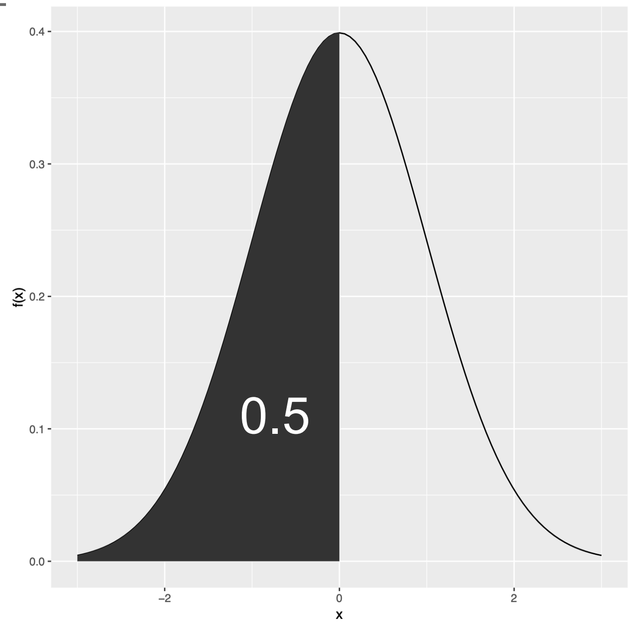
$$F(x) = \int_{-\infty}^{x} f(y)dy$$

wobei x und y Elemente des Trägers sind.

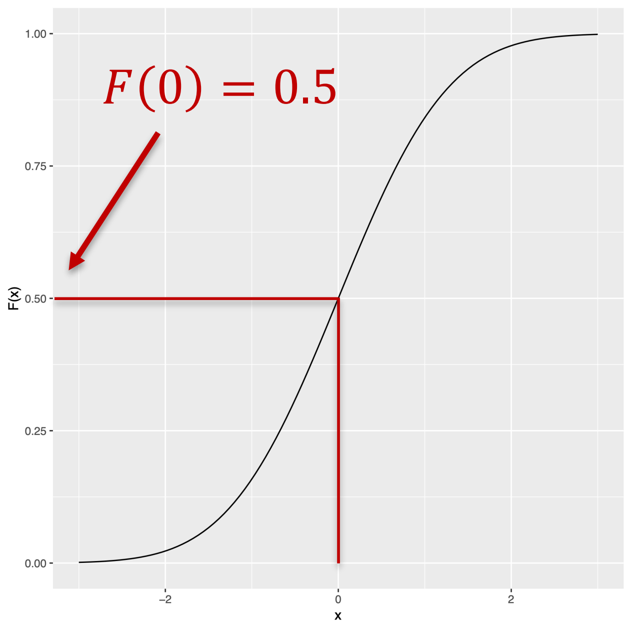
- Dieser Zusammenhang ergibt sich direkt aus der Definition der Verteilungsfunktion und den Eigenschaften der Wahrscheinlichkeitsdichtefunktion.

Beispiel: Verteilungsfunktion einer stetigen Zufallsvariable

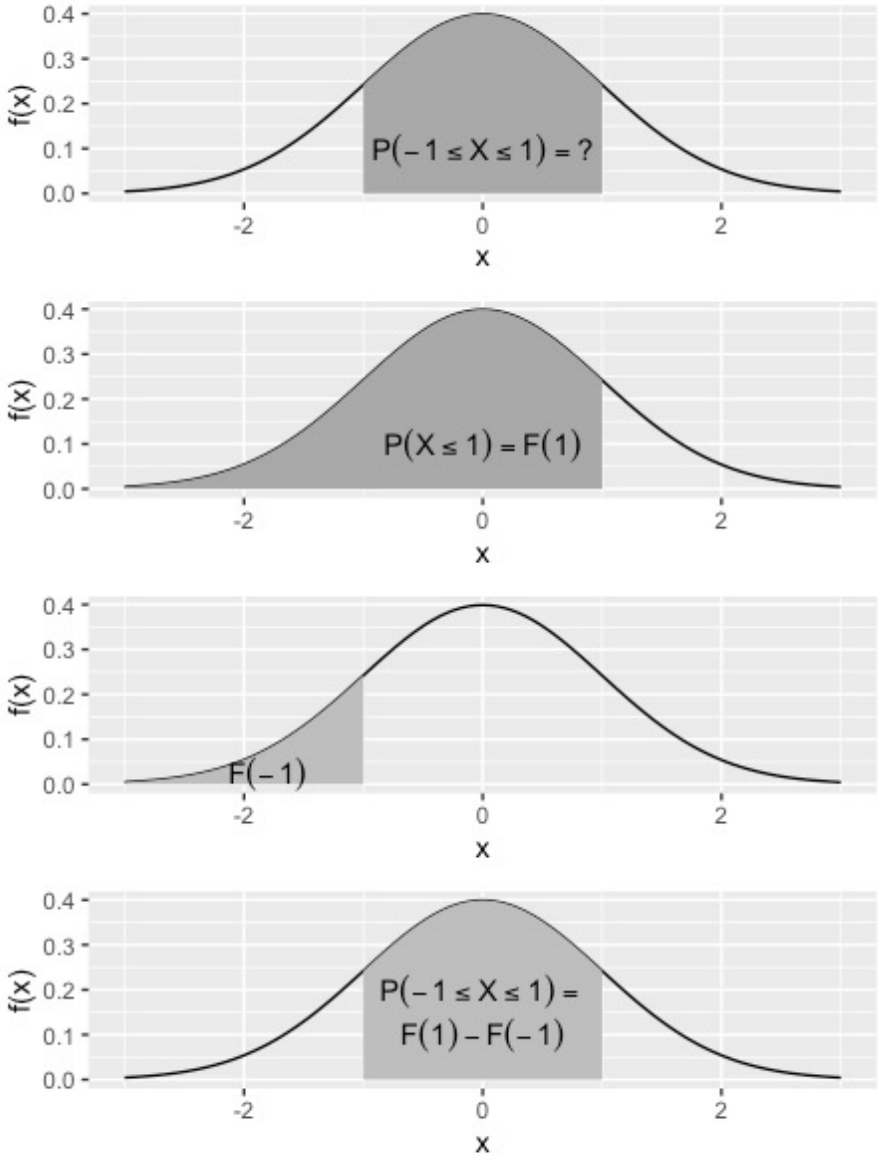
$$f(x) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right)$$



$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{y^2}{2}\right) dy$$



Beispiel: Verteilungsfunktion einer stetigen Zufallsvariable



- Mit Hilfe der Verteilungsfunktion lässt sich jedes beliebige Intervall problemlos als Kombination anderer Intervalle darstellen:
- Zum Beispiel das Intervall
$$P(-1 \leq X \leq 1) =$$
$$= P(X \leq 1) - P(X \leq -1)$$
$$= F(1) - F(-1)$$
- Falls man die Verteilungsfunktion kennt, ist also die Berechnung vieler Wahrscheinlichkeiten sehr einfach!

Zusammenfassung

- Zufallsvariablen bilden Eigenschaften eines Zufallsexperiments ab, die für beliebige Fragestellungen von Interesse sind.
- Die Wahrscheinlichkeiten für Ereignisse der Zufallsvariablen können dann direkt aus den Wahrscheinlichkeiten für das Zufallsexperiment abgeleitet werden.
- Zufallsvariablen können diskret oder stetig sein.
- Die Wahrscheinlichkeitsfunktion gibt bei diskreten ZV die Wahrscheinlichkeiten der Elementarereignisse an. Daraus können dann die Wahrscheinlichkeiten beliebiger Ereignisse additiv gebildet werden.
- Die Wahrscheinlichkeitsdichtefunktion gibt bei stetigen ZV die Wahrscheinlichkeitsdichte an, die selbst erstmal keine intuitive Interpretation besitzt. Die Wahrscheinlichkeiten für beliebige Ereignisse werden aber aus der entsprechenden Fläche unter der Dichtefunktion errechnet.
- Die Verteilungsfunktion verhält sich zur Wahrscheinlichkeits(dichte)funktion wie die kumulierte Häufigkeit zur Häufigkeit. Gerade bei stetigen ZV ist die Verteilungsfunktion praktisch um Wahrscheinlichkeiten zu errechnen, ohne Integrale der Dichtefunktion lösen zu müssen.