# Wahrscheinlichkeitsrechnung und mathematische Statistik für Informatiker

Vorlesung im Wintersemester 2018/19
an der TU Dortmund

Jun.-Prof. Dr. Andreas Groll

WiSe 18/19, Fakultät Statistik, TU Dortmund

# Einleitung

# 1.1 WRUMS für Informatiker

Jun.-Prof. Dr. Andreas Groll
Mathegebäude, Raum 216a
E-mail: groll@statistik.tu-dortmund.de

- Vorlesung
  - Termin: Do 08:30 - 10:00
  - Hörsaal: HG II - HS3

# 1.1 WRUMS für Informatiker

Jun.-Prof. Dr. Andreas Groll
Mathegebäude, Raum 216a
E-mail: groll@statistik.tu-dortmund.de

- Vorlesung
  - Termin: Do 08:30 - 10:00
  - Hörsaal: HG II - HS3

Hendrik van der Wurp
Mathegebäude, Raum 226
E-mail: vanderwurp@statistik.tu-dortmund.de

- Übung (2 Gruppen)
  - Termin: Do 18:00 - 19:30 (i.d.R. alle 2 Wochen)
  - Gruppe 1: ab 18.10.2018
  - Gruppe 2: ab 25.10.2018
  - Raum: EF50 - HS1

# 1.1 WRUMS für Informatiker

Jun.-Prof. Dr. Andreas Groll
Mathegebäude, Raum 216a
E-mail: groll@statistik.tu-dortmund.de

- Vorlesung
  - Termin: Do 08:30 - 10:00
  - Hörsaal: HG II - HS3

Hendrik van der Wurp
Mathegebäude, Raum 226
E-mail: vanderwurp@statistik.tu-dortmund.de

- Übung (2 Gruppen)
  - Termin: Do 18:00 - 19:30 (i.d.R. alle 2 Wochen)
  - Gruppe 1: ab 18.10.2018
  - Gruppe 2: ab 25.10.2018
  - Raum: EF50 - HS1

- Klausurtermine
  - Klausur: Mo (18.02.2019) 16-18 Uhr, Räume: Audimax & EF50 - HS3 & Mathe E28 & Mathe E29
  - Nachklausur: Fr (29.03.2019) 8-10 Uhr, Räume: Audimax & Mathe E28 & Mathe E29

# 1.1 WRUMS für Informatiker

Jun.-Prof. Dr. Andreas Groll
Mathegebäude, Raum 216a
E-mail: groll@statistik.tu-dortmund.de

- Vorlesung
  - Termin: Do 08:30 - 10:00
  - Hörsaal: HG II - HS3

Hendrik van der Wurp
Mathegebäude, Raum 226
E-mail: vanderwurp@statistik.tu-dortmund.de

- Übung (2 Gruppen)
  - Termin: Do 18:00 - 19:30 (i.d.R. alle 2 Wochen)
  - Gruppe 1: ab 18.10.2018
  - Gruppe 2: ab 25.10.2018
  - Raum: EF50 - HS1

- Klausurtermine
  - Klausur: Mo (18.02.2019) 16-18 Uhr, Räume: Audimax & EF50 - HS3 & Mathe E28 & Mathe E29
  - Nachklausur: Fr (29.03.2019) 8-10 Uhr, Räume: Audimax & Mathe E28 & Mathe E29
- Voraussetzung für Klausur
  Regelmäßige Teilnahme Ü, selbstständige Bearbeitung der Übungsaufgaben

# 1.1 WRUMS für Informatiker

Jun.-Prof. Dr. Andreas Groll
Mathegebäude, Raum 216a
E-mail: groll@statistik.tu-dortmund.de

- Vorlesung
  - Termin: Do 08:30 - 10:00
  - Hörsaal: HG II - HS3

Hendrik van der Wurp
Mathegebäude, Raum 226
E-mail: vanderwurp@statistik.tu-dortmund.de

- Übung (2 Gruppen)
  - Termin: Do 18:00 - 19:30 (i.d.R. alle 2 Wochen)
  - Gruppe 1: ab 18.10.2018
  - Gruppe 2: ab 25.10.2018
  - Raum: EF50 - HS1

- Klausurtermine
  - Klausur: Mo (18.02.2019) 16-18 Uhr, Räume: Audimax & EF50 - HS3 & Mathe E28 & Mathe E29
  - Nachklausur: Fr (29.03.2019) 8-10 Uhr, Räume: Audimax & Mathe E28 & Mathe E29
- Voraussetzung für Klausur
  Regelmäßige Teilnahme Ü, selbstständige Bearbeitung der Übungsaufgaben
- Moodle
  **Website**: `https://moodle.tu-dortmund.de/course/view.php?id=13183`
  **Passwort**: wrums1819

# 1.2 Übersicht

- Motivation
- Merkmale und Datentypen
- Univariate Daten
  - Tabellarische und grafische Darstellung
- Statistische Kennzahlen
  - für die Lage
  - für die Streuung
- Bivariate Daten
  - Tabellarische und grafische Darstellung
  - Zusammenhangsmaße
  - Lineare Regression
- Mengentheoretische Grundlagen
- Wahrscheinlichkeitsmaße und -räume
- Zufallsvariablen und deren Verteilungen
- Wichtige Wahrscheinlichkeitsverteilungen
- Bedingte Wahrscheinlichkeiten und stochastische Unabhängigkeit
- Erwartungswert und Varianz
- Weitere wahrscheinlichkeitstheoretische Kennzahlen
- Markoffketten
- Statistische Tests
  - Normalverteilung
  - Test bei nicht-normalverteilten Daten

# 1.3 Motivation

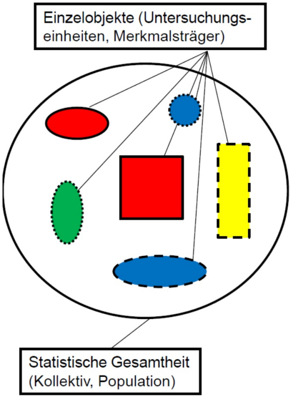
- Statistische Methoden spielen in der Informatik an vielen Stellen eine große Rolle

- Beispiele:
    - Laufzeiten von Algorithmen mit stochastischem Input
    - Stochastische Algorithmen
    - Spieltheorie
    - Ausfälle von Datenverbindungen oder Hardwarekomponenten
    - Automatische Übersetzung
    - Assoziationsregeln, Bilderkennung, Signalanalyse
    - Statistische Lernverfahren

- Diese Vorlesung behandelt nicht alle diese Themen, sondern die dazu notwendigen statistischen Grundlagen.

# Univariate Daten

# 2.1 Merkmale und Datentypen




| Merkmal | Merkmals-ausprägungen | Wertebereich |
|---|---|---|
| Form | Ellipse, Ellipse, Ellipse, Rechteck, Rechteck, Ellipse | {Ellipse, Rechteck} |
| Farbe | Rot, Blau, Grün, Rot, Gelb, Blau | {Blau, Gelb, Grün, Rot} |
| Linienart | Durchgängig, Gepunktet, Gepunktet, Durchgängig, Gestrichelt, Gestrichelt | {Gepunktet, Gestrichelt, Durchgängig} |
| Breite in cm | 2, 1, 1, 2, 1, 3 | $(0, \infty)$ |
| Höhe in cm | 1, 1, 2, 2, 3, 1 | $(0, \infty)$ |

# 2.1 Merkmale und Datentypen

## Datentypen

| Skalentyp | mögliche Aussagen | Im Beispiel |
|---|---|---|
| **qualitativ** | | |
| Nominal | Gleich / Verschieden | Farbe, Form (binär, dichotom) |
| Ordinal | Größer / Kleiner | Linienart |
| **quantitativ / metrisch** | | |
| Intervall | Differenzen gleich / verschieden | (Breite, Höhe) |
| Verhältnis | Verhältnisse gleich / verschieden | Breite, Höhe |

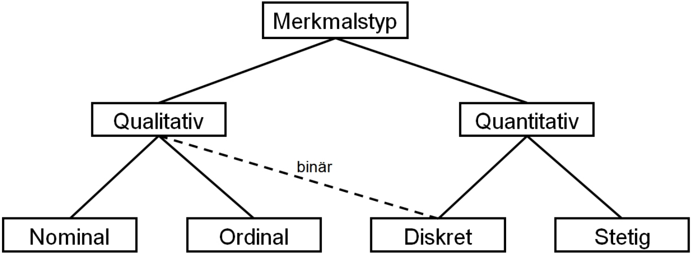
# 2.1 Merkmale und Datentypen

## Datentypen

| **Datentyp** | **Anzahl der Ausprägungen** | **Im Beispiel** |
|---|---|---|
| Diskret | Endlich *oder* abzählbar unendlich viele | Form<br>Breite, Höhe<br>(wenn grob bemessen) |
| Stetig | Überabzählbar viele | Breite, Höhe<br>(wenn beliebig fein bemessen) |

# 2.1 Merkmale und Datentypen

## Datentypen




- Qualitativ heißt immer diskret
- Skalenniveau wird von links nach rechts immer höher

# 2.1 Merkmale und Datentypen

- Unter Inkaufnahme von Informationsverlust können Merkmale in andere Skalenniveaus überführt und entsprechend analysiert werden
  - stetig in diskret (runden, genaue Werte gehen verloren)
  - diskret quantitativ in ordinal (Abstände gehen verloren)
  - ordinal in nominal (Ordnung geht verloren)
- Dieses Vorgehen kann generell auch sinnvoll sein (z.B. bei Linearitätsverletzung)

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

## Qualitative Daten

| | |
|---|---|
| $M_N = \{e_1, \ldots, e_N\}$ | Population bestehend aus Objekten $e_1, \ldots, e_N$ |
| $X$ | Nominales bzw. ordinales Merkmal |
| $x \in W_X$ | Merkmalsausprägungen von $X$ |
| $W_X = \{x(j) \mid j = 1, \ldots, J\}$ <br> $= \{x(1), \ldots, x(J)\}$ | Wertebereich von $X$ mit <br> Merkmalsausprägungen $x(j)$, $j = 1, \ldots, J$ |
| $D_N = \{x_n \mid n = 1, \ldots, N\}$ <br> $= \{x_1, \ldots, x_N\}$ | Urliste aus der Messung von $X$ in der <br> Population $M_N$, d.h. $x_n = X(e_n)$, $n = 1, \ldots, N$ |
| $x(1) < x(2) < \ldots < x(J)$ | falls $X$ ordinal |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

Objekte

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

5 Variablen

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

Qualitative Daten

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

$D_{N;1}$, N=12

$X_1$ = Bearbeiter(in)

$W_{X1}$ = {Kai, Miriam, Oliver, Tina}

$J_1 = 4$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

$D_{N;2}$, N=12

$X_2$ = Aufgabe

$W_{X2}$ = {Abfrage, Export, Verknüpfung}

$J_2 = 3$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

$D_{N;3}$, N=12

$X_3$ = Version

$W_{X3} = \{1.1, 1.2, 2.0\}$, $1.1 < 1.2 < 2.0$

$J_3 = 3$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Deskriptive Auswertung

**Absolute Häufigkeit** $N_j$ von $x(j)$: $N_j = N[x(j)] = \sum_{i=1}^{N} d_i(j)$, mit $d_i(j) := I_{x(e_i)=x(j)}$

Damit gilt $\sum_{j=1}^{J} N_j = N$

| $x_1(1)$ Kai | $x_1(2)$ Miriam | $x_1(3)$ Oliver | $x_1(4)$ Tina | Σ |
|---|---|---|---|---|
| II | III | IIII | III | ~~IIII~~ ~~IIII~~ II |
| 2 | 3 | 4 | 3 | 12 |

| i | $X_1(e_i)$ | $d_{1i}(1)$ | $d_{1i}(2)$ | $d_{1i}(3)$ | $d_{1i}(4)$ |
|---|---|---|---|---|---|
| 1 | Kai | 1 | 0 | 0 | 0 |
| 2 | Kai | 1 | 0 | 0 | 0 |
| 3 | Miriam | 0 | 1 | 0 | 0 |
| 4 | Tina | 0 | 0 | 0 | 1 |
| 5 | Oliver | 0 | 0 | 1 | 0 |
| 6 | Tina | 0 | 0 | 0 | 1 |
| 7 | Tina | 0 | 0 | 0 | 1 |
| 8 | Miriam | 0 | 1 | 0 | 0 |
| 9 | Miriam | 0 | 1 | 0 | 0 |
| 10 | Oliver | 0 | 0 | 1 | 0 |
| 11 | Oliver | 0 | 0 | 1 | 0 |
| 12 | Oliver | 0 | 0 | 1 | 0 |
| Σ | | 2 | 3 | 4 | 3 |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Deskriptive Auswertung

**Relative Häufigkeit** $f_j$ von $x(j)$: $f_j = \frac{N_j}{N}$

Damit gilt $\sum_{j=1}^{J} f_j = 1$

| $x_1(1)$ Kai | $x_1(2)$ Miriam | $x_1(3)$ Oliver | $x_1(4)$ Tina | Σ |
|---|---|---|---|---|
| | | | | |
| 2/12 ≈ 0.17 | 3/12 = 0.25 | 4/12 ≈ 0.33 | 3/12 = 0.25 | 12/12 = 1 |

| i | $X_1(e_i)$ | $d_{1i}(1)$ | $d_{1i}(2)$ | $d_{1i}(3)$ | $d_{1i}(4)$ |
|---|---|---|---|---|---|
| 1 | Kai | 1 | 0 | 0 | 0 |
| 2 | Kai | 1 | 0 | 0 | 0 |
| 3 | Miriam | 0 | 1 | 0 | 0 |
| 4 | Tina | 0 | 0 | 0 | 1 |
| 5 | Oliver | 0 | 0 | 1 | 0 |
| 6 | Tina | 0 | 0 | 0 | 1 |
| 7 | Tina | 0 | 0 | 0 | 1 |
| 8 | Miriam | 0 | 1 | 0 | 0 |
| 9 | Miriam | 0 | 1 | 0 | 0 |
| 10 | Oliver | 0 | 0 | 1 | 0 |
| 11 | Oliver | 0 | 0 | 1 | 0 |
| 12 | Oliver | 0 | 0 | 1 | 0 |
| Σ/12 | | 0.17 | 0.25 | 0.33 | 0.25 |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Deskriptive Auswertung

Tabellarische Darstellung absoluter und relativer Häufigkeiten

| **Ausprägung** | **Absolute Häufigkeit** | **Relative Häufigkeit** |
|---|---|---|
| $x(1)$ | $N_1$ | $f_1 = N_1/N$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $x(J)$ | $N_J$ | $f_J = N_J/N$ |
| | $\sum_{j=1}^{J} N_j = N$ | $\sum_{j=1}^{J} f_j = 1$ |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Qualitative Daten: Deskriptive Auswertung

Tabellarische Darstellung absoluter und relativer Häufigkeiten

| Bearbeiter(in) | | |
|---|---|---|
| **Ausprägung** | **Absolute Häufigkeit** | **Relative Häufigkeit** |
| Kai | 2 | 0.17 |
| Miriam | 3 | 0.25 |
| Oliver | 4 | 0.33 |
| Tina | 3 | 0.25 |
| | 12 | 1 |

| Aufgabe | | |
|---|---|---|
| **Ausprägung** | **Absolute Häufigkeit** | **Relative Häufigkeit** |
| Abfrage | 2 | 0.17 |
| Export | 6 | 0.5 |
| Verknüpfung | 4 | 0.33 |
| | 12 | 1 |

| Version | | |
|---|---|---|
| **Ausprägung** | **Absolute Häufigkeit** | **Relative Häufigkeit** |
| 1.1 | 3 | 0.25 |
| 1.2 | 6 | 0.5 |
| 2.0 | 3 | 0.25 |
| | 12 | 1 |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ diskrete Daten

| | |
|---|---|
| $M_N = \{e_1, \ldots, e_N\}$ | Population bestehend aus Objekten $e_1, \ldots, e_N$ |
| $X$ | Quantitatives Merkmal |
| $x \in W_X$ | Merkmalsausprägungen von $X$ |
| $W_X = \{x(j) \mid j = 1, \ldots, J\}$ $= \{x(1), \ldots, x(J)\}$ | Wertebereich von $X$ mit Merkmalsausprägungen $x(j)$, $j = 1, \ldots, J$ |
| $D_N = \{x_n \mid n = 1, \ldots, N\}$ $= \{x_1, \ldots, x_N\}$ | Urliste aus der Messung von $X$ in der Population $M_N$, d.h. $x_n = X(e_n)$, $n = 1, \ldots, N$ |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ diskrete Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

$D_{N;4}$, N=12

$X_4$ = Anzahl Clicks

$W_{X4} = \{0, 1, \ldots, 10, 11, 12, 13, 14, 15, 16, 17, 18, \ldots, \infty\}$

$J_4 = \infty$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ diskrete Daten: Deskriptive Auswertung

**Absolute Häufigkeit** $N_j$ und **relative Häufigkeit** $f_j$ analog zu qualitativen Daten

**Relative Summenhäufigkeit** $s_j = \sum_{k=1}^{j} f_k = \frac{\#\{x_n | x_n \leq x(j)\}}{N}$

| **Ausprägung** | **Absolute Häufigkeit** | **Relative Häufigkeit** | **Relative Summenhäufigkeit** |
|---|---|---|---|
| $x(1)$ | $N_1$ | $f_1 = N_1/N$ | $f_1$ |
| $x(2)$ | $N_2$ | $f_2 = N_2/N$ | $f_1 + f_2$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $x(J-1)$ | $N_{J-1}$ | $f_{J-1} = N_{J-1}/N$ | $f_1 + \ldots + f_{J-1}$ |
| $x(J)$ | $N_J$ | $f_J = N_J/N$ | $f_1 + \ldots + f_J = 1$ |
| | $\sum_{j=1}^{J} N_J = N$ | $\sum_{j=1}^{J} f_j = 1$ | |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ diskrete Daten: Deskriptive Auswertung

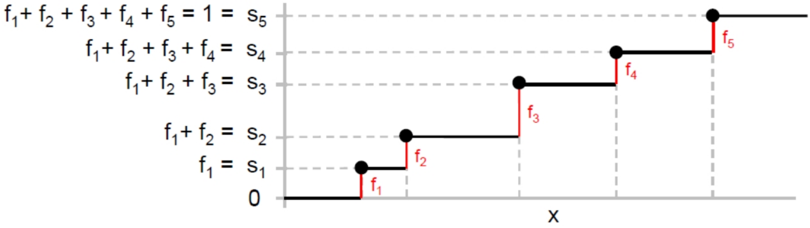
| Anzahl Clicks | | | |
|---|---|---|---|
| **Ausprägung** | **Absolute Häufigkeit** | **Relative Häufigkeit** | **Relative Summen-häufigkeit** |
| 0 - 9 | 0 | 0 | 0 |
| 10 | 2 | 0.167 | 0.167 |
| 11 | 1 | 0.083 | 0.25 |
| 12 | 2 | 0.167 | 0.417 |
| 13 | 1 | 0.083 | 0.5 |
| 14 | 2 | 0.167 | 0.667 |
| 15 | 1 | 0.083 | 0.75 |
| 16 | 1 | 0.083 | 0.833 |
| 17 | 1 | 0.083 | 0.917 |
| 18 | 1 | 0.083 | 1 |
| 19 - ∞ | 0 | 0 | 1 |
| | 12 | 1 | |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ diskrete Daten: Deskriptive Auswertung

Grafische Darstellung: **Empirische Verteilungsfunktion**

$$F_N(X) = \begin{cases} 0 & \text{, falls } x < x(1) \\ s_j = \sum_{k=1}^{j} f_k, \text{ mit } j = \max\{\tilde{j} | x(\tilde{j}) \leq x\} & \text{, falls } x(1) \leq x \end{cases}$$



# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ stetige Daten:

| | |
|---|---|
| $M_N = \{e_1, \ldots, e_N\}$ | Population bestehend aus Objekten $e_1, \ldots, e_N$ |
| $X$ | Quantitatives Merkmal |
| $x \in W_X$ | Merkmalsausprägungen von $X$ |
| $W_X = (-\infty, \infty) = \bigcup_{j=1}^{J} K_j$ | Klassierter (kategorisierter) Wertebereich von $X$ |
| $K_j = (v_{j-1}, v_j],\ j = 1, \ldots, J-1$<br>$K_J = (v_{J-1}, v_J)$ | Merkmalsklassen mit Klassengrenzen<br>$-\infty = v_0 < v_1 < \ldots < v_{J-1} < v_J = \infty$ |
| $D_N = \{x_n \mid n = 1, \ldots, N\}$<br>$= \{x_1, \ldots, x_N\}$ | Urliste aus der Messung von $X$ in der Population $M_N$, d.h. $x_n = X(e_n),\ n = 1, \ldots, N$ |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

## Quantitativ stetige Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

$D_{N;5}$, $N = 12$

$X_5$=Bearbeitungszeit

$$W_{X_5} = (-\infty, \infty) = (-\infty, 4] \bigcup (4, 5] \bigcup \ldots \bigcup (7, 8] \bigcup (8, \infty) = (-\infty, 4] \bigcup (\bigcup_{j=1}^{4} (j+3, j+4]) \bigcup (8, \infty)$$

$J_5 = 6$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ stetige Daten: Deskriptive Auswertung

**Klassierte Häufigkeitsverteilung**

| **Klasse $K_j$** | **Absolute Häufigkeit** | **Relative Häufigkeit** | **Relative Summenhäufigkeit** |
|---|---|---|---|
| $K_1 = (v_0, v_1]$ | $N(K_1)$ | $f(K_1) = N(K_1)/N$ | $f(K_1)$ |
| $K_2 = (v_1, v_2]$ | $N(K_2)$ | $f(K_2) = N(K_2)/N$ | $f(K_1) + f(K_2)$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $K_{J-1} = (v_{J-2}, v_{J-1}]$ | $N(K_{J-1})$ | $f(K_{J-1}) = N(K_{J-1})/N$ | $f(K_1) + \ldots + f(K_{J-1})$ |
| $K_J = (v_{J-1}, v_J)$ | $N(K_J)$ | $f(K_J) = N(K_J)/N$ | $f(K_1) + \ldots + f(K_J) = 1$ |
| | $\sum_{j=1}^{J} N(K_j) = N$ | $\sum_{j=1}^{J} f(K_j) = 1$ | |

$$N(K_j) = \#\{x \mid x \in K_j\} = \#\{x \mid v_{j-1} < x \leq v_j\}$$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

## Quantitativ stetige Daten: Deskriptive Auswertung

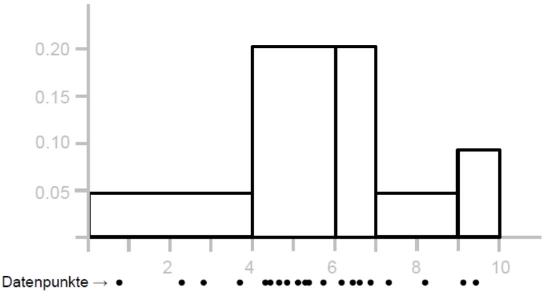
| Bearbeitungszeit | | | |
|---|---|---|---|
| **Klasse** | **Absolute Häufigkeit** | **Relative Häufigkeit** | **Relative Summen-häufigkeit** |
| $K_1 = (-\infty, 4]$ | 5 | 0.417 | 0.417 |
| $K_2 = (4, 5]$ | 3 | 0.250 | 0.667 |
| $K_3 = (5, 6]$ | 0 | 0.000 | 0.667 |
| $K_4 = (6, 7]$ | 2 | 0.167 | 0.833 |
| $K_5 = (7, 8]$ | 1 | 0.083 | 0.917 |
| $K_6 = (8, \infty)$ | 1 | 0.083 | 1 |
| | 12 | 1 | |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ stetige Daten: Deskriptive Auswertung

Grafische Darstellung: **Histogramm**

Aufbauend auf klassierter Häufigkeitsverteilung, allerdings $v_0 \neq -\infty$ und $v_J \neq \infty$



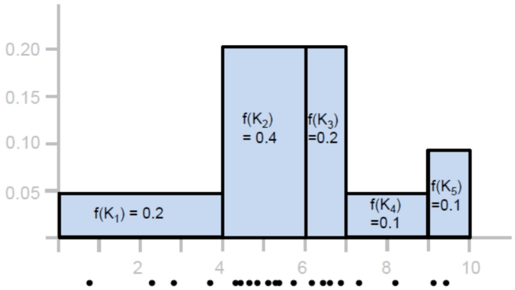

Flächen der Rechtecke zu $K_j$ entsprechen $f(K_j)$.
Rechteckbreiten sind gegeben durch $b_j = v_j - v_{j-1}$.
Damit ergeben sich als Rechteckhöhen $h_j = f(K_j)/b_j$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ stetige Daten: Deskriptive Auswertung

Grafische Darstellung: **Histogramm**

Aufbauend auf klassierter Häufigkeitsverteilung, allerdings $v_0 \neq -\infty$ und $v_J \neq \infty$



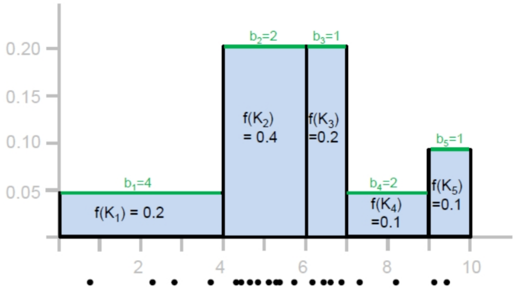

Flächen der Rechtecke zu $K_j$ entsprechen $f(K_j)$.
Rechteckbreiten sind gegeben durch $b_j = v_j - v_{j-1}$.
Damit ergeben sich als Rechteckhöhen $h_j = f(K_j)/b_j$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ stetige Daten: Deskriptive Auswertung

Grafische Darstellung: **Histogramm**

Aufbauend auf klassierter Häufigkeitsverteilung, allerdings $v_0 \neq -\infty$ und $v_J \neq \infty$



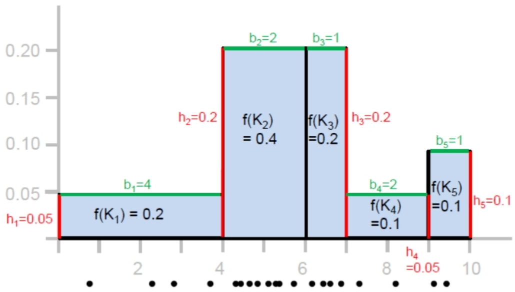

Flächen der Rechtecke zu $K_j$ entsprechen $f(K_j)$.
Rechteckbreiten sind gegeben durch $b_j = v_j - v_{j-1}$.
Damit ergeben sich als Rechteckhöhen $h_j = f(K_j)/b_j$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ stetige Daten: Deskriptive Auswertung

Grafische Darstellung: **Histogramm**

Aufbauend auf klassierter Häufigkeitsverteilung, allerdings $v_0 \neq -\infty$ und $v_J \neq \infty$



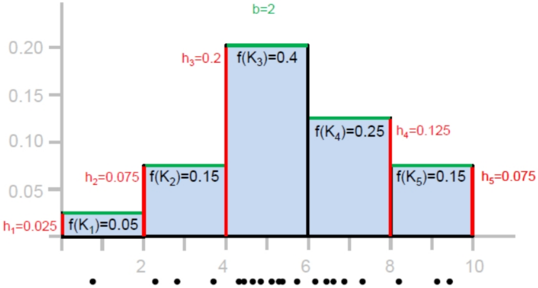

Flächen der Rechtecke zu $K_j$ entsprechen $f(K_j)$.
Rechteckbreiten sind gegeben durch $b_j = v_j - v_{j-1}$.
Damit ergeben sich als Rechteckhöhen $h_j = f(K_j)/b_j$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ stetige Daten: Deskriptive Auswertung

Grafische Darstellung: **Histogramm**

Üblicherweise gleiche Klassenbreiten




Flächen der Rechtecke zu $K_j$ entsprechen $f(K_j)$. Rechteckbreiten sind gegeben durch $b = v_j - v_{j-1}$. Damit ergeben sich als Rechteckhöhen $h_j = f(K_j)/b$

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Quantitativ stetige Daten: Deskriptive Auswertung

Grafische Darstellung: **Histogramm**

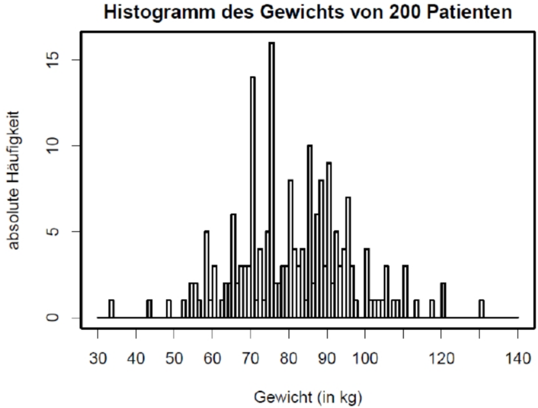
- Beispiel Patientendaten: Gewicht (in kg); NA: fehlender Wert (Not Available) $\rightarrow$ Zufällige Auswahl des Gewichts von 200 Patienten:

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 85 | 70 | 75 | 70 | 92 | 88 | 68 | 101 | 74 | 80 | 87 | 68 | 95 | 33 | 75 | 117 |
| 105 | 88 | 76 | 82 | 107 | 92 | 87 | 91 | 83 | 80 | 85 | 95 | 75 | 60 | 85 | 75 |
| 73 | 58 | 93 | 70 | 100 | 94 | 100 | 75 | 80 | 85 | 87 | 43 | 90 | 92 | 89 | NA |
| 100 | 96 | 58 | 72 | 77 | 83 | 48 | 74 | 90 | 58 | 78 | 75 | 56 | 70 | 75 | 70 |
| 67 | 95 | 74 | 88 | 70 | 68 | 66 | 102 | 72 | 74 | 113 | 72 | 81 | 75 | 55 | 60 |
| 75 | 90 | 71 | 93 | NA | 94 | 75 | 89 | 90 | 80 | 52 | 90 | 105 | 90 | 82 | 80 |
| 83 | 80 | 89 | 70 | 67 | 92 | 108 | 58 | 75 | 75 | 110 | 85 | 58 | 74 | 93 | 97 |
| 65 | 83 | 110 | 87 | 81 | 64 | 103 | 120 | 65 | 85 | 79 | 95 | 110 | 70 | 90 | 85 |
| 94 | 88 | 88 | 130 | 70 | 69 | 78 | 100 | 88 | 86 | 85 | 76 | 60 | 79 | 90 | 88 |
| 104 | 69 | 96 | 59 | 75 | NA | 75 | 66 | 70 | 86 | 80 | 65 | 94 | 72 | 62 | 75 |
| 105 | 91 | 79 | 88 | 80 | 85 | 69 | 87 | 54 | 96 | 70 | 82 | 70 | 95 | 78 | 95 |
| 95 | 84 | 70 | 90 | 65 | 67 | 85 | NA | 92 | 87 | 63 | 120 | 65 | 55 | 65 | 81 |
| NA | 54 | 81 | 63 | 64 | 77 | 70 | 75 | | | | | | | | |

# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Grafische Darstellung: **Histogramm**

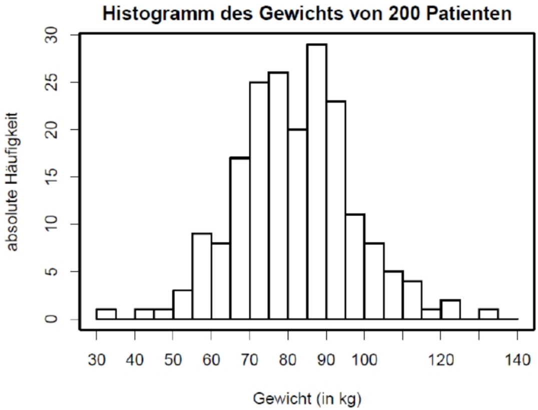
- Patientendaten: Klassenbreite 1 kg führt zu unruhigem Bild, auffällig: Häufungen bei Vielfachen von 5 kg



# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Grafische Darstellung: **Histogramm**

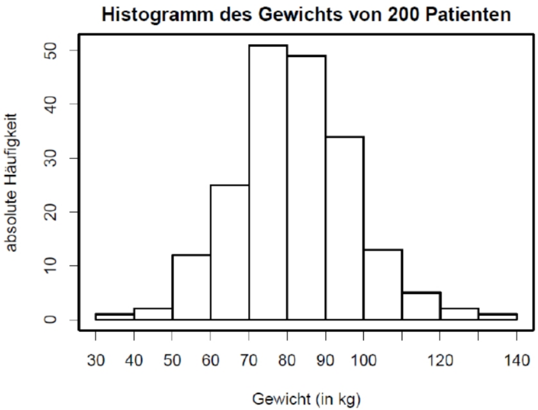
- Patientendaten: Klassenbreite 5 kg



# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

Grafische Darstellung: **Histogramm**

- Patientendaten: Klassenbreite 10 kg



# 2.2 Tabellarische und grafische Darstellung von univariaten Daten

- Bei qualitativen Merkmalen ist ein sogenanntes **Stabdiagramm (Balkendiagramm)** etabliert
  - Pro Merkmalsausprägung wird ein schmaler Stab (Balken) mit der absoluten oder relativen Häufigkeit über dem Merkmalswert gezeichnet
  - Merkmalsausprägungen werden für qualitative Merkmale gleichabständig auf der $x$-Achse gezeichnet
  - Stäbe sind immer (im Gegensatz zu Kästen beim Histogramm) voneinander separiert!
- Zur Visualisierung von Klassenanteilen an einer Gesamtheit wird häufig ein **Kuchen- bzw. Kreis-Diagramm** verwendet.
  - Dabei wird ein Kreis so in Sektoren aufgeteilt, dass die Sektorflächen proportional zu den absoluten (bzw. relativen) Häufigkeiten sind
  - Kreissegmente (Winkel) sind viel schlechter vergleichbar als Stäbe/Balken, deshalb besser Stabdiagramme verwenden!

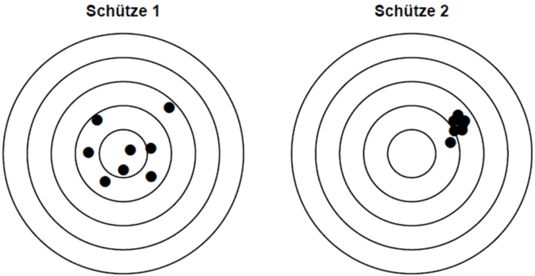
# Statistische Kennzahlen

# 3.1 Statistische Kennzahlen für die Lage

- Nach der passenden grafischen Darstellung der Werte eines Merkmals, nun (algebraische) Charakterisierungen der Verteilung solcher Werte.
- Ziel ist es, die Verteilung durch möglichst wenige Maßzahlen zu beschreiben.
  1. Wo liegt die Mitte der Werte?
     Repräsentative Charakterisierung einer Verteilung durch eine Zahl: **Lagemaß**
  2. Wie streuen die Werte um die Mitte?
     Charakterisierung der Größe der Unsicherheit (=Streuung) der Merkmalswerte: **Streuungsmaß**
- Später: Vergleich verschiedener Gesamtheiten miteinander mit Hilfe der Maßzahlen

# 3.1 Statistische Kennzahlen für die Lage

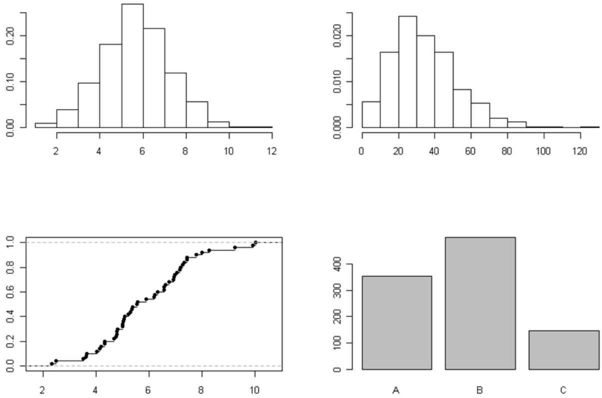
- Beispiel: Welcher Schütze schießt besser?




- Schütze 1: Lage gut, Streuung schlecht
- Schütze 2: Lage schlecht, Streuung gut

# 3.1 Statistische Kennzahlen für die Lage

Bisher: geringe Informationsverdichtung durch Verteilungsbeschreibung

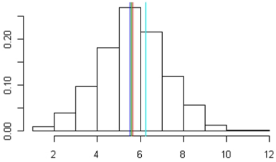


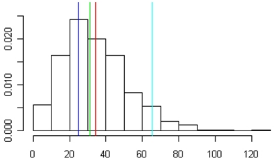

Beispiel

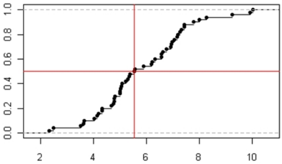
- Histogramm
- Empirische Verteilungsfunktion
- Stabdiagramm

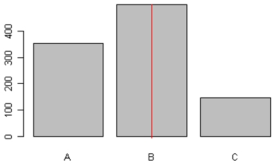
# 3.1 Statistische Kennzahlen für die Lage

Bisher: geringe Informationsverdichtung durch Verteilungsbeschreibung
Jetzt: stärkere Zusammenfassung der Daten auf ihr „Zentrum“



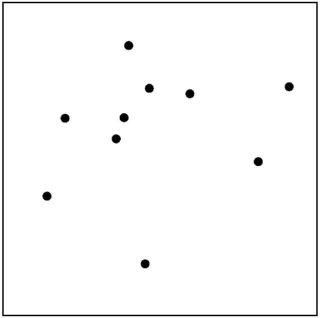




Farbige Linien repräsentieren das Zentrum

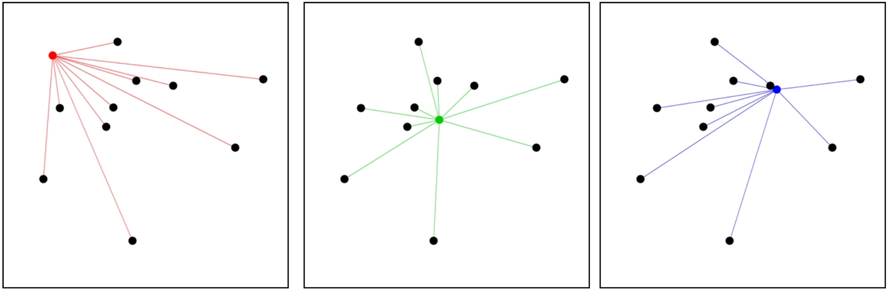
# 3.1 Statistische Kennzahlen für die Lage

Bisher: geringe Informationsverdichtung durch Verteilungsbeschreibung
Jetzt: stärkere Zusammenfassung der Daten auf ihr „Zentrum“
Unterschiedliche Definitionen von „Zentrum“.
Allgemein: repräsentative Merkmalsausprägung, von der alle beobachteten Werte möglichst wenig abweichen



# 3.1 Statistische Kennzahlen für die Lage

Bisher: geringe Informationsverdichtung durch Verteilungsbeschreibung
Jetzt: stärkere Zusammenfassung der Daten auf ihr „Zentrum"
Unterschiedliche Definitionen von „Zentrum".
Allgemein: repräsentative Merkmalsausprägung, von der alle beobachteten Werte möglichst wenig abweichen



# 3.1 Statistische Kennzahlen für die Lage

- Charakterisierung der Merkmalswerte auf einer Gesamtheit durch eine einzige Zahl: Lagemaße
- Lagemaß = „Mitte der Merkmalswerte“
- Auswahl des geeigneten Lagemaßes hängt vom Skalenniveau ab
- Wichtigste Beispiele
  - Arithmetisches Mittel: Klassischer Mittelwert
    - Reagiert am empfindlichsten auf „Ausreißer“, d.h. wenn für die Verteilung einige ungewöhnlich große oder kleine Werte vorliegen
  - Median: Zentralwert, mittlerer Wert in der geordneten Stichprobe
    - Liegt nicht unbedingt in der Mitte der Merkmalswerte, ist aber dennoch oft ein guter „Repräsentant“
    - Ist nicht unbedingt eindeutig
  - Modalwert: Häufigster Wert in der Stichprobe
    - Ist nicht unbedingt eindeutig
    - Bei stetigen Merkmalen meist erst nach Klassierung geeignet

# 3.1 Statistische Kennzahlen für die Lage

## Lagemaße:

### **Arithmetisches Mittel** = Mittelwert (mean)

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

### **Median** = „Zentralwert“ = 50%-Wert: $\text{med}_x$

Der Median ist derjenige Wert, für den 50% der Merkmalswerte größer oder gleich und 50% kleiner oder gleich sind.

Der Median ist der mittlere Wert der Rangliste:

$$\text{med}_x := \begin{cases} x_{\left(\frac{n+1}{2}\right)} & n \text{ ungerade} \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} & n \text{ gerade} \end{cases}$$

### **Modalwert / Modus** = häufigster Wert: $\text{mod}_x$

Der Modalwert ist derjenige Merkmalswert, der am häufigsten vorkommt.

# 3.1 Statistische Kennzahlen für die Lage

- $p$-Quantil $Q_p = \tilde{x}_p$
  - Verallgemeinerung des Medians (50%-Wert) auf beliebige Prozentzahlen (100%-Werte)
  - Nützliche Mittel zur Beschreibung einer Rangliste $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$

Ein $p$-**Quantil** $Q_p$, $p \in [0, 1]$, ist eine Zahl, für die $100 \cdot p\,\%$ der Merkmalswerte einer Gesamtheit kleiner oder gleich sind und $100 \cdot (1 - p)\,\%$ größer oder gleich.

Genauer könnte man für $Q_p$ z.B. Folgendes fordern:

$Q_p \geq$ größtem Merkmalswert einer Gesamtheit, der $\geq 100 \cdot p\,\%$ der Merkmalswerte ist und
$Q_p \leq$ nächstgrößerem Merkmalswert der Gesamtheit, also

$$x_{(\lfloor np \rfloor)} \leq Q_p \leq x_{(\lfloor np \rfloor + 1)}.$$

# 3.1 Statistische Kennzahlen für die Lage

Die folgende Berechnungsmethode für Quantile entspricht der obigen Berechnung des Medians.

$p$-Quantil Berechnung: „Standard“ (Nicht in R, dort `type` = 2 wählen.)

$$Q_p := \begin{cases} x_{(j)}, & j := \lceil np \rceil,\ np \text{ nicht ganzzahlig} \\ \frac{x_{(j)} + x_{(j+1)}}{2}, & j := np,\ np \text{ ganzzahlig} \end{cases}$$

**Bezeichnung**

- Anstelle von $p$-**Quantil** sagt man auch $100 \cdot p\,(\%)$-**Perzentil** oder **(1-$p$)-Fraktil**.
- 0.25- bzw. 0.75-Quantile heißen auch unteres bzw. oberes **Quartil**: unteres Quartil $q_4 = 0.25$-Quantil; oberes Quartil $q^4 = 0.75$-Quantil.

# 3.1 Statistische Kennzahlen für die Lage

- Nominale Daten
  - Gesucht: $x^*$, für das Abweichung zwischen $x^*$ und $x_1, \ldots, x_N$ minimal ist
  - Mit nominellen Ausprägungen kann keine sinnvolle Abweichung berechnet werden
  - Dummykodierung führt auf den Modalwert $x(j^*)$

| $i$ | $x_i$ |
|---|---|
| 1 | A |
| 2 | C |
| $\vdots$ | $\vdots$ |
| N | B |

| $i$ | $x_i$ | $d_i(1)$ | $d_i(2)$ | $d_i(3)$ |
|---|---|---|---|---|
| 1 | A | 1 | 0 | 0 |
| 2 | C | 0 | 0 | 1 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| N | B | 0 | 1 | 0 |
| $\sum$ | | $N_1$ | $N_2$ | $N_3$ |

# 3.1 Statistische Kennzahlen für die Lage

## Nominale Daten

### Modalwert

Beispiel **Bearbeitung von Softwareaufgaben**

Die Modalwerte lauten

$x_1(j^*) = \text{Oliver}$

$x_2(j^*) = \text{Export}$

$x_3(j^*) = 1.2$

**Bearbeiter(in)**

| Ausprägung | Absolute Häufigkeit | Relative Häufigkeit |
|---|---|---|
| Kai | 2 | 0.17 |
| Miriam | 3 | 0.25 |
| Oliver | 4 | 0.33 |
| Tina | 3 | 0.25 |
| | 12 | 1 |

**Aufgabe**

| Ausprägung | Absolute Häufigkeit | Relative Häufigkeit |
|---|---|---|
| Abfrage | 2 | 0.17 |
| Export | 6 | 0.5 |
| Verknüpfung | 4 | 0.33 |
| | 12 | 1 |

**Version**

| Ausprägung | Absolute Häufigkeit | Relative Häufigkeit |
|---|---|---|
| 1.1 | 3 | 0.25 |
| 1.2 | 6 | 0.5 |
| 2.0 | 3 | 0.25 |
| | 12 | 1 |

# 3.1 Statistische Kennzahlen für die Lage

## Ordinale Daten

$x_1, \ldots, x_N$

$x_i \in W_X, \; i = 1, \ldots, N$

$W_X = \{x(j) | j = 1, \ldots, J\} = \{x(1), \ldots, x(J)\}$

$x(1) < x(2) < \ldots < x(J)$

| $i$ | $x_i$ |
|---|---|
| 1 | $x(3)$ |
| 2 | $x(2)$ |
| 3 | $x(1)$ |
| 4 | $x(1)$ |
| 5 | $x(3)$ |

$\xrightarrow{\text{Geordnete Liste}}$

| $k$ | $x_{(k)}$ |
|---|---|
| 1 | $x(1)$ |
| 2 | $x(1)$ |
| 3 | $x(2)$ |
| 4 | $x(3)$ |
| 5 | $x(3)$ |

| | |
|---|---|
| Urliste | $x_1, \ldots, x_N$ |
| Geordnete Liste | $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(N)}$ |
| | $x_{(k)} = x_{i_k}$ |

mit $i_k = \min[\operatorname{argmin}_{i^*}(x_{i^*} | i^* \in \{1, \ldots, N\} \setminus \{i_1, \ldots, i_{k-1}\})]$, $k = 1, \ldots, N$

$x_{(k)}$ wird $k$-ter **Rangwert** genannt, erster und letzter Rangwert $x_{(1)}$ und $x_{(N)}$ heißen **Minimum** und **Maximum**.

# 3.1 Statistische Kennzahlen für die Lage

Ordinale Daten

$x_1, \ldots, x_N$

$x_i \in W_X,\ i = 1, \ldots, N$

$W_X = \{x(j) | j = 1, \ldots, J\} = \{x(1), \ldots, x(J)\}$

$x(1) < x(2) < \ldots < x(J)$

$x_{(k)}$ wird $k$-ter **Rangwert** genannt, erster und letzter Rangwert $x_{(1)}$ und $x_{(n)}$ heißen Minimum und Maximum.

$$R(x_i) = \frac{1}{\# K^*} \sum_{k^* \in K^*} k^* \text{ mit } K^* = \{k^* | x_{(k^*)} = x_i\}$$

$R(x_i)$ ist der **Rang** von $x_i$

| $i$ | $x_i$ | $R(x_i)$ |
|---|---|---|
| 1 | $x(3)$ | 4.5 |
| 2 | $x(2)$ | 3 |
| 3 | $x(1)$ | 1.5 |
| 4 | $x(1)$ | 1.5 |
| 5 | $x(3)$ | 4.5 |

↑ Ränge

| $k$ | $x_{(k)}$ |
|---|---|
| 1 | $x(1)$ |
| 2 | $x(1)$ |
| 3 | $x(2)$ |
| 4 | $x(3)$ |
| 5 | $x(3)$ |

# 3.1 Statistische Kennzahlen für die Lage

## Ordinale Daten

Beispiel **Bearbeitungen von Softwareaufgaben**

| $i$ | $\text{Version}_i$ |
|---|---|
| 1 | 1.1 |
| 2 | 1.2 |
| 3 | 1.1 |
| 4 | 1.2 |
| 5 | 2.0 |
| 6 | 1.2 |
| 7 | 1.2 |
| 8 | 1.2 |
| 9 | 1.2 |
| 10 | 1.1 |
| 11 | 2.0 |
| 12 | 2.0 |

Geordnete Liste →

| $k$ | $\text{Version}_{(k)}$ |
|---|---|
| 1 | 1.1 |
| 2 | 1.1 |
| 3 | 1.1 |
| 4 | 1.2 |
| 5 | 1.2 |
| 6 | 1.2 |
| 7 | 1.2 |
| 8 | 1.2 |
| 9 | 1.2 |
| 10 | 2.0 |
| 11 | 2.0 |
| 12 | 2.0 |

Ränge →

$$\frac{1}{3}\sum_{s=1}^{3} s = 2$$

$$\frac{1}{6}\sum_{s=4}^{9} s = 6.5$$

$$\frac{1}{3}\sum_{s=10}^{12} s = 11$$

| $i$ | $\text{Version}_i$ | $R(\text{Version}_i)$ |
|---|---|---|
| 1 | 1.1 | 2 |
| 2 | 1.2 | 6.5 |
| 3 | 1.1 | 2 |
| 4 | 1.2 | 6.5 |
| 5 | 2.0 | 11 |
| 6 | 1.2 | 6.5 |
| 7 | 1.2 | 6.5 |
| 8 | 1.2 | 6.5 |
| 9 | 1.2 | 6.5 |
| 10 | 1.1 | 2 |
| 11 | 2.0 | 11 |
| 12 | 2.0 | 11 |

# 3.1 Statistische Kennzahlen für die Lage

Quantitative Daten

$x_1, \ldots, x_N$

$x_i \in W_X,\ i = 1, \ldots, N$

$W_X = \{x(j) | j = 1, \ldots, J\} = \{x(1), \ldots, x(J)\}$
bzw. $W_X = (-\infty, \infty)$

Der Median minimiert die Summe der absoluten Abweichungen

$$\Delta_a(x) = \sum_{i=1}^{N} |x_i - x|$$

Der Mittelwert minimiert die Summe der quadratischen Abweichungen

$$\Delta(x) = \sum_{i=1}^{N} (x_i - x)^2$$

# 3.1 Statistische Kennzahlen für die Lage

## Quantitative Daten

Generell gilt: $\Delta(x) = \sum_{i=1}^{N} (x_i - x)^2$ ist minimal für $x = \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$

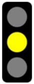
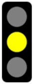
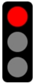
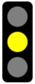
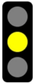
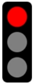
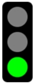
Beweis $\forall x \in \mathbb{R}$:

$$\begin{aligned}
\Delta(x) &= \sum_{i=1}^{N} (x_i - x)^2 = \sum_{i=1}^{N} [(x_i - \bar{x}) + (\bar{x} - x)]^2 \\
&= \sum_{i=1}^{N} (x_i - \bar{x})^2 + 2(\bar{x} - x) \underbrace{\sum_{i=1}^{N} (x_i - \bar{x})}_{=0\ (*)} + \underbrace{\sum_{i=1}^{N} (\bar{x} - x)^2}_{=N(\bar{x}-x)^2} \\
&= \Delta(\bar{x}) + \underbrace{N(\bar{x} - x)^2}_{\geq 0} \geq \Delta(\bar{x})
\end{aligned}$$

$$(*) \sum_{i=1}^{N} (x_i - \bar{x}) = \sum_{i=1}^{N} \left(x_i - \frac{1}{N} \sum_{l=1}^{N} x_l\right) = \sum_{i=1}^{N} x_i - \frac{1}{N} \sum_{i=1}^{N} \sum_{l=1}^{N} x_l = \sum_{i=1}^{N} x_i - \frac{1}{N} N \sum_{l=1}^{N} x_l = 0$$

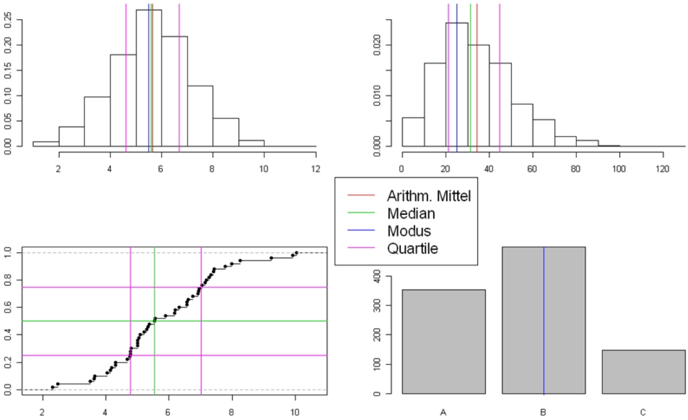
# 3.1 Statistische Kennzahlen für die Lage

Zusammenfassung: Welche Maßzahlen sind bei welchem Skalenniveau geeignet?

| **Skalenniveau →**<br>**↓ Lagemaß** | **Nominal** | **Ordinal** | **Quantitativ** |
|---|---|---|---|
| **Modus** |  |  – Informations-verlust |  – Nur für klassierte Daten |
| **Median** |  |  – Geringe Aussagekraft für kleine J |  + Robust<br>– Informations-verlust<br>– Hohe Streubreite |
| **Arithmetisches Mittel** |  – Nur für J = 2 |  |  – Ausreißeranfällig<br>+ Informations-nutzung<br>+ Geringe Streubreite |

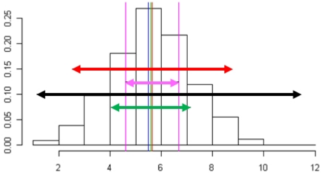
# 3.2 Statistische Kennzahlen für die Streuung

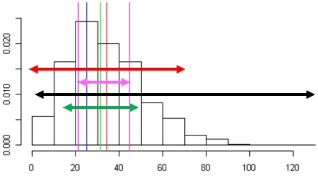
Bisher: Beschreibung von Häufigkeitsverteilung und Lage

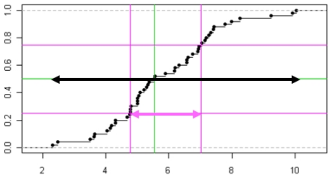


# 3.2 Statistische Kennzahlen für die Streuung

Bisher: Beschreibung von Häufigkeitsverteilung und Lage
Jetzt: Beschreibung der mittleren Variation um die Lage




Allgemein: Streuung desto höher, je schlechter sich konkrete Werte vorhersagen lassen.

# 3.2 Statistische Kennzahlen für die Streuung

## Streuungsmaße:

### empirische Varianz und Standardabweichung

- Varianz: „Durchschnitt“ der quadrierten Abweichungen vom arithmetischen Mittel

$$\mathrm{var}_x = s_x^2 := \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{(n-1)} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

- Standardabweichung: Wurzel aus der Varianz

$$s_x := \sqrt{\mathrm{var}_x}$$

### Interquartilsabstand (interquartile range)

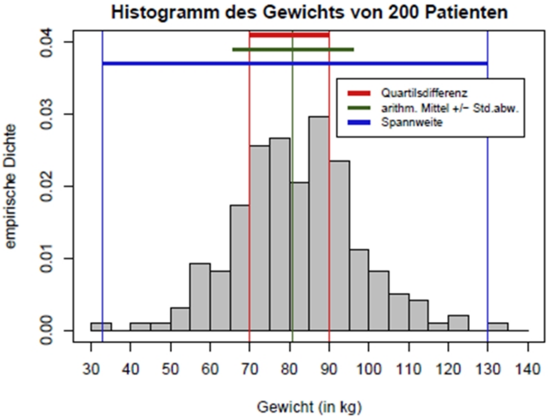
$$\mathrm{qd}_x := q^4 - q_4$$

### Spannweite (range)

$$R_x := \max(x) - \min(x) = x_{(n)} - x_{(1)}$$

# 3.2 Statistische Kennzahlen für die Streuung

- Beispiel 1: Gewicht von 200 Patienten
  $s_x = 15.14$ kg, $\mathrm{qd}_x = 20$ kg, $R_x = 97$ kg



# 3.2 Statistische Kennzahlen für die Streuung

## Streuungsmaße:

### **Variationskoeffizient** (relative Standardabweichung)

$$v_x := \frac{s_x}{\bar{x}}$$

### **Mittlere absolute Medianabweichung MD**

(von „Mean Deviation from the Median“)

$$\mathrm{md}_x := \frac{1}{n} \sum_{i=1}^{n} |x_i - \mathrm{med}_x|$$

### **Mediane absolute Medianabweichung MAD**

(von „Median Absolute Deviation“)

$$\mathrm{mad}_x := \mathrm{med}(|x_i - \mathrm{med}_x|)$$

# 3.2 Statistische Kennzahlen für die Streuung

## Nominale Daten

$x_1, \dots, x_N$

$x_i \in W_X,\ i = 1, \dots, N$

$W_X = \{x(j) | j = 1, \dots, J\}$
$= \{x(1), \dots, x(J)\}$

| $i$ | $x_i$ |
|---|---|
| 1 | A |
| 2 | C |
| $\vdots$ | $\vdots$ |
| N | B |

Rechnen nur sinnvoll mit Dummyvariablen bzw. Häufigkeiten

| $i$ | $x_i$ | $d_i(1)$ | $d_i(2)$ | $d_i(3)$ |
|---|---|---|---|---|
| 1 | A | 1 | 0 | 0 |
| 2 | C | 0 | 0 | 1 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| N | B | 0 | 1 | 0 |
| $\sum$ | | $N_1$ | $N_2$ | $N_3$ |

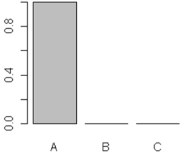
# 3.2 Statistische Kennzahlen für die Streuung

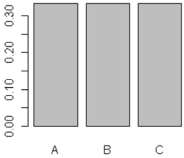
## Nominale Daten

Allgemein: Streuung ist desto höher, je schlechter sich konkrete Werte vorhersagen lassen.

Nominale Merkmalsausprägungen lassen sich um so besser vorhersagen, je häufiger eine bestimmte Kategorie vorkommt.

Geringste Streuung , falls es ein j gibt mit $f_j = 1$. →




← Höchste Streuung , falls $f_j = 1/J$, j=1,...,J.

# 3.2 Statistische Kennzahlen für die Streuung

Nominale Daten

Geringe Streuung, falls es ein $j$ gibt mit $f_j = 1$.

Höchste Streuung, falls $f_j = 1/J,\ j = 1, \ldots, J$

**Simpson's** $D$

$$D = 1 - \sum_{j=1}^{J} f_j^2$$

$D$ entspricht dem Anteil von Paaren mit unterschiedlichen Merkmalsausprägungen an allen aus der Urliste bildbaren Beobachtungspaaren:

$$D = \frac{\#\{(i,k) \in \{1,\ldots,N\} \times \{1,\ldots,N\} | x_i \neq x_k\}}{N^2}$$

Beispiel

| **i** | $\mathbf{x_i}$ |
|---|---|
| 1 | A |
| 2 | B |
| 3 | A |
| 4 | C |

$$D = 1 - \left( \frac{2^2 + 1^2 + 1^2}{4^2} \right) = 1 - \frac{6}{16} = \frac{5}{8}$$

$$= \frac{\#\{(1,2),(1,4),(2,1),(2,3),(2,4),(3,2),(3,4),(4,1),(4,2),(4,3)\}}{4^2}$$

# 3.2 Statistische Kennzahlen für die Streuung

Nominale Daten

Geringste Streuung, falls es ein $j$ gibt mit $f_j = 1$.

Höchste Streuung, falls $f_j = 1/J$, $j = 1, \ldots, J$.

**Simpson's** $D$

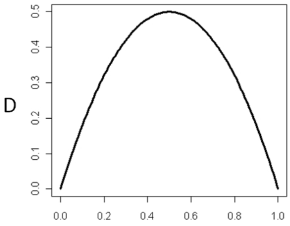
$$D = 1 - \sum_{j=1}^{J} f_j^2$$

$$\boxed{0} \leq D \leq \boxed{1 - \frac{1}{J}}$$

$D = 0$ für $\max[(f_1, \ldots, f_J)] = 1$

$D = 1 - \frac{1}{J}$ für $f_1 = \ldots = f_J = \frac{1}{J}$

Beispiel $J = 2$




$f_1 = 1 - f_2$

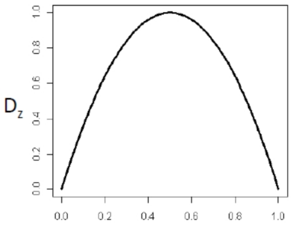
# 3.2 Statistische Kennzahlen für die Streuung

Nominale Daten

Geringste Streuung, falls es ein $j$ gibt mit $f_j = 1$.

Höchste Streuung, falls $f_j = 1/J$, $j = 1, \ldots, J$.

Beispiel $J = 2$




$f_1 = 1 - f_2$

**Simpson's $D_z$ (Normierte Version)**

$$D_z = \frac{J(1-\sum_{j=1}^{J} f_j^2)}{J-1}$$

$$\boxed{0} \leq D_z \leq \boxed{1}$$

$D_z = 0$ für $\max[(f_1, \ldots, f_J)] = 1$

$D_z = 1$ für $f_1 = \ldots = f_J = \frac{1}{J}$

# 3.2 Statistische Kennzahlen für die Streuung

Nominale Daten

Informationstheorie: Ein Ereignis liefert desto mehr Information, je geringer seine Eintrittswahrscheinlichkeit ist.

Kodierung der Elementarereignisse in Bits, Beispiel **Kaffeebestellung**

| | | |
|---|---|---|
| 1. Bit: | 0 = Tasse | 1 = Kännchen |
| 2. Bit | 0 = Schwarz | 1 = mit Milch |
| 3. Bit: | 0 = Süßstoff | 1 = Zucker |

8 mögliche Bestellungen: 000, 001, 010, 011, 100, 101, 110, 111

Beträgt die Wahrscheinlichkeit einer Teilmenge dieser Bestellungen $p = 1/8$, wird genau eine Bestellung ausgewählt und man erhält Information über alle $3 = -\log_2(1/8)$ Bits, falls die Teilmenge ausgewählt wird.

Wird dagegen die Menge möglicher Bestellungen auf 50%, z.B. alle Bestellungen mit Kännchen eingegrenzt, also $p = 1/4$, so erhält man Information über $2 = -\log_2(1/4)$ Bits.

# 3.2 Statistische Kennzahlen für die Streuung

## Nominale Daten

Die Information einer Merkmalsausprägung $x(j)$ in Bits kann also allgemein definiert werden durch -$\log_2(f_j)$.

Der Informationsgehalt des gesamten Merkmals $x$ ergibt sich durch die **Entropie** genannte erwartete Information $H(x)$ von $x$:

$$H(x) = -\sum_{j=1}^{J} f_j \log_2(f_j)$$

Beispiel **Kaffeebestellung**: Sei $x_F$ die Antwort auf eine bestimmte Frage $F$


$F$=„Möchten Sie Ihren Kaffee mit Milch und Zucker?“

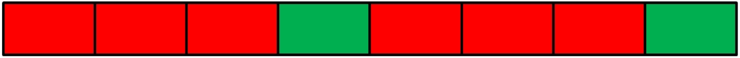
$$x_F(0) = \text{„Nein“},\ x_F(1) = \text{„Ja“},\ f_1 = 6/8,\ f_2 = 2/8,$$

$$H(x_F) = -(6/8)\cdot\log_2(6/8) - (2/8)\cdot\log_2(2/8) = \boxed{0.8113}$$

# 3.2 Statistische Kennzahlen für die Streuung

Nominale Daten

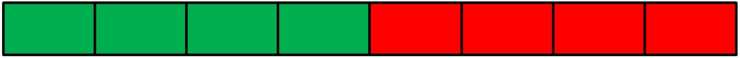
**Entropie** von $x$: $$H(x) = -\sum_{j=1}^{J} f_j \log_2(f_j)$$ Beispiel **Kaffeebestellung**



000 001 010 011 100 101 110 111

$F$=„Möchten Sie Ihren Kaffee mit Milch und Zucker?“

$x_F(0) =$„Nein“, $x_F(1) =$„Ja“, $f_1 = 6/8$, $f_2 = 2/8$, $H(x_F) = \boxed{0.8113}$



000 001 010 011 100 101 110 111

$F$ = „Möchten Sie Ihren Kaffee in der Tasse?“

$x_F(0) =$„Nein“, $x_F(1) =$„Ja“, $f_1 = 4/8$, $f_2 = 4/8$,

$$H(x_F) = -(4/8) \cdot \log_2(4/8) - (4/8) \cdot \log_2(4/8) = \boxed{1}$$

# 3.2 Statistische Kennzahlen für die Streuung

- Entropie gibt also die Information an, die man im Mittel durch Kenntnis der tatsächlichen Ausprägung erhält, wenn man vorher nur die Verteilung kannte. Ist diese hoch, konnte man den Wert vorher schlecht vorhersagen $\Rightarrow$ hohe Streuung.
- Ist der Informationszugewinn gering, konnte man vorher schon gut prognostizieren.
- Beispiel „Wer wird Millionär“
    - Kandidat ist sicher = geringe Streuung, keine weitere Information durch Joker
    - Kandidat ist unsicher = hohe Streuung, erhofft Informationsgewinn durch Publikumsjoker
    - Ist hier die Streuung hoch, weiterer Informationsgewinn durch Einzelbefragungsjoker

# 3.2 Statistische Kennzahlen für die Streuung

Nominale Daten

**Entropie** von $x$: $$H(x) = -\sum_{j=1}^{J} f_j \cdot \log_2(f_j)$$

Die Entropie ist ein sinnvolles Maß für die Streuung, denn sie erfüllt die Forderungen:

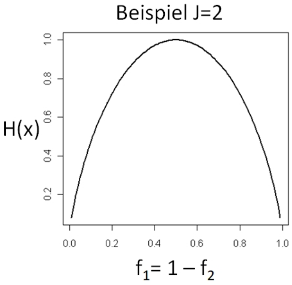
Geringe Streuung, falls es ein $j$ gibt mit $f_j = 1$.

Höchste Streuung, falls $f_j = 1/J$, $j = 1, \ldots, J$.

$$\boxed{0} < H(x) \leq \boxed{\log_2(J)}$$

$$\lim[H(x)] = 0 \text{ für } \max[(f_1, \ldots, f_J)] \to 1$$

$$H(x) = \log_2(J) \text{ für } f_1 = \ldots = f_J = \frac{1}{J}$$



# 3.2 Statistische Kennzahlen für die Streuung

Nominale Daten

**Normierte Entropie** von $x$: $$H_n(x) = -\sum_{j=1}^{J} f_j \cdot \frac{\log_2(f_j)}{\log_2(J)}$$

Die Entropie ist ein sinnvolles Maß für die Streuung, denn sie erfüllt die Forderungen:

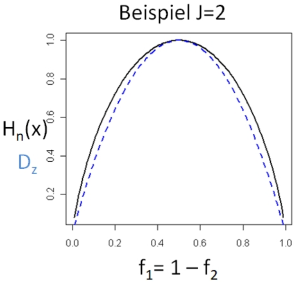
$\boxed{\text{Geringe Streuung, falls es ein } j \text{ gibt mit } f_j = 1.}$

$\boxed{\text{Höchste Streuung, falls } f_j = 1/J,\ j = 1, \ldots, J.}$

$\boxed{0} < H_n(x) \leq \boxed{1}$

$\boxed{\lim[H_n(x)] = 0 \text{ für } \max[(f_1, \ldots, f_J)] \to 1}$

$\boxed{H_n(x) = 1 \text{ für } f_1 = \ldots = f_J = \frac{1}{J}}$



# 3.2 Statistische Kennzahlen für die Streuung

## Nominale Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Merkmal | $D_z$ | $H_n(x)$ |
|---|---|---|
| Bearbeiter(in) | $4/3\cdot(1-0.17^2-0.25^2-0.33^2-0.25^2) = \mathbf{0.9815}$ | $[-0.17\cdot\log_2(0.17)-0.25\cdot\log_2(0.25)-0.33\cdot\log_2(0.33)-0.25\cdot\log_2(0.25)]/\log_2(4) = \mathbf{0.9796}$ |
| Aufgabe | $3/2\cdot(1-0.17^2-0.5^2-0.33^2) = \mathbf{0.9167}$ | $[-0.17\cdot\log_2(0.17)-0.5\cdot\log_2(0.5)-0.33\cdot\log_2(0.33)]/\log_2(3) = \mathbf{0.9206}$ |
| Version | $3/2\cdot(1-0.25^2-0.5^2-0.25^2) = \mathbf{0.9375}$ | $[-0.25\cdot\log_2(0.25)-0.5\cdot\log_2(0.5)-0.25\cdot\log_2(0.25)]/\log_2(3) = \mathbf{0.9464}$ |

**Bearbeiter(in)**

| Ausprägung | Absolute Häufigkeit | Relative Häufigkeit |
|---|---|---|
| Kai | 2 | 0.17 |
| Miriam | 3 | 0.25 |
| Oliver | 4 | 0.33 |
| Tina | 3 | 0.25 |
| | 12 | 1 |

**Aufgabe**

| Ausprägung | Absolute Häufigkeit | Relative Häufigkeit |
|---|---|---|
| Abfrage | 2 | 0.17 |
| Export | 6 | 0.5 |
| Verknüpfung | 4 | 0.33 |
| | 12 | 1 |

**Version**

| Ausprägung | Absolute Häufigkeit | Relative Häufigkeit |
|---|---|---|
| 1.1 | 3 | 0.25 |
| 1.2 | 6 | 0.5 |
| 2.0 | 3 | 0.25 |
| | 12 | 1 |

# 3.2 Statistische Kennzahlen für die Streuung

Ordinale Daten

$x_1, \ldots, x_N$

$x_i \in W_x,\ i = 1, \ldots, N$

$W_X = \{x(j) |\ j = 1, \ldots, J\} = \{x(1), \ldots, x(J)\}$

$x(1) < x(2) < \ldots < x(J)$

| $i$ | $x_i$ |
|---|---|
| 1 | $x(3)$ |
| 2 | $x(2)$ |
| 3 | $x(1)$ |
| 4 | $x(1)$ |
| 5 | $x(3)$ |

$\xrightarrow{\text{Geordnete Liste}}$

| $k$ | $x_{(k)}$ |
|---|---|
| 1 | $x(1)$ |
| 2 | $x(1)$ |
| 3 | $x(2)$ |
| 4 | $x(3)$ |
| 5 | $x(3)$ |

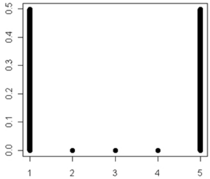
Simpson's $D$ und $H(x)$ sind anwendbar, allerdings wird die Information der Kategorienordnung nicht genutzt.

# 3.2 Statistische Kennzahlen für die Streuung

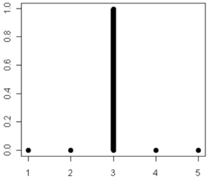
## Ordinale Daten

Allgemein: Streuung deste höher, je schlechter konkrete Werte sich vorhersagen lassen.

Werte lassen sich umso besser vorhersagen, je stärker sie sich um den Median verdichten.




Geringste Streuung für $N(\tilde{x}_{0.5}) = N \rightarrow$




$\leftarrow$ Höchste Streuung für $N(\tilde{x}_0) = N(\tilde{x}_1) = N/2$

Nicht mehr höchste Streuung bei ausgeglichener Belegung, da die Kategorien unterschiedlich weit von der Mitte entfernt sind. Höchste Streuung bei maximaler Entfernung zur Mitte, also bei gleichmäßiger Konzentration an Minimum und Maximum.

# 3.2 Statistische Kennzahlen für die Streuung

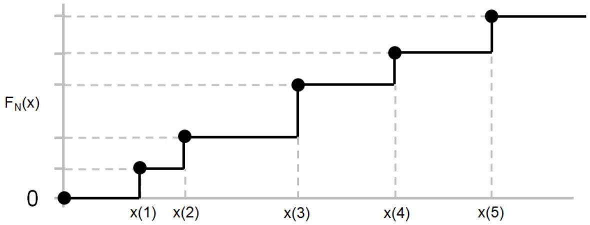
Ordinale Daten

Geringe Streuung für $N(\tilde{x}_{0.5}) = N$

Höchste Streuung für $N(\tilde{x}_0) = N(\tilde{x}_1) = N/2$

**Dispersionsindex nach Leti**

$$D_L = \sum_{j=1}^{J-1} F_N[x(j)] \cdot (1 - F_N[x(j)])$$

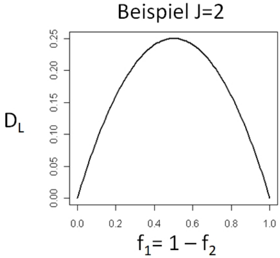


# 3.2 Statistische Kennzahlen für die Streuung

Ordinale Daten

$\boxed{\text{Geringe Streuung für } N(\tilde{x}_{0.5}) = N}$

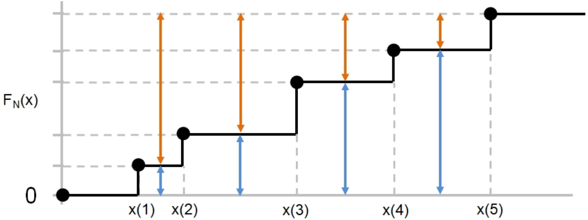
$\boxed{\text{Höchste Streuung für } N(\tilde{x}_0) = N(\tilde{x}_1) = N/2}$




**Dispersionsindex nach Leti**

$$D_L = \sum_{j=1}^{J-1} \boxed{F_N[x(j)]} \cdot \boxed{(1 - F_N[x(j)])}$$

$$\boxed{0} \leq D_L \leq \boxed{\frac{J-1}{4}}$$



# 3.2 Statistische Kennzahlen für die Streuung

Ordinale Daten

$\boxed{\text{Geringe Streuung für } N(\tilde{x}_{0.5}) = N}$

$\boxed{\text{Höchste Streuung für } N(\tilde{x}_0) = N(\tilde{x}_1) = N/2}$

**Normierter Dispersionsindex nach Leti**

$$D_{LZ} = \tfrac{4}{J-1} \sum_{j=1}^{J-1} F_N[x(j)] \cdot (1 - F_N[x(j)])$$

$$\boxed{0} \leq D_{LZ} \leq \boxed{1}$$

**Für** $J = 2$ **gilt** $D_Z = D_{LZ}$,

d.h. normierte Versionen von Simpson und Leti sind äquivalent.

Beweis: $\boxed{D_{LZ}} =$

$$= \tfrac{4}{2-1} \sum_{j=1}^{1} F_N[x(j)](1 - F_N[x(j)])$$

$$= 4 \cdot (f_1(1 - f_1)) \; = \; 2 \cdot (2f_1 - 2f_1^2)$$

$$= 2(1 - f_1^2 - 1 + 2f_1 - f_1^2)$$

$$= 2(1 - [f_1^2 + (1 - f_1)^2])$$

$$= \frac{2\left(1 - \sum_{j=1}^{2} f_j^2\right)}{2-1} = \boxed{D_Z} \quad \square$$

# 3.2 Statistische Kennzahlen für die Streuung

Quantitative Daten

$x_1, \ldots, x_N$

$x_i \in W_x,\ i = 1, \ldots, N$

$W_X = \{x(j) |\, j = 1, \ldots, J\} = \{x(1), \ldots, x(J)\}$

bzw. $W_x = (-\infty, \infty)$

Allgemein: Streuung desto höher, je schlechter konkrete Werte sich vorhersagen lassen.

Werte lassen sich umso besser vorhersagen, je stärker sie sich um das jeweilige Lagemaß verdichten.

# 3.2 Statistische Kennzahlen für die Streuung

Quantitative Daten

Werte lassen sich umso besser vorhersagen, je stärker sie sich um das jeweilige Lagemaß verdichten.

Lagemaß: Arithmetisches Mittel

Streuungsmaß:

**Varianz (mittlere quadratische Abweichung**

$$s_x^2 = \frac{1}{N-1}\sum_{n=1}^{N}(x_n - \bar{x})^2 \qquad \left(\text{bzw. } d_x^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})^2\right.$$

**Standardabweichung**

$$s = \sqrt{s_x^2} = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x_n - \bar{x})^2}$$

# 3.2 Statistische Kennzahlen für die Streuung

## Quantitative Daten

Von Streuungsparametern abgeleitete Größen für verhältnisskalierte Merkmale

**Quartilskoeffizient**

$$Q_{\text{koeff}} = \frac{2Q}{\tilde{x}_{0.25} + \tilde{x}_{0.75}} = \frac{2(\tilde{x}_{0.75} - \tilde{x}_{0.25})}{\tilde{x}_{0.25} + \tilde{x}_{0.75}} = (\tilde{x}_{0.75} - \tilde{x}_{0.25}) \Big/ \left(\frac{\tilde{x}_{0.25} + \tilde{x}_{0.75}}{2}\right)$$

**Variationskoeffizient**

$$V_x = \frac{s_x}{\bar{x}}$$

# 3.2 Statistische Kennzahlen für die Streuung

Quantitative Daten: **Berechnung der Varianz aus Häufigkeitsverteilung**

$$s_x^2 = \frac{N}{N-1}\sum_{j=1}^{J} f_j \cdot \left[x(j) - \sum_{k=1}^{J} f_k \cdot x(k)\right]^2 = \frac{N}{N-1}\sum_{j=1}^{J} f_j \cdot (x(j) - \bar{x})^2$$

Beweis:

$$s_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_{(i)} - \bar{x})^2 = \frac{(x_{(1)} - \bar{x})^2}{N-1} + \ldots + \frac{(x_{(N)} - \bar{x})^2}{N-1}$$

$$= \underbrace{\frac{(x(1) - \bar{x})^2}{N-1} + \ldots + \frac{(x(1) - \bar{x})^2}{N-1}}_{f_1 \cdot N\text{-mal}} + \ldots + \underbrace{\frac{(x(J) - \bar{x})^2}{N-1} + \ldots + \frac{(x(J) - \bar{x})^2}{N-1}}_{f_J \cdot N\text{-mal}}$$

$$= \frac{N}{N-1} f_1 \cdot (x(1) - \bar{x})^2 + \ldots + \frac{N}{N-1} f_J \cdot (x(J) - \bar{x})^2 = \frac{N}{N-1}\sum_{j=1}^{J} f_j \cdot (x(j) - \bar{x})^2 \quad \square$$

# 3.2 Statistische Kennzahlen für die Streuung

Quantitative Daten: **Varianz von Lineartransformationen**

$$y = ax + b \Rightarrow s_y^2 = a^2 s_x^2$$

Beweis:

$$\boxed{s_y^2} = \frac{1}{N-1}\sum_{n=1}^{N}(y_n - \bar{y})^2 = \frac{1}{N-1}\sum_{n=1}^{N}[ax_n + b - (a\bar{x} + b)]^2$$

$$= \frac{1}{N-1}\sum_{n=1}^{N}(ax_n - a\bar{x})^2$$

$$= a^2\frac{1}{N-1}\sum_{n=1}^{N}(x_n - \bar{x})^2 = \boxed{a^2 s_x^2} \quad \square$$

$$\bar{y} = \overline{ax + b} = \frac{1}{N}\sum_{n=1}^{N}(ax_n + b) = a\frac{1}{N}\sum_{n=1}^{N}x_n + \frac{bN}{N} = a\bar{x} + b$$

# 3.2 Statistische Kennzahlen für die Streuung

Quantitative Daten: **Verschiebungssatz von Steiner**

$$d_x^2 = \left(\frac{1}{N}\sum_{n=1}^{N}(x_n - b)^2\right) \quad \text{speziell für } b = 0: \; d_x^2 = \overline{x^2} - \bar{x}^2$$

Beweis:

$$\boxed{d_x^2} = \frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})^2 = \frac{1}{N}\sum_{n=1}^{N}[(x_n - b) + (b - \bar{x})]^2$$

$$= \frac{1}{N}\sum_{n=1}^{N}[(x_n - b)^2 + 2(x_n - b)(b - \bar{x}) + (b - \bar{x})^2]$$

$$= \frac{1}{N}\sum_{n=1}^{N}(x_n - b)^2 + 2(b - x)\frac{1}{N}\sum_{n=1}^{N}(x_n - b) + \frac{1}{N}\sum_{n=1}^{N}(b - \bar{x})^2$$

$$= \frac{1}{N}\sum_{n=1}^{N}(x_n - b)^2 - 2(\bar{x} - b)^2 + (\bar{x} - b)^2 = \boxed{\frac{1}{N}\sum_{n=1}^{N}(x_n - b)^2 - (\bar{x} - b)^2} \quad \square$$

# 3.2 Statistische Kennzahlen für die Streuung

Quantitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| k | Anzahl Clicks$_{(k)}$ |
|---|---|
| 1 | 10 |
| 2 | 10 |
| 3 | 11 |
| 4 | 12 |
| 5 | 12 |
| 6 | 13 |
| 7 | 14 |
| 8 | 14 |
| 9 | 15 |
| 10 | 16 |
| 11 | 17 |
| 12 | 18 |

$$s^2_{x_4} = \frac{2\cdot(10-13.5)^2+1\cdot(11-13.5)^2+2\cdot(12-13.5)^2}{11} + \frac{1\cdot(13-13.5)^2+2\cdot(14-13.5)^2+1\cdot(15-13.5)^2}{11} + \frac{1\cdot(16-13.5)^2+1\cdot(17-13.5)^2+1\cdot(18-13.5)^2}{11} = \boxed{7}$$

$$\bar{x}_4 = \boxed{13.5}$$

$$V_4 = \frac{\sqrt{7}}{13.5} = \boxed{0.196}$$

$$Q_{koeff;4} = \frac{2\cdot 4}{11.5+15.5} = \boxed{0.296}$$

$$\tilde{x}_{4;0.25} = 11.5$$

$$\tilde{x}_{4;0.75} = 15.5$$

$$\boxed{Q_4 = 4}$$

$$\boxed{R_4 = 8}$$

# 3.2 Statistische Kennzahlen für die Streuung

## Quantitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

$$s_{x_5}^2 = \frac{(3.2-5.075)^2+(3.6-5.075)^2+2\cdot(3.7-5.075)^2}{11} + \frac{(3.9-5.075)^2+(4.2-5.075)^2+(4.5-5.075)^2}{11} + \frac{(4.9-5.075)^2+(6.1-5.075)^2+(6.6-5.075)^2}{11} + \frac{(8.0-5.075)^2+(8.5-5.075)^2}{11} = \boxed{3.24}$$

$$\bar{x}_5 = \boxed{5.075}$$

$$V_5 = \frac{\sqrt{3.24}}{5.075} = \boxed{0.355}$$

| k | Bearbeitungszeit$_{(k)}$ |
|---|---|
| 1 | 3.2 |
| 2 | 3.6 |
| 3 | 3.7 |
| 4 | 3.7 |
| 5 | 3.9 |
| 6 | 4.2 |
| 7 | 4.5 |
| 8 | 4.9 |
| 9 | 6.1 |
| 10 | 6.6 |
| 11 | 8.0 |
| 12 | 8.5 |

$$Q_{koeff;5} = \frac{2\cdot 2.65}{3.7+6.35} = \boxed{0.527}$$

$$\boxed{\tilde{x}_{5;0.25} = 3.7}$$

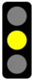
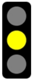
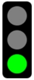
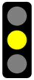
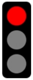
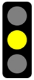
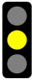
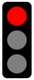
$$\boxed{\tilde{x}_{5;0.75} = 6.35}$$

$$\boxed{Q_5 = 2.65}$$

$$\boxed{R_5 = 5.3}$$

# 3.2 Statistische Kennzahlen für die Streuung

Zusammenfassung: Welche Maßzahlen sind bei welchem Skalenniveau geeignet?

| **Skalenniveau →**<br>**↓ Streuungsmaß** | **Nominal** | **Ordinal** | **Quantitativ** |
|---|---|---|---|
| **Simpson's D/ Entropie** |  |  – Informations-verlust |  – Nur für klassierte Daten |
| **Leti's D** |  – Nur für J = 2 |  |  – Nur für klassierte Daten |
| **MAD/ Spannweite/ Quartilsdifferenz** |  |  – Geringe Aussagekraft für kleine J |  + Robust<br>– Informations-verlust<br>– Hohe Streubreite |
| **Varianz/ Standardabweichung Variationskoeffizient** |  – Nur für J = 2 |  |  – Ausreißeranfällig<br>+ Informations-nutzung<br>+ Geringe Streubreite |

# Bivariate Daten

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

Bisher: Betrachtung einzelner Merkmale $X$

Jetzt: Betrachtung von Merkmalspaaren $(X, Y)$

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version | Anzahl Clicks | Bearbeitungszeit |
|---|---|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 | 14 | 8.0 |
| $e_2$ | Kai | Verknüpfung | 1.2 | 12 | 4.9 |
| $e_3$ | Miriam | Export | 1.1 | 12 | 6.6 |
| $e_4$ | Tina | Verknüpfung | 1.2 | 13 | 3.2 |
| $e_5$ | Oliver | Export | 2.0 | 17 | 3.9 |
| $e_6$ | Tina | Export | 1.2 | 11 | 4.5 |
| $e_7$ | Tina | Verknüpfung | 1.2 | 14 | 6.1 |
| $e_8$ | Miriam | Export | 1.2 | 10 | 3.7 |
| $e_9$ | Miriam | Export | 1.2 | 10 | 4.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 | 18 | 8.5 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 | 16 | 3.6 |
| $e_{12}$ | Oliver | Abfrage | 2.0 | 15 | 3.7 |

X=Aufgabe

Y=Anzahl Clicks

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

## Nominale Daten

Univariate Urlisten:

$x_1, \ldots, x_N$
$y_1, \ldots, y_N$

Univariate Wertebereiche:

$x_i \in W_X,\ y_i \in W_Y,\ i = 1, \ldots, N$

$W_X = \{x(j) | j = 1, \ldots, J\} = \{x(1), \ldots, x(J)\}$

$W_Y = \{Y(k) | k = 1, \ldots, K\} = \{y(1), \ldots, y(K)\}$

| i | $x_i$ | $y_i$ |
|---|---|---|
| 1 | A | D |
| 2 | C | E |
| ... | ... | |
| N | B | E |

Bivariate Urliste:
$(x_1, y_1), \ldots, (x_N, y_N)$

Bivariater Wertebereich:
$(x_i, y_i) \in W_{XY} = W_X \times W_Y = \{(x[1], y[1]), \ldots, (x[1], y[K]), (x[2], y[1]), \ldots, (x[J], y[K])\}$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

## Nominale Daten

$x_1, \ldots, x_N;\ y_1, \ldots, y_N$ $\quad (x_1, y_1), \ldots, (x_N, y_N)$ $\quad d_i(j) = I_{x(e_i)=x(j)}$

$x_i \in W_X,\ y_i \in W_Y$ $\quad (x_i, y_i) \in W_{XY} = W_X \times W_Y$ $\quad r_i(k) = I_{y(e_i)=y(k)}$

| i | $x_i$ | $y_i$ |
|---|---|---|
| 1 | A | D |
| 2 | C | E |
| ... | ... | |
| N | B | E |

Dummykodierung →

| i | $x_i$ | $y_i$ | $d_i(1)$ | $d_i(2)$ | $d_i(3)$ | $r_i(1)$ | $r_i(2)$ |
|---|---|---|---|---|---|---|---|
| 1 | A | D | 1 | 0 | 0 | 1 | 0 |
| 2 | C | E | 0 | 0 | 1 | 0 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| N | B | E | 0 | 1 | 0 | 0 | 1 |
| Σ | | | $N_{1\bullet}$ | $N_{2\bullet}$ | $N_{3\bullet}$ | $N_{\bullet 1}$ | $N_{\bullet 2}$ |

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

## Nominale Daten

Häufigkeitsverteilung eines bivariaten Merkmals

$(x_i, y_i) \in W_{XY} = W_X \times W_Y, \; i = 1, \ldots, N$

$$W_{XY} = \{(x[j], y[k]) | j = 1, \ldots, J; k = 1, \ldots, K\}$$

$$= \begin{Bmatrix} (x[1], y[1]) & \ldots & (x[1], y[K]) \\ (x[2], y[1]) & \ldots & (x[2], y[K]) \\ \vdots & \ddots & \vdots \\ (x[J], y[1]) & \ldots & (x[J], y[K]) \end{Bmatrix}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

<u>Nominale Daten</u>

**Gemeinsame absolute Häufigkeitsverteilung von $x$ und $y$**

$$N_{jk} = N\big((x[j], y[k])\big),\ j = 1, \dots, J;\ k = 1, \dots, K \qquad \begin{pmatrix} N_{11} & \dots & N_{1K} \\ N_{21} & \dots & N_{2K} \\ \vdots & \ddots & \vdots \\ N_{J1} & \dots & N_{JK} \end{pmatrix}$$

**Gemeinsame relative Häufigkeitsverteilung von $x$ und $y$**

$$f_{jk} = \frac{N_{jk}}{N},\ j = 1, \dots, J;\ k = 1, \dots, K \qquad \begin{pmatrix} f_{11} & \dots & f_{1k} \\ f_{21} & \dots & f_{2k} \\ \vdots & \ddots & \vdots \\ f_{j1} & \dots & f_{jk} \end{pmatrix}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Häufigkeitsverteilung eines bivariaten Merkmals

| i | $x_i$ | $y_i$ | $d_i(1)$ | $d_i(2)$ | $d_i(3)$ | $r_i(1)$ | $r_i(2)$ |
|---|---|---|---|---|---|---|---|
| 1 | A | D | 1 | 0 | 0 | 1 | 0 |
| 2 | C | E | 0 | 0 | 1 | 0 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| N | B | E | 0 | 1 | 0 | 0 | 1 |
| Σ | | | $N_{1\bullet}$ | $N_{2\bullet}$ | $N_{3\bullet}$ | $N_{\bullet 1}$ | $N_{\bullet 2}$ |

$$
\begin{aligned}
N_{jk} &= N\big((x[j], y[k])\big) \\
&= \sum_{i\in\{l|r_l(k)=1\}} d_i(j) = \sum_{i\in\{l|d_l(j)=1\}} r_i(k) \\
&= \sum_{i=1}^{N} d_i(j)\cdot r_i(k)
\end{aligned}
$$

$$
\begin{aligned}
N_{j\bullet} &= \sum_{i=1}^{N} d_i(j) = \sum_{i\in\{l|r_l(1)=1\}} d_i(j) + \sum_{i\in\{l|r_l(2)=1\}} d_i(j) + \ldots + \sum_{i\in\{l|r_l(K)=1\}} d_i(j) \\
&= \sum_{k=1}^{K}\sum_{l=1}^{N} d_l(j)\cdot r_l(k) = \sum_{k=1}^{K} N_{jk}
\end{aligned}
$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

Absolute Häufigkeiten

| | | Y | | | | $\Sigma$ |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | $\Sigma$ |
| X | x(1) | $N_{11}$ | $N_{12}$ | … | $N_{1K}$ | $N_{1\bullet}$ |
| | x(2) | $N_{21}$ | $N_{22}$ | … | $N_{2K}$ | $N_{2\bullet}$ |
| | … | … | … | … | … | … |
| | x(J) | $N_{J1}$ | $N_{J2}$ | … | $N_{JK}$ | $N_{J\bullet}$ |
| | $\Sigma$ | $N_{\bullet 1}$ | $N_{\bullet 2}$ | … | $N_{\bullet K}$ | N |

$$N_{j\bullet} = \sum_{k=1}^{K} N_{jk}$$

$$N_{\bullet k} = \sum_{j=1}^{J} N_{jk}$$

$$N = \sum_{j=1}^{J} \sum_{k=1}^{K} N_{jk}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

**Gemeinsame absolute Häufigkeitsverteilung von $X$ und $Y$**

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | Σ |
| X | x(1) | $N_{11}$ | $N_{12}$ | … | $N_{1K}$ | $N_{1\bullet}$ |
| | x(2) | $N_{21}$ | $N_{22}$ | … | $N_{2K}$ | $N_{2\bullet}$ |
| | … | … | … | … | … | … |
| | x(J) | $N_{J1}$ | $N_{J2}$ | … | $N_{JK}$ | $N_{J\bullet}$ |
| | Σ | $N_{\bullet 1}$ | $N_{\bullet 2}$ | … | $N_{\bullet K}$ | N |

$$N_{j\bullet} = \sum_{k=1}^{K} N_{jk}$$

$$N_{\bullet k} = \sum_{j=1}^{J} N_{jk}$$

$$N = \sum_{j=1}^{J} \sum_{k=1}^{K} N_{jk}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | Σ |
| X | x(1) | $N_{11}$ | $N_{12}$ | … | $N_{1K}$ | $N_{1\cdot}$ |
| | x(2) | $N_{21}$ | $N_{22}$ | … | $N_{2K}$ | $N_{2\cdot}$ |
| | … | … | … | … | … | … |
| | x(J) | $N_{J1}$ | $N_{J2}$ | … | $N_{JK}$ | $N_{J\cdot}$ |
| | Σ | $N_{\cdot 1}$ | $N_{\cdot 2}$ | … | $N_{\cdot K}$ | N |

**Absolute Randhäufigkeitsverteilung** von X

**Absolute Randhäufigkeitsverteilung** von Y

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | Σ |
| X | x(1) | $N_{11}/N$ | $N_{12}/N$ | ... | $N_{1K}/N$ | $N_{1\bullet}/N$ |
| | x(2) | $N_{21}/N$ | $N_{22}/N$ | ... | $N_{2K}/N$ | $N_{2\bullet}/N$ |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | $N_{J1}/N$ | $N_{J2}/N$ | ... | $N_{JK}/N$ | $N_{J\bullet}/N$ |
| | Σ | $N_{\bullet 1}/N$ | $N_{\bullet 2}/N$ | ... | $N_{\bullet K}/N$ | N/N |

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

Relative Häufigkeiten

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | $\Sigma$ |
| X | x(1) | $f_{11}$ | $f_{12}$ | … | $f_{1K}$ | $f_{1\bullet}$ |
| | x(2) | $f_{21}$ | $f_{22}$ | … | $f_{2K}$ | $f_{2\bullet}$ |
| | … | … | … | … | … | … |
| | x(J) | $f_{J1}$ | $f_{J2}$ | … | $f_{JK}$ | $f_{J\bullet}$ |
| | $\Sigma$ | $f_{\bullet 1}$ | $f_{\bullet 2}$ | … | $f_{\bullet K}$ | 1 |

$$f_{j\bullet} = \sum_{k=1}^{K} f_{jk}$$

$$f_{\bullet k} = \sum_{j=1}^{J} f_{jk}$$

$$1 = \sum_{j=1}^{J} \sum_{k=1}^{K} f_{jk}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

**Gemeinsame relative Häufigkeitsverteilung $f_{XY}$ von $X$ und $Y$**

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | Σ |
| X | x(1) | $f_{11}$ | $f_{12}$ | ... | $f_{1K}$ | $f_{1\bullet}$ |
| | x(2) | $f_{21}$ | $f_{22}$ | ... | $f_{2K}$ | $f_{2\bullet}$ |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | $f_{J1}$ | $f_{J2}$ | ... | $f_{JK}$ | $f_{J\bullet}$ |
| | Σ | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | 1 |

$$\boxed{f_{XY} = \{f_{jk} \mid j = 1,\dots,J;\ \ k = 1,\dots,K\}}$$

$$f_{j\bullet} = \sum_{k=1}^{K} f_{jk}$$

$$f_{\bullet k} = \sum_{j=1}^{J} f_{jk}$$

$$1 = \sum_{j=1}^{J} \sum_{k=1}^{K} f_{jk}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | Σ |
| X | x(1) | $f_{11}$ | $f_{12}$ | … | $f_{1K}$ | $f_{1\bullet}$ |
| | x(2) | $f_{21}$ | $f_{22}$ | … | $f_{2K}$ | $f_{2\bullet}$ |
| | … | … | … | … | … | … |
| | x(J) | $f_{J1}$ | $f_{J2}$ | … | $f_{JK}$ | $f_{J\bullet}$ |
| | Σ | $f_{\bullet 1}$ | $f_{\bullet 2}$ | … | $f_{\bullet K}$ | 1 |

$$f_{X\bullet} = \{f_{j\bullet} \mid j = 1,\dots,J\}$$

**Relative Randhäufigkeitsverteilung**
$f_{X\bullet}$ von X

**Relative Randhäufigkeitsverteilung**
$f_{\bullet Y}$ von Y

$$f_{\bullet Y} = \{f_{\bullet k} \mid k = 1,\dots,K\}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

Wie lautet die Verteilung von $Y$ im Teildatensatz, für den $X = x(2)$ gilt?

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | Σ |
| X | x(1) | | | | | |
| | x(2) | $N_{21}$ | $N_{22}$ | … | $N_{2K}$ | $N_{2\bullet}$ |
| | … | | | | | |
| | x(J) | | | | | |
| | | | | | | |

Dieser Datensatz hat Umfang $N_{2\bullet}$

Absolute Häufigkeitsverteilung:

$N_{y;k|2}=N_{2k}$ , k=1,…,K

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

Wie lautet die Verteilung von $Y$ im Teildatensatz, für den $X = x(2)$ gilt?

| | | Y | | | | Σ |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | Σ |
| X | x(1) | | | | | |
| | x(2) | $N_{21}/N_{2\cdot}$ | $N_{22}/N_{2\cdot}$ | … | $N_{2K}/N_{2\cdot}$ | $N_{2\cdot}/N_{2\cdot}$ |
| | … | | | | | |
| | x(J) | | | | | |
| | | | | | | |

Dieser Datensatz hat Umfang $N_{2\bullet}$

Relative Häufigkeitsverteilung:

$f_{y;k|2} = N_{y;2k}/N_{2\bullet} = f_{2k}/f_{2\bullet}$, k=1,...,K

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

Bedingte Verteilung von $Y$ gegeben $X$, d.h. von $Y|X = k$

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | Σ |
| X | x(1) | $f_{11}/f_{1\bullet}$ | $f_{12}/f_{1\bullet}$ | … | $f_{1K}/f_{1\bullet}$ | 1 |
| | x(2) | $f_{21}/f_{2\bullet}$ | $f_{22}/f_{2\bullet}$ | … | $f_{2K}/f_{2\bullet}$ | 1 |
| | … | … | … | … | … | … |
| | x(J) | $f_{J1}/f_{J\bullet}$ | $f_{J2}/f_{J\bullet}$ | … | $f_{JK}/f_{J\bullet}$ | 1 |
| | Σ | | | | | J |

**Bedingte Verteilung**
von Y gegeben X=x(2)

$$f_{y;k|2} = f_{2k}/f_{2\bullet}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Darstellung einer bivariaten Häufigkeitsverteilung

**Kontingenztafel**

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | … | y(K) | Σ |
| X | x(1) | $f_{11}/f_{1\bullet}$ | $f_{12}/f_{1\bullet}$ | … | $f_{1K}/f_{1\bullet}$ | 1 |
| | x(2) | $f_{21}/f_{2\bullet}$ | $f_{22}/f_{2\bullet}$ | … | $f_{2K}/f_{2\bullet}$ | 1 |
| | … | … | … | … | … | … |
| | x(J) | $f_{J1}/f_{J\bullet}$ | $f_{J2}/f_{J\bullet}$ | … | $f_{JK}/f_{J\bullet}$ | 1 |
| | Σ | | | | | J |

**Bedingte Verteilung**
$f_{Y|X}$ von Y gegeben X

$$f_{y;k|j} = f_{jk}/f_{j\bullet}$$

$$\boxed{f_{Y|X} = \{f_{y;k|j} \mid j = 1,\ldots,J;\ \ k = 1,\ldots,K\}}$$

**Bedingte Verteilung**
$f_{X|Y}$ von X gegeben Y

$$f_{x;j|k} = f_{jk}/f_{\bullet k}$$

$$\boxed{f_{X|Y} = \{f_{x;j|k} \mid j = 1,\ldots,J;\ \ k = 1,\ldots,K\}}$$

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

| Be-arbeiter(in) | | Aufgabe |
|---|---|---|
| Kai | | Export |
| Kai | | Verknüpfung |
| Miriam | | Export |
| Tina | | Verknüpfung |
| Oliver | | Export |
| Tina | | Export |
| Tina | | Verknüpfung |
| Miriam | | Export |
| Miriam | | Export |
| Oliver | | Abfrage |
| Oliver | | Verknüpfung |
| Oliver | | Abfrage |

Absolute Häufigkeiten

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bear-bei-ter(in) | Kai | 0 | 1 | 1 | 2 |
| | Miriam | 0 | 3 | 0 | 3 |
| | Oliver | 2 | 1 | 1 | 4 |
| | Tina | 0 | 1 | 2 | 3 |
| | Σ | 2 | 6 | 4 | 12 |

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

| Be-arbeiter(in) | | Aufgabe |
|---|---|---|
| Kai | | Export |
| Kai | | Verknüpfung |
| Miriam | | Export |
| Tina | | Verknüpfung |
| Oliver | | Export |
| Tina | | Export |
| Tina | | Verknüpfung |
| Miriam | | Export |
| Miriam | | Export |
| Oliver | | Abfrage |
| Oliver | | Verknüpfung |
| Oliver | | Abfrage |

Relative Häufigkeiten

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bear-bei-ter(in) | Kai | 0 | 1/12 | 1/12 | 2/12 |
| | Miriam | 0 | 3/12 | 0 | 3/12 |
| | Oliver | 2/12 | 1/12 | 1/12 | 4/12 |
| | Tina | 0 | 1/12 | 2/12 | 3/12 |
| | Σ | 2/12 | 6/12 | 4/12 | 1 |

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

| Be-arbeiter(in) | | Aufgabe |
|---|---|---|
| Kai | | Export |
| Kai | | Verknüpfung |
| Miriam | | Export |
| Tina | | Verknüpfung |
| Oliver | | Export |
| Tina | | Export |
| Tina | | Verknüpfung |
| Miriam | | Export |
| Miriam | | Export |
| Oliver | | Abfrage |
| Oliver | | Verknüpfung |
| Oliver | | Abfrage |

Relative Häufigkeiten Aufgabe bedingt auf Bearbeiter(in)

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bear-bei-ter(in) | Kai | 0 | 1/2 | 1/2 | 1 |
| | Miriam | 0 | 1 | 0 | 1 |
| | Oliver | 2/4 | 1/4 | 1/4 | 1 |
| | Tina | 0 | 1/3 | 2/3 | 1 |
| | Σ | | | | 4 |

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

| Be-arbeiter(in) | | Aufgabe |
|---|---|---|
| Kai | | Export |
| Kai | | Verknüpfung |
| Miriam | | Export |
| Tina | | Verknüpfung |
| Oliver | | Export |
| Tina | | Export |
| Tina | | Verknüpfung |
| Miriam | | Export |
| Miriam | | Export |
| Oliver | | Abfrage |
| Oliver | | Verknüpfung |
| Oliver | | Abfrage |

Relative Häufigkeiten Bearbeiter(in) bedingt auf Aufgabe

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bear-bei-ter(in) | Kai | 0 | 1/6 | 1/4 | |
| | Miriam | 0 | 1/2 | 0 | |
| | Oliver | 1 | 1/6 | 1/4 | |
| | Tina | 0 | 1/6 | 1/2 | |
| | Σ | 1 | 1 | 1 | 3 |

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Nominale Daten**: Beispiel in R:
*Überlebende der Titanic*

Der **Mosaikplot**

Code in R:

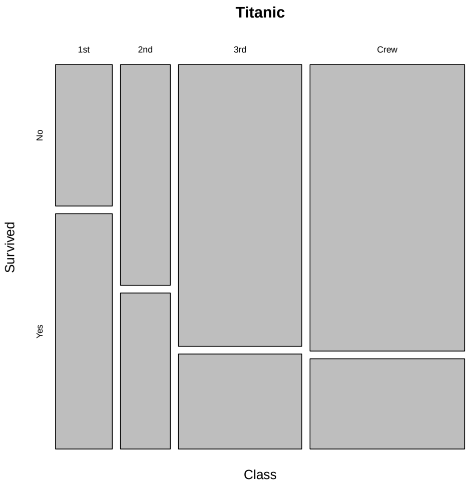
```
mosaicplot(∼ Class + Survived,
data = Titanic)
```

Rechteckbreiten entsprechen $f_{\bullet c}$

Rechteckhöhen entsprechen $f_{s|c}$

Rechteckflächen entsprechen

$f_{sc} = f_{s|c} \cdot f_{\bullet c}$



# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

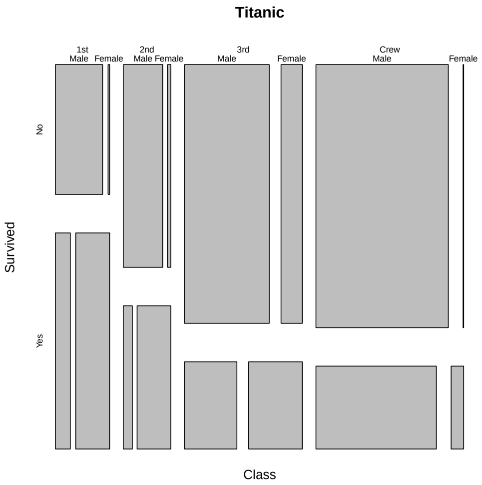
**Nominale Daten**: Beispiel in R:
*Überlebende der Titanic*

Der **Mosaikplot**

Code in R:

```
mosaicplot(∼ Class + Survived +
Sex, data = Titanic)
```

Zusätzliche Einteilung der Flächen nach Geschlecht



# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

Ordinale Daten:

- Kontingenztafeln und Mosaikplots mit geordneten Kategorien

Quantitative Daten:

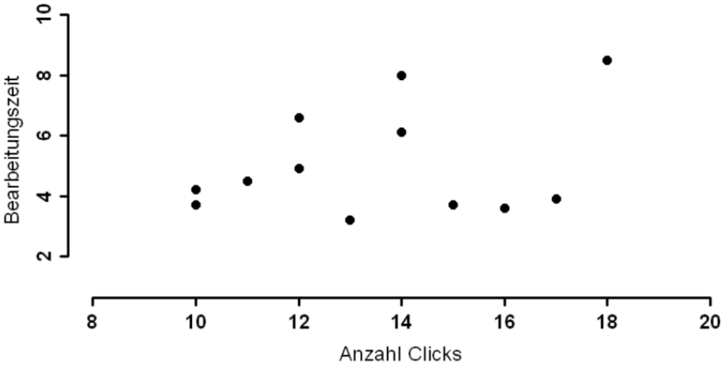
- Kontingenztafeln und Mosaikplots mit klassierten Daten
- Streudiagramme!

# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Quantitative Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

### Streudiagramm

Darstellung der Punktepaare $(x_i, y_i)$ in einem kartesischen Koordinatensystem

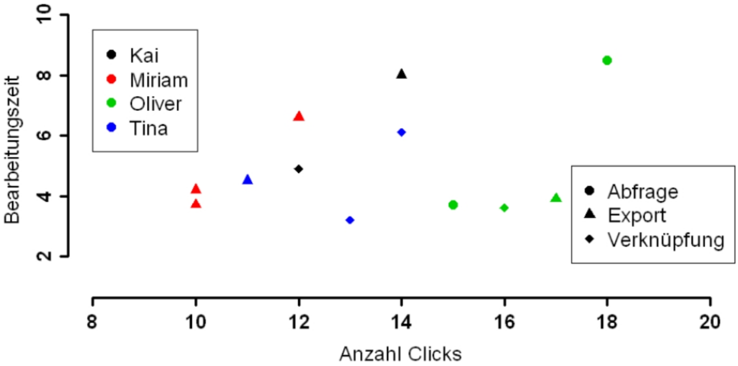


# 4.1 Bivariate Daten: Tabellarische und grafische Darstellungen

**Quantitative Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

## Streudiagramm

Darstellung der Punktepaare $(x_i, y_i)$ in einem kartesischen Koordinatensystem



# 4.2 Bivariate Daten: Zusammenhangsmaße

- Erinnerung: Allgemeine Eigenschaft der Streuung univariater Daten: Streuung von $X$ desto höher, je schlechter sich konkrete Werte vorhersagen lassen.
  - Bisher: Vorhersage der Werte von $X$ durch einzelne Lageparameter.
  - **Jetzt: Vorhersage der Werte von $Y$ unter Verwendung der Werte von $X$.**

- Allgemein: Zusammenhang ( $=$ Korrelation) zwischen $X$ und $Y$ desto größer, je besser sich der Wert von $Y$ unter Kenntnis des Werts von $X$ vorhersagen lässt (oder umgekehrt).

- **Wichtige Unterscheidung**
  - Korrelation bedeutet nicht notwendig Kausalität (Beziehung zwischen *Ursache* und *Wirkung* oder *Aktion* und *Reaktion*)

# 4.2 Bivariate Daten: Zusammenhangsmaße

## Korrelation und Kausalität

Es gilt:

$X$ ist Ursache von $Y \Rightarrow X$ und $Y$ korrelieren

$$\boxed{X} \longrightarrow \boxed{Y}$$
$$\Downarrow$$
$$\boxed{X} \longleftrightarrow \boxed{Y}$$

Aber:

$X$ und $Y$ korrelieren $\not\Rightarrow X$ ist Ursache von $Y$

$$\boxed{X} \longrightarrow \boxed{Y}$$
$$\not\Uparrow \text{ (gilt nicht)}$$
$$\boxed{X} \longleftrightarrow \boxed{Y}$$

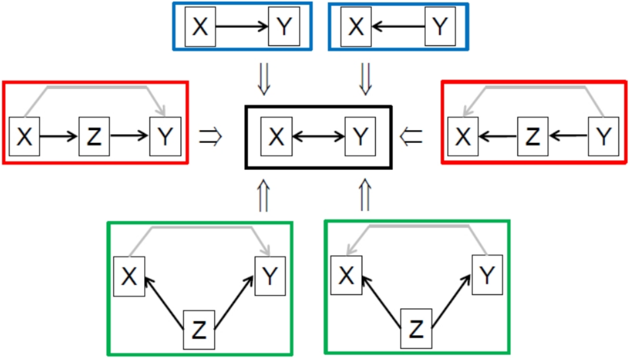
# 4.2 Bivariate Daten: Zusammenhangsmaße

## Korrelation und Kausalität

$X$ ist Ursache von $Y \Rightarrow X$ und $Y$ korrelieren

$X$ und $Y$ korrelieren $\not\Rightarrow X$ ist Ursache von $Y$

Verschiedene Korrelationsquellen möglich



# 4.2 Bivariate Daten: Zusammenhangsmaße

## Nominale Daten

Zusammenhang (=Korrelation) zwischen $X$ und $Y$ desto größer, je besser sich der Wert von $Y$ unter Kenntnis des Werts von $X$ vorhersagen lässt (oder umgekehrt).

| | | Y | | | | Σ |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | Σ |
| X | x(1) | $f_{y;1\|1}$ | $f_{y;2\|1}$ | ... | $f_{y;K\|1}$ | 1 |
| | x(2) | $f_{y;1\|2}$ | $f_{y;2\|2}$ | ... | $f_{y;K\|2}$ | 1 |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | $f_{y;1\|J}$ | $f_{y;2\|J}$ | ... | $f_{y;K\|J}$ | 1 |
| | | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | |

Wert von Y lässt sich bei Kenntnis von X umso besser vorhersagen, je stärker die bedingte Verteilung $f_{Y|X}$ von Y gegeben X von der Randverteilung $f_{\bullet Y}$ von Y abweicht.

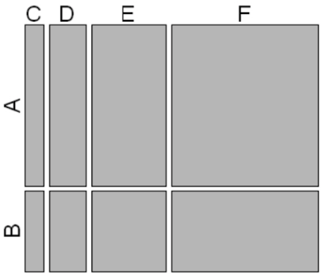
# 4.2 Bivariate Daten: Zusammenhangsmaße

## Nominale Daten

Wert von $Y$ lässt sich bei Kenntnis von $X$ umso besser vorhersagen, je stärker die bedingte Verteilung $f_{Y|X}$ von $Y$ gegeben $X$ von der Randverteilung $f_{\bullet Y}$ von $Y$ abweicht.

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | Σ |
| X | x(1) | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | 1 |
| | x(2) | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | 1 |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | 1 |
| | | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | |

Zusammenhang minimal, falls $f_{y;k|j} = f_{\bullet j}$ für alle $j \in \{1, \ldots, J\}$ und $k \in \{1, \ldots, K\}$



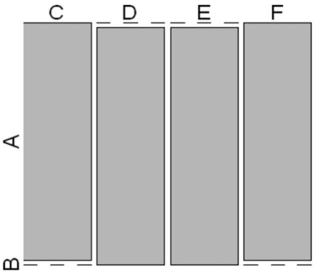
# 4.2 Bivariate Daten: Zusammenhangsmaße

## Nominale Daten

Wert von $Y$ lässt sich bei Kenntnis von $X$ umso besser vorhersagen, je stärker die bedingte Verteilung $f_{Y|X}$ von $Y$ gegeben $X$ von der Randverteilung $f_{\bullet Y}$ von $Y$ abweicht.

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | Σ |
| X | x(1) | 0 | 1 | ... | 0 | 1 |
| | x(2) | 0 | 0 | ... | 1 | 1 |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | 1 | 0 | ... | 0 | 1 |
| | | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | |

Zusammenhang maximal, falls es für alle $j \in \{1, \dots, J\}$ ein $k \in \{1, \dots, K\}$ mit $f_{y;k|j} = 1$ gibt



# 4.2 Bivariate Daten: Zusammenhangsmaße

## Nominale Daten

Wert von $Y$ lässt sich bei Kenntnis von $X$ umso besser vorhersagen, je stärker die bedingte Verteilung $f_{Y|X}$ von $Y$ gegeben $X$ von der Randverteilung $f_{\bullet Y}$ von $Y$ abweicht.

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | Σ |
| X | x(1) | $f_{y;1\|1}$ | $f_{y;2\|1}$ | ... | $f_{y;K\|1}$ | 1 |
| | x(2) | $f_{y;1\|2}$ | $f_{y;2\|2}$ | ... | $f_{y;K\|2}$ | 1 |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | $f_{y;1\|J}$ | $f_{y;2\|J}$ | ... | $f_{y;K\|J}$ | 1 |
| | | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | |

Ein Maß, dass desto größer wird, je größer die Abweichung der bedingten Verteilung $f_{Y|X}$ von der Randverteilung $f_{\bullet Y}$ ist, ist also ein sinnvolles Zusammenhangsmaß.

# 4.2 Bivariate Daten: Zusammenhangsmaße

## Nominale Daten

Ein Maß, dass desto größer wird, je größer die Abweichung der bedingten Verteilung $f_{Y|X}$ von der Randverteilung $Y_{\bullet Y}$ ist, ist also ein sinnvolles Zusammenhangsmaß.

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | Σ |
| X | x(1) | $f_{0;11}$ | $f_{0;12}$ | ... | $f_{0;1K}$ | $f_{1\bullet}$ |
| | x(2) | $f_{0;21}$ | $f_{0;22}$ | ... | $f_{0;2K}$ | $f_{2\bullet}$ |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | $f_{0;J1}$ | $f_{0;J2}$ | ... | $f_{0;JK}$ | $f_{J\bullet}$ |
| | Σ | $f_{\bullet 1}$ | $f_{\bullet 2}$ | ... | $f_{\bullet K}$ | 1 |

Wären bedingte und Randverteilung identisch, so würde ein Anteil von von $f_{0;jk} = f_{\bullet k} \cdot f_{j\bullet}$ an den N Daten in Kategorie (x(j), y(k)) fallen.

Dieser Fall wird als **empirische Unabhängigkeit** von X und Y bezeichnet.

# 4.2 Bivariate Daten: Zusammenhangsmaße

## Nominale Daten

Ein Maß, dass desto größer wird, je größer die Abweichung der bedingten Verteilung $f_{Y|X}$ von der Randverteilung $Y_{\bullet Y}$ ist, ist also ein sinnvolles Zusammenhangsmaß.

| | | Y | | | | Σ |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | |
| X | x(1) | $v_{11}$ | $v_{12}$ | ... | $v_{1K}$ | $N_{1\bullet}$ |
| | x(2) | $v_{21}$ | $v_{22}$ | ... | $v_{2K}$ | $N_{2\bullet}$ |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | $v_{J1}$ | $v_{J2}$ | ... | $v_{JK}$ | $N_{J\bullet}$ |
| | Σ | $N_{\bullet 1}$ | $N_{\bullet 2}$ | ... | $N_{\bullet K}$ | N |

Somit würden bei Unabhängigkeit

$$v_{jk} = f_{\bullet k} \cdot f_{j\bullet} \cdot N = \frac{N_{\bullet k} \cdot N_{j\bullet} \cdot N}{N \cdot N} = \frac{N_{\bullet k} \cdot N_{j\bullet}}{N}$$

Beobachtungen in Kategorie (x(j), x(k)) erwartet.

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Nominale Daten** Je größer die beobachteten Anzahlen $N_{jk}$ von den erwarteten $v_{jk}$ abweichen, desto mehr unterscheiden sich bedingte und Randverteilungen. Ein Maß, dass auf der quadratischen Abweichung der erwarteten von den beobachteten Häufigkeiten basiert, ist die **$\chi^2$-Größe**

$$\chi^2 = \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{(N_{jk} - v_{jk})^2}{v_{jk}}, \quad v_{jk} = \frac{N_{j\bullet} N_{\bullet k}}{N}$$

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | y(1) | y(2) | ... | y(K) | Σ |
| | x(1) | $(N_{11}-v_{11})^2$ | $(N_{12}-v_{12})^2$ | ... | $(N_{1K}-v_{1K})^2$ | $N_{1\bullet}$ |
| X | x(2) | $(N_{21}-v_{21})^2$ | $(N_{22}-v_{22})^2$ | ... | $(N_{2K}-v_{2K})^2$ | $N_{2\bullet}$ |
| | ... | ... | ... | ... | ... | ... |
| | x(J) | $(N_{J1}-v_{J1})^2$ | $(N_{J2}-v_{J2})^2$ | ... | $(N_{JK}-v_{JK})^2$ | $N_{J\bullet}$ |
| | Σ | $N_{\bullet 1}$ | $N_{\bullet 2}$ | ... | $N_{\bullet K}$ | N |

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Nominale Daten**: die **$\chi^2$-Größe**

$$\chi^2 = \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{(N_{jk} - v_{jk})^2}{v_{jk}}, \quad v_{jk} = \frac{N_{j\bullet}N_{\bullet k}}{N}$$

Die $\chi^2$-Größe erfüllt die Forderung, desto größer zu werden, je größer die Abweichung der bedingten Verteilung $f_{Y|X}$ von der Randverteilung $f_{\cdot Y}$ ist.

$$\boxed{\chi^2} = \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{\left(N_{jk} - \frac{N_{j\bullet}N_{\bullet k}}{N}\right)^2 N}{N_{j\bullet}N_{\bullet k}} = \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{(f_{jk}N - f_{j\bullet}f_{\bullet k}N)^2}{f_{j\bullet}f_{\bullet k}N}$$

$$= \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{N(f_{jk} - f_{j\bullet}f_{\bullet k})^2}{f_{j\bullet}f_{\bullet k}} = \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{Nf_{j\bullet}^2\left(\frac{f_{jk}}{f_{j\bullet}} - f_{\bullet k}\right)^2}{f_{j\bullet}f_{\bullet k}}$$

$$= \boxed{\sum_{j=1}^{J}\sum_{k=1}^{K} \frac{Nf_{j\bullet}(f_{y;k|j} - f_{\bullet k})^2}{f_{\bullet k}}}$$

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Nominale Daten**: die $\chi^2$-**Größe**

$$\chi^2 = \sum_{j=1}^{J}\sum_{k=1}^{K}\frac{(N_{jk}-v_{jk})^2}{v_{jk}} = N\left(\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{N_{jk}^2}{N_{j\bullet}N_{\bullet k}}-1\right),\ v_{jk}=\frac{N_{j\bullet}N_{\bullet k}}{N}$$

Es gilt: $\boxed{0 \leq \chi^2 \leq N(\min\{J,K\}-1)}$

Beweis:

$0 \leq \chi^2$ klar wegen $N_{j\bullet} > 0,\ N_{\bullet k} > 0,\ (N_{jk}-v_{jk})^2 \geq 0$

$0 = \chi^2$, wenn $N_{jk} = v_{jk}$, d.h. wenn alle bedingten Häufigkeiten den unter Unabhängigkeit erwarteten Häufigkeiten entsprechen.

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Nominale Daten**: die $\chi^2$-**Größe**

$$\chi^2 = \sum_{j=1}^{J}\sum_{k=1}^{K}\frac{(N_{jk}-v_{jk})^2}{v_{jk}} = N\Big(\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{N_{jk}^2}{N_{j\bullet}N_{\bullet k}}-1\Big),\ v_{jk}=\frac{N_{j\bullet}N_{\bullet k}}{N}$$

Wann gilt: $\boxed{\chi^2 = N(\min\{J,K\}-1)}$ ?

**Beweisskizze**: Sei o.B.d.A. $K \le J$.

Dann gilt für alle $k = 1,\ldots,K$ und $j = 1,\ldots,J$ mit $N_{jk} > 0$:

$$\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{N_{jk}^2}{N_{j\bullet}N_{\bullet k}} = K \Leftrightarrow \frac{N_{jk}}{N_{j\bullet}} = 1 \quad \text{für ein } k_j,$$

d.h. $\chi^2$ wird maximal, wenn es zu jedem $j$ ein $k_j$ mit $f_{y;k_j|j} = 1$ gibt.

# 4.2 Bivariate Daten: Zusammenhangsmaße

<u>**Nominale Daten**</u>: die $\chi^2$-**Größe**

$$\chi^2 = \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{(N_{jk} - v_{jk})^2}{v_{jk}} = N\left(\sum_{j=1}^{J} \sum_{k=1}^{K} \frac{N_{jk}^2}{N_{j\bullet} N_{\bullet k}} - 1\right), \; v_{jk} = \frac{N_{j\bullet} N_{\bullet k}}{N}$$

Es gilt: $0 \leq \chi^2 \leq N(\min\{J, K\} - 1)$

**(Korrigierter) Kontingenzkoeffizient nach Pearson:**

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N} \frac{\min\{J, K\}}{\min\{J, K\} - 1}} \in [0, 1]$$

Eliminiert Abhängigkeit des Koeffizienten vom Stichprobenumfang $N$ und von der Dimension $\min\{J, K\}$.

# 4.2 Bivariate Daten: Zusammenhangsmaße

<u>**Nominale Daten**</u>: Beispiel **Bearbeitungen von Softwareaufgaben**

$N_{jk}$

| b | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bearbei-ter(in) | Kai | 0 | 1 | 1 | 2 |
| | Miriam | 0 | 3 | 0 | 3 |
| | Oliver | 2 | 1 | 1 | 4 |
| | Tina | 0 | 1 | 2 | 3 |
| | Σ | 2 | 6 | 4 | 12 |

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Nominale Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

$v_{jk}$

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bearbei-ter(in) | Kai | 0<br>2·2/12=**1/3** | 1<br>2·6/12=**1** | 1<br>2·4/12=**2/3** | 2 |
| | Miriam | 0<br>3·2/12=**1/2** | 3<br>3·6/12=**3/2** | 0<br>3·4/12=**1** | 3 |
| | Oliver | 2<br>4·2/12=**2/3** | 1<br>4·6/12=**2** | 1<br>4·4/12=**4/3** | 4 |
| | Tina | 0<br>3·2/12=**1/2** | 1<br>3·6/12=**3/2** | 2<br>3·4/12=**1** | 3 |
| | Σ | 2 | 6 | 4 | 12 |

# 4.2 Bivariate Daten: Zusammenhangsmaße

**<u>Nominale Daten</u>**: Beispiel **Bearbeitungen von Softwareaufgaben**

$(N_{jk} - v_{jk})^2$

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bearbeiter(in) | Kai | 0<br>$(0-1/3)^2$=**1/9** | 1<br>$(1-1)^2$=**0** | 1<br>$(1-2/3)^2$=**1/9** | 2 |
| | Miriam | 0<br>$(0-1/2)^2$=**1/4** | 3<br>$(3-3/2)^2$=**9/4** | 0<br>$(0-1)^2$=**1** | 3 |
| | Oliver | 2<br>$(2-2/3)^2$=**16/9** | 1<br>$(1-2)^2$=**1** | 1<br>$(1-4/3)^2$=**1/9** | 4 |
| | Tina | 0<br>$(0-1/2)^2$=**1/4** | 1<br>$(1-3/2)^2$=**1/4** | 2<br>$(2-1)^2$=**1** | 3 |
| | Σ | 2 | 6 | 4 | 12 |

# 4.2 Bivariate Daten: Zusammenhangsmaße

**<u>Nominale Daten</u>**: Beispiel **Bearbeitungen von Softwareaufgaben**

$(N_{jk} - v_{jk})^2 / v_{jk}$

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bearbei-ter(in) | Kai | 0<br>1·3/(9·1)=**1/3** | 1<br>0/1=**0** | 1<br>1·3/(9·2)=**1/6** | 2 |
| | Miriam | 0<br>1·2/(4·1)=**1/2** | 3<br>9·2/(4·3)=**3/2** | 0<br>1/1=**1** | 3 |
| | Oliver | 2<br>16·3/(9·2)=**8/3** | 1<br>**1/2** | 1<br>1·3/(9·4)=**1/12** | 4 |
| | Tina | 0<br>1·2/(4·1)=**1/2** | 1<br>1·2/(4·3)=**1/6** | 2<br>1/1=**1** | 3 |
| | Σ | 2 | 6 | 4 | 12 |

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Nominale Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

$$\chi^2 = \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{(N_{jk} - v_{jk})^2}{v_{jk}}$$
$$= \frac{1}{12}(4+6+32+6+0+18+6+2+2+12+1+12) = \frac{101}{12} \approx 8.417$$

$(N_{jk} - v_{jk})^2 / v_{jk}$

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bearbei-ter(in) | Kai | 1/3 | 0 | 1/6 | 2 |
| | Miriam | 1/2 | 3/2 | 1 | 3 |
| | Oliver | 8/3 | 1/2 | 1/12 | 4 |
| | Tina | 1/2 | 1/6 | 1 | 3 |
| | Σ | 2 | 6 | 4 | 12 |

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Nominale Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

$$\chi^2 = \frac{101}{12}, \quad C = \sqrt{\frac{\chi^2}{\chi^2 + N} \frac{\min\{J,K\}}{\min\{J,K\}-1}} = \sqrt{\frac{101 \cdot 12}{12 \cdot 245} \cdot \frac{3}{2}} = \sqrt{\frac{303}{490}} \approx 0.786$$

$(N_{jk} - v_{jk})^2 / v_{jk}$

| | | Aufgabe | | | |
|---|---|---|---|---|---|
| | | Abfrage | Export | Verknüpfung | Σ |
| Bearbei-ter(in) | Kai | 1/3 | 0 | 1/6 | 2 |
| | Miriam | 1/2 | 3/2 | 1 | 3 |
| | Oliver | 8/3 | 1/2 | 1/12 | 4 |
| | Tina | 1/2 | 1/6 | 1 | 3 |
| | Σ | 2 | 6 | 4 | 12 |

# 4.2 Bivariate Daten: Zusammenhangsmaße

## Ordinale Daten

Allgemein: Zusammenhang (=Korrelation) zwischen $Y$ und $X$ desto größer, je besser sich der Wert von $Y$ unter Kenntnis des Werts von $X$ vorhersagen lässt (oder umgekehrt).

Wert von $Y$ lässt sich bei Kenntnis von $X$ umso besser vorhersagen, je mehr ein hoher Wert von $X$ einen hohen Wert von $Y$ impliziert (**positiver Zusammenhang**) bzw. je mehr ein hoher Wert von $X$ einen niedrigen Wert von Y impliziert (**negativer Zusammenhang**).

Ein sinnvolles Zusammenhangsmaß für ordinale Daten sollte also im Absolutwert hoch sein, wenn hohe **Ränge** von $X$ mit hohen bzw. niedrigen **Rängen** von $Y$ einhergehen und niedrig, wenn Paare von hohen und hohen, hohen und niedrigen, niedrigen und hohen sowie niedrigen und niedrigen $X$- und $Y$-**Rängen** in gleichem Maße auftreten.

# 4.2 Bivariate Daten: Zusammenhangsmaße

## Quantitative Daten

Allgemein: Zusammenhang (=Korrelation) zwischen $Y$ und $X$ desto größer, je besser sich der Wert von $Y$ unter Kenntnis des Werts von $X$ vorhersagen lässt (oder umgekehrt).

Wert von $Y$ lässt sich bei Kenntnis von $X$ umso besser vorhersagen, je mehr ein hoher **Wert** von $X$ einen hohen **Wert** von $Y$ impliziert (positiver Zusammenhang) bzw. je mehr ein hoher **Wert** von $X$ einen niedrigen **Wert** von Y impliziert (negativer Zusammenhang).

Ein sinnvolles Zusammenhangsmaß für ordinale Daten sollte also im Absolutwert hoch sein, wenn hohe **Werte** von $X$ mit hohen bzw. niedrigen **Werten** von $Y$ einhergehen und niedrig, wenn Paare von hohen und hohen, hohen und niedrigen, niedrigen und hohen sowie niedrigen und niedrigen $X$- und $Y$-**Werten** in gleichem Maße auftreten.

# 4.2 Bivariate Daten: Zusammenhangsmaße

## Quantitative Daten

Allgemein: Zusammenhang (=Korrelation) zwischen $Y$ und $X$ desto größer, je besser sich der Wert von $Y$ unter Kenntnis des Werts von $X$ vorhersagen lässt (oder umgekehrt).

$$\textbf{Kovarianz:} \quad s_{xy} = \frac{1}{N-1} \sum_{n=1}^{N} (x_n - \bar{x})(y_n - \bar{y})$$

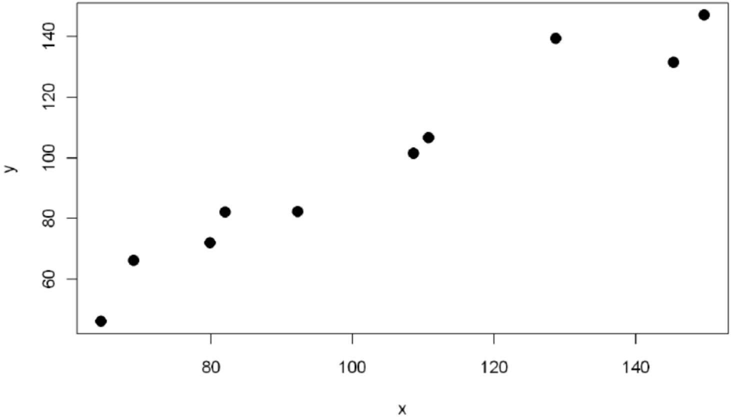
$s_{xy} > 0$, wenn hohe Werte von $X$ in hohem Maße mit hohen Werten von $Y$ einhergehen (Positive Korrelation)

$s_{xy} < 0$, wenn hohe Werte von $X$ in hohem Maße mit niedrigen Werten von $Y$ einhergehen (Negative Korrelation)

$s_{xy} = 0$, wenn hohe Werte von $X$ in gleichem Maße mit hohen Werten wie mit niedrigen Werten von $Y$ einhergehen (Unkorreliertheit)

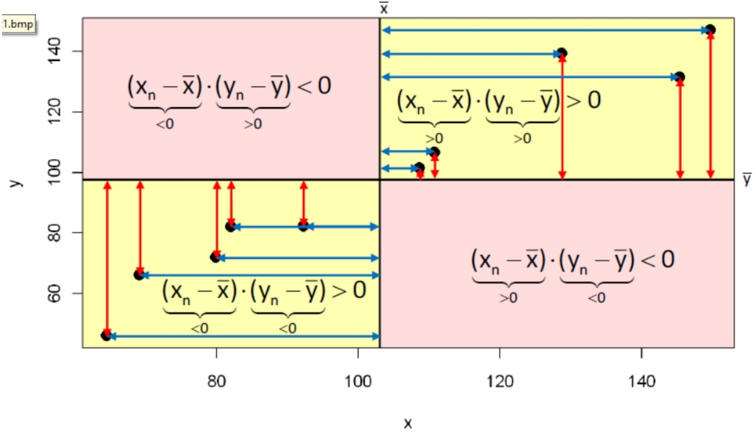
# 4.2 Bivariate Daten: Zusammenhangsmaße

**Quantitative Daten**: **Kovarianz:** $s_{xy} = \frac{1}{N-1}\sum_{n=1}^{N}(x_n - \bar{x})(y_n - \bar{y})$



# 4.2 Bivariate Daten: Zusammenhangsmaße

**Quantitative Daten**: **Kovarianz:** $s_{xy} = \frac{1}{N-1} \sum_{n=1}^{N} \boxed{(x_n - \bar{x})} \boxed{(y_n - \bar{y})}$



# 4.2 Bivariate Daten: Zusammenhangsmaße

**Quantitative Daten**:

**Kovarianz:**

$$s_{xy} = \frac{1}{N-1}\sum_{n=1}^{N}(x_n-\bar{x})(y_n-\bar{y}) = \frac{1}{N-1}\left(\sum_{n=1}^{N} x_n y_n - N\bar{x}\bar{y}\right) = \frac{N}{N-1}(\overline{xy}-\bar{x}\cdot\bar{y})$$

Beweis analog zu Beweis von $d_x^2 = \overline{x^2} - \bar{x}^2$ :

$$\boxed{s_{xy}^2} = \frac{1}{N-1}\sum_{n=1}^{N}(x_n-\bar{x})(y_n-\bar{y}) = \frac{1}{N-1}\sum_{n=1}^{N}(x_n y_n - x_n\bar{y} - \bar{x}y_n + \bar{x}\cdot\bar{y})$$

$$= \frac{1}{N-1}\sum_{n=1}^{N} x_n y_n - \frac{1}{N-1}\left(\sum_{n=1}^{N} x_n\right)\bar{y} - \bar{x}\frac{1}{N-1}\left(\sum_{n=1}^{N} y_n\right) + \frac{N}{N-1}\bar{x}\cdot\bar{y}$$

$$= \frac{N}{N-1}\overline{xy} - \frac{N}{N-1}\bar{x}\cdot\bar{y} - \frac{N}{N-1}\bar{x}\cdot\bar{y} + \frac{N}{N-1}\bar{x}\cdot\bar{y} = \boxed{\frac{N}{N-1}(\overline{xy}-\bar{x}\cdot\bar{y})} \quad \square$$

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Quantitative Daten**: **Kovarianz:** $-s_x s_y \leq s_{xy} \leq s_x s_y$

Beweis: Spezialfall der Cauchy-Schwarz-Ungleichung:

Für $(a_n, b_n) \in \mathbb{R}^2$ gilt:

$$\left(\sum_{n=1}^{N} a_n b_n\right)^2 \leq \sum_{n=1}^{N} a_n^2 \cdot \sum_{n=1}^{N} b_n^2 \Longrightarrow \left(\sum_{n=1}^{N} (x_n - \bar{x})(y_n - \bar{y})\right)^2 \leq \sum_{n=1}^{N} (x_n - \bar{x})^2 \cdot \sum_{n=1}^{N} (y_n - \bar{y})^2$$

$$\Leftrightarrow -\sqrt{\sum_{n=1}^{N} (x_n - \bar{x})^2 \cdot \sum_{n=1}^{N} (y_n - \bar{y})^2} \leq \left(\sum_{n=1}^{N} (x_n - \bar{x})(y_n - \bar{y})\right) \leq \sqrt{\sum_{n=1}^{N} (x_n - \bar{x})^2 \cdot \sum_{n=1}^{N} (y_n - \bar{y})^2}$$

$$\Leftrightarrow -\sqrt{\frac{\sum_{n=1}^{N} (x_n - \bar{x})^2}{N-1}} \sqrt{\frac{\sum_{n=1}^{N} (y_n - \bar{y})^2}{N-1}} \leq \frac{\left(\sum_{n=1}^{N} (x_n - \bar{x})(y_n - \bar{y})\right)}{N-1} \leq \sqrt{\frac{\sum_{n=1}^{N} (x_n - \bar{x})^2}{N-1}} \sqrt{\frac{\sum_{n=1}^{N} (y_n - \bar{y})^2}{N-1}}$$

$$\Leftrightarrow -s_x s_y \leq s_{xy} \leq s_x s_y \quad \square$$

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Quantitative Daten: Korrelationskoeffizient nach Bravais-Pearson**

$$r_{xy} = \frac{s_{xy}}{s_x s_y} \qquad -s_x s_y \leq s_{xy} \leq s_x s_y \Longrightarrow -1 \leq r_{xy} \leq 1$$

Gleichheitsbedingung bei der Cauchy-Schwarz-Ungleichung:

Für $(a_n, b_n) \in \mathbb{R}^2$ gilt:

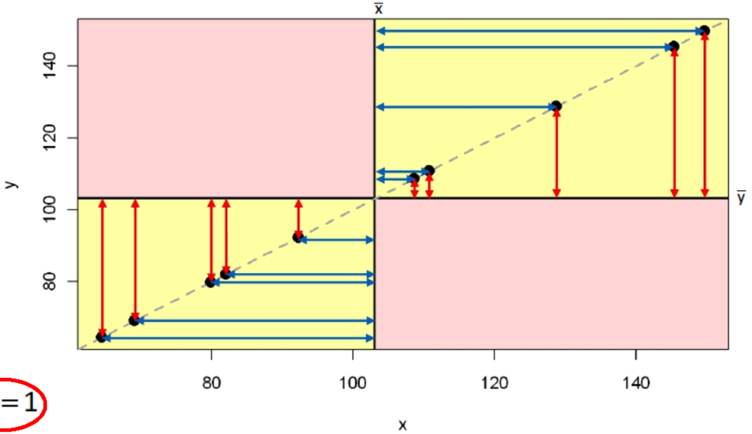
$$\left(\sum_{n=1}^{N} a_n b_n\right)^2 = \sum_{n=1}^{N} a_n^2 \cdot \sum_{n=1}^{N} b_n^2 \Leftrightarrow \text{es gibt eine Konstante } d \text{ mit } b_n = d \cdot a_n \ \forall n$$

$$\begin{aligned} \Longrightarrow r_{xy} \in \{-1, 1\} &\Leftrightarrow (y_n - \bar{y}) = d \cdot (x_n - \bar{x}) \\ &\Leftrightarrow y_n = c + d \cdot x_n \quad \text{mit } c = \bar{y} - d\bar{x} \end{aligned}$$

Das heißt, $|r_{xy}|$ ist genau dann 1, wenn alle $x_n$ und $y_n$ auf einer Geraden liegen.

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Quantitative Daten**: **Kovarianz:** $s_{xy} = \frac{1}{N-1} \sum_{n=1}^{N} \boxed{(x_n - \bar{x})} \boxed{(c + dx_n - c + d\bar{x})}$

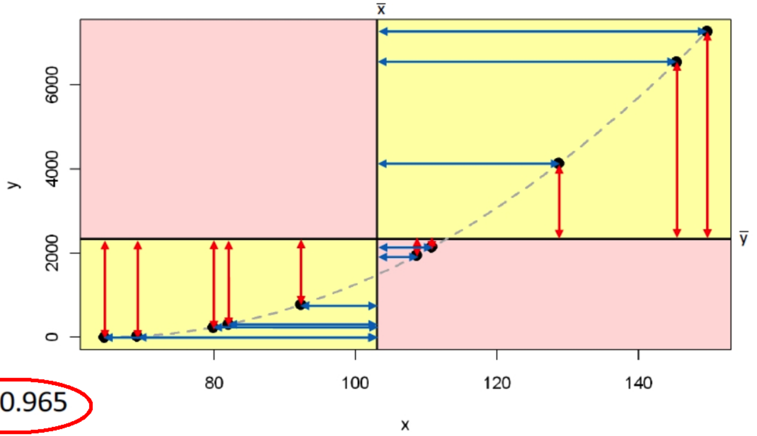



$r_{xy} = 1$

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Quantitative Daten: Korrelationskoeffizient nach Bravais-Pearson**

Nicht-linearer monotoner Zusammenhang

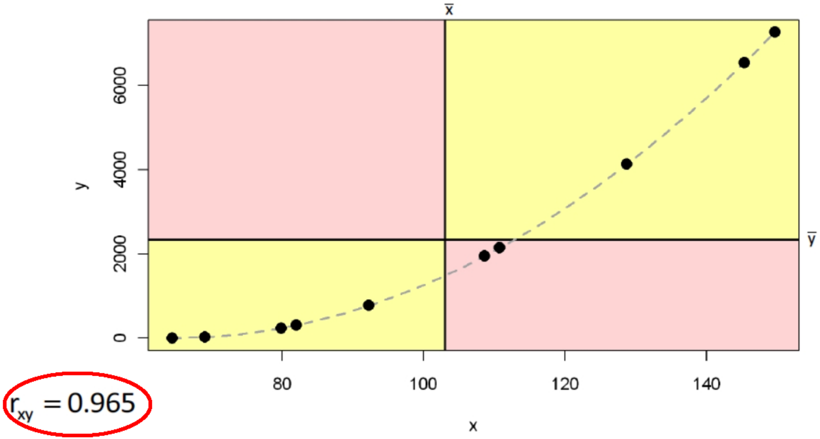



$r_{xy} = 0.965$

# 4.2 Bivariate Daten: Zusammenhangsmaße

**Ordinale/Quantitative Daten**: Nicht-linearer monotoner Zusammenhang

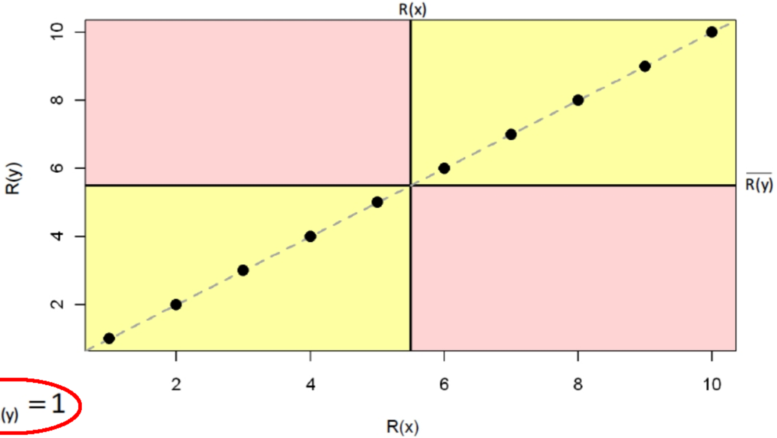
Übergang zu Rängen



# 4.2 Bivariate Daten: Zusammenhangsmaße

**Ordinale/Quantitative Daten**: Nicht-linearer monotoner Zusammenhang

Übergang zu Rängen

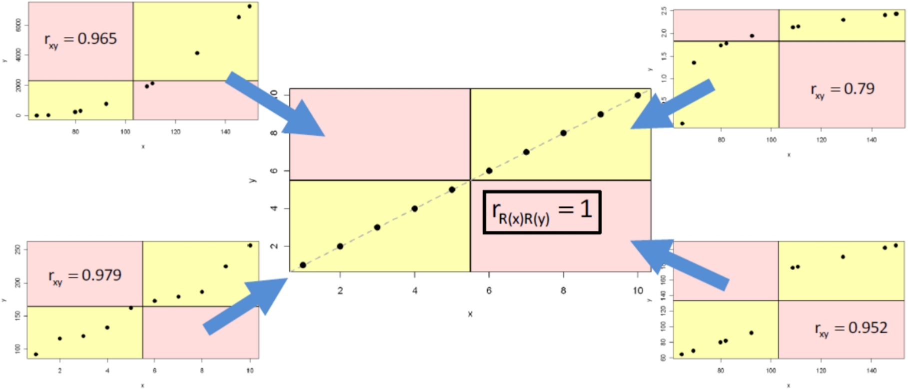



$r_{R(x)R(y)} = 1$

# 4.2 Bivariate Daten: Zusammenhangsmaße

## Ordinale/Quantitative Daten

Absolute Korrelation von Rängen bei monotonem Zusammenhang immer 1



# 4.2 Bivariate Daten: Zusammenhangsmaße

## Ordinale/Quantitative Daten

Falls $X$ und $Y$ mindestens ordinales Skalenniveau haben, so wird der Bravais-Pearson-Korrelationskoeffizient der Ränge $R(X)$ und $R(X)$ von $X$ und $Y$ der **Spearman'sche Rangkorrelationskoeffizient** $r_{xy}^{\mathrm{Sp}}$ von $X$ und $Y$ genannt:

$$r_{xy}^{\mathrm{Sp}} = r_{R(X)R(Y)} = \frac{s_{R(X)R(Y)}}{s_{R(X)}s_{R(Y)}} = \frac{\sum_{n=1}^{N}\left(R(x_n) - \overline{R(x)}\right)\left(R(y_n) - \overline{R(y)}\right)}{\sqrt{\sum_{n=1}^{N}\left(R(x_n) - \overline{R(x)}\right)^2 \sum_{n=1}^{N}\left(R(y_n) - \overline{R(y)}\right)^2}}$$

# 4.2 Bivariate Daten: Zusammenhangsmaße

## Ordinale/Quantitative Daten

### Spearman'scher Rangkorrelationskoeffizient

Falls keine Bindungen auftreten, d.h. $R(x_j) \neq R(x_k)$ und $R(y_j) \neq R(y_k)$ für alle $j \neq k$, so gilt:

$$r_{xy}^{\mathrm{Sp}} = 1 - \frac{6}{N(N^2-1)} \sum_{n=1}^{N} \big(R(x_n) - R(y_n)\big)^2$$

Beweisansatz:

$$\sum_{n=1}^{N} R(x_n) = \sum_{n=1}^{N} R(y_n) = \sum_{n=1}^{N} n = \frac{N(N+1)}{2}$$

$$\text{und} \sum_{n=1}^{N} R(x_n)^2 = \sum_{n=1}^{N} R(y_n)^2 = \sum_{n=1}^{N} n^2 = \frac{N(N+1)(2N+1)}{6}$$

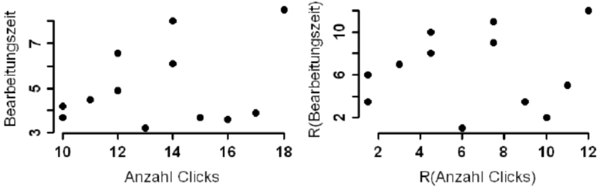
# 4.2 Bivariate Daten: Zusammenhangsmaße

**Ordinale/Quantitative Daten**: Beispiel **Bearbeitungen von Softwareaufgaben**

| Anzahl Clicks | Rang | Bearbeitungszeit | Rang |
|---|---|---|---|
| 14 | 7.5 | 8.0 | 11 |
| 12 | 4.5 | 4.9 | 8 |
| 12 | 4.5 | 6.6 | 10 |
| 13 | 6 | 3.2 | 1 |
| 17 | 11 | 3.9 | 5 |
| 11 | 3 | 4.5 | 7 |
| 14 | 7.5 | 6.1 | 9 |
| 10 | 1.5 | 3.7 | 3.5 |
| 10 | 1.5 | 4.2 | 6 |
| 18 | 12 | 8.5 | 12 |
| 16 | 10 | 3.6 | 2 |
| 15 | 9 | 3.7 | 3.5 |

$\bar{x}_4 = 13.5$, $s^2_{x_4} = 7$

$\bar{x}_5 = 5.075$, $s^2_{x_5} = 3.24$




$$\boxed{r_{x_4x_5}} = [(0.5\cdot 2.925)+(-1.5\cdot -0.175)+(-1.5\cdot 1.525)+(-0.5\cdot -1.875)+(3.5\cdot -1.175)+(-2.5\cdot -0.575)+(0.5\cdot 1.025)+(-3.5\cdot -1.375)+(-3.5\cdot -0.875)+(4.5\cdot 3.425)+(2.5\cdot -1.475)+(1.5\cdot -1.375)]/(11\cdot\sqrt{7\cdot 3.24}) = \boxed{0.301}$$

$$\boxed{r^{Sp}_{x_4x_5}} = [(1\cdot 4.5)+(-2\cdot 1.5)+(-2\cdot 3.5)+(-0.5\cdot -5.5)+(4.5\cdot -1.5)+(-3.5\cdot 0.5)+(1\cdot 2.5)+(-5\cdot -3)+(-5\cdot -0.5)+(5.5\cdot 5.5)+(3.5\cdot -4.5)+(2.5\cdot -3)]/(39.4525) = \boxed{0.111}$$

# 4.3 Bivariate Daten: Lineare Regression

## Korrelation und Linearität:

Der Korrelationskoeffizient ist auch deshalb so beliebt, weil er ein *Maß für die Linearität eines Zusammenhangs* darstellt.

- Es gilt $r_{xy} = 1$, genau wenn die Punkte $(x_i, y_i)$ auf einer Geraden liegen, und es gilt $r_{xy} = 0$, wenn keine lineare Beziehung besteht.
- Um den Grad der Linearität eines Zusammenhangs quantifizieren zu können, ist es notwendig, sich auf ein Optimalitätskriterium zu einigen, nach dem man eine „optimal an die Punkte angepasste Gerade“ bestimmt.
- Das beliebteste Kriterium ist das Prinzip der Kleinsten Quadrate, nach dem die Gerade so bestimmt wird, dass die Quadratsumme derjenigen Abstände der Punkte von der Geraden minimal werden, die senkrecht zu der $x$-Achse gemessen werden.

# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: Erinnerung

Allgemein: Zusammenhang (=Korrelation) zwischen $Y$ und $X$ desto größer, je besser sich der Wert von $Y$ unter Kenntnis des Werts von $X$ **vorhersagen** lässt (oder umgekehrt).

Bravais-Pearson-Korrelationskoeffizient misst linearen Zusammenhang.

Wie lässt sich der lineare Zusammenhang zur Vorhersage nutzen?

# 4.3 Bivariate Daten: Lineare Regression

## Quantitative Daten

$|r_{xy}| = 1 \iff y_n = c + dx_n$ für≠ $n = 1, \dots, N$

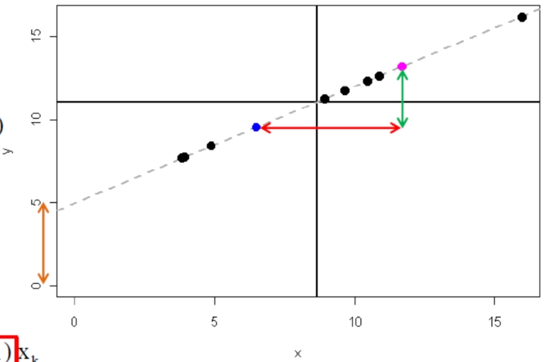
Für beliebiges (j,k) mit j≠k:

$$\boxed{y_j = c + dx_j}$$

$$\boxed{y_k = c + dx_k}$$

$$\Rightarrow \quad y_j - y_k = (c + dx_j) - (c + dx_k)$$
$$= d(x_j - x_k)$$
$$\Leftrightarrow \quad \boxed{d} = \boxed{(y_j - y_k)} / \boxed{(x_j - x_k)}$$

$$y_k = c + dx_k$$
$$\Leftrightarrow \quad \boxed{c} = y_k - dx_k$$
$$= y_k - \boxed{(y_j - y_k)} / \boxed{(x_j - x_k)}\, x_k$$




Perfekte Vorhersage durch Einsetzen in die Gleichung.

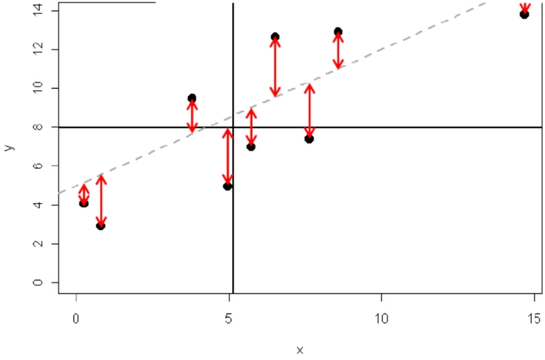
# 4.3 Bivariate Daten: Lineare Regression

## Quantitative Daten

$0 < |r_{xy}| < 1 \iff y_n = c + dx_n + \boxed{\varepsilon_n}$ für n=1,...,N

Vorhersagefehler

$\boxed{\varepsilon_n} = y_n - c - dx_n$



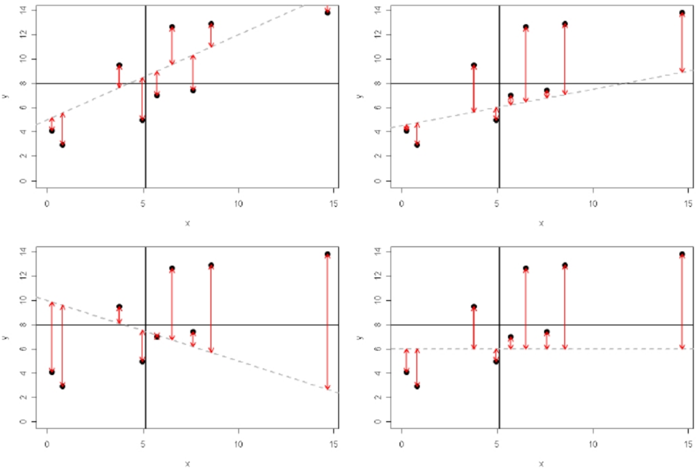
# 4.3 Bivariate Daten: Lineare Regression

## Quantitative Daten: **Methode der kleinsten Quadrate**

Koeffizienten c und d so bestimmen, dass Fehlerquadrat-summe

$$Q(c,d) = \boxed{\sum_{n=1}^{N} \varepsilon_n^2}$$

minimal wird.



# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: **Methode der kleinsten Quadrate**

Die Fehlerquadratsumme $Q(c,d) = \sum_{n=1}^{N} (Y_n - c - dx_n)^2$ ist minimal für

$$d = \frac{s_{xy}}{s_x^2} \quad \text{und} \quad c = \bar{y} - \frac{s_{xy}}{s_x^2}\bar{x}$$

Beweis:

$$\frac{\partial}{\partial c} Q(c,d) = \sum_{n=1}^{N} 2(c + dx_n - y_n) = 2Nc + 2dN\bar{x} - 2N\bar{y} \stackrel{!}{=} 0 \quad \Leftrightarrow c + d\bar{x} - \bar{y} = 0$$

$$\frac{\partial}{\partial d} Q(c,d) = \sum_{n=1}^{N} 2(c + dx_n - y_n)x_n = 2Nc\bar{x} + 2d\sum_{n=1}^{N} x_n^2 - 2\sum_{n=1}^{N} x_n y_n \stackrel{!}{=} 0$$

$$\Leftrightarrow cN\bar{x} + d\sum_{n=1}^{N} x_n^2 - \sum_{n=1}^{N} x_n y_n = 0$$

# 4.3 Bivariate Daten: Lineare Regression

Beweis (Fortsetzung)

$$(1)\ c+d\bar{x}-\bar{y}=0 \Leftrightarrow c=\bar{y}-d\bar{x} \quad (2)\ cN\bar{x}+d\sum_{n=1}^{N}x_n^2-\sum_{n=1}^{N}x_ny_n=0$$

$$(1) \text{ in } (2)\ (\bar{y}-d\bar{x})N\bar{x}+d\sum_{n=1}^{N}x_n^2-\sum_{n=1}^{N}x_ny_n=0$$

$$\Leftrightarrow d\left(\sum_{n=1}^{N}x_n^2-N\bar{x}^2\right)=\sum_{n=1}^{N}x_ny_n-N\bar{x}\cdot\bar{y}$$

$$\Leftrightarrow d=\frac{\sum\limits_{n=1}^{N}x_ny_n-N\bar{x}\cdot\bar{y}}{\left(\sum\limits_{n=1}^{N}x_n^2-N\bar{x}^2\right)}=\frac{\frac{N}{N-1}(\overline{xy}-\bar{x}\cdot\bar{y})}{\frac{N}{N-1}\left(\frac{1}{N}\sum\limits_{n=1}^{N}x_n^2-\bar{x}^2\right)}=\frac{s_{xy}}{s_x^2}\ (3)$$

$$(3) \text{ in } (1)\ c=y-\frac{s_{xy}}{s_x^2}\bar{x}$$

# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: **Methode der kleinsten Quadrate**

Beweis (Fortsetzung)

$$\frac{\partial}{\partial c}Q(c,d) = 2Nc + 2dN\bar{x} - 2N\bar{y}, \quad \frac{\partial}{\partial d}Q(c,d) = 2Nc\bar{x} + 2d\sum_{n=1}^{N}x_n^2 - 2\sum_{n=1}^{N}x_n y_n$$

$$\frac{\partial^2}{\partial c \partial c}Q(c,d) = 2N, \; \frac{\partial^2}{\partial c \partial d}Q(c,d) = 2\sum_{n=1}^{N}x_n, \; \frac{\partial^2}{\partial d \partial d}Q(c,d) = 2\sum_{n=1}^{N}x_n^2$$

$$\det\begin{bmatrix} 2N & 2\sum\limits_{n=1}^{N}x_n \\ 2\sum\limits_{n=1}^{N}x_n & 2\sum\limits_{n=1}^{N}x_n^2 \end{bmatrix} = 4N\sum_{n=1}^{N}x_n^2 - 4\left(\sum_{n=1}^{N}x_n\right)^2 = 4(N-1)Ns_x^2 > 0 \quad \square$$

# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: **Methode der kleinsten Quadrate**

Je größer die absolute Korrelation, desto kleiner die Fehlerquadratsumme

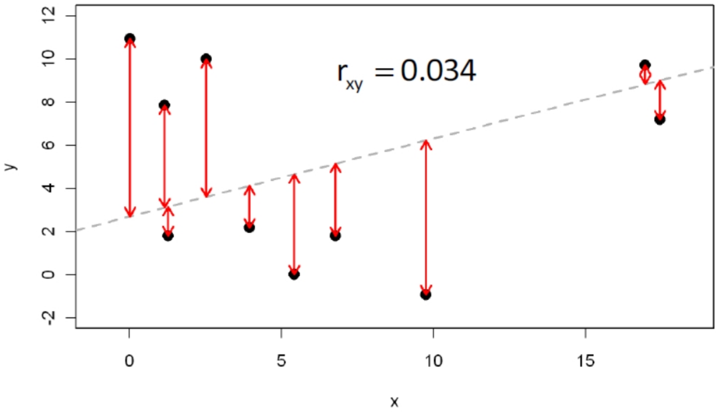
$$d = \frac{s_{xy}}{s_x^2} \text{ und } c = \bar{y} - \frac{s_{xy}}{s_x^2}\bar{x}$$

$$\begin{aligned}
\sum_{n=1}^{N} \varepsilon_n^2 &= \sum_{n=1}^{N} \Big( y_n - (\bar{y} - \frac{s_{xy}}{s_x^2}\bar{x}) - \frac{s_{xy}}{s_x^2} x_n \Big)^2 = \sum_{n=1}^{N} \Big( y_n - (\bar{y} - r_{xy}\frac{s_y}{s_x}\bar{x}) - r_{xy}\frac{s_y}{s_x} x_n \Big)^2 \\
&= \sum_{n=1}^{N} \Big( (y_n - \bar{y}) - r_{xy}\frac{s_y}{s_x}(x_n - \bar{x}) \Big)^2 \\
&= \sum_{n=1}^{N} \Big( (y_n - \bar{y})^2 - 2 r_{xy}\frac{s_y}{s_x}(y_n - \bar{y})(x_n - \bar{x}) + \Big( r_{xy}\frac{s_y}{s_x} \Big)^2 (x_n - \bar{x})^2 \Big) \\
&= (N-1) \cdot \Big( s_y^2 - 2 r_{xy}\frac{s_y}{s_x} s_{xy} + \Big( r_{xy}\frac{s_y}{s_x} \Big)^2 s_x^2 \Big) = (N-1) \cdot (s_y^2 - 2 r_{xy}^2 s_y^2 + r_{xy}^2 s_y^2) \\
&= (N-1) \cdot (s_y^2 - r_{xy}^2 s_y^2) \quad \square
\end{aligned}$$

# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: **Methode der kleinsten Quadrate**

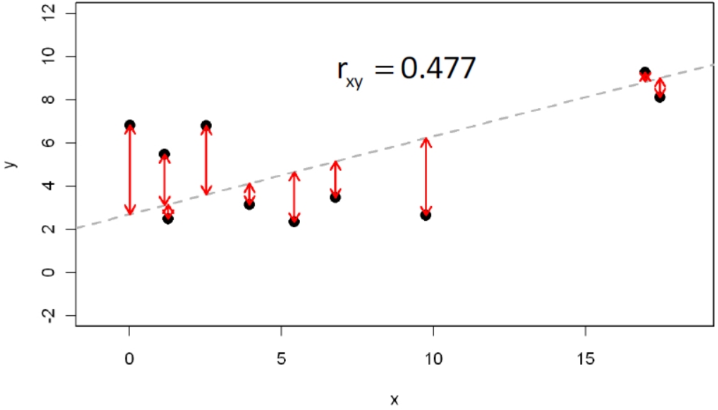
Je größer die absolute Korrelation, desto kleiner die Fehlerquadratsumme



# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: **Methode der kleinsten Quadrate**

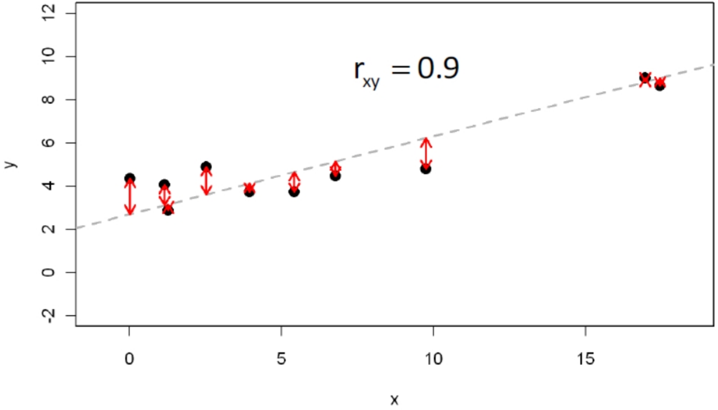
Je größer die absolute Korrelation, desto kleiner die Fehlerquadratsumme



# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: **Methode der kleinsten Quadrate**

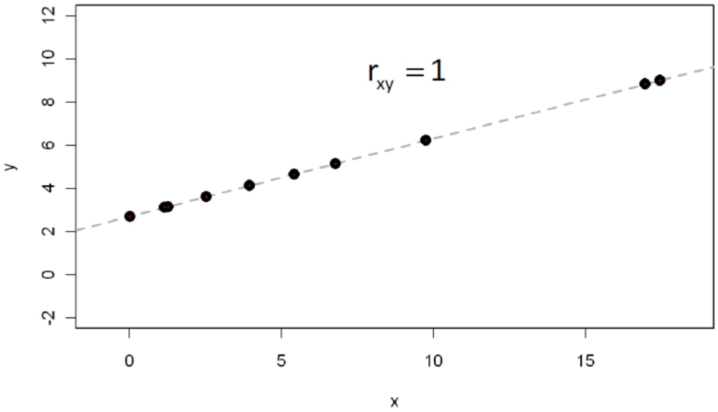
Je größer die absolute Korrelation, desto kleiner die Fehlerquadratsumme



# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: **Methode der kleinsten Quadrate**

Je größer die absolute Korrelation, desto kleiner die Fehlerquadratsumme



# 4.3 Bivariate Daten: Lineare Regression

Quantitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

| Anzahl Clicks | Bear-beitungszeit | c+d$x_4$ | ε |
|---|---|---|---|
| 14 | 8.0 | 5.177 | 2.823 |
| 12 | 4.9 | 4.768 | 0.132 |
| 12 | 6.6 | 4.768 | 1.832 |
| 13 | 3.2 | 4.973 | −1.773 |
| 17 | 3.9 | 5.791 | −1.891 |
| 11 | 4.5 | 4.564 | −0.064 |
| 14 | 6.1 | 5.177 | 0.922 |
| 10 | 3.7 | 4.359 | −0.659 |
| 10 | 4.2 | 4.359 | −0.159 |
| 18 | 8.5 | 5.995 | 2.505 |
| 16 | 3.6 | 5.586 | −1.986 |
| 15 | 3.7 | 5.382 | −1.682 |

$$\bar{x}_4 = 13.5 \qquad s^2_{x_4} = 7$$

$$\bar{x}_5 = 5.075 \qquad s^2_{x_5} = 3.24$$

$$r_{x_4 x_5} = 0.301$$

$$\boxed{x_5 = c + d x_4 + \varepsilon}$$

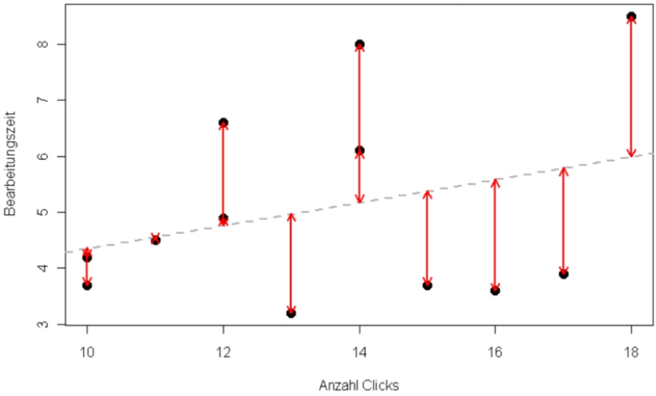
$$\boxed{c} = \bar{x}_5 - r_{x_4 x_5}\frac{s_{x_5}}{s_{x_4}}\bar{x}_4 = 5.075 - 0.301\sqrt{\frac{3.24}{7}}\,13.5 = \boxed{2.314}$$

$$\boxed{d} = r_{x_4 x_5}\frac{s_{x_5}}{s_{x_4}} = 0.301\sqrt{\frac{3.24}{7}} = \boxed{0.205}$$

# 4.3 Bivariate Daten: Lineare Regression

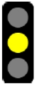
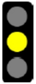
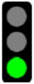
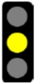
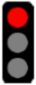
Quantitative Daten: Beispiel **Bearbeitungen von Softwareaufgaben**

$$x_5 = 2.314 + 0.205 x_4 + \boxed{\varepsilon}$$



# 4.3 Bivariate Daten: Lineare Regression

## Zusammenfassung

| Skalenniveau → / ↓Zusammenhangsmaß | Nominal | Ordinal | Quantitativ |
|---|---|---|---|
| χ²-Größe/ Kontingenzkoeffizient nach Pearson |  |  – Informationsverlust |  – Nur für klassierte Daten |
| Rangkorrelationskoeffizient nach Spearman |  – Nur für J = 2 |  |  + Robust<br>+ Allg. Zusammenhang<br>– Informationsverlust |
| Korrelationskoeff. nach Bravais-Pearson/lin. Regr. |  – Nur für J = 2 |  |  – Ausreißeranfällig<br>– Lin. Zusammenhang<br>+ Informationsnutzung |

# Wahrscheinlichkeitstheorie

# 5.0 Wahrscheinlichkeitstheorie

## Wahrscheinlichkeitstheorie

- Teilgebiet der Mathematik, dass sich mit Wahrscheinlichkeiten und der Analyse zufälliger Prozesse beschäftigt
- Mathematische Abstraktion von nicht-deterministischen Ereignissen

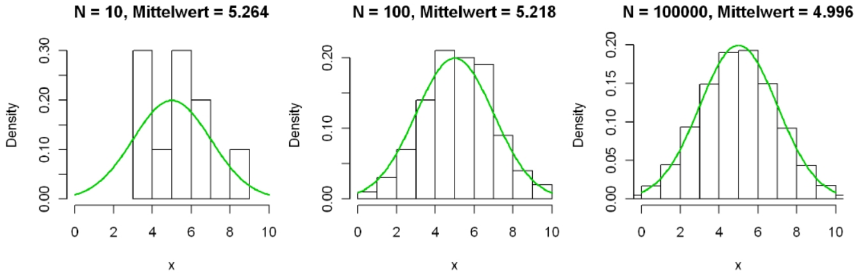
## Unterscheidung von Empirie und Theorie

- Bisher: reine Beschreibung von Lage, Streuung und Zusammenhang von Daten ohne Berücksichtigung ihrer Entstehung
- Jetzt: Interpretation der Daten als Realisationen von Zufallsvariablen und Beschreibung von deren Wahrscheinlichkeitsverteilungen
- Auf Basis dieser Wahrscheinlichkeitsverteilungen lassen sich dann Aussagen über nicht betrachtete oder zukünftige Daten machen

# 5.0 Beschreibung des Zufalls

Vergleich unterschiedlich großer Stichproben aus der gleichen Wahrscheinlichkeitsverteilung

Beispiel Normalverteilung: Gesetz der großen Zahlen



# 5.1 Mengentheoretische Grundlagen

## Elementare Begriffe

| | |
|---|---|
| **Zufallsexperiment** | Datenerhebungsprozess mit nicht vorhersagbarem Ausgang |
| **Ergebnis** $\omega$ | Elementarer Ausgang eines Zufallsexperiments |
| **Grundraum** $\Omega$ | Menge aller möglichen Ergebnisse<br>$\Omega = \{\omega \vert \omega \text{ ist Ergebnis des Zufallsexperiments}\}$ |
| **Ereignis** $A$ | Menge von Ergebnissen, d.h. Teilmenge von $\Omega$ |
| **Elementarereignis** | Einelementiges Ereignis |

# 5.1 Mengentheoretische Grundlagen

Beispiel

| | |
|---|---|
| **Zufallsexperiment** | Einfacher Würfelwurf |
| **Ergebnisse** | $\omega_1 = 1,\ \omega_2 = 2,\ \omega_3 = 3,\ \omega_4 = 4,\ \omega_5 = 5,\ \omega_6 = 6$ |
| **Grundraum** | $\Omega = \{1, 2, 3, 4, 5, 6\}$ |
| **Ereignisse** | $A = \{2, 4, 6\},\ B = \{1, 3, 5\},\ C = \{1, 2, 3, 4, 5\},$<br>$D = \{3, 4, 5, 6\},\ E = \{2, 3, 5\},\ F = \{1, 2, 3, 4, 5, 6\} = \Omega$ |
| **Elementarereignisse** | $\{1\},\ \{2\},\ \{3\},\ \{4\},\ \{5\},\ \{6\}$ |

# 5.1 Mengentheoretische Grundlagen

## Beispiele

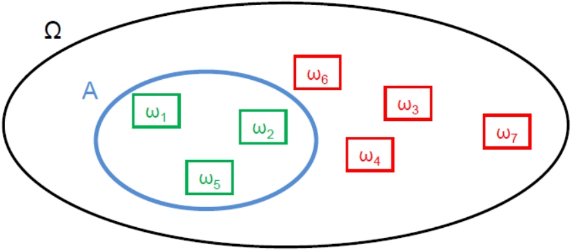
| **Experiment** | **Grundraum** $\Omega$ | **Ergebnis** $\omega$ |
|---|---|---|
| Roulette | $\{0, 1, \ldots, 36\}$ | Zahlenfeld der Kugel |
| Würfeln: Warten auf 6 | $\mathbb{N} \bigcup \{\infty\}$ | Anzahl Würfe bis zur ersten 6 |
| 6 aus 49 | $\{(\omega_1, \ldots, \omega_6) \mid 1 \leq \omega_1 < \ldots < \omega_6 \leq 49\}$ | Geordnete Nummern der gezogenen Kugeln |
| Einzelne Serveranfrage | $[t_{\text{min}}, t_{\text{max}}]$ | Anfragezeitpunkt $t$ |
| Mausaktivität | $\{\omega : [t_{\text{min}}, t_{\text{max}}] \rightarrow (1, \ldots, 600) \times (1, \ldots, 800) \times (0, 1, 2)\}$ | Koordinaten und Clickzustand (nicht, links, rechts) des Mauszeigers zu jeder Zeit |
| Wartezeit bis zur nächsten Serveranfrage | $[0, \infty)$ | Zeit zwischen zwei Anfragen |

# 5.1 Mengentheoretische Grundlagen

## Bezeichnungen

$$\omega \in A$$

Ergebnis ω ist im Ereignis A enthalten

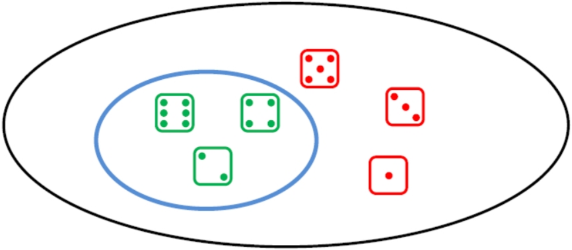


# 5.1 Mengentheoretische Grundlagen

Bezeichnungen: Beispiel Würfelwurf

$$2 \in \{2, 4, 6\}$$

Augenzahl 2 ist gerade Zahl



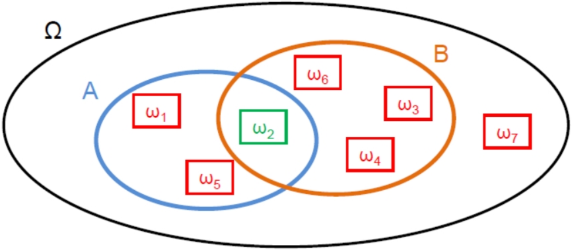
# 5.1 Mengentheoretische Grundlagen

## Bezeichnungen

**Schnittereignis** zweier Mengen

$$A \cap B = \{\omega \in \Omega | \omega \in A \text{ und } \omega \in B\}$$

Ergebnis ω ist in Ereignis A **und** Ereignis B enthalten



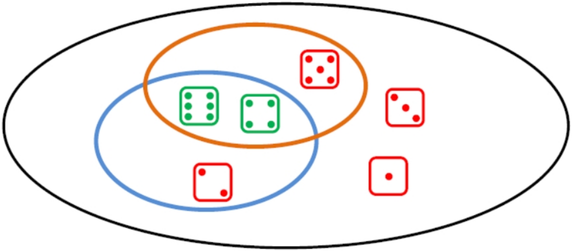
# 5.1 Mengentheoretische Grundlagen

Bezeichnungen: Beispiel Würfelwurf

**Schnittereignis** zweier Mengen

$$4 \in \{2, 4, 6\} \cap \{4, 5, 6\}$$

Augenzahl 4 ist gerade Zahl **und** größer als 3



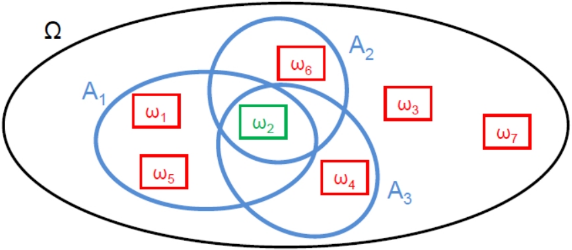
# 5.1 Mengentheoretische Grundlagen

## Bezeichnungen

**Schnittereignis** beliebig vieler Mengen

$$\bigcap_{i \in I} A_i = \{\omega \in \Omega | \omega \in A_i \text{ für } i \in I\}$$

Ergebnis ω ist in allen Ereignissen $A_i$ enthalten



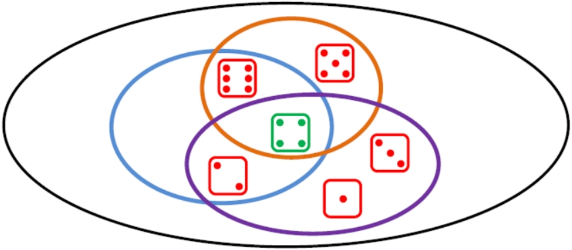
# 5.1 Mengentheoretische Grundlagen

Bezeichnungen: Beispiel Würfelwurf

**Schnittereignis** beliebig vieler Mengen

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

Augenzahl 4 ist gerade Zahl **und** größer als 3 **und** kleiner als 5



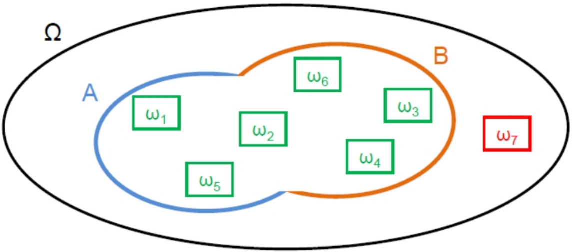
# 5.1 Mengentheoretische Grundlagen

## Bezeichnungen

**Vereinigungsereignis** zweier Mengen

$$A \cup B = \{\omega \in \Omega | \omega \in A \text{ und/oder } \omega \in B\}$$

Ergebnis ω ist in Ereignis A **und/oder** Ereignis B enthalten



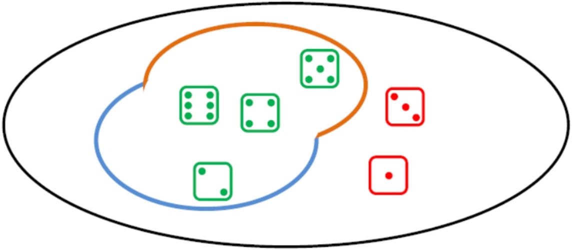
# 5.1 Mengentheoretische Grundlagen

Bezeichnungen: Beispiel Würfelwurf

**Vereinigungsereignis** zweier Mengen

$$2 \in \{2,4,6\} \cup \{4,5,6\}$$

Augenzahl 2 ist gerade Zahl



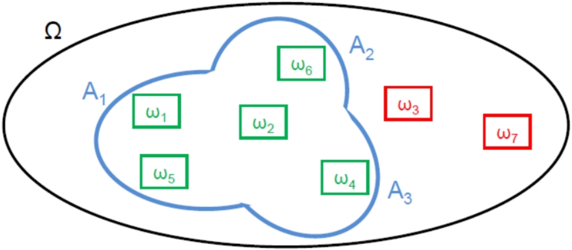
# 5.1 Mengentheoretische Grundlagen

## Bezeichnungen

**Vereinigungsereignis** beliebig vieler Mengen

$$\bigcup_{i \in I} = \{\omega \in \Omega | \omega \in A_i \text{ für mindestens ein } i \in I\}$$

Ergebnis ω ist in mindestens einem $A_i$ enthalten



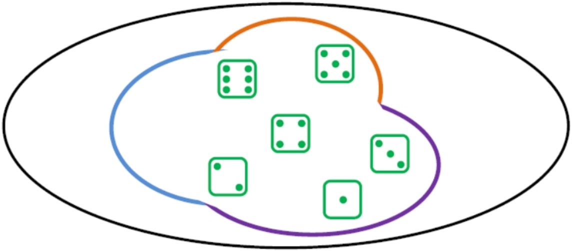
# 5.1 Mengentheoretische Grundlagen

Bezeichnungen: Beispiel Würfelwurf

**Vereinigungsereignis** beliebig vieler Mengen

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

Augenzahl 5 ist größer als 3



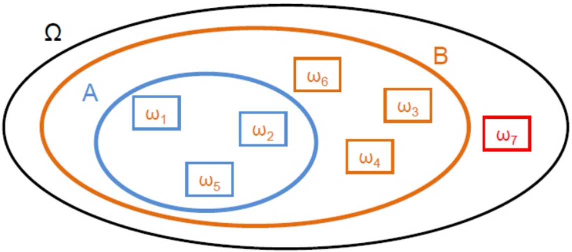
# 5.1 Mengentheoretische Grundlagen

## Bezeichnungen

### Teilereignis

$$A \subset B \text{ (bzw. } A \subseteq B\text{)}$$

Ereignis A ist in Ereignis B enthalten, aus Ereignis A folgt Ereignis B



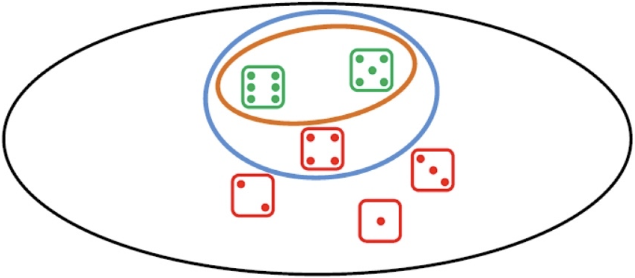
# 5.1 Mengentheoretische Grundlagen

Bezeichnungen: Beispiel Würfelwurf

**Teilereignis**

$$\{5, 6\} \subset \{4, 5, 6\}$$

Augenzahl 5 ist größer als 4 und damit auch größer als 3.



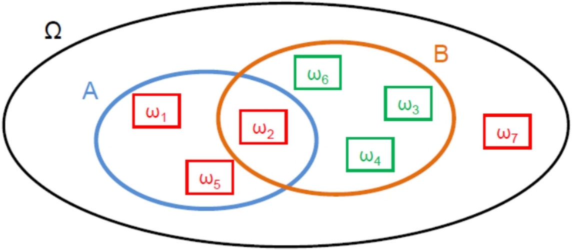
# 5.1 Mengentheoretische Grundlagen

## Bezeichnungen

### Differenzereignis

$$B \setminus A = \{\omega \in \Omega | \omega \in B \text{ und } \omega \notin A\}$$

Ergebnis ω ist in Ereignis B, aber nicht in Ereignis A enthalten



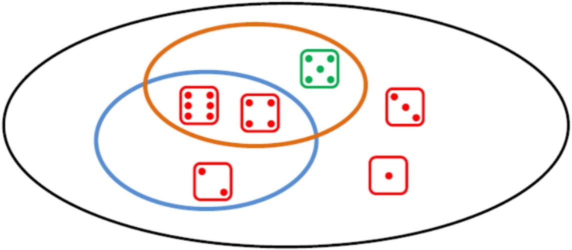
# 5.1 Mengentheoretische Grundlagen

Bezeichnungen: Beispiel Würfelwurf

**Differenzereignis**

$$5 \in \{4, 5, 6\} \setminus \{2, 4, 6\}$$

Augenzahl 5 ist größer als 3, aber nicht gerade Zahl



# 5.1 Mengentheoretische Grundlagen

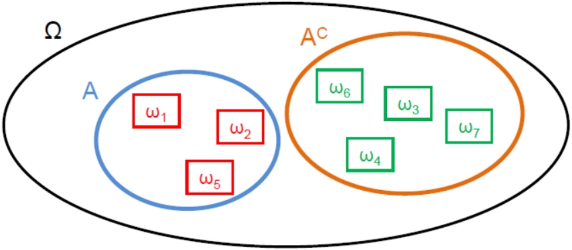
## Bezeichnungen

### Komplementärereignis

$$A^c = \Omega \setminus A = \{\omega \in \Omega | \omega \notin A\}$$

Ergebnis ω ist in Ereignis $A^C$ enthalten ⇔ Ergebnis ω ist nicht in Ereignis A enthalten

Das Ereignis $A^C$ heißt **Komplement** bzw. **Gegenereignis** von bzw. zu A



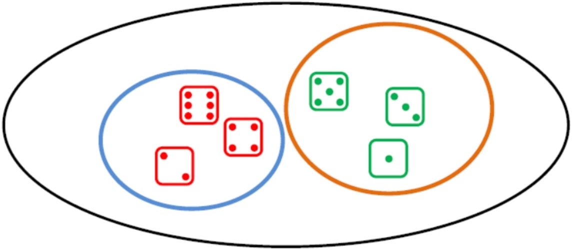
# 5.1 Mengentheoretische Grundlagen

Bezeichnungen: Beispiel Würfelwurf

**Komplementärereignis**

$$5 \in \{2, 4, 6\}^c$$

Augenzahl 5 ist nicht gerade Zahl



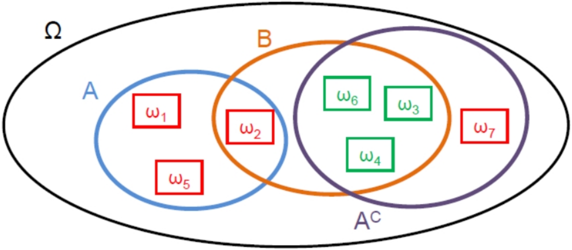
# 5.1 Mengentheoretische Grundlagen

## Regeln

**Differenzereignis und Komplementärereignis**

$$B \setminus A = \{\omega \in \Omega | \omega \in B \text{ und } \omega \notin A\} = B \cap A^c = B \setminus (A \cap B)$$

Ergebnis ω ist in Ereignis B, aber nicht in Ereignis A enthalten

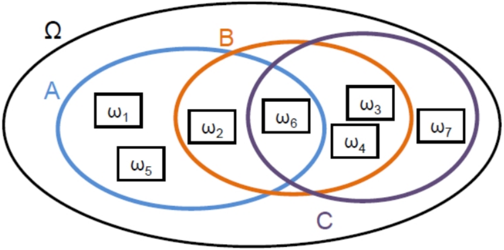


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Distributivgesetz**

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

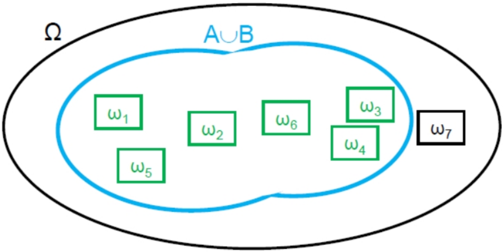


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Distributivgesetz**

$$\boxed{(A \cup B)} \cap C = (A \cap C) \cup (B \cap C)$$

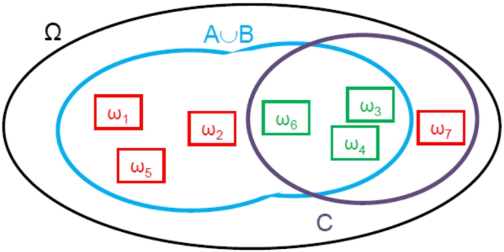


# 5.1 Mengentheoretische Grundlagen

## Regeln

### Distributivgesetz

$$\boxed{\boxed{(A \cup B)} \cap \boxed{C}} = (A \cap C) \cup (B \cap C)$$

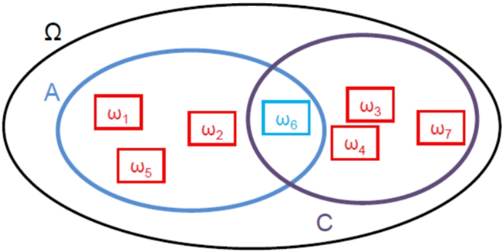


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Distributivgesetz**

$$(A \cup B) \cap C = \boxed{(A \cap C)} \cup (B \cap C)$$

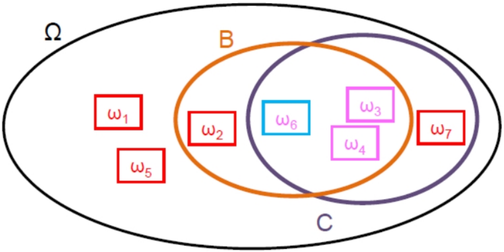


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Distributivgesetz**

$$(A\cup B)\cap C = \boxed{(A\cap C)} \cup \boxed{(B\cap C)}$$

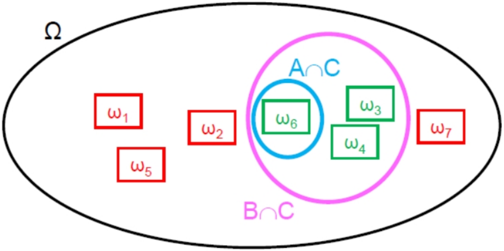


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Distributivgesetz**

$$(A \cup B) \cap C = \boxed{\boxed{(A \cap C)} \cup \boxed{(B \cap C)}}$$

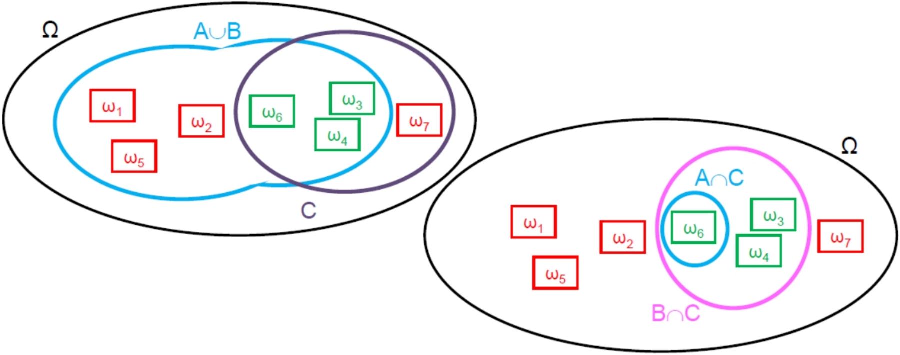


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Distributivgesetz**

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

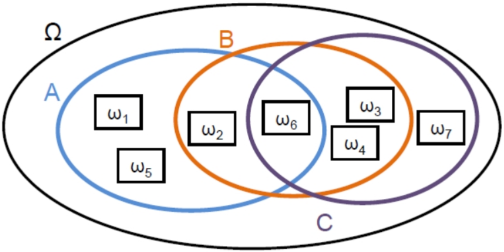


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Distributivgesetz**

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

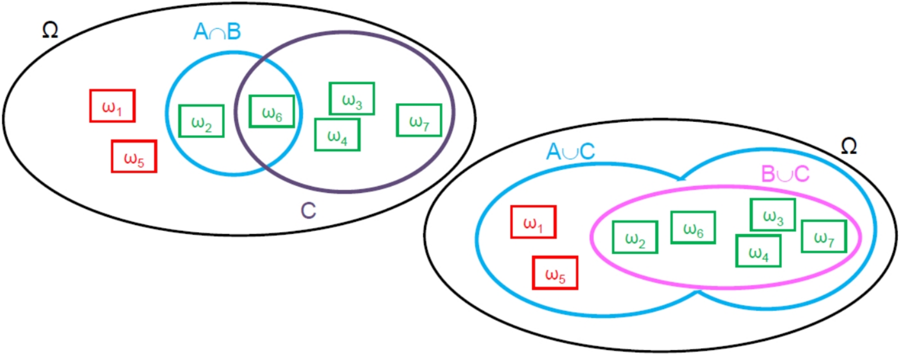


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Distributivgesetz**

$$(A\cap B)\cup C = (A\cup C)\cap(B\cup C)$$

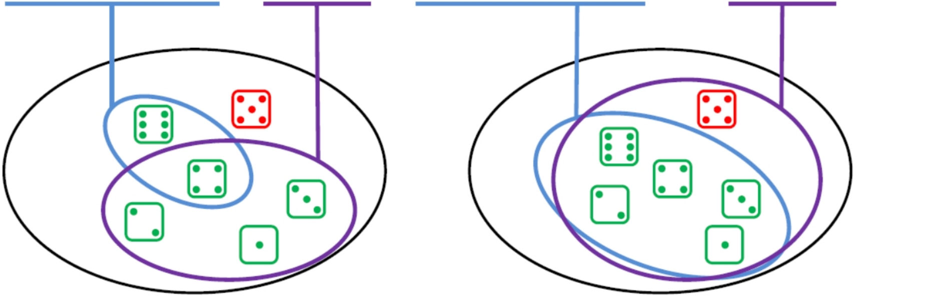


# 5.1 Mengentheoretische Grundlagen

Regeln: Beispiel Würfelwurf

**Distributivgesetz**

$$(\{2,4,6\} \cap \{4,5,6\}) \cup \{1,2,3,4\} = (\{2,4,6\} \cup \{1,2,3,4\}) \cap (\{4,5,6\} \cup \{1,2,3,4\})$$

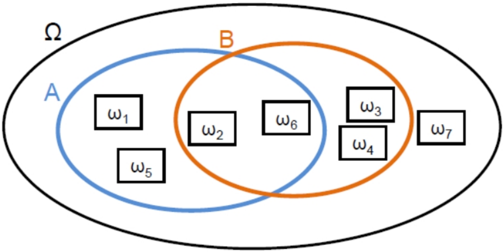


# 5.1 Mengentheoretische Grundlagen

Regeln

**Regeln von de Morgan**

$$(A \cap B)^C = A^C \cup B^C \quad , \quad (A \cup B)^C = A^C \cap B^C$$

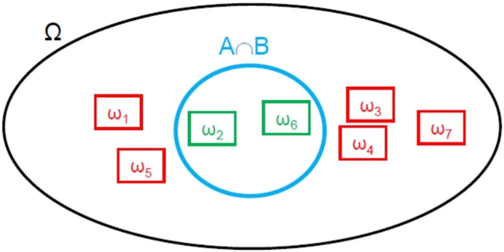


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Regeln von de Morgan**

$$\boxed{(A \cap B)}^C = A^C \cup B^C$$

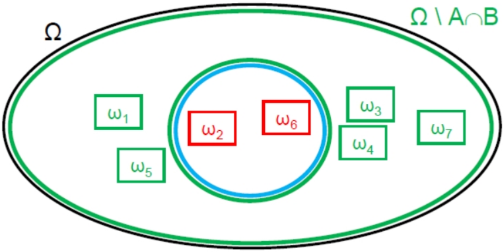


# 5.1 Mengentheoretische Grundlagen

Regeln

**Regeln von de Morgan**

$$\boxed{(A \cap B)^C} = A^C \cup B^C$$

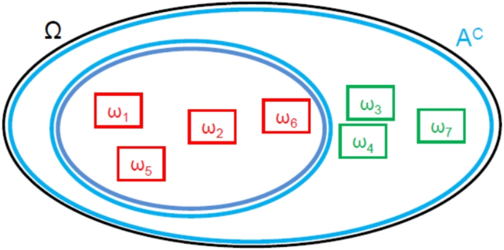


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Regeln von de Morgan**

$$(A \cap B)^C = \boxed{A^C} \cup B^C$$

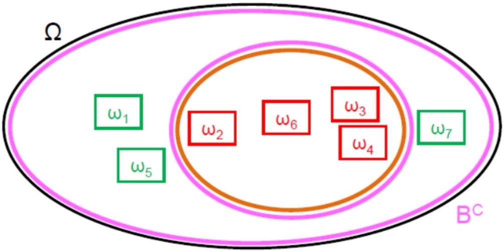


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Regeln von de Morgan**

$$(A \cap B)^C = A^C \cup \boxed{B^C}$$

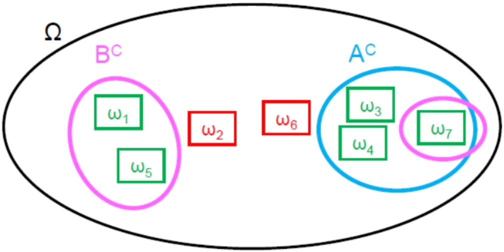


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Regeln von de Morgan**

$$(A \cap B)^C = \boxed{\boxed{A^C} \cup \boxed{B^C}}$$

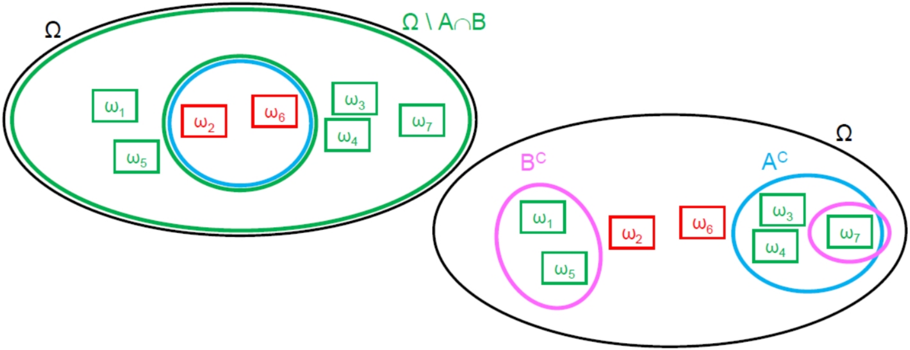


# 5.1 Mengentheoretische Grundlagen

## Regeln

### Regeln von de Morgan

$$(A \cap B)^C = A^C \cup B^C$$

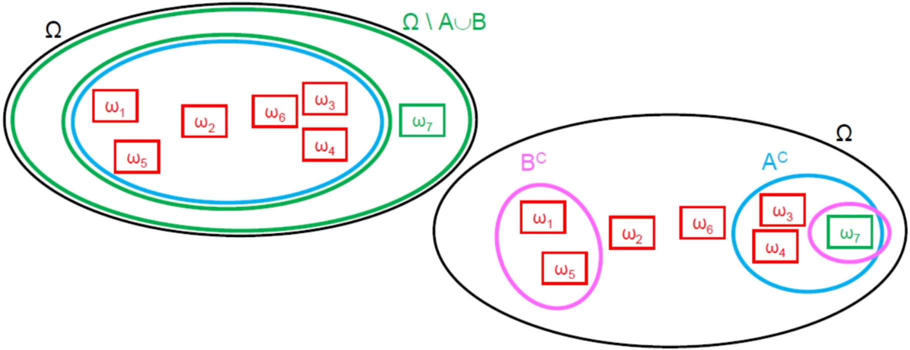


# 5.1 Mengentheoretische Grundlagen

## Regeln

**Regeln von de Morgan**

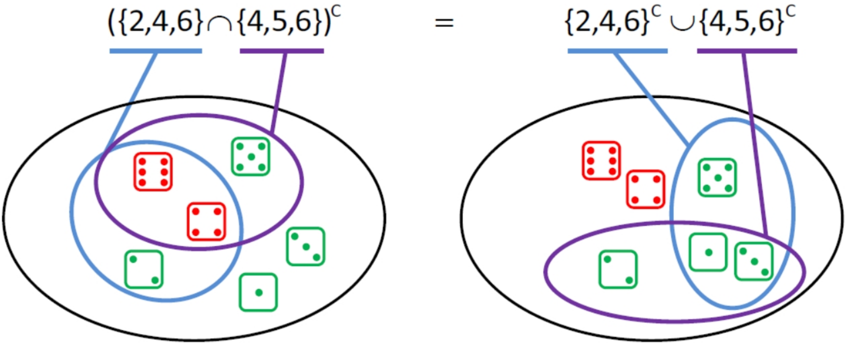
$$\boxed{\boxed{(A \cup B)}^C} = \boxed{\boxed{A^C} \cap \boxed{B^C}}$$



# 5.1 Mengentheoretische Grundlagen

Regeln: Beispiel Würfelwurf

**Regeln von de Morgan**



# 5.1 Mengentheoretische Grundlagen

## Beispiel Mausaktivität

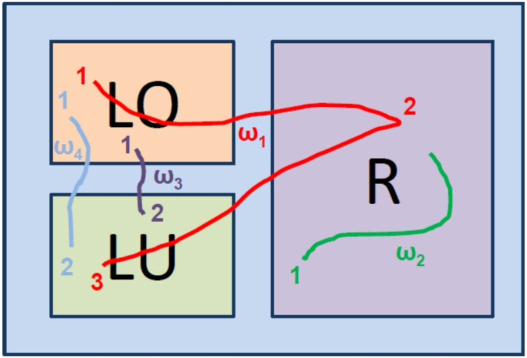
### Ereignisbeispiele

$A \supset \{\omega_1, \omega_3, \omega_4\}$
„Letzter Click auf LU“

$B \supset \{\omega_3, \omega_4\}$
„Mauszeiger immer in linker Hälfte“

$C \supset \{\omega_2\}$
„Nur einmal geclickt“

$D \supset \{\omega_1, \omega_2\}$
„R wurde geclickt“



# 5.1 Mengentheoretische Grundlagen

## Zusammenfassung Bezeichnungen

| **Mathematische Schreibweise** | **Ausformulierte Schreibweise** |
| --- | --- |
| $\omega \in A$ | Ergebnis $\omega$ ist in Ereignis $A$ enthalten |
| $A \cap B$ | **Schnittereignis**: Menge aller Ergebnisse, die in $A$ und $B$ enthalten sind |
| $A \cap B = \emptyset$ | $A$ und $B$ sind **disjunkt**: es gibt kein Ergebnis, das in $A$ und $B$ enthalten ist |
| $A \cup B$ | **Vereinigungsereignis**: Menge aller Ergebnisse, die in $A$ **und/oder** $B$ enthalten sind |
| $A \subseteq B$ | $A$ ist **Teilereignis** von $B$: Alle in $A$ enthaltenen Ergebnise sind auch in $B$ enthalten |
| $B \setminus A$ | **Differenzereignis**: Menge der Ergebnisse, die in $B$, aber nicht in $A$ enthalten sind |
| $A^c = \Omega \setminus A$ | **Komplementärereignis**: Menge aller Ergebnisse, die nicht in $A$ enthalten sind |

# 5.1 Mengentheoretische Grundlagen

## Zusammenfassung Regeln

| Mathematische Schreibweise | Ausformulierte Schreibweise |
|---|---|
| $(A \cup B) \cap C = (A \cap B) \cup (B \cap C)$ | **Distributivgesetze** Die Schnittmenge einer zwei Mengen $A$ und $B$ vereinigenden Menge mit einer weiteren Menge $C$ ist gleich der Vereinigung der beiden aus $C$ und jeweils einer der beiden Mengen $A$ und $B$ gebildeten Schnittmengen. |
| $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ | Die Vereinigung der Schnittmenge zweier Mengen $A$ und $B$ mit einer weiteren Menge $C$ ist gleich der Schnittmenge der beiden aus $C$ und jeweils einer der beiden Mengen $A$ und $B$ gebildeten Vereinigungen |

# 5.1 Mengentheoretische Grundlagen

## Zusammenfassung Regeln

| **Mathematische Schreibweise** | **Ausformulierte Schreibweise** |
|---|---|
| $(A \cap B)^c = A^c \cup B^c$ | **Regeln von de Morgan** Das Komplementärereignis der Schnittmenge zweier Mengen ist gleich der Vereinigung der Komplementärereignisse der zwei Mengen. |
| $(A \cup B)^c = A^c \cap B^c$ | Das Komplementärereignis der Vereinigung zweier Mengen ist gleich der Schnittmenge der Komplementärereignisse der zwei Mengen |

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

Zuletzt: Interpretation von Daten als Realisationen von Zufallsvariablen.
Mengentheoretische Grundlagen zur Ordnung von Ergebnissen und Ereignissen

| | | |
|---|---|---|
| **Ergebnis und Ereignis** $\omega \in A$ | **Teilereignis** $A \subseteq B$ | **Vereinigungsereignis** $A \cup B$ |
| **Schnittereignis** $A \cap B$ | **Differenzereignis** $B \setminus A$ | **Distributivgesetze** $(A \cup B) \cap C = (A \cap B) \cup (B \cap C)$ <br> $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ |
| **Disjunkte Ereignisse** $A \cap B = \emptyset$ | **Komplementärereignis** $A^c = \Omega \setminus A$ | **Regeln von de Morgan** $(A \cap B)^c = A^c \cup B^c$ <br> $(A \cup B)^c = A^c \cap B^c$ |

Jetzt: Zuordnung von Wahrscheinlichkeiten zu Ergebnissen und Ereignissen

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Kolmogorov-Axiome, Wahrscheinlichkeitsmaß**

Seien $\Omega$ ein Grundraum und $\mathcal{A}$ die Menge aller Ereignisse über $\Omega$.
Dann heißt die Abbildung

$$P: \mathcal{A} \to [0,1], \quad A \mapsto P(A),$$

ein **Wahrscheinlichkeitsmaß**, falls sie folgende Eigenschaften (**Kolmogorov-Axiome**) besitzt:

1. $0 \leq P(A) \leq 1$ für jedes Ereignis $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{A}$

Der Wert $P(A)$ für ein Ereignis $A$ heißt **Wahrscheinlichkeit** von $A$.

Das Tripel $(\Omega, \mathcal{A}, P)$ heißt **Wahrscheinlichkeitsraum**

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Kolmogorov-Axiome, Wahrscheinlichkeitsmaß**: Beispiel Würfelwurf

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

$$\begin{aligned}\mathcal{A} = \Big\{&\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\},\\
&\{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{1,6\}, \{2,3\}, \{2,4\}, \{2,5\}, \{2,6\}, \{3,4\}, \{3,5\}, \{3,6\},\\
&\{4,5\}, \{4,6\}, \{5,6\}, \{1\},\\
&\{1,2,3\}, \{1,2,4\}, \{1,2,5\}, \{1,2,6\}, \{1,3,4\}, \{1,3,5\}, \{1,3,6\}, \{1,4,5\}, \{1,4,6\}, \{1,5,6\},\\
&\{2,3,4\}, \{2,3,5\}, \{2,3,6\}, \{2,4,5\}, \{2,4,6\}, \{2,5,6\}, \{3,4,5\}, \{3,4,6\}, \{3,5,6\}, \{4,5,6\},\\
&\{1,2,3,4\}, \{1,2,3,5\}, \{1,2,3,6\}, \{1,2,4,5\}, \{1,2,4,6\}, \{1,2,5,6\}, \{1,2,5,6\}, \{1,3,4,5\},\\
&\{1,3,4,6\}, \{1,3,5,6\}, \{1,4,5,6\}, \{2,3,4,5\}, \{2,3,4,6\}, \{2,4,5,6\}, \{3,4,5,6\},\\
&\{1,2,3,4,5\}, \{1,2,3,4,6\}, \{1,2,4,5,6\}, \{1,3,4,5,6\}, \{2,3,4,5,6\},\\
&\{1,2,3,4,5,6\}\Big\}\end{aligned}$$

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = 1/6$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Kolmogorov-Axiome, Wahrscheinlichkeitsmaß**: Beispiel Würfelwurf (Fortsetzung)

1. $0 \leq P(A) \leq 6/6 = 1$ für alle $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarw. disj. Ereignisse $A_i \in \mathcal{A}$,

   insb. für $A_i = \{i\}$, $i = 1, \ldots, 6$, $A_i = \emptyset$, $i > 6$ : $P\left(\bigcup_{i=1}^{\infty} \{i\}\right) = P\left(\bigcup_{i=1}^{6} \{i\}\right)$

   $= P(\Omega) = \sum_{i=1}^{6} P(i) = 6/6 = 1$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Kolmogorov-Axiome, Wahrscheinlichkeitsmaß**:

Beispiel Mausaktivität, interpolierte $x$-Position des Mauszeigers zu stetiger Zeit $t$

$\Omega = [1, 800]$

$$\mathcal{A} = \Big\{\emptyset, \, \{(a,b) | a \in \Omega, b \in \Omega, a \leq b\}, \Big\{\bigcup_{c=1}^{2} (a_c, b_c) | a_c \in \Omega, b_c \in \Omega, a_c \leq b_c\Big\}, \ldots,$$
$$\Big\{\bigcup_{c=1}^{\infty} (a_c, b_c) | a_c \in \Omega, b_c \in \Omega, a_c \leq b_c\Big\}, \{[a,b)|...\}, \ldots \{(a,b]|...\}, \{[a,b]|...\}, \ldots\Big\}$$

$$P([a,b]) = (b-a)/799, \; a \leq b$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Kolmogorov-Axiome, Wahrscheinlichkeitsmaß**:

Beispiel Mausaktivität, interpolierte $x$-Position des Mauszeigers zu stetiger Zeit $t$ (Fortsetzung)

1. $0 \leq P(A) \leq 799/799 = 1$ für alle $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} P(A_i)$ für alle paarw. disj. Ereignisse $A_i \in \mathcal{A}$,
   z.B. für
   $A_1 = 1, A_2 = (1, 400), A_3 = 400, A_4 = (400, 800) A_5 = 800, A_i = \emptyset,\ i > 5:$
   $P\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = P(1 \cup (1, 400) \cup 400 \cup (400, 800) \cup 800) = P(\Omega) = \sum\limits_{i=1}^{5} P(i)$
   $= \frac{0+399+0+400+0}{799} = 1$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen**

$P : \mathcal{A} \to [0,1], \quad A \mapsto P(A)$

1. $0 \leq P(A) \leq 1$ für jedes Ereignis $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{A}$

$$\boxed{\textbf{(i)}\ A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)}$$

Beweis:

Setze $A_1 = A$, $A_2 = B$, $A_i = \emptyset$, für $i > 2$

$$\boxed{P(A \cup B)} = P\left(\bigcup_{i=1}^{\infty} A_i\right) \underset{3.}{=} \sum_{i=1}^{\infty} P(A_i) = P(A_1) + \sum_{i=2}^{\infty} P(A_i)$$

$$\underset{3.}{=} P(A_1) + \left(\bigcup_{i=2}^{\infty} A_i\right) = \boxed{P(A) + P(B)} \qquad \square$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen**

$P: \mathcal{A} \to [0,1], \quad A \mapsto P(A)$

1. $0 \leq P(A) \leq 1$ für jedes Ereignis $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{A}$

$$\boxed{\textbf{(ii)}\ A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)}$$

Beweis:

$$P(B) = P((B \setminus A) \cup A) \underset{(i)}{=} P(B \setminus A) + P(A) \Rightarrow \boxed{P(B \setminus A) = P(B) - P(A)} \quad \square$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen**

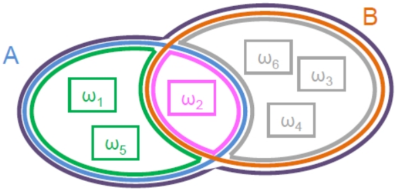
$P: \mathcal{A} \to [0,1], \quad A \mapsto P(A)$

1. $0 \leq P(A) \leq 1$ für jedes Ereignis $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{A}$

$$\boxed{\textbf{(iii)}\ P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

Beweis:

$A \cup B = [A \setminus (A \cap B)] \cup [B \setminus (A \cap B)] \cup [A \cap B]$



# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

Beweis: (Fortsetzung)

$$
\begin{aligned}
A \cup B &= [A \setminus (A \cap B)] \cup [B \setminus (A \cap B)] \cup [A \cap B] \\
&\Rightarrow \boxed{P([A \cup B])} = P([A \setminus (A \cap B)] \cup [B \setminus (A \cap B)] \cup [A \cap B]) \\
&\underset{(i)}{=} P([A \setminus (A \cap B)]) + P([B \setminus (A \cap B)]) + P(A \cap B) \\
&\underset{(ii)}{=} P(A) - P(A \cap B) + P(B) - P(A \cap B) + P(A \cap B) \\
&= \boxed{P(A) + P(B) - P(A \cap B)} \quad \square
\end{aligned}
$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

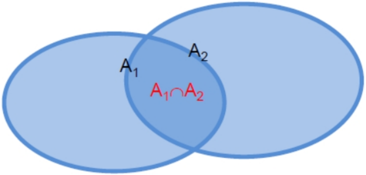
(iv) **Poincaré-Sylvesterformel**

$$P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$$

Am Beispiel $N = 2$:

$$P(A_1 \cup A_2) = (-1)^{1+1} \cdot P(A_1) + (-1)^{1+1} \cdot P(A_2) + (-1)^{2+1} \cdot P(A_1 \cap A_2)$$

$$= \boxed{P(A_1)} + \boxed{P(A_2)} - \boxed{P(A_1 \cap A_2)}$$



# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen**

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) **Poincaré-Sylvesterformel**

$$P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$$

Für $N = 3$:

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= (-1)^{1+1} \cdot P(A_1) + (-1)^{1+1} \cdot P(A_2) + (-1)^{1+1} \cdot P(A_3) \\ &\quad + (-1)^{2+1} \cdot P(A_1 \cap A_2) + (-1)^{2+1} \cdot P(A_1 \cap A_3) + (-1)^{2+1} \cdot P(A_2 \cap A_3) \\ &\quad + (-1)^{3+1} \cdot P(A_1 \cap A_2 \cap A_3) \\ &= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ &\quad + P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen**

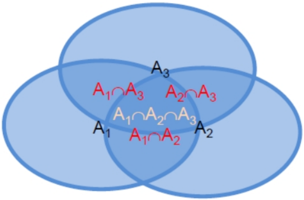
(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) **Poincaré-Sylvesterformel**

$$P\Big(\bigcup_{n=1}^{N} A_N\Big) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$$

Für $N = 3$: $\quad P(A_1 \cup A_2 \cup A_3) =$

$$\boxed{P(A_1)} + \boxed{P(A_2)} + \boxed{P(A_3)} - \boxed{P(A_1 \cap A_2} - \boxed{P(A_1 \cap A_3} - \boxed{P(A_2 \cap A_3}$$
$$+ \boxed{P(A_1 \cap A_2 \cap A_3}$$



# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen**

$P: \mathcal{A} \to [0,1], \quad A \mapsto P(A)$

1. $0 \le P(A) \le 1$ für jedes Ereignis $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{A}$

$\boxed{\textbf{(v)}\ P(A^c) = 1 - P(A)}$

Beweis:

$$\boxed{P(A^c)} = P(\Omega \setminus A) \underset{(ii)}{=} P(\Omega) - P(A) = \boxed{1 - P(A)} \quad \square$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen**

$P: \mathcal{A} \to [0,1], \quad A \mapsto P(A)$

1. $0 \leq P(A) \leq 1$ für jedes Ereignis $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{A}$

$\boxed{\textbf{(vi)} \;\; P(\emptyset) = 0}$

Beweis:

$$\boxed{P(\emptyset)} = P(\Omega^c) \underset{(v)}{=} 1 - P(\Omega) = \boxed{0} \qquad \square$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen**

$P : \mathcal{A} \to [0,1], \quad A \mapsto P(A)$

1. $0 \leq P(A) \leq 1$ für jedes Ereignis $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{A}$

$\boxed{\textbf{(vii)}\ A \subseteq B \Rightarrow P(A) \leq P(B)}$

Beweis:

$$\boxed{A \subseteq B} \Rightarrow P(B \setminus A) \underset{(ii)}{=} P(B) - P(A)$$

$$\Rightarrow \boxed{P(A)} = P(B) - \underbrace{P(B \setminus A)}_{\geq 0} \boxed{\leq P(B)} \qquad \square$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Diskreter Wahrscheinlichkeitsraum

Seien $\Omega = \{\omega_1, \omega_2, \ldots\}$ ein **endlicher** oder **abzählbarer unendlicher** Grundraum und $P$ ein Wahrscheinlichkeitsmaß auf $\Omega$. Dann heißt $(\Omega, \mathcal{A}, P)$ **diskreter Wahrscheinlichkeitsraum**.

Für beliebiges Ereignis $A \in \mathcal{A}$ gilt dann nach (i):

$$P(A) = P\Big(\bigcup_{i:\omega_i \in A} \{\omega_i\}\Big) = \sum_{i:\omega_i \in A} P(\{\omega_i\})$$

## Laplace-Raum

Treten die Elemente von endlichem $\Omega = \{\omega_1, \ldots, \omega_{|\Omega|}\}$ aus einem diskreten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ alle mit der selben Wahrscheinlichkeit auf, d.h. gilt $P(\{\omega_i\}) = 1/|\Omega|$ für $i = 1, \ldots, |\Omega|$, so wird $(\Omega, \mathcal{A}, P)$ auch **Laplace-Raum** genannt und die Wahrscheinlichkeit für ein Ereignis $A \in \mathcal{A}$ kann durch $P(A) = |A|/|\Omega|$ angegeben werden.

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

**Eigenschaften von Wahrscheinlichkeitsmaßen; Diskreter Wahrscheinlichkeitsraum**: Beispiel: *Bearbeitungen von Softwareaufgaben*

| Bearbeitung | Bearbeiter(in) | Aufgabe | Version |
|---|---|---|---|
| $e_1$ | Kai | Export | 1.1 |
| $e_2$ | Kai | Verknüpfung | 1.2 |
| $e_3$ | Miriam | Export | 1.1 |
| $e_4$ | Tina | Verknüpfung | 1.2 |
| $e_5$ | Oliver | Export | 2.0 |
| $e_6$ | Tina | Export | 1.2 |
| $e_7$ | Tina | Verknüpfung | 1.2 |
| $e_8$ | Miriam | Export | 1.2 |
| $e_9$ | Miriam | Export | 1.2 |
| $e_{10}$ | Oliver | Abfrage | 1.1 |
| $e_{11}$ | Oliver | Verknüpfung | 2.0 |
| $e_{12}$ | Oliver | Abfrage | 2.0 |

Zufällige Auswahl einer Bearbeitung
$\rightarrow$ Ergebnis $\omega \in \{e_1, \ldots, e_{12}\} = \Omega$

Elementarwahrscheinlichkeiten
$P(\{e_i\}) = 1/12, \; i = 1, \ldots, 12$

Ereignisse

1. Bearbeiter männlich
   $A_1 = \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\}$
2. Gestellte Aufgabe Export
   $A_2 = \{e_1, e_3, e_5, e_6, e_8, e_9\}$
3. Verwendete Version 2.0
   $A_3 = \{e_5, e_{11}, e_{12}\}$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

(i) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(ii) $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) $P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$

(v) $P(A^c) = 1 - P(A)$

(vi) $P(\emptyset) = 0$

(vii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

$$\omega \in \{e_1, \ldots, e_{12}\} = \Omega$$
$$P(\{e_i\}) = 1/12,\ i = 1, \ldots, 12$$
$$A_1 = \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\}$$
$$A_2 = \{e_1, e_3, e_5, e_6, e_8, e_9\}$$
$$A_3 = \{e_5, e_{11}, e_{12}\}$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

(i) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(ii) $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) $P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$

(v) $P(A^c) = 1 - P(A)$

(vi) $P(\emptyset) = 0$

(vii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

$\omega \in \{e_1, \ldots, e_{12}\} = \Omega$
$P(\{e_i\}) = 1/12,\ i = 1, \ldots, 12$

$A_1 = \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\}$
$A_2 = \{e_1, e_3, e_5, e_6, e_8, e_9\}$
$A_3 = \{e_5, e_{11}, e_{12}\}$

$$\begin{aligned}
P(A_1) &= P(\{e_1\} \cup \{e_2\} \cup \{e_5\} \cup \{e_{10}\} \cup \{e_{11}\} \cup \{e_{12}\}) \\
&= P(\{e_1\}) + P(\{e_2\}) + P(\{e_5\}) \\
&\quad + P(\{e_{10}\}) + P(\{e_{11}\}) + P(\{e_{12}\}) = 6/12 = 1/2 \\
P(A_2) &= P(\{e_1\} \cup \{e_3\} \cup \{e_5\} \cup \{e_6\} \cup \{e_8\} \cup \{e_9\}) \\
&= P(\{e_1\}) + P(\{e_3\}) + P(\{e_5\}) \\
&\quad + P(\{e_6\}) + P(\{e_8\}) + P(\{e_9\}) = 6/12 = 1/2 \\
P(A_3) &= P(\{e_5\} \cup \{e_{11}\} \cup \{e_{12}\}) \\
&= P(\{e_5\}) + P(\{e_{11}\}) + P(\{e_{12}\}) = 3/12 = 1/4
\end{aligned}$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

(i) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(ii) $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) $P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$

(v) $P(A^c) = 1 - P(A)$

(vi) $P(\emptyset) = 0$

(vii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

$\omega \in \{e_1, \ldots, e_{12}\} = \Omega$
$P(\{e_i\}) = 1/12,\ i = 1, \ldots, 12$

$A_1 = \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\}$
$A_2 = \{e_1, e_3, e_5, e_6, e_8, e_9\}$
$A_3 = \{e_5, e_{11}, e_{12}\}$

$P(A_1) = 1/2$
$P(A_2) = 1/2$
$P(A_3) = 1/4$

Wahrscheinlichkeit für eine Bearbeitung, die von einem Mann mit einer anderen Version als 2.0 durchgeführt wurde

$$A_3 = \{e_5, e_{11}, e_{12}\} \subset \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\} = A_1$$
$$\Rightarrow \text{(ii) } P(A_1 \setminus A_3) = P(A_1) - P(A_3) = 1/2 - 1/4 = 1/4$$
$$\Rightarrow \text{(vii) } 1/4 = P(A_3) \leq P(A_1) = 1/2$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

(i) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(ii) $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) $P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \le n_1 < \ldots < n_m \le N} P(A_{n_1} \cap \ldots \cap A_{n_m})$

(v) $P(A^c) = 1 - P(A)$

(vi) $P(\emptyset) = 0$

(vii) $A \subseteq B \Rightarrow P(A) \le P(B)$

$\omega \in \{e_1, \ldots, e_{12}\} = \Omega$
$P(\{e_i\}) = 1/12, \; i = 1, \ldots, 12$

$A_1 = \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\}$
$A_2 = \{e_1, e_3, e_5, e_6, e_8, e_9\}$
$A_3 = \{e_5, e_{11}, e_{12}\}$

$P(A_1) = 1/2$
$P(A_2) = 1/2$
$P(A_3) = 1/4$

Wahrscheinlichkeit für eine Bearbeitung, die Aufgabe Export hatte und/oder von einem Mann durchgeführt wurde

$$\begin{aligned}
P(A_1 \cup A_2) &= P(A_1) + P(A_2) - P(A_1 \cap A_2) \\
&= 1/2 + 1/2 - P(\{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\} \\
&\quad \cap \{e_1, e_3, e_5, e_6, e_8, e_9\}) \\
&= 1 - P(\{e_1, e_5\}) = 1 - P(\{e_1\} \cup \{e_5\}) \\
&= 1 - (P(\{e_1\}) + P(\{e_5\})) = 1 - 2/12 \\
&= 10/12 = 5/6
\end{aligned}$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

(i) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(ii) $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) $P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$

(v) $P(A^c) = 1 - P(A)$

(vi) $P(\emptyset) = 0$

(vii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

$\omega \in \{e_1, \ldots, e_{12}\} = \Omega$
$P(\{e_i\}) = 1/12,\ i = 1, \ldots, 12$

$A_1 = \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\}$
$A_2 = \{e_1, e_3, e_5, e_6, e_8, e_9\}$
$A_3 = \{e_5, e_{11}, e_{12}\}$

$P(A_1) = 1/2$
$P(A_2) = 1/2$
$P(A_3) = 1/4$

W'keit für eine Bearbeitung, die Aufgabe Export hatte und/oder von einem Mann und /oder mit Version 2.0 durchgeführt wurde

$$
\begin{aligned}
&P(A_1 \cup A_2 \cup A_3)\\
&= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3)\\
&\quad - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)\\
&= 1/2 + 1/2 - 1/4 - P(\{e_1, e_5\}) - P(\{e_5, e_{11}, e_{12}\})\\
&\quad - P(\{e_5\}) + P(\{e_5\})\\
&= 15/12 - 2/12 - 3/12 - 1/12 + 1/12 = 10/12 = 5/6
\end{aligned}
$$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

(i) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(ii) $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) $P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$

$\boxed{\text{(v) } P(A^c) = 1 - P(A)}$

(vi) $P(\emptyset) = 0$

(vii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

$\omega \in \{e_1, \ldots, e_{12}\} = \Omega$
$P(\{e_i\}) = 1/12,\ i = 1, \ldots, 12$

$A_1 = \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\}$
$A_2 = \{e_1, e_3, e_5, e_6, e_8, e_9\}$
$A_3 = \{e_5, e_{11}, e_{12}\}$

$P(A_1) = 1/2$
$P(A_2) = 1/2$
$P(A_3) = 1/4$

W'keit für eine Bearbeitung, die weder Aufgabe Export hatte noch von einem Mann noch mit Version 2.0 durchgeführt wurde

Mit (v): $P([A_1 \cup A_2 \cup A_3]^c) = 1 - P(A_1 \cup A_2 \cup A_3)$
$= 1 - 5/6 = 1/6$

Mit de Morgan: $P([A_1 \cup A_2 \cup A_3]^c) = P(A_1^c \cap A_2^c \cap A_3^c)$

$= P(\{e_3, e_4, e_6, e_7, e_8, e_9\} \cap \{e_2, e_4, e_7, e_{10}, e_{11}, e_{12}\}$
$\cap \{e_1, e_2, e_3, e_4, e_6, e_7, e_8, e_9, e_{10}\})$

$= P(\{e_4, e_7\}) = 2/12 = 1/6$

# 5.2 Wahrscheinlichkeitsmaße, Wahrscheinlichkeitsräume

## Eigenschaften von Wahrscheinlichkeitsmaßen

(i) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(ii) $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$

(iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(iv) $P\left(\bigcup_{n=1}^{N} A_N\right) = \sum_{m=1}^{N} (-1)^{m+1} \sum_{1 \leq n_1 < \ldots < n_m \leq N} P(A_{n_1} \cap \ldots \cap A_{n_m})$

(v) $P(A^c) = 1 - P(A)$

(vi) $P(\emptyset) = 0$

(vii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

$\omega \in \{e_1, \ldots, e_{12}\} = \Omega$

$P(\{e_i\}) = 1/12, \ i = 1, \ldots, 12$

$A_1 = \{e_1, e_2, e_5, e_{10}, e_{11}, e_{12}\}$

$A_2 = \{e_1, e_3, e_5, e_6, e_8, e_9\}$

$A_3 = \{e_5, e_{11}, e_{12}\}$

$P(A_1) = 1/2$

$P(A_2) = 1/2$

$P(A_3) = 1/4$

W'keit für eine Bearbeitung, die mit Version 2.0 von einer Frau durchgeführt wurde

$P(A_1^c \cap A_3) = P(\{e_3, e_4, e_6, e_7, e_8, e_9\} \cap \{e_5, e_{11}, e_{12}\})$
$= P(\emptyset) = 0$

# Zufallsvariablen und deren Verteilung

# 6.0 Zufallsvariablen und deren Verteilung

## Erinnerung

| | |
|---|---|
| **Zufallsexperiment** | Datenerhebungsprozess mit nicht vorhersagbarem Ausgang |
| **Ergebnis** $\omega$ | Elementarer Ausgang eines Zufallsexperiments |
| **Grundraum** $\Omega$ | Menge aller möglichen Ergebnisse<br>$\Omega = \{\omega \,\vert\, \omega \text{ ist Ergebnis des Zufallsexperiments}\}$ |

## Zufallsvariable

Eine Abbildung, die jedem Ergebnis eines Zufallsexperiments eine reelle Zahl zuordnet, wird **Zufallsvariable (ZV)** genannt. Ein konkreter Wert $x = X(\omega)$ heißt **Realisation** der Zufallsvariable $X$.

$$X : \Omega \to \mathbb{R} \qquad \omega \mapsto X(\omega)$$

# 6.0 Zufallsvariablen und deren Verteilung

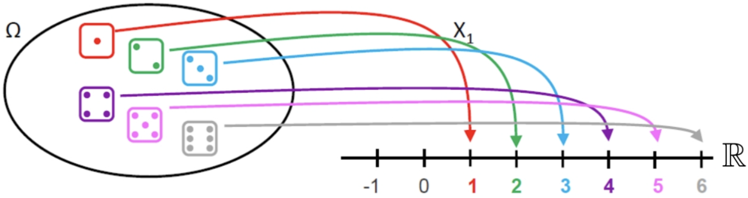
## Zufallsvariable

Eine Abbildung, die jedem Ergebnis eines Zufallsexperiments eine reelle Zahl zuordnet, wird **Zufallsvariable** genannt. Ein konkreter Wert $x = X(\omega)$ heißt **Realisation** der Zufallsvariable $X$.

$$X : \Omega \to \mathbb{R} \qquad \omega \mapsto X(\omega)$$

## Beispiel Würfelwurf

Zufallsvariable Augenzahl: $X_1(\omega) = \omega$



# 6.0 Zufallsvariablen und deren Verteilung

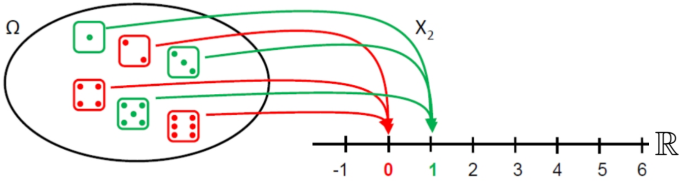
**Zufallsvariable**

Eine Abbildung, die jedem Ergebnis eines Zufallsexperiments eine reelle Zahl zuordnet, wird **Zufallsvariable** genannt. Ein konkreter Wert $x = X(\omega)$ heißt **Realisation** der Zufallsvariable $X$.

$$X : \Omega \to \mathbb{R} \qquad \omega \mapsto X(\omega)$$

**Beispiel Würfelwurf**: $X_2(\omega_i) = 1$, falls $i$-ter Wurf Kopf, $X_2(\omega_i) = 0$ sonst

Zufallsvariable Gerade/Ungerade: $X_2(\omega) \in \{0, 1\}$



# 6.0 Zufallsvariablen und deren Verteilung

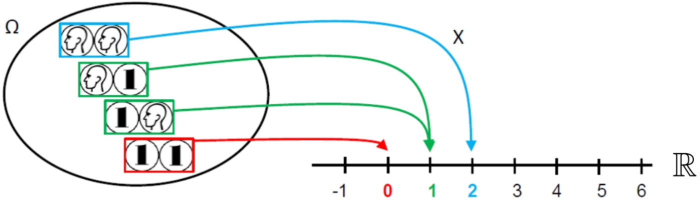
## Zufallsvariable

Eine Abbildung, die jedem Ergebnis eines Zufallsexperiments eine reelle Zahl zuordnet, wird **Zufallsvariable** genannt. Ein konkreter Wert $x = X(\omega)$ heißt **Realisation** der Zufallsvariable $X$.

$$X : \Omega \to \mathbb{R} \qquad \omega \mapsto X(\omega)$$

## Beispiel zweifacher Münzwurf

Zufallsvariable Anzahl Kopf: $X\big((\omega_1, \omega_2)\big) = \omega_1 + \omega_2$

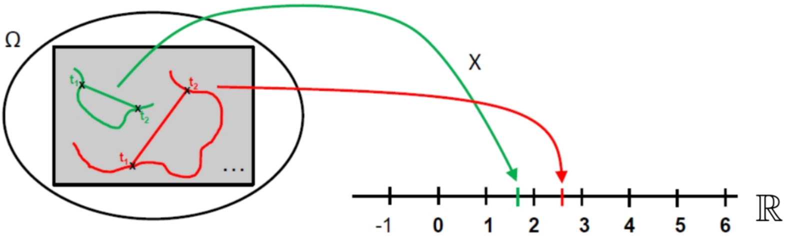


# 6.0 Zufallsvariablen und deren Verteilung

**Beispiel Mausaktivität**: $\omega(t) = [x(t), y(t), c(t)]$

ZV: Distanz zwischen ersten 2 Mausclicks

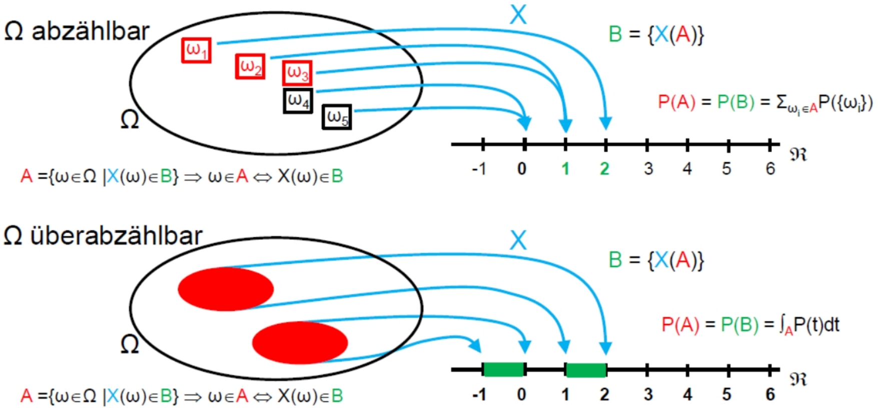
$$\boxed{\begin{gathered} X(\omega) = \sqrt{[x(t_2) - x(t_1)]^2 + [y(t_2) - y(t_1)]^2} \\ t_1 = \min(t|c(t) > 0)\; t_2 = \min(t|c(t) > 0, t > t_1) \end{gathered}}$$



# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung eindimensionaler Zufallsvariablen

Die durch die Zufallsvariable definierte Abbildung von beliebigem Grundraum $\Omega$ auf die reellen Zahlen erlaubt die Zuordnung von Wahrscheinlichkeiten zu Teilmengen von $\mathbb{R}$.



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

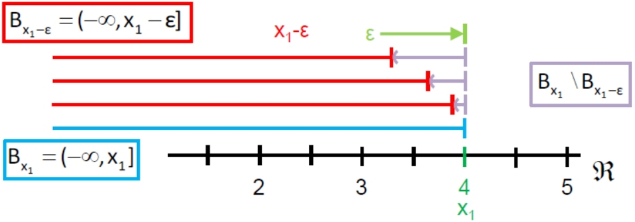
Die **Wahrscheinlichkeitsverteilung** oder kurz **Verteilung** einer Zufallsvariablen $X$ ist definiert durch

$$P^X(B) = P(X \in B) = P(\{\omega \in \Omega | X(\omega) \in B\}),\ B \subseteq \mathbb{R}$$

Diese Verteilung ist eindeutig definiert, wenn $P^X(B_x)$ für jedes Intervall der Form $B_x = (-\infty, x]$ bekannt ist:

$$\boxed{B = \{x_1\}} = \lim_{\epsilon \downarrow 0}(\{B_{x_1} \setminus B_{x_1 - \epsilon}\}) \Rightarrow P^X(B) = \lim_{\epsilon \downarrow 0}[P^X(B_{x_1}) - P^X(B_{x_1-\epsilon})],$$

da $B_{x_1-\epsilon} \subset B_{x_1}$



# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung eindimensionaler Zufallsvariablen

Die **Wahrscheinlichkeitsverteilung** oder kurz **Verteilung** einer Zufallsvariablen $X$ ist definiert durch

$$P^X(B) = P(X \in B) = P(\{\omega \in \Omega | X(\omega) \in B\}),\ B \subseteq \mathbb{R}$$

Diese Verteilung ist eindeutig definiert, wenn $P^X(B_x)$ für jedes Intervall der Form $B_x = (-\infty, x]$ bekannt ist:

$$\boxed{B = \{x_1\}} = \lim_{\epsilon \downarrow 0}(\{B_{x_1} \setminus B_{x_1 - \epsilon}\}) \Rightarrow P^X(B) = \lim_{\epsilon \downarrow 0}[P^X(B_{x_1}) - P^X(B_{x_1 - \epsilon})],$$

da $B_{x_1 - \epsilon} \subset B_{x_1}$

$$x_1 \neq \ldots \neq x_k : \boxed{B = \{x_1, \ldots, x_k\}} = \bigcup_{i=1}^{k} \lim_{\epsilon \downarrow 0}(\{B_{x_i} \setminus B_{x_i - \epsilon}\})$$

$$\Rightarrow P^X(B) = \sum_{i=1}^{k} \lim_{\epsilon \downarrow 0}[P^X(B_{x_i}) - P^X(B_{x_i - \epsilon})]$$

# 6.0 Zufallsvariablen und deren Verteilung

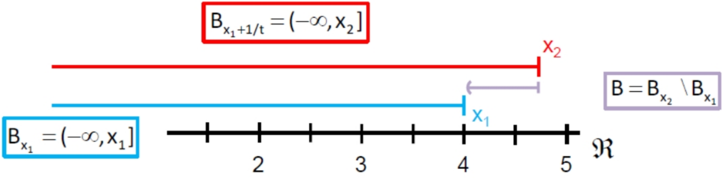
**Verteilung eindimensionaler Zufallsvariablen**

Die **Wahrscheinlichkeitsverteilung** oder kurz **Verteilung** einer Zufallsvariablen $X$ ist definiert durch

$$P^X(B) = P(X \in B) = P(\{\omega \in \Omega | X(\omega) \in B\}),\ B \subseteq \mathbb{R}$$

Diese Verteilung ist eindeutig definiert, wenn $P^X(B_x)$ für jedes Intervall der Form $B_x = (-\infty, x]$ bekannt ist:

$x_1 < x2 : \boxed{B = (x_1, x_2]} = B_{x_2} \setminus B_{x_1} \Rightarrow P^X(B) = P^X(B_{x_2}) - P^X(B_{x_1})$, da $B_{x_1} \subset B_{x_2}$



# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung eindimensionaler Zufallsvariablen

**Wahrscheinlichkeitsverteilung**: $P^X(B) = P(X \in B) = P(\{\omega \in \Omega | X(\omega) \in B\}),\ B \subseteq \mathbb{R}$

$\rightarrow$ eindeutig definiert, wenn $P^X(B_x)$ für jedes Intervall der Form $B_x = (-\infty, x]$ bekannt:

$$\boxed{B = \{x_1\}} = \lim_{\epsilon \downarrow 0}(\{B_{x_1} \setminus B_{x_1 - \epsilon}\}) \Rightarrow P^X(B) = \lim_{\epsilon \downarrow 0}[P^X(B_{x_1}) - P^X(B_{x_1 - \epsilon})],$$

da $B_{x_1 - \epsilon} \subset B_{x_1}$

$$x_1 \neq \ldots \neq x_k : \boxed{B = \{x_1, \ldots, x_k\}} = \bigcup_{i=1}^{k} \lim_{\epsilon \downarrow 0}(\{B_{x_i} \setminus B_{x_i - \epsilon}\})$$

$$\Rightarrow P^X(B) = \sum_{i=1}^{k} \lim_{\epsilon \downarrow 0}[P^X(B_{x_i}) - P^X(B_{x_i - \epsilon})]$$

$$x_1 < x_2 : \boxed{B = (x_1, x_2]} = B_{x_2} \setminus B_{x_1} \Rightarrow P^X(B) = P^X(B_{x_2}) - P^X(B_{x_1}),\ \text{da } B_{x_1} \subset B_{x_2}$$

Beliebige Ereignisse lassen sich dann aus den halboffenen Intervallen durch Schnitte und Vereinigungen konstruieren.

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

Die Funktion $F = F^X : \mathbb{R} \to [0, 1]$ mit

$$F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}),\ x \in \mathbb{R}$$

wird **Verteilungsfunktion** genannt.

Die Entsprechung der Verteilungsfunktion in der deskriptiven Statistik ist die empirische Verteilungsfunktion, bei der an die Stelle von Wahrscheinlichkeiten kumulierte relative Häufigkeiten treten.

$$F_N(x) = \begin{cases} 0 & \text{falls } x < x(1) \\ s_j = \frac{\#\{x_n | x_n \leq x(j)\}}{N} \text{ mit } j = \max\{\tilde{j} | x(\tilde{j}) \leq x\} & \text{falls } x(1) \leq x \end{cases}$$

$$= \frac{\#\{x_n | x_n \leq x\}}{N}$$

# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung eindimensionaler Zufallsvariablen

$F = F^X : \mathbb{R} \to [0,1]$ mit

$$F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \; x \in \mathbb{R}$$

## Eigenschaften der Verteilungsfunktion

$$\boxed{\text{(A)} \lim_{x\to-\infty} F(x) = 0, \; \lim_{x\to\infty} F(x) = 1}$$

Beweis:

$$\begin{aligned}
\lim_{x\to-\infty} F(x) &= \lim_{x\to-\infty} P(\{\omega \in \Omega | X(\omega) \in (-\infty, x] \cap \mathbb{R}\}) \\
&= P(\{\omega \in \Omega | X(\omega) \in \{-\infty\} \cap \mathbb{R}\}) = P(\{\omega \in \Omega | X(\omega) = \emptyset\}) \\
&\overset{(*)}{=} P(\emptyset) = 0 \qquad (*) \quad [\omega \in \Omega \Rightarrow X(\omega) \in \mathbb{R}] \Leftrightarrow [X(\omega) \notin \mathbb{R} \Rightarrow \omega \notin \Omega]
\end{aligned}$$

$$\begin{aligned}
\lim_{x\to\infty} F(x) &= \lim_{x\to\infty} P(\{\omega \in \Omega | X(\omega) \in (-\infty, x] \cap \mathbb{R}\}) \\
&= P(\{\omega \in \Omega | X(\omega) \in \mathbb{R}) = P(\Omega) = 1 \quad \square
\end{aligned}$$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

$F = F^X : \mathbb{R} \to [0,1]$ mit

$$F(X) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \; x \in \mathbb{R}$$

**Eigenschaften der Verteilungsfunktion**

(A) $\lim_{x \to -\infty} F(x) = 0, \; \lim_{x \to \infty} F(x) = 1$

$\boxed{\text{(B) } x < y \Rightarrow F(x) < F(y)}$

Beweis:

$$F(x) = P(A) \text{ mit } A = \{\omega \in \Omega | X(\omega) \leq x\}$$
$$F(y) = P(B) \text{ mit } B = \{\omega \in \Omega | X(\omega) \leq y\}$$

$$\boxed{x < y} \Rightarrow A \subseteq B \Rightarrow P(A) \leq P(B) \Leftrightarrow \boxed{F(x) \leq F(y)}$$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

$F = F^X : \mathbb{R} \to [0,1]$ mit

$$F(X) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \ x \in \mathbb{R}$$

**Eigenschaften der Verteilungsfunktion**

(A) $\lim\limits_{x \to -\infty} F(x) = 0, \ \lim\limits_{x \to \infty} F(x) = 1$

(B) $x < y \Rightarrow F(x) < F(y)$

$\boxed{\text{(C) } \lim\limits_{x \downarrow z} F(x) = F(z)}$

Beweis:

Setze $A_n = \{\omega \in \Omega | X(\omega) \in (-\infty, z + 1/n]\}, \ A_0 = \Omega$

$$\Rightarrow A = \bigcap_{n=1}^{\infty} A_n = \{\omega \in \Omega | X(\omega) \in (-\infty, z]\}, \ A_n \subset A_{n-1}, \ A_{n-1}^c \subset A_n^c, \ n = 1, 2, \ldots$$

# 6.0 Zufallsvariablen und deren Verteilung

Beweis (Fortsetzung)

$$\boxed{F(z)} = P(A) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = 1 - P\left(\bigcup_{n=1}^{\infty} A_n^c\right) = 1 - \sum_{n=1}^{\infty} P(A_n^c \setminus A_{n-1}^c)$$

$$= 1 - \lim_{N\uparrow\infty} \sum_{n=1}^{N} P(A_n^c \setminus A_{n-1}^c) = 1 - \lim_{N\uparrow\infty} P(A_N^c) = \lim_{N\uparrow\infty} P(A_N)$$

$$= \boxed{\lim_{x\downarrow z} F(x)} \quad \square$$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

$F = F^X : \mathbb{R} \to [0,1]$ mit

$$F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \ x \in \mathbb{R}$$

**Eigenschaften der Verteilungsfunktion**

(A) $\lim\limits_{x \to -\infty} F(x) = 0, \ \lim\limits_{x \to \infty} F(x) = 1$

(B) $x < y \Rightarrow F(x) < F(y)$

(C) $\lim\limits_{x \downarrow z} F(x) = F(z)$

(D) $\boxed{P(a < X \leq b) = F(b) - F(a)}$

Beweis:

Setze $A = \{\omega \in \Omega | X(\omega) \in (-\infty, a]\}$ und $B = \{\omega \in \Omega | X(\omega) \in (-\infty, b]\}$

$$\Rightarrow \boxed{P(a < X \leq b)} = P(\{\omega \in \Omega | X(\omega) \in (a, b]\}) = P(B \setminus A) \underset{A \subseteq B}{=} P(B) - P(A)$$

$$= P(X \leq b) - P(X \leq a) = \boxed{F(b) - F(a)} \quad \square$$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

$F = F^X : \mathbb{R} \to [0,1]$ mit

$$F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \; x \in \mathbb{R}$$

**Eigenschaften der Verteilungsfunktion**

(A) $\lim\limits_{x \to -\infty} F(x) = 0, \; \lim\limits_{x \to \infty} F(x) = 1$

(B) $x < y \Rightarrow F(x) < F(y)$

(C) $\lim\limits_{x \downarrow z} F(x) = F(z)$

(D) $P(a < X \leq b) = F(b) - F(a)$

$\boxed{\text{(E) } P(X > a) = 1 - F(a)}$

Beweis:

Setze $A = \{\omega \in \Omega | X(\omega) \leq a\} \Rightarrow A^c = \{\omega \in \Omega | X(\omega) > a\}$

$\Rightarrow \boxed{P(X > a)} = P(A^c) = 1 - P(A) = \boxed{1 - F(a)} \quad \square$

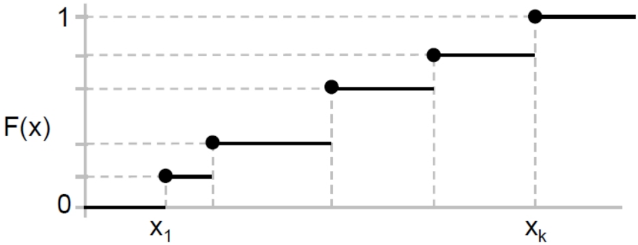
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar)

$$\Omega = \{\omega_1, \ldots, \omega_n\} \Rightarrow X \in \{x_1, \ldots, x_k\} \text{ mit } -\infty < x_1 < \ldots < x_k < \infty,\ k \leq n$$

$F = F^X : \mathbb{R} \to [0, 1]$ mit
$F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}),\ x \in \mathbb{R}$



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar)

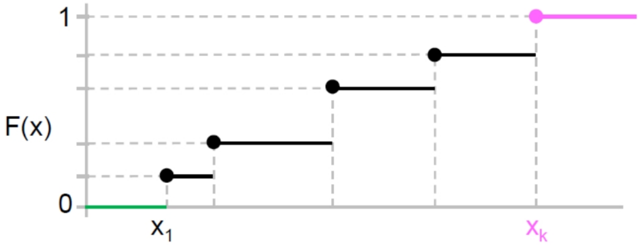
$$\Omega = \{\omega_1, \ldots, \omega_n\} \Rightarrow X \in \{x_1, \ldots, x_k\} \text{ mit } -\infty < x_1 < \ldots < x_k < \infty, \; k \leq n$$

$$\boxed{(A) \quad \lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to +\infty} F(x) = 1}$$

$F(x) = P(A_x)$ mit $A_x = \{\omega \in \Omega \,|\, X(\omega) \in (-\infty, x] \cap \{x_1, \ldots x_k\}\}$

$$\boxed{x < x_1 \Rightarrow A_x = \varnothing \Rightarrow P(A_x) = 0}$$

$$\boxed{x \geq x_k \Rightarrow A_x = \Omega \Rightarrow P(A_x) = 1}$$



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

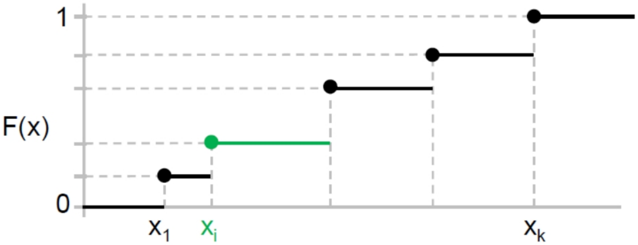
**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar)

$$\Omega = \{\omega_1, \ldots, \omega_n\} \Rightarrow X \in \{x_1, \ldots, x_k\} \text{ mit } -\infty < x_1 < \ldots < x_k < \infty, \; k \leq n$$

$$\boxed{\text{(C)} \quad \lim_{x \downarrow z} F(x) = F(z)}$$

$$F(x) = P(A_x) \text{ mit } A_x = \{\omega \in \Omega \,|\, X(\omega) \in (-\infty, x] \cap \{x_1, \ldots x_k\}\}$$

$$\boxed{\begin{aligned} i = 1, \ldots, n-1: \; x_i \leq x < x_{i+1} \; &\Rightarrow A_x = \{\omega \in \Omega \,|\, X(\omega) \in \{x_1, \ldots, x_i\}\} \\ &\Rightarrow P(A_x) = F(x_i) \end{aligned}}$$



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

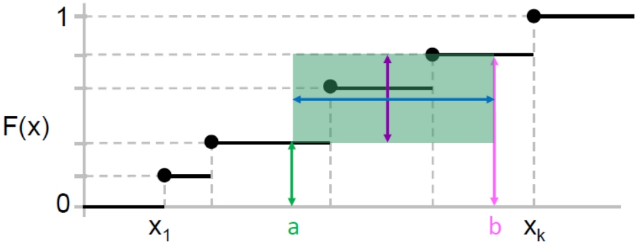
**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar)

$$\Omega = \{\omega_1, \ldots, \omega_n\} \Rightarrow X \in \{x_1, \ldots, x_k\} \text{ mit } -\infty < x_1 < \ldots < x_k < \infty, \; k \leq n$$

$$\boxed{\text{(D)} \;\; P(a < X \leq b) = F(b) - F(a)}$$

$$A_b \setminus A_a = \{\omega \in \Omega \mid X(\omega) \in \{x_1, \ldots x_k\}, a < X(\omega) \leq b\}$$

$$\boxed{P(\boxed{a < X \leq b})} = P(A_b \setminus A_a) = F(b) - F(a)$$



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

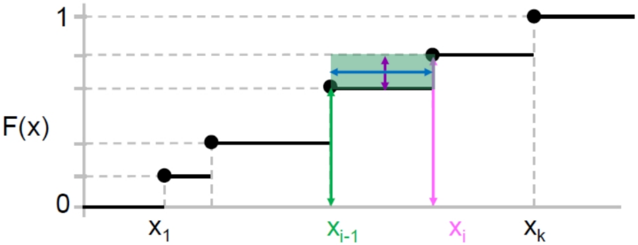
**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar)

$$\Omega = \{\omega_1, \ldots, \omega_n\} \Rightarrow X \in \{x_1, \ldots, x_k\} \text{ mit } -\infty < x_1 < \ldots < x_k < \infty, \; k \leq n$$

$\boxed{\text{(D) } i = 1,\ldots,n: P(x_{i-1} < X \leq x_i)}$ $\quad (x_0 = -\infty)$

$A_{x_i} \setminus A_{x_{i-1}} = \{\omega \in \Omega \,|\, X(\omega) \in \{x_i\}\}$

$\boxed{P(\boxed{x_{i-1} < X \leq x_i})} = P(A_{x_i} \setminus A_{x_{i-1}}) = P(X = x_i) = p_i$



# 6.0 Zufallsvariablen und deren Verteilung

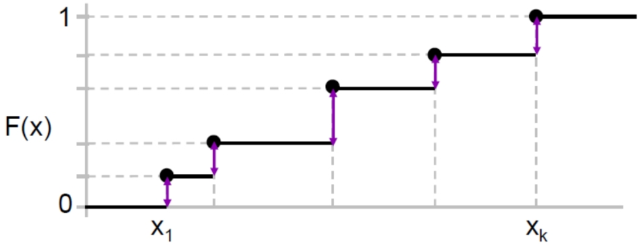
**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar)

$$\Omega = \{\omega_1, \ldots, \omega_n\} \Rightarrow X \in \{x_1, \ldots, x_k\} \text{ mit } -\infty < x_1 < \ldots < x_k < \infty, \; k \leq n$$

$\boxed{\text{(D) } i = 1,\ldots,n: P(x_{i-1} < X \leq x_i)}$ $\quad A_{x_i} \setminus A_{x_{i-1}} = \{\omega \in \Omega \mid X(\omega) \in \{x_i\}\}$

$(x_0 = -\infty)$ $\quad P(x_{i-1} < X \leq x_i) = P(A_{x_i} \setminus A_{x_{i-1}}) = P(X = x_i) = \boxed{p(x_i)}$



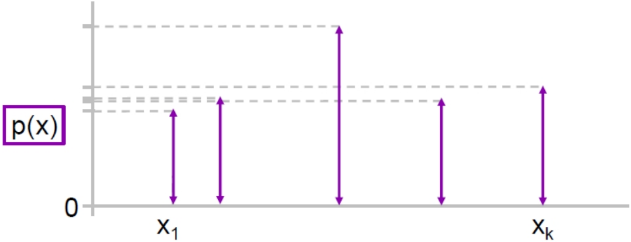
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar)

$$\Omega = \{\omega_1, \ldots, \omega_n\} \Rightarrow X \in \{x_1, \ldots, x_k\} \text{ mit } -\infty < x_1 < \ldots < x_k < \infty, \; k \leq n$$

Die Funktion: $p : \mathbb{R} \to [0, 1]$ mit $p(x) = P(X = x)$ heißt **Zähldichte von X**.



# 6.0 Zufallsvariablen und deren Verteilung

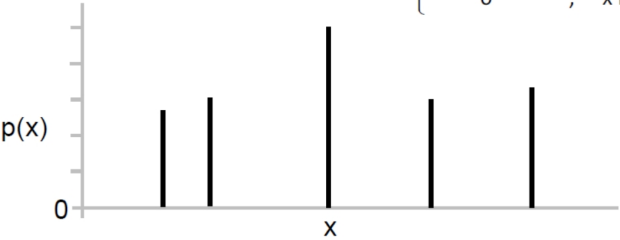
**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar)

$$\Omega = \{\omega_1, \ldots, \omega_n\} \Rightarrow X \in \{x_1, \ldots, x_k\} \text{ mit } -\infty < x_1 < \ldots < x_k < \infty, \; k \leq n$$

Die Funktion: $p : \mathbb{R} \to [0,1]$ mit $p(x) = P(X = x)$ heißt **Zähldichte von X**.

$$p(x) = \begin{cases} F(x_i) - F(x_{i-1}) & , \quad x \cap \{x_1, \ldots, x_k\} = \{x_i\} \\ 0 & , \quad x \cap \{x_1, \ldots, x_k\} = \emptyset \end{cases}$$

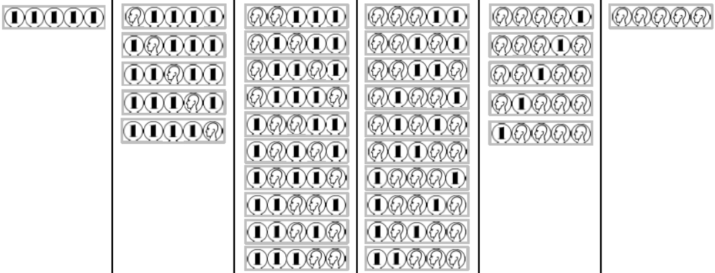


# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen, diskrete Verteilungsfunktion**

Beispiel: Anzahl Kopf beim **5-fachen Münzwurf**

**Zähldichte**

| x | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $A_x = \{\omega \in \Omega \mid X(\omega) = x\}$ |  | | | | | |
| $p(x) = P(X=x) = \|A_x\|/\|\Omega\|$ | 1/32 | 5/32 | 10/32 | 10/32 | 5/32 | 1/32 |

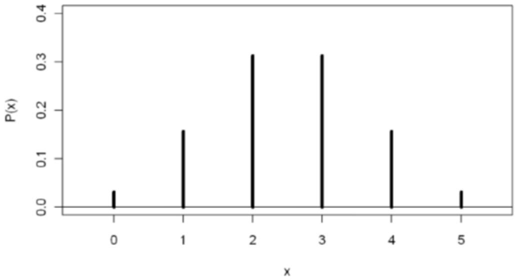
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen, diskrete Verteilungsfunktion**

Beispiel: Anzahl Kopf beim **5-fachen Münzwurf**

**Zähldichte**

| x | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| p(x)=P(X=x) | 1/32 | 5/32 | 10/32 | 10/32 | 5/32 | 1/32 |



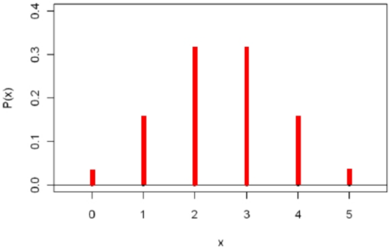
# 6.0 Zufallsvariablen und deren Verteilung

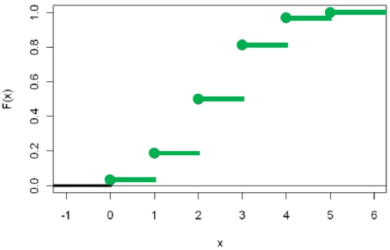
**Verteilung eindimensionaler Zufallsvariablen, diskrete Verteilungsfunktion**

Beispiel: Anzahl Kopf beim **5-fachen Münzwurf**

**Zähldichte und Verteilungsfunktion**

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $p(x)=P(X=x)$ | 1/32 | 5/32 | 10/32 | 10/32 | 5/32 | 1/32 |
| $F(x)=P(X\leq x)=\sum_{i=0}^{x} p(i)$ | 1/32 | 6/32 | 16/32 | 26/32 | 31/32 | 32/32 |



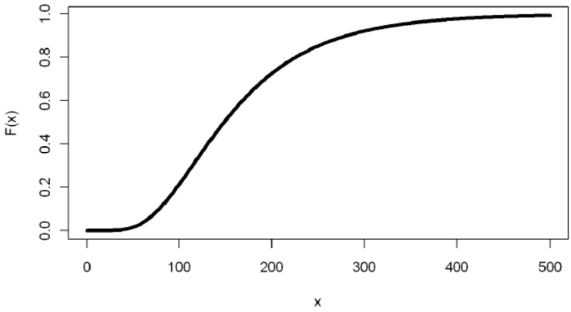



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**
**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$\omega \in \Omega: \; X(\omega) \in B, \; B \subseteq \mathbb{R}$

$$F = F^X : \mathbb{R} \to [0,1] \text{ mit } F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \; x \in \mathbb{R}$$



# 6.0 Zufallsvariablen und deren Verteilung

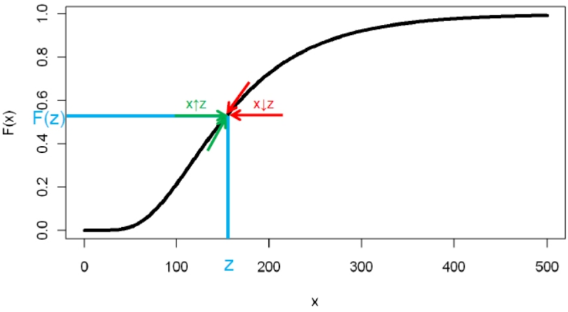
**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$\omega \in \Omega : \; X(\omega) \in B, \; B \subseteq \mathbb{R}$

$$F = F^X : \mathbb{R} \to [0,1] \text{ mit } F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \; x \in \mathbb{R}$$

$\boxed{\text{(C)} \lim_{x \downarrow z} F(x) = \boxed{F(z)}} \boxed{= \lim_{x \uparrow z} F(x)}$



# 6.0 Zufallsvariablen und deren Verteilung

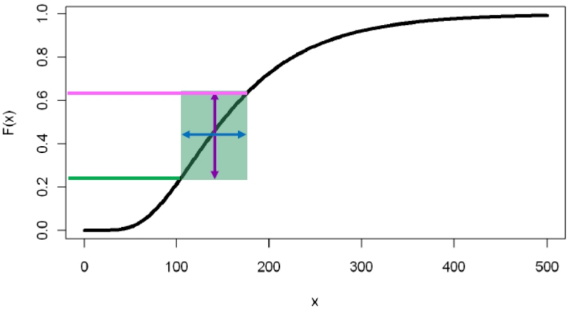
**Verteilung eindimensionaler Zufallsvariablen**
**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$\omega \in \Omega: \; X(\omega) \in B, \; B \subseteq \mathbb{R}$

$$F = F^X : \mathbb{R} \to [0,1] \text{ mit } F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \; x \in \mathbb{R}$$

(C) $\lim\limits_{x \downarrow z} F(x) = F(z) = \lim\limits_{x \uparrow z} F(x)$

$$\boxed{\text{(D)} \; \boxed{P(\boxed{a < X \leq b})} = \boxed{F(b)} - \boxed{F(a)}}$$



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$\omega \in \Omega : \; X(\omega) \in B, \; B \subseteq \mathbb{R}$

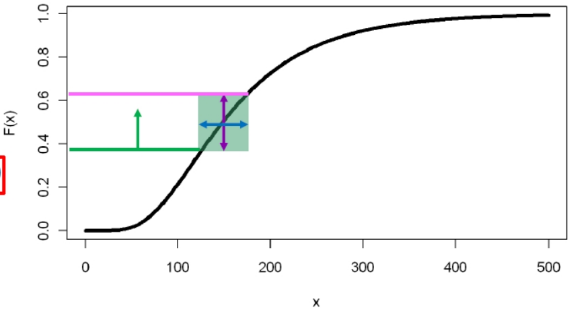
$$F = F^X : \mathbb{R} \to [0,1] \text{ mit } F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \; x \in \mathbb{R}$$

(C) $\lim_{x \downarrow z} F(x) = F(z) = \lim_{x \uparrow z} F(x)$

(D) $P(a < X \leq b) = F(b) - F(a)$

$\boxed{P(X = b)} = \lim_{a \uparrow b} P(a < X \leq b)$

$= F(b) - \lim_{a \uparrow b} F(a) = F(b) - F(b) \boxed{= 0}$



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**
**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$\omega \in \Omega : \; X(\omega) \in B, \; B \subseteq \mathbb{R}$

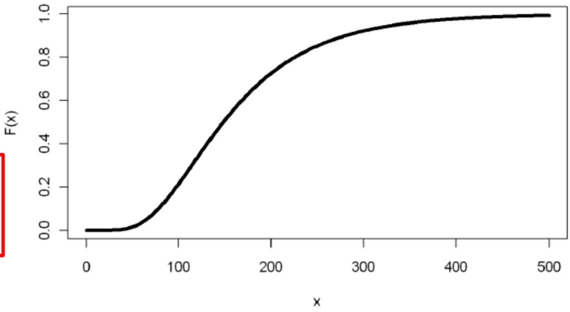
$$F = F^X : \mathbb{R} \to [0,1] \text{ mit } F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \; x \in \mathbb{R}$$

(C) $\lim_{x \downarrow z} F(x) = F(z) = \lim_{x \uparrow z} F(x)$

(D) $P(a < X \leq b) = F(b) - F(a)$

(F) $P(X = x) = 0, x \in \Re$

$$\boxed{\begin{aligned}\text{(G) } P(a < X \leq b) &= P(a \leq X \leq b) \\ &= P(a \leq X < b) = P(a < X < b) \\ &= F(b) - F(a)\end{aligned}}$$



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**
**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$\omega \in \Omega : \ X(\omega) \in B, \ B \subseteq \mathbb{R}$

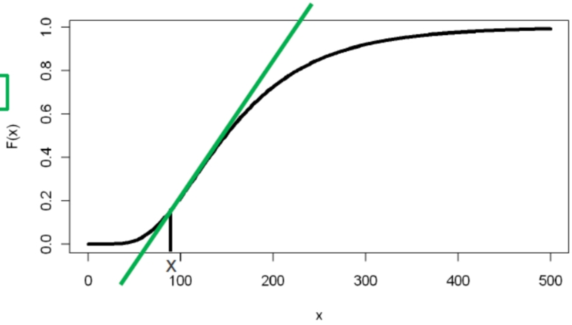
$$F = F^X : \mathbb{R} \to [0,1] \text{ mit } F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}), \ x \in \mathbb{R}$$

$$\lim_{a \to b} P(a < X \leq b) = 0$$

$$\lim_{c \downarrow 0} \frac{F(x+c) - F(x)}{c} = \boxed{F'(x) = f(x)}$$

Die Funktion f(x) wird **Dichtefunktion** bzw. **Dichte** von X genannt.

Sie beschreibt die Steigung (Grad der Verdichtung) der Verteilung X

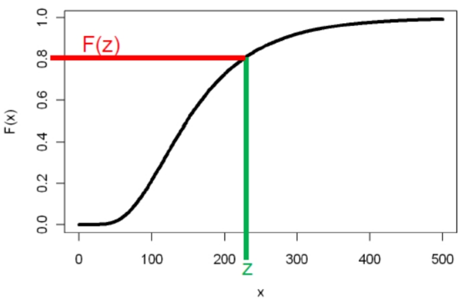


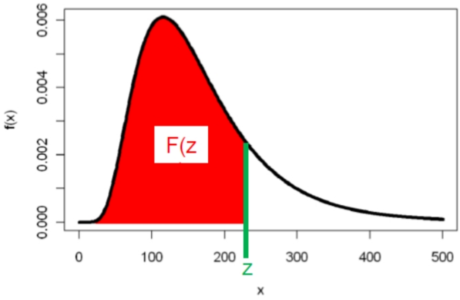
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$$F'(x) = f(x), \quad F(x) = \int_{-\infty}^{x} f(t)dt, \, x \in \Re, \quad \int_{-\infty}^{\infty} f(t)dt = 1 \qquad \boxed{P(X \le \boxed{z})} = \boxed{F(\boxed{z})} = \int_{-\infty}^{z} f(t)dt$$



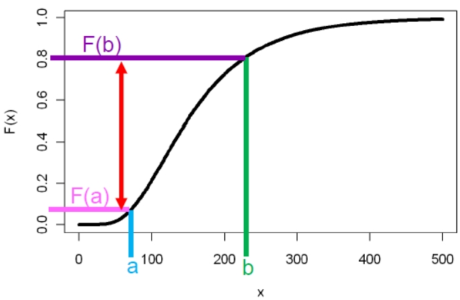
# 6.0 Zufallsvariablen und deren Verteilung

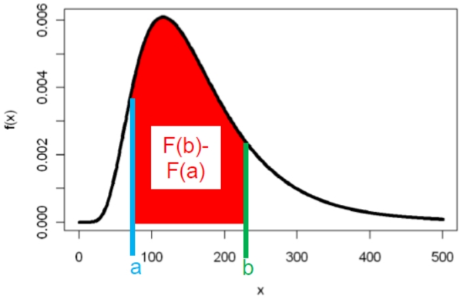
**Verteilung eindimensionaler Zufallsvariablen**

**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$$F'(x) = f(x), \quad F(x) = \int_{-\infty}^{x} f(t)dt, x \in \Re, \quad \int_{-\infty}^{\infty} f(t)dt = 1$$

$$\boxed{P(\boxed{a} < X \le \boxed{b})} = \boxed{\boxed{F(b)} - \boxed{F(a)}} = \int_{a}^{b} f(t)dt$$



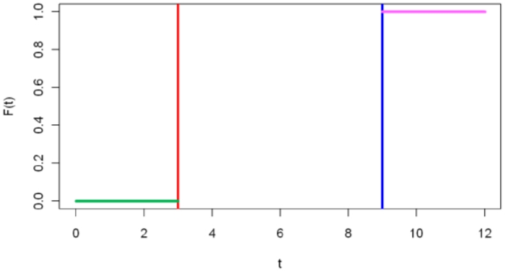
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen, stetige Verteilungsfunktion**

Beispiel: **Mausaktivität, exakter Zeitpunkt $T$ des ersten Mausclicks**

Annahme: $T$ fällt in jedes Intervall gleicher Länge $c$ zwischen $t_{\text{min}}$ und $t_{\text{max}}$ mit derselben Wahrscheinlichkeit

$$P(T < t_{min}) = 0 = F(t_{min}) \Rightarrow \boxed{F(t) = 0, t \leq \boxed{t_{min}}} \qquad P(T > t_{max}) = 0 = 1 - F(t_{max}) \Rightarrow \boxed{F(t) = 1, t \geq \boxed{t_{max}}}$$



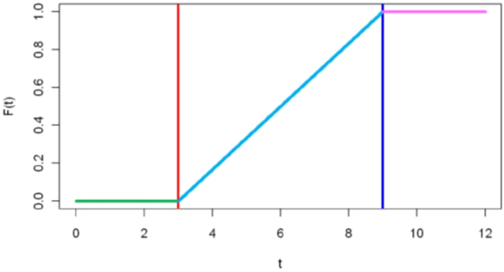
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen, stetige Verteilungsfunktion**

Beispiel: **Mausaktivität, exakter Zeitpunkt $T$ des ersten Mausclicks**

Annahme: $T$ fällt in jedes Intervall gleicher Länge $c$ zwischen $t_{\min}$ und $t_{\max}$ mit derselben Wahrscheinlichkeit

$$F(t) = 0,\ t \le \boxed{t_{\min}} \qquad F(t) = \frac{t - t_{\min}}{t_{\max} - t_{\min}},\ \boxed{t_{\min}} < t < \boxed{t_{\max}} \qquad F(t) = 1,\ t \ge \boxed{t_{\max}}$$




Wahrscheinlichkeitsdichte

$t \le t_{\min}$: $F'(t) = f(t) = \partial 0 / \partial t = 0$

$t_{\min} < t < t_{\max}$ :

$$F'(t) = f(t) \quad = \partial\left(\frac{t - t_{\min}}{t_{\max} - t_{\min}}\right) / \partial t = \frac{1}{t_{\max} - t_{\min}}$$

$t > t_{\max}$ : $F'(t) = f(t) = \partial 1 / \partial t = 0$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung eindimensionaler Zufallsvariablen, stetige Verteilungsfunktion**

Beispiel: **Mausaktivität, exakter Zeitpunkt $T$ des ersten Mausclicks**

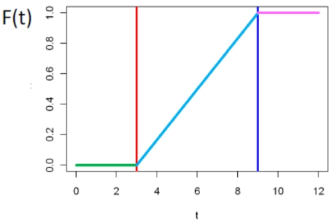
Annahme: $T$ fällt in jedes Intervall gleicher Länge $c$ zwischen $t_{\min}$ und $t_{\max}$ mit derselben Wahrscheinlichkeit

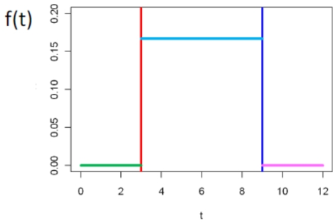
$F(t) = 0, t \leq t_{min}$ $\quad$ $F(t) = 1, t \geq t_{max}$

$$F(t) = \frac{t - t_{min}}{t_{max} - t_{min}}, \; t_{min} < t < t_{max}$$

$f(t) = 0, t \leq t_{min}$ $\quad$ $f(t) = 0, t \geq t_{max}$

$$f(t) = \frac{1}{t_{max} - t_{min}}, \; t_{min} < t < t_{max}$$



# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

Die **Wahrscheinlichkeitsverteilung** oder kurz **Verteilung** einer zweidimensionalen Zufallsvariablen $(X, Y)$ ist definiert durch

$$P^{(X,Y)}(B) = P((X,Y) \in B) = P(\{\omega \in \Omega | (X(\omega), Y(\omega)) \in B\}), \; B \subseteq \mathbb{R}^2$$

Die Funktion $F = F^{(X,Y)} : \mathbb{R}^2 \to [0,1]$ mit

$$\begin{aligned} F(x,y) &= P^{(X,Y)}((-\infty, x] \times (-\infty, y]) = P(X \leq x, Y \leq y) \\ &= P(\{\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y\}), \; x, y \in \mathbb{R}, \end{aligned}$$

wird **Verteilungsfunktion** von $(X, Y)$ genannt.

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

$P^{(X,Y)}(B) = P((X,Y) \in B) = P(\{\omega \in \Omega | (X(\omega), Y(\omega)) \in B\}),\ B \subseteq \mathbb{R}^2$

$F(x,y) = P(X \leq x, Y \leq y) = P(\{\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y\}),\ x, y \in \mathbb{R}$

Eigenschaften

1. $\lim_{x \to -\infty} F(x,y) = \lim_{y \to -\infty} F(x,y) = \lim_{x,y \to -\infty} F(x,y) = 0$

Beweis:

$$A = \{\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y\} = A_x \cap A_y \text{ mit } A_x = \{\omega \in \Omega | X(\omega) \leq x\}$$
$$A_y = \{\omega \in \Omega | Y(\omega) \leq y\}$$

$$F(x,y) = P(A) = P(A_x \cap A_y) = 1 - P(A_x^c \cup A_y^c)$$

$$\boxed{\lim_{x \to -\infty} F(x,y)} = 1 - (P(A_{-\infty}^c \cup A_y^c) = 1 - P(\Omega \cup A_y^c)$$

$$= 1 - [P(\Omega) + P(A_y^c) - P(A_y^c)] = 1 - 1 = \boxed{0} \quad \square$$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

$P^{(X,Y)}(B) = P((X,Y) \in B) = P(\{\omega \in \Omega | (X(\omega), Y(\omega)) \in B\}),\ B \subseteq \mathbb{R}^2$

$F(x,y) = P(X \leq x, Y \leq y) = P(\{\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y\}),\ x, y \in \mathbb{R}$

Eigenschaften

1. $\lim\limits_{x \to -\infty} F(x,y) = \lim\limits_{y \to -\infty} F(x,y) = \lim\limits_{x,y \to -\infty} F(x,y) = 0,\ \boxed{\lim\limits_{x,y \to \infty} F(x,y) = 1}$

2. $\boxed{\lim\limits_{y \to \infty} F(x,y) = F^X(x),\ \lim\limits_{x \to \infty} F(x,y) = F^Y(y)}$

Beweis:

$$A = \{\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y\} = A_x \cap A_y \text{ mit } \begin{aligned} A_x &= \{\omega \in \Omega | X(\omega) \leq x\} \\ A_y &= \{\omega \in \Omega | Y(\omega) \leq y\} \end{aligned}$$

$$F(x,y) = P(A) = P(A_x \cap A_y) = 1 - P(A_x^c \cup A_y^c)$$

$$\boxed{\lim_{x \to \infty} F(x,y)} = 1 - P(A_\infty^c \cup A_y^c) = 1 - P(\emptyset \cup A_y^c) = 1 - P(A_y^c) = P(A_y) = \boxed{F^Y(y)}$$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

$P^{(X,Y)}(B) = P((X,Y) \in B) = P(\{\omega \in \Omega | (X(\omega), Y(\omega)) \in B\}),\ B \subseteq \mathbb{R}^2$

$F(x,y) = P(X \leq x, Y \leq y) = P(\{\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y\}),\ x, y \in \mathbb{R}$

Eigenschaften

1. $\lim\limits_{x \to -\infty} F(x,y) = \lim\limits_{y \to -\infty} F(x,y) = \lim\limits_{x,y \to -\infty} F(x,y) = 0,\ \boxed{\lim\limits_{x,y \to \infty} F(x,y) = 1}$

$\boxed{2.\ \lim\limits_{y \to \infty} F(x,y) = F^X(x),\ \lim\limits_{x \to \infty} F(x,y) = F^Y(y)}$

Beweis (Fortsetzung):

$$\boxed{\lim_{x \to \infty} F(x,y) = F^Y(y)}$$

$$\boxed{\lim_{x,y \to \infty} F(x,y)} = \lim_{y \to \infty} F^Y(y) = \boxed{1}$$

$\square$

Beweis für $\lim\limits_{y \to \infty} F(x,y) = F^X(x)$ analog.

$F^X(x)$ und $F^Y(y)$ heißen **Randverteilungen** von $X$ und $Y$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

$P^{(X,Y)}(B) = P((X,Y) \in B) = P(\{\omega \in \Omega | (X(\omega), Y(\omega)) \in B\}),\ B \subseteq \mathbb{R}^2$

$F(x,y) = P(X \leq x, Y \leq y) = P(\{\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y\}),\ x, y \in \mathbb{R}$

Eigenschaften

1. $\lim\limits_{x \to -\infty} F(x,y) = \lim\limits_{y \to -\infty} F(x,y) = \lim\limits_{x,y \to -\infty} F(x,y) = 0,\ \lim\limits_{x,y \to \infty} F(x,y) = 1$
2. $\lim\limits_{y \to \infty} F(x,y) = F^X(x),\ \lim\limits_{x \to \infty} F(x,y) = F^Y(y)$
3. $x_1 < x_2 \Rightarrow F(x_1,y) \leq F(x_2,y),\ y_1 < y_2 \Rightarrow F(x,y_1) \leq F(x,y_2)$

Beweis

$F(x_i,y) = P(A_i)$ mit $A_i = \{\omega \in \Omega | X(\omega) \leq x_i, Y(\omega) \leq y\}$

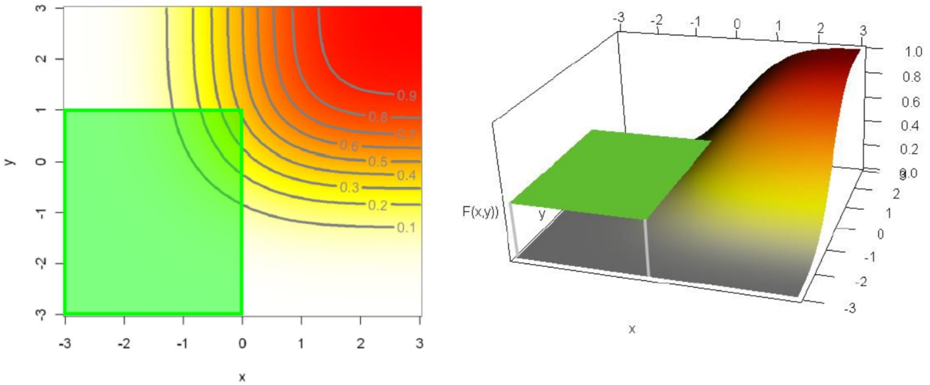
Beweis für $F(x,y_1)$ analog

$\boxed{x_1 < x_2} \Rightarrow A_1 \subseteq A_2 \Rightarrow P(A_1) \leq P(A_2) \Leftrightarrow \boxed{F(x_1,y) \leq F(x_2,y)}$ $\square$

# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung mehrdimensionaler Zufallsvariablen

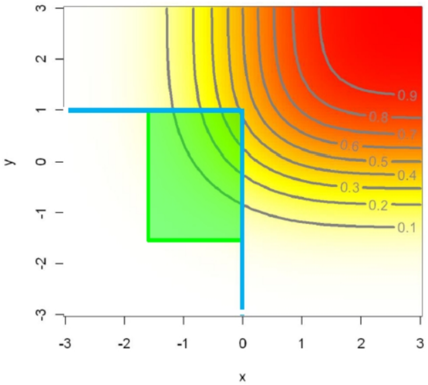
$$\boxed{F(x,y)} = \boxed{P(\boxed{X \leq x\,,\; Y \leq y})}\,,\quad x,y \in \Re$$



# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung mehrdimensionaler Zufallsvariablen

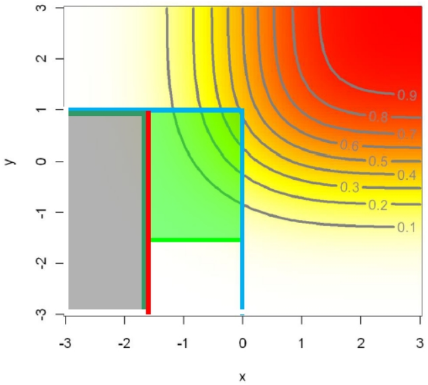
$$P(\boxed{x_1} < X \leq \boxed{x_2}\ ,\ \boxed{y_1} < Y \leq \boxed{y_2}) = P(X \leq \boxed{x_2}\ ,\ Y \leq \boxed{y_2})$$



# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung mehrdimensionaler Zufallsvariablen

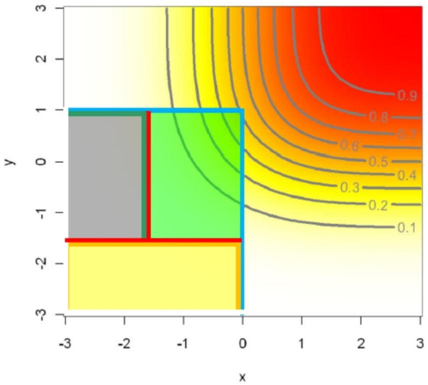
$$P(x_1 < X \leq x_2\, ,\, y_1 < Y \leq y_2) = P(X \leq x_2\, ,\, Y \leq y_2) - P(X \leq x_1\, ,\, Y \leq y_2)$$



# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung mehrdimensionaler Zufallsvariablen

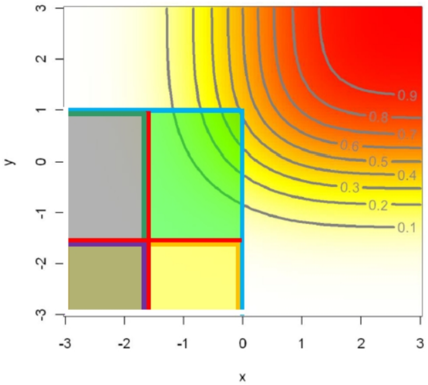
$$P(\boxed{x_1} < X \leq \boxed{x_2}, \boxed{y_1} < Y \leq \boxed{y_2}) = \boxed{P(X \leq \boxed{x_2}, Y \leq \boxed{y_2})}$$
$$- \boxed{P(X \leq \boxed{x_1}, Y \leq \boxed{y_2})}$$
$$- \boxed{P(X \leq \boxed{x_2}, Y \leq \boxed{y_1})}$$
$$+ \boxed{P(X \leq \boxed{x_1}, Y \leq \boxed{y_1})}$$



# 6.0 Zufallsvariablen und deren Verteilung

## Verteilung mehrdimensionaler Zufallsvariablen

$$
\begin{aligned}
P(x_1 < X \leq x_2\,,\ y_1 < Y \leq y_2) &= P(X \leq x_2\,,\ Y \leq y_2) \\
&- P(X \leq x_1\,,\ Y \leq y_2) \\
&- P(X \leq x_2\,,\ Y \leq y_1) \\
&+ P(X \leq x_1\,,\ Y \leq y_1) \\
&= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)
\end{aligned}
$$

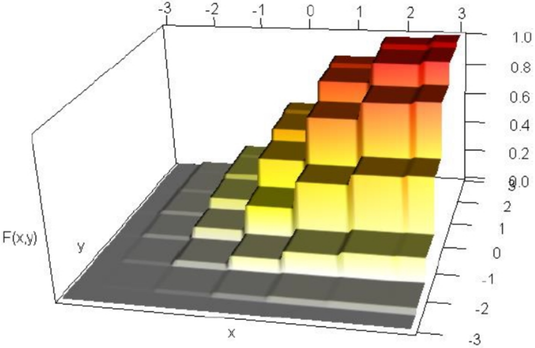


# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar): $\Omega = \{\omega_1, \ldots, \omega_n\}$

$$\Rightarrow X \in \{X(\omega_1), \ldots, X(\omega_n)\} = \{x_1, \ldots, x_n\} \text{ mit } -\infty < x_1 \leq \ldots \leq x_n < \infty$$
$$Y \in \{Y(\omega_1), \ldots, Y(\omega_n)\} = \{y_1, \ldots, y_n\} \text{ mit } -\infty < y_1 \leq \ldots \leq y_n < \infty$$



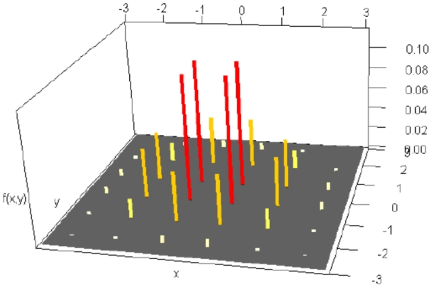
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

**Spezialfall diskrete Verteilungsfunktion** ($\Omega$ abzählbar): $\Omega = \{\omega_1, \ldots, \omega_n\}$

$$\Rightarrow X \in \{X(\omega_1), \ldots, X(\omega_n)\} = \{x_1, \ldots, x_n\} \text{ mit } -\infty < x_1 \leq \ldots \leq x_n < \infty$$
$$Y \in \{Y(\omega_1), \ldots, Y(\omega_n)\} = \{y_1, \ldots, y_n\} \text{ mit } -\infty < y_1 \leq \ldots \leq y_n < \infty$$

Die Funktion $p: \mathbb{R}^2 \to [0,1]$ mit $p(x,y) = P(X = x, Y = y)$ heißt **Zähldichte von** $(X, Y)$



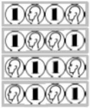
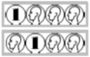

$$p(x,y) = \begin{cases} F(x_i,y_i) - F(x_{i-1},y_i) & , \quad x_i \in \{x_1,\ldots,x_n\} \\ -F(x_i,y_{i-1}) + F(x_{i-1},y_{i-1}) & \quad y_i \in \{y_1,\ldots,y_n\} \\ 0 & , \quad \text{sonst} \end{cases}$$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen, diskrete Verteilungsfunktion**

Beispiel: **4-facher Münzwurf**, $X$ =Anzahl Kopf nach 4 Würfen, $Y$ =Anzahl Kopf nach 2 Würfen

**Zähldichte**

| ↓y x→ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 |  |  |  | | |
| 1 | |  |  |  | |
| 2 | | |  |  |  |

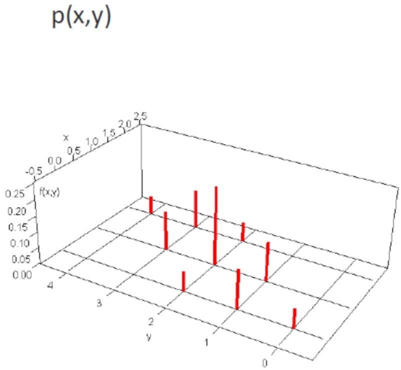
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen, diskrete Verteilungsfunktion**

Beispiel: **4-facher Münzwurf**, $X$ =Anzahl Kopf nach 4 Würfen, $Y$ =Anzahl Kopf nach 2 Würfen

**Zähldichte**

| ↓y x→ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| **0** | 1/16 | 2/16 | 1/16 | | |
| **1** | | 2/16 | 4/16 | 2/16 | |
| **2** | | | 1/16 | 2/16 | 1/16 |



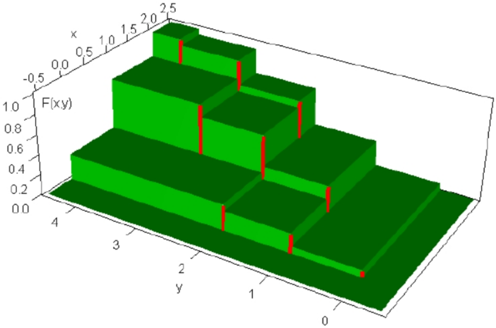
# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen, diskrete Verteilungsfunktion**

Beispiel: **4-facher Münzwurf**, $X$ =Anzahl Kopf nach 4 Würfen, $Y$ =Anzahl Kopf nach 2 Würfen

**Zähldichte**

| ↓y x→ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| **0** | 1/16<br>**1/16** | 2/16<br>**3/16** | 1/16<br>**4/16** | | |
| **1** | | 2/16<br>**5/16** | 4/16<br>**10/16** | 2/16<br>**12/16** | |
| **2** | | | 1/16<br>**11/16** | 2/16<br>**15/16** | 1/16<br>**16/16** |




p(x,y)
**F(x,y)**

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)

$\omega \in \Omega : \ (X(\omega), Y(\omega)) \in B, \ B \subseteq \mathbb{R}^2$
$F = F^{XY} : \mathbb{R}^2 \to [0,1]$ mit
$F(x,y) = P(X \leq x, Y \leq y) = P(\{\omega \in \Omega | X(\omega) \leq x, Y(\omega) \leq y\}), \ x, y \in \mathbb{R}$

Die Funktion $f : \mathbb{R}^2 \to [0,1]$ mit $f(x,y) = \frac{\delta^2 F(x,y)}{\delta x \delta y}$
heißt die **gemeinsame Dichtefunktion** von $X$ und $Y$.

Es gilt:

$$F(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)\, dt\, ds, \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s,t)\, dt\, ds = 1$$

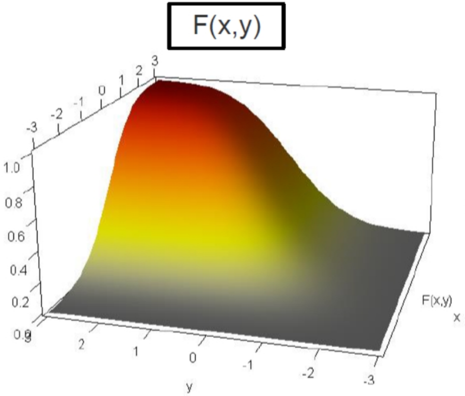
Die **Randdichten** $f^X$ und $f^Y$ von $X$ und $Y$ sind definiert durch

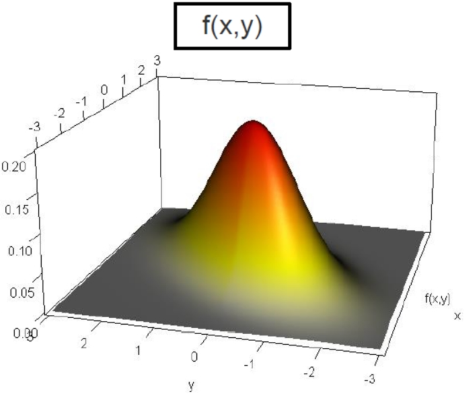
$$f^X(x) = \int_{-\infty}^{\infty} f(x,t)\, dt \text{ und } f^Y(y) = \int_{-\infty}^{\infty} f(s,y)\, ds$$

# 6.0 Zufallsvariablen und deren Verteilung

**Verteilung mehrdimensionaler Zufallsvariablen**

**Spezialfall stetige Verteilungsfunktion** ($\Omega$ überabzählbar)



# Zufallsvariablen und deren Verteilung

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

**Wahrscheinlichkeitsverteilung** einer Zufallsvariablen $X$:

$$P^X(B) = P(X \in B) = P(\{\omega \in \Omega | X(\omega) \in B\}),\ B \subseteq \mathbb{R}$$

**Verteilungsfunktion** von $X$

$$F = F^X : \mathbb{R} \to [0,1]$$
$$\text{mit } F(x) = P^X((-\infty, x]) = P(X \leq x) = P(\{\omega \in \Omega | X(\omega) \leq x\}),\ x \in \mathbb{R}$$

**Träger** $T_X$ einer diskreten Wahrscheinlichkeitsverteilung $P^X$

$$T_X = \{x | P(X = x) > 0\}$$

**Träger** $T_X$ einer stetigen Wahrscheinlichkeitsverteilung $P^X$

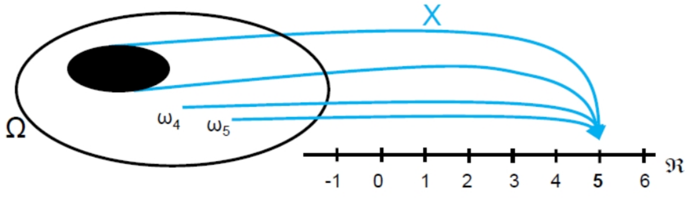
$$T_X = \{x | f(x) > 0\}$$

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

### Einpunktverteilung $\varepsilon_a$

$\omega \in \Omega \Rightarrow X(\omega) = a$



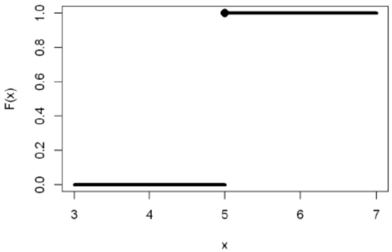

### Bestimmung der Dichte über Laplace-Raum

Setze $\Omega' = \{\omega_1\}$, $\omega \in \Omega' \Rightarrow \omega = \omega_1 \Rightarrow X(\omega_1) = a \Rightarrow p(X = a) = 1/|\Omega'| = 1$

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

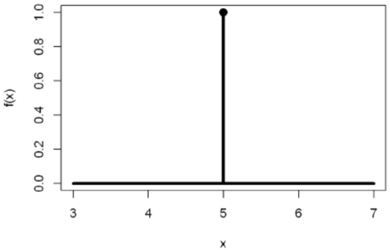
**Einpunktverteilung** $\varepsilon_a$




**Verteilungsfunktion:** $F(x) = \mathtt{I}(a \leq x)$

**Zähldichte:** $p(x) = \mathtt{I}(a = x)$

**Träger:** $T_X = \{a\}$

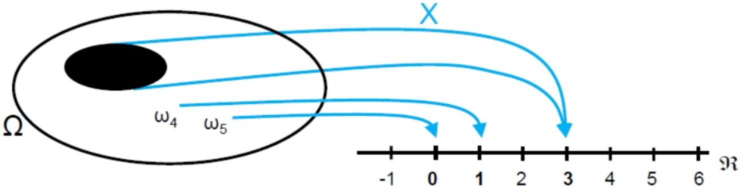


# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

### Diskrete Gleichverteilung $G(x_1, \ldots, x_n)$

$$P(\{\omega \in \Omega | X(\omega) = x_1\}) = P(\{\omega \in \Omega | X(\omega))x_2\}) = \ldots = P(\{\omega \in \Omega | X(\omega) = x_n\})$$




### Bestimmung der Dichte über Laplace-Raum

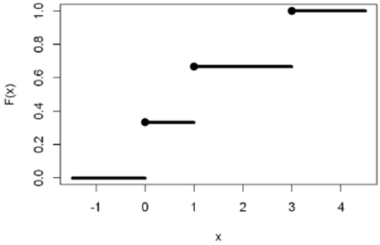
$$\text{Setze } \Omega' = \{\omega_1, \ldots, \omega_n\},\ \omega \in \Omega' \Rightarrow \omega = \omega_i \Rightarrow X(\omega) = x_i \Rightarrow p(X = x_i)$$
$$= 1/|\Omega'| = 1 = n$$

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

### Diskrete Gleichverteilung

$G(x_1, \dots, x_n)$




**Verteilungsfunktion:**

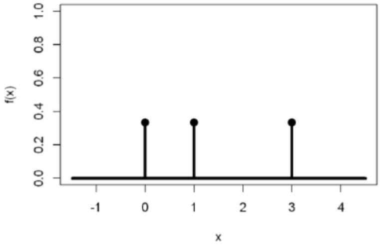
$$F(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(x_i \leq x)$$

**Zähldichte:**

$$p(x) = \frac{1}{n} \cdot \mathbb{I}(x \in \{x_1, \dots, x_n\})$$

**Träger:**

$$T_X = \{x_1, \dots, x_n\}$$

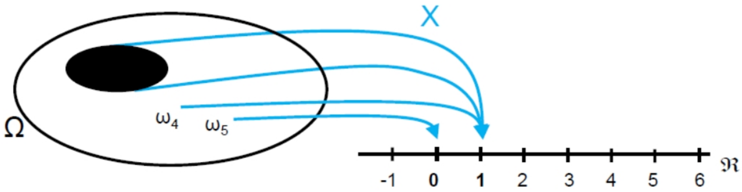


# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Bernoulli-Verteilung** $B(1, p)$

$$P(\{\omega \in \Omega | X(\omega) = 1\}) = p, \quad P(\{\omega \in \Omega | X(\omega) = 0\}) = 1 - p$$




Der Fall $X(\omega) = 1$ wird Erfolg genannt, entsprechend ist $p$ die Erfolgswahrscheinlichkeit.

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Bernoulli-Verteilung** $B(1, p)$
$P(\{\omega \in \Omega | X(\omega) = 1\}) = p, \; P(\{\omega \in \Omega | X(\omega) = 0\}) = 1 - p$

**Bestimmung der Dichte über Laplace-Raum**

Sei $p$ o.B.d.A. rational,
d.h. es gibt $r, g \in \mathbb{N}, r < g$, mit $p = r/g$

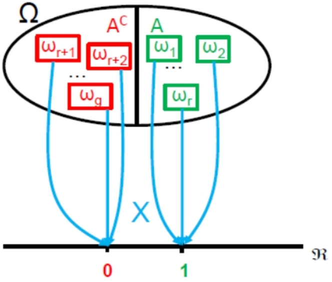
Setze $\Omega' = \{\omega_1, \ldots, \omega_g\}$

$\boxed{A = \{\omega_1, \ldots, \omega_r\}}, \; \boxed{A^c = \{\omega_{r+1}, \ldots, \omega_g\}},$

$X(\omega) = \mathrm{I}(\omega \in A)$

$$\Rightarrow p(X = 1) = \boxed{P(A)} = \sum_{i=1}^{r} 1/|\Omega'| = r/g = p$$

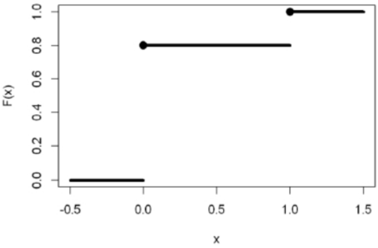
$$p(X = 0) = \boxed{P(A^c)} = 1 - r/g = 1 - p$$



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Bernoulli-Verteilung** $B(1, p)$,
$0 < p < 1$



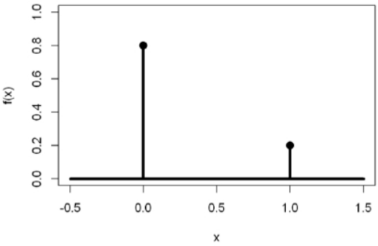

**Verteilungsfunktion:**

$$F(x) = (1 - p) \cdot \mathtt{I}(0 \leq x) + p \cdot \mathtt{I}(1 \leq x)$$

**Zähldichte:**

$$p(x) = p \cdot \mathtt{I}(x = 1) + (1 - p) \cdot \mathtt{I}(x = 0)$$

**Träger:** $T_X = \{0, 1\}$



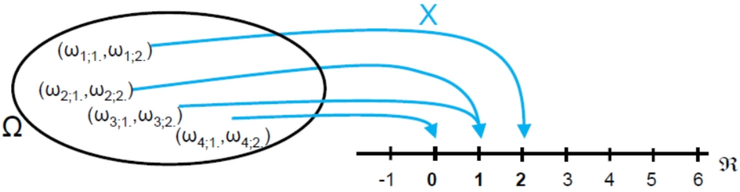
# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Binomialverteilung** $B(n, p)$

$$P(\{\omega = \{\omega_{1.}, \dots, \omega_{n.}\} \in \Omega = (\Omega')^n | X(\omega) = \sum_{i=1}^{n} Y(\omega_{i.}) = k\}),$$

wobei $Y \sim B(1, p)$, ist die W'keit für genau $k$ Erfolge in $n$ unabhängigen Zufallsexperimenten mit $B(1, p)$-verteilter Zufallsvariable



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Binomialverteilung** $B(n, p)$

**Bestimmung der Dichte über Laplace-Raum**

$r, g \in \mathbb{N}$, $p = r/g$, $\Omega' = \{\omega_1, \ldots, \omega_g\}$, $A = \{\omega_1, \ldots, \omega_r\}$, $A^c = \{\omega_{r+1}, \ldots, \omega_g\}$,
$Y(\omega_i) = \mathtt{I}(\omega_i \in A),\ i = 1, \ldots, n$

Elementarereignisse insgesamt: $|\Omega| = |\Omega'|^n = g^n$

Elementarereignisse mit Erfolg in ersten $k$ Versuchen, danach Misserfolge:
$|A^k \times (A^c)^{n-k}| = r^k \cdot (g - r)^{n-k}$

Dieselbe Anzahl ergibt sich für jede Verteilung der $k$ Erfolge auf die $n$ Experimente.

Anzahl der Anordnungen der $k$ Erfolge in den $n$ Versuchen: $n! =$ Anzahl der Reihenfolgen-Permutationen

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Binomialverteilung** $B(n, p)$

Die Reihenfolge der Erfolge/Misserfolge untereinander spielt dabei keine Rolle; es wird also jedes Elementarereignis $k! \cdot (n-k)!$ mal gezählt.

Binomialkoeffizient:

Anzahl Elementarereignisse mit insgesamt $k$ Erfolgen $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Wahrscheinlichkeit für $k$ Erfolge:

$$\boxed{P(X=k)} = \frac{r^k \cdot (g-r)^{n-k} \cdot \binom{n}{k}}{g^n} = \frac{r^k \cdot (g-r)^{n-k} \cdot \binom{n}{k}}{g^k g^{n-k}}$$

$$= \binom{n}{k} \left(\frac{r}{g}\right)^k \left(\frac{g-r}{g}\right)^{n-k} = \boxed{\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}}$$

$\boxed{\text{Beispiel } n=3, r=3, g=5, k=2 \Rightarrow P(X=2) = 54/125}$

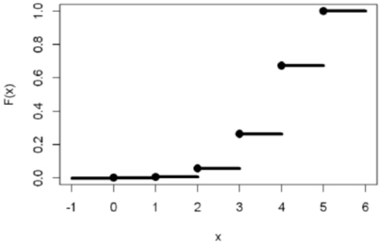
# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Binomialverteilung** $B(n, p)$,
$0 < p < 1, \; n \in \mathbb{N}$

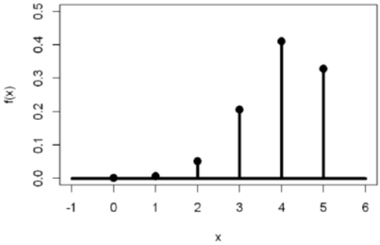
**Verteilungsfunktion:**

$$F(x) = \sum_{i=0}^{n} \mathbb{I}(i \leq x) \cdot \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$$




**Zähldichte:**

$$p(x) = \mathbb{I}(x \in \{0, \ldots, n\}) \cdot \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$$




**Träger:** $T_X = \{0, \ldots, n\}$

**Beispiel:** $X =$ Anzahl Treffer bei 5 Strafstößen

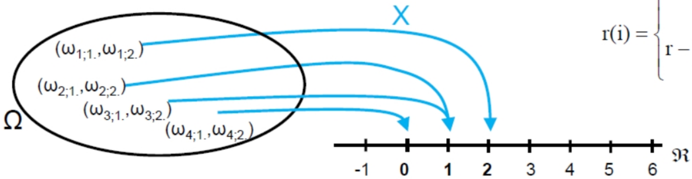
# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Hypergeometrsiche Verteilung** Hyp$(n, r, s)$

$$P(\{\omega = (\omega_{1,.}, \ldots, \omega_{n,.}) \in \Omega = \Omega' \times \Omega' \setminus \{\omega_{1,.}\} \times \ldots \times \Omega' \setminus \{\omega_{1,.}, \ldots, \omega_{n-1,.}\}| \quad X(\omega) = \sum_{i=1}^{n} Y_i(\omega) = k\}), \qquad Y_i \sim B\left(1, \tfrac{r(i)}{r+s-i+1}\right)$$

ist die W'keit für genau $k$ Erfolge in $n$ Zufallsexperimenten mit jeweils $B\left(1, \tfrac{r(i)}{r+s-i+1}\right)$-verteilter Zufallsvariable




$$r(i) = \begin{cases} r & , \quad i = 1 \\ r - \sum_{j=1}^{i-1} Y_j & , \quad i > 1 \end{cases}$$

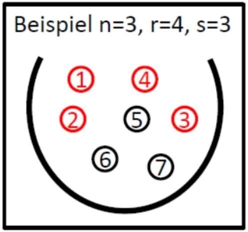
# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

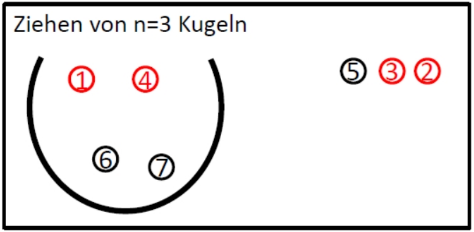
## Diskrete Wahrscheinlichkeitsverteilungen

**Hypergeometrsiche Verteilung** Hyp$(n, r, s)$

Urnenmodell

$k =$ Anzahl rote Kugeln nach $n$-maligem Ziehen aus einer Urne mit $r$ roten und $s$ schwarzen Kugeln ohne Zurücklegen



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

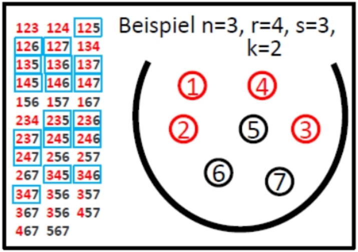
**Hypergeometrsiche Verteilung** $\mathrm{Hyp}(n, r, s)$

Urnenmodell

$k =$ Anzahl rote Kugeln nach $n$-maligem Ziehen aus einer Urne mit $r$ roten und $s$ schwarzen Kugeln ohne Zurücklegen

$\Omega$ sind alle Möglichkeiten, die $n$ „Ziehungen“ auf die $r + s$ Kugeln zu verteilen.
Davon gibt es $|\Omega| = \binom{r+s}{n}$.

Die günstigen Fälle sind alle, in denen $k$ rote und $n - k$ schwarze Kugeln auf die $n$ gezogenen verteilt werden.
Davon gibt es $|\{\omega \in \Omega | X(\omega) = k\}| = \binom{r}{k} \cdot \binom{s}{n-k}$.



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Hypergeometrsiche Verteilung** $\text{Hyp}(n, r, s)$

Urnenmodell

$k =$ Anzahl rote Kugeln nach $n$-maligem Ziehen aus einer Urne mit $r$ roten und $s$ schwarzen Kugeln ohne Zurücklegen

Es gibt $|\Omega| = \binom{r+s}{n}$ Möglichkeiten, die $n$ „Ziehungen" auf die $r + s$ Kugeln zu verteilen.

Davon enthalten $|\{\omega \in \Omega | X(\omega) = k\}| = \binom{r}{k} \cdot \binom{s}{n-k}$ Kombinationen $k$ rote Kugeln.

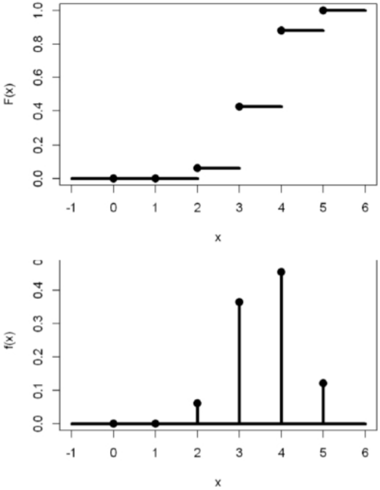
Die W'keit für $k$ rote Kugeln beträgt also

$$P(X = k) = \frac{\binom{r}{k} \cdot \binom{s}{n-k}}{\binom{r+s}{n}}$$

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Hypergeometrische Verteilung** $\text{Hyp}(n, r, s)$,
$0 < n < r + s, \;\; n \in \mathbb{N}$




**Verteilungsfunktion:**

$$F(x) = \sum_{i=0}^{n} \mathbb{I}(i \leq x) \cdot \frac{\binom{r}{i} \cdot \binom{s}{n-i}}{\binom{r+s}{n}}$$

**Zähldichte:**

$$p(x) = \mathbb{I}(x \in \{\max(0, n-s), \ldots, \min(n, r)\}) \cdot \frac{\binom{r}{x} \cdot \binom{s}{n-x}}{\binom{r+s}{n}}$$

**Träger:** $T_X = \{\max(0, n-s), \ldots, \min(n, r)\}$

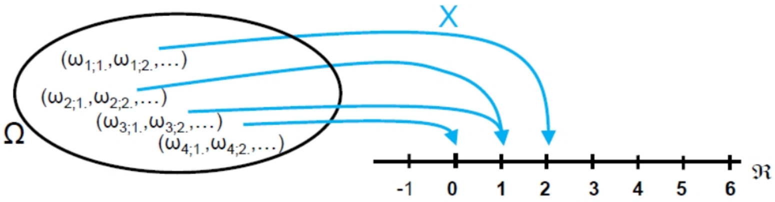
# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

### Geometrische Verteilung Geo($p$)

$P(\{\omega = (\omega_{1.}, \omega_{2.}, \ldots) \in \Omega = \Omega'^{\infty} | X(\omega) = \min[i | Y_i(\omega_{i.}) = 1] - 1 = k\})$,
$Y_i \sim B(1, p)$

ist die W'keit, den ersten Erfolg im $(k+1)$-ten Zufallsexperiment mit jeweils $B(1, p)$-verteilter Zufallsvariable zu erhalten.



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Geometrische Verteilung** Geo($p$)

$r, g \in \mathbb{N}$, $p = r/g$, $\Omega' = \{\omega_1, \ldots, \omega_g\}$, $A = \{\omega_1, \ldots, \omega_r\}$, $A^c = \{\omega_{r+1}, \ldots, \omega_g\}$,
$Y(\omega_i) = \mathtt{I}(\omega_i \in A),\ i = 1, \ldots, n$

Für festes $k$ relevanter Grundraum $\quad \Omega'_{k+1} = |\Omega'|^{k+1} \Rightarrow |\Omega'_{k+1}| = g^{k+1}$

Günstige Elementarereignisse sind diejenigen mit ausschließlich Misserfolgen in den ersten $k$ Versuchen, danach ein Erfolg:

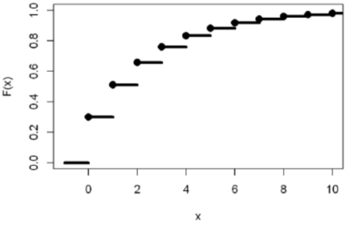
$$\left|\left(A^c\right)^k \times A\right| = r \cdot (g-r)^k$$

$$\Rightarrow \boxed{P(X = k)} = \frac{r \cdot (g-r)^k}{g^{k+1}} = \frac{r \cdot (g-r)^k}{g \cdot g^k} = \left(\frac{r}{g}\right) \cdot \left(\frac{g-r}{g}\right)^k = \boxed{p \cdot (1-p)^k}$$

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Geometrische Verteilung** Geo($p$), $0 < p < 1$

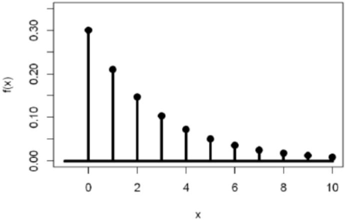



**Verteilungsfunktion:**

$$F(x) = 1 - (1-p)^{x+1}$$

**Zähldichte:**

$$p(x) = \mathbb{I}(x \in \mathbb{N} \cup 0) \cdot p \cdot (1-p)^x$$




**Träger:** $T_X = \mathbb{N} \cup 0$

**Beispiel:** $X =$ Anzahl verschossener Strafstöße bis zum ersten verwandelten

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

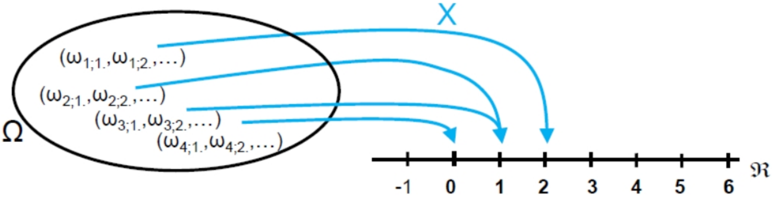
## Diskrete Wahrscheinlichkeitsverteilungen

**Poisson-Verteilung** Poi($\lambda$) $\quad P(\{\omega = (\omega_{1.}, \omega_{2.}, \ldots) \in \Omega = \Omega'^{\infty} | X(\omega) = k\})$

$$= \lim_{n\to\infty} P(\{\omega = (\omega_{1.}, \ldots, \omega_{n.}) \in \Omega = \Omega'^{n} | X(\omega) = \sum_{i=1}^{n} Y_n(\omega_{i.}) = k\}),$$

$Y_n \sim B(1, \lambda/n)$

ist die Wahrscheinlichkeit, in einem festen Zeitraum der Länge $t$ genau $k$ Ereignisse zu beobachten, wenn die mittlere Anzahl an Ereignissen im Zeitraum $t$ durch den Wert von $\lambda$ gegeben ist und ein Erfolg zu jedem Zeitpunkt im betrachteten Zeitraum gleichwahrscheinlich ist.



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

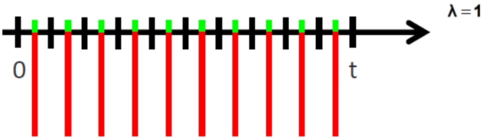
## Diskrete Wahrscheinlichkeitsverteilungen

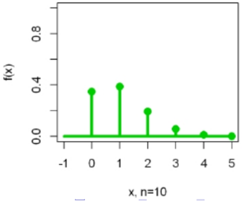
**Poisson-Verteilung** Poi($\lambda$) $\quad P(\{\omega = (\omega_{1.}, \omega_{2.}, \ldots) \in \Omega = \Omega'^{\infty} | X(\omega) = k\})$

$$= \lim_{n\to\infty} P(\{\omega = (\omega_{1.}, \ldots, \omega_{n.}) \in \Omega = \Omega'^{n} | X(\omega) = \sum_{i=1}^{n} Y_n(\omega_{i.}) = k\}),\ Y_n \sim B(1, \lambda/n)$$

ist die Wahrscheinlichkeit, in einem festen Zeitraum der Länge $t$ genau $k$ Ereignisse zu beobachten, wenn die mittlere Anzahl an Ereignissen im Zeitraum $t$ durch den Wert von $\lambda$ gegeben ist und ein Erfolg zu jedem Zeitpunkt im betrachteten Zeitraum gleichwahrscheinlich ist.

Durchführung von $n$ Bernoulli-Experimenten im Zeitraum $t$:
$p(x) = \mathtt{I}(x \in \{0, \ldots, n\}) \cdot \binom{n}{x} \cdot (\frac{\lambda}{n})^x \cdot (1 - \frac{\lambda}{n})^{n-x}$



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

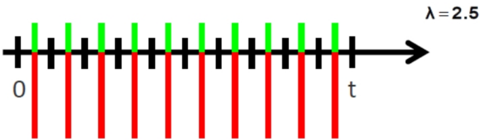
## Diskrete Wahrscheinlichkeitsverteilungen

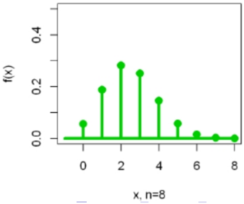
**Poisson-Verteilung** Poi($\lambda$) $\quad P(\{\omega = (\omega_{1.}, \omega_{2.}, \ldots) \in \Omega = \Omega'^{\infty} | X(\omega) = k\})$

$$= \lim_{n\to\infty} P(\{\omega = (\omega_{1.}, \ldots, \omega_{n.}) \in \Omega = \Omega'^{n} | X(\omega) = \sum_{i=1}^{n} Y_n(\omega_{i.}) = k\}),\ Y_n \sim B(1, \lambda/n)$$

ist die Wahrscheinlichkeit, in einem festen Zeitraum der Länge $t$ genau $k$ Ereignisse zu beobachten, wenn die mittlere Anzahl an Ereignissen im Zeitraum $t$ durch den Wert von $\lambda$ gegeben ist und ein Erfolg zu jedem Zeitpunkt im betrachteten Zeitraum gleichwahrscheinlich ist.

Durchführung von $n$ Bernoulli-Experimenten im Zeitraum $t$:
$p(x) = \mathbb{I}(x \in \{0, \ldots, n\}) \cdot \binom{n}{x} \cdot (\frac{\lambda}{n})^x \cdot (1 - \frac{\lambda}{n})^{n-x}$



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Poisson-Verteilung** Poi($\lambda$)

$$
\begin{aligned}
\boxed{p(x)} &= \lim_{n\to\infty}\left[\mathtt{I}(x\in\{0,\ldots,n\})\cdot\binom{n}{x}\cdot\left(\frac{\lambda}{n}\right)^{x}\cdot\left(1-\frac{\lambda}{n}\right)^{n-x}\right]\\
&= \mathtt{I}(x\in\mathbb{N}\cup 0)\cdot\lim_{n\to\infty}\left[\frac{n!}{x!(n-x)!}\cdot\left(\frac{\lambda}{n}\right)^{x}\cdot\left(1-\frac{\lambda}{n}\right)^{n-x}\right]\\
&= \mathtt{I}(x\in\mathbb{N}\cup 0)\cdot\lim_{n\to\infty}\left[\frac{n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-x+1)}{x!}\cdot\frac{\lambda^{x}}{n^{x}}\right.\\
&\quad\left.\cdot\left(1-\frac{\lambda}{n}\right)^{n}\cdot\left(1-\frac{\lambda}{n}\right)^{-x}\right]\\
&= \mathtt{I}(x\in\mathbb{N}\cup 0)\cdot\left(\frac{\lambda^{x}}{x!}\right)\cdot\lim_{n\to\infty}\left(\frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdot\ldots\cdot\frac{n-x+1}{n}\right)\\
&\quad\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n}\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x}\\
&= \mathtt{I}(x\in\mathbb{N}\cup 0)\cdot\left(\frac{\lambda^{x}}{x!}\right)\cdot 1\cdot e^{-\lambda}\cdot 1=\boxed{\frac{\lambda^{x}e^{-\lambda}}{x!}}
\end{aligned}
$$

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Poisson-Verteilung** Poi($\lambda$), $\lambda > 0$

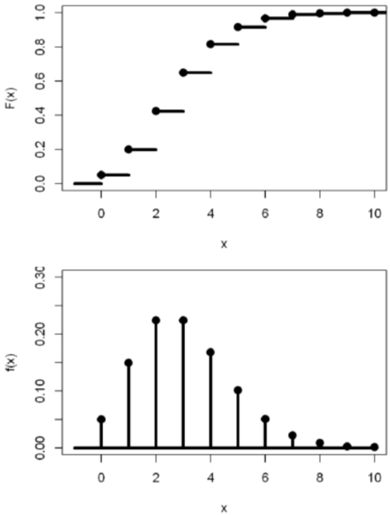
**Verteilungsfunktion:**

$$F(x) = e^{-\lambda} \cdot \sum_{i=0}^{n} \mathtt{I}(i \leq x) \cdot \frac{\lambda^x}{x!}$$

**Zähldichte:** $p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$

**Träger:** $T_X = \mathbb{N} \cup 0$

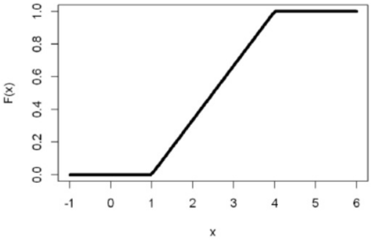
**Beispiel:** $X$ = Anzahl Tore pro Bundesliga-Spieltag



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

**Rechteckverteilung** $R(a, b)$, $a < b$

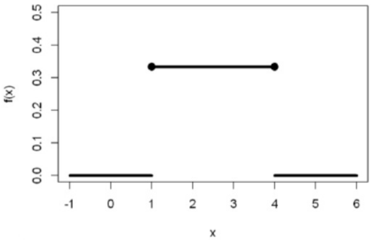



**Verteilungsfunktion:**

$$F(x) = \mathtt{I}(a \leq x) \cdot \frac{\min(x, b) - a}{b - a}$$

**Dichtefunktion:**

$$f(x) = \frac{\mathtt{I}(a \leq x \leq b)}{b - a}$$




**Träger:** $T_X = [a, b]$

**Beispiel:** $X$ = Zeitpunkt beliebiger Spielunterbrechung zwischen 10. und 20. Spielminute

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

**Exponentialverteilung** $\text{Exp}(\lambda)$, $\lambda > 0$

**Verteilungsfunktion:**

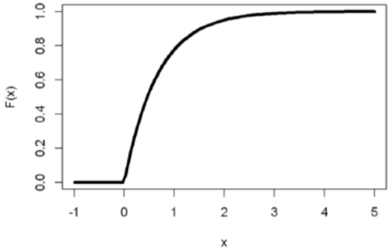
$$F(x) = \mathtt{I}(0 \leq x) \cdot (1 - e^{-(\lambda \cdot x)})$$

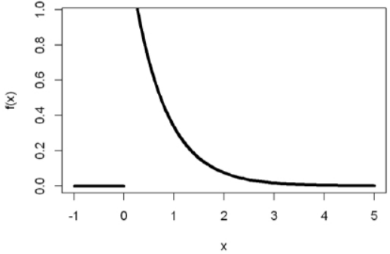
**Dichtefunktion:**

$$f(x) = \mathtt{I}(0 \leq x) \cdot \lambda e^{-(\lambda \cdot x)}$$

**Träger:** $T_X = [0, \infty)$

**Beispiel:** $X$ = Zeit zwischen zwei Spielunterbrechungen



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

### Weibull-Verteilung $W(\alpha, \beta)$

$P(\{\omega \in \Omega | X(\omega) \leq x)\}$
ist die Wahrscheinlichkeit, höchstens $x$ Zeiteinheiten auf das erste Ereignis zu warten, wenn die Hazardrate

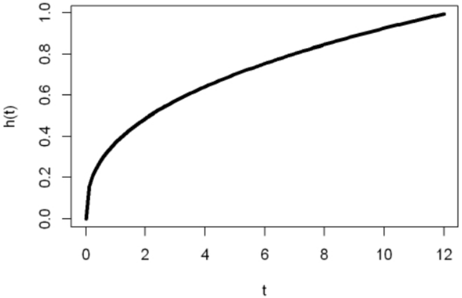
$$h(t) = \lim_{\delta \to \infty} \frac{P(t \leq T \leq t + \delta t)}{\delta t}$$

durch die Funktion

$$h(t) = e^{(v + \rho \cdot \ln(t))}$$

mit $v = \ln(\alpha) - \alpha \cdot \ln(\beta)$ und $\rho = \alpha - 1$

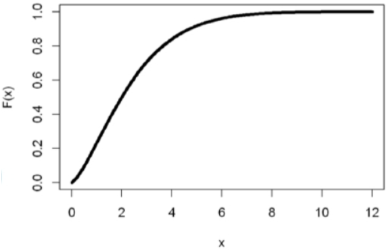
beschrieben wird.



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

**Weibull-Verteilung** $W(\alpha, \beta)$,
$\alpha > 0$, $\beta > 0$

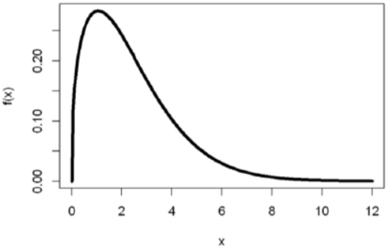



**Verteilungsfunktion:**

$$F(x) = \mathbb{I}(0 \leq x) \cdot (1 - e^{-(x/\beta)^\alpha})$$

**Dichtefunktion:**

$$f(x) = \mathbb{I}(0 \leq x) \cdot \left( \frac{\alpha}{\beta} \cdot \left( \frac{x}{\beta} \right)^{\alpha - 1} e^{-(x/\beta)^\alpha} \right)$$




**Träger:** $T_X = [0, \infty)$

**Beispiel:** $X$ = Zeitpunkt des Tors zum 1:0, wenn das Hinspiel 0:1 verloren wurde

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

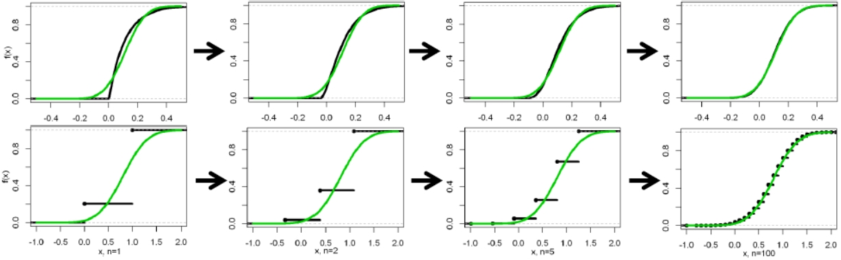
## Stetige Wahrscheinlichkeitsverteilungen

**Normalverteilung** $N(\mu, \sigma^2)$ $\quad P(\{\omega \in \Omega | X(\omega) \leq x\})$

$$= \lim_{n\to\infty} P\left(\{\omega \in \Omega'^n | \left(\frac{1}{n}\sum_{i=1}^{n}\left[\frac{Y(\omega_i)-E(X)}{\sqrt{\mathrm{Var}(X)}}\right]\right) \cdot \sqrt{\mathrm{Var}(X)\cdot n} + E(X) \leq x\}\right),$$

$$Y \sim F_0(E(X) = \mu, \mathrm{Var}(X) = \sigma^2)$$

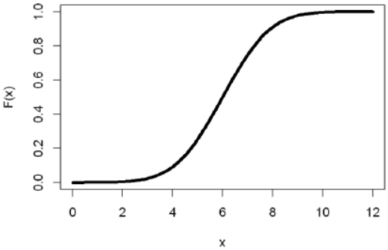
ist die Grenzw'keit für unendliches $n$, dass der Mittelwert aus $n$ standardisierten Zufallsvariablen, die aus der selben Verteilung stammen, den Wert $[x - E(X)]/[\mathrm{Var}(X \cdot n)^{0.5}]$ nicht übersteigt.



# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

**Normalverteilung** $N(\mu, \sigma^2)$, $\sigma > 0$

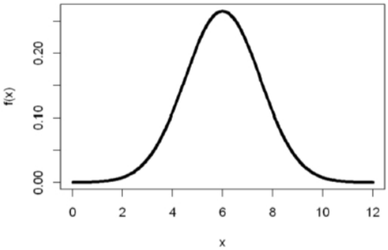



**Verteilungsfunktion:**

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

**Dichtefunktion:**

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




**Träger:** $T_X = \mathbb{R}$

**Beispiel:** $X =$ Nettospielzeit pro Bundesligaspielzeit

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

### $\chi^2$-Verteilung $\chi^2_f$ mit $f$ Freiheitsgraden

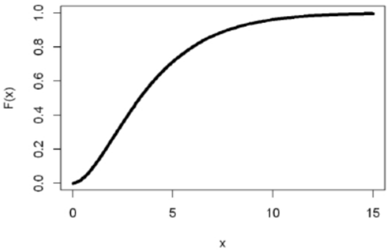
$$P(\{\omega \in \Omega | X(\omega) \leq x\}) = P\left(\{\omega \in \Omega'^f | \sum_{i=1}^{f} Y^2 \leq x\}\right), \; Y \sim N(0,1)$$

ist die Wahrscheinlichkeit, dass die Summe von $f$ quadrierten, standardnormalverteilten Zufallsvariablen den Wert $x$ nicht übersteigt.

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

$\chi^2$-**Verteilung**, $f \in \mathbb{N}$ $\quad (= \Gamma[0.5, f/2])$

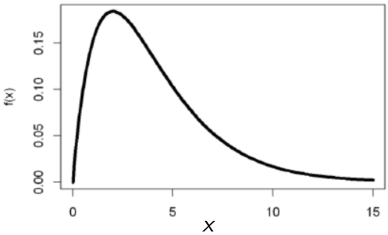



**Verteilungsfunktion:**

$$F(x) = \int_0^x \frac{1}{2^{f/2}\Gamma(f/2)} t^{f/2-1} e^{-t/2} dt$$

**Dichtefunktion:**

$$f(x) = \mathrm{I}(0 < x) \cdot \frac{1}{2^{f/2}\Gamma(f/2)} t^{f/2-1} e^{-x/2}$$




**Träger:** $T_X = [0, \infty)$

**Beispiel:** $X = \chi^2$-Statistik zwischen Anzahl gelber Karten pro Spiel in Hin- und Rückrunde

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

### $F$-**Verteilung** $F_{f_1,f_2}$ **mit** $f_1$ **und** $f_2$ **Freiheitsgraden**

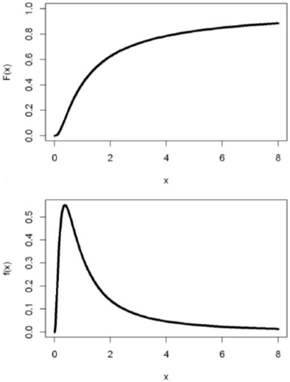
$$P(\{\omega \in \Omega | X(\omega) \leq x\}) = P\Big(\{\omega \in \Omega'^f | [Y \cdot f_2]/[Z \cdot f_1] \leq x\}\Big),\ Y \sim \chi^2_{f_1},\ Z \sim \chi^2_{f_2}$$

ist die Wahrscheinlichkeit, dass der Quotient zweier durch ihre Anzahl an Freiheitsgraden dividierter $\chi^2$-verteilter Zufallsvariablen den Wert $x$ nicht übersteigt.

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

$F$-**Verteilung** $F_{f_1,f_2}$ $f_1 \in \mathbb{N}$, $f_2 \in \mathbb{N}$




**Verteilungsfunktion:**

$$F(x) = \int_0^x \frac{\Gamma\left(\frac{f_1+f_2}{2}\right)}{\Gamma\left(\frac{f_1}{2}\right)\cdot\Gamma\left(\frac{f_2}{2}\right)} \cdot \left(\frac{f_1}{f_2}\right)^{f_1/2} \frac{t^{f_2/2-1}}{(1+f_1\, t/f_2)^{(f_1+f_2)/2}}\, dt$$

**Dichtefunktion:**

$$f(x) = \mathtt{I}(0<x)\cdot\frac{\Gamma\left(\frac{f_1+f_2}{2}\right)}{\Gamma\left(\frac{f_1}{2}\right)\cdot\Gamma\left(\frac{f_2}{2}\right)} \cdot \left(\frac{f_1}{f_2}\right)^{f_1/2} \frac{x^{f_2/2-1}}{(1+f_1\, x/f_2)^{(f_1+f_2)/2}}$$

**Träger:** $T_X = [0, \infty)$

**Beispiel:** $X =$ Verhältnis der mittleren Varianz der Zuschauerzahlen pro Spiel an jeweils einem Spieltag zur Varianz der mittleren Zuschauerzahlen pro Spieltag

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

### $t$-Verteilung $t_f$ mit $f$ Freiheitsgraden

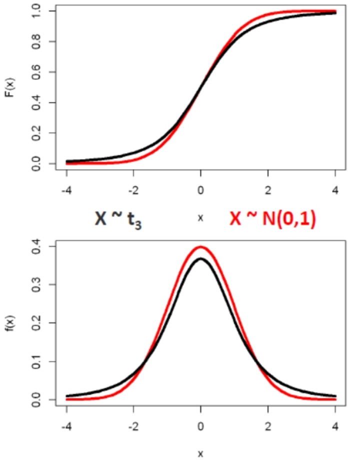
$$P(\{\omega \in \Omega | X(\omega) \leq x\}) = P\left(\{\omega \in \Omega'^{f+1} | \frac{Z}{\sqrt{Y/f}} \leq x\}\right), \ Z \sim N(0,1), \ Y \sim \chi^2_f$$

ist die Wahrscheinlichkeit, dass der Quotient einer standardnormalverteilten Zufallsvariable und der Wurzel aus einer durch die Anzahl ihrer Freiheitsgrade dividierter $\chi^2$-verteilter Zufallsvariablen den Wert $x$ nicht übersteigt.

# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Stetige Wahrscheinlichkeitsverteilungen

### $t$-Verteilung $t_f$ $f \in \mathbb{N}$




**Verteilungsfunktion:**

$$F(x) = \int_0^x \frac{\Gamma\left(\frac{f+1}{2}\right)}{\sqrt{f\pi} \cdot \Gamma\left(\frac{f}{2}\right)} \cdot \left(1 + \frac{t^2}{f}\right)^{-(f+1)/2} dt$$

**Dichtefunktion:**

$$f(x) = \frac{\Gamma\left(\frac{f+1}{2}\right)}{\sqrt{f\pi} \cdot \Gamma\left(\frac{f}{2}\right)} \cdot \left(1 + \frac{x^2}{f}\right)^{-(f+1)/2}$$

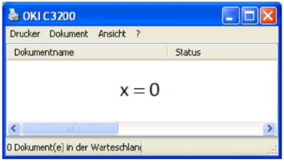
**Träger:** $T_X = \mathbb{R}$

**Beispiel:** $X =$ Geeignet skalierte Differenz aus mittlerer Nettospielzeit an geraden und ungeraden Spieltagen

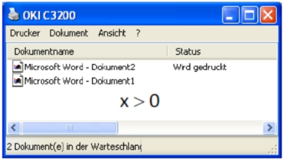
# 7.1 Wichtige Wahrscheinlichkeitsverteilungen

## Kombinationen aus diskreten und stetigen Wahrscheinlichkeitsverteilungen

Beispiel **Wartezeit bis zum Start eines Druckauftrags**

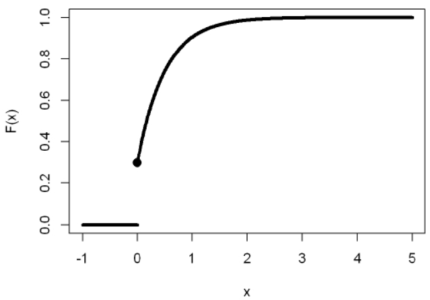



$P(X = 0) = 1 - p$




$P(0 < X < x) = p - pe^{-\lambda x}$

**Verteilungsfunktion:** $F(X) = \mathbf{I}(0 \leq x) \cdot (1 - pe^{-\lambda x})$



# 7.2 Zwei- und Mehrdimensionale W'keitsverteilungen

## Diskrete Wahrscheinlichkeitsverteilungen

**Multinomialverteilung Mult**$(n, p_1, \ldots, p_r)$

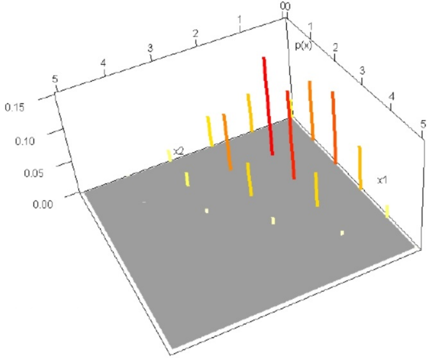
$$n \in \mathbb{N},\ 0 \leq p_i \leq 1, i = 1, \ldots, r,\ \sum_{i=1}^{r} p_i = 1$$

**Zähldichte:**

$$p(x_1, \ldots, x_r) = \mathrm{I}(x \in T_X) \cdot \frac{n!}{\prod_{i=1}^{r} x_i!} \cdot \prod_{i=1}^{r} p_i$$

**Träger:**

$$T_X = \left\{ x = (x_1, \ldots, x_r) \in (\mathbb{N} \cup \{0\})^r \,\middle|\, \sum_{i=1}^{r} x_i = n \right\}$$




**Beispiel:** $X$ = Vektor mit Anzahlen geschossener Tore in einer Saison für alle Spieler im Kader

# 7.2 Zwei- und Mehrdimensionale W'keitsverteilungen

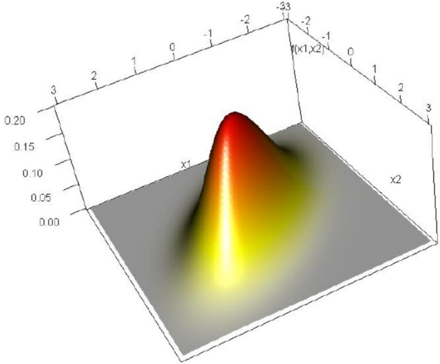
## Stetige Wahrscheinlichkeitsverteilungen

**Multivariate Normalverteilung** $N(\mu, \Sigma)$

$\mu \in \mathbb{R}^r$, $\Sigma \in \mathbb{R}^{r \times r}$, $\Sigma$ p.s.d

**Dichtefunktion:**

$$f(x_1, \ldots, x_r) = \frac{1}{(2\pi)^{k/2}\sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)}$$




**Träger:** $T_X = \mathbb{R}^r$

**Beispiel:** $X =$ Zuschauerzahl, Etat und Summe der Spielergehälter

# Wahrscheinlichkeits-theoretische Kennzahlen

# 8.1 Erwartungswert und Varianz

Zuletzt: Abbildung von Ereignissen auf reelle Zahlen durch Zufallsvariablen; mathematische Formulierung ihrer W'keiten durch Verteilungsfunktionen

$$\begin{aligned} X : \Omega &\to \mathbb{R} \\ \omega &\mapsto X(\omega) \end{aligned} \qquad \begin{aligned} &F^X \in \mathcal{F} \\ &F^X : \mathbb{R} \to [0,1] \\ &\qquad x \mapsto F^X(x) \end{aligned}$$

Jetzt: **Charakterisierung der Verteilungen durch einzelne Parameter**

$$\begin{aligned} \theta : &\mathcal{F} \to \mathbb{R} \\ &F^X \mapsto \theta(F^X) \end{aligned}$$

# 8.1 Erwartungswert und Varianz

## Diskrete Wahrscheinlichkeitsverteilungen

Der **Erwartungswert** einer diskret verteilten Zufallsvariable $X$ mit Zähldichte $p(x)$ und Träger $T_X = \{x_1, x_2, \ldots\}$ ist definiert durch

$$\mathrm{E}[X] = \sum_{j=1}^{J} p(x_j) \cdot x_j, \ J \in \mathbb{N} \cup \{\infty\}$$

Die **Varianz** von $X$ ist definiert durch

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = \sum_{j=1}^{J} p(x_j) \cdot (x_j - \mathrm{E}[X])^2, \ J \in \mathbb{N} \cup \{\infty\}$$

Die **Standardabweichung** von $X$ ist definiert durch

$$\sqrt{\mathrm{Var}(X)}$$

# 8.1 Erwartungswert und Varianz

## Diskrete Wahrscheinlichkeitsverteilungen

Der **Erwartungswert** ist der Wert, den man „im Mittel“ annimmt.
Beispiel einmaliger Würfelwurf:

$$\begin{aligned}\mathrm{E}[X] &= P(X=1)\cdot 1 + P(X=2)\cdot 2 + P(X=3)\cdot 3\\ &\quad + P(X=4)\cdot 4 + P(X=5)\cdot 5 + P(X=6)\cdot 6\\ &= \frac{1}{6}(1+2+3+4+5+6) = 3.5\end{aligned}$$

Die **Varianz** ist die erwartete quadratische Abweichung vom Erwartungswert, die „Streuung“ um den Erwartungswert.
Beispiel einmaliger Würfelwurf:

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}[(X-\mathrm{E}[X])^2] = \mathrm{E}[(X-3.5)^2]\\ &= P(X=1)\cdot(1-3.5)^2 + P(X=2)\cdot(2-3.5)^2 + P(X=3)\cdot(3-3.5)^2\\ &\quad + P(X=4)\cdot(4-3.5)^2 + P(X=5)\cdot(5-3.5)^2 + P(X=6)\cdot(6-3.5)^2\\ &= \frac{1}{6}((-2.5)^2 + (-1.5)^2 + (-0.5)^2 + 0.5^2 + 1.5^2 + 2.5^2) = \frac{17.5}{6} = \frac{35}{12}\end{aligned}$$

# 8.1 Erwartungswert und Varianz

## Diskrete Wahrscheinlichkeitsverteilungen

Ist $h : \mathbb{R} \to \mathbb{R}$ eine Funktion, so gilt für den Erwartungswert der transformierten Zufallsvariable $h(X)$:

$$\mathrm{E}[h(X)] = \sum_{j=1}^{J} h(x_j) \cdot p(x_j), \; J \in \mathbb{N} \cup \{\infty\}$$

Für $h : x \mapsto x$ ergibt sich für $\mathrm{E}[h(X)]$ damit der Erwartungswert von $X$ und für $h : x \mapsto (x - \mathrm{E}[X])^2$ die Varianz von $X$.

Der Wert, der sich für $h : x \mapsto x^k$ ergibt, wird $k$-tes **Moment** von $X$ genannt:

$$m_k(X) = \mathrm{E}[X^k] = \sum_{j=1}^{J} x_j^k \cdot p(x_j), \; J \in \mathbb{N} \cup \{\infty\}.$$

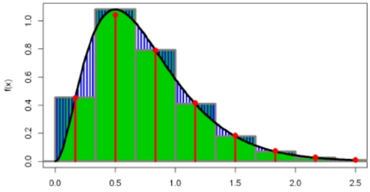
Das $k$-te Moment der um den Erwartungswert zentrierten Zufallsvariable $X - \mathrm{E}[X]$ heißt $k$-tes **zentrales Moment**:

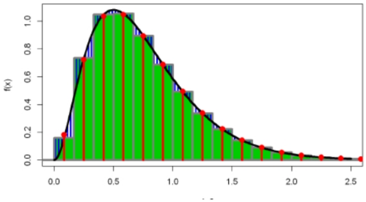
$$\mu_k(X) = \mathrm{E}[(X - \mathrm{E}[X])^k] = \sum_{j=1}^{J} (x_j - \mathrm{E}[X])^k \cdot p(x_j), \; J \in \mathbb{N} \cup \{\infty\}$$

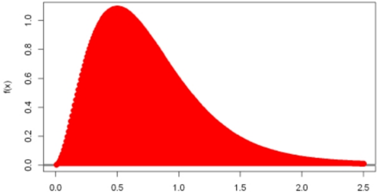
# 8.1 Erwartungswert und Varianz

## Stetige Wahrscheinlichkeitsverteilungen

Approximation stetiger Dichtefunktionen von $X$ durch Zähldichte diskretisierter Zufallsvariable




Integral ist die Fläche unter der Kurve und kann als Grenzwert von immer dünneren Rechtecken berechnet werden

# 8.1 Erwartungswert und Varianz

## Stetige Wahrscheinlichkeitsverteilungen

Der **Erwartungswert** einer stetig verteilten Zufallsvariable $X$ mit Dichtefunktion $f(\cdot)$ ist definiert durch

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} t \cdot f(t) dt$$

Die **Varianz** von $X$ ist definiert durch

$$\mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2] = \int_{-\infty}^{\infty} (t - \mathrm{E}[X])^2 \cdot f(t) dt$$

Die **Standardabweichung** von $X$ ist definiert durch

$$\sqrt{\mathrm{Var}(X)}$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Erwartungswerten

Falls die folgenden Erwartungswerte von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(1) $\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$

(2) $\mathrm{E}[aX+b] = a\,\mathrm{E}[X] + b,\ a, b \in \mathbb{R}$

(3) $\mathrm{E}\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i\,\mathrm{E}[X_i] + b,\ a_1, \ldots, a_n, b \in \mathbb{R}$

(4) $X$ und $Y$ st.u. $\Rightarrow \mathrm{E}[XY] = \mathrm{E}[X] \cdot \mathrm{E}[Y]$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Erwartungswerten

Falls die folgenden Erwartungswerte von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(1) $\boxed{\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]}$

Beweis:

$$\boxed{\mathrm{E}[X+Y]} = \sum_{x+y \in T_{X+Y}} [x+y] \cdot P(X+Y = x+y) = \sum_{\omega \in \Omega} [X(\omega) + Y(\omega)] \cdot P(\{\omega\})$$

$$= \sum_{\omega \in \Omega} x(\omega) \cdot P(\{\omega\}) + \sum_{\omega \in \Omega} Y(\omega) \cdot P(\{\omega\})$$

$$= \sum_{x \in T_X} x \cdot P(X = x) + \sum_{y \in T_Y} P(Y = y) = \boxed{\mathrm{E}[X] + \mathrm{E}[Y]} \quad \square$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Erwartungswerten

Falls die folgenden Erwartungswerte von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(2) $\boxed{\mathrm{E}[aX+b] = a\,\mathrm{E}[X] + b, \;\; a, b \in \mathbb{R}}$

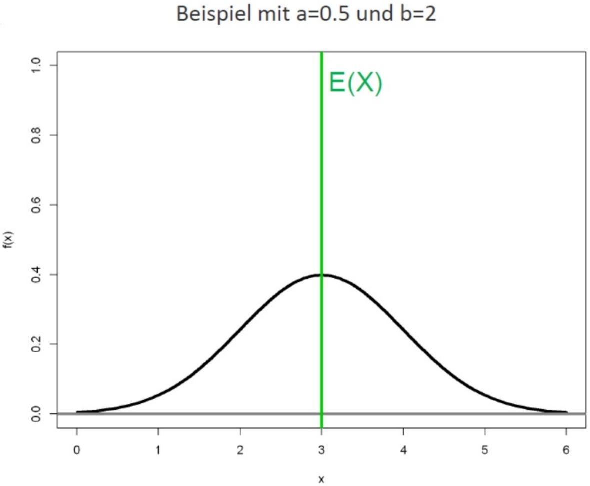
Beweis:

$$\begin{aligned}
\boxed{\mathrm{E}[aX+b]} &= \int_{-\infty}^{\infty} (at+b)\cdot f(t)dt = \int_{-\infty}^{\infty} at\cdot f(t)dt + \int_{-\infty}^{\infty} b\cdot f(t)dt \\
&= a\cdot\int_{-\infty}^{\infty} t\cdot f(t)dt + b\cdot\int_{-\infty}^{\infty} f(t)dt = a\cdot \mathrm{E}[X] + b\cdot 1 \\
&= \boxed{a\,\mathrm{E}[X] + b}
\end{aligned}$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Erwartungswerten

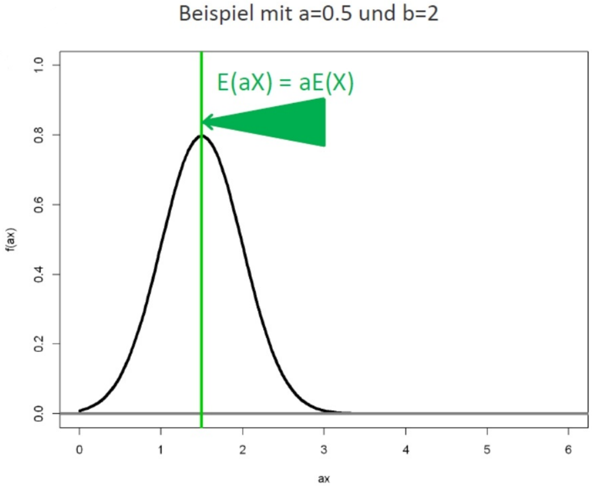
$\mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b$,
$a, b \in \mathbb{R}$



# 8.1 Erwartungswert und Varianz

## Eigenschaften von Erwartungswerten

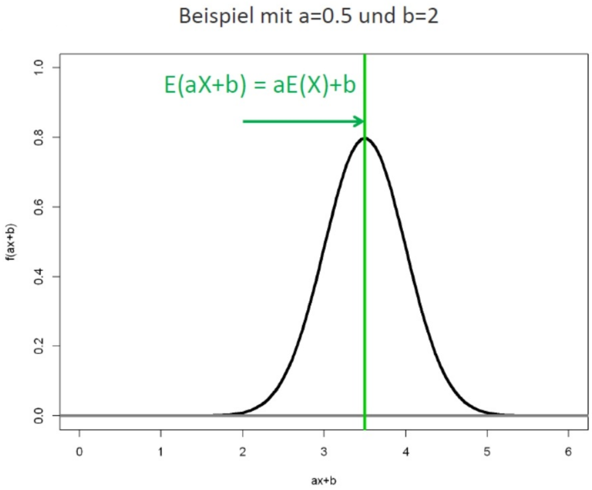
$$\mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b,$$
$$a, b \in \mathbb{R}$$



# 8.1 Erwartungswert und Varianz

## Eigenschaften von Erwartungswerten

$\mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b,$
$a, b \in \mathbb{R}$



# 8.1 Erwartungswert und Varianz

## Eigenschaften von Erwartungswerten

Falls die folgenden Erwartungswerte von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

$$(3)\quad \boxed{\mathrm{E}\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n} a_i \,\mathrm{E}[X_i] + b,\ \ a_1, \ldots, a_n, b \in \mathbb{R}}$$

Beweis:

$$\boxed{\mathrm{E}\left[\sum_{i=1}^{n} a_i X_i + b\right]} \underset{(2)}{=} \mathrm{E}\left[\sum_{i=1}^{n} a_i X_i\right] + b \underset{(1)}{=} \sum_{i=1}^{n} \mathrm{E}[a_i X_i] + b$$

$$\underset{(2)}{=} \boxed{\sum_{i=1}^{n} a_i \,\mathrm{E}[X_i] + b} \qquad \square$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Erwartungswerten

Falls die folgenden Erwartungswerte von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(4) $\boxed{X \text{ und } Y \text{ st.u.} \Rightarrow \mathrm{E}[XY] = \mathrm{E}[X] \cdot \mathrm{E}[Y]}$

Beweis:

$$\boxed{X \text{ und } Y \text{ st.u.}} \Rightarrow f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

$$\Rightarrow \boxed{\mathrm{E}[XY]} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u \cdot v \cdot f_{XY}(u,v)du\,dv = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u \cdot v \cdot f_X(u) \cdot f_Y(v)du\,dv$$

$$= \int_{-\infty}^{\infty} u \cdot f_X(u)du \int_{-\infty}^{\infty} v \cdot f_Y(v)dv = \boxed{\mathrm{E}[X] \cdot \mathrm{E}[Y]}$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

Falls die folgenden Varianzen von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(A) $\text{Var}(X) \geq 0$

(B) $\text{Var}(X) = 0 \Leftrightarrow X \sim \varepsilon_a,\ a \in \mathbb{R}$

(C) $\text{Var}(aX + b) = a^2\,\text{Var}(X)$

(D) $X$ und $Y$ st.u. $\Rightarrow \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

(E) $X_1, \ldots, X_n$ st.u., $a_1, \ldots, a_n \in \mathbb{R}$
$$\Rightarrow \text{Var}\left(\sum_{i=1}^{n} a_i X_i + b\right) = \sum_{i=1}^{n} a_i^2\,\text{Var}(X)$$

(A) $\boxed{\text{Var}(X) \geq 0}$ ist klar, da Quadrat und Dichte immer größer 0

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

Falls die folgenden Varianzen von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(B) $\boxed{\mathrm{Var}(X) = 0 \Leftrightarrow X \sim \varepsilon_a,\ a \in \mathbb{R}}$

Beweis:

$$„\Leftarrow\text{“}:\ X \sim \varepsilon_a \Rightarrow \mathrm{E}[X] = a \Rightarrow X - \mathrm{E}[X] = 0 \Rightarrow \boxed{\mathrm{Var}(X) = 0}$$

$$„\Rightarrow\text{“}:\ \mathrm{Var}(X) = 0 \Leftrightarrow \int\limits_{-\infty}^{\infty} (t - \mathrm{E}[X])^2 \cdot f(t)dt = 0$$

$$\Rightarrow \text{ für alle } t \in \mathbb{R}:\ (t - \mathrm{E}[X])^2 = 0 \text{ oder } f(t) = 0$$

$$(t - \mathrm{E}[X])^2 = 0 \Leftrightarrow t = \mathrm{E}[X] \Rightarrow \text{ für alle } t \neq \mathrm{E}[X] \text{ muss gelten}:\ f(t) = 0$$

$$\underset{\int\limits_{-\infty}^{\infty} f(t)dt = 1}{\Rightarrow} \boxed{X \sim \varepsilon_{\mathrm{E}[X]}} \qquad \square$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

Falls die folgenden Varianzen von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(C) $\boxed{\operatorname{Var}(aX+b) = a^2 \operatorname{Var}(X)}$

Beweis:

$$\boxed{\operatorname{Var}(aX+b)} = \mathrm{E}[(aX+b-\mathrm{E}[aX+b])^2]$$
$$\underset{\mathrm{E}[aX+b]=a\,\mathrm{E}[X]+b}{=} \mathrm{E}[(aX+b-a\,\mathrm{E}[X]-b)^2] = \mathrm{E}[(a\cdot(X-\mathrm{E}[X]))^2]$$
$$= \mathrm{E}[a^2\cdot(X-\mathrm{E}[X])^2] = a^2\cdot\mathrm{E}[(X-\mathrm{E}[X])^2] = \boxed{a^2\operatorname{Var}(X)} \qquad \square$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

Falls die folgenden Varianzen von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(D) $\boxed{X \text{ und } Y \text{ st.u.} \Rightarrow \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)}$

Beweis: $X$ und $Y$ st.u. $\Rightarrow \text{E}[XY] = \text{E}[X]\,\text{E}[Y]$ $(*)$

$$\begin{aligned}
\text{Var}(X+Y) &= \text{E}[(X+Y-\text{E}[X+Y])^2] = \text{E}[(X+Y-\text{E}[X]-\text{E}[Y])^2] \\
&= \text{E}[((X-\text{E}[X]) + (Y-\text{E}[Y]))^2] \\
&= \text{E}\left[(X-\text{E}[X])^2 + (Y-\text{E}[Y])^2 + 2(X-\text{E}[X])\cdot(Y-\text{E}[Y])\right] \\
&= \text{Var}(X) + \text{Var}(Y) + 2R
\end{aligned}$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

Falls die folgenden Varianzen von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(D) $\boxed{X \text{ und } Y \text{ st.u.} \Rightarrow \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)}$

Beweis: (Fortsetzung)

$$
\begin{aligned}
R &= \mathrm{E}\left[(X - \mathrm{E}[X]) \cdot (Y - \mathrm{E}[Y])\right] = \mathrm{E}\left[XY - \mathrm{E}[X]Y - X\,\mathrm{E}[Y] + \mathrm{E}[X]\,\mathrm{E}[Y]\right] \\
&= \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y] - \mathrm{E}[X]\,\mathrm{E}[Y] + \mathrm{E}[X]\,\mathrm{E}[Y] \underset{(*)}{=} 0
\end{aligned}
$$

$$
\Rightarrow \boxed{\mathrm{Var}(X+Y)} = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2R \underset{R=0}{=} \boxed{\mathrm{Var}(X) + \mathrm{Var}(Y)} \quad \square
$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

Falls die folgenden Varianzen von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(F) **Verschiebungssatz von Steiner**:

$$a \in \mathbb{R} \Rightarrow \mathrm{Var}(X) = \mathrm{E}[(X-a)^2] - (\mathrm{E}[X]-a)^2,$$

speziell für $a = 0 \Rightarrow \mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$

(G) **Tschebyscheff-Ungleichung**:

$$P(|X - \mathrm{E}[X]| > \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}, \ \varepsilon \in (0, \infty)$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

Falls die folgenden Varianzen von stetig oder diskret verteilten Zufallsvariablen existieren, so gelten folgende Eigenschaften:

(F) **Verschiebungssatz von Steiner**:

$$a \in \mathbb{R} \Rightarrow \mathrm{Var}(X) = \mathrm{E}[(X-a)^2] - (\mathrm{E}[X]-a)^2,$$

speziell für $a = 0 \Rightarrow \mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$

Beweis:

$$\begin{aligned}
\boxed{\mathrm{Var}(X)} &= \mathrm{E}[(X-\mathrm{E}[X])^2] = \mathrm{E}[((X-a)+(a-\mathrm{E}[X]))^2] \\
&= \mathrm{E}[(X-a)^2 + 2(a-\mathrm{E}[X])(X-a) + (a-\mathrm{E}[X])^2] \\
&= \mathrm{E}[(X-a)^2] + 2(a-\mathrm{E}[X])(\mathrm{E}[X]-a) + (a-\mathrm{E}[X])^2 \\
&= \mathrm{E}[(X-a)^2] - 2(a-\mathrm{E}[X])^2 + (a-\mathrm{E}[X])^2 \\
&= \boxed{\mathrm{E}[(X-a)^2] - (a-\mathrm{E}[X])^2}
\end{aligned}$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

(G) **Tschebyscheff-Ungleichung**:

$$P(|X - \mathrm{E}[X]| > \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2},\ \varepsilon \in (0, \infty)$$

Beweis:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (t - \mathrm{E}[X])^2 f_X(t)dt \underset{(t-\mathrm{E}[X])^2 f_X(t) \geq 0}{\geq} \int_{t:(t-\mathrm{E}[X])^2 > \varepsilon^2} (t - \mathrm{E}[X])^2 f_X(t)dt$$

$$\geq \int_{t:(t-\mathrm{E}[X])^2 > \varepsilon^2} \varepsilon^2 f_X(t)dt = \varepsilon^2 \int_{t:(t-\mathrm{E}[X])^2 > \varepsilon^2} f_X(t)dt$$

$$\underset{(*)}{=} \varepsilon^2 \cdot P[(X > \mathrm{E}[X] + \varepsilon) \cup (X < \mathrm{E}[X] - \varepsilon)]$$

$$= \varepsilon^2 \cdot P[|X - \mathrm{E}[X]| > \varepsilon] \quad \square$$

$$(*)(t - \mathrm{E}[X])^2 > \varepsilon^2 \Leftrightarrow t > \mathrm{E}[X] + \varepsilon \cup t < \mathrm{E}[X] - \varepsilon$$

# 8.1 Erwartungswert und Varianz

## Eigenschaften von Varianzen

Anwendung der Tschebyscheff-Ungleichung: Abschätzung verteilungsunabhängiger Unsicherheitsbereiche

$$P(|X - \mathrm{E}[X]| > \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}, \; \varepsilon \in (0, \infty)$$

$$\text{Setze } \varepsilon := r\sqrt{\mathrm{Var}(X)} \Rightarrow P(|X - \mathrm{E}[X]| > r\sqrt{\mathrm{Var}(X)}) \leq \frac{1}{r^2}$$

$$\Leftrightarrow P(\mathrm{E}[X] - r\sqrt{\mathrm{Var}(X)} \leq X \leq \mathrm{E}[X] + r\sqrt{\mathrm{Var}(X)}) \geq 1 - \frac{1}{r^2}$$

Die Wahrscheinlichkeit, dass eine Realisation von $X$ in einem symmetrischen Intervall der Breite von $r$ Standardabweichungen fällt, beträgt also unabhängig von der Verteilung von $X$ mindestens $1 - 1/r^2$.

# 8.1 Erwartungswert und Varianz

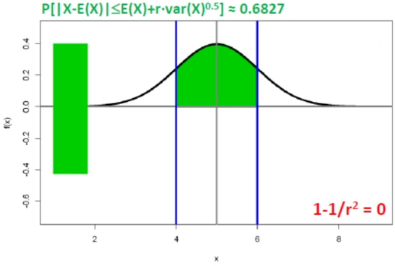
**Eigenschaften von Varianzen** Anwendung der Tschebyscheff-Ungleichung:

Abschätzung verteilungsunabhängiger Unsicherheitsbereiche

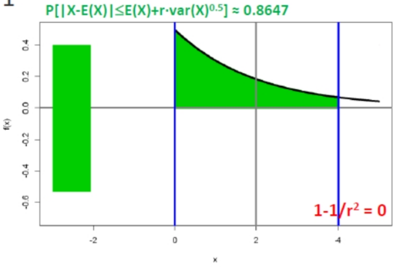
$$P\left(\mathrm{E}[X] - r\sqrt{\mathrm{Var}(X)} \leq X \leq \mathrm{E}[X] + r\sqrt{\mathrm{Var}(X)}\right) \geq 1 - \frac{1}{r^2}$$

Die Wahrscheinlichkeit, dass eine Realisation von $X$ in einem symmetrischen Intervall der Breite von $r$ Standardabweichungen fällt, beträgt also unabhängig von der Verteilung von $X$ mindestens $1 - 1/r^2$.

r = 1




N(5,1)




Exp(0.5)

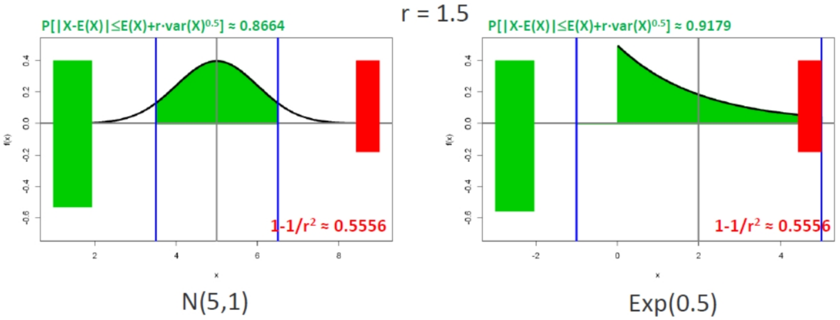
# 8.1 Erwartungswert und Varianz

**Eigenschaften von Varianzen** Anwendung der Tschebyscheff-Ungleichung:

Abschätzung verteilungsunabhängiger Unsicherheitsbereiche

$$P\left(\mathrm{E}[X] - r\sqrt{\mathrm{Var}(X)} \leq X \leq \mathrm{E}[X] + r\sqrt{\mathrm{Var}(X)}\right) \geq 1 - \frac{1}{r^2}$$

Die Wahrscheinlichkeit, dass eine Realisation von $X$ in einem symmetrischen Intervall der Breite von $r$ Standardabweichungen fällt, beträgt also unabhängig von der Verteilung von $X$ mindestens $1 - 1/r^2$.



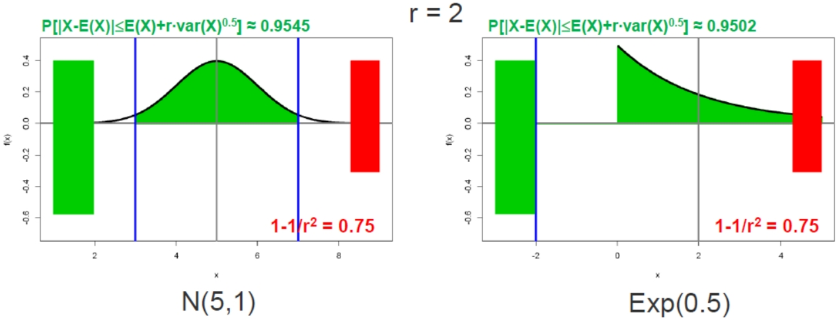
# 8.1 Erwartungswert und Varianz

**Eigenschaften von Varianzen** Anwendung der Tschebyscheff-Ungleichung:

Abschätzung verteilungsunabhängiger Unsicherheitsbereiche

$$P\left(\mathrm{E}[X] - r\sqrt{\mathrm{Var}(X)} \leq X \leq \mathrm{E}[X] + r\sqrt{\mathrm{Var}(X)}\right) \geq 1 - \frac{1}{r^2}$$

Die Wahrscheinlichkeit, dass eine Realisation von $X$ in einem symmetrischen Intervall der Breite von $r$ Standardabweichungen fällt, beträgt also unabhängig von der Verteilung von $X$ mindestens $1 - 1/r^2$.



# 8.1 Erwartungswert und Varianz

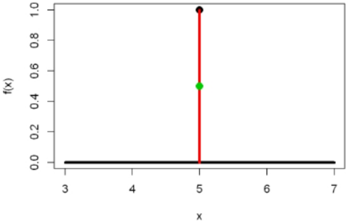
## Diskrete Wahrscheinlichkeitsverteilungen

$$\mathrm{E}[X] = \sum_{j=1}^{J} p(x_j) \cdot x_j \qquad \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

Beispiel **Einpunktverteilung** $\varepsilon_a$

**Zähldichte**: $p(x) = \mathtt{I}(a = x)$

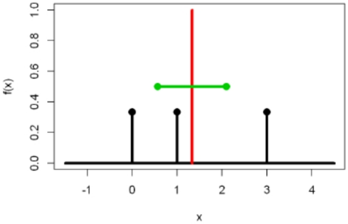
$\boxed{\mathrm{E}[X] = a}$ $\boxed{\sqrt{\mathrm{Var}(X)} = 0}$




Bsp. **Diskrete Gleichverteilung** $G(x_1, \ldots, x_n)$

**Zähldichte**: $p(x) = \frac{1}{n} \cdot \mathtt{I}(x \in (\{x_1, \ldots, x_n\})$

$\boxed{\mathrm{E}[X] = \bar{x}}$ $\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{s_X^2}}$



# 8.1 Erwartungswert und Varianz

## Diskrete Wahrscheinlichkeitsverteilungen

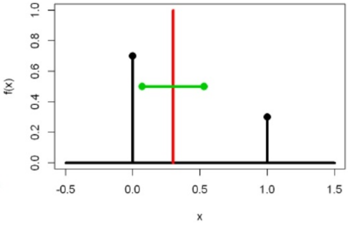
$$\mathrm{E}[X] = \sum_{j=1}^{J} p(x_j) \cdot x_j \qquad \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

Bsp. **Bernoulli-Verteilung** $\mathrm{Ber}(p)$

**Zähldichte**:

$$p(x) = p \cdot \mathtt{I}(x = 1) + (1 - p) \cdot \mathtt{I}(x = 0)$$

$$\boxed{\mathrm{E}[X] = p} \quad \boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{p(1-p)}}$$

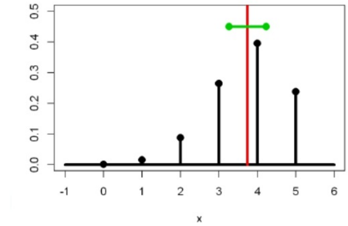



Bsp. **Binomialverteilung** $\mathrm{Bin}(n, p)$

**Zähldichte**:

$$p(x) = \mathtt{I}(x \in \{0, \dots, n\}) \cdot \binom{n}{x} \cdot p^x \cdot (1 - p)^{n-x}$$

$$\boxed{\mathrm{E}[X] = np} \quad \boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{np(1-p)}}$$



# 8.1 Erwartungswert und Varianz

## Diskrete Wahrscheinlichkeitsverteilungen

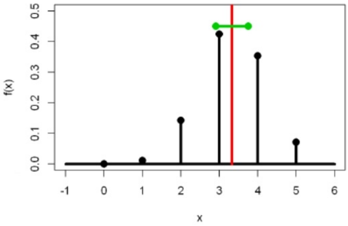
$$\mathrm{E}[X] = \sum_{j=1}^{J} p(x_j) \cdot x_j \qquad \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

Bsp. **Hypergeometrische Verteilung** $\mathrm{Hyp}(n, r, s)$

**Zähldichte**:

$$p(x) = \mathtt{I}(x \in \{\max(0, n-1), \ldots, \min(n, r)\}) \cdot \frac{\binom{r}{x} \cdot \binom{s}{n-x}}{\binom{r+s}{x}}$$

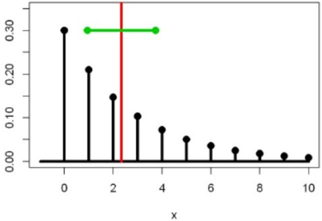
$$\boxed{\mathrm{E}[X] = n\frac{r}{r+s}} \qquad \boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{n\frac{r}{r+s}\,\frac{s}{r+s}\,\frac{r+s-n}{r+s-1}}}$$




Bsp. **Geometrische Verteilung** $\mathrm{Geo}(p)$

**Zähldichte**: $p(x) = \mathtt{I}(x \in \mathbb{N} \cup 0) \cdot p \cdot (1-p)^x$

$$\boxed{\mathrm{E}[X] = \frac{1-p}{p}} \qquad \boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\frac{1-p}{p^2}}}$$



# 8.1 Erwartungswert und Varianz

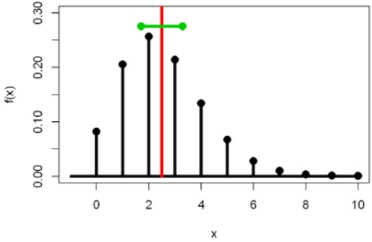
## Diskrete Wahrscheinlichkeitsverteilungen

$$E[X] = \sum_{j=1}^{J} p(x_j) \cdot x_j \qquad \mathrm{Var}(X) = E[(X - E[X])^2]$$

Bsp. **Poisson-Verteilung** Poi($\lambda$)

**Zähldichte**: $p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$

$\boxed{E[X] = \lambda}$ $\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\lambda}}$



# 8.1 Erwartungswert und Varianz

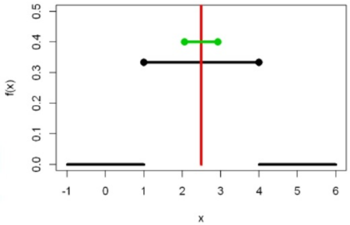
## Stetige Wahrscheinlichkeitsverteilungen

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} t \cdot f(t)dt \qquad \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

Bsp. **Rechteckverteilung** $R(a, b)$

**Dichtefunktion**: $f(x) = \frac{\mathtt{I}(a \leq x \leq b)}{b-a}$

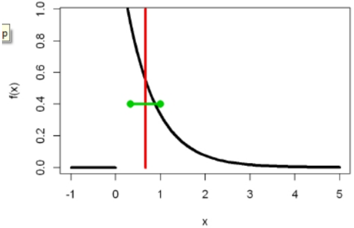
$$\boxed{\mathrm{E}[X] = \frac{a+b}{2}} \quad \boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\frac{(b-a)^2}{12}}}$$




Bsp. **Exponentialverteilung** $\mathrm{Exp}(\lambda)$

**Dichtefunktion**: $f(x) = \mathtt{I}(0 \leq x) \cdot \lambda e^{-(\lambda \cdot x)}$

$$\boxed{\mathrm{E}[X] = \frac{1}{\lambda}} \quad \boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\frac{1}{\lambda^2}}}$$



# 8.1 Erwartungswert und Varianz

## Stetige Wahrscheinlichkeitsverteilungen

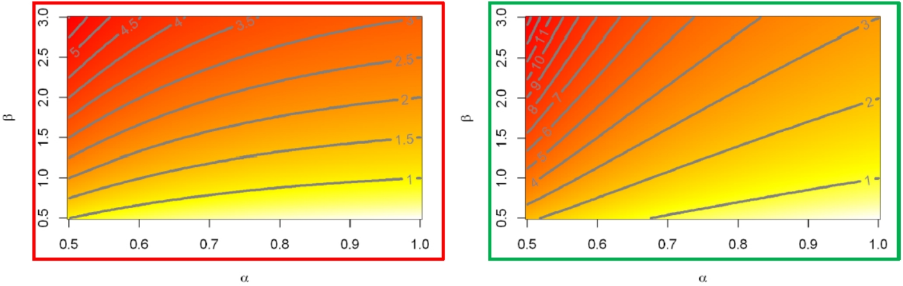
Bsp. **Weibull-Verteilung** $W(\alpha,\beta)$

**Dichtefunktion**:

$$f(x) = \mathtt{I}(0 \leq x) \cdot \left(\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-(x/\beta)^{\alpha}}\right)$$

$$\boxed{\mathrm{E}[X] = \frac{\beta}{\alpha}\Gamma\left(\frac{1}{\alpha}\right)}$$

$$\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\beta^2\left(\frac{2}{\alpha}\Gamma\left[\frac{2}{\alpha}\right] - \left[\frac{1}{\alpha}\Gamma\left(\frac{1}{\alpha}\right)\right]^2\right)}}$$



# 8.1 Erwartungswert und Varianz

## Stetige Wahrscheinlichkeitsverteilungen

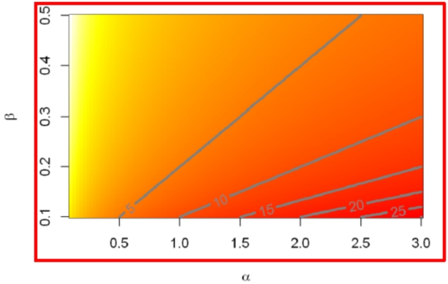
Bsp. **Gamma-Verteilung** $\Gamma(\alpha, \beta)$

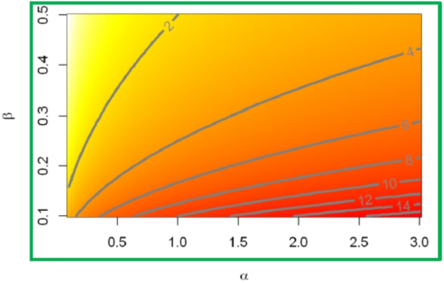
**Dichtefunktion**:

$f(x) = \mathbb{I}(0 \leq x) \cdot \frac{\alpha^{\beta}}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x}$

$\boxed{\mathrm{E}[X] = \frac{\alpha}{\beta}}$

$\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\frac{\alpha}{\beta^2}}}$



# 8.1 Erwartungswert und Varianz

## Stetige Wahrscheinlichkeitsverteilungen

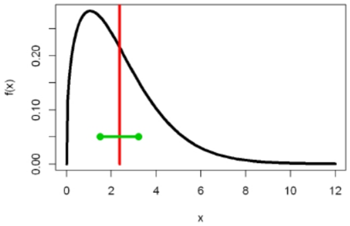
$$\mathrm{E}[X] = \int_{-\infty}^{\infty} t \cdot f(t)dt \qquad \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

Bsp. **Weibull-Verteilung** $W(\alpha, \beta)$

**Dichtefunktion**:

$$f(x) = \mathrm{I}(0 \leq x) \cdot \left( \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha - 1} e^{-(x/\beta)^{\alpha}} \right)$$

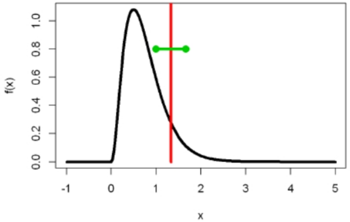
$\boxed{\mathrm{E}[X] = \frac{\beta}{\alpha}\Gamma\left(\frac{1}{\alpha}\right)}$ $\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\beta^2\left(\frac{2}{\alpha}\Gamma\left[\frac{2}{\alpha}\right] - \left[\frac{1}{\alpha}\Gamma\left(\frac{1}{\alpha}\right)\right]^2\right)}}$




Bsp. **Gamma-Verteilung** $\Gamma(\alpha, \beta)$

**Dichtefunktion**: $f(x) = \mathrm{I}(0 \leq x) \cdot \frac{\alpha^{\beta}}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x}$

$\boxed{\mathrm{E}[X] = \frac{\alpha}{\beta}}$ $\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\frac{\alpha}{\beta^2}}}$



# 8.1 Erwartungswert und Varianz

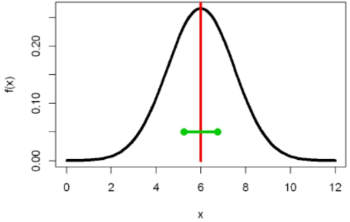
## Stetige Wahrscheinlichkeitsverteilungen

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} t \cdot f(t)dt \qquad \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

Bsp. **Normalverteilung** $N(\mu, \sigma^2)$

**Dichtefunktion**: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$\boxed{\mathrm{E}[X] = \mu}$ $\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\sigma^2}}$

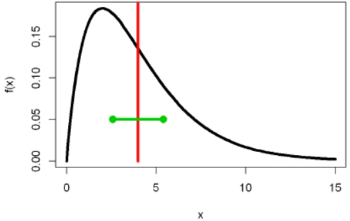



Bsp. $\chi^2$-**Verteilung** $\chi^2_f$

**Dichtefunktion**:

$f(x) = \mathrm{I}(0 < x) \cdot \frac{1}{2^{f/2}\Gamma(f/2)} x^{f/2-1} e^{-x/2}$

$\boxed{\mathrm{E}[X] = f}$ $\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{2f}}$



# 8.1 Erwartungswert und Varianz

## Stetige Wahrscheinlichkeitsverteilungen

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} t \cdot f(t)dt \quad \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2[$$

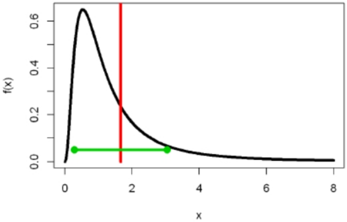
Bsp. $F$-**Verteilung** $F_{f_1,f_2}$

**Dichtefunktion**:

$$f(x) = \mathrm{I}(0 < x) \cdot \frac{\Gamma\left(\frac{f_1+f_2}{2}\right)}{\Gamma\left(\frac{f_1}{2}\right)\Gamma\left(\frac{f_2}{2}\right)} \left(\frac{f_1}{f_2}\right)^{f_1/2} \frac{x^{f_1/2-1}}{(1+f_1x/f_2)^{(f_1+f_2)/2}}$$

$$\boxed{\mathrm{E}[X] = \frac{f_2}{f_2-2},\ f_2 > 2}$$

$$\boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\frac{2f_2^2(f_1+f_2-2)}{f_1(f_2-2)^2(f_2-4)}},\ f_2 > 4}$$



# 8.1 Erwartungswert und Varianz

## Stetige Wahrscheinlichkeitsverteilungen

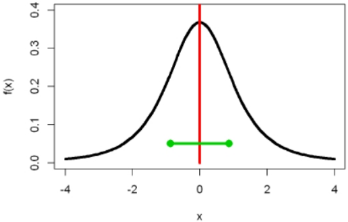
$$\mathrm{E}[X] = \int_{-\infty}^{\infty} t \cdot f(t)dt \quad \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

Bsp. $t$-**Verteilung** $t_f$

**Dichtefunktion**:

$$f(x) = \frac{\Gamma\left(\frac{f+1}{2}\right)}{\sqrt{f\pi}\cdot\Gamma\left(\frac{f}{2}\right)} \cdot \left(1 + \frac{x^2}{f}\right)^{-(f+1)/2-x/2}$$

$$\boxed{\mathrm{E}[X] = 0} \quad \boxed{\sqrt{\mathrm{Var}(X)} = \sqrt{\frac{f}{f-2}},\ f > 2}$$



# 8.1 Erwartungswert und Varianz

## Schwaches Gesetz der großen Zahlen

Es gelte: $X = (X_1, \ldots, X_N)$, $X_i \sim F^X$ u.i.v.,
$\mathrm{E}[X_i] = \mu$, $\mathrm{Var}(X_i) = \sigma^2$, $i = 1, \ldots, N$, $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$

Dann gilt:

$$\lim_{N\to\infty} [P(|\bar{X} - \mu| > \varepsilon)] = 0$$

Beweis:

1. $$\mathrm{E}[\bar{X}] = \mathrm{E}\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] = \frac{1}{N}\mathrm{E}\left[\sum_{i=1}^{N} X_i\right] = \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}[X_i] = \frac{1}{N}\sum_{i=1}^{N}\mu = \frac{1}{N}N\mu = \mu$$

2. $$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \frac{1}{N^2}\mathrm{Var}\left(\sum_{i=1}^{N} X_i\right) \underset{X_1,\ldots,X_N \text{ st.u.}}{=} \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{Var}(X_i)$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sigma^2 = \frac{1}{N^2}N\sigma^2 = \frac{\sigma^2}{N}$$

# 8.1 Erwartungswert und Varianz

## Schwaches Gesetz der großen Zahlen

Beweis: (Fortsetzung)

Einsetzen in Tschebyscheff-Ungleichung:

$$P(|\bar{X} - \mathrm{E}[\bar{X}]| > \varepsilon) \leq \frac{\mathrm{Var}(\bar{X})}{\varepsilon^2} \Rightarrow P(|\bar{X} - \mu| > \varepsilon) \leq \frac{\sigma^2}{N\varepsilon^2} \underset{N\to\infty}{\to} 0$$

$$\Rightarrow \boxed{\lim_{N\to\infty} P(|\bar{X} - \mu| > \varepsilon) = 0} \quad \square$$

# 8.1 Erwartungswert und Varianz

## Nicht existente Erwartungswerte oder Varianzen

Erwartungswert, Varianz und höhere Momente müssen nicht existieren.

**Beispiel**:

Sei $X$ stetig verteilte Zufallsvariable mit Dichtefunktion $f_X(x) = \mathbb{I}(x \geq 1) \cdot \frac{1}{x^2}$

Dann ist $f_X$ tatsächlich Dichte, denn es gilt:

$F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = \int\limits_{1}^{x} f_X(t)dt = -\frac{1}{x} + 1 = 1 - \frac{1}{x}$ und damit

$\int\limits_{-\infty}^{\infty} f_X(t)dt = \lim\limits_{x\to\infty}(1 - 1/x) = 1$

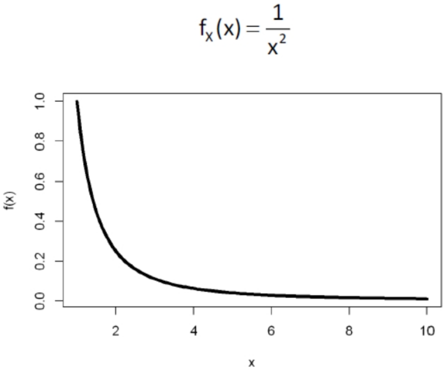
Allerdings gilt auch: $\mathrm{E}[X] = \int\limits_{-\infty}^{\infty} tf_X(t)dt = \int\limits_{1}^{\infty} t\frac{1}{t^2}dt = \lim\limits_{b\to\infty}(log(b)) = \infty$

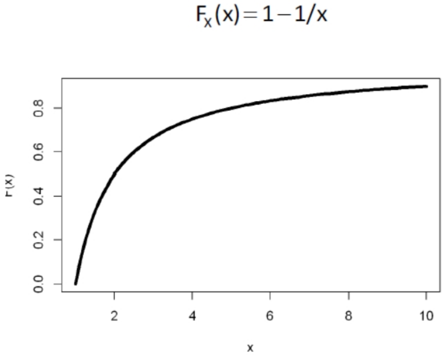
# 8.1 Erwartungswert und Varianz

## Nicht existente Erwartungswerte oder Varianzen

Erwartungswert, Varianz und höhere Momente müssen nicht existieren.

**Beispiel**:



# 8.2 Weitere wahrscheinlichkeitstheoretische Kennzahlen

## $p$-Quantile

Sei $X$ eine Zufallsvariable mit Verteilungsfunktion $F$.

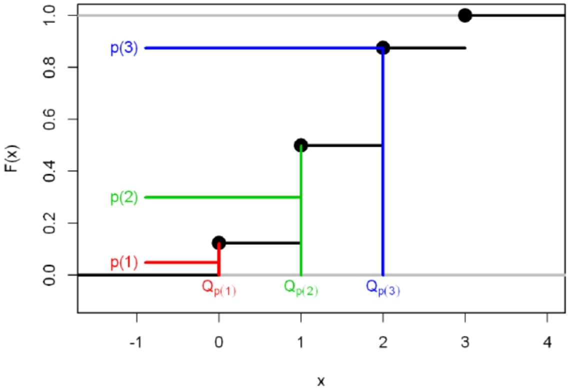
Für eine Zahl $p \in (0, 1)$ wird das $p$-**Quantil** $Q_p = Q_p(X)$ der durch die Verteilungsfunktion $F = F^X$ festgelegten Verteilung $P^X$ definiert durch den kleinsten Wert $x \in \mathbb{R}$, für den gilt:

$$F(x) \geq p$$

Das 0.5-Quantil heißt **Median**, das 0.25-Quantil **unteres Quartil** und das 0.75-Quantil **oberes Quartil**.

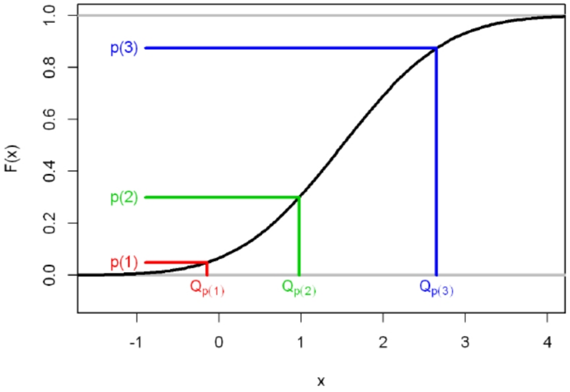
# 8.2 Weitere wahrscheinlichkeitstheoretische Kennzahlen

**Quantilsbestimmng über Verteilungsfunktion** (diskreter Fall)



# 8.2 Weitere wahrscheinlichkeitstheoretische Kennzahlen

**Quantilsbestimmng über Verteilungsfunktion** (stetiger Fall)



# 8.2 Weitere wahrscheinlichkeitstheoretische Kennzahlen

## Abhängigkeitsmaße

**Erinnerung:** $X$ und $Y$ stochastisch unabhängig $\Rightarrow \mathrm{E}[XY] = \mathrm{E}[X] \cdot \mathrm{E}[Y]$

Umgekehrt gilt somit: $\mathrm{E}[XY] \neq \mathrm{E}[X] \cdot \mathrm{E}[Y] \Rightarrow X$ und $Y$ stochastisch abhängig

Seien $X$ und $Y$ Zufallsvariablen. Dann heißt

$$\mathrm{Cov}(X, Y) = \mathrm{E}[(X - \mathrm{E}[X]) \cdot (Y - \mathrm{E}[Y])]$$

**Kovarianz** von $X$ und $Y$. Die Größe

$$\mathrm{Cor}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}$$

heißt **Korrelation** von $X$ und $Y$.

# 8.2 Weitere wahrscheinlichkeitstheoretische Kennzahlen

## Abhängigkeitsmaße

$$\mathrm{Cov}(X,Y) = \mathrm{E}[(X - \mathrm{E}[X]) \cdot (Y - \mathrm{E}[Y])] \qquad \mathrm{Cor}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}$$

## Eigenschaften

(i) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X), \quad \mathrm{Cor}(X,Y) = \mathrm{Cor}(Y,X)$

(ii) $\mathrm{Cov}(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]$

(iii) $-1 \le \mathrm{Cor}(X,Y) \le 1$

(iv) $\mathrm{Cor}(X,Y) < 0 \Leftrightarrow X$ und $Y$ sind negativ korreliert
$\mathrm{Cor}(X,Y) > 0 \Leftrightarrow X$ und $Y$ sind positiv korreliert
$\mathrm{Cor}(X,Y) = 0 \Leftrightarrow X$ und $Y$ sind unkorreliert

# 8.2 Weitere wahrscheinlichkeitstheoretische Kennzahlen

## Abhängigkeitsmaße

$$\mathrm{Cov}(X, Y) = \mathrm{E}[(X - \mathrm{E}[X]) \cdot (Y - \mathrm{E}(Y))] \qquad \mathrm{Cor}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}$$

## Eigenschaften

(v) $X$ und $Y$ stoch. unabh. $\Rightarrow$ $X$ und $Y$ unkorreliert
$X$ und $Y$ unkorreliert $\not\Rightarrow$ $X$ und $Y$ stoch. unabh.

(vi) $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2 \cdot \mathrm{Cov}(X, Y)$

(vii) $\mathrm{Cov}(aX + b, cY + d) = ac \cdot \mathrm{Cov}(X, Y)$

(viii) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$

Beweise der Eigenschaften analog zu den Beweisen der gleichen Eigenschaften für die empirische Kovarianz und empirische Korrelation.

# Bedingte Wahrscheinlichkeiten und stochastische Unabhängigkeit

# 9.1 Bedingte Wahrscheinlichkeiten

Bisher: Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$

Grundraum $\Omega$, Menge aller Ereignisse $\mathcal{A}$ auf $\Omega$, Wahrscheinlichkeitsmaß $P$ auf $\Omega$

$P : \mathcal{A} \to [0,1],\ A \mapsto P(A)$

1. $0 \leq P(A) \leq 1$ für jedes Ereignis $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{A}$

# 9.1 Bedingte Wahrscheinlichkeiten

Jetzt: Wahrscheinlichkeitsraum $(B, \mathcal{B}, P_B)$

Einschränkung des Grundraums auf Ereignis $B \subset \Omega$, Wahrscheinlichkeitsmaß $P_B$ auf $B$

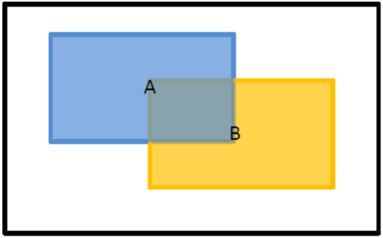
$P : \mathcal{B} \to [0,1], \; A \mapsto P_B(A)$

1. $0 \leq P(A) \leq 1$ für jedes Ereignis $A \in \mathcal{B}$
2. $P(A) = 1$
3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für alle paarweise disjunkten Ereignisse $A_i \in \mathcal{B}$

# 9.1 Bedingte Wahrscheinlichkeiten

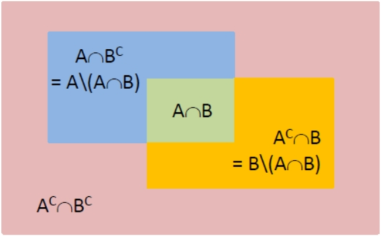
Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$

$\Omega$



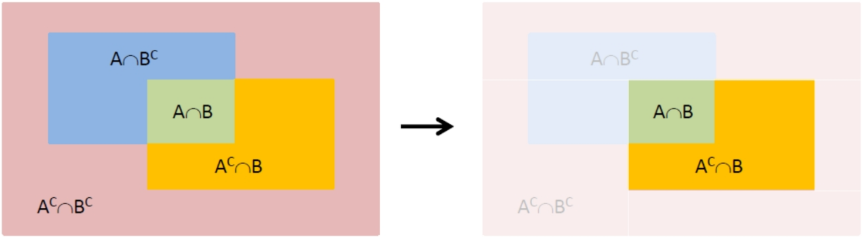
# 9.1 Bedingte Wahrscheinlichkeiten

Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$



# 9.1 Bedingte Wahrscheinlichkeiten

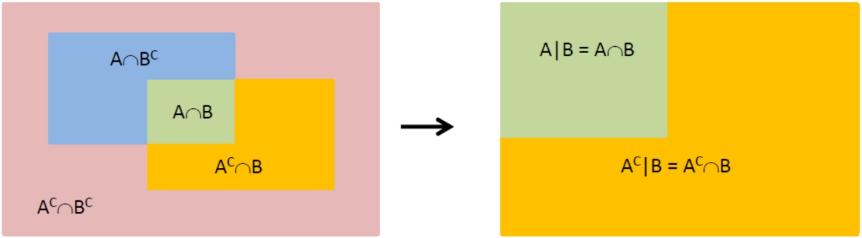
Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ Wahrscheinlichkeitsraum $(B, \mathcal{B}, P_B)$




$A \cap B^C \subset B^C$
$A^C \cap B^C \subset B^C$

# 9.1 Bedingte Wahrscheinlichkeiten

Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ Wahrscheinlichkeitsraum $(B, \mathcal{B}, P_B)$




$A \cap B^C \subset B^C$
$A^C \cap B^C \subset B^C$

# 9.1 Bedingte Wahrscheinlichkeiten

Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ Wahrscheinlichkeitsraum $(B, \mathcal{B}, P_B)$


$$
\begin{aligned}
1 &= P(\Omega) = P_\Omega(\Omega) = P(\Omega \mid \Omega) \\
&= P\left[\boxed{A \cap B^C}\right] + P\left[\boxed{A^C \cap B^C}\right] \\
&+ P\left[\boxed{A \cap B}\right] + P\left[\boxed{A^C \cap B}\right] \\
&= P(\boxed{C_1}) + P(\boxed{C_2}) + P(\boxed{C_3}) + P(\boxed{C_4})
\end{aligned}
$$

$$
\begin{aligned}
1 &= P_B(B) = P(B \mid B) \\
&= P\left[\boxed{A \mid B}\right] + P\left[\boxed{A^C \mid B}\right] \\
&= P_B\left[\boxed{A \cap B}\right] + P_B\left[\boxed{A^C \cap B}\right] \\
&= P_B(\boxed{C_3}) + P_B(\boxed{C_4})
\end{aligned}
$$

# 9.1 Bedingte Wahrscheinlichkeiten

Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ Wahrscheinlichkeitsraum $(B, \mathcal{B}, P_B)$


Die Wahrscheinlichkeit

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

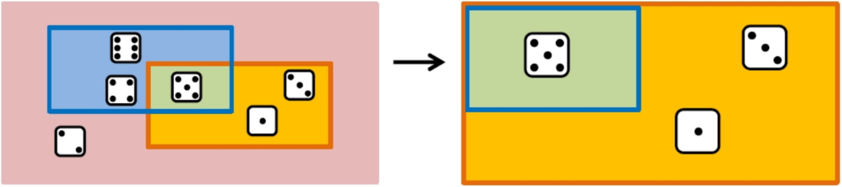
heißt **bedingte Wahrscheinlichkeit** von $A$ gegeben $B$.

# 9.1 Bedingte Wahrscheinlichkeiten

Beispiel: **einfacher Würfelwurf** $A =$ Zahl größer 3

$B =$ Zahl ungerade

Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ Wahrscheinlichkeitsraum $(B, \mathcal{B}, P_B)$



$P_{(1,\dots,6)}(A) = 3/6$ $P_{(1,\dots,6)}(B) = 3/6$

$P_{(1,\dots,6)}(A \cap B) = 1/6$

$P_{(1,3,5)}(B) = P_{(1,\dots,6)}(B \cap B) / P_{(1,\dots,6)}(B) = 1$

$P_{(1,3,5)}(A) = P_{(1,\dots,6)}(A \cap B) / P_{(1,\dots,6)}(B)$

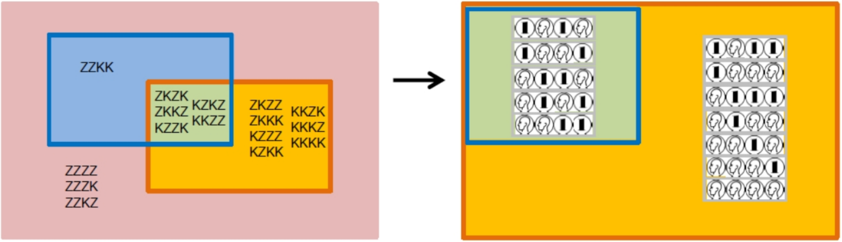
$= (1/6)/(3/6) = 1/3$

# 9.1 Bedingte Wahrscheinlichkeiten

Beispiel: **vierfacher Münzwurf** $A =$ genau zweimal Kopf nach vier Würfen

$B =$ mindestens einmal Kopf nach zwei Würfen

Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ Wahrscheinlichkeitsraum $(B, \mathcal{B}, P_B)$




$P_\Omega(A) = 6/16 = 0.375$ $P_\Omega(B) = 12/16$

$P_\Omega(A \cap B) = 5/16$

$P_B(A) = P_\Omega(A \cap B) / P_\Omega(B)$

$= (5/16)/(12/16) = 5/12 \approx 0.417$

# 9.1 Bedingte Wahrscheinlichkeiten

Umformung der Definitionsformel

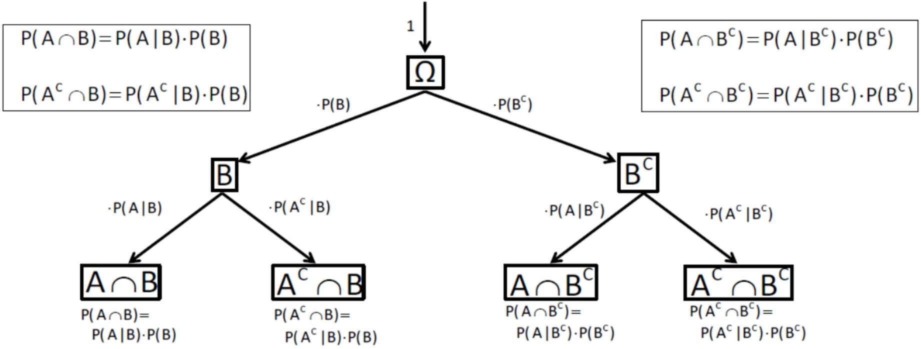
$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B) \cdot P(B)$$

$$\boxed{\begin{aligned} P(A \cap B) &= P(A|B) \cdot P(B) \\ P(A^c \cap B) &= P(A^c|B) \cdot P(B) \\ P(A \cap B^c) &= P(A|B^c) \cdot P(B^c) \\ P(A^c \cap B^c) &= P(A^c|B^c) \cdot P(B^c) \end{aligned}}$$

# 9.1 Bedingte Wahrscheinlichkeiten

Umformung der Definitionsformel

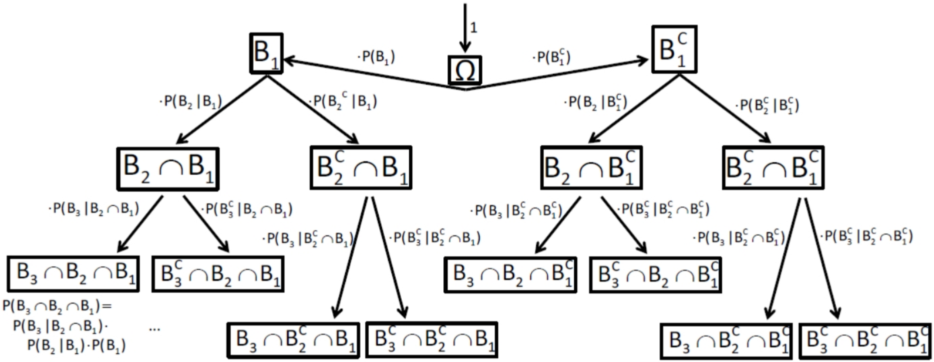
Darstellung der Wahrscheinlichkeit für Schnittereignisse in Ereignisbaum



# 9.1 Bedingte Wahrscheinlichkeiten

Umformung der Definitionsformel

Darstellung der Wahrscheinlichkeit für Schnittereignisse in Ereignisbaum

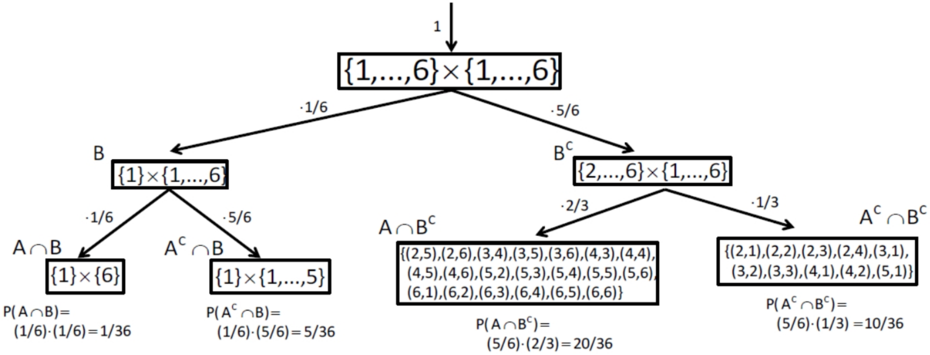


# 9.1 Bedingte Wahrscheinlichkeiten

Ereignisbaum: Beispiel **zweifacher Würfelwurf**

$A =$ Gesamtaugenzahl ist größer als 6

$B =$ erster Wurf ergibt Augenzahl 1



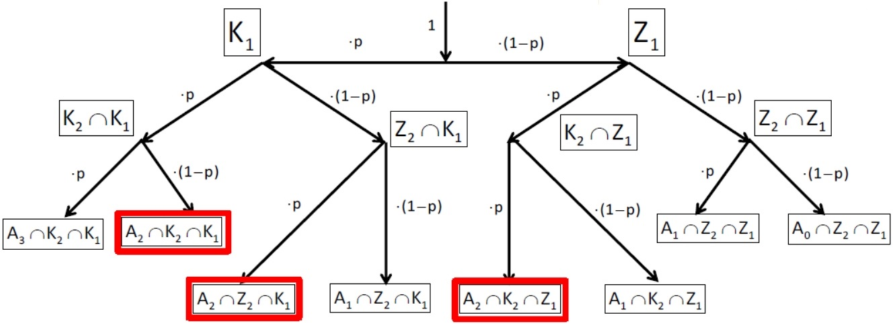
# 9.1 Bedingte Wahrscheinlichkeiten

Ergebnisbaum: Beispiel **dreifacher Münzwurf mit unfairer Münze**

$P$ : W'keit für Kopf

$K_i$ : $i$-ter Wurf Kopf, $i = 1, \ldots, 4$

$A_n$ : Gesamtzahl Kopf $= n$, $n = 0, \ldots, 3$



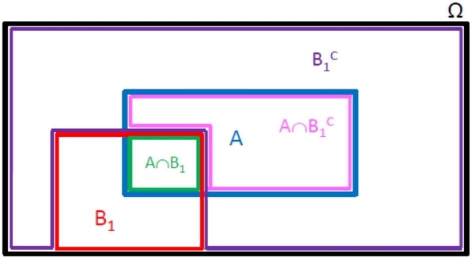

$$P(\boxed{A_2}) = p \cdot p \cdot (1-p) + p \cdot (1-p) \cdot p + (1-p) \cdot p \cdot p = \binom{3}{2} \cdot p^2 (1-p)^{3-2}$$

# 9.1 Bedingte Wahrscheinlichkeiten

## Umformung der Definitionsformel

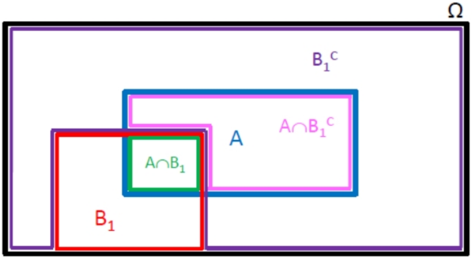
$$P(A\,|\,B_1)=\frac{P(A\cap B_1)}{P(B_1)} \Rightarrow P(\boxed{A\cap B_1})=P(\boxed{A\,|\,B_1})\cdot P(\boxed{B_1})$$

$$P(A\,|\,B_1^C)=\frac{P(A\cap B_1^C)}{P(B_1^C)} \Rightarrow P(\boxed{A\cap B_1^C})=P(\boxed{A\,|\,B_1^C})\cdot P(\boxed{B_1^C})$$



# 9.1 Bedingte Wahrscheinlichkeiten

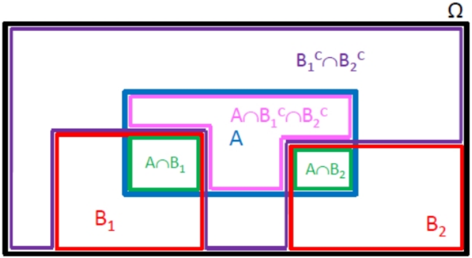
$$P(\boxed{A}) = P(\boxed{A \cap B_1}) + P(\boxed{A \cap B_1^c}) = P(\boxed{A \mid B_1}) \cdot P(\boxed{B_1}) + P(\boxed{A \mid B_1^c}) \cdot P(\boxed{B_1^c})$$



# 9.1 Bedingte Wahrscheinlichkeiten

$$B_1 \cap B_2 = \varnothing \Rightarrow P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_1^c \cap B_2^c)$$

$$= P(A \mid B_1) \cdot P(B_1) + P(A \mid B_2) \cdot P(B_2) + P(A \mid B_1^c \cap B_2^c) \cdot P(B_1^c \cap B_2^c)$$

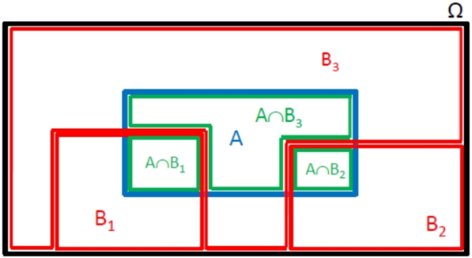


# 9.1 Bedingte Wahrscheinlichkeiten

$$B_3 = B_1^C \cap B_2^C \Rightarrow B_1 \cap B_2 = \emptyset, B_1 \cap B_3 = \emptyset, B_2 \cap B_3 = \emptyset, \; B_1 \cup B_2 \cup B_3 = \Omega$$

$$\Rightarrow P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3)$$

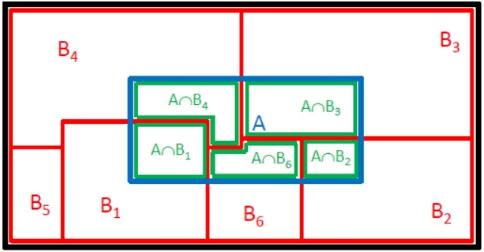
$$= P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + P(A|B_3) \cdot P(B_3)$$



# 9.1 Bedingte Wahrscheinlichkeiten

## Satz von der totalen Wahrscheinlichkeit

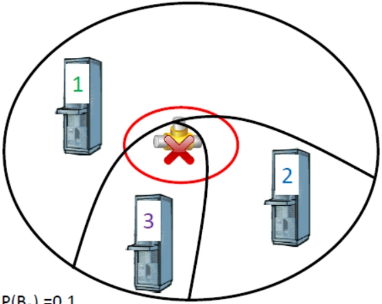
$$i \neq j \Rightarrow B_i \cap B_j = \emptyset, \ \bigcup_{i=1}^{n} B_i = \Omega \ \Rightarrow \ P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$$



# 9.1 Bedingte Wahrscheinlichkeiten

## Satz von der totalen Wahrscheinlichkeit

Beispiel: **Ausfall einer Internetverbindung**




A = Internetverbindung fällt aus

$B_1$ = Verbindung mit Knoten 1

$B_2$ = Verbindung mit Knoten 2

$B_3$ = Verbindung mit Knoten 3

W'keiten für Einwahl in Knoten 1, 2 oder 3: $P(B_1) = 0.7$, $P(B_2) = 0.2$, $P(B_3) = 0.1$

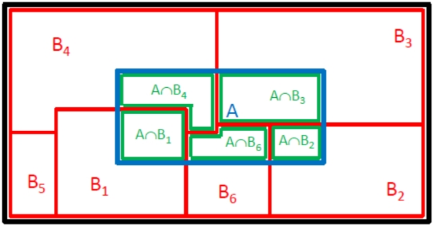
W'keiten für Verbindungsausfall für Knoten 1, 2 oder 3 : $P(A|B_1) = 0.02$, $P(A|B_2) = 0.04$, $P(A|B_3) = 0.06$

$\Rightarrow$ W'keit für Verbindungsausfall: $P(A) = 0.02 \cdot 0.7 + 0.04 \cdot 0.2 + 0.06 \cdot 0.1 = 0.028$

# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von der totalen Wahrscheinlichkeit** $i \neq j \Rightarrow B_i \cap B_j = \emptyset, \; \bigcup_{i=1}^{n} B_i = \Omega$

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

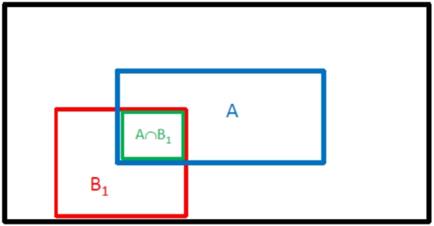


# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von der totalen Wahrscheinlichkeit** $i \neq j \Rightarrow B_i \cap B_j = \emptyset, \ \bigcup_{i=1}^{n} B_i = \Omega$

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

$$\begin{aligned} P(A \cap B_i) &= P(A|B_i) \cdot P(B_i) \\ &= P(B_i|A) \cdot P(A) \end{aligned}$$



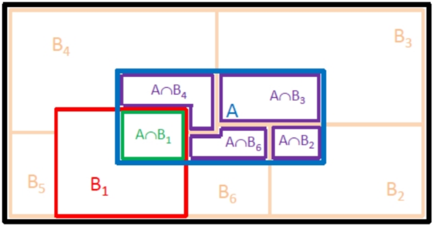
# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von der totalen Wahrscheinlichkeit** $\quad i \neq j \Rightarrow B_i \cap B_j = \emptyset, \ \bigcup_{i=1}^{n} B_i = \Omega$

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

$$\begin{aligned} P(A \cap B_i) &= P(A|B_i) \cdot P(B_i) \\ &= P(B_i|A) \cdot P(A) \end{aligned}$$

$$\begin{aligned} \Rightarrow P(B_i|A) &= \frac{P(A \cap B_i)}{P(A)} \\ &= \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j=1}^{n} P(A|B_j) \cdot P(B_j)} \end{aligned}$$



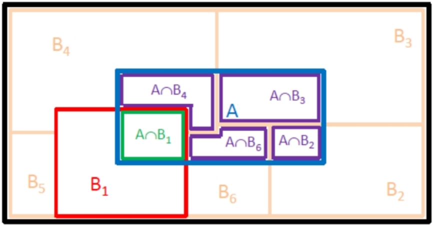
# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes** $\quad i \neq j \Rightarrow B_i \cap B_j = \emptyset, \ \bigcup_{i=1}^{n} B_i = \Omega$

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j=1}^{n} P(A|B_j) \cdot P(B_j)}$$

$P(B_1), \ldots, P(B_n)$ heißen **a-priori-Wahrscheinlichkeiten**,

$P(B_1|A), \ldots, P(B_n|A)$ heißen **a-posteriori-Wahrscheinlichkeiten**



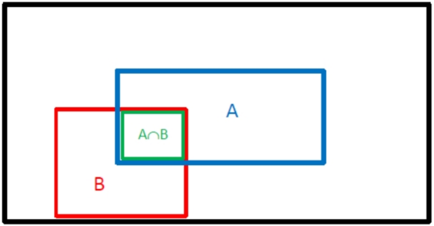
# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes** $\quad i \neq j \Rightarrow B_i \cap B_j = \emptyset, \; \bigcup_{i=1}^{n} B_i = \Omega$

$$\Rightarrow P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum\limits_{j=1}^{n} P(A|B_j) \cdot P(B_j)}$$

Insbesondere

$$A, B \in \Omega \Rightarrow P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$



# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes**, Beispiel **Spam-Filter** $P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$

Ereignis A: „Mail enthält das Wort *Maximalgewinn*“ $\Rightarrow$ Klassifiziere Mail als Spam

Ereignis B: „Mail ist Spam“

$P(A|B)$ = **Sensitivität** = W'keit, Spam als solchen zu klassifizieren

$P(A^c|B^c)$ = **Spezifität** = W'keit, normale Mails nicht als Spam zu klassifizieren

| | $B$ : Mail ist Spam | $B^c$ : Mail ist kein Spam |
|---|---|---|
| $A$ : *Maximalgewinn* in Mail | $P(A\|B)$ | $P(A\|B^c) = 1 - P(A^c\|B^c)$ |
| $A^c$ : *Maximalgewinn* nicht in Mail | $P(A^c\|B) = 1 - P(A\|B)$ | $P(A^c\|B^c)$ |

# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes**, Beispiel **Spam-Filter** $P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$

$P(A|B)$ = **Sensitivität** = W'keit, Spam als solchen zu klassifizieren

$P(A^c|B^c)$ = **Spezifität** = W'keit, normale Mails nicht als Spam zu klassifizieren

Gesucht:

$P(B|A)$ = W'keit, dass klassifizierte Mail Spam ist

Im Satz von Bayes werden $P(A)$ und $P(B)$ benötigt.

$$P(A) = P(A|B) \cdot P(B) + P(A|B^c) \cdot (1 - P(B))$$

Für die Berechnung von $P(B|A)$ ist also die Angabe der Prävalenz $P(B)$ ausreichend.

# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes**, Beispiel **Spam-Filter**

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

Gegeben:

$P(A|B) =$ Sensitivität

$P(A^c|B^c) =$ Spezifiztät

$P(B) =$ Prävalenz

| | $B$ : Mail ist Spam | $B^c$ : Mail ist kein Spam |
|---|---|---|
| $A$ : *Maximalgewinn* in Mail | $P(A\|B)$ | $P(A\|B^c)$ |
| $A^c$ : *Maximalgewinn* nicht in Mail | $P(A^c\|B)$ | $P(A^c\|B^c)$ |

$$\Rightarrow \boxed{P(B|A)} = \frac{\boxed{P(A|B)} \cdot \boxed{P(B)}}{\boxed{P(A|B)} \cdot \boxed{P(B)} + [1 - \boxed{P(A^C|B^C)}] \cdot [1 - \boxed{P(B)}]}$$

# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes**, Beispiel **Spam-Filter** $P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$

Gegeben:

$P(A|B)$ = Sensitivität =0.95

$P(A^c|B^c)$ = Spezifiztät =0.98

$P(B)$ = Prävalenz =0.3

| | $B$ : Mail ist Spam | $B^c$ : Mail ist kein Spam |
|---|---|---|
| $A$ : *Maximalgewinn* in Mail | 0.95 | 0.02 |
| $A^c$ : *Maximalgewinn* nicht in Mail | 0.05 | 0.98 |

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A|B) \cdot P(B) + [1 - P(A^c|B^c)] \cdot [1 - P(B)]}$$

$$= \frac{0.95 \cdot 0.3}{0{,}95 \cdot 0.3 + (1 - 0.98) \cdot (1 - 0.3)} = \frac{0.285}{0.299} \approx 0.9532$$

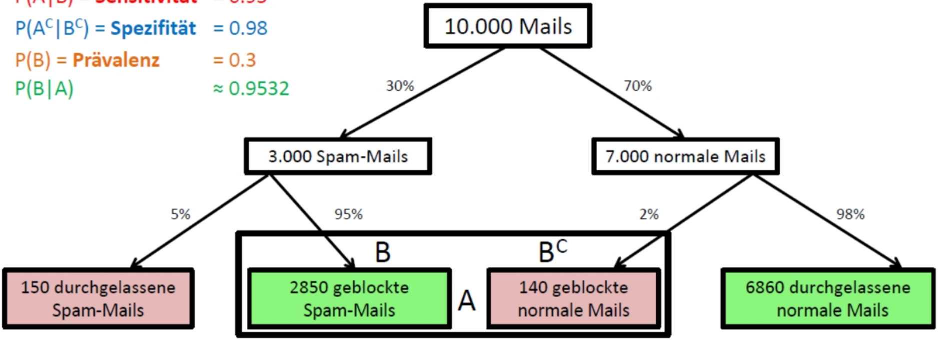
$\Rightarrow$ Wenn 30% der Mails Spam sind, sind bei einer Sensitivität von 95% und bei einer Spezifität von 98% ca. 95.3% der als Spam klassifizierten Mails tatsächlich Spam

# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes**, Beispiel **Spam-Filter**

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

P(A|B) = **Sensitivität** = 0.95
P(A^C|B^C) = **Spezifität** = 0.98
P(B) = **Prävalenz** = 0.3
P(B|A) ≈ 0.9532




P(B|A) = 2850/(2850+140) ≈ 95.3%

# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes**, Beispiel **Spam-Filter** $P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$

Gegeben:

$P(A|B)$ = Sensitivität =0.95

$P(A^c|B^c)$ = Spezifiztät =0.98

$P(B)$ = Prävalenz =0.005

| | $B$ : Mail ist Spam | $B^c$ : Mail ist kein Spam |
|---|---|---|
| $A$ : *Maximalgewinn* in Mail | 0.95 | 0.02 |
| $A^c$ : *Maximalgewinn* nicht in Mail | 0.05 | 0.98 |

$$P(B|A) = \frac{P(A|B) \cdot P_\Omega(B)}{P(A|B) \cdot P_\Omega(B) + [1 - P(A^c|B^c)] \cdot [1 - P\Omega(B)]}$$

$$= \frac{0.95 \cdot 0.005}{0{,}95 \cdot 0.005 + (1 - 0.98) \cdot (1 - 0.005)} = \frac{0.00475}{0.02465} \approx 0.1927$$

$\Rightarrow$ Wenn 0.5% der Mails Spam sind, sind bei einer Sensitivität von 95% und bei einer Spezifität von 98% ca. 19.3% der als Spam klassifizierten Mails tatsächlich Spam

# 9.1 Bedingte Wahrscheinlichkeiten

**Satz von Bayes**, Beispiel **Spam-Filter**

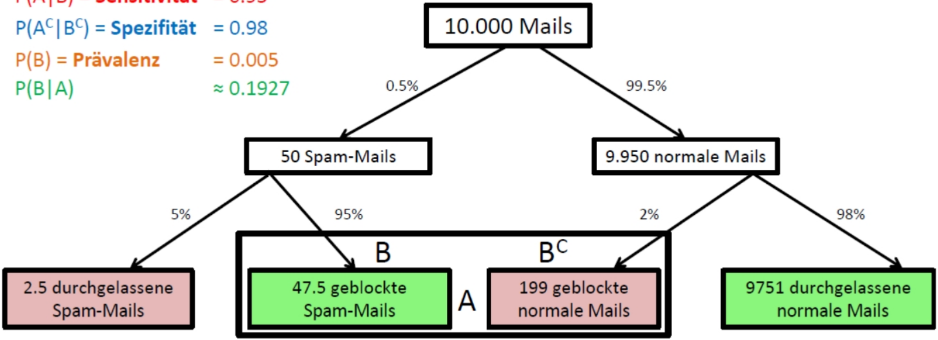
$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

P(A|B) = **Sensitivität** = 0.95

P(A$^C$|B$^C$) = **Spezifität** = 0.98

P(B) = **Prävalenz** = 0.005

P(B|A) ≈ 0.1927




P(B|A) = 47.5/(47.5+199) ≈ 19.27%

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Zwei Ereignisse $A$ und $B$ aus $(\Omega, \mathcal{A}, P)$ heißen **stochastisch unabhängig**, wenn gilt:

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

Sind $P(B) > 0$ und $P(A)$ stochastisch unabhängig, so sind bedingte und unbedingte Wahrscheinlichkeit von $A$ gleich:

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

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Eine Menge von $n$ Ereignissen $A_1, \dots, A_n$ aus $(\Omega, A, P)$ heißt **paarweise stochastisch unabhängig**, wenn gilt:

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j),\ j \neq i,\ j = 1, \dots, n,\ i = 1, \dots, n$$

Eine Menge von $n$ Ereignissen aus $(\Omega, A, P)$ heißt **gemeinsam stochastisch unabhängig**, wenn gilt:

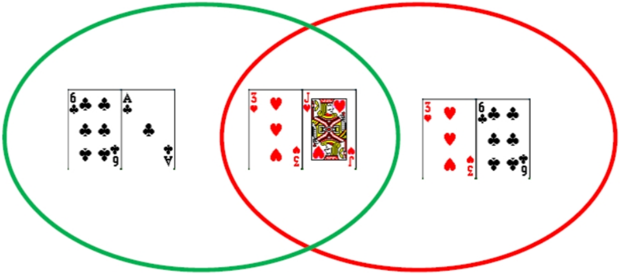
$$P\Big(\bigcap_{j=1}^{s} A_{i_j}\Big) = \prod_{j=1}^{s} P(A_{i_j}),\ \{i_1, \dots, i_s\} \subseteq \{1, \dots, n\}$$

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Beispiel: **wiederholtes Ziehen aus gemischtem Skatspiel**

A: zweite Karte hat gleiche Symbolfarbe wie erste Karte

B: erste Karte ist rot



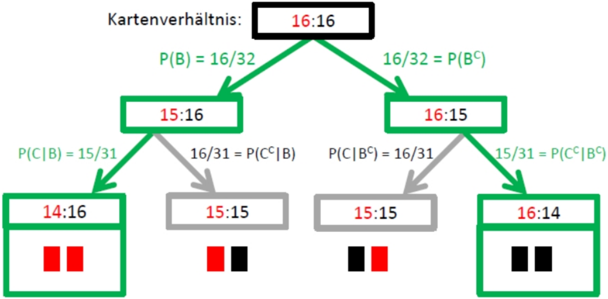
# 9.2 Stochastische Unabhängigkeit von Ereignissen

Beispiel: **wiederholtes Ziehen aus gemischtem Skatspiel**

A: zweite Karte hat gleiche Symbolfarbe wie erste Karte

B: erste Karte ist rot

C: zweite Karte ist rot




$$P(\boxed{A}) = \frac{16\cdot 15}{32\cdot 31} + \frac{16\cdot 15}{32\cdot 31} = \frac{15}{31}$$

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Beispiel: **wiederholtes Ziehen aus gemischtem Skatspiel**

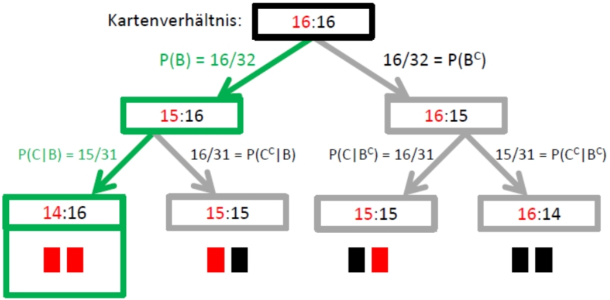
A: zweite Karte hat gleiche Symbolfarbe wie erste Karte

B: erste Karte ist rot

C: zweite Karte ist rot

$$P(\boxed{A}) = \frac{15}{31}$$

$$P(\boxed{A} \mid \boxed{B}) = \left(\frac{16 \cdot 15}{32 \cdot 31}\right) \Big/ \left(\frac{16}{32}\right) = \frac{15}{31}$$



# 9.2 Stochastische Unabhängigkeit von Ereignissen

Beispiel: **wiederholtes Ziehen aus gemischtem Skatspiel**

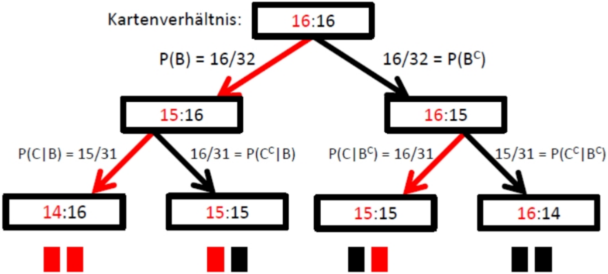
A: zweite Karte hat gleiche Symbolfarbe wie erste Karte

B: erste Karte ist rot

C: zweite Karte ist rot

$$P(A) = \frac{15}{31} = P(A \mid B)$$

$\Rightarrow$ A und B sind stochastisch unabhängig



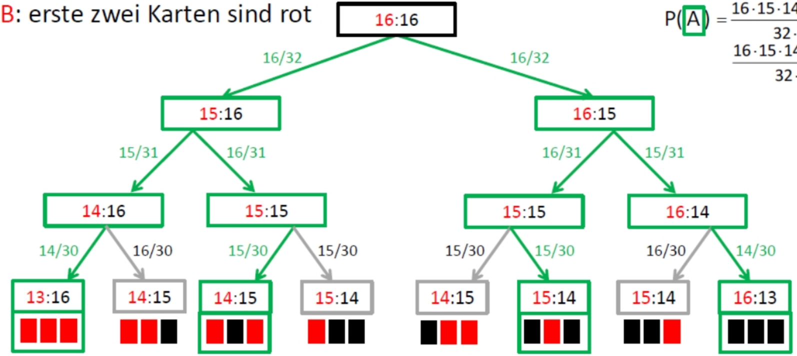
# 9.2 Stochastische Unabhängigkeit von Ereignissen

Beispiel: **wiederholtes Ziehen aus gemischtem Skatspiel**

A: dritte Karte hat gleiche Symbolfarbe wie erste Karte

B: erste zwei Karten sind rot

$$P(\boxed{A}) = \frac{16\cdot 15\cdot 14+16\cdot 16\cdot 15}{32\cdot 31\cdot 30} + \frac{16\cdot 15\cdot 14+16\cdot 16\cdot 15}{32\cdot 31\cdot 30} \approx 0.4839$$



# 9.2 Stochastische Unabhängigkeit von Ereignissen

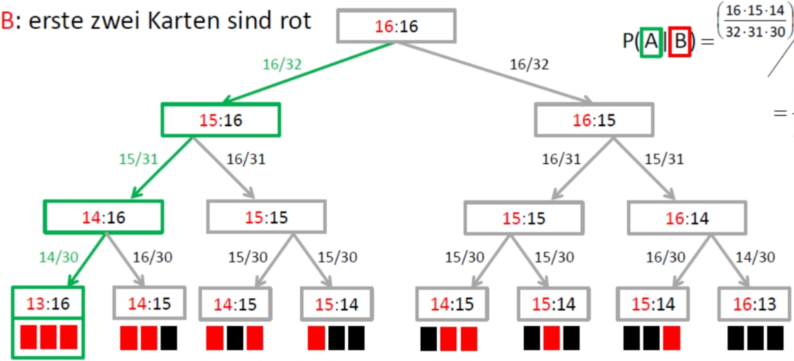
Beispiel: **wiederholtes Ziehen aus gemischtem Skatspiel**

$P(\boxed{A}) \approx 0.4839$

A: dritte Karte hat gleiche Symbolfarbe wie erste Karte

B: erste zwei Karten sind rot

$$P(\boxed{A}|\boxed{B}) = \left(\frac{16\cdot 15\cdot 14}{32\cdot 31\cdot 30}\right) \Big/ \left(\frac{16\cdot 15}{32\cdot 31}\right) = \frac{14}{30} \approx 0.4667$$



# 9.2 Stochastische Unabhängigkeit von Ereignissen

Beispiel: **wiederholtes Ziehen aus gemischtem Skatspiel**

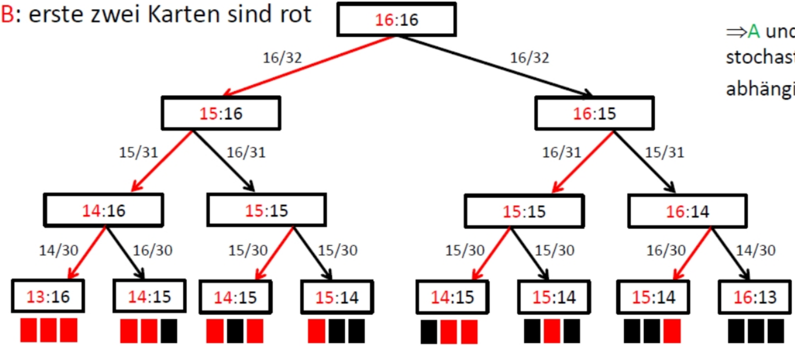
A: dritte Karte hat gleiche Symbolfarbe wie erste Karte

B: erste zwei Karten sind rot

$P(A) \approx 0.4839$

$\neq 0.4667 = P(A \mid B)$

$\Rightarrow$ A und B sind stochastisch abhängig



# 9.2 Stochastische Unabhängigkeit von Ereignissen

Seien $X$ und $Y$ Zufallsvariablen mit Verteilungsfunktionen $F^X$, $F^Y$ und $F^{(X,Y)}$.

Dann heißen $X$ und $Y$ stochastisch unabhängig, falls

$$F^{(X,Y)} = F^X(x) \cdot F^Y(y) \text{ für alle } x, y \in \mathbb{R}$$

Die Zufallsvariablen $X_1, \ldots, X_n$ heißen stochastisch unabhängig, falls

$$F^{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) = \prod_{i=1}^{n} F^{X_i}(x_i) \text{ für alle } x_1, \ldots, x_n \in \mathbb{R}$$

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Seien $X$ und $Y$ Zufallsvariablen mit Verteilungsfunktionen $F^X$, $F^Y$ und $F^{(X,Y)}$.

Dann heißen $X$ und $Y$ stochastisch unabhängig, falls

$$F^{(X,Y)}(x,y) = F^X(x) \cdot F^Y(y) \text{ für alle } x, y \in \mathbb{R}$$

$$\Rightarrow F^{(X,Y)}(x,y) = \boxed{P(A \cap B)} = \boxed{P(A) \cdot P(B)} = F^X(x) \cdot F^Y(y)$$

$$\text{mit } A = \{\omega \in \Omega | X(\omega) \leq x\},\ B = \{\omega \in \Omega | Y(\omega) \leq y\}$$

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Seien $X$ und $Y$ Zufallsvariablen mit Verteilungsfunktionen $F^X$, $F^Y$ und $F^{(X,Y)}$.

Dann heißen $X$ und $Y$ stochastisch unabhängig, falls

$$F^{(X,Y)} = F^X(x) \cdot F^Y(y) \text{ für alle } x, y \in \mathbb{R}$$

$F^{(X,Y)}$ diskret

$$\begin{aligned}
\Rightarrow \boxed{p(x_i, y_j)} &= P(A \cap B) \\
&= F^{(X,Y)}(x_i, y_j) - F^{(X,Y)}(x_{i-1}, y_j) - F^{(X,Y)}(x_i, y_{j-1}) \\
&\quad + F^{(X,Y)}(x_{i-1}, y_{j-1}) \\
&= F^X(x_i)F^Y(y_j) - F^X(x_{i-1})F^Y(y_j) - F^X(x_i)F^Y(y_{j-1}) \\
&\quad + F^X(x_{i-1})F^Y(y_{j-1}) \\
&= \left[F^X(x_i) - F^X(x_{i-1})\right] \cdot \left[F^Y(y_i) - F^Y(y_{j-1})\right] = \boxed{p(x_i) \cdot p(y_j)}
\end{aligned}$$

$$\text{mit } A = \{\omega \in \Omega | X(\omega) \le x_i\},\ B = \{\omega \in \Omega | Y(\omega) \le y_j\}$$

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Seien $X$ und $Y$ Zufallsvariablen mit Verteilungsfunktionen $F^X$, $F^Y$ und $F^{(X,Y)}$.

Dann heißen $X$ und $Y$ stochastisch unabhängig, falls

$$F^{(X,Y)} = F^X(x) \cdot F^Y(y) \text{ für alle } x, y \in \mathbb{R}$$

Seien $X$ und $Y$ Zufallsvariablen mit Verteilungsfunktionen $F^X$, $F^Y$ und $F^{(X,Y)}$.

Dann heißen $X$ und $Y$ stochastisch unabhängig, falls $F^{(X,Y)}$ stetig

$$\Rightarrow \boxed{f^{(X,Y)}(x,y)} = \frac{\delta F^{(X,Y)}(x,y)}{\delta x \delta y} = \frac{\delta[F^X(x) \cdot F^Y(y)]}{\delta x \delta y} = \frac{\delta F^X(x) \cdot \delta F^Y(y)}{\delta x \delta y}$$
$$= \boxed{f^X(x) \cdot f^Y(y)}$$

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Beispiel: **Multinomialverteilung**

$(X, Y) \sim \text{Mult}(2, 0.5, 0.5)$

$$\begin{aligned} \Rightarrow p^{XY}(0,2) &= \frac{2}{0! \cdot 2!} \cdot 0.5^2 = \frac{2}{1 \cdot 2} \cdot \frac{1}{4} = \frac{1}{4} \\ p^{XY}(1,1) &= \frac{2}{1! \cdot 1!} \cdot 0.5^2 = \frac{2}{1 \cdot 1} \cdot \frac{1}{4} = \frac{1}{2} \\ p^{XY}(2,0) &= \frac{2}{2! \cdot 0!} \cdot 0.5^2 = \frac{2}{2 \cdot 1} \cdot \frac{1}{4} = \frac{1}{4} \end{aligned}$$

# 9.2 Stochastische Unabhängigkeit von Ereignissen

Beispiel: **Multinomialverteilung**

$(X, Y) \sim \text{Mult}(2, 0.5, 0.5)$

$$p^{XY}(0,2) = \frac{1}{4} \quad p^X(0) = p^Y(2) = \frac{1}{4} \quad p^X(0) \cdot p^Y(2) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16} \neq \frac{1}{4} = p^{XY}(0,2)$$

$$p^{XY}(1,1) = \frac{1}{2} \quad p^X(1) = p^Y(1) = \frac{1}{2} \quad p^X(1) \cdot p^Y(1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \neq \frac{1}{2} = p^{XY}(1,1)$$

$$p^{XY}(2,0) = \frac{1}{4} \quad p^X(2) = p^Y(0) = \frac{1}{4} \quad p^X(2) \cdot p^Y(0) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16} \neq \frac{1}{4} = p^{XY}(2,0)$$

$\Rightarrow$ $X$ und $Y$ sind stochastisch abhängig