



Statistik

Relative Häufigkeit, Anteil & Wahrscheinlichkeit

Relative Häufigkeit, Anteil und Wahrscheinlichkeit



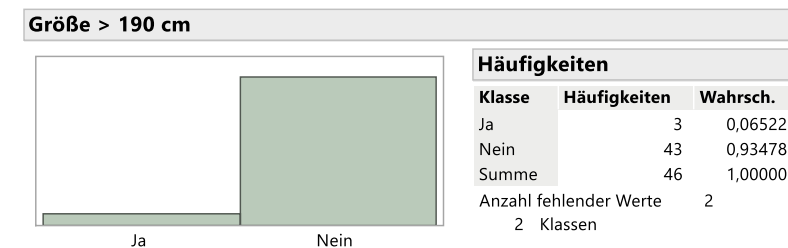
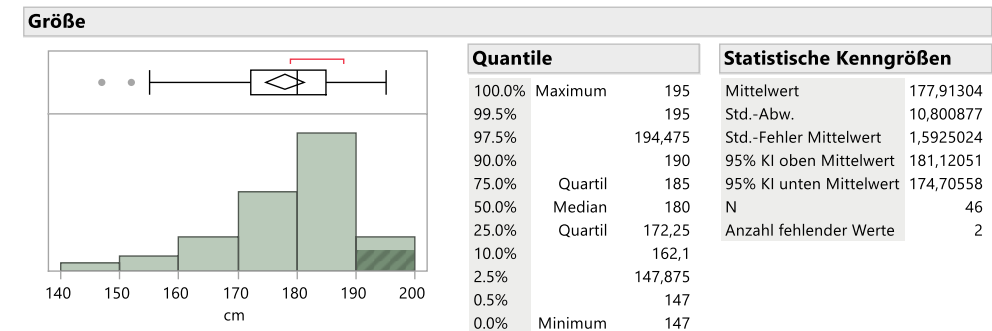
> Relative Häufigkeit Studierende Größe > 190 cm?

– Absolute Häufigkeit:

$$h(x > 190) = 3$$

– Relative Häufigkeit:

$$f(x > 190) = \frac{h(x > 190)}{n} = \frac{3}{46} \approx 0,065$$



> Wie wahrscheinlich ist es, „blind“ (= zufällig) mit einem Ball im Vorlesungssaal eine Person größer als 190 cm zu treffen?

Relative Häufigkeit, Anteil und Wahrscheinlichkeit



> Normalverteilte Zufallsgröße Größe $\sim N(171,6; 9,6^2)$

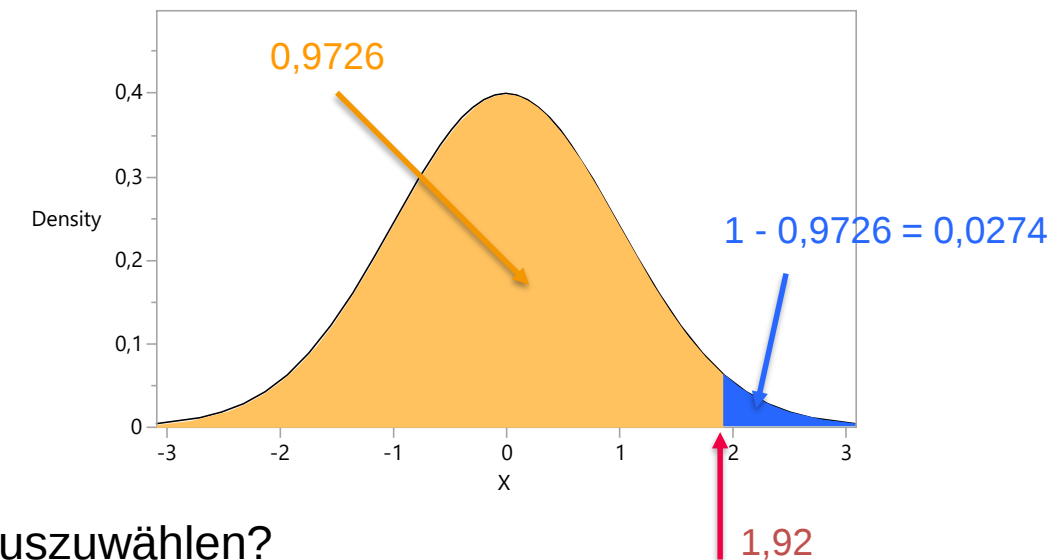
> Anteil Bevölkerung Größe > 190 cm?

– Z-Transformation

$$z = \frac{190 - 171,6}{9,6} \approx 1,92$$

– $P(X > 190) = 0,0274$

– Anteil $\approx 2,7\%$



> Wahrscheinlichkeit zufällig eine Person größer als 190 cm auszuwählen?

$\Rightarrow 0,0274$ oder $\sim 2,7\%$

> Unter 100 zufällig ausgewählten Personen

$\Rightarrow \sim 2-3$ größer 190 cm



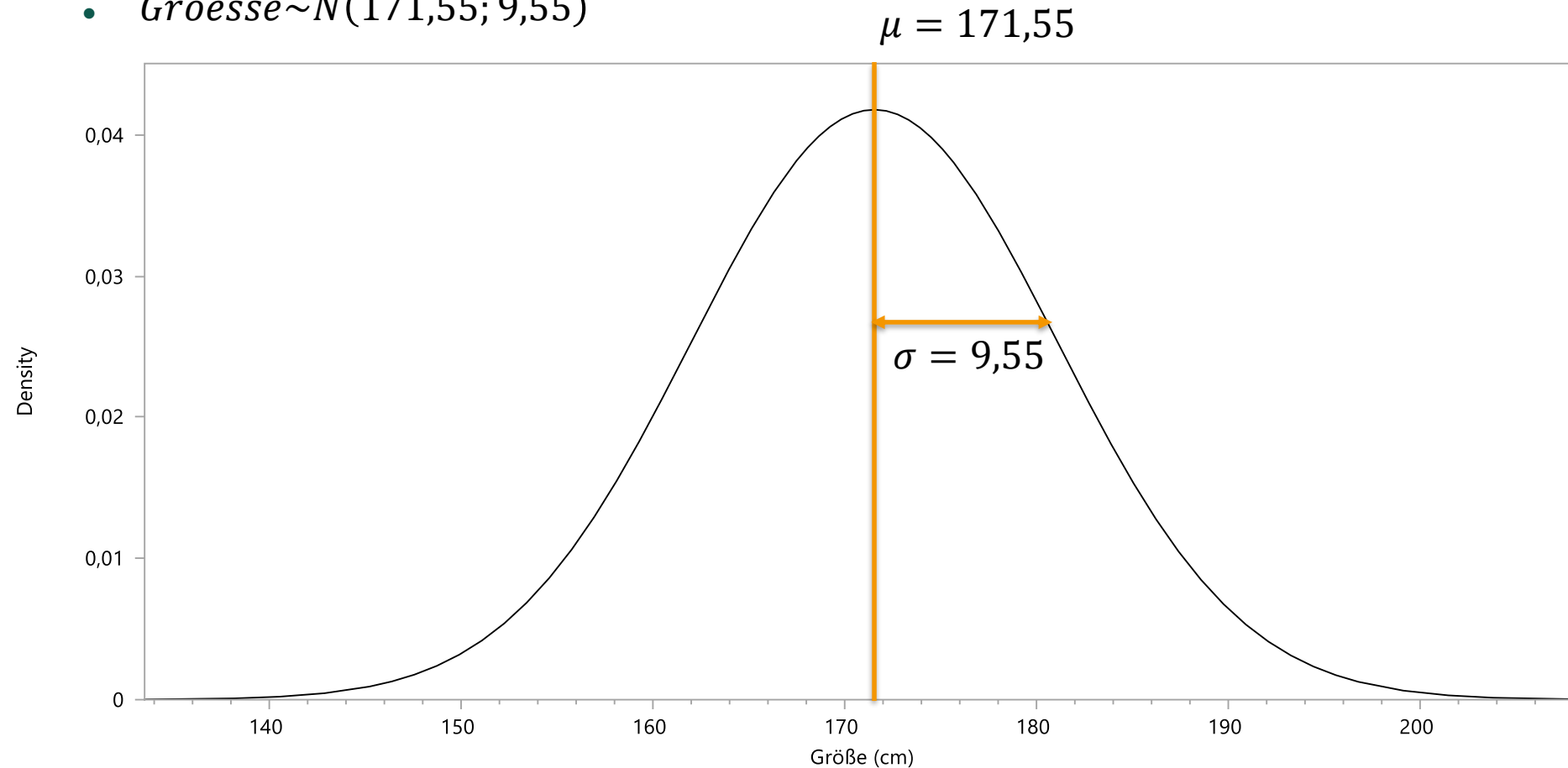
Statistik

Wahrscheinlichkeitsintervalle und Punktschätzer

Wahrscheinlichkeitsintervall für Punktschätzer

Verteilung Größe Deutschland

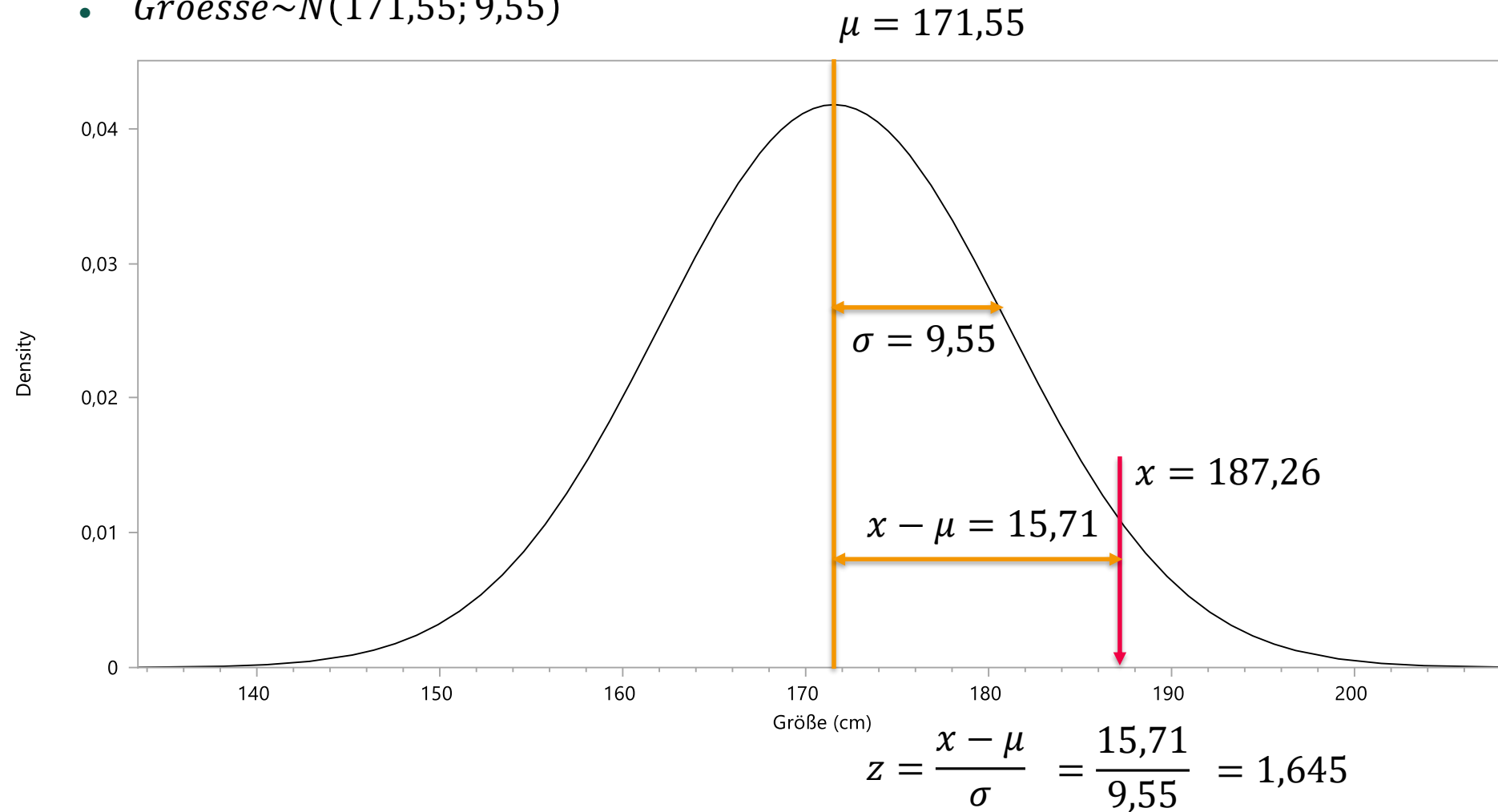
- $Groesse \sim N(171,55; 9,55)$



Wahrscheinlichkeitsintervall für Punktschätzer

Verteilung Größe Deutschland

- $Groesse \sim N(171,55; 9,55)$

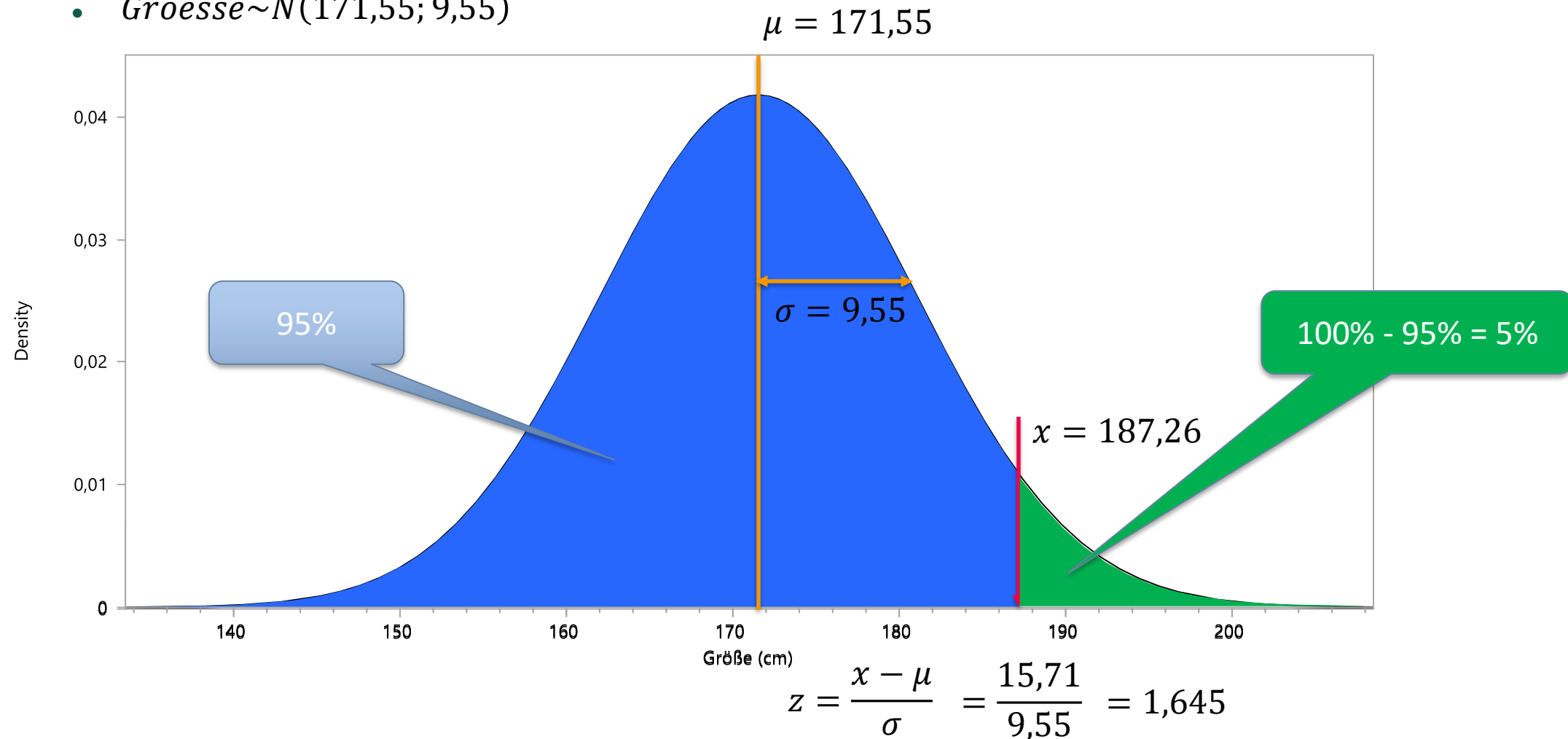


Wahrscheinlichkeitsintervall für Punktschätzer

Verteilung Größe Deutschland



- $Groesse \sim N(171,55; 9,55)$

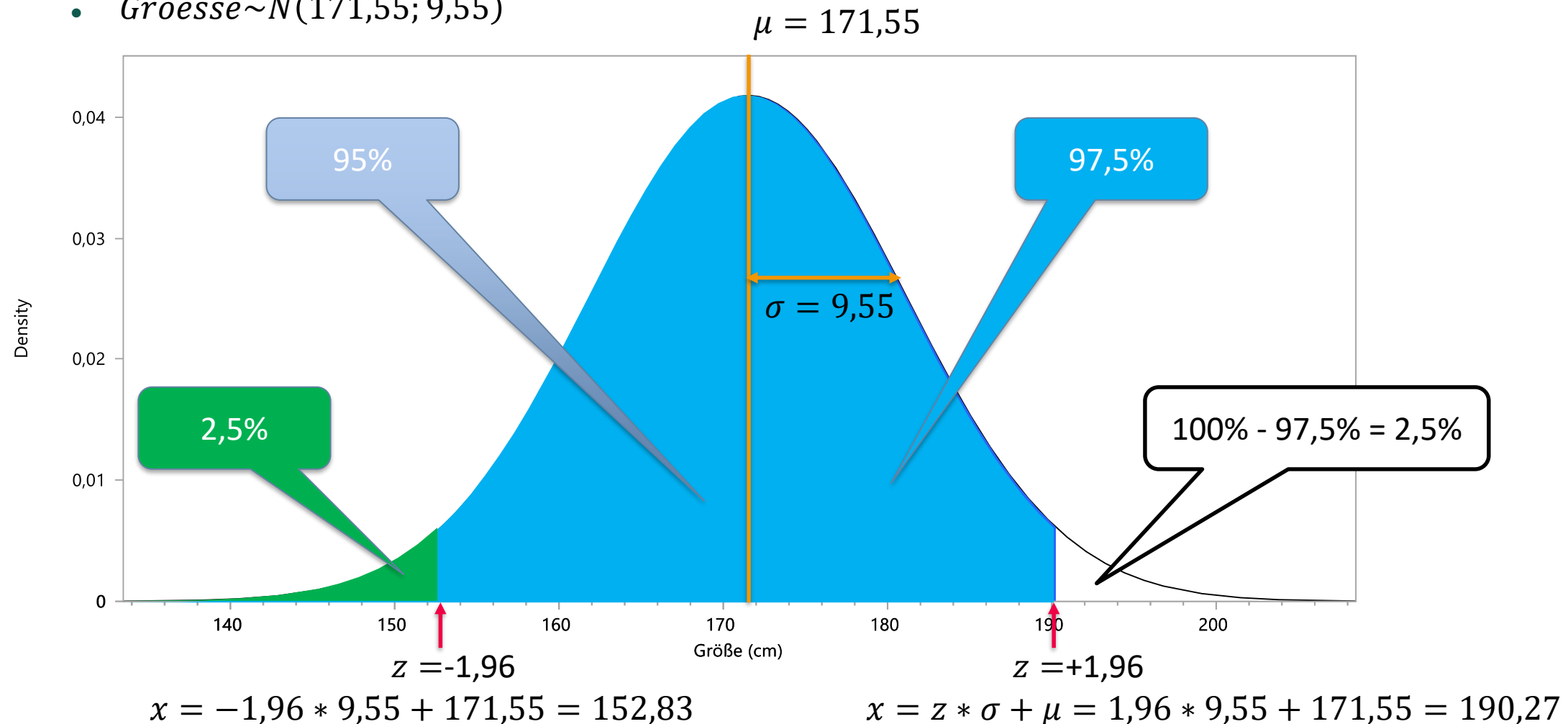


Wahrscheinlichkeitsintervall für Punktschätzer

Verteilung Größe Deutschland



- $Groesse \sim N(171,55; 9,55)$



k-σ-Bereiche, Sigma-Regeln



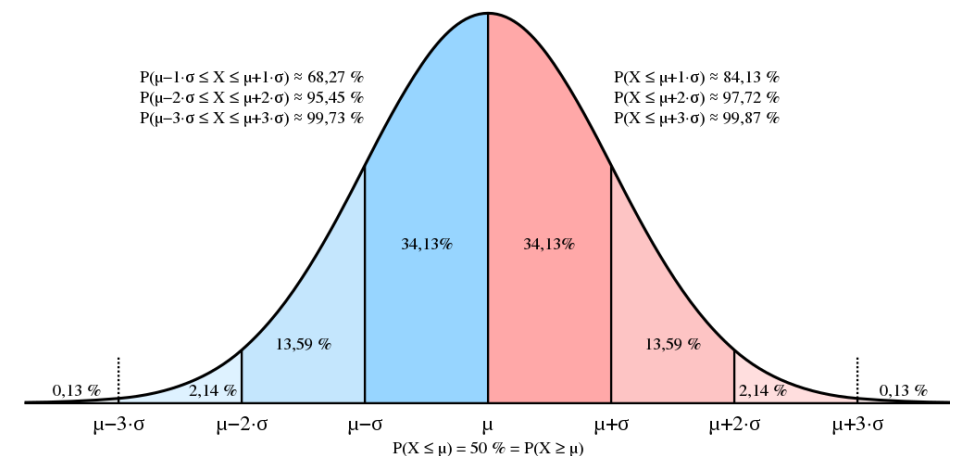
> Ermittlung eines Intervalls für gegebene Wahrscheinlichkeit

- Normalverteilte Zufallsgrößen $X \sim N(\mu, \sigma^2)$
 - Symmetrisch um Erwartungswert μ
- Intervall i.d.R. durch Vielfachens der Standardabweichung σ zum Erwartungswert μ
 - Z-Transformation (siehe frühere Folie)

$$P(|X - \mu| \leq k * \sigma) = P(\mu - k * \sigma \leq X \leq \mu + k * \sigma)$$

> Sigma-Regeln

Intervall I	$P(X \in I)$	Intervall I	$P(X \in I)$
$[\mu - \sigma; \mu + \sigma]$	$\approx 0,683$	$[\mu - 1,64\sigma; \mu + 1,64\sigma]$	$\approx 0,90$
$[\mu - 2\sigma; \mu + 2\sigma]$	$\approx 0,954$	$[\mu - 1,96\sigma; \mu + 1,96\sigma]$	$\approx 0,95$
$[\mu - 3\sigma; \mu + 3\sigma]$	$\approx 0,997$	$[\mu - 2,58\sigma; \mu + 2,58\sigma]$	$\approx 0,99$



Sigma-Regeln, k- σ -Bereiche

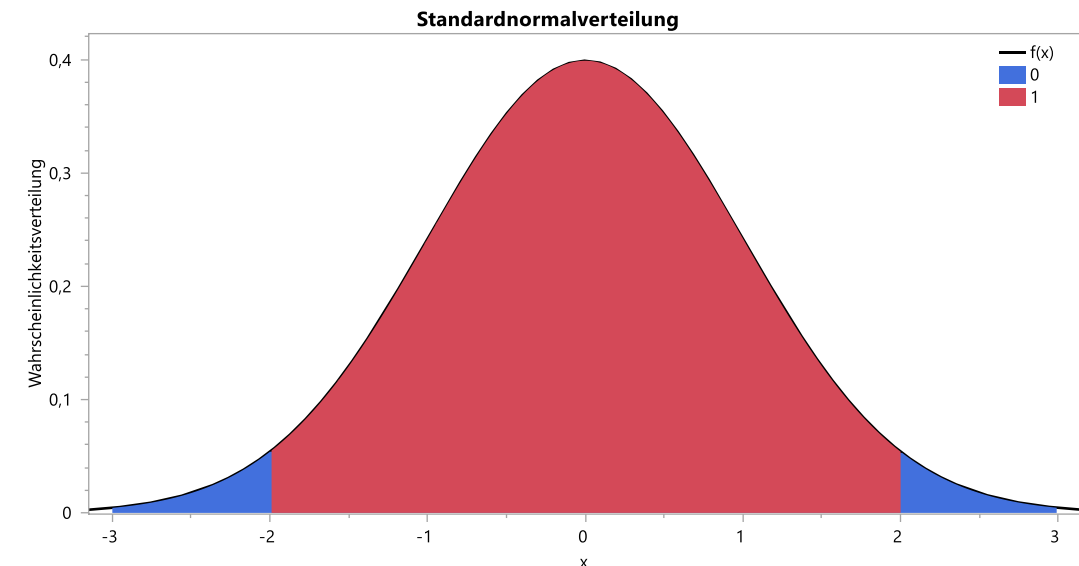
Beispiele

- > Geben Sie zu einer $N(200, 75^2)$ -verteilten Zufallsgröße das um den Erwartungswert symmetrische Intervall an, welches eine Wahrscheinlichkeit von 0,95 hat
- Gesucht: Vielfaches k von σ , so dass $[\mu - k\sigma; \mu + k\sigma] = 0,95$
 - $\mu = 200, \sigma = 75$
 - Die Zufallsgröße weicht dann mit der Wahrscheinlichkeit 0,95 höchstens um $k * \sigma$ vom Erwartungswert μ ab

$$P(|X - 200| \leq k * 75) = 0,95$$

$$\Leftrightarrow P(200 - k * 75 \leq X \leq 200 + k * 75) = 0,95$$

$$\Rightarrow P(-k \leq X \leq k) = 0,95$$



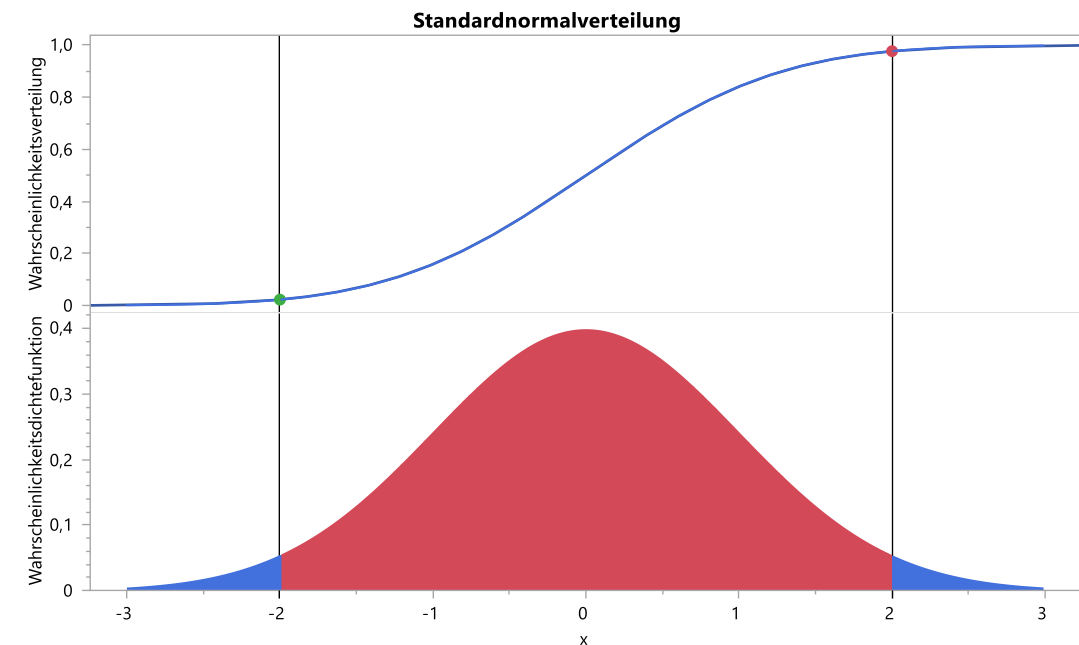
Sigma-Regeln, k- σ -Bereiche

Beispiele

- > Geben Sie zu einer $N(200, 75^2)$ -verteilten Zufallsgröße das um den Erwartungswert symmetrische Intervall an, welches eine Wahrscheinlichkeit von 0,95 hat
- Gesucht: Vielfaches k von σ , so dass $[\mu - k\sigma; \mu + k\sigma] = 0,95$, $\mu = 200$, $\sigma = 75$
 - Die Zufallsgröße weicht dann mit der Wahrscheinlichkeit 0,95 höchstens um $k * \sigma$ vom Erwartungswert μ ab

$$\Rightarrow P(-k \leq X \leq k) = 0,95$$

$$\Leftrightarrow P(X \leq k) - P(X \leq -k) = 0,95$$



Sigma-Regeln, k- σ -Bereiche

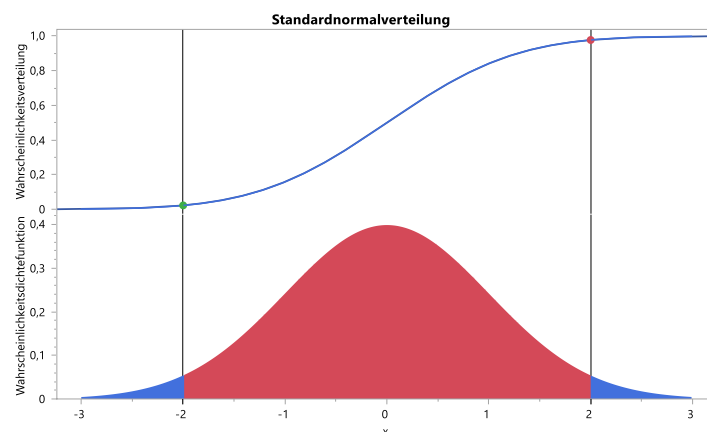
Beispiele

> Geben Sie zu einer $N(200, 75^2)$ -verteilten Zufallsgröße das um den Erwartungswert symmetrische Intervall an, welches eine Wahrscheinlichkeit von 0,95 hat

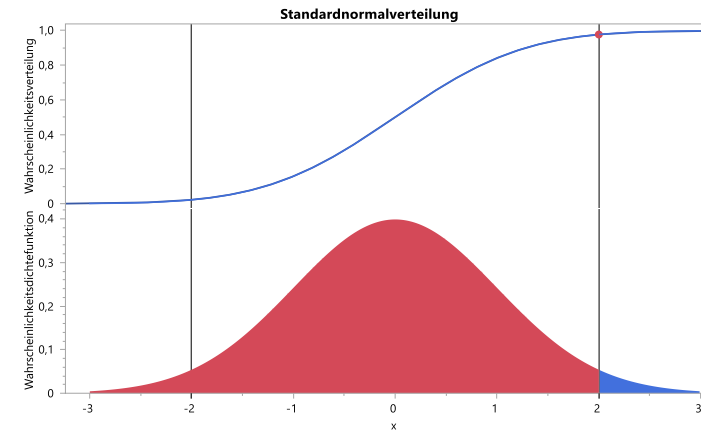
- Gesucht: Vielfaches k von σ , so dass $[\mu - k\sigma; \mu + k\sigma] = 0,95$, $\mu = 200$, $\sigma = 75$
- Die Zufallsgröße weicht dann mit der Wahrscheinlichkeit 0,95 höchstens um $k * \sigma$ vom Erwartungswert μ ab

$$\Rightarrow P(X \leq k) - P(X \leq -k) = 0,95$$

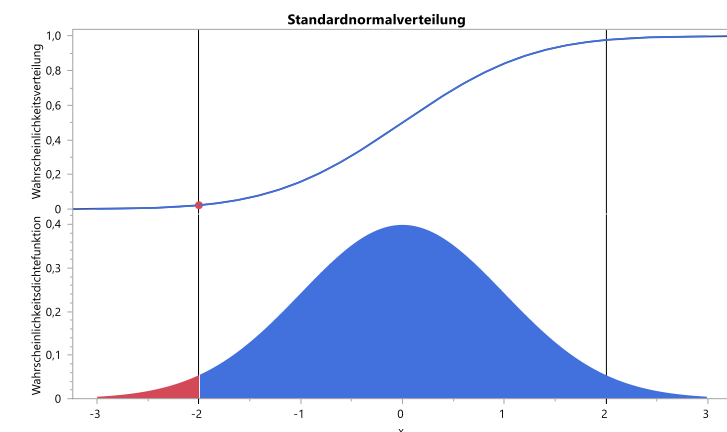
$$\Leftrightarrow \Phi(k) - \Phi(-k) = 0,95$$



=



-



Sigma-Regeln, k - σ -Bereiche

Beispiele



> Geben Sie zu einer $N(200, 75^2)$ -verteilten Zufallsgröße das um den Erwartungswert symmetrische Intervall an, welches eine Wahrscheinlichkeit von 0,95 hat

- Gesucht: Vielfaches k von σ , so dass $[\mu - k\sigma; \mu + k\sigma] = 0,95$, $\mu = 200$, $\sigma = 75$
- Die Zufallsgröße weicht dann mit der Wahrscheinlichkeit 0,95 höchstens um $k * \sigma$ vom Erwartungswert μ ab

$$\Rightarrow P(X \leq k) - P(X \leq -k) = 0,95$$

$$\Leftrightarrow \Phi(k) - \Phi(-k) = 0,95$$

$$\Leftrightarrow \Phi(k) - (1 - \Phi(k)) = 0,95$$

$$\Leftrightarrow 2 * \Phi(k) - 1 = 0,95 \Leftrightarrow 2 * \Phi(k) = 1,95$$

$$\Leftrightarrow \Phi(k) = \frac{1,95}{2} = 0,975$$

$$\Rightarrow k = \Phi^{-1}(0,975) = 1,96$$

$$\Rightarrow [200 - 1,96 * 75; 200 + 1,96 * \sigma] = [53; 347]$$



von
nd
al-
tigter
Rolle.
Schaus
die,
scheinungen
Merkmalen

Statistik

Hohe Wahrscheinlichkeit garantiert keinen Erfolg

99,9% Erfolgswahrscheinlichkeit

- Eine Raumfahrtbehörde entwickelt eine Rakete.
- Die Ingenieure/-innen sagen: "Unsere Startsysteme haben eine Erfolgswahrscheinlichkeit von 99,9 %".
- Klinkt extrem zuverlässig!?
 - Aber:
 - Bei **1000 Starts** wird **im Schnitt 1 Fehlstart** erwartet
 - Falls Du an Bord genau der Rakete mit Fehlstart bist
 - Ist die Wahrscheinlichkeit eines Fehlers 100%!
- Selbst eine hohe Wahrscheinlichkeit garantiert keinen Erfolg im Einzelfall



<https://www.businessinsider.de/tech/spacex-rakete-explodiert-kurz-nach-dem-start-ein-erfolg-fuer-elon-musk/>



von
nd
al-
tigter
Rolle.
Schaus
die,
scheinungen
Merkmalen

Statistik

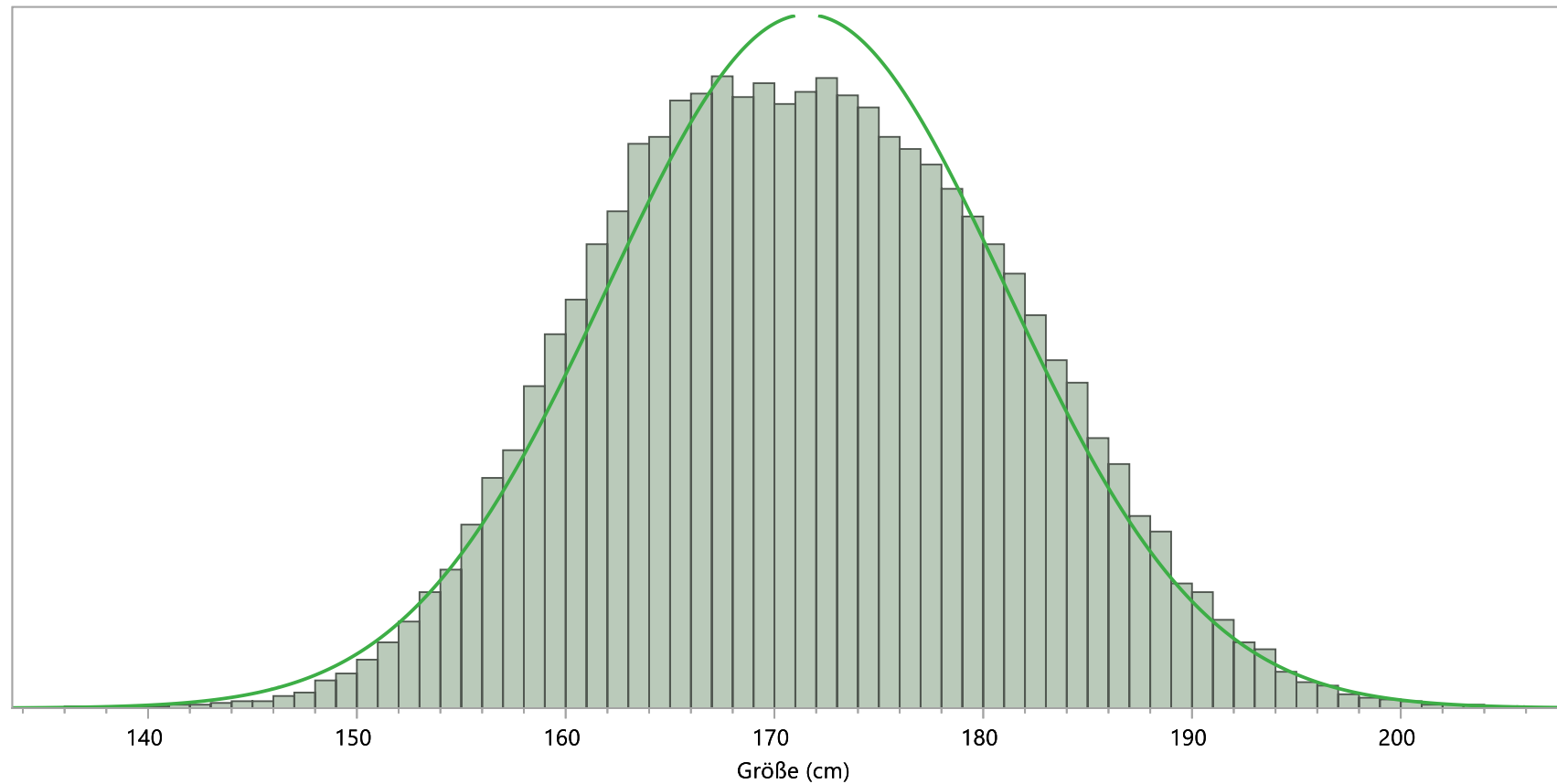
Unwahrscheinlich heißt nicht unmöglich

NAKO

Verteilung Größe Deutschland



- $Groesse \sim N(171,55; 9,55)$



Normalverteilung



- Warum gibt es keine Menschen > 240 cm, obwohl die Wahrscheinlichkeit dafür nicht Null ist?
 - Tatsächlich gibt es Menschen > 240 cm:
 - Robert Wadlow (1918 – 1940), *The Alton Giant*: 272 cm
 - Wahrscheinlichkeit für Menschen > 240 cm:
 - $P(X > 240) = 1 - N_{171,55;9,55}(240) \cong 1 - N_{0;1}(7,17)$
 $\cong 1 - 0,99999999999996182 \cong 0,00000000000003818$
 - Häufigkeit bei 8.000.000.000?
 - $0,00000000000003818 * 8.000.000.000 = 0,0031$
 - Häufigkeit bei allen bis heute jemals geborenen Menschen (~160.000.000.000)
 - 0,061
 - Die Biologie schlägt zu: Häufig genetische Veränderungen, die frühen Tod verursachen



von
nd
al-
tigter
Rolle.
Schaus
die,
scheinung
Merkmalen

Statistik

Eine Wahrscheinlichkeit von Null heißt nicht unmöglich

Dartscheiben-Paradoxon

- Dart-Spieler wirft Pfeil auf Dartscheibe
 - Trifft Dartscheibe immer
 - Jeder Punkt gleichwahrscheinlich: p
- Unendlich viele Punkte
 - Wenn unendlich genau gemessen
- Wahrscheinlichkeit genau einen Punkt zu treffen?
 - $P = \frac{1}{\text{Anzahl Punkte}}$
 - $\lim_{\text{Anzahl Punkte} \rightarrow \infty} \frac{1}{\text{Anzahl Punkte}} = 0$
⇒ Wie kann dann Scheibe getroffen werden?
- ⇒ $P > 0$
 - Aber: Wenn $P > 0 \Rightarrow \sum p \rightarrow \infty \neq 1$
- Ereignisraum unendlich
 - Wahrscheinlichkeit von 0 bedeutet nicht, dass Ereignis niemals eintritt – sondern, dass es „fasst sicher“ nicht eintritt
- Im Gegensatz: Ereignisraum endlich (z. B. Würfeln)
 - Wahrscheinlichkeit eine 7 zu Würfeln = 0 ⇒ Ereignis tritt niemals ein



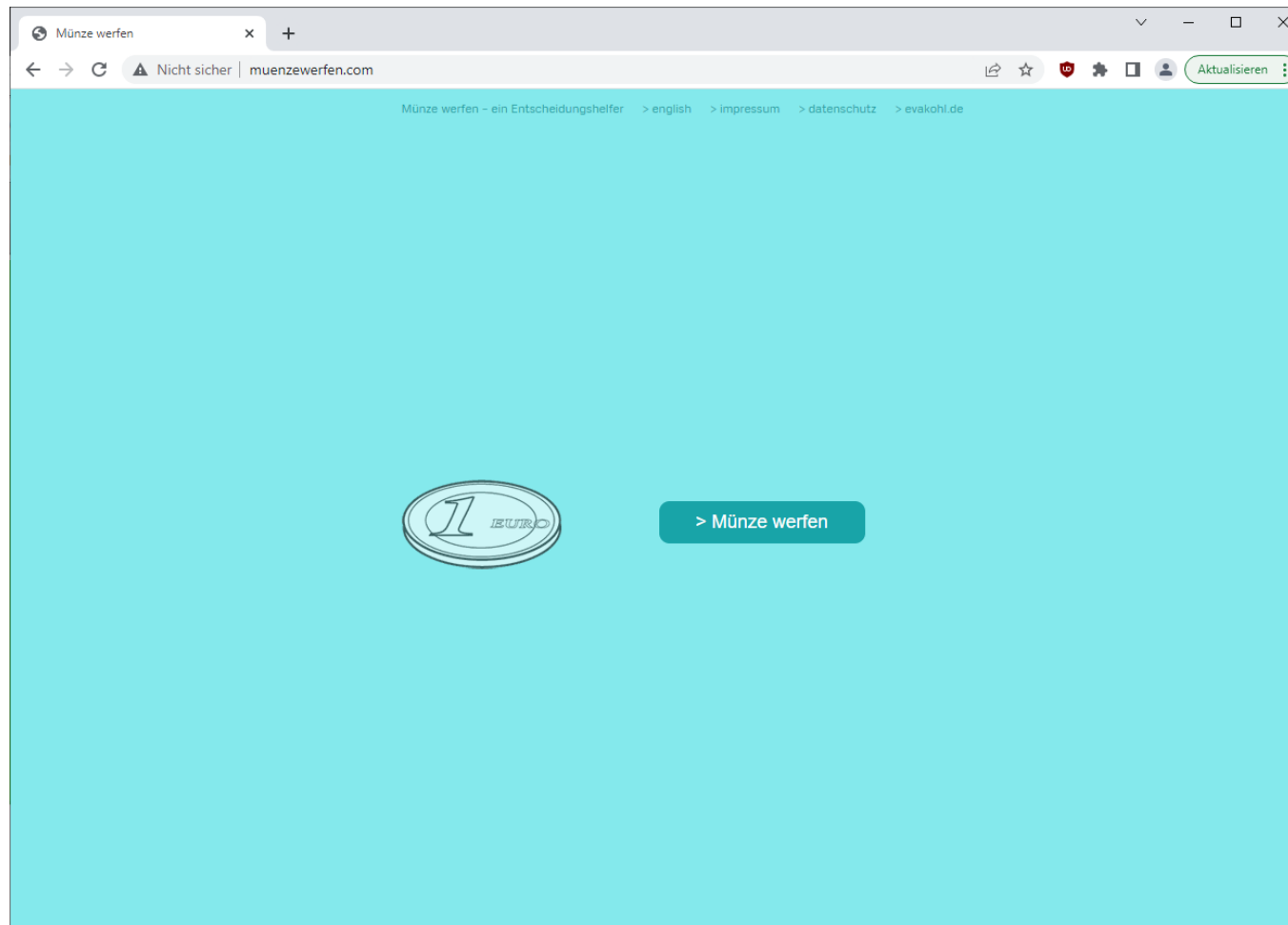
© WRAGG / GETTY IMAGES / ISTOCK (AUSSCHNITT)

<https://www.spektrum.de/kolumne/masstheorie-eine-wahrscheinlichkeit-von-null-heisst-nicht-unmoeglich/2092452>

Statistik

Theoretische Wahrscheinlichkeit

Münzwurf



<http://muenzewerfen.com/>

Münzwurf

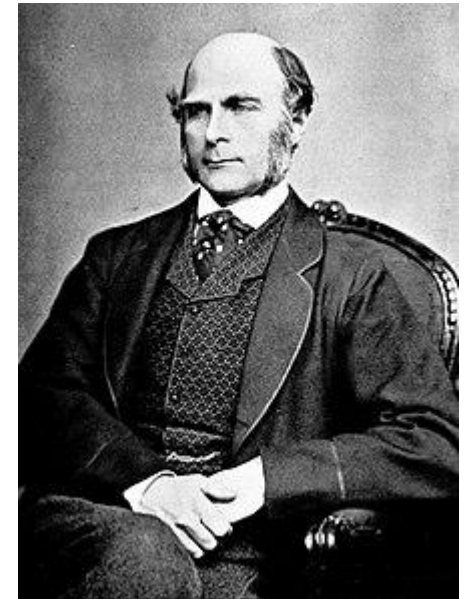
Gedankenexperiment

- Es wurde eine Münze 100 Mal geworfen
 - 100 Mal lag „Kopf“ oben
- Nun wird die Münze noch einmal geworfen
 - Wie hoch ist die Wahrscheinlichkeit, dass nun „Zahl“ oben liegt?
- Jeder Münzwurf ist von den vorherigen unabhängig!
 - Die Münze „weiß nicht“, was vorher oben lag!



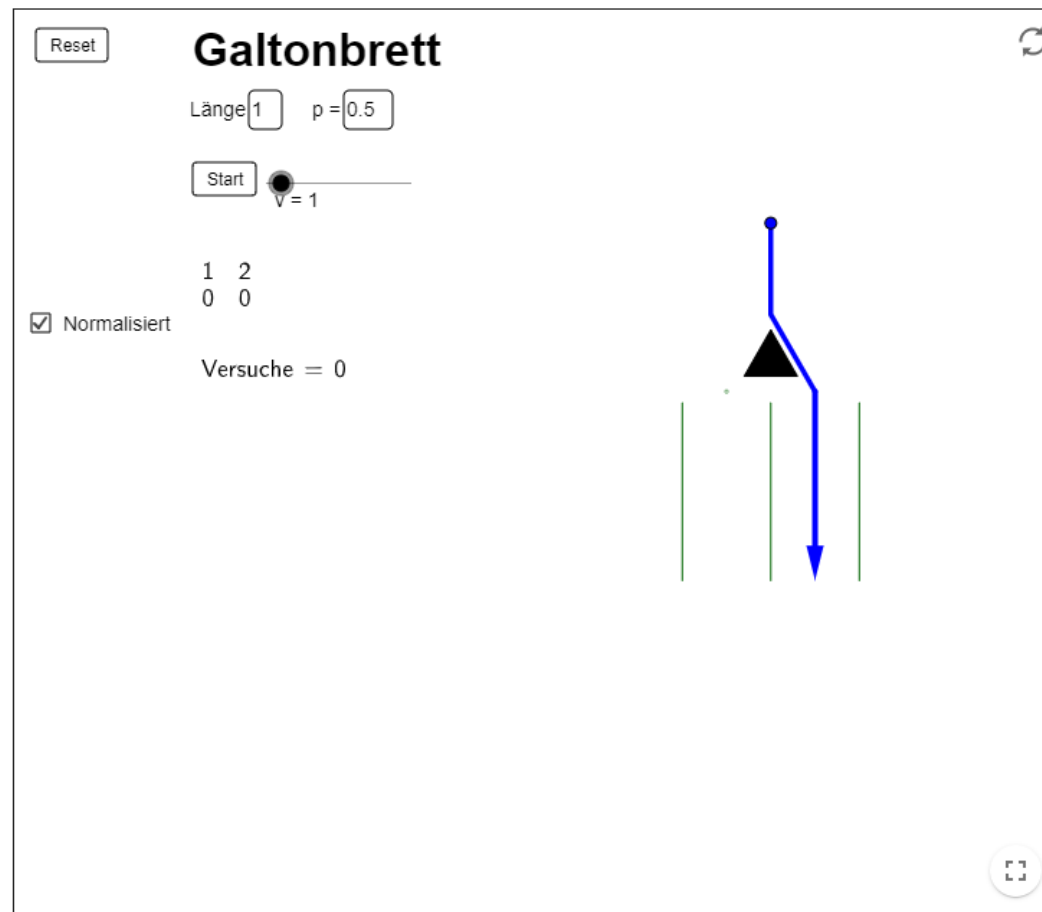
Galtonbrett

- Zufallsbrett, Galtonsches Nagelbrett
- Mechanisches Modell zur Veranschaulichung von Wahrscheinlichkeitsverteilungen
- Nach Sir Francis Galton (1822-1911), Naturforscher, Schriftsteller, Geograph, Meteorologe
 - Vater der Daktyloskopie
 - Mitbegründer der Differenzialpsychologie und der experimentellen Psychologie
 - Versuch die Dummheit der Masse zu beweisen (1906) schlug fehl: Vox populi („Stimme des Volkes“)
 - Cousin von Charles Darwin
 - Einer der Väter der Eugenik mit dem Ziel „der Verbesserung der menschlichen Rasse“



Galtonbrett

Simulation Münzwurf



<https://www.geogebra.org/m/tcR9U7zG>





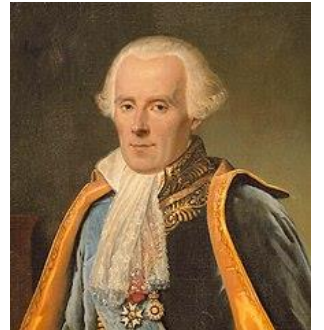
von
nd
al-
tigter
Rolle.
Schaus
die,
enerscheinung
Merkmalen

Statistik

Laplace Experiment

Laplace Experiment

- Zufallsexperiment
- Zufällige Auswahl
- Endliche Anzahl Ergebnisse
 - Jedes Ergebnis gleichwahrscheinlich (Fairer Würfel, Faire Münze)
 - Mögliches Ergebnis ist **Elementarereignis** ω (Würfel: Augenzahl, Münze: Kopf oder Zahl)
 - Menge aller möglichen Ergebnisse
 - **Ergebnismenge** Ω
 - Münze: $\Omega = \{Kopf, Zahl\}$
 - Würfel: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- Pierre-Simon Laplace (1749 – 1827)
Mathematiker, Physiker, Astronom; Mitglied im
Komitee für Maße und Gewichte
- Wahrscheinlichkeit eines Elementarereignisses nach Laplace-Formel



Laplace Formel

$$p(A) = \frac{\text{Anzahl der guenstigen Ereignisse}}{\text{Anzahl aller moeglichen Ereignisse}} = \frac{|A|}{|\Omega|}$$

Summe der Wahrscheinlichkeiten = 1

Statistik

Gleichverteilung

Gleichverteilung

- Verteilung der Wahrscheinlichkeiten von Ereignissen wenn jedes Ereignis gleichwahrscheinlich
- Diskrete Wahrscheinlichkeitsverteilung mit Parameter n

- Wahrscheinlichkeitsfunktion

$$P(X = x) = f(x) = \begin{cases} \frac{1}{n} & \text{für } x = x_i; i \in \{1, \dots, n\} \\ 0 & \text{sonst} \end{cases}$$

- Verteilungsfunktion

$$P(X \leq x) = F(x) = \begin{cases} 0 & ; \text{falls } x < 1 \\ \frac{\lfloor x \rfloor}{n} & ; \text{falls } 1 \leq x < n \\ 1 & ; \text{falls } x \geq n \end{cases} \quad \lfloor \cdot \rfloor : \text{Abrundungsfunktion}$$

- Erwartungswert

$$E(x) = \frac{1}{n} * \sum_{i=1}^n x_i$$

- Varianz

$$Var(x) = \frac{1}{n} * \sum_{i=1}^n (x_i - E(x))^2$$

- Standardabweichung

$$\sigma(x) = \sqrt{Var(x)}$$

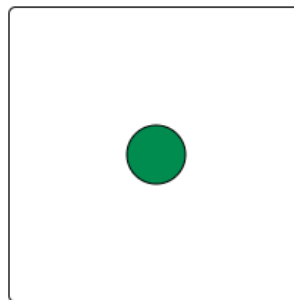
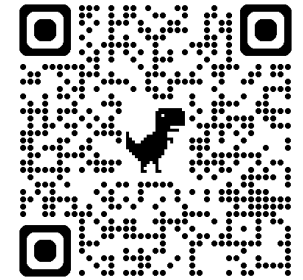
Würfeln

- Wahrscheinlichkeitsfunktion

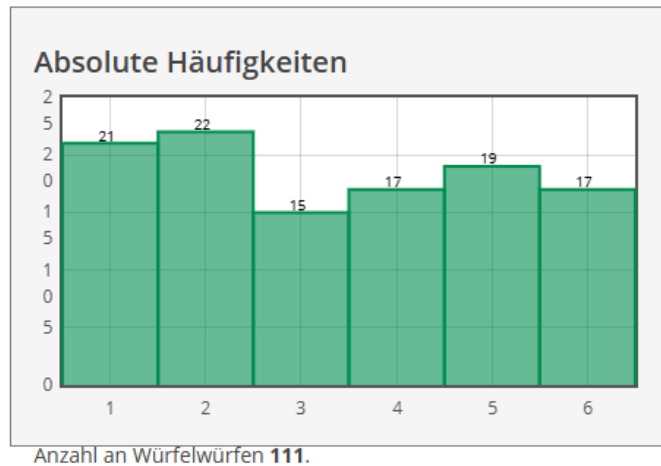
- $$P(X = x) = f(x) = \begin{cases} \frac{1}{6} & \text{für } x = x_i; i \in \{1, \dots, 6\} \\ 0 & \text{sonst} \end{cases}$$

- Erwartungswert

- $$E(x) = \frac{1}{6} * \sum_{i=1}^6 x_i = 3,5; i \in \{1, \dots, 6\}$$



Automatisch würfeln an/aus
1 x Manuell würfeln
Anzeige zurücksetzen



<https://www.mathematik.tu-clausthal.de/interaktiv/simulation/wuerfelsimulator>

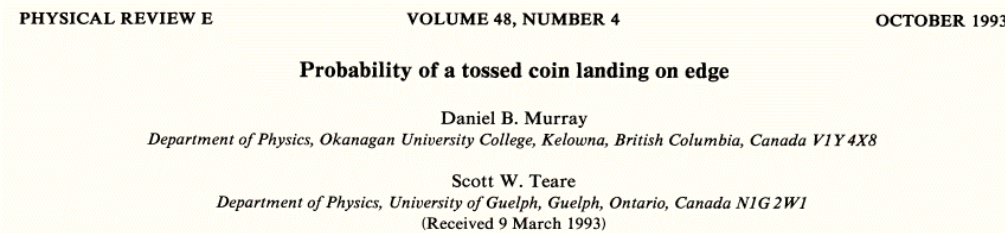
Statistik

Bernoulli Versuch

Münzwurf

Gedankenexperiment

- Wahrscheinlichkeit, dass die Münze auf der Kante steht
 - 1:6000



<https://journals.aps.org/pre/pdf/10.1103/PhysRevE.48.2547>

- Ist die Wahrscheinlichkeit wirklich 50%?
 - Physik sagt nein - es sind ~51%

SIAM REVIEW
Vol. 49, No. 2, pp. 211–235

© 2007 Society for Industrial and Applied Mathematics

Dynamical Bias in the Coin Toss*

Persi Diaconis[†]
Susan Holmes[‡]
Richard Montgomery[§]

<https://epubs.siam.org/doi/abs/10.1137/S0036144504446436>



- Was sagt die Empirie?
 - Scientists destroy illusion that coin toss flips are 50-50
 - ~51%!

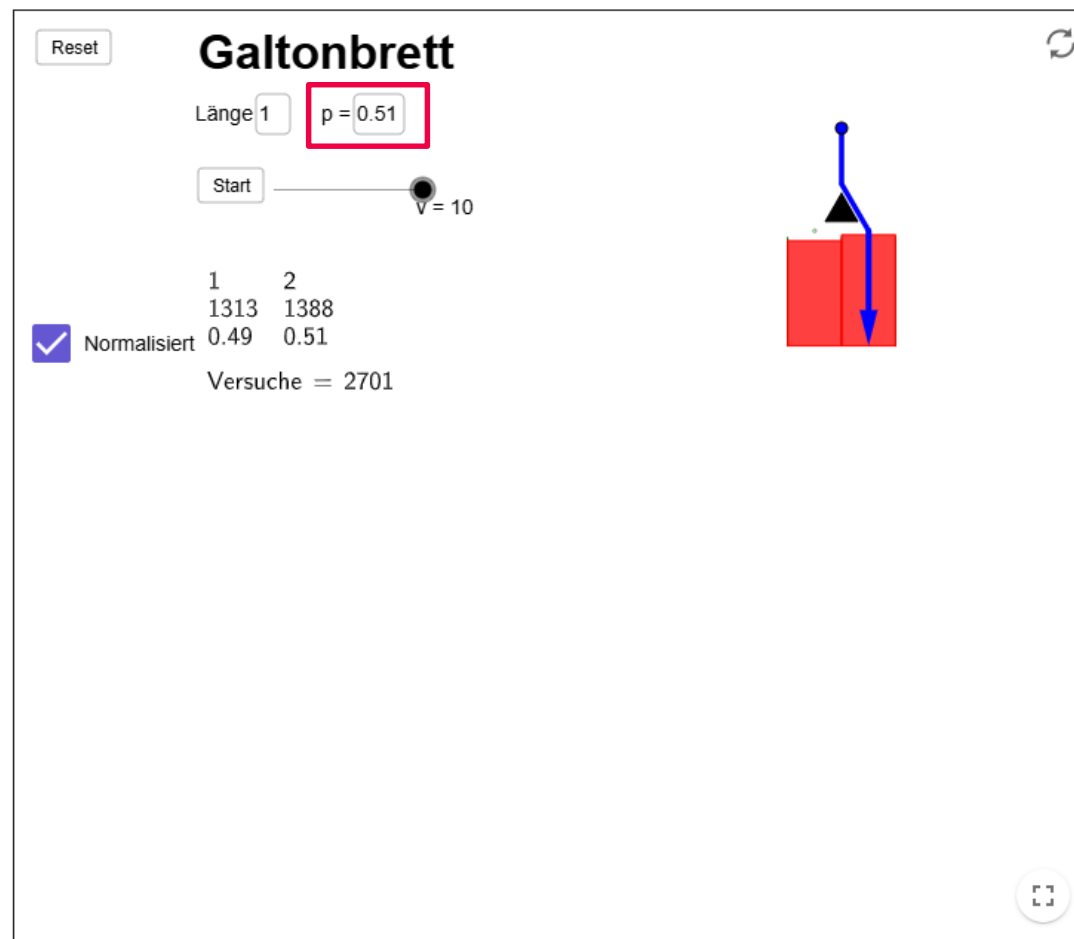
FAIR COINS TEND TO LAND ON THE SAME SIDE THEY STARTED:
EVIDENCE FROM 350,757 FLIPS

František Bartoš^{1,*}, Alexandra Sarafoglou¹, Henrik R. Godmann¹, Amir Sahrani¹, David Klein Leunk²,
Pierre Y. Gu², David Voss², Kaleem Ullah², Malte J. Zoubek¹, Franziska Nippold,
Frederik Aust¹, Felipe F. Vieira¹, Chris-Gabriel Islami^{2,3}, Anton J. Zoubek², Sara Shabani⁵,
Jonas Petter¹, Ingeborg B. Roos⁶, Adam Finnemann^{1,10}, Aaron B. Loh^{2,12}, Madlen F. Hoffstadt¹,
Jason Nak⁷, Jill de Ron², Koen Derks¹², Karoline Huth^{2,13}, Sjoerd Terpstra¹⁴,
Thomas Bastelica^{15,16}, Magda Matetovici^{2,17}, Vincent L. Ott², Andreea S. Zetea², Katharina Karnbach²,
Michelle C. Donzallaz¹, Arne John², Roy M. Moore², Franziska Assion^{1,8}, Riet van Bork²,
Theresa E. Leidinger², Xiaochang Zhao², Adrian Karami Motaghi², Ting Pan^{1,9}, Hannah Armstrong²,
Tianqi Peng², Mara Bialas²⁰, Joyce Y.-C. Pang², Bohan Fu², Shujun Yang²,
Xiaoyi Lin², Dana Steifter², Miklos Bognar²¹, Balazs Aczel²², and Eric-Jan Wagenmakers¹

<https://arxiv.org/abs/2310.04153>

Galtonbrett

Simulation Münzwurf – „Realistische“ Münze



<https://www.geogebra.org/m/tcR9U7zG>



Erfolg und Misserfolg

- > Gummibärchen
 - 11 x Rot
 - 27 x Grün
- > Wahrscheinlichkeit, blind ein rotes Gummibärchen zu ziehen?

$$- p(\text{rot}) = ?$$

$$> p(\text{rot}) = \frac{n(\text{rot})}{n(\text{rot}) + n(\text{grün})} = \frac{n(\text{rot})}{N}$$

$$> p(\text{rot}) = \frac{11}{11+27} = \frac{11}{38} = 0,29$$



Erfolg und Misserfolg

- > Gummibärchen
 - 11 x Rot
 - 27 x Grün

- > Rotes Gummibärchen gezogen
 - Erfolg!
 - $p(Rot) = \frac{11}{38} = 0,29$

- > Grünes Gummibärchen gezogen
 - Misserfolg!
 - $p(Grün) = \frac{27}{38} = 0,71$
 $= 1 - p(Rot)$
 - $q = 1 - p$



Erfolg und Misserfolg

Bernoulli-Versuch

> Zufallsversuch mit genau zwei möglichen Ergebnissen (z. B. Erfolg und Misserfolg)

- Wahrscheinlichkeit für Erfolg

$$p$$

- Wahrscheinlichkeit für Misserfolg (Gegenwahrscheinlichkeit)

$$q = 1 - p$$

> Benannt nach Jakob I Bernoulli (1654/55 – 1705),
Mathematiker und Physiker

> Beispiele: Münzwurf, Würfeln einer Sechs, ...

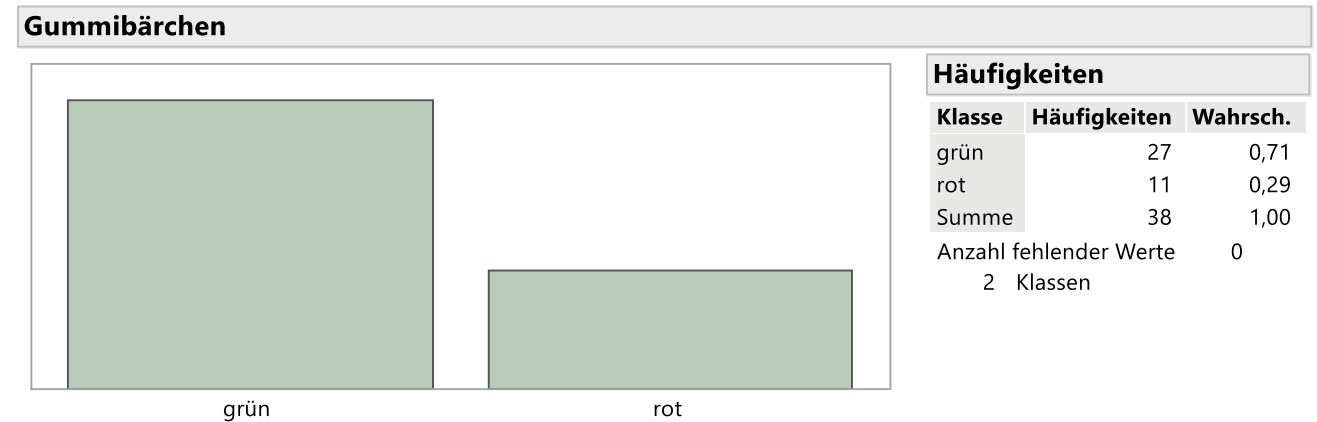


Statistik

Bernoulli-Verteilung

Bernoulli-Verteilung

- > Beschreibung zufälliger Ereignisse mit nur zwei möglichen Ausgängen
- > Diskrete Verteilung
 - Null-Eins-, Boole-Verteilung
- > Beispiele
 - Werfen einer Münze
$$p = \frac{1}{2}, q = 1 - p = \frac{1}{2}$$
 - Würfeln einer Sechs
$$p = \frac{1}{6}, q = \frac{5}{6}$$
 - Gummibärchen (Rot = Erfolg, Grün = Misserfolg)
$$p = \frac{11}{38}, q = \frac{27}{38}$$



Bernoulli-Verteilung

> Diskrete Verteilung

- Beschreibung zufälliger Ereignisse mit nur zwei möglichen Ausgängen
- Null-Eins-, Boole-Verteilung
- $X \sim Ber(p); x \in \{0; 1\}$

> Wahrscheinlichkeitsfunktion

$$f(x) = P(X = x) = \begin{cases} 1 - p, & x = 0 & (\text{Misserfolg}) \\ p, & x = 1 & (\text{Erfolg}) \\ 0, & \text{sonst} \end{cases}$$

> Verteilungsfunktion

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 0 \\ 1 - p, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

> Erwartungswert

$$E(x) = p$$

(Erfolgswahrscheinlichkeit)

> Varianz

$$Var(x) = p * (1 - p) = pq$$

> Standardabweichung

$$\sigma_x = \sqrt{pq}$$

Bestimmung von π mit Bernoulli-Verteilung



> Berechnung mittels Flächenformel

- Flächeninhalt Kreis: $A_{Kreis} = \pi r^2$
- Flächeninhalt Quadrat: $A_{Quadrat} = (2r)^2 = 4r^2$
- Verhältnis Kreis- zu Quadratfläche:

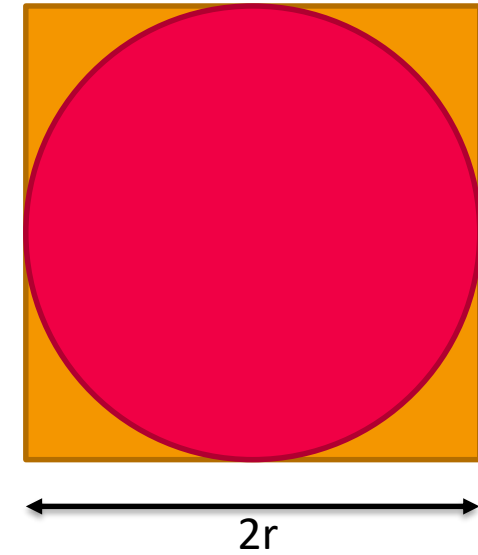
$$\Rightarrow \pi = 4 \frac{A_{Kreis}}{A_{Quadrat}} \quad \frac{A_{Kreis}}{A_{Quadrat}} = \frac{\pi r^2}{4r^2} = \frac{\pi}{4}$$

> Statistische Bestimmung (Monte-Carlo-Algorithmus)

- Punkte „regnen“ zufällig auf Quadrat
- Jeder Punkt gleichwahrscheinlich

> Bestimmung Punkte innerhalb (Erfolg) und außerhalb (Misserfolg) Kreis: $d \leq \sqrt{x^2 + y^2}$

- Anteil Punkte innerhalb Kreis approximiert $\frac{\pi}{4} \Rightarrow$ Bernoulli-Verteilung
- Genauigkeit steigt mit der Anzahl der Punkte



In Quadrat eingeschriebener Kreis

Statistik

Binomialverteilung

Binomialverteilung

- > Gummibärchen
 - 11 x Rot
 - 27 x Grün
- > Wahrscheinlichkeit zufällig 5-mal ($k = 5$) bei 8 Versuchen ($n = 8$) mit Zurücklegen ein rotes Gummibärchen zu ziehen?
 - Pech: 0-mal
 - Glück: 5-mal (k -mal)
 - $0 \leq k \leq 5$
- > $x_i = \begin{cases} 0 & \text{falls Rot bei } i - \text{tem Ziehen} \\ 1 & \text{falls nicht Rot bei } i - \text{tem Ziehen} \end{cases}$
- > Anzahl positiver Versuchsausgänge bei n -facher Wiederholung mit konstanter Wahrscheinlichkeit

$$X = \sum_{i=1}^n x_i$$



Binomialverteilung



> Beschreibt Verteilung der Anzahl der Erfolge bei einer Serie von gleichartigen und **unabhängigen** Versuchen mit jeweils genau zwei möglichen Ergebnissen (z. B. Erfolg/Misserfolg) $\Rightarrow$ Bernoulli-Versuch!

> Wiederholte Bernoulli-Versuche
– Bernoulli-Kette

> Wahrscheinlichkeitsfunktion

> Diskrete Wahrscheinlichkeitsverteilung

$$P(X = k) = f(k) = \begin{cases} \binom{n}{k} * p^k * (1 - p)^{(n-k)} & \text{falls } k \in \{0, 1, \dots, n\} \\ 0 & \text{sonst} \end{cases}$$

> Drei Parameter

- k: Länge der Serie
- n: Anzahl der Versuche
- p: Wahrscheinlichkeit des Ereignisses

$$\binom{n}{k} = \frac{n!}{k! * (n-k)!} \quad \text{Binomialkoeffizient (n über k)}$$

> Wahrscheinlichkeitsverteilung (kumulierte Wahrscheinlichkeit)

$$P(X \leq k) = F(k) = \sum_{i=0}^k \binom{n}{i} * p^i * (1 - p)^{(n-i)}$$

Binomialverteilung

- > Gummibärchen
 - 11 x Rot
 - 27 x Grün
- > Wahrscheinlichkeit zufällig 5-mal ($k = 5$) bei 8 Versuchen ($n = 8$) mit Zurücklegen ein rotes Gummibärchen zu ziehen?

$$> f(5) = \binom{8}{5} * \left(\frac{11}{38}\right)^5 * \left(1 - \frac{11}{38}\right)^{(8-5)}$$

$$> f(5) = \frac{8!}{5! * (8-5)!} * \left(\frac{11}{38}\right)^5 * \left(1 - \frac{11}{38}\right)^{(8-5)}$$

$$> f(5) = \frac{40320}{120 * 6} * 0,002 * 0,3587$$

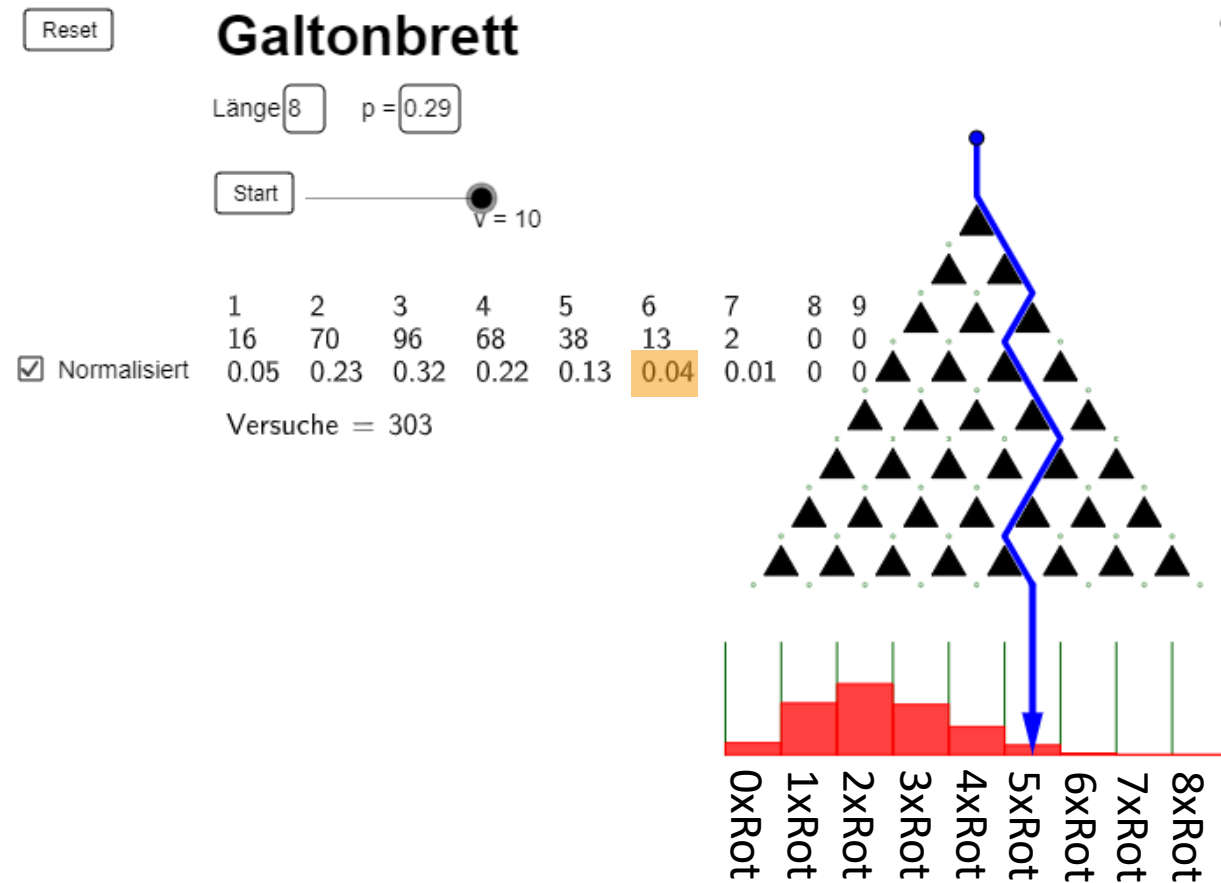
$$> f(5) = 0,0402$$

$$> B\left(5 \mid \frac{11}{38}, 8\right) = B\left(5, \frac{11}{38}, 8\right) = B_{\frac{11}{38}, 8}(5)$$



Galtonbrett

Simulation Gummibärchensituation



$$p(Rot) = 0,29$$

$$p(Grün) = 0,71 = 1 - p(Rot)$$

$$n = 8$$

$$k = 5$$

$$P(X = 5) = 0,0402$$

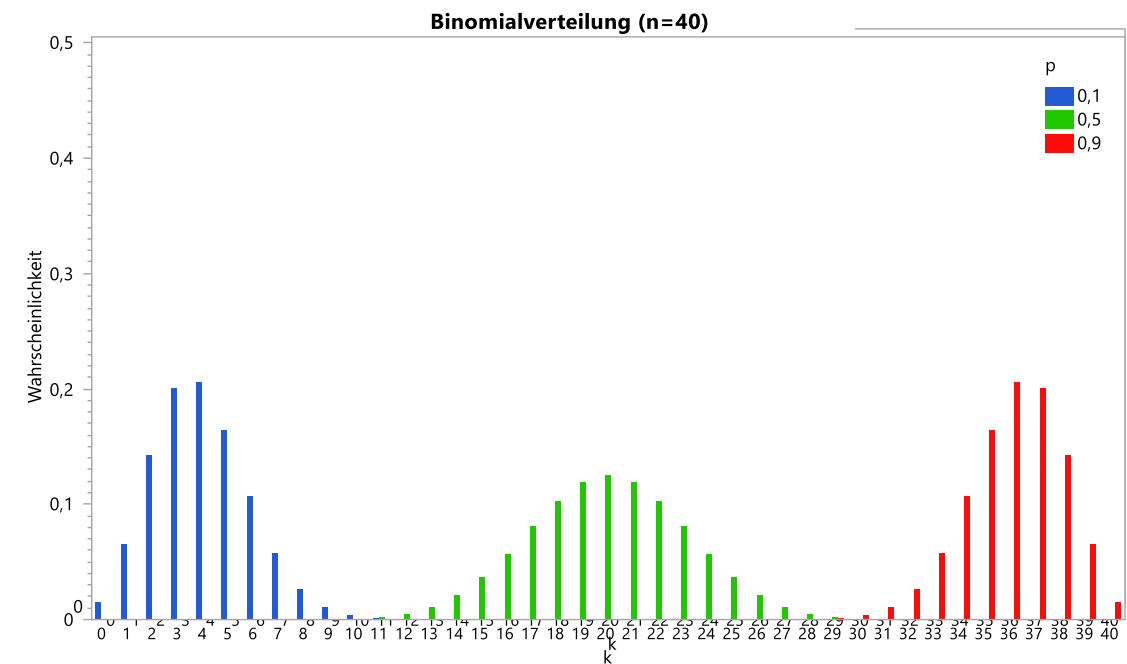
<https://www.geogebra.org/m/tcR9U7zG>

Binomialverteilung

Form und Symmetrien – Einfluss von p



- > Maximum liegt bei Erwartungswert $\mu = n * p$
 - Je größer p , desto größer μ , linksschief
 - Je kleiner p , desto kleiner μ , rechtsschief
 - Symmetrisch, wenn $p = 0,5$
- > Wenn $p = 0,5$
 - Erwartungswert „in der Mitte“: $n/2$
 - Dann: Binomialverteilung symmetrisch um Erwartungswert
 - $B(k|0,5; n) = B(n - k|0,5; n)$



Galtonbrett

Simulation Bernoulli-Kette / Binomialverteilung ($p = 0,5$)



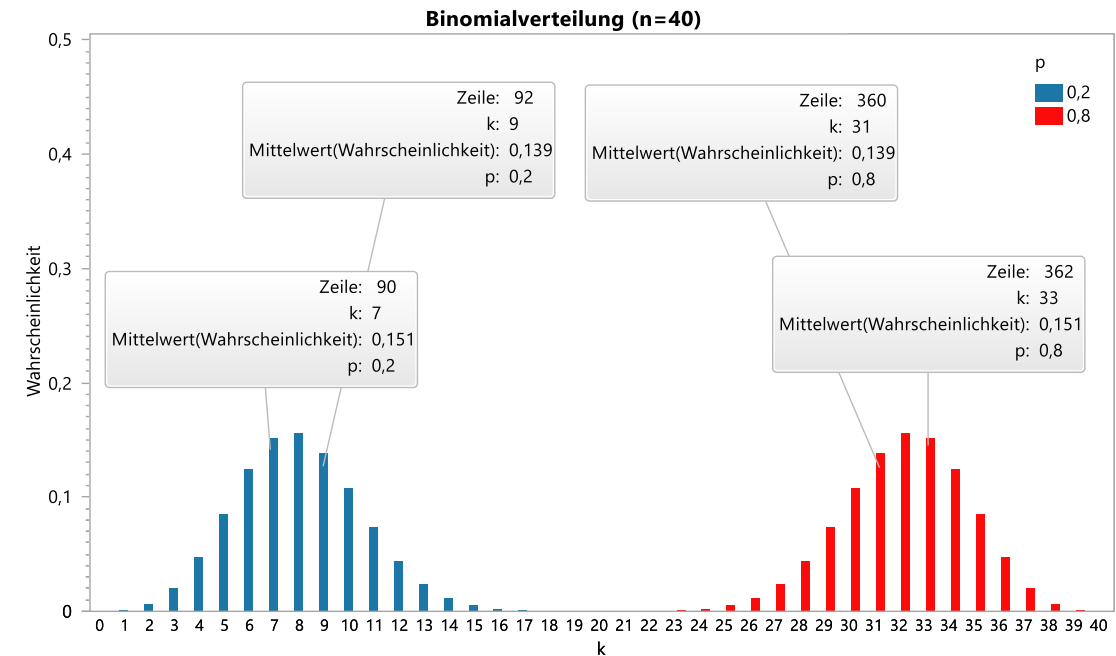
<https://www.youtube.com/watch?v=4HpvBZnHOVI>

Binomialverteilung

Form und Symmetrien – Symmetrie um $n/2$

> Bei gleichen n , Symmetrie um $n/2$

- $B(7|0,2; 40) \approx 0,1513$
- $B(33|0,8; 40) \approx 0,1513$
- $B(33|0,8; 40) = B(40 - 7|1 - 0,2; 40)$
- $B(k|p; n) = B(n - k|1 - p; n)$



$$B(k|0,2; 40):$$

$$\mu = 40 * 0,2 = 8$$

$$B(k|0,8; 40):$$

$$\mu = 40 * 0,8 = 32$$

Binomialverteilung

Form und Symmetrien – Einfluss von n

> Wenn $n \rightarrow \infty$

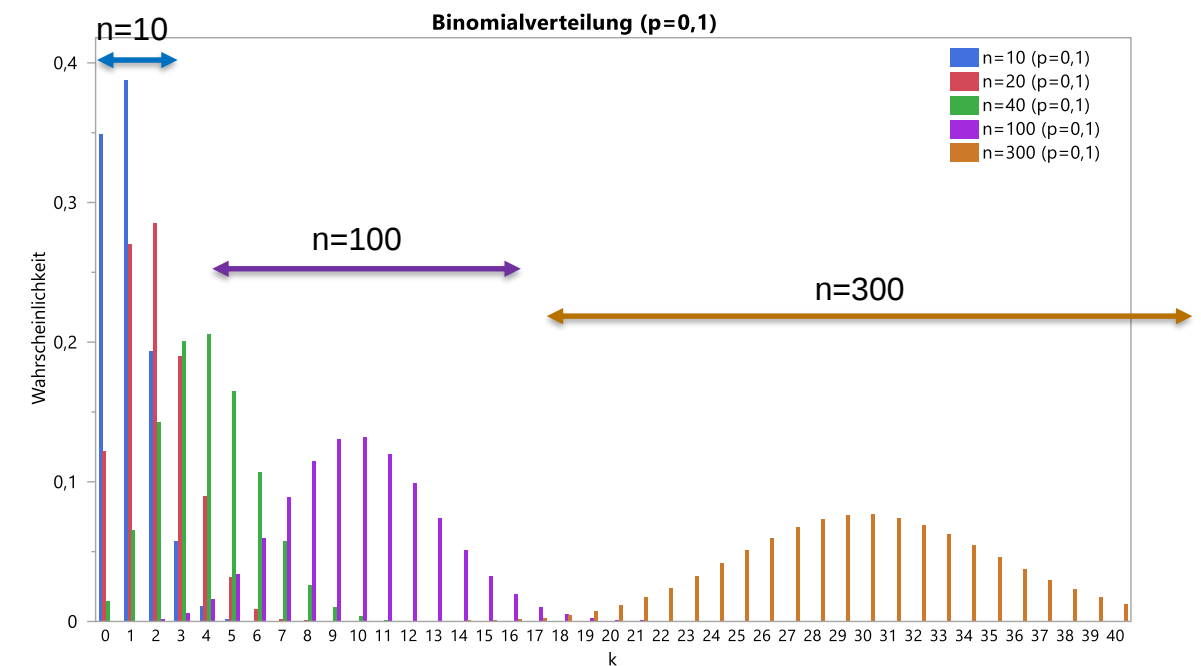
- „Breite“ wird größer
- Annähernd symmetrisch um Erwartungswert $\mu = n * p$
- Unabhängig von p

> Satz von Moivre-Laplace

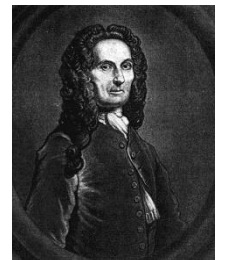
- $X \sim B(p; n)$ für $n \rightarrow \infty$ standardnormalverteilt
- $\lim_{n \rightarrow \infty} P\left(\frac{X - np}{\sqrt{np*(1-p)}} \leq x\right) = \Phi(x)$
- Theoretische Grundlage für Normal-Approximation

> Laplace Bedingung:

- $Var(X) = n * p * (1 - p) > 9 \Leftrightarrow SD(X) > 3$



Abraham de Moivre
(1667-1754), Mathematiker
Freund von Isaac Newton
Keine Professur, starb verarmt



Binomialverteilung



> Diskrete Wahrscheinlichkeitsverteilung

> Voraussetzung

- feste Anzahl Versuche
- **zwei** mögliche Ergebnisse
- **unabhängige** Versuche
- gleiche Wahrscheinlichkeiten

> Wahrscheinlichkeitsfunktion

$$B(k|p, n) = \binom{n}{k} * p^k * (1 - p)^{(n-k)}$$

> Verteilungsfunktion

$$F_X(x) = P(X \leq x) = \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} * p^k * (1 - p)^{(n-k)}$$

> Erwartungswert

$$E(X) = \sum_{k=0}^n k * \binom{n}{k} * p^k * (1 - p)^{n-k} \\ = n * p$$

> Varianz

$$Var(X) = \sum_{k=0}^n k^2 * P(X = k) - (n * p)^2 \\ = n * p * (1 - p) = n * p * q$$

> Standardabweichung

$$SD(X) = \sqrt{Var(X)} = \sqrt{n * p * (1 - p)}$$

> Variationskoeffizient

$$VarK(X) = \sqrt{\frac{1-p}{np}}$$

> Schiefe

$$v(X) = \sqrt{\frac{1-2p}{np(1-p)}}$$

Binomialverteilung

Anwendungsbeispiele: Roulette

- 37 Zahlen, 0 – 36
- Alle Ereignisse gleichwahrscheinlich

⇒ Gleichverteilung

- Laplace-Formel: $P(A) = \frac{|A|}{|\Omega|} = \frac{1}{n} = \frac{1}{37}$
- Erwartungswert: $E(x) = \frac{1}{n} \sum_{i=1}^n x_i = 18$
- Varianz: $Var(x) = \frac{1}{n} \sum_{i=1}^n (x_i - E(x))^2 = 114$
- SD: $SD(x) = \sqrt{Var(x)} = 10,68$



Ralf Roletschek, CC BY-SA 3.0

		0		
	1	2	3	
PASSE	4	5	6	MANQUE
	7	8	9	
	10	11	12	
	13	14	15	
PAIR	16	17	18	IMPAIR
	19	20	21	
	22	23	24	
	25	26	27	
◆	28	29	30	◆
	31	32	33	
	34	35	36	
12 ^P 12 ^M 12 ^D				12 ^D 12 ^M 12 ^P

Binomialverteilung

Anwendungsbeispiele: Roulette

- Einfache Chancen (Rouge/Noir, Pair/Impair, Manque/Passe)
 - Zwei mögliche Ergebnisse (Erfolg/Misserfolg)

⇒ Bernoulli-Versuch

- Erfolgswahrscheinlichkeit: $P(x) = \sum_1^k \frac{1}{n} = \frac{18}{37}$
- Misserfolgsw.: $P(\bar{x}) = 1 - P(x) = \frac{19}{37}$
- Erwartungswert: $E(x) = P(x) = \frac{18}{37}$
- Varianz: $\text{Var}(x) = P(x) * (1 - P(x)) = \frac{342}{1369}$
- SD: $\text{SD}(x) = \sqrt{\text{Var}(x)} = 0,4998$



Ralf Roletschek, CC BY-SA 3.0

Binomialverteilung

Anwendungsbeispiele: Roulette

- > 37 Nummern (0, 1-36)
- > Einfache Chancen: Rot, Schwarz, Gerade, Ungerade, Niedrig, Hoch
 - Jeweils 18 Nummern
- > Sie bekommen als Willkommensgeschenk vom Kasino 100 € um damit genau 10-mal 10 € zu setzen
- > Sie setzen bei jedem Durchgang auf Rot
- > Wie hoch ist die Chance, **genau** 5-mal zu gewinnen?



Wikipedia, Ralf Roletschek, CC BY-SA 3.0

Binomialverteilung

Anwendungsbeispiele: Roulette



> Zwei Möglichkeiten:

- Erfolg (Rot), Misserfolg (Nicht Rot)
- ⇒ $p(\text{Erfolg}) = p(\text{Rot}) = \frac{18}{37}$ (Erfolgswahrscheinlichkeit)
- ⇒ Bernoulli-Versuch

> 10 Durchgänge ($n = 10$)

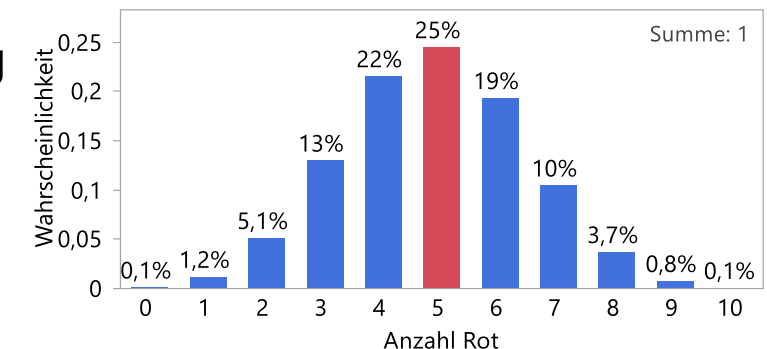
- **Durchgänge unabhängig**
 - **Gleiche Wahrscheinlichkeit bei jedem Durchgang**
- ⇒ Bernoulli-Kette
- ⇒ Binomialverteilung: $X \sim B(k, p, n)$

> Chance auf genau 5 Gewinne ($k = 5$)?

$$\begin{aligned} \Rightarrow P(X = 5) &= \binom{n}{k} * p^k * (1 - p)^{n-k} \Rightarrow f(5) = \binom{10}{5} * \left(\frac{18}{37}\right)^5 * \left(1 - \frac{18}{37}\right)^{10-5} \\ &= 0,2425 \end{aligned}$$



Wikipedia, Ralf Roletschek, CC BY-SA 3.0



Binomialverteilung

Anwendungsbeispiele: Roulette

- > 37 Nummern (0, 1-36)
- > Einfache Chancen: Rot, Schwarz, Gerade, Ungerade, Niedrig, Hoch
 - Jeweils 18 Nummern
- > Sie bekommen als Willkommensgeschenk vom Kasino 100 € um damit genau 10-mal 10 € zu setzen



Wikipedia, Ralf Roletschek, CC BY-SA 3.0

- > Was ist der Erwartungswert für Rot?
 - $E(X) = n * p$
 $\Rightarrow E(Rot) = 10 * \frac{18}{37} \approx 4,86$

- > Wie lautet die Standardabweichung?

$$\begin{aligned} - \quad SD(X) &= \sqrt{n * p * q} = \sqrt{n * p * (1 - p)} \\ \Rightarrow SD(Rot) &= \sqrt{10 * \frac{18}{37} * (1 - \frac{18}{37})} \\ &= \sqrt{10 * \frac{18}{37} * \frac{19}{37}} \\ &\approx 1,5086 \end{aligned}$$

Binomialverteilung

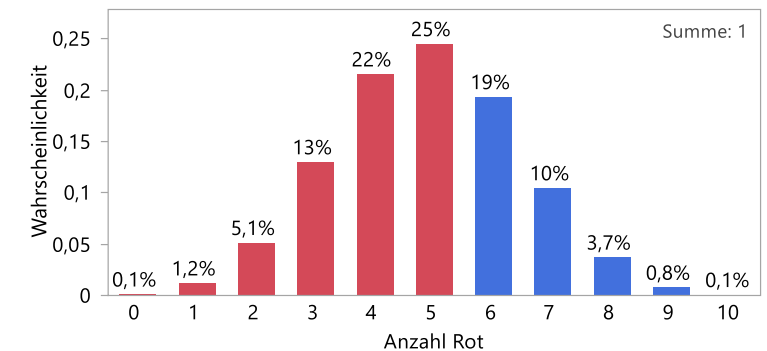
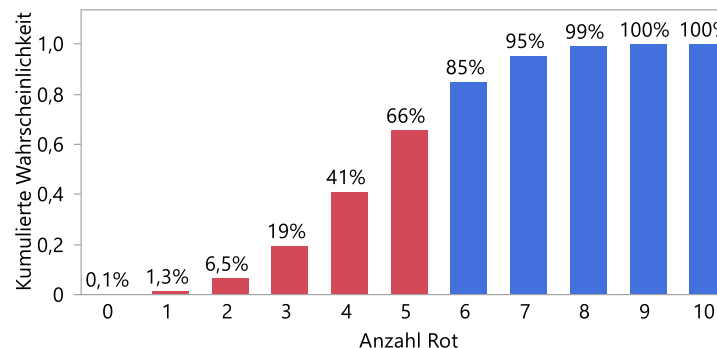
Anwendungsbeispiele: Roulette

- > Sie setzen bei jedem Durchgang auf Rot
- > Wie hoch ist die Chance, **bis zu** 5-mal zu gewinnen?
 - $\Rightarrow P(X \leq 5) = ?$
 - $\Rightarrow P(X \leq 5) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$
 - $= 0,0013 + 0,0121 + 0,0515 + 0,1301 + 0,2157 + 0,2452$
 - $= 0,6558$



Wikipedia, Ralf Roletschek, CC BY-SA 3.0

- > Verteilungsfunktion



Roulette

Martingalespiel

- Strategie im Glückspiel / Spielsystem
 - Bekannt seit 18. Jahrhundert
 - Benannt nach französischer Stadt Martigues
 - Einwohner/-innen galten als „etwas naiv“
 - Anwendung z. B. beim Roulette (Rot/Schwarz, Gerade/Ungerade, ...)
- Erhöhen des Einsatzes im Verlustfall, typischerweise Verdoppeln (*Doublieren*)
 - Spekulation mit Gesetz des Ausgleichs (*Equilibre*)
 - **Spielerfehlschluss** (*Gambler's fallacy*)
 - ⇒ Jede Runde ist unabhängiges Zufallsexperiment (unabhängig von vorhergehenden Runden)
 - Sinnlos, auf Ausgleich zu warten!

Roulette

Martingalespiel



- Spieler spielt Martingale auf einfache Chance: Rouge
 - Annahme: Zéro gilt als Noir (Verlust)
 - Starteinsatz 10 €, Verdoppelung bei jedem Verlieren
 - Maximaleinsatz: 20.480 € $\Rightarrow$ Maximal 12 Spiele (12x Verlieren)
- Wahrscheinlichkeit einzelnes Spiel zu verlieren
 - Bernoulli-Versuch: $p(\text{Verlieren}) = \frac{19}{37} \approx 0,514$
- Wahrscheinlichkeit 12 Spiele in Folge zu verlieren
 - Bernoulli-Kette $\Rightarrow$ Binomial-Verteilung $B(k|p, n)$
 - $p(X = 12) = B\left(12 \mid \frac{19}{37}, 12\right) = \binom{12}{12} * \frac{19^{12}}{37^{12}} * \left(1 - \frac{19}{37}\right)^{12-12} \approx 0,00034 \approx \mathbf{0,034\%}$
 JMP JSL: `Binomial Probability( 19 / 37, 12, 12 )`
- Wahrscheinlichkeit mindestens 1 Spiel von 12 zu gewinnen
 - $p(X \geq 1) = \sum_{i=1}^{12} B\left(i \mid \frac{18}{37}, 12\right) \approx 0,99966 \approx \mathbf{99,966\%}$
 JMP JSL: `Summation( i = 1, 12, Binomial Probability( 18 / 37, 12, i ) )`
 - Alternativ (Gegenwahrscheinlichkeit): $p(X \geq 1) = B\left(X \geq 1 \mid \frac{18}{37}, 12\right) = 1 - B\left(12 \mid \frac{19}{37}, 12\right)$

Roulette

Martingalespiel



- Verlust wird durch Verdoppelung im nächsten Spiel jeweils ausgeglichen (falls man gewinnt!)
- Wahrscheinlichkeit Martingale mit Gewinn zu beenden: 99,966%
- Scheinbar hervorragende Chancen, aber:
 - Gewinn bei glücklichem Abschluss: 10 €
 - Maximaler Verlust: 40.950 €!
- Insgesamt:
 - Erfolgswahrscheinlichkeit (min. 1 Gewinn): 0,99966
 - Misserfolgswahrscheinlichkeit (kein Gewinn): 0,00034
 - Erwartungswert Gesamtgewinn:
$$E(\text{Gesamt}) = 0,99966 * 10\text{€} - 0,00034 * 40.950\text{€} \approx -3,77\text{€}$$
$$\Rightarrow \text{Für Spieler/-in negativ!}$$
$$\Rightarrow \approx 3,77\% \text{ des Starteinsatzes}$$

Roulette

Martingalespiel

Aus: Giacomo Casanova (1725 - 1798), „Geschichte meines Lebens“

Vorher bat M. M. mich noch, in ihr Kasino zu gehen, dort Geld zu holen und mit ihr auf Halbpant zu spielen. Ich tat es und nahm alles Geld, das ich fand. Damit spielte ich die Martingale, indem ich stets die Sätze verdoppelte; ich gewann bis zum Ende des Karnevals täglich. Ich hatte das Glück, niemals die sechste Karte zu verlieren; und wenn mir das passiert wäre, so hätte ich kein Spielkapital mehr gehabt; denn dieser sechste Satz betrug zweitausend Zechinen. Ich freute mich, den Schatz meiner teuren Geliebten vermehrt zu haben.



Ich spielte immer noch meine Martingale, aber so unglücklich, dass ich bald keine Zechine mehr hatte. Da ich auf gemeinsame Rechnung mit M. M. spielte, musste ich ihr Rechenschaft über den Stand meiner Finanzen ablegen.

Auf ihr Drängen verkaufte ich nach und nach alle ihre Diamanten. Den Erlös verlor ich wieder; sie behielt für sich nur fünfhundert Zechinen für den Fall der Not zurück. Von Entführung war keine Rede mehr; denn wie hätten wir uns mittellos durch die Welt schlagen sollen?

Binomialverteilung

Anwendungsbeispiele: Anzahl benötigter Server

- > Ihre Firma betreibt eine Serverfarm für einen Internetdienst. Im Zuge einer Modernisierung sollen neue Server für den Betrieb angeschafft werden. Aus den Spezifikationen der neuen Server wissen Sie, dass auf einem Server 10 Dienstinstanzen parallel betrieben werden können. Aus früherer Erfahrung wissen Sie, dass von Ihren 2000 Kunden durchschnittlich 15 % gleichzeitig auf den Dienst zugreifen. Ziel ist es, dass alle Kunden ohne Verzögerung auf den Dienst zugreifen können, d.h. dass für bei jedem Zugriff mit großer Wahrscheinlichkeit eine Dienstinstanz frei ist.
 - a) Welche Verteilung ist zur Abschätzung der Wahrscheinlichkeit, dass Dienstinstanzen frei sind, geeignet?
 - b) Welche Anforderungen müssen dazu erfüllt sein?
 - c) Welcher Wert der Verteilung könnte für die zu im Mittel zu erwartenden Kundenzugriffe zur Abschätzung der Anzahl von benötigten Servern verwendet werden? Welche Anzahl von Servern ergibt sich daraus?
 - d) Mit welcher Wahrscheinlichkeit können mit dieser Anzahl von Servern alle Kundenzugriffe ohne Verzögerung bedient werden?
 - e) Um die Wahrscheinlichkeit zu erhöhen, soll die Anzahl der Server um eine Standardabweichung der Verteilung erhöht werden. Wie viele Server werden eingeplant?
 - f) Wie hoch ist damit die Wahrscheinlichkeit?
 - g) Wie viele Server werden benötigt, um die Wahrscheinlichkeit auf über 95 % zu erhöhen?

Binomialverteilung

Anwendungsbeispiele: Anzahl benötigter Server

- a) Binomialverteilung
- b) Feste Anzahl $n = 2000$, konstante Wahrscheinlichkeit $p = 0.15$, Versuche (Zugriffe) unabhängig, nur zwei Ergebnisse (verfügbar, nicht verfügbar)
- c) Erwartungswert $E(X) = n * p = 2000 * 0.15 = 300$ Dienstanstanzen $\triangleq$ 30 Server (10 Instanzen/Server)
- d) $P(X \leq 300) = F(300; 2000; 0.15) = \sum_{k=0}^{300} B(k|0.15, 2000) = \sum_{k=0}^{300} \binom{n}{k} * p^k * (1-p)^{n-k} \approx 0.52$
Die Wahrscheinlichkeit beträgt ca. 52 %
- e) $SD(X) = \sqrt{n * p * (1-p)} = \sqrt{2000 * 0.15 * (1-0.15)} \approx 15.97 \triangleq 1.6 \Rightarrow 2 \text{ Server (10 Instanzen/Server)}$
 $\Rightarrow 30 + 2 = 32 \text{ Server}$
- f) $P(X \leq 32 * 10) = F(320; 2000; 0.15) = \sum_{k=0}^{320} B(k|0.15, 2000) \approx 0.90$
Die Wahrscheinlichkeit beträgt ca. 90 %
- g) $P(X \leq 33 * 10) \approx 0.97 \Rightarrow 97 \% \text{ bei 33 Servern}$

Statistik

Hypergeometrische Verteilung

Hypergeometrische Verteilung

- Gummibärchen
 - 11 rot
 - 27 grün
- Wahrscheinlichkeit zufällig 5-mal ohne Zurücklegen bei 8 Versuchen ein rotes Gummibärchen zu ziehen?
 - Gesamtzahl Gummibärchen
 $N = 11 + 27 = 38$
 - Anzahl rote Gummibärchen (Erfolg)
 $M = 11$
 - Umfang Stichprobe
 $n = 8$
 - Anzahl angestrebte rote Gummibärchen
 $k = 5$



Hypergeometrische Verteilung

- Wahrscheinlichkeit zufällig 5-mal ohne Zurücklegen bei 8 Versuchen ein rotes Gummibärchen (GB) zu ziehen?
- $$P(k) = \frac{\text{Anzahl Kombinationen } k \text{ rote} * (n-k) \text{ grüne GB}}{\text{Anzahl Kombinationen } n \text{ GB beliebiger Farbe}}$$

$$= h(k|N, M, n)$$
- Anzahl Kombinationen k rote GB $\binom{M}{k}$
- Anzahl Kombinationen n-k grüne GB $\binom{N-M}{n-k}$
- Anzahl Kombinationen k rote und n-k grüne GB $\binom{M}{k} * \binom{N-M}{n-k}$
- Anzahl beliebige Kombinationen $\binom{N}{n}$
- $$P(k) = h(k|N, M, n) = \frac{\binom{M}{k} * \binom{N-M}{n-k}}{\binom{N}{n}}$$



Hypergeometrische Verteilung



- Wahrscheinlichkeit zufällig 5-mal **ohne Zurücklegen** bei 8 Versuchen ein rotes Gummibärchen (GB) zu ziehen?
- $$P(k) = \frac{\text{Anzahl Kombinationen } 5 \text{ rote} * (8-5) \text{ grüne GB}}{\text{Anzahl Kombinationen } 8 \text{ GB beliebiger Farbe}}$$

$$= h(5|38,11,8)$$
- Anzahl Kombinationen k rote GB $\binom{11}{5} = 462$
- Anzahl Kombinationen n-k grüne GB $\binom{38-11}{8-5} = 2.925$
- Anzahl Kombinationen k rote und n-k grüne GB

$$\binom{11}{5} * \binom{38-11}{8-5} = 1.351.350$$
- Anzahl beliebige Kombinationen $\binom{38}{8} = 48.903.492$
- $$P(5) = \frac{\binom{11}{5} * \binom{38-11}{8-5}}{\binom{38}{8}} = \frac{462 * 2.925}{48.903.492} = \frac{1.351.350}{48.903.492} \cong 0,028$$



Hypergeometrische Verteilung

- Verteilung Anzahl Erfolge bei Serie von gleichartigen und unabhängigen Versuchen mit jeweils genau zwei möglichen Ergebnissen (Erfolg/Misserfolg) ohne Zurücklegen

- Diskrete Wahrscheinlichkeitsverteilung

- Parameter

- k: Anzahl der zu ziehenden günstigen Elemente
- n: Anzahl gezogener Elemente
- M: Anzahl der günstigen Elemente
- N: Gesamtzahl Elemente

- Wahrscheinlichkeitsfunktion

$$P(X = k) = h(k|N, M, n) = \frac{\binom{M}{k} * \binom{N-M}{n-k}}{\binom{N}{n}}$$

- Verteilungsfunktion

$$H(X \leq k) = H(k|N, M, n) = \sum_{y=0}^k h(y|N, M, n) = \sum_{y=0}^k \frac{\binom{M}{y} * \binom{N-M}{n-y}}{\binom{N}{n}}$$

- Modus $\left\lfloor \frac{(n+1)*(M+1)}{N+2} \right\rfloor$ | $\lfloor \cdot \rfloor$: Gaussklammer, Abrunden

- Erwartungswert $E(X) = \sum_{k=0}^n k * \frac{\binom{M}{k} * \binom{N-M}{n-k}}{\binom{N}{n}} = n * \frac{M}{N}$

- Varianz $Var(X) = \sum_{k=0}^n k^2 * \frac{\binom{M}{k} * \binom{N-M}{n-k}}{\binom{N}{n}} - \left(n * \frac{M}{N} \right)$
 $= n * \frac{M}{N} * \left(1 - \frac{M}{N} \right) * \frac{N-n}{N-1}$

- Standardabweichung $SD(X) = \sqrt{Var(X)}$

Triggerwarnung

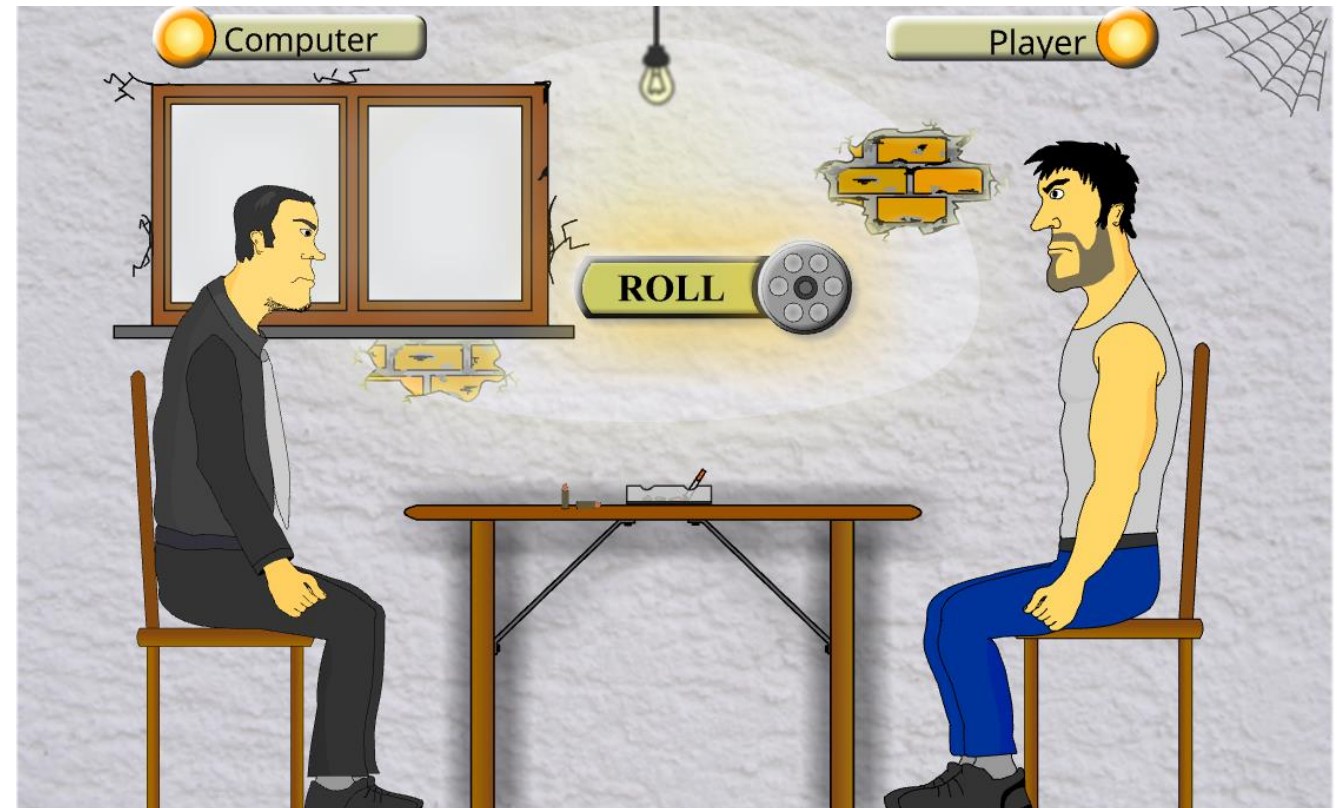
Auf den folgenden Folien wird ein äußerst dummes Spiel zur Demonstration von Wahrscheinlichkeiten in sarkastischer Form comicartig dargestellt

Bitte versuchen Sie das nicht Zuhause!

Russisches Roulette



- Gewinnwahrscheinlichkeit bei Russischem Roulett
- Hypergeometrische Wahrscheinlichkeit
 - Nur richtig, falls ohne Beachtung der Reihenfolge!
- Social and Racial Correlates of Russian Roulette
- Wattestäbchen in den Kopf geschossen: Männer spielten Russisch Roulette



<https://www.silvergames.com/de/russian-roulette-game>

Statistik

Poisson-Verteilung

Poisson-Verteilung

- > 38 Gummibärchen
- > Durchschnittlicher Verzehr
 - 5 Gummibärchen pro Stunde
- > Wahrscheinlichkeit zwischen 21 und 22 Uhr
10 Gummibärchen zu essen?
- > Durchschnittlicher Verzehr
$$\lambda = \frac{5}{1} = 5$$
- > Wahrscheinlichkeit für Verzehr von 10 Gummibärchen
in einer Stunde?



Poisson-Verteilung

Verteilung seltener Ereignisse

- > Verteilung der Anzahl von unabhängigen Ereignissen mit **konstanter mittlerer Rate** in einem **festen Zeitintervall** oder **räumlichen Gebiet** eintreten
- > Diskrete Wahrscheinlichkeitsverteilung
- > Benannt nach Siméon Denis Poisson (1781 – 1840, Mathematiker und Physiker)
 - Publikation im letzten Studienjahr von hoher Qualität, Abschluss ohne Abschlussprüfung, Anstellung als Repetitor
 - Theoretiker: „Poisson ... war ... zufrieden, mit den Wechselfällen physikalischer Forschung völlig unvertraut zu sein. Es ist unwahrscheinlich, dass er jemals eine experimentelle Messung versuchte, noch per Hand eine Versuchsanordnung entwarf.“
J. R. Hofman: Poisson's 1812 Electricity Memoir. In: André-Marie Ampère (Cambridge 1995)
 - Poisson: „Das Leben ist nur zu zwei Dingen gut: um Mathematik zu machen und sie zu lehren.“



Poisson-Verteilung

> Anwendung für z. B. „seltene Ereignisse“

- Ladislaus von Bortkewitsch (1868 – 1931, Statistiker und Ökonom):
Beschreibung der Anzahl der Todesfälle durch Hufschlag in einzelnen Kavallerie-Einheiten der preußischen Armee pro Jahr durch Poisson-Verteilung
- Druckfehler je Buchseite
- Anzahl radioaktiver Zerfälle je Zeitintervall (Zeitintervall $\ll$ Halbwertszeit)
- Blitzeinschläge pro Fläche pro Jahr

> Parameter λ : mittlere Rate, Erwartungswert

$$- \quad P(X = x) = \frac{\lambda^x}{x!} * e^{-\lambda}$$

Poisson-Verteilung



- > 38 Gummibärchen
- > Durchschnittlicher Verzehr
 - 5 Gummibärchen pro Stunde
- > Wahrscheinlichkeit zwischen 21 und 22 Uhr
10 Gummibärchen zu essen?
- > Mittlere Rate: $\lambda = \frac{5}{1} = 5$
- > $P(X = x) = \frac{\lambda^x}{x!} * e^{-\lambda}$
- > $P(X = 10) = \frac{5^{10}}{10!} * e^{-5} = \frac{9.765.625}{3.628.800} * 0,006738$
- > $P(X = 10) = 0,01813$



Poisson-Verteilung

- Verteilung der Anzahl von unabhängigen Ereignissen mit **konstanter mittlerer Rate** in einem **festen Zeitintervall** oder **räumlichen Gebiet** eintreten

- Diskrete Wahrscheinlichkeitsverteilung

- Parameter
 - λ : Mittlere Rate

- Wahrscheinlichkeitsfunktion

$$P_{\lambda}(k) = \frac{\lambda^k}{k!} * e^{-\lambda}$$

- Verteilungsfunktion

$$F_{\lambda}(k) = \sum_{k=0}^n P_{\lambda}(k)$$

- Erwartungswert $E(X) = \lambda$

- Varianz $\text{Var}(X) = \lambda$

- Standardabweichung $SD(X) = \sqrt{\text{Var}(X)} = \sqrt{\lambda}$

- Variationskoeffizient $\text{VarK}(X) = \frac{\sqrt{\text{Var}(X)}}{E(X)} = \frac{1}{\sqrt{\lambda}}$

- Schiefe $v(X) = \frac{1}{\sqrt{\lambda}}$

Poisson-Verteilung und Binomialverteilung

Poisson-Approximation



- > Approximation Binomialverteilung durch Poisson-Verteilung möglich wenn
 - n groß (große Stichprobe)
 - p klein (kleine Wahrscheinlichkeit)

- > $E_{Binomial}(X) = n * p \rightarrow \lambda$ für $n \rightarrow \infty$ und $p \rightarrow 0$
 - $0 < p * n < \infty$

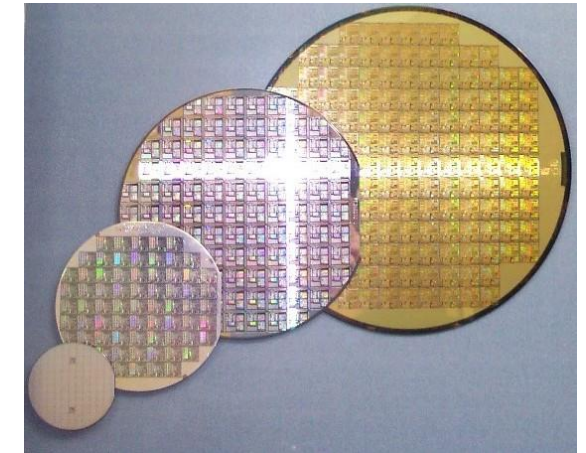
- > Faustregel
 - $n \geq 50$
 - $p \leq 0,05$

- > Vorteile
 - Hohe Werte in Tabelle für Binomialverteilung nicht enthalten
 - Hoher Rechenaufwand für Binomialkoeffizient bei großen Werten
 - Poisson-Verteilung unendlich abzählbar

Poisson-Verteilung

Anwendungsbeispiele: Chipfertigung

- > Ihre Firma stellt integrierte Chips (ICs) her. Dazu werden Siliziumwafer belichtet. Auf einem 5-Zoll-Wafer können dabei 1000 ICs aufgebracht werden. Durch Fehler, z. B. bei der Belichtung, gibt es eine durchschnittliche Fehlerrate von 144 ICs je 100 Wafer.
 - a) Welche Verteilung kann für die Abschätzung der Wahrscheinlichkeit fehlerhafter Chips verwendet werden?
 - b) Wie ist der Erwartungswert der Verteilung?
 - c) Wie ist die Varianz der Verteilung?
 - d) Wie ist die Standardabweichung der Verteilung?
 - e) Wie hoch ist die Wahrscheinlichkeit für die Produktion eines Wafers mit keinen fehlerhaften ICs?
 - f) Wie hoch ist die Wahrscheinlichkeit für die Produktion eines Wafers mit höchstens 2 fehlerhaften ICs?



Wikipedia, German Wikipediabiatch,
CC BY-SA 3.0

Poisson-Verteilung

Anwendungsbeispiele: Chipfertigung

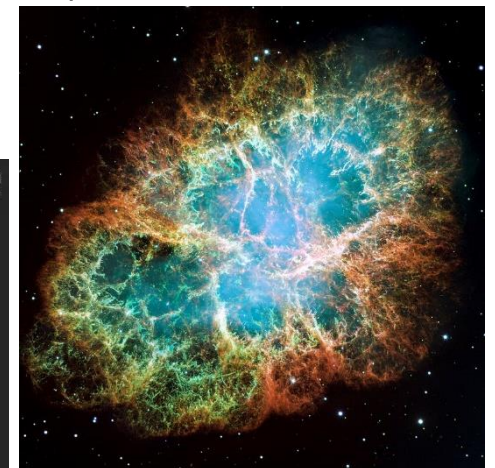
- a) Poisson-Verteilung
- b) Erwartungswert pro Einheit; 1 Wafer = 1 Einheit; $\lambda = \frac{144}{100} = 1,44$
- c) $Var(X) = \lambda = 1,44$
- d) $SD(X) = \sqrt{Var(X)} = \sqrt{1,44} = 1,2$
- e) $P_{\lambda}(k) = \frac{\lambda^k}{k!} * e^{-\lambda} = P_{1,44}(0) = \frac{1,44^0}{0!} * e^{-1,44} = \frac{1}{1} * 0,2369 \approx 0,2369$
Die Wahrscheinlichkeit für einen Wafer ohne fehlerhafte ICs beträgt ca. 23,7 %
- f) $P_{1,44}(k \leq 2) = P_{1,44}(0) + P_{1,44}(1) + P_{1,44}(2) = 0,2369 + 0,3419 + 0,2457 = 0,8245$
Die Wahrscheinlichkeit für einen Wafer mit höchstens 2 fehlerhaften ICs beträgt ca. 82,5 %

Multimessenger-Astronomie

- Supernovae (Explosion massereicher Sterne am Ende ihrer Entwicklung) sehr interessant für Forschung
- Vor Aufleuchten massiver Ausbruch an Neutrinos
 - Neutrinos erreichen Erde vor Photonen
 - Gelegenheit zum Ausrichten von Teleskopen
- Neutrinos schwierig zu detektieren
 - Geringe Masse
 - Ungeladen, keine elektromagnetische Wechselwirkung
 - Wechselwirkung nur über schwache Wechselwirkung
- Neutrino-Observatorium IceCube (Antarktis)



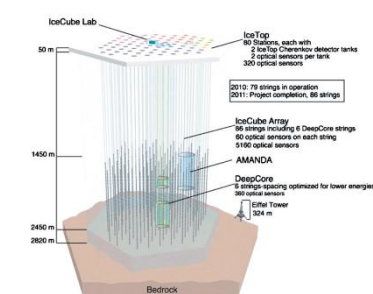
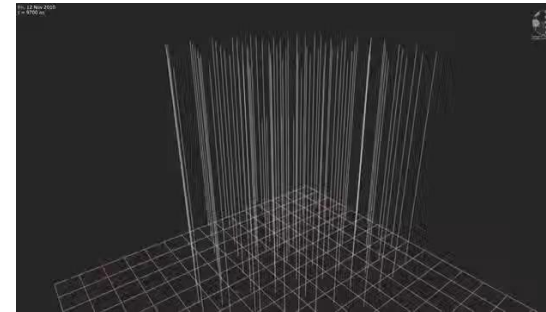
Supernova 1994D in Galaxie NGC 4526



Krebsnebel



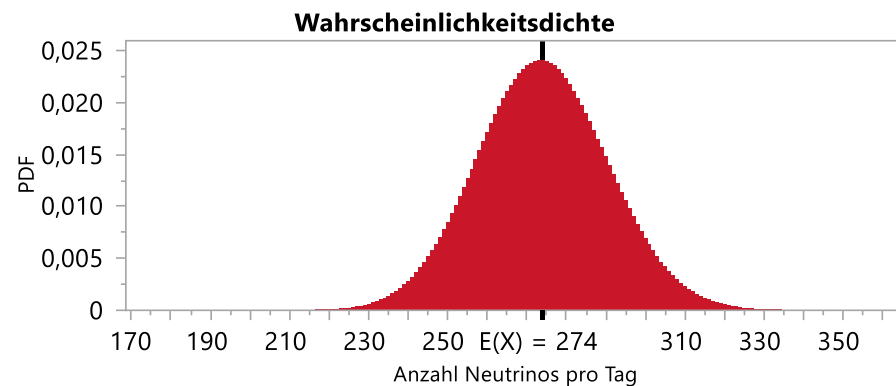
Cmichel67, CC BY-SA 4.0

IceCube Science Team,
Francis Halzen, CC BY 3.0Taavi Adamberg [youtube.com/watch?v=G1fE3ijplOw](https://www.youtube.com/watch?v=G1fE3ijplOw)
CC BY-SA 3.0

Multimessenger-Astronomie

Poisson-Verteilung

- Durchschnittliche Detektion von 100.000 Neutrinos pro Jahr
 - Ursprung u.a. Sonne, supermassive Schwarze Löcher, Supernovae, usw.
 - Neutrinoemissionen „stetig“
 - Konstante mittlere Rate pro Tag $\lambda = \frac{100.000}{365} \Rightarrow$ Poisson-Verteilung
 - Erwartungswert pro Tag $E(X) = \lambda \approx 274$
- Wie groß muss die Anzahl der Neutrinos sein, damit die Wahrscheinlichkeit, dass diese zufällig aufgetreten sind („Hintergrundrauschen“), kleiner oder gleich 5% ist?

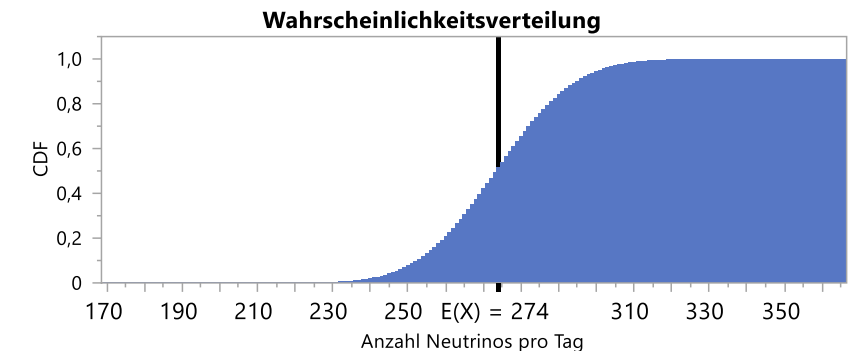


Multimessenger-Astronomie

Poisson-Verteilung



- Poisson-Verteilung: $\lambda = \frac{100.000}{365}$
 - Wahrscheinlichkeit: $P_{\lambda}(k) = \frac{\lambda^k}{k!} e^{-\lambda}$
- Wahrscheinlichkeit (0,05) gegeben, gesucht k (= Quantil)
 - Umkehrfunktion der Poisson-Verteilung benötigt
 - Nutzung Wahrscheinlichkeitsverteilung: $F_{\lambda}(k) = \sum_{i=0}^k \frac{\lambda^i}{i!} e^{-\lambda} = 1$
 - Ab welchem k gilt: $1 - \sum_{i=0}^k \frac{\lambda^i}{i!} e^{-\lambda} = 0,05$?
 - JMP JSL: `Poisson Quantile( 275, 1-0.05 )`
- $\Rightarrow k = 303$: Treten an einem Tag 303 oder mehr Neutrinos auf: Hinweis auf Supernova!



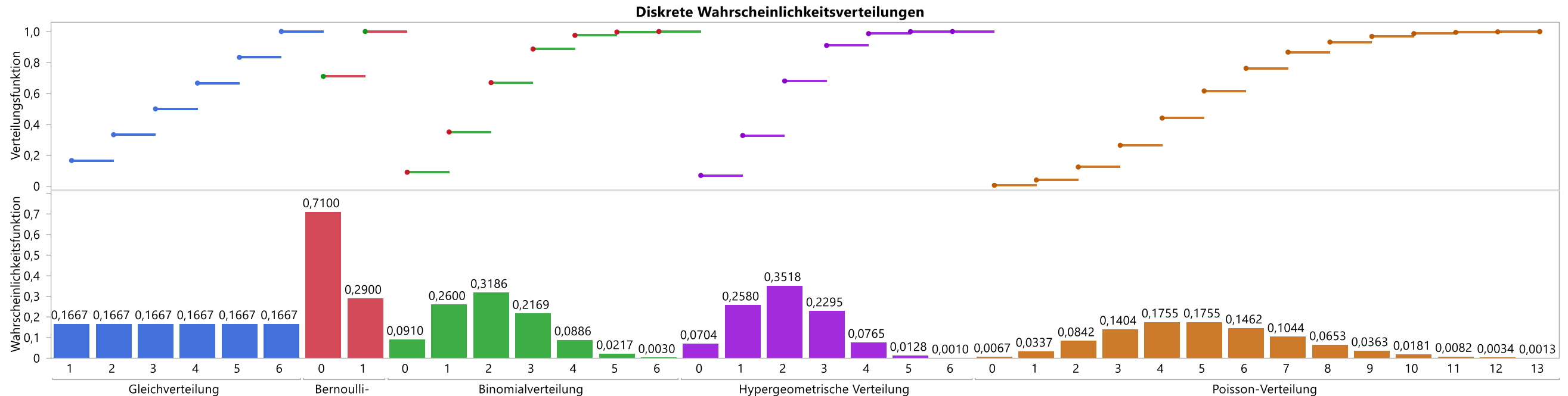


Statistik

Übersicht diskrete Wahrscheinlichkeitsverteilungen

Übersicht diskrete Wahrscheinlichkeitsverteilungen

Gleichverteilung, Bernoulli-Verteilung, Binomialverteilung, Hypergeometrische Verteilung, Poisson-Verteilung



Gleichverteilung

$$P(X = x) = \frac{1}{n}$$

Parameter

- Anzahl Elementarereignisse n

$$E(x) = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\sigma(x) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - E(x))^2}$$

Bernoulli-Verteilung

$$f(x|p) = p^x (1-p)^{1-x}$$

Parameter

- Erfolgswahrscheinlichkeit p

$$E(x) = p$$

$$\sigma(x) = \sqrt{pq} = \sqrt{p(1-p)}$$

Binomialverteilung

$$B(k|p, n) = \binom{n}{k} p^k (1-p)^{n-k}$$

Parameter

- Anzahl Erfolge k
- Anzahl Versuche n
- Erfolgswahrscheinlichkeit p

$$E(x) = n * p$$

$$\sigma(x) = \sqrt{npq} = \sqrt{np * (1-p)}$$

Hypergeometrische Verteilung

$$h(k|N, M, n) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$$

Parameter

- Anzahl Erfolge k
- Anzahl Versuche n
- Anzahl Elemente N
- Anzahl Elemente mit gewünschter Eigenschaft M

$$E(X) = n \frac{M}{N}$$

$$\sigma(x) = \sqrt{n \frac{M}{N} * \left(1 - \frac{M}{N}\right) * \frac{N-n}{N-1}}$$

Poisson-Verteilung

$$P_{\lambda}(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Parameter

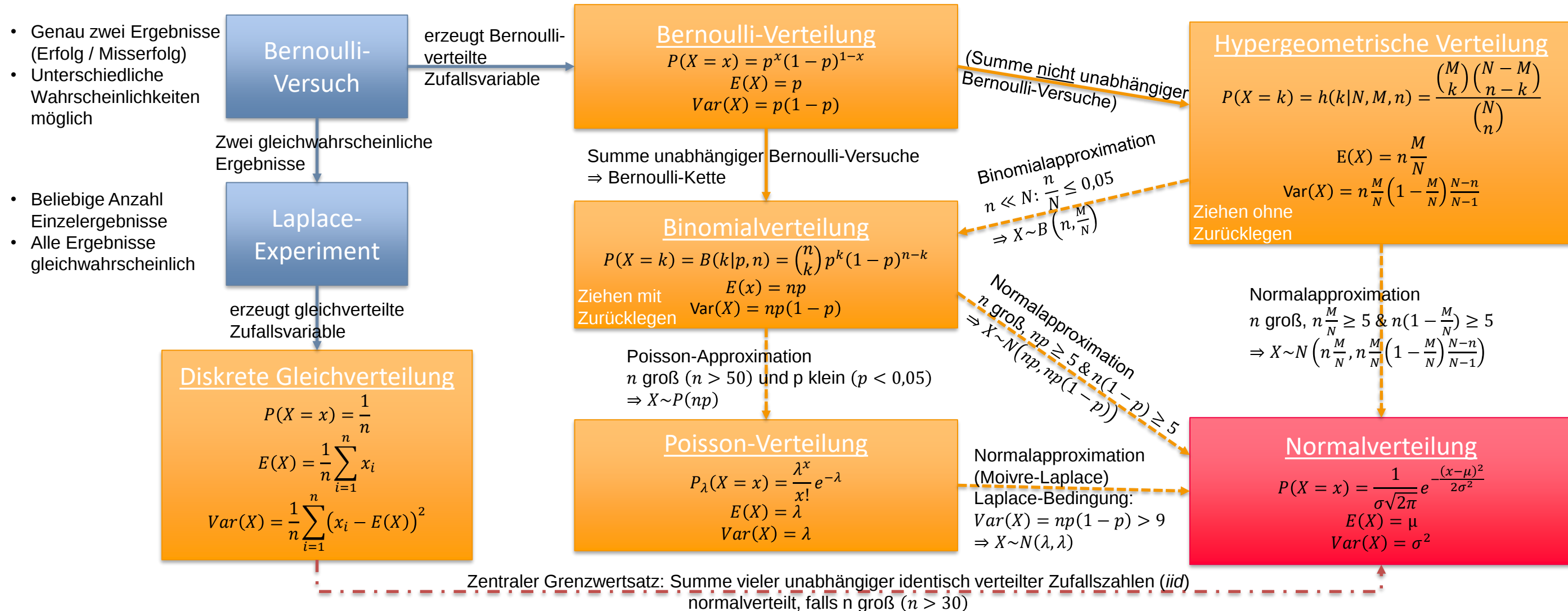
- Rate λ

$$E(x) = \lambda$$

$$\sigma(x) = \lambda$$

Übersicht (diskrete) Wahrscheinlichkeitsverteilungen

Gleichverteilung, Bernoulli-Verteilung, Binomialverteilung, Hypergeometrische Verteilung, Poisson-Verteilung, Normalverteilung





von
nd
al-
tigter
Rolle.
Schaus
die,
senerscheinung
Merkmalen

Statistik

Take Home

Take Home

- Relative Häufigkeit, Anteil und Wahrscheinlichkeit
- Wahrscheinlichkeitsintervalle und Punktschätzer
- Theoretische Wahrscheinlichkeit
 - Laplace-Experiment
 - Galton-Brett
- Diskrete vs. stetige Wahrscheinlichkeitsverteilungen
- Theoretische Wahrscheinlichkeitsverteilungen
 - Gleichverteilung
 - Bernoulli-Verteilung
 - Bernoulli-Versuch
 - Binomial-Verteilung
 - Bernoulli-Kette
 - Hypergeometrische Verteilung
 - Poisson-Verteilung