1 Statistische Grundlagen (Teil 1)

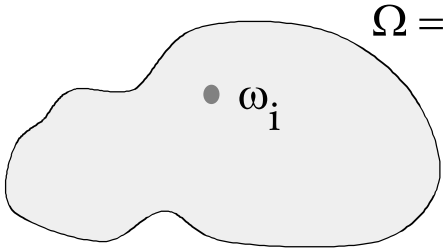
Zentraler Punkt der Mathematik ist die zahlenmäßige (quantitative) Erfassung der Welt. Während der Übergang von körperlich vorhandenen Dingen zu abstrakten Größen wie Länge, Masse oder Zeit zügig gemeistert wurde, blieb die quantitative Erfassung von Wahrscheinlichkeiten lange Zeit problematisch. Der Begriff "Wahrscheinlichkeit" wird seit 1933 durch die Axiome von Kolmogorow exakt implizit definiert. Damit wurde jedoch nicht geklärt, was Wahrscheinlichkeit ist, sondern nur herausgearbeitet, welche Eigenschaften strukturell Wahrscheinlichkeiten auszeichnen. Die Interpretation der Axiome bleibt eine offene Frage. Hier bestehen weiterhin unterschiedliche Auffassungen.

Quelle: https://de.wikipedia.org/wiki/Geschichte_der_Wahrscheinlichkeitsrechnung

1.1 Zufallsereignis und Wahrscheinlichkeit

1.1.1 Zufallsereignisse

$$\Omega = \{\omega_1, \omega_2, \ldots\}$$



Ω : Ereignisraum

ω_i : Elementarereignis

Bsp.: Würfel $\Omega = \{1,2,3,4,5,6\}$

1.1.2 Wahrscheinlichkeitsbegriff

a) Wahrscheinlichkeit als Grenzwert der relativen Häufigkeit

n : Experimente,

A_i : ausgewähltes Merkmal, Menge von elementaren Ereignissen

n_i : Anzahl der Experimente, bei denen A_i festgestellt wird

relative Häufigkeit für A_i $\qquad h(A_i) = \dfrac{n_i}{n}$

Wahrscheinlichkeit A_i

$$\boxed{P(A_i) = \lim_{n \to \infty} \frac{n_i}{n}} \qquad (2.1)$$

Eigenschaften:

- $0 \le h(A_i) \le 1 \quad \Rightarrow \quad 0 \le P(A_i) \le 1$

- disjunkte Merkmale A_i , A_j $(A_i \cap A_j = \varnothing)$

  $$h(A_i \cup A_j) = h(A_i) + h(A_j) \quad \Rightarrow \quad P(A_i \cup A_j) = P(A_i) + P(A_j)$$

- Vollständigkeitsrelation

  $$\sum_i h(\omega_i) = 1 \quad \Rightarrow \quad \sum_i P(\omega_i) = 1$$

b) Laplace-Wahrscheinlichkeit

„A priori"-Wahrscheinlichkeit aufgrund symmetrischer Eigenschaften, z.B. Münze, Würfel, etc.

Experiment mit m gleichwertigen Alternativen
(m : Anzahl aller Alternativen)

m_i : Anzahl der Alternativen für A_i

Wahrscheinlichkeit für A_i wird definiert als: $$\boxed{P(A_i) = \frac{m_i}{m}} \qquad (2.2)$$

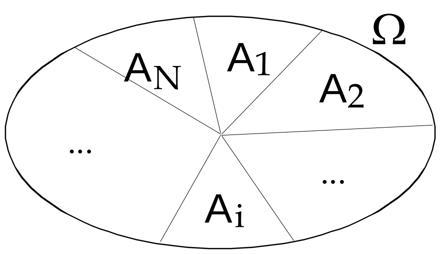
Beispiel Würfel :

- m=6 Alternativen (gleichwertig bei vollkommen symmetrischen Würfeln)
- A_i : Augenzahl ungerade

$$m_i = 3 \Rightarrow P(A_i) = \frac{3}{6} = 0{,}5$$

1.1.3 Wichtige Begriffe und Gesetze

a) Vollständiges Ereignissystem



Gegeben:

$$M = \{A_1, A_2, ..., A_N\} \ , \qquad A_i \cap A_j = \emptyset \quad \forall\, i \neq j$$

$$\Rightarrow \quad P(M) = \sum_{i=1}^{N} P(A_i) \qquad (2.3)$$

Die Menge von disjunkten Ereignissen $M = \{A_1, \dots, A_N\}$ bildet ein *vollständiges Ereignissystem* bzw. eine Partition, falls deren Vereinigung den gesamten Ereignisraum abdeckt, d.h. $\Omega = \bigcup_{1 \le i \le N} A_i$

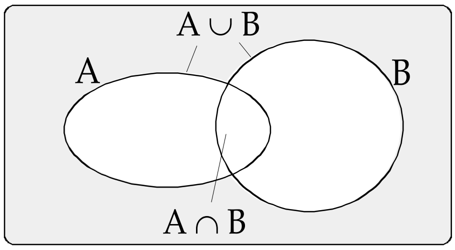
b) Verbundereignis und Verbundwahrscheinlichkeit

- A , B : zwei (nicht notwendigerweise disjunkte) Ereignisse

  Verbundereignis : $(A \cap B) = (A, B)$

  Verbundwahrscheinlichkeit : $P(A \cap B) = P(A, B) = P(B, A)$

  Weiter gilt: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (2.4)

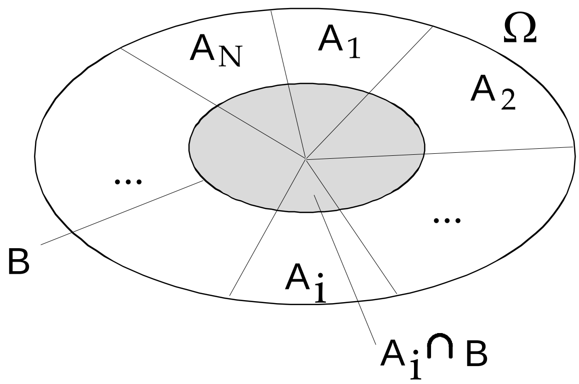


Anwendungsbeispiele:

- A: „Augenzahl maximal 3“
- B: „Augenzahl gerade“

- Falls $M = \{A_1, \dots, A_N\}$ ein vollständiges Ereignissystem bildet, dann gilt:

$$P(B) = \sum_{i=1}^{N} P(A_i, B) \tag{2.5}$$



c) Bedingte Wahrscheinlichkeit

- Bedingtes Ereignis $(A\,|\,B)$
  Eintritt des Ereignisses A unter der Bedingung, dass Ereignis B eintritt; Voraussetzung ist $P(B) > 0$.

- Bedingte Wahrscheinlichkeit

$$P(A\,|\,B) = \frac{P(A,B)}{P(B)} \quad , \quad \text{d.h.} \quad P(A\,|\,B) \geq P(A,B) \tag{2.6}$$

- Simpson Paradoxon
  Eigenschaften einer Gesamtmenge stehen mindestens teilweise im Widerspruch zu den Eigenschaften der Teilmengen.

  Beispiel:
  http://www.spiegel.de/wissenschaft/mensch/simpsons-paradoxon-diese-statistik-kann-nicht-stimmen-oder-doch-a-1068204.html

  - Bewerbungen an einer Hochschule: 500 Frauen und 500 Männer
  - Zulassungen: 240 Frauen (48% Quote) und 300 Männer (60% Quote)
    $\Rightarrow$ In der Gesamtmenge wurden mehr Männer zuglassen
  - 2 Studiengänge
    1. Studiengang: 360 Studienplätze
       Bewerbungen: 100 Frauen und 400 Männer
       Zulassungen: 80 Frauen (80% Quote) und 280 Männer (70% Quote)
    2. Studiengang: 180 Studienplätze
       Bewerbungen: 400 Frauen und 100 Männer
       Zulassungen: 160 Frauen (40% Quote) und 20 Männer (20% Quote)

    $\Rightarrow$ In beiden Studiengängen wurden mehr Frauen zuglassen.

d) Statistische Unabhängigkeit

Ereignisse A und B sind statistisch unabhängig, wenn gilt:

$$P(A \,|\, B) = P(A) \quad \text{oder} \quad P(A, B) = P(A) \cdot P(B) \tag{2.7}$$

e) Gesetz der totalen Wahrscheinlichkeit

Es sei $\{A_i\,,\ i = 1,\dots,N\}$ ein vollständiges Ereignissystem und $B \subset \Omega$. Dann gilt:

$$P(B) = \sum_{i=1}^{N} P(A_i, B) = \sum_{i=1}^{N} P(B \,|\, A_i) \cdot P(A_i)$$

Aus der zweimaligen Anwendung von (2.6) erhalten wir

$$P(A_i, B) = P(B \,|\, A_i) \cdot P(A_i) = P(A_i | B) \cdot P(B)$$

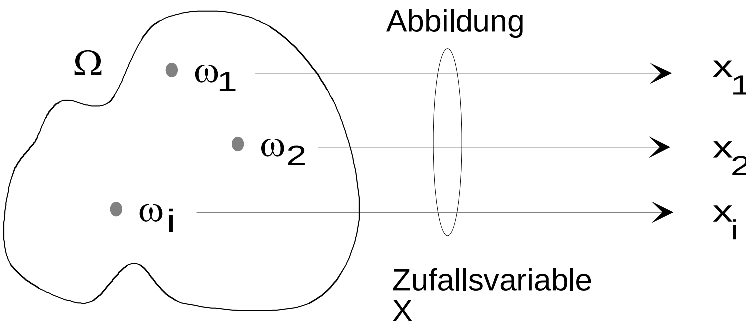
woraus sich die <u>Bayes-Formel</u> herleitet:

$$P(A_i \,|\, B) = \frac{P(B \,|\, A_i) \cdot P(A_i)}{P(B)} = \frac{P(B \,|\, A_i) \cdot P(A_i)}{\sum\limits_{j=1}^{N} P(B \,|\, A_j) \cdot P(A_j)} \tag{2.8}$$

Sie hilft $P(A_i \,|\, B)$ zu berechnen, wenn nur $P(B \,|\, A_i)$ und $P(A_i)$ bekannt sind.

1.2 Zufallsvariable, Verteilung und Verteilungsfunktion

1.2.1 Zufallsvariable



- Zufallsvariable (ZV)
  - Offiziell: Funktion, die jedem Elementarereignis ω_i eine reelle Zahl zuordnet.
  - Intuitiv: Variable, deren Wert (eventuell mehrmals) zufällig nach einer gegebenen (Wahrscheinlichkeits-)Verteilung(sfunktion) ausgewürfelt wird.

- Diskrete Zufallsvariable
  - ZV X hat diskreten Wertebereich (meist nur ganzzahlige Werte)
  - Beispiel: Wurf mit zwei Würfeln, $X \in \{2, 3, ..., 12\}$

- Kontinuierliche Zufallsvariable
  - ZV A hat reellwertigen Wertebereich
  - Beispiel: Durchlaufzeit eines Datenpakets im Rechnernetz, $A \in (150\text{ms}, 700\text{ms}]$

1.2.2 Verteilung

a) Definition

- X : diskrete Zufallsvariable

$$x(i) = P(X = i) \quad , \qquad i = 0,1,\ldots,X_{max} \qquad \text{(Verteilung)} \tag{2.9}$$

$$\sum_{i=0}^{X_{max}} x(i) = 1 \qquad \text{(Vollständigkeitsrelation)} \tag{2.10}$$

b) Beispiele einfacher Verteilungen

- Bernoulli-Verteilung
  - Bernoulli-Versuch: Experiment mit zwei möglichen Ergebnissen
  - X: ZV für den Ausgang eines Bernoulli-Versuchs
  - Verteilung: $x(i) = \begin{cases} 1-p & i = 0\ (Misserfolg) \\ p & i = 1\ (Erfolg) \end{cases}$ (2.11)
  - Wertebereich: $i \in \{0,1\}$
  - Beispiel: Übertragung eines Bits als Bernoulli-Versuch
    - Bitfehlerwahrscheinlichkeit $p_b \Rightarrow p = 1 - p_b$

- Binomial-Verteilung
  - X: ZV für die Anzahl von Erfolgen bei n Bernoulli-Versuchen mit Erfolgswahrscheinlichkeit p
  - Verteilung: $x(i) = \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$ (2.12)
    - $\binom{n}{i} = \frac{n!}{i! \cdot (n-i)!}$ verschiedene Muster mit i Erfolgen
    - Jedes Muster mit i Erfolgen tritt auf mit Wahrscheinlichkeit $p^i \cdot (1-p)^{n-i}$
  - Wertebereich: $0 \leq i \leq n$
  - Beispiel: Fehlerwahrscheinlichkeit p_p eines Datenpakets mit n Bits
    - Gegeben: Bitfehlerwahrscheinlichkeit p_b
    - Wahrscheinlichkeit für i Bitfehler im Paket wird durch Binomial-Verteilung beschrieben
    - Paketfehlerwahrscheinlichkeit $p_p = \sum_{1 \leq i \leq n} x(i) = 1 - x(0) = 1 - (1 - p_b)^n \neq n \cdot p_b$
    - Für $p_b \ll 1$ gilt: $(1 - p_b)^n \approx 1 - n \cdot p_b \Rightarrow\ p_p \approx n \cdot p_b$

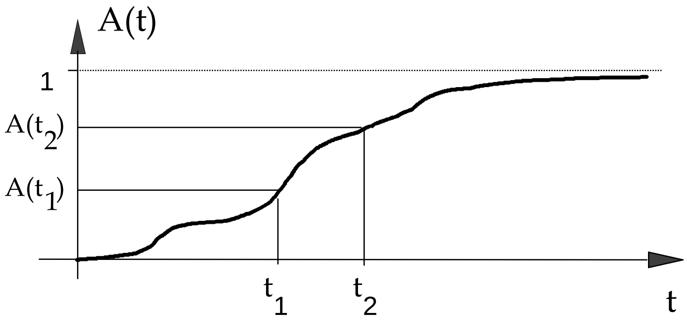
- Geometrische Verteilung
  - X: ZV für die Anzahl von Fehlversuchen bei wiederholten Bernoulli-Versuchen mit Erfolgswahrscheinlichkeit p bis zum ersten Erfolg
  - Verteilung: $x(i) = (1-p)^i \cdot p$ (2.13)
  - Wertebereich: $0 \leq i < \infty$
  - Beispiel: Anzahl von fehlerhaften Übertragungsversuchen Y beim Versenden eines Datenpaketes mit Fehlerwahrscheinlichkeit p_p
    - $y(i) = {p_p}^i \cdot (1 - p_p)$

1.2.3 Verteilungsfunktion und –dichtefunktion

A : beliebige (diskrete oder kontinuierliche) Zufallsvariable

a) Verteilungsfunktion (VF)

- $A(t) = P(A \leq t)$ (Verteilungsfunktion) (2.14)
- Visualisierung



- Eigenschaften
  - $t_1 < t_2 \quad \Rightarrow \quad A(t_1) \leq A(t_2)$ (Monotonieeigenschaft)
  - $t_1 < t_2 \quad \Rightarrow \quad P(t_1 < A \leq t_2) = A(t_2) - A(t_1)$
  - $A(-\infty) = 0$
  - $A(\infty) = 1$
- $A^c(t) = P(A > t) = 1 - A(t)$ (komplementäre Verteilungsfunktion)

b) Verteilungsdichtefunktion (VDF)

- Nur bei kontinuierlichen ZV anwendbar
- Pendant zur Verteilung von diskreten ZV
- $a(t) = \frac{d}{dt} A(t)$ (Verteilungsdichtefunktion) (2.15)
- $\int_{-\infty}^{\infty} a(t)\, dt = 1$ (Vollständigkeitsrelation) (2.16)
- $A(t) = \int_{-\infty}^{t} a(\tau)\, d\tau$ (Zusammenhang zw. VDF und VF)

- Beispiel: Exponentielle VF
    - VF: $A(t) = 1 - e^{-\lambda \cdot t}$
    - VDF: $a(t) = \lambda \cdot e^{-\lambda \cdot t}$

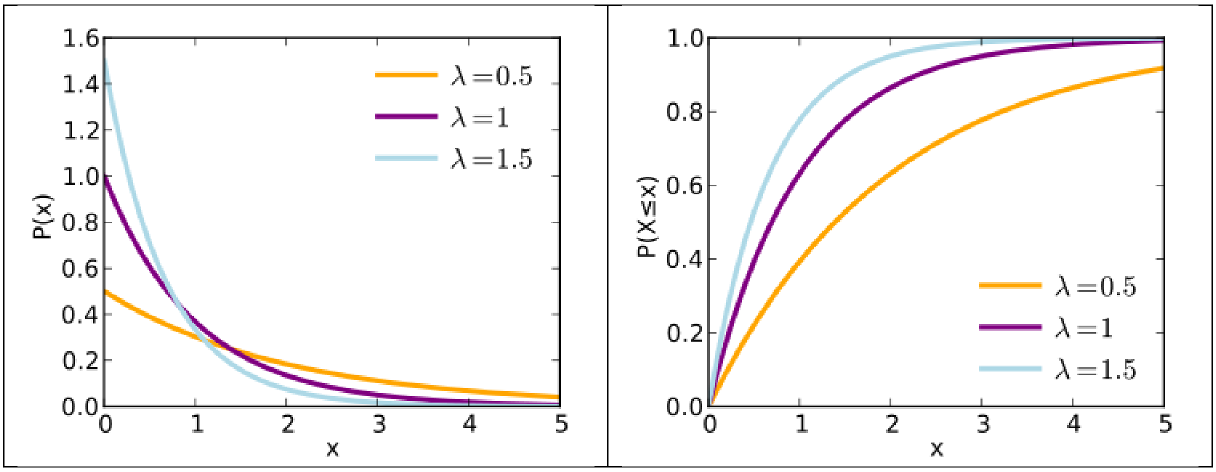


Figure 1: VDF und VF der exponentiellen Verteilung (Quelle: Wikipedia)

1.2.4 Special Operations on Distributions

Applicable to independent random variables.
Examples: X_i = number of pips, A_i= exponentially distributed duration

1.2.4.1 Sum of Two Random Variables (Convolution, „Faltung“)

1.2.4.1.1 Discrete Distributions

- X_1, X_2: random variables
- Wanted: distribution of $X = X_1 + X_2$
- $P(X = i) = x(i) = \sum_j x_1(j) \cdot x_2(i - j) = (x_1 * x_2)(i)$

1.2.4.1.2 Continuous Distributions

- A_1, A_2: random variables
- Wanted: probability density function of $A = A_1 + A_2$
- $a(t) = \int_{-\infty}^{\infty} a_1(\tau) \cdot a_2(t - \tau) d\tau = (a_1 * a_2)(t)$

1.2.4.2 Difference of Two Random Variables

1.2.4.2.1 Discrete Distributions

- X_1, X_2: random variables
- Wanted: distribution of $X = X_1 - X_2$
- $P(X = i) = x(i) = \sum_j x_1(i + j) \cdot x_2(j)$

1.2.4.2.2 Continuous Distributions

- A_1, A_2: random variables
- Wanted: probability density function of $A = A_1 - A_2$
- $a(t) = \int_{-\infty}^{\infty} a_1(t + \tau) \cdot a_2(\tau) d\tau$

1.2.4.3 Maximum of Several Random Variables

- A_i: random variables
- Wanted: distribution function of $A = \max(A_0, \dots, A_{k-1})$
- $k = 2$: $A(t) = A_1(t) \cdot A_2(t)$
  Proof: $A(t) = P(max(A_1, A_2) \leq t) = P(A_1 \leq t) \cdot P(A_2 \leq t) = A_1(t) \cdot A_2(t)$
- k=2: $a(t) = a_1(t) \cdot A_2(t) + a_2(t) \cdot A_1(t)$
  Proof: $a(t) = \frac{dA(t)}{dt}$
- $A(t) = \prod_{0 \leq i < k} A_i(t)$

1.2.4.4 *Minimum of Several Random Variables*

- A_i: random variables
- Wanted: distribution function of $A = \min(A_0, \dots, A_{k-1})$
- $k = 2$: $A^c(t) = A_1^c(t) \cdot A_2^c(t)$, $a(t) = a_1(t) \cdot A_2^c(t) + a_2(t) \cdot A_1^c(t)$
- $A^c(t) = \prod_{0 \le i < k} A_i^c(t)$
- In other words: $A(t) = 1 - \prod_{0 \le i < k}\left(1 - A_i(t)\right)$

1.3 Statistische Größen von Verteilungen

1.3.1 Erwartungswert

- ZV A mit VDF a(t)
- Reellwertige Funktion g(A) der ZV A

$$\boxed{E[g(A)] = \int_{-\infty}^{+\infty} g(t) \cdot a(t)\, dt} \qquad \text{(Erwartungswert von g(A))} \tag{2.17}$$

- Analog für diskrete ZV X mit Verteilung *x(i)*: $E[g(X)] = \sum_i g(i) \cdot x(i)$

1.3.2 Erwartungswert (erstes Moment) einer Zufallsvariablen

- $g(A) = A \quad \Rightarrow \quad m_1 = E[A] = \int_{-\infty}^{+\infty} t \cdot a(t)\, dt \qquad (2.18)$
- Analog für diskrete ZV X mit Verteilung *x(i)*: $E[X] = \sum_i i \cdot x(i)$

1.3.3 Verallgemeinerung: Gewöhnliche Momente

- k-tes gewöhnliches Moment von A

$$g(A) = A^k \quad \Rightarrow \quad m_k = E[A^k] = \int_{-\infty}^{+\infty} t^k \cdot a(t)\, dt \tag{2.19}$$

- Analog für diskrete ZV X mit Verteilung *x(i)*: $m_k = E[X^k] = \sum_i i^k \cdot x(i)$
- Das erste gewöhnliche Moment heißt Erwartungswert E[A] von A (mean, expected value).

1.3.4 Zentrale Momente

- **k-tes zentrales Moment von A**

$$g(A) = (A - m_1)^k \quad \Rightarrow \quad \mu_k = E[(A - m_1)^k] = \int_{-\infty}^{+\infty} (t - m_1)^k \cdot a(t)\, dt \tag{2.20}$$

- Analog für diskrete ZV X mit Verteilung *x(i)*:

$$\mu_k = E[(X - m_1)^k] = \sum_i (i - m_1)^k \cdot x(i)$$

- **Varianz (k=2) (variance)**

$$
\begin{aligned}
VAR[A] = \mu_2 &= E[(A - m_1)^2] = E[A^2 - 2Am_1 + m_1^2] \\
&= E[A^2] - 2m_1E[A] + m_1^2 \\
&= m_2 - m_1^2
\end{aligned}
\tag{2.21}
$$

- **Standardabweichung (standard deviation)**

$$\sigma_A = \sqrt{VAR[A]} \tag{2.22}$$

- **Variationskoeffizient (coefficient of variation)**

$$c_{var}[A] = c_A = \frac{\sigma_A}{m_1} = \frac{\sqrt{VAR[A]}}{E[A]} \tag{2.23}$$

Der Variationskoeffizient macht nur für positive A Sinn, denn sonst kann der Mittelwert E[A] null oder negativ werden.

- **Schiefe (skewness)**

$$\upsilon_A = \frac{E\left[(A - E[A])^3\right]}{\sigma_A^3} \tag{2.24}$$

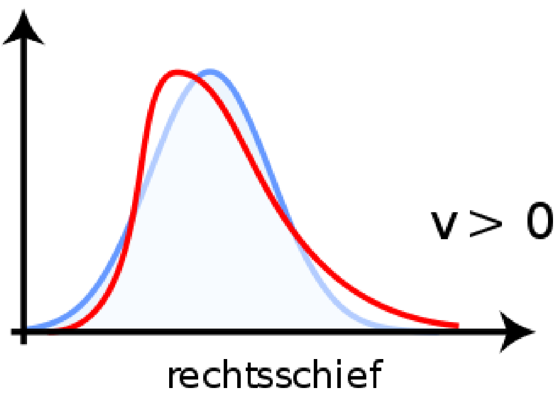
symmetrische Verteilung: $\upsilon_A = 0$

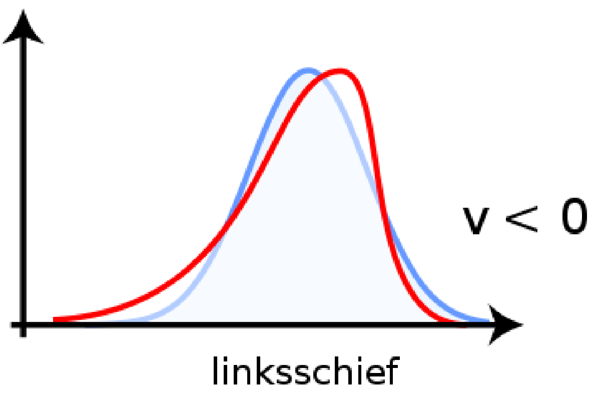
$\upsilon_A > 0$: rechtsschief

$\upsilon_A < 0$: linksschief

x0	p(x0)	x1	p(x1)	E[A]	E[A²]	E[A³]	VAR[A]	stddev[A]	E[(A-E[A])³]	skewness[A]
-1	0,1	1	0,9	0,8	1	0,8	0,36	0,6	-0,576	-2,666666667
-1	0,5	1	0,5	0	1	0	1	1	0	0
-1	0,9	1	0,1	-0,8	1	-0,8	0,36	0,6	0,576	2,666666667

Beispiele (Quelle: `http://de.wikipedia.org/wiki/Schiefe_(Statistik)`)





1.3.5 Beispiele

1.3.5.1 *Bernoulli-Verteilung*

$$E[X] = \sum_i i \cdot x(i) = p$$

$$E[X^2] = \sum_i i^2 \cdot x(i) = p$$

$$VAR[X] = E[X^2] - E[X]^2 = p - p^2 = p \cdot (1-p)$$

$$\sigma[X] = \sqrt{p \cdot (1-p)}$$

$$c_{var}[X] = \frac{\sqrt{p \cdot (1-p)}}{p} = \sqrt{\frac{(1-p)}{p}}$$

1.3.5.2 *Geometrische Verteilung*

$$E[X] = \sum_{i=0}^{\infty} i \cdot x(i) = \sum_{i=0}^{\infty} i \cdot (1-p)^i \cdot p = p \cdot (1-p) \cdot \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1}$$

$$= p \cdot (1-p) \cdot \left(-\sum_{i=0}^{\infty} (1-p)^i\right)' = p \cdot (1-p) \cdot \left(\frac{-1}{p}\right)' = p \cdot (1-p) \cdot \frac{1}{p^2} = \frac{(1-p)}{p}$$

$$E[X^2] = \sum_{i=0}^{\infty} i^2 \cdot x(i) = \sum_{i=0}^{\infty} i^2 \cdot (1-p)^i \cdot p = \sum_{i=0}^{\infty} i^2 \cdot q^i \cdot (1-q)$$

$$= (1-p)^2 \cdot p \cdot \left(\sum_{i=0}^{\infty} i^2 \cdot (1-p)^{i-2}\right)$$

$$= (1-p)^2 \cdot p \cdot \left(\sum_{i=0}^{\infty} (i-1) \cdot i \cdot (1-p)^{i-2} + \sum_{i=0}^{\infty} i \cdot (1-p)^{i-2}\right)$$

$$= (1-p)^2 \cdot p \cdot \left(\sum_{i=0}^{\infty} (1-p)^i\right)'' + \sum_{i=0}^{\infty} i \cdot (1-p)^i \cdot p = (1-p)^2 \cdot p \cdot \left(\frac{1}{p}\right)'' + \frac{(1-p)}{p}$$

$$= (1-p)^2 \cdot p \cdot \frac{2}{p^3} + \frac{(1-p)}{p} = \frac{2 \cdot (1-p)^2 + p \cdot (1-p)}{p^2} = \frac{(1-p)^2 + (1-p)}{p^2}$$

$$VAR[X] = E[X^2] - E[X]^2 = \frac{(1-p)^2 + (1-p)}{p^2} - \left(\frac{(1-p)}{p}\right)^2 = \frac{(1-p)}{p^2}$$

$$\sigma[X] = \frac{\sqrt{(1-p)}}{p}$$

$$c_{var}[X] = \frac{\sigma[X]}{E[X]} = \frac{\sqrt{(1-p)}}{p} \cdot \frac{p}{1-p} = \frac{1}{\sqrt{1-p}}$$

1.3.5.3 Exponentielle Verteilung

$$E[A] = \int_0^\infty t \cdot a(t)dt = \int_0^\infty t \cdot \lambda \cdot e^{-\lambda \cdot t} dt = \left[t \cdot \left(-e^{-\lambda \cdot t}\right)\right]_0^\infty - \int_0^\infty \left(-e^{-\lambda \cdot t}\right)dt =$$

$$0 - 0 - \left(\frac{1}{\lambda}\left[e^{-\lambda \cdot t}\right]_0^\infty\right) = \frac{1}{\lambda}$$

$$E[A^2] = \int_0^\infty t^2 \cdot a(t)dt = \int_0^\infty t^2 \cdot \lambda \cdot e^{-\lambda \cdot t} dt = \cdots = \frac{2}{\lambda^2}$$

$$VAR[A] = E[A^2] - E[A]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$\sigma[A] = \sqrt{\frac{1}{\lambda^2}} = \frac{1}{\lambda}$$

$$c_{var}[A] = \frac{\sigma[A]}{E[A]} = \frac{1}{\lambda} \Big/ \frac{1}{\lambda} = 1$$

1.3.6 Skalierungseigenschaften

- C=b · A
    - E[C]=b · E[A] (2.27a)
    - VAR[C]=b² · VAR[A] (2.27b)
- C=A+B
    - E[C]=E[A]+E[B] (immer) (2.27c)
    - VAR[C]=VAR[A]+VAR[B] (falls A und B unabhängig) (2.27d)

Beweis siehe Tran-Gia, Kap. 1.2.5 c)

1.3.7 p-Quantil t_p (im Englischen auch p-percentile)

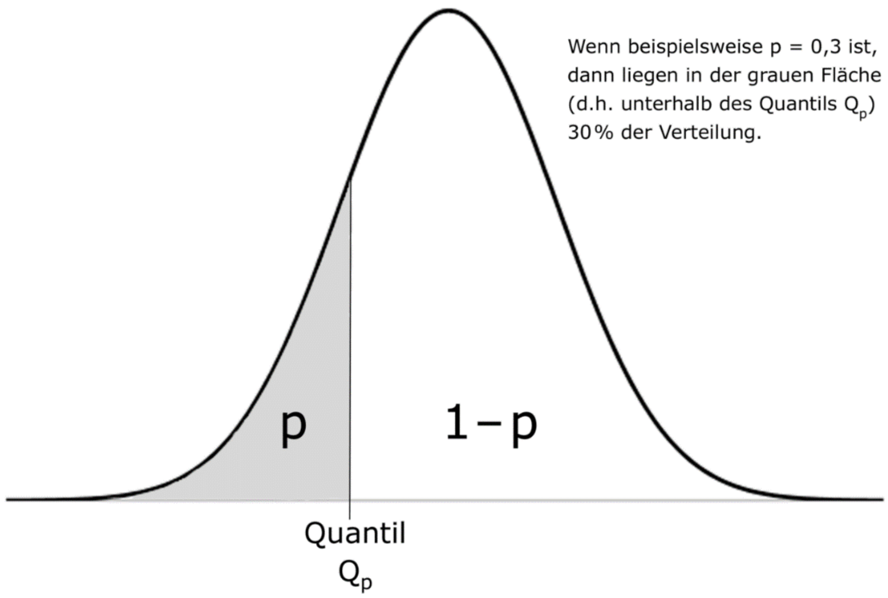


Abbildung 1: Das p-Quantil (t_p oder Q_p) visualisiert anhand der Verteilungsdichtefunktion. Quelle: https://upload.wikimedia.org/wikipedia/commons/b/ba/Normalverteilung.png

- $t_p = \inf_{t \in R} \left(A(t) \geq p \right) =: A^{-1}(p)$ für p $(0 < p < 1)$
  - Wird auch Q_p genannt
- Bei kontinuierlichen ZVs ist eine echte Umkehrung möglich, da A(t) bijektiv ist.
- Spezielle Quantile
  - Das 0.5-Quantil heißt Median.
    - Median und Erwartungswert stimmen für symmetrische Verteilungen überein.
    - Der Median kann aussagekräftiger sein als der Mittelwert, wenn dieser durch Ausreißer stark beeinflusst wird.
  - Das 0.25 bzw 0.75-Quantil heißt unteres bzw. oberes Quartil.
- Einhaltung von Grenzwerten
  - Mittelwert nicht aussagekräftig
  - Harte Obergrenze möglicherweise schwer einzuhalten
  - Grenzwert darf mit 0.1% überschritten werden: Das 99.9% Quantil muss kleiner als der Grenzwert sein. Vorteil: Grenzwert wird fast immer eingehalten, das technische System ist aber i.d.R. aber wesentlich einfacher als eines, welche strikte Obergrenzen einhalten muss.

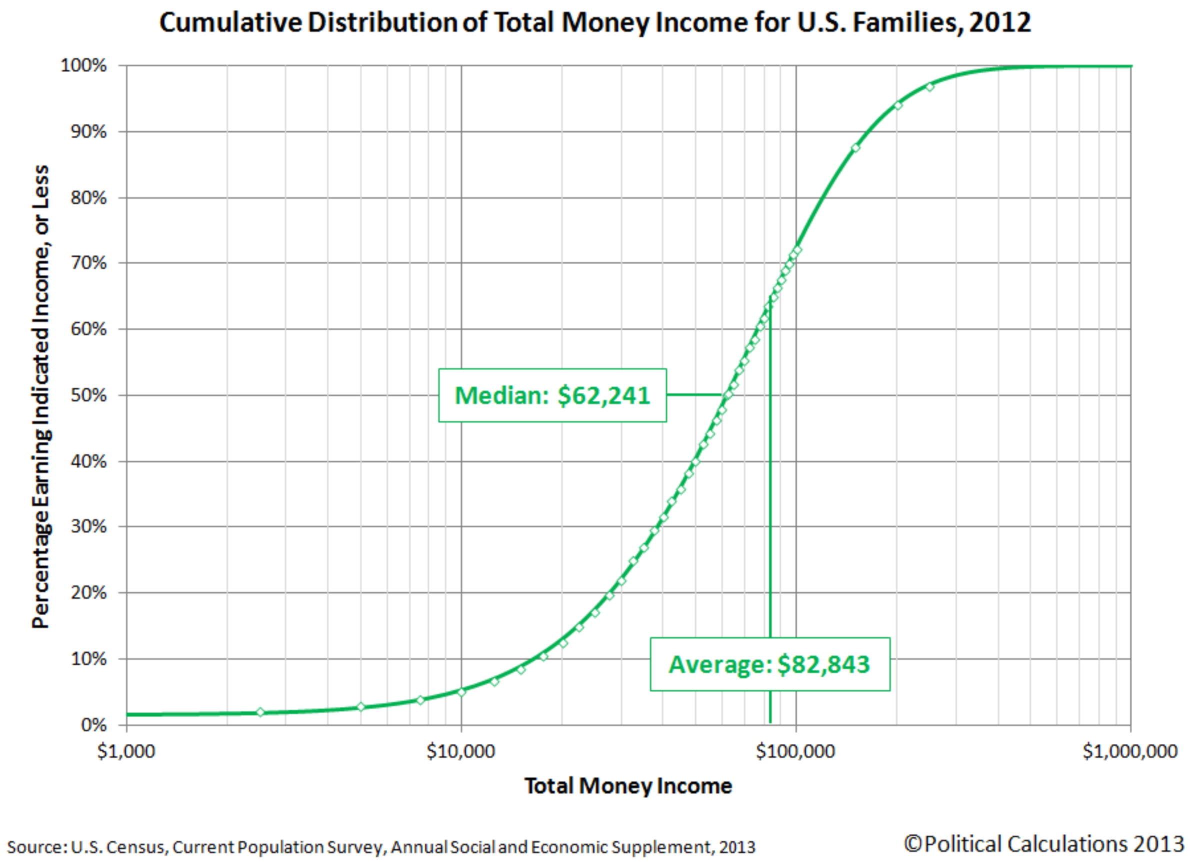


Abbildung 2: Das 50%-Quantil (Median) anhand einer Verteilungsfunktion. Quelle: http://politicalcalculations.blogspot.de/2013/09/what-is-your-us-income-percentile.html

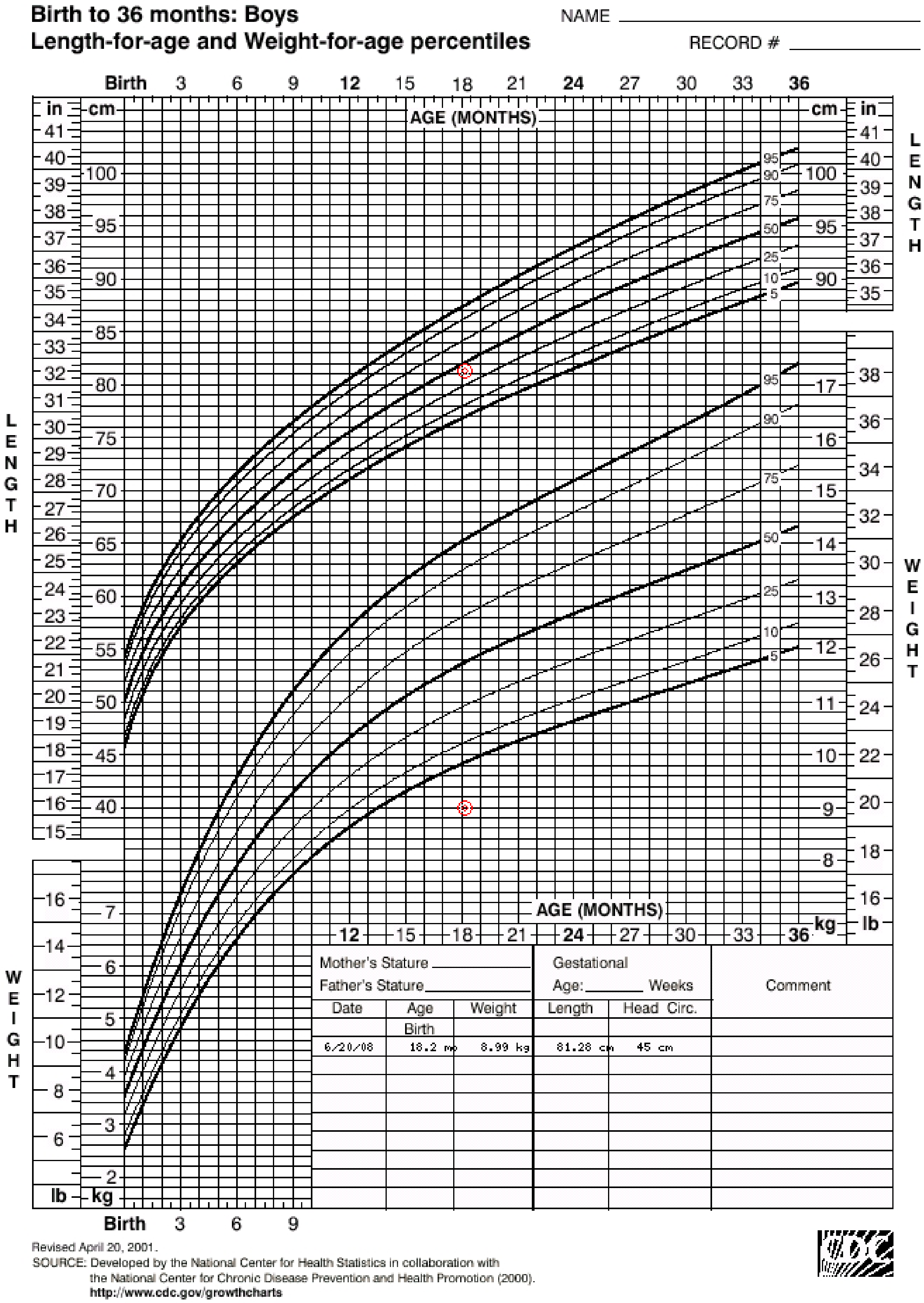


Abbildung 3: 5%, 10%, 25%, 50%, 75%, 90% und 95% Quantile von Körperlänge und –gewicht bei Jungen im Alter zwischen 0 und 36 Monaten.
Quelle: https://higherhighslowerlows.files.wordpress.com/2010/09/growth-chart.png

1.4 Statistische Größen von Stichproben (LK 4.3, 4.4, 6.4.1)

Definition:
Ein stochastischer Prozess ist eine geordnete Menge von Tupeln $\{X(t), t\}$, $X(t)\in\Xi$, $t\in\Gamma$ mit Zustandsraum Ξ und Indexmenge Γ.
Der Prozess ist zeitdiskret falls $\Gamma=N_0$ (=Betrachtungszeitpunkte) ist und zeitkontinuierlich falls $\Gamma=R_0$ (=zeitlicher Verlauf) ist.

Problem: Keine Verteilung gegeben sondern nur eine Stichprobenmenge $X_0, X_1, \ldots, X_{n-1}$ mit mit $X_i=X(t_i)$ und $t_i\in[0, T]$.
Statistische Größen müssen anders berechnet werden.

1.4.1 Allgemeine Empirische Momente

$$\overline{X^k} = \begin{cases} \overline{X^k(n)} = \frac{1}{n} \cdot \sum\limits_{0\le i<n} X_i^k & \text{für zeitdiskreten Prozess} \\ \overline{X^k(T)} = \frac{1}{T} \cdot \int\limits_0^T X(t)^k \, dt & \text{für zeitkontinuierlichen Prozess} \end{cases} \tag{2.28}$$

wobei n bzw. T die Größe der Stichprobe widerspiegelt.

1.4.2 Empirischer Mittelwert (sample mean)

$$\overline{X} = \overline{X^1} \tag{2.29}$$

1.4.3 Empirische Varianz (sample variance)

$$S^2 = \begin{cases} S^2(n) = \frac{1}{n-1} \cdot \sum\limits_{0\le i<n} \left(X_i - \overline{X}\right)^2 = \frac{n}{n-1} \cdot \left(\overline{X^2} - \overline{X^1}^2\right) & \text{für zeitdiskreten Prozess} \\ S^2(T) = \frac{1}{T} \int\limits_0^T \left(X(t) - \overline{X}\right)^2 dt = \overline{X^2} - \overline{X^1}^2 & \text{für zeitkontinuierlichen Prozess} \end{cases}$$

(2.30)

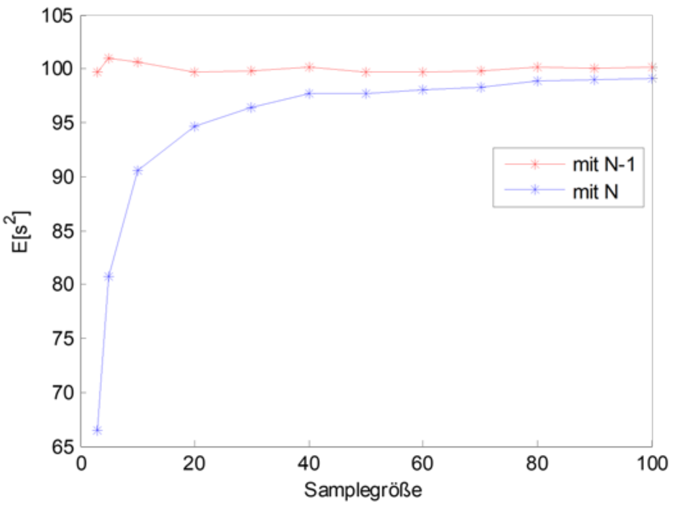


Abbildung 4: Einfluss der Erwartungstreue der Varianzschätzer auf die empirische Varianz.

1.4.4 Empirische Schiefe (sample skewness)

$$\hat{v} \quad = \frac{\overline{(X-\bar{X})^3}}{(S^2)^{\frac{3}{2}}} = \frac{n}{(n-1)\cdot(n-2)} \cdot \sum_{i=0}^{n} \left(\frac{(X_i-\bar{X})}{\sqrt{S^2}}\right)^3 \qquad (2.31)$$

Siehe: http://de.wikipedia.org/wiki/Schiefe_(Statistik)

1.5 Messobjekte in Simulationen

In Simulationsprogrammen werden Statistikobjekte angelegt, denen im Laufe einer Simulation Stichproben mit Zeitstempel zur Auswertung übergeben werden. Es gibt diskrete und zeitkontinuierliche Messwerte. Diese müssen bei der Datenerfassung unterschiedlich behandelt werden.

- http://simul.iro.umontreal.ca/ssj/doc/html/umontreal/iro/lecuyer/stat/StatProbe.html
- http://desmoj.sourceforge.net/tutorial/statistics/0.html

1.5.1 Erfassungsobjekte für Momente von diskreten Messwerten

- „Discrete counters (DC)“, „tally“ (Strichliste)
- Beispiele: Wartezeiten, Paketgrößen, Zwischenankunftszeiten
- Übergabe der Stichproben X_i an das Statistikobjekt
- Interne Datenhaltung: $\sum_{0 \leq i < n} X_i^k$ und n
- Berechnung der statistischen Größen nach 2.28 und 2.30

1.5.2 Erfassungsobjekte für Verteilungen von diskreten Messwerten (LK 6.4.2)

- Histogramm mit Gesamtwertebereich: $[a_0, b_{m-1})$
- Enthält m (z.B $m=2^k$) Zielintervalle mit je einer Zählvariable (bin)
- Array von bins entsprechend dem Wertebereich Range(j)= $[a_j, b_j)$
  - Breite äquidistanter Intervalle: $\Delta = \frac{b_{m-1} - a_0}{m}$
  - $a_j = a_0 + j * \Delta$
  - $b_j = a_0 + (j+1) * \Delta$
- Anzahl der bisherigen Messwerte n wird gespeichert
- Aktionen bei Übergabe der Stichprobe X_i
  - falls $X_i \in$ Range(j), inkrementiere bin(j) um 1,
  - inkrementiere n um 1
- Ausgabe als Liste mit Tupeln
  - (Range(j), h_j) mit relativer Häufigkeit: h_j=bin(j)/n
  - Diskrete empirische Verteilung: $((b_j + a_j)/2,\ h_j)$
  - Kontinuierliche empirische Verteilungsdichtefunktion:

$$h(x) = \begin{cases} 0 & \text{for} \quad x < a_0 \\ \dfrac{h_j}{b_j - a_j} & \text{for} \quad j \in Range(j) \\ 0 & \text{for} \quad b_{m-1} \leq x \end{cases}$$

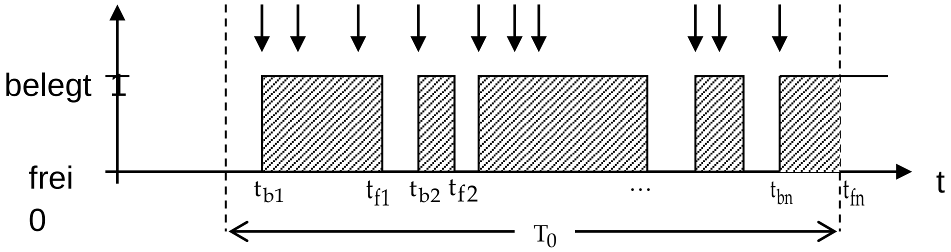
  - Fläche unter Histogramm ist 1.
- Problem: Wahl von m
  - m klein: Informationsverlust, m groß: Form unregelmäßig
  - Lösung
    - Wähle m anfangs groß
    - Aggregiere benachbarte bins um die bin-Werte für m/2 Intervalle zu erhalten
    - Führe Verfahren solange durch bis ein „vernünftiges“ Bild eintritt
- Zähler kann durch geeignete Vererbung für Erfassung von Momenten erweitert werden

1.5.3 Erfassungsobjekte für Momente von zeitkontinuierlichen Messwerten

- „Time-weighting counters", „accumulate"

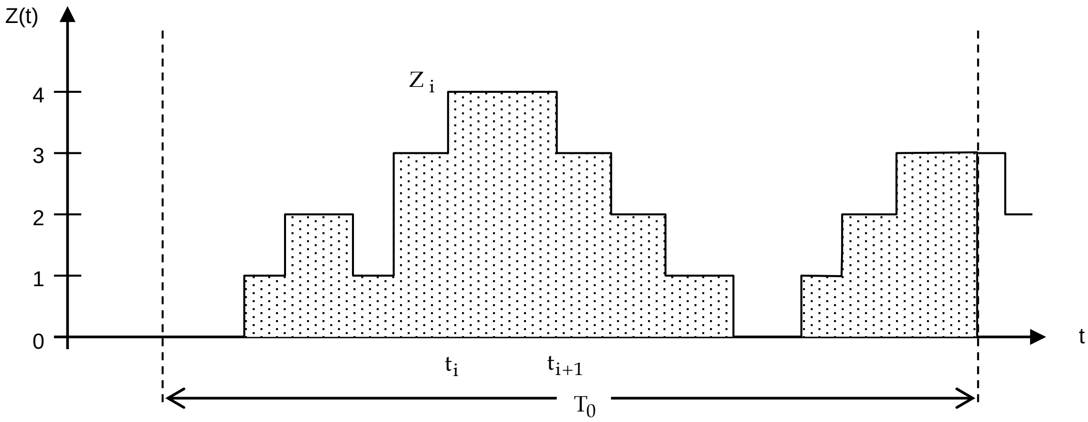
1.5.3.1 Motivation

- Beobachtung zeitkontinuierlicher Prozesse
  - Mittelwert: $\overline{Z} = \frac{1}{T_0}\int_{T_0} Z(t)\,dt$
- Beispiel: Prozessoreinheit
  - Gesuchte Größe: Auslastung $\frac{1}{T_0}\sum_{i=1}^{n}\left(t_{fi} - t_{bi}\right)$



- Beispiel: Warteschlange
  - Gesuchte Größe: Mittlere Warteschlangenlänge

$$\bar{Z} = \frac{1}{T_0} \cdot \sum_{0 \le i < n} Z_i \cdot (t_{i+1} - t_i)$$



1.5.3.2 Implementation of Time-Weighting Counters (TWCs)

- Datenhaltung
  - Summe Σ^k der k-ten Potenzen der beobachteten Messwerte gewichtet mit der Zeit, zum Zeitpunkt $t_n \leq t < t_{n+1}$:

    $$\Sigma^k = \sum_{0 \leq i < n} (X_i)^k \cdot (t_{i+1} - t_i)$$
  - Beginn der Datenerfassung t_{start},
  - Zeit der letzten Stichprobenübergabe t_{last}
  - Letzte übergebene Stichprobe X_{last}
  - Ende der Datenerfassung t_{stop}
- Aktionen zu Beginn der Statistikerhebung zum Zeitpunkt t_0
  - $X_{last} = X_0$, $t_{start} = t_0$, $t_{last} = t_0$
- Aktionen bei Übergabe der Stichprobe X_i zum Änderungszeitpunkt t_i
  - $\Sigma^k = \Sigma^k + (X_{last})^k \cdot (t_i - t_{last})$
  - $X_{last} = X_i$, $t_{last} = t_i$
- Aktionen bei Ende der Statistikerhebung zum Zeitpunkt t
  - $t_{stop} = t$
  - $\Sigma^k = \Sigma^k + (X_{last})^k \cdot \left(t_{stop} - t_{last}\right)$
- Berechnung der Momente
  - $\overline{\mathrm{X}^k} = \frac{\Sigma^k}{t_{stop} - t_{start}}$
  - Darauf basierend Berechnung weiterer statistischen Größen, z.B. der empirischen Varianz nach Gleichung 2.30

1.5.4 Erfassungsobjekte für Verteilungen von zeitkontinuierlichen Messwerten

- Datenhaltung
  - Wie bei Histogramm für diskrete Messwerte: Zähler bin(j) für Intervalle mit Range(j),
  - Wie bei Erfassung von Momenten zeitkontinuierlicher Messwerte: $t_{start}, t_{last}, X_{last}, t_{stop}$
- Aktionen zu Beginn der Statistikerhebung zum Zeitpunkt t_0
  - $X_{last} = X_0, t_{start} = t_0, t_{last} = t_0$
- Aktionen bei Übergabe der Stichprobe X_i zum Änderungszeitpunkt t_i
  - If $X_{last} \in Range(j)$, increment $bin(j)$ by $(t_i - t_{last})$
  - $X_{last} = X_i, t_{last} = t_i$
- Aktionen bei Ende der Statistikerhebung zum Zeitpunkt t_{stop}
  - If $X_{last} \in Range(j)$, increment $bin(j)$ by $(t - t_{last})$
  - $t_{stop} = t$
- Berechnung der relativen Häufigkeiten
  - $h_j = \frac{bin(j)}{t_{stop} - t_{start}}$