

Vorlesungsmaterial

von
Prof. Dr. Horst Völz

Zufall

Dieses Material beruht u.a. auf dem Buch

Völz, H.: Völz, H.: Wissen - Erkennen - Information. Allgemeine Grundlagen
für Naturwissenschaft, Technik und Medizin. Shaker Verlag, Aachen 2001

Es ist neben anderem auch vollständig auf der CD enthalten

Völz, H.: Wissen - Erkennen - Information. Datenspeicher von der Steinzeit bis
ins 21. Jahrhundert. Digitale Bibliothek Bd. 159, Berlin 2007

Der Inhalt wurde erheblich geändert, ergänzt und die Bilder farbig umgesetzt

Bei Angabe der Quelle ist das Material zum privaten Gebrauch voll nutzbar
Bei kommerzieller Nutzung bzw. in Publikationen usw. ist eine Abstimmung mit mir notwendig
Bilder sind in höherer Qualität ca. 2000×3000 Pixel oder *.cdr Version 12 verfügbar

Dieses Material wurde heruntergeladen von
aes.cs.tu-berlin.de/voelz bzw. <http://www.medienwissenschaft.hu-berlin.de/vlz>
Email: [hvoelz\(at\)fpk.cs.tu-berlin.de](mailto:hvoelz(at)fpk.cs.tu-berlin.de) bzw. [h.voelz\(at\)online.de](mailto:h.voelz(at)online.de)
Prof. Dr. Horst Völz, Koppenstr. 59, 10243 Berlin, Tel./Fax 030 288 617 08

Etymologie

Zufall: 15. Jh. zuoval: Anfall, Angriff, das Zuteilwerden; 16. Jh. starkes Andrängen, Zulauf, Parteinahme, Beistand, Beifall, Ereignis, Vorkommnis; 17. Jh. zufallen, zuteil werden

Wichtige, dazu gehörende Begriffe

Determinismus: Lateinisch *determinare* begrenzen, beschränken, beschließen, *de-* von, weg; *terminus*, Ziel, Ende, Grenzzeichen, Grenze, Termin

Häufigkeit: von Haufen $\approx$ viel, Menge

Kausalität: Lateinisch *causalitas* Ursächlichkeit

Korrelation, Lateinisch *correlatio* Wechselbeziehung, *con-* mit; *relatio* Bericht(erstattung), *relativus* bezüglich

Probabilistik: Lateinisch *prohabilitas* bezieht sich auf Urteile, Aussagen: Verb *probare*; Adjektiv *probus* = gut, ehrlich; Auch **Probe**, englisch to prove, to probe

Rationalismus: Lateinisch *ratio* Rechnung, Berechnung, Rechenschaft, Denken, Vernunft, Grund, Maß, gesetzmäßig, Ordnung, Methode, Prinzip; *rationalis* vernünftig, vernunftgemäß, mit Vernunft begabt, schließend, folgernd, berechenbar

Risiko: Italienisch *rischio*, spanisch *risco* Klippe

Statistik: Lateinisch *statisticus* staatswissenschaftlich

Stochastik: Griechisch *stochazesthai* mutmaßen, vermuten

Wahrscheinlichkeit: von wahr und scheinen, $\rightarrow$ Wahrsager

Voraussetzungen

Aussagen betreffen Objekte $\approx$ Systeme

Ereignisse $\approx$ Geschehen $\approx$ Veränderungen: geschehen oder werden durch Handlungen bewirkt

Zeit nur vorher (A) und nachher (B), kein Zeitmaßstab erforderlich

Für $A \rightarrow B$ kann es, aber muss es nicht Ursachen oder Gesetze geben

B folgt aus A $\Rightarrow$ Gesetz $\neq$ Kausalität

Ursache kann auch etwas Anderes sein

Es kann auch Nebeneinflüsse C geben

Menschliches Handeln führt zu speziellen Varianten

Beachten: Vergangenheit steht fest, meist nicht zu ergründen Zufall oder Determinismus

Naturgeschehen, Abfolge geht auf andere Ursachen zurück			
1	$A \Rightarrow B$	Tag $\rightarrow$ Nacht Jahreszeiten	A ist <i>nicht</i> Ursache von B, Rückwärts folgern eindeutig
Naturgeschehen, Abfolge geht überwiegend auf eine Ursache zurück			
2	$A \Rightarrow B$ $C \Rightarrow B$	Blitz $\rightarrow$ Donner	A ist <i>eindeutig</i> Ursache B, B ist anders kaum möglich, z.B. Theater, Schallaufzeichnung
Naturgeschehen, Abfolge kann einer Ursache entsprechen			
3	$A \Rightarrow B$ $A \Rightarrow C$	Wolke $\rightarrow$ Regen	A <i>kann</i> zu B führen, B <i>setzt</i> rückwärts A voraus
4	$A \Rightarrow B$ $C \Rightarrow B$	Regen $\rightarrow$ Pfütze	A <i>kann</i> , aber muss nicht B bewirken, B kann <i>auch anders</i> entstehen, z.B. Rohrbruch, Wasser verschüttet
Naturgeschehen, Ergebnis der Abfolge setzt Ursache voraus			
5	$A \Rightarrow B \Rightarrow C$	Knospe $\rightarrow$ Blüte $\rightarrow$ Frucht	A kann B und dann C bewirken, C setzt A und B, B setzt A voraus
Menschliches Handeln mit absichtlicher Ursache und Wirkung			
6	$A \Rightarrow B$	Klopfen $\rightarrow$ Schall	wenn ich A tue, dann <i>folgt</i> B, B erfolgt wegen A
7	$A \Rightarrow B$	Würfeln $\rightarrow$ 6	wenn ich A tue, dann <i>kann</i> B folgen, B erfolgt wegen A

Zufall_Geschehen.cdr h. vözl 5.3.08

Menschliches Handeln

1. um **Etwas zu bewirken:** Klopfen $\rightarrow$ Schall, Umkehrung möglich

2. **Zufall fordern**, Würfeln. Keine Voraussage, welche Augenzahl, aber **Wahrscheinlichkeit**
Paradoxes Erzwingen: eine „1“ zaghaft eine „6“ kräftig würfeln?!
3. **Absoluter Zufall**: ARISTOTELES (384 – 322 v.Chr.): Beim Lochgraben für eine Pflanze finden wir einen Schatz! Das ist Zufall und Glück.

Beobachter interpretiert das Geschehen und ist fest davon überzeugt

Handlungen sind erklärbar aus dem **Charakter** und bestimmte **Umstände**, $\Rightarrow$ **freier Wille**

Erwartungen können aber enttäuscht werden. SIGISMUND SCHLOMO FREUD (1856 – 1939) $\rightarrow$ Witz und Tragik

Menschen fällt es **schwer an Zufall zu glauben**

Annahme: Es existieren unbekannte oder nicht erkannte Ursachen

ALBERT EINSTEIN (1879 – 1955): „**Gott würfelt nicht!**“

Quantentheorie ist jedoch seit mehr als fünfzig Jahren der absolute Zufall eindeutig und unabänderlich bewiesen

LAPLACE'sche Dämon

Rationales Ursache-Wirkungs-Denken erreichte um 1700 seinen Höhepunkt

1776 PIERRE SIMON LAPLACE (1749 - 1827) die Idee des Dämons

kennt alle **Gesetze** der Natur + zu einem frei wählbaren Zeitpunkt die **Orte** und **Geschwindigkeiten** aller Teilchen

kann dann das gesamte Weltgeschehen der Vergangenheit und Zukunft exakt zu berechnen

(Für ihn) gäbe es dann keine Geheimnisse mehr, keinen freien Willen, keine Verantwortung

Die Welt läuft wie eine Maschine voll determiniert ab

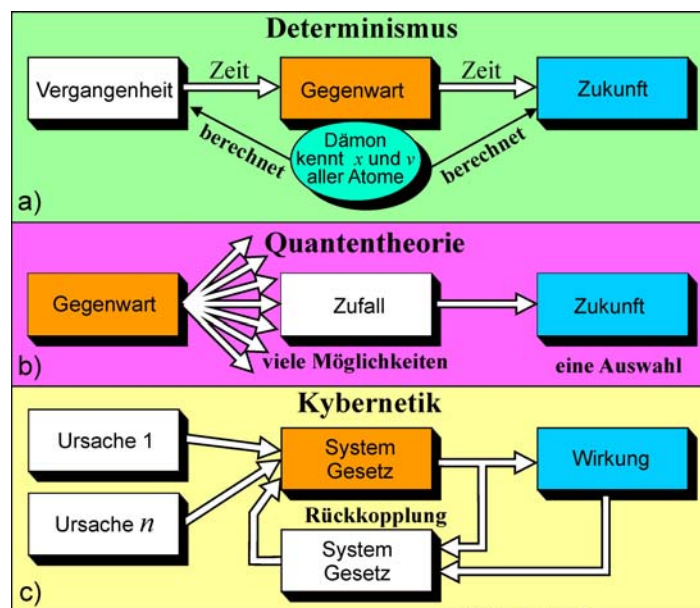
Einleitung zu seiner Wahrscheinlichkeitsrechnung, in Vorlesung von 1795:

„Eine Intelligenz, der in einem gegebenen Zeitpunkt alle in der Natur wirkenden Kräfte bekannt wären und ebenso die entsprechenden Lagen aller Dinge, aus denen die Welt besteht, könnte, wenn sie umfassend genug wäre, alle diese Daten der Analyse zu unterwerfen, in einer und derselben Formel die Bewegungen der größten Körper des Weltalls und die der leichtesten Atome zusammenfassen; nichts wäre für sie ungewiß, und die Zukunft wie die Vergangenheit wäre ihren Augen gegenwärtig.“

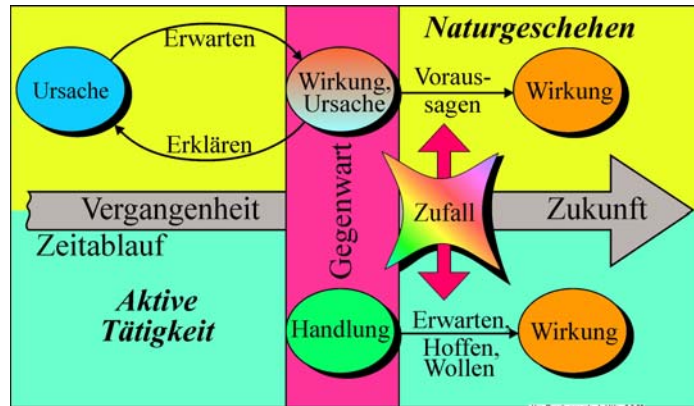
Der menschliche Geist liefert in der Vollkommenheit, die er der Astronomie zu geben wußte, eine schwache Skizze dieser Intelligenz. ... Alle seine Anstrengungen in dem Suchen nach Wahrheit zielen dahin, ihn unaufhörlich jener Intelligenz zu nähern, die wir geschildert haben, aber er wird immer unendlich weit von ihr entfernt bleiben.“

Bedingungen und Folgerungen für die Möglichkeit eines Laplaceschen Dämons

Wenn die Welt	und wenn der Dämon	dann wären
- deterministisch und ständig wäre, - nur wechselwirkende Teilchen enthielte und - die Newtonschen Bewegungsgleichungen uneingeschränkt gültig wären.	- alle Naturgesetze kennen würde, - alle Rand- und Anfangsbedingungen mit absoluter Genauigkeit zu irgendeinem einem Zeitpunkt kennen würde, - alle diese Daten speichern könnte sowie - leistungsfähig genug wäre, um mit den Daten alle Gleichungen exakt und hinreichend schnell lösen zu können.	- der Lauf der Welt in allen Einzelheiten eindeutig bestimmt (gleiche Ursachen erzeugen immer gleiche Wirkungen) und - alle Ereignisse der Vergangenheit und Zukunft korrekt berechenbar.



Kybernetik von NORBERT WIENER (1894 – 1964) zusätzlich Wirkung → Ursache



Etappen des Zufalls

Altertum, Wagenrennen, Kinderspiele, danach große Pause

Neubeginn ab 16. Jh. Versicherungen, Spiele (Risiko), theoretische Grundlagen

Thermodynamik, 19. Jh. Entropie, Statistik, Mittelwertbildung

Quantentheorie 20. Jh. absoluter Zufall

Altertum, v. Chr.

3000 Ägypten Marmeln bei Kindern (Astragalen)

2000 Inder vergnügen sich an Würfelspiel und Wagenrennen

ähnliches gilt für Ägypten, später für **Griechenland** und **Rom**

≈ 800 In der **Ilias** von HOMER (750 – 700) kommen Spiele mit den Knöcheln vor

≈ **600** SOLON beschreibt psychischen Schaden durch **Astragalen-Spiel**

≈ 600 Jaina-Logik **Indien** kennt statistische Hintergründe

600 - 400 Zeitreihen bei den Babyloniern

Platon (427 bis 347 v. Chr.) behandelt in Dialogen subjektive Wahrscheinlichkeit

Aristoteles (384 bis 322 v. Chr.) beschreibt Spiel mit Astragalen

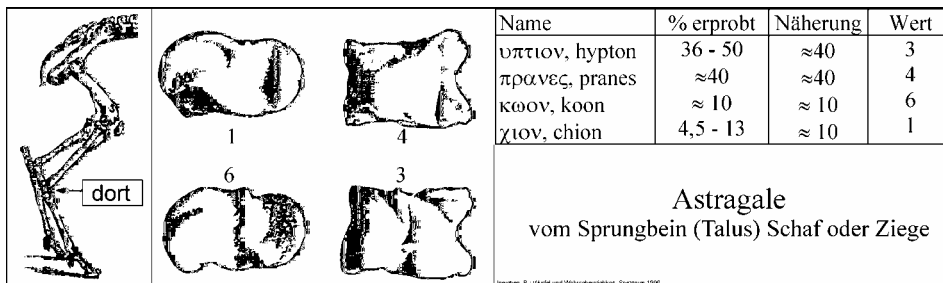
Warum Griechen keine Wahrscheinlichkeitsrechnung entwickelten

Astragale, Würfel und Lose dienen der Erforschung des göttlichen Willens, gehören den Priestern
werden auch als Amulette und Idole getragen

Jeder Astragalos ist etwas anders, daher wahrscheinlich nie „Konvergenz“ beobachtet.

Systematisches experimentieren nicht sinnvoll, wird als künstliche Einschränkung empfunden

Untersuchungen zur Kombinatorik nicht bekannt



Würfel

Sehr früh **Nüsse** des Vibhidaka-Baumes mit 5 Seitenflächen

Später beim **Jakatka-Spiel**: Tetraeder (Holz, Elfenbein, Gold) mit 1 bis 4 belegt

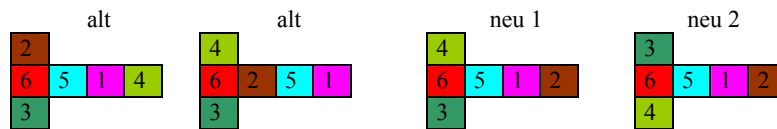
Griechen und Römer: **Astragale** (αστραγαλος) vom Schaf oder Ziege

Heutiger Würfel ermöglicht $6! = 720$ Belegungen.

Gegenüberliegende Zahlen sollen Wert 7 ergeben, nur noch 30 Varianten



jedoch eigentlich nur noch zwei typische (neue):



Neubeginn ab 16. Jh.

Versicherungen, Renten, Risiko, Spiel-Kasinos → theoretische Grundlagen, kompliziert verlaufen

1564 GEROLAMO CARDANO (1501-1576): „Liber de ludo aleae“ (Würfelspiel)

≈ **1600** gebraucht FRANCESCO ALBERTINI Begriff **Ereigniswahrscheinlichkeit**

BLAISE PASCAL (1623-1662): „De aleae Geometriae“

1654 PASCAL-FERMAT-Korrespondenz, PIERRE DE FERMAT (1601-1665)

1657 CHRISTIAAN HUYGENS (1629-1695): „De ratiociniis in ludo aleae“

ab 1660 ist die Rede vom **Wahrscheinlichen** weit verbreitet (**Emergenz**-Phänomen)

1671, 1693 es werden **Tabellen für die Rentenzahlung** erstellt

1708 P. R. DE MONTMORT (1678-1719): „Essay d’analyse sur les jeux de hasard“

1712 ABRAHAM DE MOIVRE (1667-1754): „De mensum sortis“

1713 posthume von JAKOB BERNOULLI (1655-1705): „Ars conjectandi“, gab das vermutlich erste Buch über

Wahrscheinlichkeitsrechnung heraus

1718 ABRAHAM DE MOIVRE: „The doctrine of chances“, führt mit PIERRE SIMON LAPLACE (1749 – 1827) **Normalverteilung** ein

1751 LUDWIG EULER (1707-1783) Calcul de la probabilité dans le jeu de rencontre

1763 TH. H. BAYES: Essay toward solving a Problem in the doctrine of Chances

1773 FRIEDRICH II VON PREUBEN: „sa sacrée Majesté le Hazard“. geheiligte Majestät Zufall

1787 J. G. HERDER benutzt zuerst in einem Brief das Wort „**Wahrscheinlichkeit**“

1800 ca. CARL FRIEDRICH GAUß (1777 – 1885) schuf die Fehlerrechnung

1812 PIERRE SIMON DE LAPLACE „Theorie analytique des probabilités“

1837 S. D. POISSON: Recherches sur la probabilité de jugements en matière ...

1848 CH. GOURAUD: 1. Geschichte der Wahrscheinlichkeitsrechnung

1866 J. VENN Wahrscheinlichkeit als **Grenzwert, Limes**

1919 R. V. MISES Wahrscheinlichkeit als Grenzwert, Limes

1927 H. LAMER: Lusoria Das Spielbrett

1933 A. NIKOLAJEWITSCH KOLMOGOROFF (1903-1987) **axiomatische Wahrscheinlichkeit**

1939 J. HUIZINGA: Homo ludens

Thermodynamik 19. Jh.

- RUDOLF CLAUSIUS (1822 - 1888): mechanische Arbeit ist jederzeit vollständig in Wärme umwandelbar, aber nicht umgekehrt (**Entropie** 1854, $\Delta S = Q/T$). → Zeitrichtung
In geschlossenen Systemen werden Temperaturdifferenzen ausgeglichen. **2. Hauptsatz** (1825 - 1850)
- RUDOLF CLAUSIUS, JAMES CLERK MAXWELL (1831 - 1879), LORD KELVIN und LUDWIG BOLTZMANN (1844 - 1906) entwickelten die **kinetische Gastheorie** ca. 1859 ⇔ physikalischer Ordnung
- 1865 BOLTZMANN schafft mikroskopische, atomistische Erklärung der **Entropie** angeben. $S = k \cdot \ln(W)$
- Auch im (dynamischen) **Gleichgewicht** gibt es immer statistische Schwankungen

Damit **neu in der Physik**:

- **Wahrscheinlichkeit** (W von Mikrozuständen)
- **Irreversibilität**
- **Zeitpfeil**, obwohl kein physikalisches Gesetz ihn enthält = Streben nach dem Gleichgewicht (EHRENFEST-Phänomen)

Hauptsätze Thermodynamik

0.: Streben nach Gleichgewicht, es stellt sich immer nach hinreichend langer Wartezeit ein

Im geschlossenen Systemen gleichen sich Temperaturdifferenzen immer aus

Makroskopische Größen, wie Druck, Temperatur, Volumen usw. ändern sich dann nicht mehr

1.: Energieerhaltung $\Delta U = A + Q$; Änderung der inneren Energie (Zustand)

$U = \frac{3}{2} \cdot N \cdot k \cdot T$, A = Arbeit, Q = Wärme, N = Teilchenzahl; Definition CARNOT-Prozess

Mechanische Arbeit ist jederzeit vollständig in Wärme umwandelbar, aber nicht umgekehrt, Zeitpfeil

Es gibt kein Perpetuum mobile 1. Art, das dauernd Energie erzeugt, ohne sie der Umgebung zu entziehen

2.: KELVIN 1851 = keine periodisch arbeitende Maschine, die ständig Wärme in Arbeit wandelt

Irreversibilität der natürlichen Prozesse, **Wärmemethod**

Perpetuum mobile 2. Art ist unmöglich (nicht beweisbar!)

H-Theorem BOLTZMANN Geschwindigkeitsverteilung der Moleküle → durch Stöße zunehmen bis Entropie = Max.

3.: Am absoluten Nullpunkt $T = 0$ besitzt ein System keine Anregungsenergie mehr

Absoluter Nullpunkt experimentell nie erreichbar. Dort $T \rightarrow 0 \Rightarrow \Delta S \rightarrow 0$

Kritik an BOLTZMANN

1. seine **Atom**-Annahme, die bereits in der Chemie durch JOHN DALTON (1766 – 1844) galt (Atom hat keiner gesehen!)
Widerspruch von ERNST MACH (1838 - 1916) und WILHELM FRIEDRICH OSTWALD (1853 - 1932)
2. **Statistische** Irreversibilität und Ableitung des 2. Hauptsatzes aus reversiblen Mikrozuständen
Widerspruch von HENRI POINCARÉ (1854 - 1912) und ERNST ZERMELO (1871 - 1953)

Führte schließlich zum Freitod von LUDWIG BOLTZMANN 05.09.1906
Heute sind beide Aspekte Fundament der statistischen Physik

MAX KARL ERNST LUDWIG PLANCK (1858 – 1947) stellt dazu fest: Dass sich eine neue wissenschaftliche Wahrheit normalerweise

„nicht in der Weise durchzusetzen pflegt, daß ihre Gegner überzeugt werden und sich, als belehrt erklären, sondern vielmehr dadurch, daß die Gegner allmählich aussterben und daß die heranwachsende Generation von vornherein mit der Wahrheit vertraut gemacht wird“.

Quantentheorie 20. Jh.

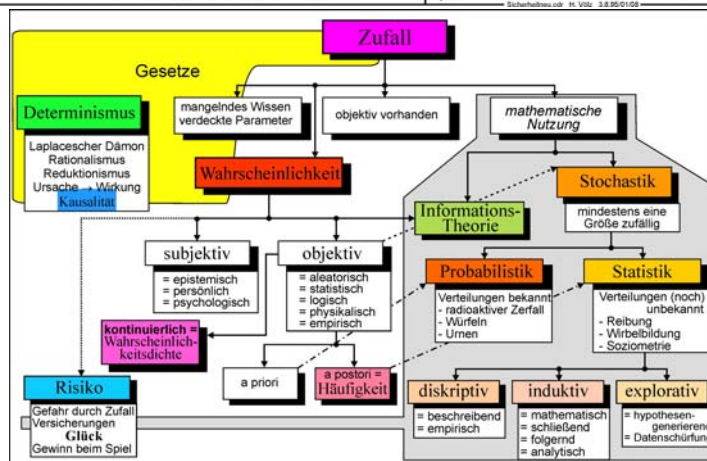
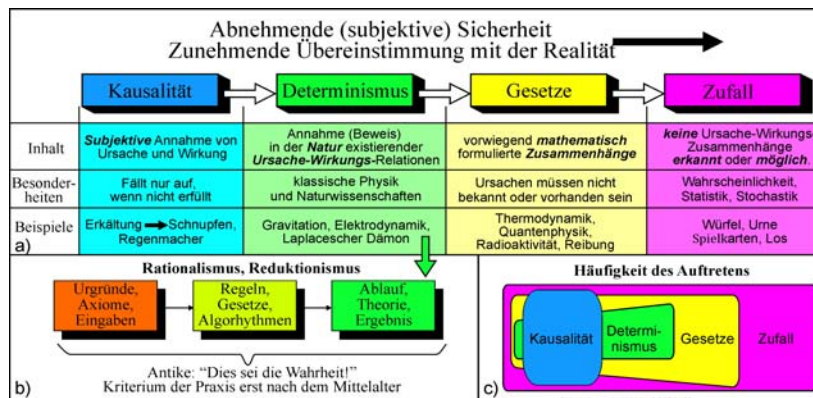
- 14.12.1900 stellt MAX PLANCK Ergebnis bzgl. Strahlung des schwarzen Körpers vor
Energie existiert nur gequantelt. Mittelbar gelangt dadurch Zufall in die Physik ein
Die ersten 3 Rationalismusregeln sind so extrem stark eingeschränkt.
PLANCK: „Formel war es ein Akt der Verzweiflung“
- 1913 NILS BOHR (1885 - 1962): Planetenmodell der Atome, Sprung = Licht, erklärt Spektren
- 1925 WERNER HEISENBERG (1901 - 1975) entwickelt Matrizenmechanik
- 1926 Brief EINSTEIN „Gott würfelt nicht“
- 1926 HEISENBERG formuliert seine Unschärfe-Relation $\Delta x \cdot \Delta p \geq h/2 \cdot \pi$
- 1926 SCHRÖDINGER formuliert Wellengleichung der Quantentheorie
- 1927 Born bestimmt Wellenfunktion der SCHRÖDINGER-Gleichung als Aufenthaltswahrscheinlichkeit
- 1938 22.12. OTTO HAHN spaltet den Uran-Kern

Informationstheorie 21. Jh.

Cloude Elwood Shannon (30.04.1916 - 24.2.2001)

A Mathematical Theory of Communication)

Bell Systems Technical Journal 27 (Juli 1948) S. 379-423 und (Oktober 1948) S.623 - 656. (eingereicht 24.3.1940)



Zufall, zufällig

Betrifft in der **Philosophie** Etwas, das durch den Verlauf äußerer Umstände und innerer Natur der Dinge bedingt und nicht notwendig ist sein und nicht sein kann und das Gegenteil von Notwendigkeit ist, die unbeeinflussbar (obligatorisch) erfolgt

Als **Unberechenbarkeit** oder **Unvorhersagbarkeit**

Wie in Quantentheorie **objektiv vorhanden** und nicht beeinflussbar sein

Weil die **Gesetzmäßigkeiten** (noch) **nicht bekannt** oder in so **großer Vielfalt** vorhanden sind, dass keine Zusammenhänge Ursache → Wirkung erkennbar sind, z. B. Thermodynamik, Reibung und Wetter (verdeckte Parameter?!)

Zufall ≠ Unberechenbarkeit, Unvorhersagbarkeit schwierig abzugrenzen

Besitzt einen relativ engen Zusammenhang mit dem **freien Willen**

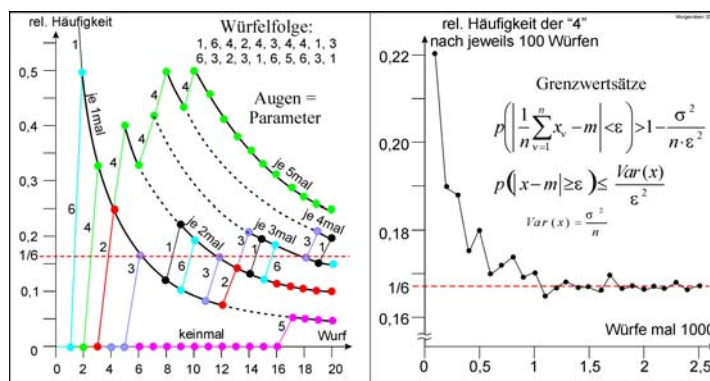
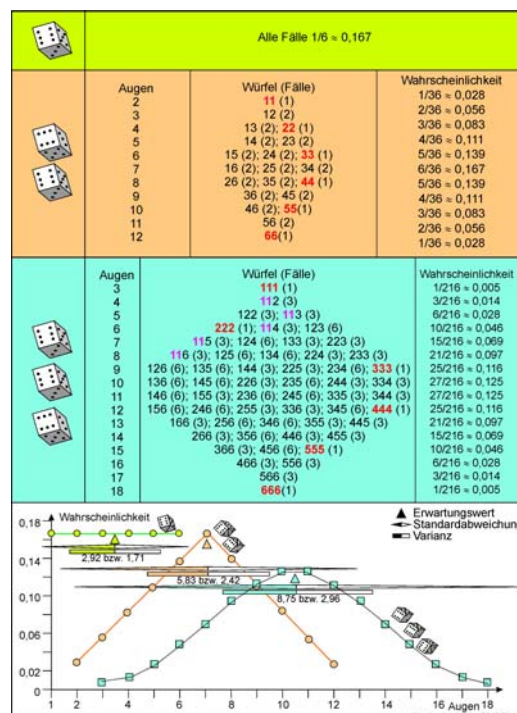
⇒ Gegenteil von eindeutiger Ursache-Wirkungs-Relation = LAPLACE'scher Dämons bzw. Determinismus

Besitzt **kein Gedächtnis**, tritt also immer wieder neu und unabhängig vom Vorher auf

Selbst wenn bereits x-Mal eine „5“ gewürfelt wurde, tritt sie beim nächsten Wurf mit $W = \frac{1}{6} \approx 0,166...$ auf

≠ **bedingte Wahrscheinlichkeit**: 9-Mal hintereinander eine „5“ zu würfeln: $W = (\frac{1}{6})^9 \approx 10^{-7}$

≈ **Urne** Kugeln entnehmen ohne sie wieder zurückzugeben: ist dann nicht mehr dieselbe wie vorher



Im täglichen Leben

vorwiegend ohne irgendwelche Gründe (Ursachen)

Bauer pflanzt einen Baum und findet dabei **zufällig** einen goldenen Mörser

Jemand geht zum Bahnhof, und trifft dort **zufällig** einen alten Freund

Wird nicht benutzt,

Wenn Etwas eintritt, was ausdrücklich vermieden werden sollte

z.B. jemand einen Nagel einschlagen will und sich dabei kräftig auf den Daumen schlägt, gilt eher als ungeschickt

Zufall $\Leftrightarrow$ Chaos

Griechisch *chaino* klaffen, sich auftun, gähnen

Mathematik, Physik = mehr oder weniger zufällige Systemzustände (Chaostheorie)

Sehr kleine Veränderungen der Anfangsbedingungen haben völlig unterschiedliche Auswirkungen

Bewegung eines doppelten Pendels kann völlig unvorhersagbar, chaotisch sein

1961 EDWARD NORTON LORENZ (*1917) bzgl. der Wettergleichungen: Butterfly- bzw. Schmetterlingseffekt

heutiger Flügelschlag eines Schmetterlings in China $\rightarrow$ morgen Orkan in den USA

Chaos in den **Schöpfungsmythen** fast aller Völker und Religionen vor, z.B. Christentum „Die Erde war wüst und leer ...“.

Wenn sich aus so einem chaotischen Anfangszustand Etwas (gesetzmäßig) entwickeln soll, 2 Bedingungen nötig:

- Er muss **reich genug** sein, um von sich aus weitere Strukturen zu erschaffen und
- er muss **arm genug** sein, damit ein Hinterfragen zum Chaos nicht sinnvoll erscheint.

Individueller Zufall

Bei Kindern gilt etwa

5-6 Jahre: Es entsteht eine Intuition für Zufälliges

7 Jahre: Es bildet sich eine Idee für den Zufall aus

11-12 Jahre: Es entsteht ein quantitativer Wahrscheinlichkeitsbegriff, u.a. in Richtung des Gesetzes der großen Zahl und der Wahrscheinlichkeit

13-14 Jahre: Schülern wurde Aussage eines Richters „**Der Angeklagte ist wahrscheinlich schuldig**“ vorgegeben.

Danach akzeptierten sie folgende Aussagen (subjektive Wahrscheinlichkeit s. u.):

Fast, aber nicht ganz gewiss	29 Schüler	52 %
Eher schuldig als unschuldig	15 Schüler	27 %
50:50 = schuldig : unschuldig	8 Schüler	15 %
vollkommen gewiss schuldig	4 Schüler	7 %

Bei **Primitiven**, die weder Schreiben noch Lesen können, ist dennoch ein Gefühl für Zufall, Wahrscheinlichkeit vorhanden

Für Erwachsene

Mit dem Zufall hängen meist auch **Erwartungen, Hoffnungen und Sorgen, Glück – Pech** zusammen.

Er wird schrittweise über **Ereignisse und Fakten** deutlich

Wichtig sind **Werkzeuge**, wie Würfel, Karten (Mischen), Lotterie, Knobeln, Roulett, Urnen usw.

Später wird bemerkt: **Vieles im Leben ist zufällig**, z.B. Geburten, Todesfälle und Unfälle (Schicksal)

Noch später: Zufall lässt sich **statistisch erfassen**

Viel später: Es gibt **vielschichtige Phänomene** mit Wahrscheinlichkeiten, z. B. Wetter, Katastrophen, Wetten; Börse und Spiele

In **frühen Kulturen**: Orakel, Los, Astragale, Götter

Auch die **Glaubhaftigkeit** von Aussagen ist wichtig; Verschwörungstheorien, Astrologie, Wahrsagen

Unvollständiges Wissen wird akzeptiert $\rightarrow$ berechtigt nicht Agnostizismus und Mystizismus

Kaum verständlich **Quantenmechanik** mit grundsätzlichen mikrophysikalischen Zufalls-Prozesse, EINSTEIN nie akzeptiert

Relativ leicht hat sich der Zufall (Wahrscheinlichkeit) für die **Informationstheorie** durchgesetzt

SOKRATES (470 – 399 v. Chr.)

Kämpfte mit „**Ich weiß, dass ich nichts weiß**“

Menschen mit Fachwissen glauben oft unberechtigt auf anderen Gebieten kompetent zu sein

Dafür hat ihn das Orakel von Delphi „**weise**“ genannt.

Für die Unvollständigkeit von Wissen, ging er mit 70 Jahren in den Tod

Das sollten auch heutige Politiker, Richter usw. akzeptieren

ODO MARQUARDT (*1928)

Vier Bücher für philosophisch-menschliche Sicht des Zufalls

Uns widerfährt etwas, was wir nicht gewollt haben

Zur **Würde** des Menschen gehört, dass er das Zufällige leiden kann; und

zu seiner **Freiheit** gehört die Anerkennung des Zufälligen

Frei ist, wer lachen und weinen kann; **Würde hat**, der lacht und weint

Ein Minimum an Chaos ist die Bedingung der Möglichkeit der **Individualität**

Unglück wird durch **Glück** kompensiert (Summe = Null = alle Menschen sind gleich!?)

Feste feiern ist menschlich; und ich glaube, es ist nur menschlich.

Menschen sind nicht absolut, sondern durch den Tod endlich

Wir leben nicht lange genug unser Tod ist stets schneller als die absolute Orientierung;

große oder gar absolute Sprünge sind unmenschlich

DESCARTES „provisorische Moral“: Ohne vorhandene **Üblichkeiten** (Traditionen, Sitten usw.) können wir nicht leben

Stets müssen mehr **Üblichkeiten** aufrechterhalten als verändert werden, Die Beweislast hat stets der Veränderer

Kunst bewältigt (vielleicht) Beliebigkeitskontingenz, die **Religion** bewältigt (vielleicht) Schicksalskontingenz

Alles erzwingt **Gewaltentrennung** nach MONTESQUIEU: Legislative, Exekutive und Jurisdiktion

Zufall $\Rightarrow$ Wahrscheinlichkeit

Englisch probability, likelihood

Zufall widerfährt uns, es kann nur im Nachhinein auf bereits Geschehenes reagiert werden

Er führt uns unsere eigene Begrenztheit vor Augen.

Er wird zuweilen mit Schicksal gleichgesetzt, Wahrscheinlichkeit eröffnet uns den **Übergang zum Möglichen**

Absoluter Zufall bedingt eine völlige **Sinnlosigkeit** der Wirklichkeit ($\Rightarrow$ Quantenphysik)

Relativer Zufall ist meist absichtsloses, regelloses Zusammentreffen von Ereignissen usw.,
die zwar im Einzelnen bedingt, aber insgesamt nicht notwendig sind ($\Rightarrow$ Thermodynamik)

Wahrscheinlichkeit

hängt mit **Erwartungen**, **Hoffnungen** und **Sorgen** zusammen

Ein **allwissender Geist** (LAPLACE Dämon) benötigt **keine Wahrscheinlichkeit**

Sie ist nur notwendig, wenn **unvollständiges** Wissen vorliegt

Dämon muss auch Nichts entscheiden, benötigt keine Logik. Er ist daher so „arm“, dass er nur zu rechnen braucht

Wahrscheinlichkeit $\Rightarrow$ Fragestellungen

Wie groß ist die Wahrscheinlichkeit,

Dass Leben entsteht?	bedingte Wahrscheinlichkeiten
Dass es außerirdischen Leben gibt?	Kombinatorik
Dass man im Lotto gewinnt?	Stabilität der Statistik
Wahrheitsgehalt eines Gerüchts usw.?	Inhaltliche Abschätzung

Gewinnen einzelner Zahlenwerte

Wie lange hält ein Autoreifen?	Messwerte, Experimente, zeitverkürzende Verfahren
Wann tritt ein Speicherfehler auf?	MTBF mean time between failure, Fehlerrate
Wann muß eine Anlage gewartet werden?	Auswahl $\Leftrightarrow$ Betriebssicherheit
Wie kann man die Sicherheit erhöhen?	kalte und heiße Reserve; Reihen-, Parallelschaltung
Wie groß ist der Durchschnittsverdienst?	Bildung von Mittelwerten

Bedeutung von Zahlenreihen

Wie groß ist der Mittelwert, die Streuung usw.?	Statistische Kenngrößen mit/ohne Gewicht/Häufigkeit
Gehorchen sie einem Gesetz?	Korrelation mit einer Gerade oder anderen Funktionen
Welches Modell ist gültig, brauchbar angepasst?	Ermittlung einer Formel, eines Gesetzes
Kann aus einer Zeitreihe etwas gefolgert werden?	Prognose, Rückblick, Beginn

Anwendungen

z.B. Fehler-Korrektur, Mutation, Zuverlässigkeit, Warteschlangen, Bedientheorie, Informationstheorie usw.	
Auftreten, Erkennen und Korrigieren von Fehlern	Zulässige und mögliche Sicherheit
Wann ist ein Fehler zu erwarten?	MTBF usw.
Wie viel Ärzte benötigt eine Unfallstation?	Untätigkeit der Ärzte $\Leftrightarrow$ Warten der Patienten
Zugdichte im Nahverkehr	Leere Plätze $\Leftrightarrow$ Stehplätze, nicht beförderte Reisende
Größe einer Vermittlung	Wartezeit eines Teilnehmers auf freien Anschluss
Was ist für Übertragung und Speicherung notwendig	Entropie, Kanalkapazität

Wahrscheinlichkeit Betrifft Objekte

Es werden meist mess- oder zählbare Daten erhoben

Typen bzgl. der zufällige Auswahl	Beispiele
<i>Eigenschaft</i> aus einer Klasse von Eigenschaften	Mensch Geburt: Junge v Mädel
<i>Ausprägung</i> einer Eigenschaft	Farbe: rot, grün, blau ...
<i>Diskreten Zahl</i> aus einem Zahlenbereich	Würfel: 1, 2, ..
<i>reellen Zahl</i> aus einem Kontinuum	Pegel eines Rauschsignals

U.a. gibt es **2 Arten von Wahrscheinlichkeit**

Lateinisch **aleator** Würfel, Glücksspieler; Griechisch **ont-** das Seiende, Wirkliche, Wahre

Griechisch **epistole**, Lateinisch **epistola** Brief

aleatorisch = statistisch = objektiv $\approx$ kartesisch $\approx$ logisch = physikalisch = empirisch $\approx$ ontisch	epistemisch = persönlich = subjektiv = psychologisch
Wird immer numerisch angegeben, z.B. in % Messen, Zählen Setzt grundsätzliche Regellosigkeit voraus (Quantentheorie) Halbwertszeit von Radioaktivität, chemischen Prozessen usw. Zufallsexperimente und Chancen bei Glücksspielen	Kann verbal benannt werden Schätzen, Behaupten Grad des Vertrauens, Glaubens Gültigkeit einer Meinung Glaubwürdigkeit einer Aussage

verbal ⇔ numerisch

Beispiel experimentell: beim Schwarzfahren erwischt zu werden und Cannabis-Gebrauch

Sehr unwahrscheinlich	eher unwahrscheinlich	Halbe-halbe	eher wahrscheinlich	sehr wahrscheinlich
9,8 %	12,8 %	34,8 %	35,2 %	64,8 %

Beispiele subjektiver Wahrscheinlichkeit

Subjektive Wahrscheinlichkeiten beeinflussen deutlich unser Verhalten

CUBE vermutet: Grossteil des Lernens besteht darin, subjektive Information an objektive anzugleichen

Meist werden kleine Wahrscheinlichkeiten zu groß und große zu klein geschätzt

Vom Gefühl her Unterschied ob bei einer Operation Sterberate 7 % oder Überlebensrate von 93 % angekündigt wird

Lotterie spielen insbesondere bei hohem Jackpot

Wetten, Lotto usw. machen auch **Freude**, verzerren Wahrscheinlichkeit zusätzlich

Wenn ein **Gewinn** lockt werden wir leicht zu Abenteurern, droht **Verlust** werden wir konservativ

Absichtlich genutzt in Spielhallen, z.B. **einarmiger Bandit** in Spielhallen

Meist viele Geräte dicht beieinander, seltener Gewinn aber viel Lärm → fast ständigen Lärm → täuscht oft Gewinne!

Wichtigen Voraussetzungen für **Glück**: Geld, andere Menschen, Selbstverwirklichung usw.

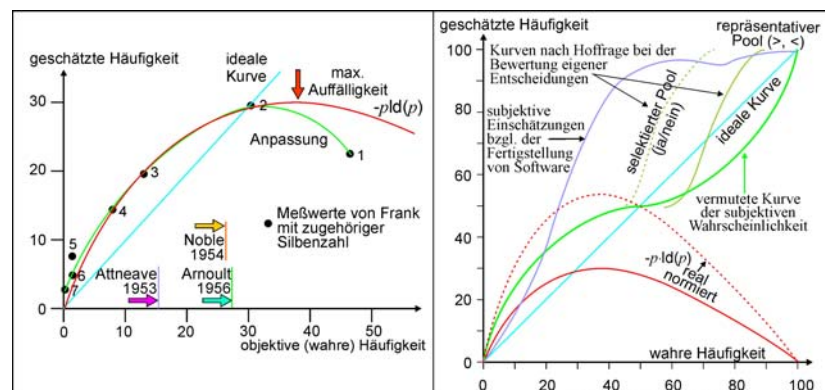
Für persönliches **Unglück**: Krankheit, Arbeitslosigkeit usw.

Widerspruch erklärt: beim Glück wird Gesundheit unterbewusst vorausgesetzt, steht nicht zur Diskussion

Experimentell u.a. HELMAR FRANK (II, 124 ff)

Textseite aus Musils „Mann ohne Eigenschaften“, Probanden mussten je Silbe auf den Tisch klopfen

Anschließend wurden sie – ohne dass sie vorher das ahnten – nach der Häufigkeit der Silbenzahl im Text befragt



Unsicherheit bei Schätzungen

Es ist frustrierender ein Flugzeug mit fünf Minuten als mit 2 Stunden zu verpassen.

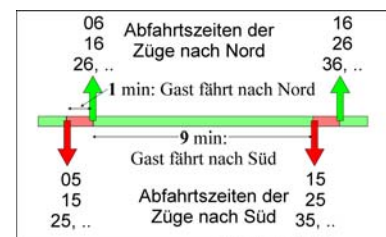
Wer eine „1“ würfeln will, würfelt vorsichtig, bei „6“ mit großem Kraftaufwand.

Beim Zusammentreffen von 24 Personen besteht 50 % Wahrscheinlichkeit → 2

Teilnehmer am gleichen Tage Geburtstag

Witz: Bombe im Flugzeug 10^{-3} . Besser: nehmen sie Bombe mit, 2 Bomben nur 10^{-6}

U-Bahn-Beispiel: 2 Freundinnen: Richtungsauswahl 1:1 gleichviel Züge, dennoch 1:5 für eine Richtung?!



Sichere Methode in Monte Carlo Geld zu gewinnen

Spiel auszuwählen, mit der Wahrscheinlichkeit $\frac{1}{2}$ zu gewinnen, also etwa rouge et noir

Einsatz = 10 €. Bei Gewinn erneut 10 € setzen, bei erneutem Gewinn wieder 10 € usw. Beim ersten Verlust Casino verlassen.

Täglich wiederholen Mögliche tägliche Verlust = 10 €, mögliche tägliche Gewinn liegt aber bei 10, 20, 30 usw. €

Ankereffekt

Für **Propaganda**, z. B. 2. Golfkrieg von 1991 Bush-Administration:

Offiziell Todesopfer unter der irakischen Zivilbevölkerung meistens auf 2, 3 oder 12 benannt

Natürlich wurden diese Zahlen von vielen nicht geglaubt → vermutet wurde vielleicht ein Faktor 100, kaum 1000

Real waren es rund 10 000, wurde erst Monate nach Kriegsende zugegeben

Funktioniert auch umgekehrt, z. B.: zum **Erreichen der Zustimmung** zu kleinen Delikten

U. a. Polizeibehörden und argwöhnische, eifersüchtige Ehepartner

Dem „Gegenüber“ werden schlimmste Vergehen vorgeworfen

Unter der Flut der Anschuldigungen, der Last angeblicher „Beweise“ gibt das Opfer irgendwann etwas zu
 Natürlich unter der Beteuerung, dass alles andere aber ganz bestimmt nicht wahr sei
 Aber genau dies wenige sollte es gestehen!

Othello-Effekt

Nach SHAKESPEARE-Drama = Beeinflussbarkeit über Szenarien
 Hinreichend lange Geschichte erzählen, die sich irgendwie logisch herzuleiten scheint
 Dabei werden immer wahrscheinlichere Schlussfolgerungen gezogen
 Auch wenn der Anfang ganz unwahrscheinlich ist, scheint das Ende bündig zu sein
 Widerspruch zu: Gesamt-Wahrscheinlichkeit = Produkt der Einzelwahrscheinlichkeiten

Einige **Ursachen für falsche Schätzungen** u.a. [Piattelli-Palmarini]

- **Selbstüberschätzung** u. a. von Experten bei komplizierten Fragen
- **Magische Denken** bei Vermutung einer Korrelation → Suche nach neuen Begründungen
- **Nachträgliche Besserwissen**: Beweisen, warum **Napoleon** hätte wissen müssen, dass er die Schlacht bei Waterloo verliert. Nach diesem Motto erfolgen sogar Entlassungen, und die Betroffenen fühlen sich auch noch schuldig!
- **Eingängigkeit von Ereignissen**: Ein Tod durch Masern wird viel wahrscheinlicher als durch Feuerwerkskörper gehalten
- **Blindheit für Wahrscheinlichkeiten**: kleine Wahrscheinlichkeiten, etwa 8 und 1 % erscheinen etwa gleich groß, ähnlich großen, wie 99% und 99,9%. Ganz extrem bei einem Risiko: entweder als vorhanden oder nicht

Ursachen ≠ Wahrscheinlichkeiten

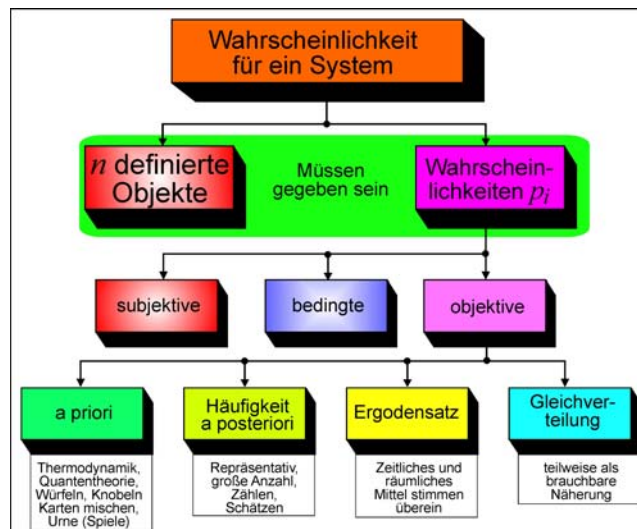
Wahrscheinlichkeiten sind immer symmetrisch, in beide Richtungen gleichwahrscheinlich
 Andernfalls würden vielleicht Störche wirklich die Kinder bringen

- In Deutschland nehmen über die Jahre Kinder und Störche gleichmäßig und hoch korreliert ab
- In ländlichen Gebieten gibt es mehr Störche und mehr Geburten

Oder, wie groß ist die Wahrscheinlichkeit dafür, dass

- die Tochter blaue Augen hat, wenn die Augen der Mutter blau sind?
- die Mutter blaue Augen hat, wenn die Augen der Tochter blau sind?

Meist wird der 1. Fall als wahrscheinlicher bewertet



Varianten der objektiven Wahrscheinlichkeit

1. **a priori**, Lateinisch **prior** früher, erst, vorder, Zahlenwert steht bereits vor dem Versuch fest
 Meist physikalisch bedingt, z. B.: Würfel, Urne, Kartenspiel, Los, Knobeln, Quantentheorie, thermisches Rauschen
2. **a posteriori**, Lateinisch **poster** nachfolgend, eigentlich **Häufigkeit**
 Durch Abzählen, erst nach dem Versuch möglich, **repräsentative** Stichproben wichtig (s.u.)
 Infolge **Gesetz der großen Zahl** nähert sie sich im Mittel nur für stationäre Erscheinungen einem **Grenzwert**
3. **Gleichverteilung**, vereinfachte Betrachtung als grobe Näherung,
 Bei Shannon-Entropie oberer Grenzwert
4. **Ergodensatz** beachten, gleichwertig serielle und parallele Ereignisse
 Zeitlicher Mittelwert muss mit Ensemble-Mittelwert übereinstimmen
 Nicht ergodische Prozesse sind z. B. Sprachen, deren Statistik in der Zeit ändert

Ergodizität bei Markow-Prozessen

- Darf der Graph nicht so in Teile zerfallen, dass man nicht zu allen Knoten zurück gelangen kann
- Muss der größte gemeinsame Teiler aller Zykluslängen (Zahl Kanten in einer geschlossenen Schleife) gleich 1 sein.
- Nichtergodische Quellen, laufen zu Stellen, die Endknoten sind, aus denen der Prozess nicht mehr heraus kann.
 Führt zu Problem der Berechnung der Übergangswahrscheinlichkeiten.

Definitionen von Wahrscheinlichkeit

klassisch	$W(A) = \frac{\text{Anzahl günstiger Elementar-Ereignisse für } A}{\text{Anzahl aller möglichen Elementar-Ereignisse}}$	setzt zumindest statistische Stabilität voraus; Häufigkeiten bringen Probleme
statistisch	Grenzwert der relativen Häufigkeiten für ∞ viele Proben	Die statistische Stabilität notwendig; führt nur zu brauchbarer (zufälliger) Näherung; Grenzwertsätze beachten!
geometrisch	Ereignisse werden als Punkte bzgl. Intervall, Fläche, Volumen aufgetragen. Der relative Anteil entspricht der Wahrscheinlichkeit	ähnelt der klassischen Definition, wenn sie geometrisch veranschaulicht wird
axiomatisch	operiert mit Mengen; 1933 von ANDREJ NIKOLAJE- WITSCH KOLMOGOROW (1903 – 1987) eingeführt	ist ziemlich abstrakt

Definition der axiomatischen Wahrscheinlichkeit

Grundbegriffe	Beispiele dafür (Würfel)
Es gibt Elementar-Ereignisse A_1, A_2 , Sie gehören einem Ereignisfeld $\mathbf{A}$ an Es gibt ein sicheres Ereignis Ω Es gibt das unmögliche Ereignis $\emptyset$	die Augenzahlen beim Würfeln $A = \{1, 2, 3, 4, 5, 6\}$ Es wird eine beliebige Zahl gewürfelt Es wird nichts erwürfelt

Auf $\mathbf{A}$ wird ein **Wahrscheinlichkeitsfeld** $(\mathbf{A}, P)$ bzw. $(\Omega, \mathbf{A}, P)$ mit den folgenden **Axiomen** definiert:

1. Zu jedem Ereignis $A_v \in \mathbf{A}$ gibt es eine reelle Zahl $P(A_v)$ mit $0 \leq P(A_v) \leq 1$, heißt **Wahrscheinlichkeit**
z. B. Würfel: $P(1)=1/6, P(2)=1/6 \dots P(6)=1/6$
2. Die Wahrscheinlichkeit des **sicheren Ergebnisses** ist 1
3. **Additionsaxiom:** $P(A_1 \cup A_2) = P(A_1) + P(A_2)$, z. B. $P(1 \cup 2) = 1/6 + 1/6 = 1/3$

Hieraus werden **alle Sätze** abgeleitet.

Oft noch die **bedingte Wahrscheinlichkeit**: A_1 tritt unter *der* Bedingung ein, dass A_2 bereits eingetreten ist:

$$P(A_1 / A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}$$

Simulation von Wahrscheinlichkeiten

Physikalische Geräte	Würfel, Münzen, Karten, Lose, Urnen, Lose ziehen, Roulett, (Knobeln)
Physikalische Effekte	thermisches Rauschen von Widerständen, Dioden usw. Quanteneffekte haben noch wenig Bedeutung
Zufallstabellen	Heute keine Bedeutung mehr, meist durch mathematische Gesetze ersetzt
Algorithmisch, Mathematische Gesetze	Es gibt Zufallsgeneratoren in vielfältiger Art Ermöglichen aber nur Pseudo-Zufall, (s. u.)

Wahrscheinlichkeitsdichte

Bisherige Betrachtungen gelten nur für **diskrete Merkmale** oder **Ereignisse**

Bei Signalen, Messungen usw. können aber auch **kontinuierliche Werte** auftreten

Hier gibt es nur **Wahrscheinlichkeitsdichten**, die z.B. bei Signalen wie folgt zu bestimmen sind:

Es werden Intervalle Δx bezüglich der Signalgröße x festgelegt

Für eine möglichst große Zeitdauer T_{max} werden die Anteile Δt bestimmt und aufsummiert

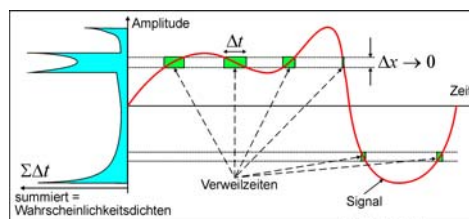
Die sich so ergebenden Werte werden dann für $\Delta x \rightarrow 0$ verfeinert

Genau genommen müsste auch der Übergang $T_{max} \rightarrow \infty$ gebildet werden

Dennoch liegt bei dieser Methode eigentlich eine **a posteriori-Dichte** vor, was repräsentativen Signalabschnitt verlangt

Für eine **apriori-Dichte** müssten theoretische Kenntnisse über das Signal existieren, z. B. MAXWELL-Verteilung

In diesem Fall ist – rein gedanklich – auch eine subjektive Wahrscheinlichkeitsdichte vorstellbar, jedoch kaum nützlich



Rechnen mit Zufallswerten

Überblick, Fachgebiete und Methoden

Rechnen mit zufälligen Werten entstand recht spät.

Im **Altertum** gab es lediglich bei den Sumerern Anfänge, betrafen wiederkehrende Zyklen, wie Mondfinsternisse
 PLATON (427 bis 347 v.Chr.) behandelt in seinen Dialogen so etwas wie eine **subjektive** Wahrscheinlichkeit.

Ab 17. Jh. **Beginn** der Mathematik: Renten, Versicherung gegen Risiko, Spielbanken-Gewinn

Im Folgenden **nur Mathematik** und **nicht Quanten-Physik** und **Informationstheorie** betrachtet

Heute Hauptgebiet

Stochastik (griechisch *stochos*, Vermutung; *stochastike (téchnē)* zum Erraten gehörend(e Kunst) $\approx$ Kunst des Mutmaßens)

Betrifft allgemein: **vom Zufall abhängige Ereignisse und Prozesse:**

Zufallsexperimente, Rentenzahlung, Lotteriegewinne, Meinungsumfragen, Messfehler usw.

Befasst sich auch mit *Kombinatorik*, *Wahrscheinlichkeitstheorie* und *mathematischen Statistik*

Es muss wenigstens eine Größe zufällig sein

z. B. Teilchen der Molekularbewegung $\Rightarrow$ Gesetze über viele Teilchen, z.B. $p \cdot V = \text{const.}$ (p = Druck, V = Volumen)

Zwei Teilgebiete von Stochastik

Probabilistik (lateinisch *probare* erproben, prüfen, untersuchen)

Es müssen die zugrunde liegenden Verteilungen bzw. Wahrscheinlichkeiten bekannt sein

z.B. Würfel, Urne, Los oder Gauß- oder Maxwell-Verteilung bzw. radioaktiver Zerfall usw.

Statistik (lateinisch *statisticum* den Staat betreffend); besonders **umfangreich**

1749 in Deutschland durch GOTTFRIED ACHENWALL (1719 – 1722) eingeführt

Jedoch erst im 19. Jh. von SIR JOHN SINCLAIR in heutiger Bedeutung benutzt

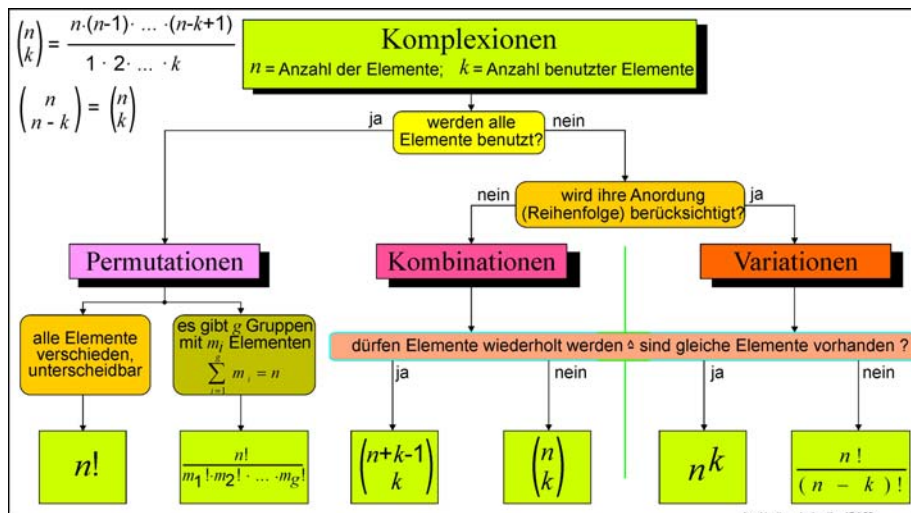
nämlich für das Sammeln und Auswerten von Daten

Verteilungen sind nicht bekannt, z.B. Reibung, Wirbelbildung, Soziologie usw.

Wichtige Verteilungen (Probabilistik)

Unterscheiden: **diskrete** $\Leftrightarrow$ **kontinuierliche** Verteilungen

Zuweilen ist es vorher notwendig die Anzahl der Klassen zu bestimmen



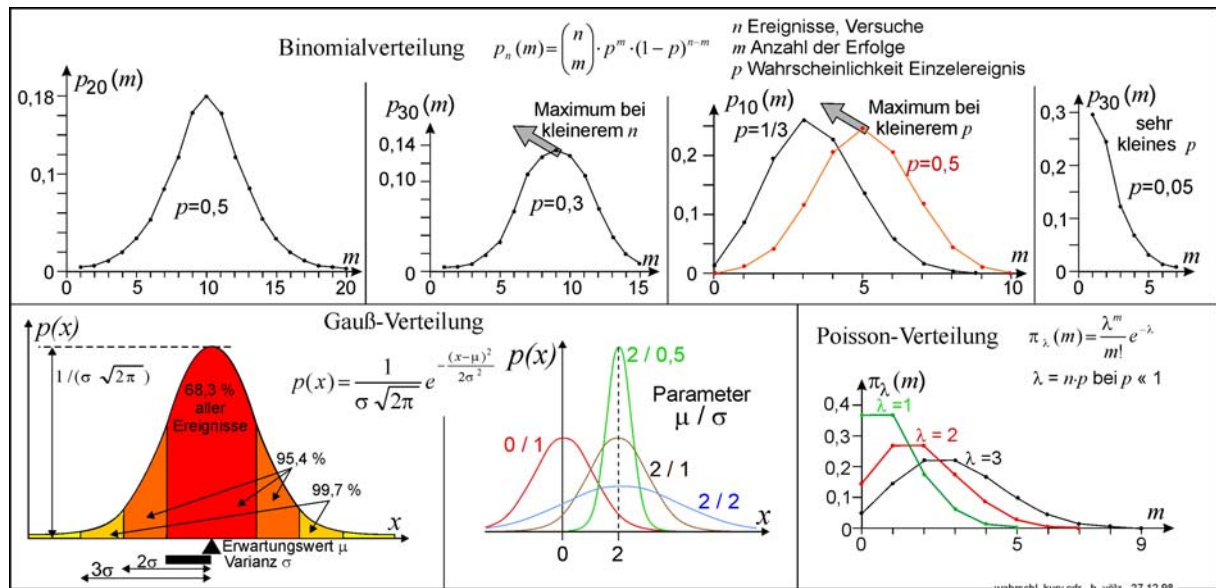
Diskrete Wahrscheinlichkeits-Verteilungen

Poisson	Binomial	Hypergeometrisch
n -stufig = n Kugelsorten Grenzfall der Binomial-Verteilung mit $n \gg 1$; $p \ll 1$; $n \cdot p = \lambda > 0$	Urne mit zwei Sorten Kugeln, p für Weiß, $(1-p)$ für Schwarz; beide als Einzelwahrscheinlichkeit bei n gezogenen Kugeln; Zahl der dabei gezogenen weißen Kugeln $x = 1, 2, \dots, n$	
	Kugeln zurücklegen	Kugeln nicht zurücklegen
$\pi_\lambda(x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}$	$P_n(x, p) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$	$H_n(x; N, M) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}$
		Urne: M weiße; $N-M$ schwarze Kugeln
$\bar{x} = \lambda$; $\sigma^2 = \lambda$	$\bar{x} = n \cdot p$; $\sigma^2 = n \cdot p \cdot (1-p)$	$\bar{x} = n \cdot \frac{M}{N}$; $\sigma^2 = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{M - n}{N - 1}$

Weitere **Sonderfälle** ohne größere Bedeutung, u.a. Bronstein S. 698ff., Göhler, S. 98; Beyer u.a. S. 61 ff.

- Null-Eins-Verteilung; $n = 1$ Binomial-Verteilung; z.B. Münzwurf
- Negative Binomial-Verteilung

- Geometrische Verteilung



Physikalische Statistik im Urnenmodell

Es sind n Kugeln auf N Urnen verteilt mit $n < N$. Es gilt $\binom{n}{k} = \binom{n}{n-k} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{1 \cdot 2 \cdot \dots \cdot k}$

Jede mögliche Aufteilung gilt als ein Elementarereignis.

	MAXWELL-BOLTZMANN	BOSE-Einstein	FERMI-Dirac
	Alle Kugeln sind unterscheidbar	Kugeln sind nicht unterscheidbar	Kugeln nicht unterscheidbar, in jeder Urne nur eine Kugel (PAULI-Prinzip)
Anzahl der möglichen Elementar-Ereignisse	N^n	$\binom{N+n-1}{n}$	$\binom{N}{n}$
Wahrscheinlichkeit für:			
in der i -ten Urne sind r_i Kugeln	$\frac{n! \cdot N^{-n}}{r_1! \cdot r_2! \cdot \dots \cdot r_N!}$	$\frac{1}{\binom{N+n-1}{n}}$	$\frac{1}{\binom{N}{n}}$
in n ausgewählten Urnen befindet sich je eine Kugel	$n! \cdot N^{-n}$	$\frac{1}{\binom{N+n-1}{n}}$	$\frac{1}{\binom{N}{n}}$
in beliebigen n Urnen je eine Kugel	$\binom{N}{n} \cdot n! \cdot \frac{1}{N^n}$	$\frac{\binom{N}{n}}{\binom{N+n-1}{n}}$	1

kontinuierliche Physik

Maxwell: Thermodynamische Geschwindigkeiten

Bose-Einstein für entartete Flüssigkeiten, wie He bei tiefen Temperaturen: kein Pauli-Verbot; nur ganzzahliger Spin

Fermi mit Pauli-Verbot, also keine zwei gleichen Elementarteilchen; auch halbzahliger Spin

$$q(W) = \left(\frac{1}{e^{\frac{W-W_F}{k \cdot T}} + x} \right) \text{ dabeigilt } x = \begin{cases} +1 \\ 0 \\ -1 \end{cases} \text{ für } \begin{cases} \text{Fermi} \\ \text{Maxwell; } W_F = 0 \\ \text{Bose-Einstein} \end{cases}$$

Kontinuierliche Verteilungen

Sehr große Vielfalt, u.a. GAUß-, Exponential-, Boltzmann-, Zipf-Verteilung

gilt für das additive Zusammenwirken vieler unabhängiger, zufälliger Ereignisse $f(x) = \frac{1}{\sigma \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	Normal-Verteilung
gilt für das multiplikative Zusammenwirken vieler unabhängiger, zufälliger Ereignisse $f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi} \cdot x} \cdot e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}$	Log-Normal-Verteilung
gilt u.a. für die Dauer von Telefongesprächen, den radioaktiven Zerfall und die Lebensdauer von Bauelementen $f(x) = \lambda \cdot e^{-\lambda \cdot x}$	Exponential-Verteilung
Verallgemeinerung der Exponential- und Weibull-Verteilung $f(x) = \frac{b^p}{\Gamma(p)} \cdot x^{p-1} \cdot e^{-b \cdot x}$	Gamma-Verteilung
wird vor allem für die Lebensdauer von Bauelementen benutzt, ist durch die Parameter α, β gut anzupassen, geht für $\alpha=1$ in Exponentialverteilung über $\lambda = 1/\beta$ $f(x) = \frac{\alpha}{\beta} \cdot \left(\frac{x}{\beta}\right)^{\alpha-1} \cdot e^{-\left(\frac{x}{\beta}\right)^\alpha}$	Weibull-Verteilung

Gauß- = Normal-Verteilung

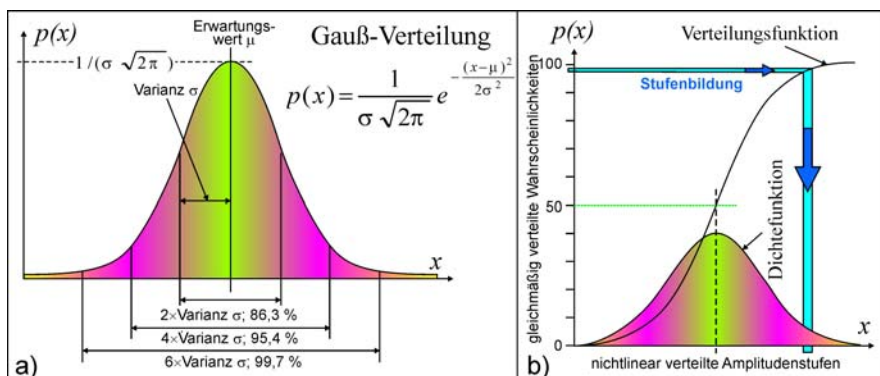
Sie ist besonders wichtig, tritt immer ein, wenn viele unabhängige Zufallsprozesse auftreten, (s. Würfel)
Aus ihrer *Normalverteilung* mit dem Erwartungswertes μ und der Varianz (Streuung) σ^2

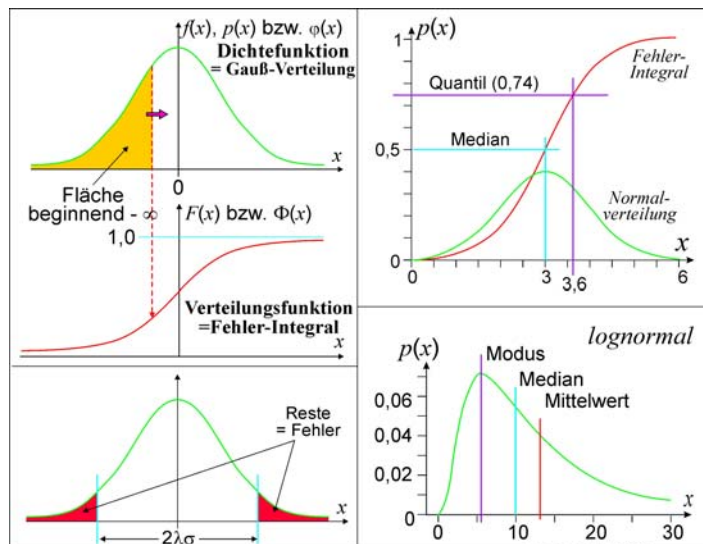
$$f_N(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}$$

Entsteht Durch Integration entsteht die **Fehlerverteilung** = **Fehlerintegral**, ist wichtig für Fehler-Rechnung

$$F_N(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \int_{-\infty}^x e^{-\frac{(z-\mu)^2}{2 \cdot \sigma^2}} dz$$

Die meisten weiteren Kenngrößen sind für beide Funktionen leicht bestimmbar.





Wahrscheinlichkeits-, statistische Maße

Es gibt n Ereignisse als Zahlenwerte $x_i \in \mathbf{A}$ mit den Wahrscheinlichkeiten $P(x_i)$

1. Mittelwerte

Erwartungswert (EX) = **Median** (m) = Zentralwert = Mittelwert = gewogenes arithmetisches Mittel:

$$\bar{x} = \sum_{i=1}^n P(x_i) \cdot x_i \text{ gleichwahrscheinlich } \bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

Geometrisches u. harmonisches Mittel meist nur gleichwahrscheinlich üblich

$$\bar{x}^o = \sqrt[n]{\prod_{i=1}^n x_i} \text{ bzw. } \bar{x}^h = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Quantil Q_p der Ordnung p ; $p=0,5 \rightarrow$ Median

$$P(x_i < Q_p) \leq p \text{ bzw. } P(x_i > Q_p) \leq 1-p$$

Modalwert (Modus) ist der häufigste Wert.

2. Streumaße

Variationsbreite = Spannweite $\delta = x_{\max} - x_{\min}$

Durchschnittliche Abweichung = zentrales Moment 1. Ordnung

$$d = \frac{1}{n} \cdot \sum_{i=1}^n |x_i - \bar{x}|$$

Die **Varianz** (D^2X) = Streuung s = Dispersion

$$s = \sum_{i=1}^n (x_i - \bar{x})^2 \cdot P(x_i) \text{ für gleichwahrscheinlich } s = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2$$

sie kann auch fortlaufend (gleichwahrscheinlich) bestimmt werden

$$s = \sum_{i=1}^n x_i^2 \cdot P(x_i) - \left(\sum_{i=1}^n x_i \cdot P(x_i) \right)^2 \text{ oder } s = \frac{1}{n-1} \cdot \left[\sum_{i=1}^n x_i^2 - \frac{1}{n} \cdot \left(\sum_{i=1}^n x_i \right)^2 \right]$$

Daraus folgt die **Standardabweichung** $\sigma = \sqrt{s}$ als Quadratwurzel aus der Varianz

3. Momente höherer Ordnung

$$\mu_k = \sum_{i=1}^n (x_i - \bar{x})^k \text{ für } k = 2, 3, 4, \dots$$

Speziell folgen daraus **Schiefe** bzw. **Exzeß**

$$\gamma = \frac{\mu_3}{\sigma^3} \text{ bzw. } \varepsilon = \frac{\mu_4}{\sigma^4} - 3$$

Weitere Verteilungen

Maxwell-Verteilung

Betrifft Geschwindigkeiten der Gasmoleküle
Die vorhandene thermische Energie sei:

$$W = \frac{m \cdot v^2}{2}$$

Die mittlere Geschwindigkeit ergibt sich dann zu

$$f(v) = \sqrt{\frac{2}{\pi}} \cdot \left(\frac{m}{k \cdot T} \right)^{3/2} \cdot v^2 \cdot e^{-m \cdot v^2 / 2 \cdot k \cdot T}$$

$$= \frac{2}{\pi} \cdot (k \cdot T)^{3/2} \cdot \sqrt{W} \cdot e^{-\frac{W}{k \cdot T}}$$

Ihr Maximum liegt bei etwa

$$v_{th} = \sqrt{v^2} = \sqrt{\frac{3 \cdot k \cdot T}{2}}$$

Für Zimmertemperatur gilt etwa

Gas	H ₂	N ₂	Cl ₂	Luft
v in m/s	1900	500	330	550

χ^2 -Verteilung

geht auf Astronomen Helmert zurück. Ihre Dichtefunktion lautet:

$$f(x) = \frac{x^{(n/2)-1} \cdot e^{-x/2}}{2^{n/2} \cdot \Gamma(n/2)}$$

Die **Testgröße** ist

$$\chi^2 = \sum_{i=1}^k \frac{(h_i - k_i)^2}{k_i}$$

- Darin sind h_i die Häufigkeiten und den gemessenen Klassen und k_i die theoretischen Häufigkeiten der Klassen.
- Auf diese Art und Weise können zwei Verteilungen bzgl. ihrer statischen Gleichheit verglichen werden.

Student-Verteilung

stammt von WILLIAM S. GOSSET (Pseudonym Student) 1907. Ihre Dichte-Funktion lautet:

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma(n/2) \cdot \sqrt{n \cdot x}} \cdot \left(\frac{1+x^2}{n}\right)^{-(n+1)/2}$$

wobei die Gammafunktion (von Euler)

$$\Gamma(x) = \int_0^{+\infty} e^{-t} \cdot t^{x-1} \cdot dt$$

Für positive ganzzahlige x gilt speziell

$$\Gamma(x) = (x-1)!$$

Anwendung des Tests ein **Beispiel**

- *Ergebnis:* 49 Proben eines Werkstoffes enthalten 2,4 % Si mit $s = 0,4$.
- *Frage:* Ist die Abweichung gegenüber dem Sollwert 3 % zufallsbedingt?
- *Ergebnis:* $t=6,6$ (statt >48) für eine Irrtumswahrscheinlichkeit 5 %

Also liegt eine systematische Abweichung vor.

Forderungen an die Statistik-Aussagen

Objektivität	unabhängig vom Standpunkt (Wünschen) des Erstellers und Auftraggebers (von Zeit und Ort)
Reliabilität	Verlässlichkeit, Genauigkeit (bei geringem zeitlichem Abstand etwa gleiches Ergebnis)
Validität	Überkontextuell gültig (in Bezug auf Hypothese)
Signifikanz	Bedeutend, aussagekräftig, wesentliche Aussagen
Relevanz	Wichtig, dem Anliegen angemessen

Kommentare zur Statistik

Von SIR WINSTON LEONARD SPENCER CHURCHILL (1874 – 1965) stammen:
Die Statistik ähnelt einem Bikini: Sie zeigt fast alles, verhüllt das Wesentliche
Ich glaube nur eine Statistik, die selbst gefälscht habe

Der britische Premier DISRAELI soll gesagt haben: „Es gibt Lügen, dreiste Lügen und Statistik“
 Angenommen, 40 % aller Autounfälle gehen auf Alkohol zurück. Dann verursachen nüchterne Autofahrer 60 % aller Unfälle

Varianten der Statistik

Deskriptive Statistik = auch **beschreibende** = empirische. Es werden **vorliegende Daten** benutzt

Sie werden beschrieben, zusammengefasst, gegliedert, daraus folgen Kennzahlen

Die Ergebnisse werden graphisch bzw. in Tabellen dargestellt

Fast immer wird auf wissenschaftliche Aussagen verzichtet.

Sondergebiete: Fehler-Rechnung, Regression (Interpolation) und Approximation

Induktive Statistik = mathematische = **schließende** = folgernde = analytische = Inferenzstatistik

Daten werden aus einer **Stichprobe** (Erhebung) für die Eigenschaften einer **Grundgesamtheit** abgeleitet

Es gibt strenge Kriterien für die Auswahl der Stichproben = **repräsentativ** für die Gesamtheit

Für die Auswertung existieren **Schätz-** und **Testverfahren**

Die Ergebnisse ermöglichen meist **verallgemeinernde Aussagen**

Erhebung ermittelt Zustände der Welt $\Leftrightarrow$ **Versuch** erforscht Ursachen, Gesetze usw.

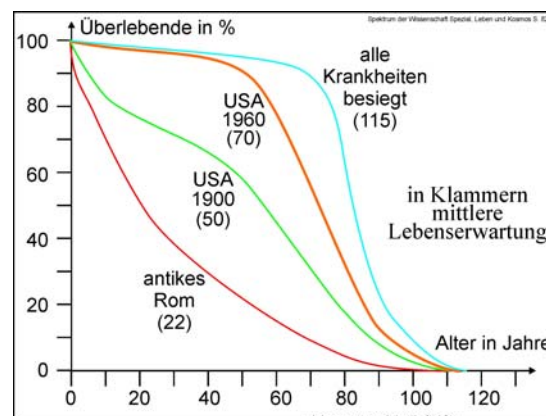
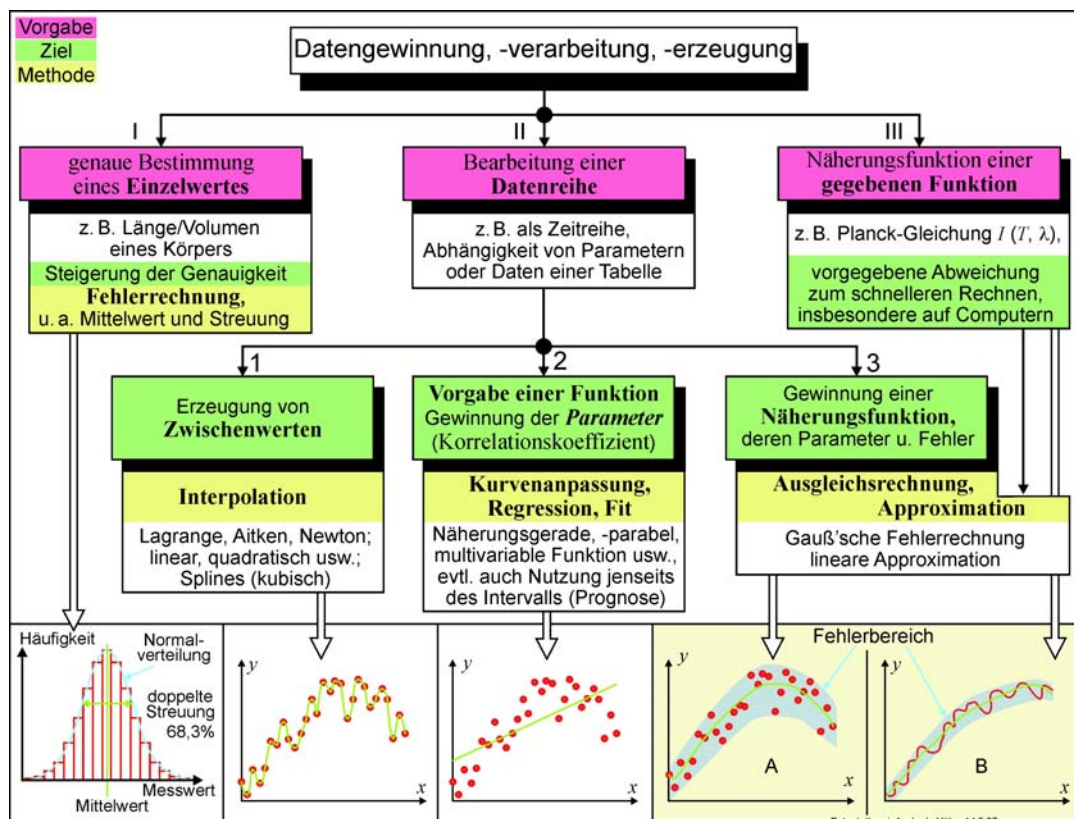
Explorative Statistik = **hypothesen-generierende** = Datenschürfung = data mining

Methodisch eine Zwischenform von deskriptiver und induktiver Statistik

Systematische Suche nach Zusammenhängen oder Unterschieden zwischen Daten in vorhandenen Datenbeständen

Ergebnisse ermöglichen **Hypothesen** für induktive Testverfahren

Mit entsprechenden (prospektiven) Versuchsplanungen können sie **statistisch bestätigt oder abgelehnt** werden



Beispiele für das Zipfsche Gesetz

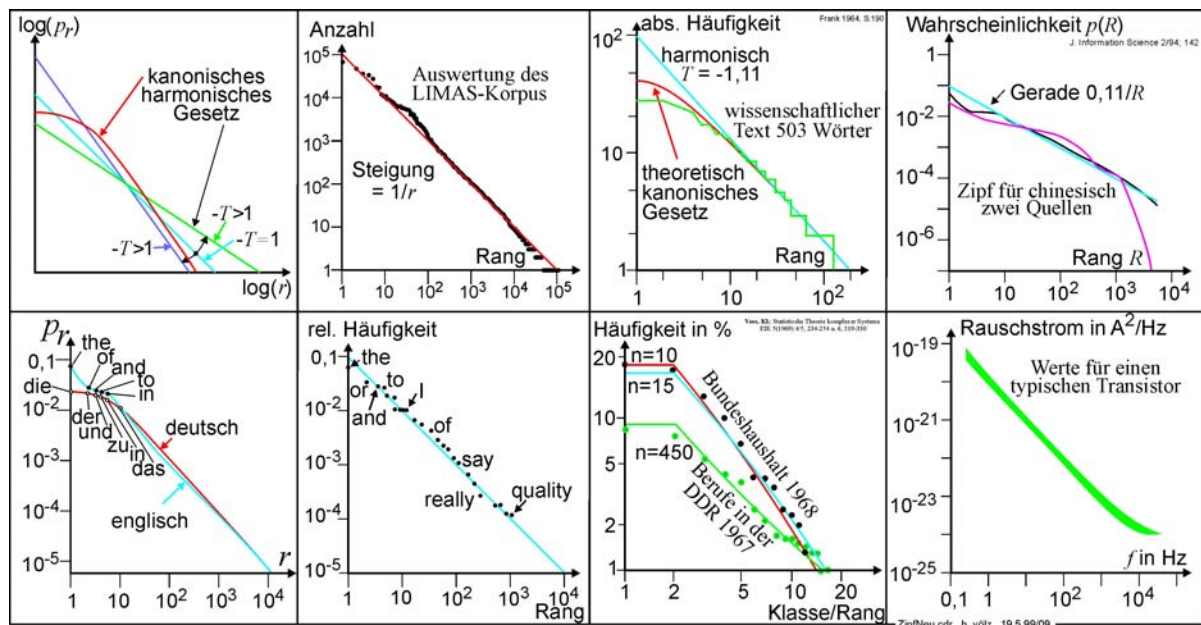
Werden Gesamtheiten in Teile zerlegt, so ergibt sich meist ein Zusammenhang bei der Anordnung nach Rängen
Wurde ursprünglich bei Texten gefunden, gilt aber auch in vielen anderen Anwendungen

Limas-Korpus stammt aus verschiedenen deutschen Texten der Gegenwartssprache von 1970

500 Textstellen mit je etwa 2000 Wortformen aus Trivialliteratur, Zeitungstexte, Romane, Fachaufsätze usw.

Gesamtumfang beträgt reichlich 10^6 , darunter 116 000 verschiedene Wortformen.

Bundeshaushalt 1968	in %	Berufe in der DDR 1967	in %
Verteidigung	22,3	Fachverkäufer	8,2; 9,12
Sozialministerium	20,7	Schlosser	7,5; 9,12
Finanzverwaltung	13,8	Maurer	5,3; 5,58
Verkehr	9,9	Elektromonteur	4,2; 3,99
Ernährung	6,7	Dreher	3,8; 3,11
Bundesschuld	4,1	Agrotechniker	2,5; 2,54
Versorgung	4,0	Stenotypistin	2,1; 2,14
Familien	3,5	Friseur	1,7; 1,85
Wirtschaftliche Zusammenarbeit	2,5	Verkehrsarbeiter	1,6; 1,63
Wissenschaft und Forschung	2,3	Krankenschwester	1,6; 1,45
Innenministerium	2,0	Rinderzüchter	1,5; 1,32
Wirtschaft	1,3	Maschinenbauer	1,4; 1,2
Sonstiges	6,9	Maler	1,3; 1,1
		Koch	1,3; 1,02
		Gärtner	1,0; 0,95
		Technische Zeichner	1,0; 0,89



ZIPF-MANDELBROT'sches Gesetz

Ursprünglich harmonisches Gesetz: 1916 Beobachtungen von ESTOP

$$p(r) \approx \frac{K}{r}$$

p = Wahrscheinlichkeit, r = Wort-Rang, K = Konstante

Anpassung an Spracheigenschaften ermöglicht die **Texttemperatur** T (Steilheit der Geraden)

$$p(r) = K \cdot r^{-1/T}$$

Werte je Textauswahl: Deutsch 0,46; 0,54; 0,66 oder 0,71; Russisch 0,88; Hebräisch 0,82; Englisch 0,72; 0,77 bzw. 0,90

GEORGE KINGSLEY **ZIPF** (1902-1950) ergänzt Formel mit Anzahl R der verschiedenen Wörter

$$p(r) \approx \frac{1}{r \cdot \ln(1,78 \cdot R)}$$

→ SHANNON-Entropie $H = 0,5 \cdot \ln(2 \cdot R \cdot \ln(2 \cdot R)) \rightarrow 12\,000$ verschiedenen Wörter $\rightarrow H \approx 9,300\,000 \rightarrow H \approx 11,5$ Bit/Wort.

BENOIT B. MANDELBROT (*1924) führt zusätzliche Konstante r_0 zum **kanonischen Gesetz** ein

$$r_0 = \frac{L}{L-1}$$

L berücksichtigt den maximalen Rang R und bewirkt eine Korrektur für die kleinen Ränge

$$p(r) = K \cdot (r + r_0)^{-1/T}$$

Für **viele Elemente eines Systems** gewinnt VOSS allgemein mit auffindbaren (ableitbaren) Konstanten Z und γ

$$p(r) = \frac{1}{Z} \cdot \frac{e^{-\gamma r}}{r}$$

M. W. ist nicht geklärt, ob das überall in der Physik, Elektronik usw. beobachtete **1/f-Rauschen** hiermit zusammenhängt

VERHULST- = logistische Gleichung

1845 entwickelte PIERRE FRANCOIS VERHULST (1804 - 1849) = Modell des gehemmten Wachstums

Zuwachs ist mit einer Konstante α proportional zur vorhandenen Anzahl N

$$\frac{dN}{dt} = \alpha \cdot N(t)$$

Integration liefert exponentielles Wachstum mit Startwerten t_0 und N_0

$$N = N_0 \cdot e^{\alpha \cdot (t - t_0)}$$

Meist gibt es jedoch endliche Ressourcen begrenzen Wachstum auf Maximum M

$$\frac{dN}{dt} = a \cdot N(t) \cdot \frac{M - N(t)}{M} = a \cdot N(t) \cdot \left(1 - \frac{N(t)}{M}\right)$$

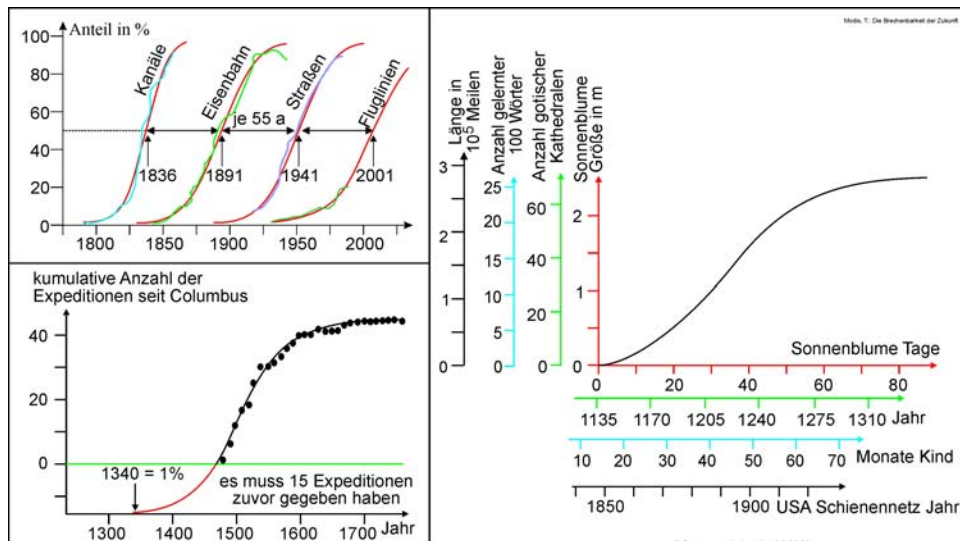
Mit einem Startwert b liefert Integration

$$N(t) = \frac{M}{1 + e^{-(at+b)}}$$

Auflösen nach $N(t)$ und Logarithmieren ergibt S-Kurve des Wachstums

$$\log\left(\frac{N(t)}{M - N(t)}\right) = a \cdot t + b$$

[Modis] enthält viele Beispiele zu **S-Kurven** und Produktionszyklen.



Spezielle Ziffernhäufigkeit

1881 fand Astronom Simon NEWCOMB:

Bei Logarithmentafeln verschmutzen die ersten Seiten, die mit 1 beginnen besonders schnell

Generell fand er, dass Zahlen, die 1 beginnen doppelt so häufig sind wie die mit 2 → NEWCOMBS Formel

1938 fand der Physiker FRANK BENFORD erneut das Gesetz von NEWCOMB an Zahlen über Gebiete, Flusslängen usw.

1961 folgten Untersuchungen von ROGER PINKHAM, die Aussagen wurden unabhängig von Maßeinheiten bestätigt

Generell ergibt sich aus Rechnungen, statistischen Erhebungen usw.

Für die erste Ziffer: $1 \approx 30\%$; $2 \approx 18\%$ langsamer Abfall bis $9 \approx 4,6\%$

Für die zweite Stelle: $0 \approx 12\%$ Abfall bis $9 \approx 8,5\%$

Um 1995 wies THEODORE HILL nach, dass wesentlich **logarithmische Verteilungen** sind

Um 2007 konnte auf dieser Grundlage MARK NIGRINI **Fälschungen bei Patientendaten** nachweisen

U.a. 13 gefälschte Schecks zwischen 6.500 und 6.599 \$ für angebliche Herzoperationen

Theoretisch folgt für die erste Ziffer: $n = 1$ bis $9 \Rightarrow \log(n+1) - \log(n)$. Beachte $\log(1) = 0$ und $\log(10) = 1$

Ziffer	1	2	3	4	5	6	7	8	9
%	30,1	17,6	12,5	9,7	7,9	6,7	5,8	5,1	4,6

Bei der zweiten Ziffer kann auch 0 Null auftreten

Es ergibt sich eine komplizierte Verteilung, welche über alle neun vorangehenden Ziffern zu mitteln ist

Für die Einhaltung des Gesetzes gibt es **zwei Bedingungen**

1. Es müssen mehrere Hundert Zahlen betrachtet werden
2. Es dürfen keine Einschränkungen auftreten, z. B. Körpergröße $\rightarrow$ beginnt immer mit 1

Besonders extrem erwiesen sich bei Untersuchungen von [Bosch] die **Lotto-Zahlen**

Er hat $6,8 \cdot 10^6$ Tippreihen eines Sonntags von 6 aus 49 untersucht:

$19 = 32\%$ und $36 = 26\%$ liegen erheblich über dem Durchschnitt

Es folgen: 9, 7, 17, 10, 11, 18, 25, 3, 32, 12

Am Ende stehen: 43, 35, 29, 44, 42, 47, 22, 15, 14, 49

Leider ist nicht nachprüfbar, wie das Ergebnis wäre, wenn nur die jeweils erste **Ziffer** betrachtet wäre

Weber-Fechner-Gesetz

Der Logarithmus bzw. die Exponentialfunktion kommt sehr oft in (Natur-) Gesetzen vor (besondere Ausarbeitung)

Psychophysik: Gesetz von ERNST HEINRICH WEBER (1795 – 1878) und GUSTAV THEODOR FECHNER (1801 – 1887).

Zwischen der Energie R eines Reizes und der psychologisch empfundenen Intensität W gilt:

$$W = K \cdot \ln(R/R_0) + c$$

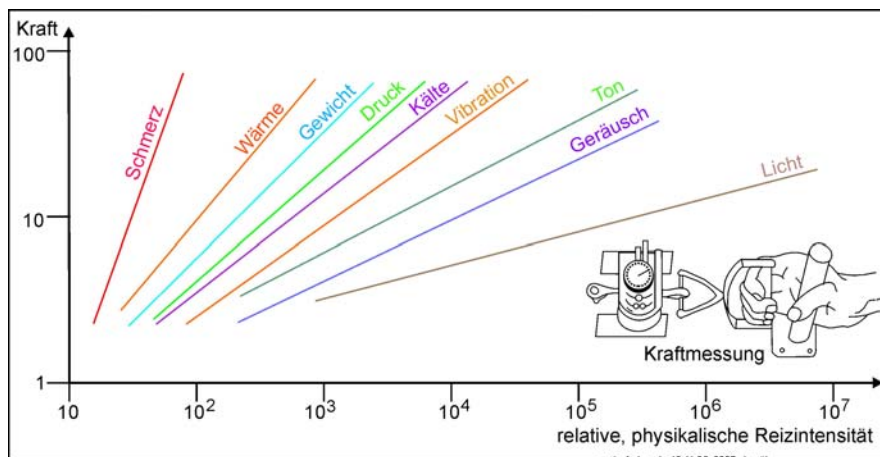
Die Konstanten K und c hängen von der Reizart ab, z.B. Licht, Schall, Wärme oder Kälte

Eine etwas geänderte, genormte Schreibweise mit der Konstanten k_s lautet

$$W_S \approx k_s \cdot (R - R_{0S})^n$$

Für den Exponenten n gilt

Reizart	Exponent n	Dynamikbereich	Unterscheidbare Stufen
Schmerz	2,1	15 dB; 1:8	≈ 5
Wärme	1,6	33 dB; 1:40	≈ 20
Schwere	1,5	24 dB; 1:17	≈ 50
Druck	1,1	20 dB; 1:10	≈ 50
Kälte	1,0	20 dB; 1:10	≈ 10
Vibration	0,95	50 dB; 1:300	≈ 100
Schall	0,6	100 dB; $1:10^5$	400 Lautstärken
Licht	0,33	130 dB; $1:3 \cdot 10^6$	600 Helligkeiten

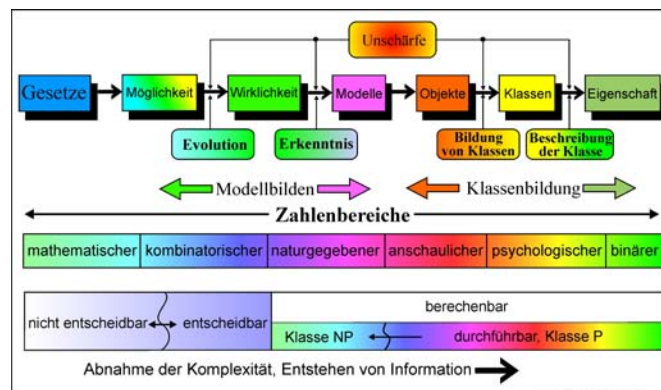
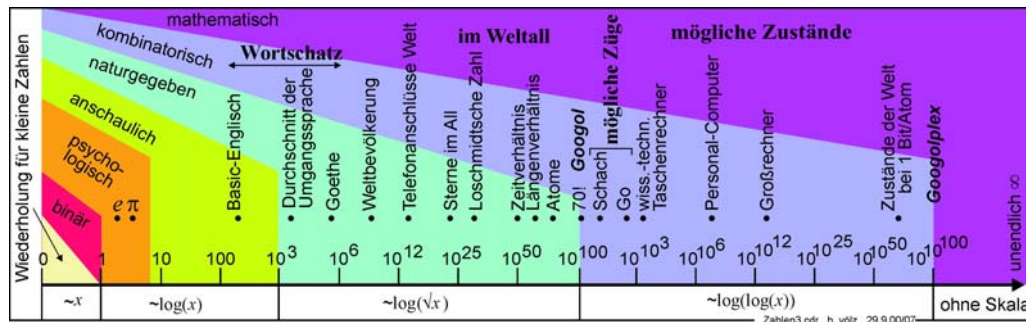


Frage

Was ist Ursache, was ist Wirkung für Logarithmus?

Natur Physiologie $\leftrightarrow$ Physiologie unsere Vorstellungen von der Welt
Bedeutung von Zahlenklassen

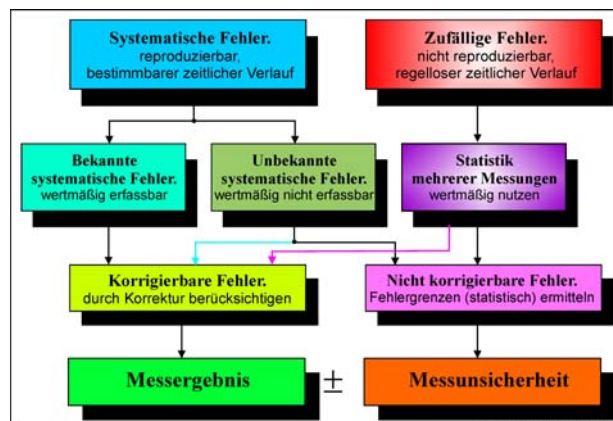
Name	Beispiel für Bezug	Zahlenbereich
Binär	Logik	0, 1
Physiologisch	Gedächtnis	1, 2, ... 7 (± 2)
Anschaulich	Vorstellung, Zählen	ca. 0,001 bis 1000
Naturgegeben	Natur, Kosmos	10^{-100} bis 10^{100}
Kombinatorisch	Spiele, Evolution	noch kleiner bzw. größer
Mathematisch	Theorie	$\varepsilon \rightarrow 0$ bzw. $n \rightarrow \infty$



Auswertung von Messdaten

Durch Messungen oder Erhebungen werden häufig mehrere Messpunkte als Punktwolke erhalten
Die Auswertung kann zu drei Varianten der Anwendung führen:

1. Für **eine Messgröße** kann bei wiederholter Messung der wahrscheinlichste Wert und Messfehler bestimmt werden
Dabei wird die Normalverteilung (Gauß-Kurve) berücksichtigt. So lassen sich grobe Fehler ausschalten (s. o.)
Angabe werden meist **Erwartungswert** = wahrscheinlichster Wert und **Varianz** bzw. Streuung
2. Bei einer **Messreihe** wird oft eine bestimmte Funktion vermutet
Hierfür kann dann eine Anpassung, Regression bzw. Fit gesucht werden
die Korrelation gibt dann die Sicherheit der Anpassung an
3. Die Daten von einer Messreihe werden benutzt um eine **Hypothese** zu bestätigen oder abzulehnen.



Regression – (Kurven-) Anpassung – Fit (-ting) - Smoothing

Betrifft drei Anwendungen:

- Der mathematische **Zusammenhang** wird als **bekannt** vorausgesetzt
Durch Mittelung werden die Parameter für den Zusammenhang bestimmt, typische Regression für $y = a + b \cdot x$
Die „Güte“ wird dann häufig über den Korrelationskoeffizient ρ^2 im Sinne eines Messfehlers bestimmt (s. u.)
- Es sind **unterschiedliche Zusammenhänge**, mathematische Funktionen möglich
Über die Korrelationskoeffizienten (oder andere Maße) wird entschieden, welche Funktion wahrscheinlich vorliegt
- Möglichkeit der **Extrapolation** über den gemessenen Bereich hinaus
So sind Voraussagen/Prognosen möglich. Auch Startpunkte, Vergangenheit sind erreichbar (s. Erkundungen)
Anwendungen häufig Ökonometrie und Zinseszinsrechnung.

Die drei Haupt-Varianten der Kurven-Anpassung

Linear	$y = a + b \cdot x$
Nichtlinear	$y = a + b \cdot x + c \cdot x^2 + d \cdot x^3 + e \cdot x^4 \dots$ oder allgemein $y = f(x)$
Mehrere Variablen	$y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + \dots$ oder allgemein $y = f(x_1, x_2, x_3, \dots)$

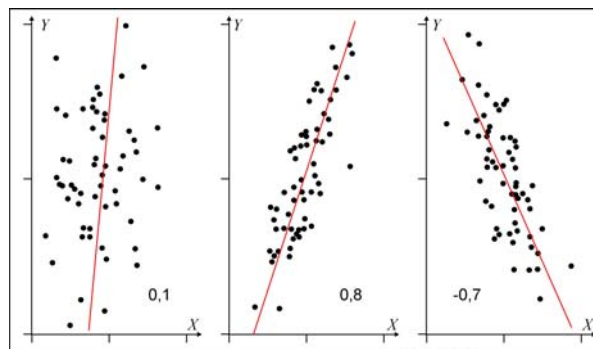
Lineare Variante

Sie ist besonders einfach, da nur zwei Konstanten a und b und der Korrelationskoeffizient r (ρ^2) zu bestimmen sind:

$$a = \frac{\sum y \cdot \sum x^2 - \sum x \cdot \sum x \cdot y}{n \cdot \sum x^2 - [\sum x]^2} = \frac{1}{n} \cdot (\sum y - b \cdot \sum x) \quad \text{und} \quad b = \frac{n \cdot \sum x \cdot y - \sum x \cdot \sum y}{n \cdot \sum x^2 - [\sum x]^2}$$

$$r = \frac{n \cdot \sum x \cdot y - \sum x \cdot \sum y}{\sqrt{(n \cdot \sum x^2 - [\sum x]^2) \cdot (n \cdot \sum y^2 - [\sum y]^2)}}$$

Es sind die Summen für 5 Größen zu bilden: $\sum x$; $\sum y$, $\sum x \cdot y$; $\sum x^2$ und $\sum y^2$



Für Messungen an Studenten ergaben sich bei [Stoyan] folgende Korrelationen

	Größe	Gewicht	Bauchumfang	Schuhgröße	Handgröße	Oberschenkelumfang
Größe	1	0,74	0,32	0,86	0,27	0,28
Gewicht		1	0,57	0,54	0,13	0,68
Bauchumfang			1	0,25	-0,14	0,50
Schuhgröße				1	0,13	0,23
Handgröße					1	0,44
Oberschenkelumfang						1

Korrelationen müssen keinen **ursächlichen** Zusammenhang ausweisen

Beispiel (s. o.) **Störche** $\Leftrightarrow$ **Kinder**, Achtung Medizin und Pharmazie

Umformung von Formeln zu linearer Regression

logarithmisch	$y = a + b \cdot \ln(x)$	$x' = \ln(x)$
exponentiell	$y = a \cdot b^x$	$y' = \ln(y) = a' + b' \cdot x$ $a' = \ln(a); b' = \ln(b)$
geometrisch	$y = a \cdot x^b$	$y' = \ln(y) = a' + b' \cdot \ln(x)$ $a' = \ln(a)$
trigonometrisch	$y = a + b \cdot \sin(x)$	$x' = \sin(x)$
reziprok	$y = a + b/x$	$x' = 1/x$
Texttemperatur T	$p(R) \approx k \cdot R^{-T}$ Beachten: R beginnt bei 0	$x = \log(R+1); y = \log(p(R+1))$ $T = -b$

Fehlerrechnung = Methode der kleinsten Quadrate = GAUß'sches Fehlerquadrat oder L_2

CARL FRIEDRICH GAUß (1777 – 1855)

Approximation erfolgt mit vorgegebenen Funktionen $\Phi_m(x)$. Bestimmt werden deren Koeffizienten (Parameter) c_m

Vereinfachung der Schreibweise: $\Phi_{m,n} = \Phi_m(x_n)$ und $f_n = f(x_n)$ eingeführt. Für Fehler gilt:

$$E^2(c_0, c_1, \dots, c_{M-1}) = \sum_{n=0}^{N-1} \left[f_n - \sum_{m=0}^{M-1} c_m \cdot \Phi_{m,n} \right]^2$$

Ihre Minimierung erfolgt durch Nullsetzen der partiellen Ableitungen nach den c_k , getrennt für alle $k = 0$ bis $M-1$:

$$\frac{\partial E^2}{\partial c_k} = -2 \sum_{n=0}^{N-1} \left[f_n - \sum_{m=0}^{M-1} c_m \cdot \Phi_{m,n} \right] \cdot \Phi_{k,n} = -2 \sum_{n=0}^{N-1} f_n \cdot \Phi_{k,n} + 2 \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} c_m \cdot \Phi_{m,n} \cdot \Phi_{k,n} = 0$$

Übersichtlicher ist die Umformung in die Schreibweise

$$\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} c_m \cdot \Phi_{m,n} \cdot \Phi_{k,n} = \sum_{n=0}^{N-1} f_n \cdot \Phi_{k,n}$$

Ermöglicht unmittelbare Überführung in die Matrixschreibweise

$$\begin{bmatrix} \sum_{n=0}^{N-1} \Phi_{0,n} \cdot \Phi_{0,n} & \dots & \sum_{n=0}^{N-1} \Phi_{0,n} \cdot \Phi_{M-1,n} \\ \vdots & \ddots & \vdots \\ \sum_{n=0}^{N-1} \Phi_{M-1,n} \cdot \Phi_{0,n} & \dots & \sum_{n=0}^{N-1} \Phi_{M-1,n} \cdot \Phi_{M-1,n} \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ \vdots \\ c_{M-1} \end{bmatrix} = \begin{bmatrix} \sum_{n=0}^{N-1} f_n \cdot \Phi_{0,n} \\ \vdots \\ \sum_{n=0}^{N-1} f_n \cdot \Phi_{M-1,n} \end{bmatrix}$$

Werden orthogonale Funktionssysteme gewählt, so existieren nur die Diagonalglieder

Für orthonormierte nehmen sie sogar den Wert 1 an. Dann ergeben sich für die c_k die einfachen Lösungen

$$c_k = \frac{\sum_{n=0}^{N-1} f_n \cdot \Phi_{k,n}}{\sum_{n=0}^{N-1} \Phi_{k,n}^2}$$

Für **orthogonale Funktionen** Φ_k muss gelten

$$\int_a^b \Phi_{k1} \cdot \Phi_{k2} = \begin{cases} 0 & \text{für } k_1 \neq k_2 \\ w^2 & \text{für } k_1 = k_2 \end{cases}$$

Orthonormiert, sind sie dann, wenn $w^2 = 1$ gilt

Beispiele sind Sin- + Cos-, LEGENDRE- und TSCHEBYSCHEFF-Funktionen

Versuchsplanung (Stichproben), Testmethoden (Hypothesen), Schätzverfahren

Beispiel Brustkrebs

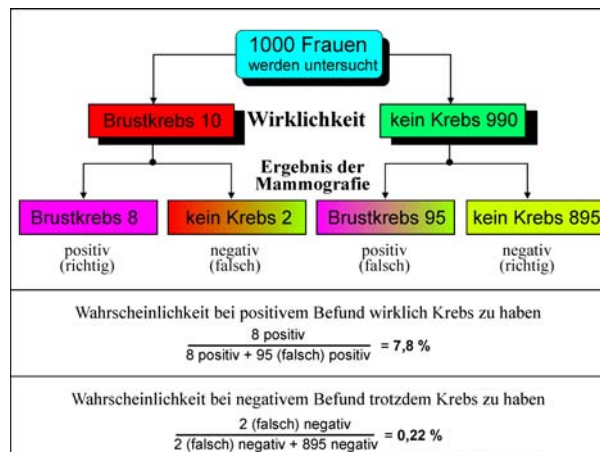
Stimmt die Diagnose Brustkrebs? Daten von 2001. Es gibt immer unvermeidliche Fehler bei Mammografie

Von 1000 Patientinnen haben im Mittel 10 Krebs. Nur bei 8 von ihnen fällt die Diagnose positiv aus

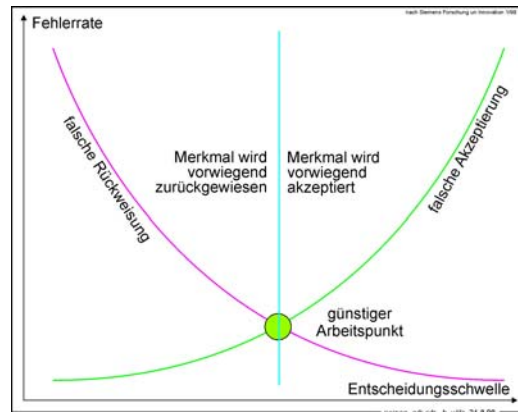
Andererseits erhalten auch 95 gesunde Frauen einen positiven Befund

103 Frauen erhalten ein beunruhigendes Ergebnis, aber nur 8 davon sind wirklich krank

Das Krebsrisiko liegt damit bei 8 zu 103 oder 7,8 %



Beispiel Biometrie



Statistische Tests = schließenden Statistik

Sie dienen der **Entscheidung** über die Gültigkeit einer Hypothese von zwei gegensätzlichen:

Grund- = **Null**-Hypothese $H_0 \Leftrightarrow$ **Arbeits**- (abzulehnende) Hypothese H_A : Frage: **Wie sicher ist H_0 ?**

Entscheidungen: *einseitig*: $\theta > \theta_0$ bzw. $\theta < \theta_0$, oder *beidseitig*: $\theta \neq \theta_0$

Wichtige **Methoden**: Normal-, t-, F- und χ^2 -Test [Stoyan]; [Gränicher]

Beispiele

- Müssen die Erzproben aus Lagerstätte A und B unterschieden werden?
- Ist der Umfang der Bäume in diesem Wald normalverteilt?
- Ist der vorliegende Würfel ideal?
- Gehören zwei Messreihen zur gleichen Grundgesamtheit?

Stichproben

Normalverteilung wird angenommen. **Grundgesamtheit** in zwei Varianten;

- **Endliche Anzahl** aller Objekte, z. B. Wähler in Deutschland, Bäume eines Waldes
- **Wertevorrat**, z. B. Messungen, die ständig wiederholt werden können; formal $\rightarrow \infty$

(sehr) kleine Anzahl, aber **repräsentativ** ausgewählt, z. B.:

- **unabhängig** vom zu testenden Ergebnis: z.B. Wahl: Menschen am 31.7. geboren (Horoskope andere Meinung)
- **Teilgrundgesamtheit**: Männer / Frauen / Kinder; Altersgruppen; Berufe usw.
- **Hinreichend großer Stichprobenumfang**

Ermittelte Irrtumswahrscheinlichkeit α

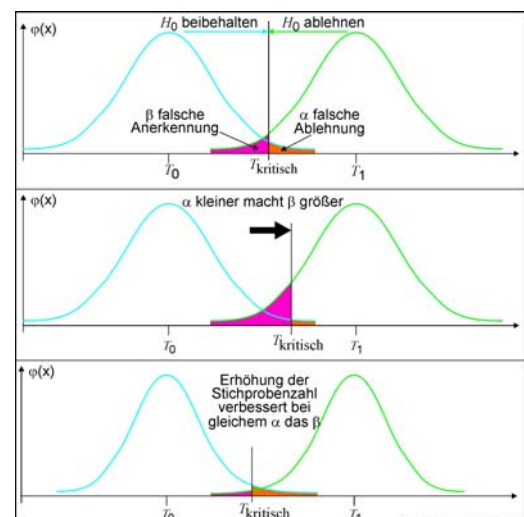
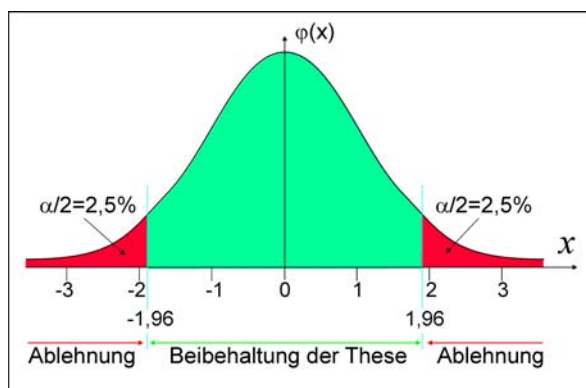
Sie bestimmt Grenze zwischen Ablehnung und Zustimmung = Fehlentscheidung *erster Art* = Irrtumswahrscheinlichkeit

α sehr klein $\Rightarrow$ immer Annahme von H_0 ; α sehr groß $\Rightarrow$ immer Annahme von H_A

$0,01 \leq \alpha \leq 0,05$ = typische, günstige Werte und

$\alpha < 0,005$ bzw. $\alpha > 0,1$ Vorsicht! Manipulation?

Fehler zweiter Art, Risiko, der tatsächliche Unterschied wird nicht gefunden, hierfür β als **Teststärke**



Erzeugung von Zufallszahlen

Zufallszahlen werden für viele Untersuchungen benötigt.
Zu ihrer **echten** Erzeugung existieren hauptsächlich zwei Varianten

1. Pegel des thermodynamischen oder Schrot-**Rauschens**
2. Rechnerinterne Uhr: Entnahme des Wertes der $\frac{1}{100}$ Sekunden
3. Adresse vom **Refresh-Register**

Wichtiger sind deterministische Algorithmen

Dabei wird nur **nur Pseudozufall** erzeugt = eine (sehr) **lange Periode N** von „**unregelmäßigen**“ Zahlen: **Eigenschaften:**
Gewünscht $N > 10^6$, oft genügen 50 000

- Es wird das **Standard-Intervall** [0; 1) benutzt
- Innerhalb der Periode dürfen sich die Zahlen **nicht wiederholen**
- Das Intervall müssen sie „gleichmäßig“ erfüllen (**Gleichverteilung**): Mittelwert $\approx 0,5$; Streuung $\approx 0,288675$
- Für **besondere Verteilungen** – z.B. normalverteilt – erfolgen nachträglich **Transformationen**

Vorteile: Zuverlässigkeit, Wiederholbarkeit, Ergänzung durch echten Zufall

Dazu wird der wählbare Startwert der Periode **seed** (englisch Samen, Saat, Keim) benutzt
Zwei Typische Befehle sind:

RND(X): hierin bedeutet

- $X < 0$ X ist seed bestimmt Startwert der Zufallsfolge
- $X = 0$ letzter Zufallswert wird wiederholt
- $X > 0$ Zufallsfolge wird fortgesetzt (X hat keine Bedeutung)

RANDOMIZE erzeugt zufälliges seed von interner Uhr oder Refresh-Register

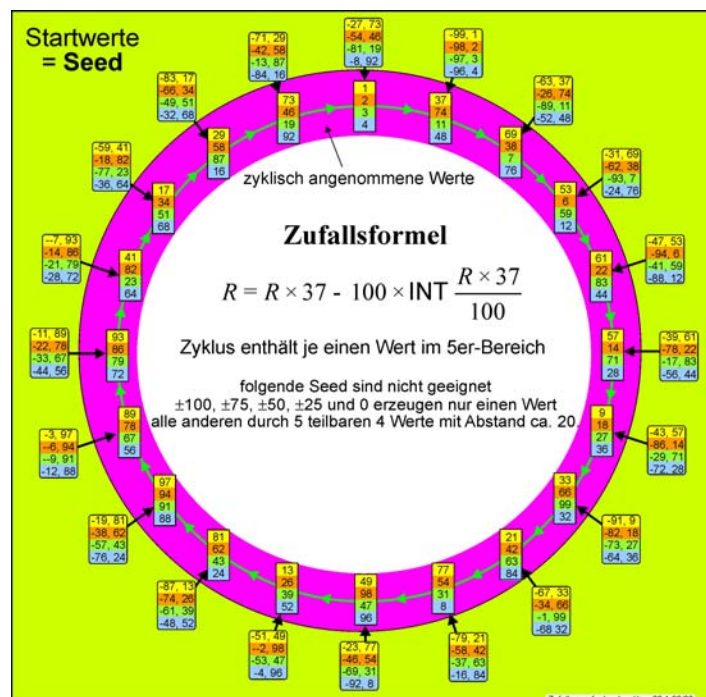


Bild N = 20 mit 4 möglichen Zyklen, einige Seed erzeugen keine Folge

Erster Zufallsgenerator JOHN VON NEUMANN (1903 - 1957), benutzt mittlere Ziffern aus Quadratzahlen < 1

Seed bzw. Startwert	Quadrat	Zufallszahl
0,9876	0,97535376	0,5353
0,5353	0,28654609	0,6546
0,6546	0,42850116	0,8501
0,8501	usw.	usw.

Methode hat sich nicht bewährt, es entstehen zuviel kleine Werte

Typische Zufallsgeneratoren

Simple Methode M Modul (Arithmetik) F Faktor; $F \approx M/2$ aber teilerfremd; Beispiel $M = 100$; $F = 37$

$x := x \cdot F$; $x := x - M \cdot \text{INT}(x/M)$; seed = -1; nicht ganzzahlig, dann eventuell sehr lange Perioden

D. Lehner (1951) Methode der *linearen Kongruenz*, $x := x \cdot b + 1$; **MOD(m)**; $m \gg 1 \approx$ Grenze des Computerbereiches

Z. B.: $x := 5^{17} \cdot x \text{ MOD}(2^{40})$; Normierung auf $[0,1)$ mit 2^{-40} , benötigt 40-stellige Zahlen, Periodenlänge 2^{38}
Abwandlung für 36-stellige Zahlen mit $\text{MOD}(2^{36})$, Periodenlänge 2^{34}

Knuth-Abwandlung: Zweier- oder Zehnerpotenz, b endet auf x_{21} , x gerade

Taschenrechner: $x := \text{FRAC}(997 \cdot x)$, $s = 0,5284163$; Periodendauer bei 10-stelligen Zahlen 500 000

$$x := \frac{4835715066 \cdot x + 99991}{199017} \text{ (TI) oder } x := \text{FRAC}[(x + \pi)^5]$$

Weitere Varianten

$z(i) := [z(i-1) + z(i-2)] \text{ MOD } m$

$m = 2^{28}$

$z(i) := [2^R \cdot z(i-1) + C] \text{ MOD } m$

$R > 2$; C gerade; $m = 2^{25}$

$z(i) := [S \cdot z(i-1)] \text{ MOD } m$

$S = 23$; $m = 10^7 + 1$

Verteilung von Zufallszahlen

Ausgang: Zufallszahlen $u \in [0, 1)$, **Ziel** Zufallszahlen gemäß Funktion $F(x) \Rightarrow$ Transformation $x = F^{-1}(u)$

Beispiel **Exponential-Verteilung** Mittelwert μ ; $z_e = -\mu \cdot \ln(x)$

$$F(x) = 1 - e^{-\lambda \cdot x}; \quad x \geq 0 \Rightarrow 1 - e^{-\lambda \cdot x} = u \Rightarrow e^{-\lambda \cdot x} = 1 - u$$

$$x = -\frac{1}{\lambda} \cdot \ln(1 - u) \text{ bzw. } x = -\frac{1}{\lambda} \cdot \ln(u)$$

Beispiel: **Normalverteilung**: Standardabweichung σ , Mittelwert m

Gleichverteilung nach $[a, b)$ durch $y = (b - a) \cdot x + a$

$$v_1 = 2 \cdot x_i - 1 \text{ und } v_2 = 2 \cdot x_{i+1} - 1 \text{ und } t = v_1^2 + v_2^2$$

Falls $t \geq 1$ Werte verwerfen und nächsten zwei x_i nehmen:

$$x_i = \sigma \cdot v_1 \cdot \sqrt{\frac{-2 \ln(s)}{t}} + m \text{ und } x_{i+1} = \sigma \cdot v_2 \cdot \sqrt{\frac{-2 \ln(s)}{t}} + m$$

Anwendungen

Spiele $\Rightarrow$ Künstliche Intelligenz ab Folie 28

Informationstheorie extra behandelt

Bediensysteme: Verlustsysteme, Warteschlangen, Verkehrsaufkommen

Bediensysteme

U. a. für Ambulatorien, Frisöre, Flughäfen, Verwaltungen, Telefonzentralen usw. [Härtler], [Gnedenko], [Stoyan]

Verlustsysteme, z. B. Telefonanschluß nur Anschluss oder nicht

Warteschlangensysteme: Möglichkeit zu warten, Größe der Schlange: endlich $\Leftrightarrow \infty$

Verkehrssysteme, mehrere Leitungen und Forderungen zur Verteilung, klassischer Telefonanschluß

Fragen betreffen: u. a.: **Wartezeiten** der Kunden und **Leerlaufzeiten** der „Geräte“

Verlustsystem liegt dann vor, wenn

- die Plätze in der Warteschlange nicht unendlich sind und $|$ oder
- wenn Kunden nur bereit sind, eine bestimmte Zeit zu warten

Die Formeln für alle Warteschlangensysteme sind recht kompliziert

Selbst bei groben Vereinfachungen bleiben sie noch unübersichtlich:

Warteschlangen

Zufallsgrößen

Anzahl der **Forderungen** je Zeiteinheit; z. B.: Zeitpunkt jeder Forderung

Mittlere **Dauer je Bedienung**, z. B.: Kunden je Stunde; Spitzenzeiten (Dauer eine Haarschnitts)

Länge der **Warteschlange** z. B. Anzahl geöffneter Schalter; Sitzplätze, verfügbarer Raum

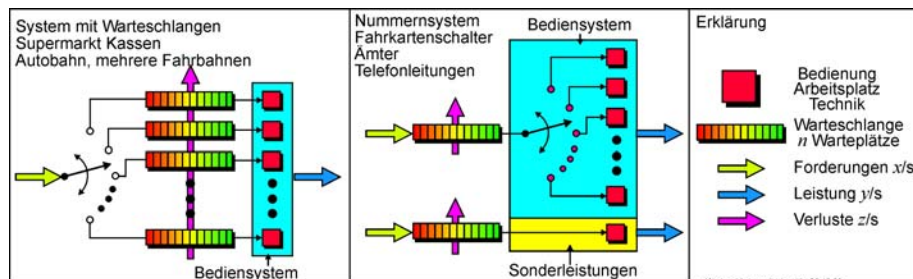
Geduld der Wartenden, Verluste möglich

Prioritäten bzw. Möglichkeit der **Verkürzung**; z. B. Sonderbehandlung (nur 3 Artikel, Sofortfahrt); durch Aufpreis

Wahrscheinlichkeit überhaupt und wann **bedient zu werden**, z. B. steht freier Anschluss zur Verfügung
Relativer **Leerlauf**, z. B. keine Kunden, Forderungen

Annahmen:

- Die Bedienungswünsche sind POISSON-verteilt
- Die **Dauer** der Bedienung ist zufällig, gemäß $e^{-\nu t}$ mit ν als Konstante
- jeder Wunsch wird von nur einem, der **n „Geräte“** befriedigt
- Die **Plätze** in der Warteschlange sind **unendlich**, k davon sind belegt (warten)
- Zwischen den Bedienungen jedes „Gerätes“ gibt es **keine Pause**



POISSON'scher Punktprozess

Prozess aus vielen unabhängigen Teilprozessen zusammensetzt, ein Teil wenig Einfluss auf die Summe hat. Eigenschaften:

- **stationär:** verhalten sich unabhängig vom aktuellen Zeitpunkt
- **nachwirkungsfrei:** Gegenwart wird in keiner Weise vom Vorangegangenen beeinflusst
- **ordinär:** In einem sehr kurzen Intervall dürfen nicht mehrere Ereignisse gleichzeitig auftreten

Beispiele: Anrufe in einer Telefonzentrale, Meldungen bei Rettungsdiensten, Ankünfte von Flugzeugen oder Frachtschiffen, Braunsche Molekularbewegung

Dann gilt

$$p_k(t) = \frac{(\lambda \cdot t)^k}{k!} \cdot e^{-\lambda \cdot t}$$

Im Zeitintervall t existieren k Prozesse mit Wahrscheinlichkeit $p_k(t)$, λ = Konstante

Beispiel Warteschlangen

Pausenzeit gehorcht der Exponentialverteilung $F_p(t) = 1 - e^{-\lambda t}$

Bedienzeiten ebenso $F_p(t) = 1 - e^{-\mu t}$

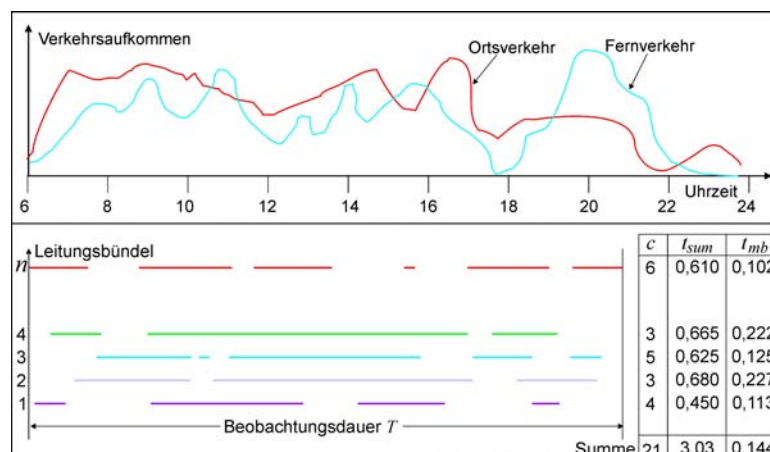
Nur ein Bediener mit $\rho = \lambda/\mu$. Nur für $\rho < 1$ stationäre Zustände

Dann müssen im Mittel warten (Länge der Schlange) $m_L = \frac{\rho^2}{1 - \rho}$

Mittlere Wartezeit beträgt $m_w = \frac{\rho}{(1 - \rho) \cdot \mu}$

Verkehrstheorie

Dient zur Bestimmung der notwendigen „Leitungen“, Kanäle. Verkehrsaufkommen in 1/4 Stunden gemessen.



Erste Publikation von A. K. ERLANG 1917

- **Verkehrswert A**, experimentell ermittelt in Erl (Erlang), statistischer Mittel-, Erwartungswert
- gilt für **Leitungsbündel** aus **n** Leitungen, Anzahl Leitungen zu einer Zeit (z.B. Hauptverkehrszeit) belegt

Wie viel Leitungen **x** sind mit welcher Wahrscheinlichkeit belegt; Erlang-Verteilung:

$$p_x = \frac{A^x / x!}{1 + \sum_{v=1}^n \frac{A^v}{v!}}$$

Wenn p_0 berechnet ist, dann rekursiv möglich

$$p_{x+1} := \frac{A}{x+1} p_x \text{ für } x=1 \text{ bis } n$$

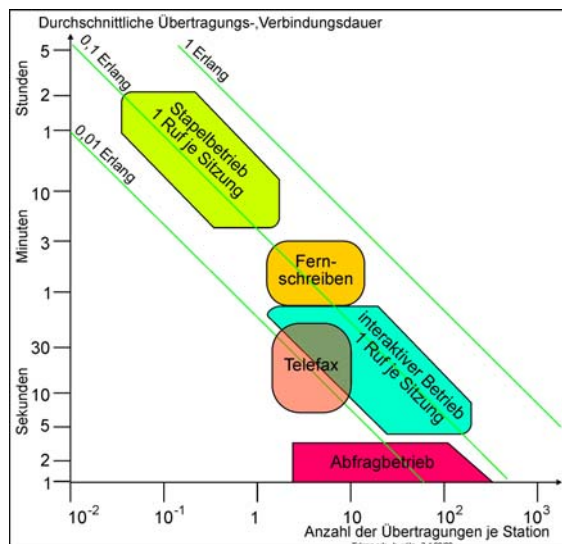
Für große x , insbesondere n , möglich Wahrscheinlichkeit $p_x > 1 \rightarrow$ tritt Verlust $B = p_x - 1$ auf.
Grenzwert für alle n Leitungen:

$$B = \frac{A^n / n!}{1 + \sum_{v=1}^n \frac{A^v}{v!}}$$

z. B.: Bündel aus 10 Leitungen, Verkehrswert 2,5 Erl gewählt $\rightarrow$ Wahrscheinlichkeiten der Belegung von x Leitungen

x	1	2	3	4	5	6	7	8	9	10
%	0,684	3,42	14,25	17,81	17,81	14,48	10,60	6,625	3,681	1,840

Die Summe = 1 (Rundungen) eins $\rightarrow$ Verlust $B = 1,85\%$ [Rumpf]



Zuverlässigkeit

Fehlerkorrektur-Verfahren bei Speicherung und Datenübertragung

Hardware- Zuverlässigkeit

Fehlertoleranz

Einfluss von Entwicklung, Testung, Aufwand usw.

Software-Zuverlässigkeit, wird hier nicht behandelt, da nicht Statistik sinnvoll

Fehlererfassung und -behandlung

Für Daten-Übertragung und -speicherung: **Fehlererkennung, -korrektur**;

Als vorbereitende und meist auf Hardware bezogen **Fehlertoleranz** und **Zuverlässigkeit**

Software-Zuverlässigkeit wird hier nicht behandelt

Mittelbar **Kryptographie**, **Personen-** und **Produkt-Identifikation**, da kaum Zufall wichtig, hier nicht behandelt

Fehler-Erkennung und -Korrektur = FEK
= EDC und EEC error detection or correction code

Fehlerwahrscheinlichkeit F_w

Fügt Datenströmen und Dateien eine „förderliche“ **Redundanz** hinzu

Damit wird eine **feststellbare Struktur** erzwungen

Ist sie gestört, verändert oder fehlt sie später gar, so sind Fehler bzw. Störungen vorhanden

Z. T. entstehen Fehler a priori nach definierten **Gesetzen der Wahrscheinlichkeit**, z.B. Thermisches Rauschen, Quanteneffekte

In anderen Fällen, kann nur über einen hinreichend großen Datenumfang eine **Häufigkeit** ermittelt werden

Typisch sind Störungen, z. B. durch Gewitter, Netzschwankungen, Fingerabdrücke auf Magnetbändern, CD, DVD usw.

Auch bei der Produktion von Speichermedien treten unregelmäßige Fehler, z.B. als **Drop-Out** auf

Viele Fehler bewirken in einem kurzen Abschnitt mehrere eng beieinander liegende Fehler = **Bündel-, Büschelfehler, Burst**

Trotz alledem ist es üblich immer von **Fehlerwahrscheinlichkeit F_w** zuspriechen, wird als **relativer Wert** z. B. 10^{-6} angegeben

Auch **messtechnisch** kann wegen endlicher Messzeit eigentlich nur eine **Fehlerhäufigkeit** bestimmt werden

Um sie zu annähernd zu bestimmen, muss aber **mindestens die hundertfache Bit-Anzahl** benutzt werden

Für Fehler von 10^{-10} also $>10^{12}$ Bit ≈ 1 TByte

Daher sind sehr kleine Fehlerwerte – wenn überhaupt – kaum in sinnvoller Zeit zu messen, sie müssen extrapoliert werden

Die **Erfahrung** zeigt, dass für Fehler-Erkennung und -Korrektur eine Fehlerwahrscheinlichkeit von **10^{-4} bis 10^{-6} optimal** ist

Für kleinere wächst der Technische (Hardware-) **Aufwand exponentiell** an

Für größere gibt es **keine brauchbaren Fehlerkorrektur-Verfahren**

Für diesen Zusammenhang ist keine theoretische Begründung bekannt, doch selbst die **Genetik** (Mutation s. u.) hält sich daran

Fehler-Arten

Bei einer Fehlerwahrscheinlichkeit von 10^{-5} kommt im statistischen Mittel auf 10^5 Bit ein Bit-Fehler

Bei einer Datei von **1 MByte** ergeben sich **≈ 800 Bit-Fehler**

Für größere Dateien und Dateiströme sind es noch entsprechend mehr

Technisch lassen sich aber **nur wenige Fehler $\approx < 10$** einigermaßen brauchbar feststellen und/oder beheben

Außerdem können Bit-Fehler den Takt stören und den dann folgenden Rest der Datei, des Datenstromes unbrauchbar machen

Deshalb arbeiten Fehlerkorrektur und -erkennung mit relativ **kleinen Blockgrößen** von etwa 100 bis 10 000 Bit

Damit bei 10^{-5} nur etwa jeder 20. Block einen Fehler besitzt, darf die Blockgröße nur ≈ 750 Byte betragen

Bei **Bursts** usw. kann innerhalb eines Blocks ein nur relativ kurzer Datei-Abschnitt

betrachtet werden, Beginn bei 1. Fehler-Bit

Dies ermöglicht besonders effektive Korrekturverfahren

Tritt in einem Block ein **1-Bit-Fehler** mit $F_w = 10^{-5}$ auf

So erscheinen **2-Bit-Fehler** mit $F_w \approx (10^{-5})^2 = 10^{-10}$

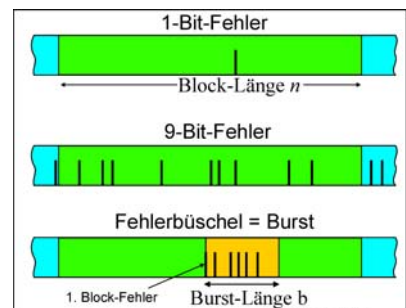
und **3-Bit-Fehler** mit $F_w \approx (10^{-5})^3 = 10^{-15}$ usw.

Fehlererkennung und -korrektur erfolgt immer nur mit **statistischer Sicherheit**

Folglich können auch – aber viel seltener – **zusätzliche Fehler entstehen**

Neben den Bit-Fehlern gibt es noch **nicht korrigierbare Synchron-Fehler**

Dann ist der ganze Block unbrauchbar. Sie werden in der FEK nicht betrachtet



Daten- = Code-Spreizung = Verwürfelung = Verschachtelung = Interleaving = Shuffling

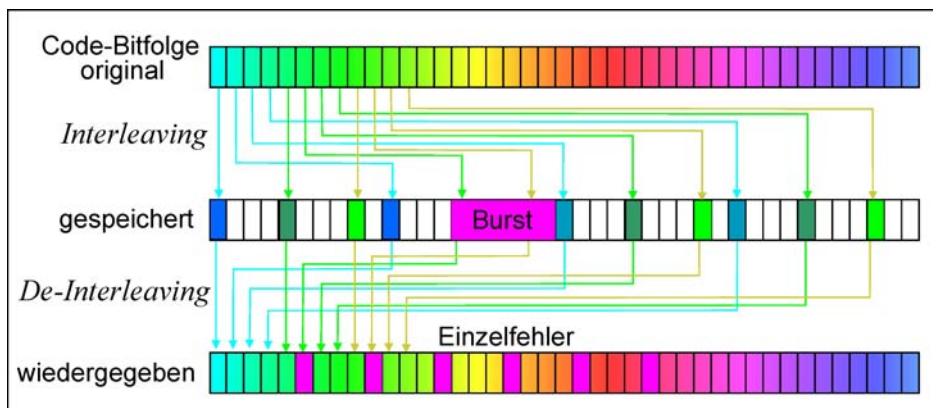
Englisch **to shuffle** schieben, schlürfen, mischen, schwerfällig gehen

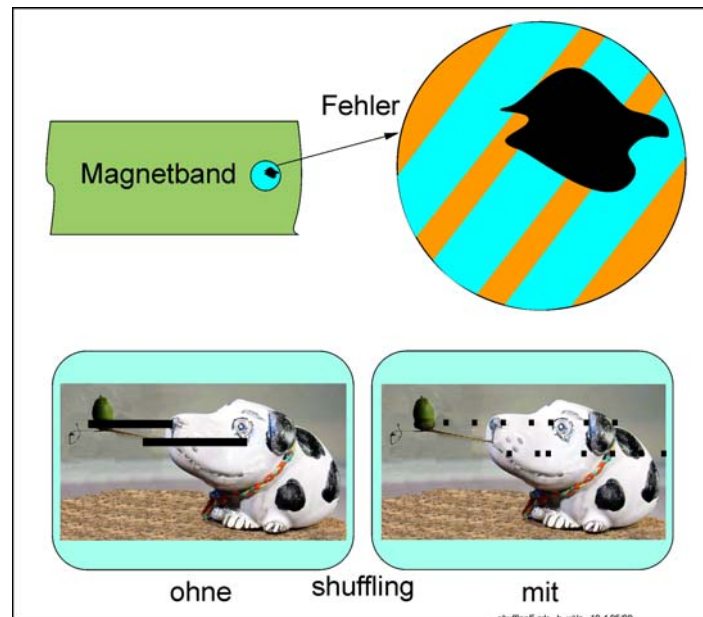
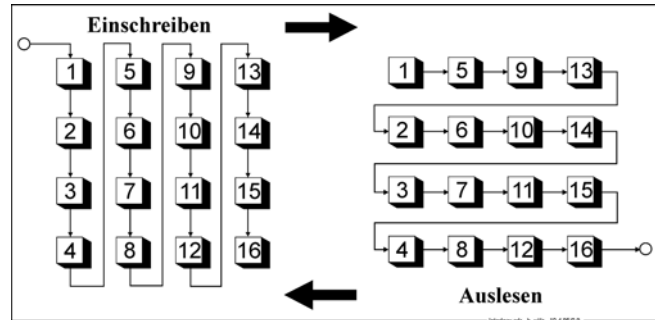
to interleave verschachteln, durchschießen, überlappen, verzahnen

Bursts sind recht häufig und bestehen aus dicht gedrängten Fehler-Anhäufungen

Hiergegen ist eine Verschachtelung sehr vorteilhaft: sie verlagert sie in verschiedene Blocks als 1-Bit-Fehler

Shuffling wird vor allem beim Fernsehen verwendet, hier sind 1-Bit-Fehler kaum noch sichtbar

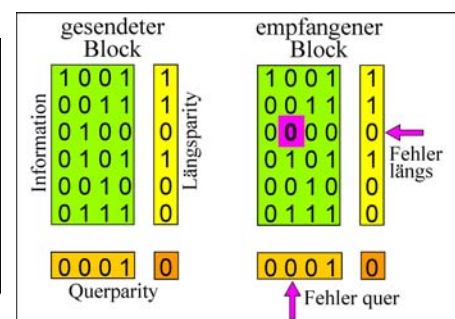
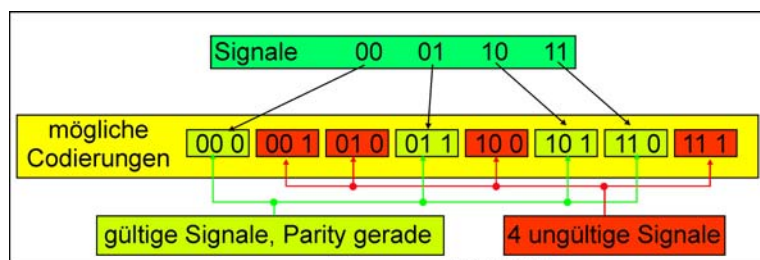




Einfache Sicherungsverfahren

Wiederholung einer Übertragung ist wenig effektiv: Hoher Aufwand und trotzdem wenn gleich immer noch nicht sicher

- **Gleiches Gewicht** genau festgelegte Anzahl von "1"
z. B. für $4 \times 1 \rightarrow 110011, 101011, 110110, 10101010, 1111$ usw.
- **Symmetrie, z. B.:** gültig: 110011, 101101, 011110 usw. vor- und rückwärts gleich (Bar-Codes)
- **Parity** gezählt werden die 1 mod 2; zusätzliche wird 1 oder 0 angehängt gerade / ungerade Parity
mehrere Parity-Bit möglich aber unüblich, erkennen dann mehrere Fehler, korrigierbar wenige
Blockparity x; y korrigiert im Block genau einen Fehler
- **komplexe Verfahren** (extra behandelt, mathematisch schwierig)



Ziel und Prinzip der FEK

Ein **Block** besitzt die **Länge** n , in ihm sind i Bit für **Information** und $k = n - i$ Kontroll-Bit für **FEK** enthalten

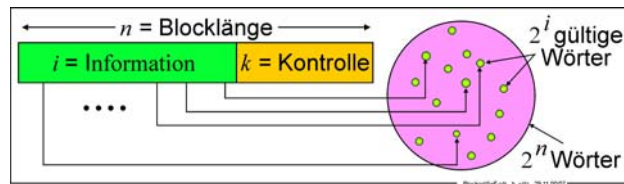
Die Informations-Bit müssen dabei im Block weder zusammenhängend vorliegen noch direkt enthalten sein

Die i Informations-Bit ermöglichen 2^i unterscheidbare **Signale**

Sie sind auf eine **Auswahl** der 2^n mögliche **Code-Wörter** abzubilden

Die Auswahl muss eine **gut erkennbare Struktur** schaffen

Dies ist das **mathematisch schwierige Problem** der FEK, deshalb Entwickler-Namen für viele Codes
Z. B. HAMMING-, BOSE-CHAUDHURI-HOCQUENGHEM- (BCH), FIRE-Code usw.

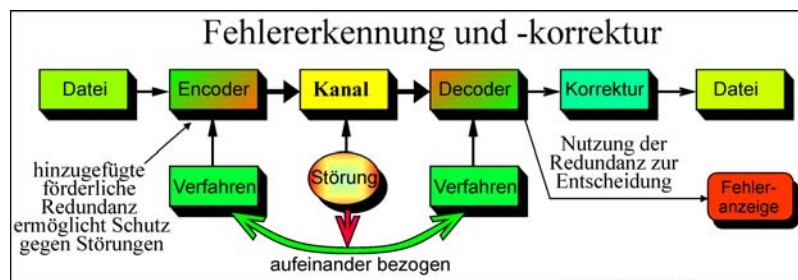


Methoden der FEK

Die Abbildung der Informations-bit auf die auf die „gültigen“ Code-Wörter heißt **Encodierung**
Im Kanal treten Störungen auf, die entsprechend F_W einige gültige **Code-Wörter verändern**

1. Stark überwiegend treten dabei ungültige Wörter auf
2. Sehr selten auch gültige, diese Veränderung kann weder erkannt noch korrigiert werden

Auf der Wiedergabe-Seite erfolgt die **Decodierung** zu den ursprünglichen Informations-Bit
Dabei kann für die technische Auslegung zwischen den Anteilen von EDC und ECC ausgewählt werden

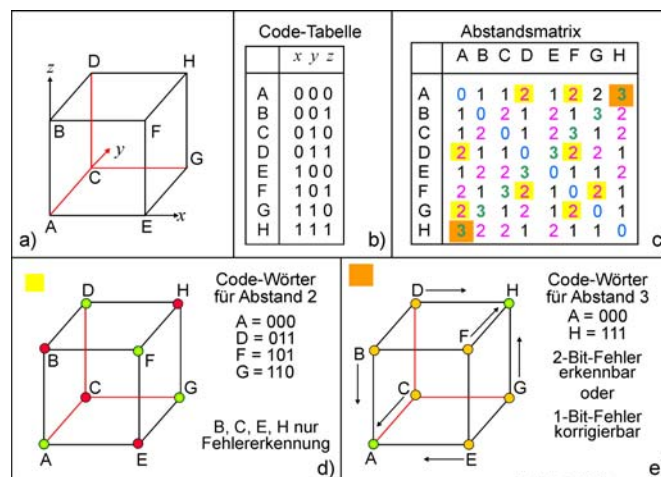


Beispiel für 3-Bit-Wörter

$n=3 \Rightarrow$ bit 1 = x-, Bit 2 = y-, Bit 3 = z-Achse $\Rightarrow$ Codewörter = 8 Ecken von **Drahtwürfel**

3 Varianten für Auswahl von gültigen Wörtern mit Bit-Abständen:

- 8 gültige (1-Bit-Abstand): keine FEK möglich
- 4 gültige (2-Bit-Abstand) 4 ungültige Wörter ermöglichen Erkennung von 1-, 3-, 5-, ... Bit-Fehlern
2-, 4-, 6- ... Bitfehler werden nicht nur nicht erkannt, sondern führen zu **falschen gültigen** Wörtern!
- 2 gültige (3-Bit-Abstand) ermöglichen Varianten:
 - a) Erkennung von 1-, 2-, 4-, ...Bit-Fehlern
 - b) Korrektur von 1-Bit-Fehlern, aber keine Erkennung 2-Bit-Fehlern, werden **falsch korrigiert**!
 Generell gibt es wieder Bit-Fehler, die weder erkannt noch korrigiert werden



Der Hamming-Abstand

Bild vom Würfel $\rightarrow$ **n Dimensionen** für n-Bit-Wörter nicht anschaulich möglich

Daher einen Ausschnitt in Ebene projiziert zeigt: Abstand zwischen gültigen Wörtern ist unterschiedlich

Minimaler Abstand zwischen allen gültigen Wörtern ist **Hamming-Abstand h** (RICHARD WESLEY HAMMING; 1915 – 1998)

Er bestimmt die **Leistungsfähigkeit des Codes**, der gültigen Code-Wörter

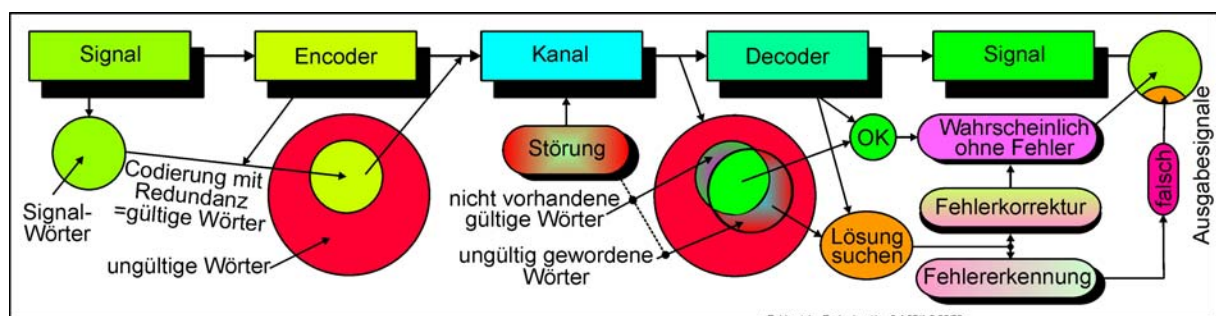
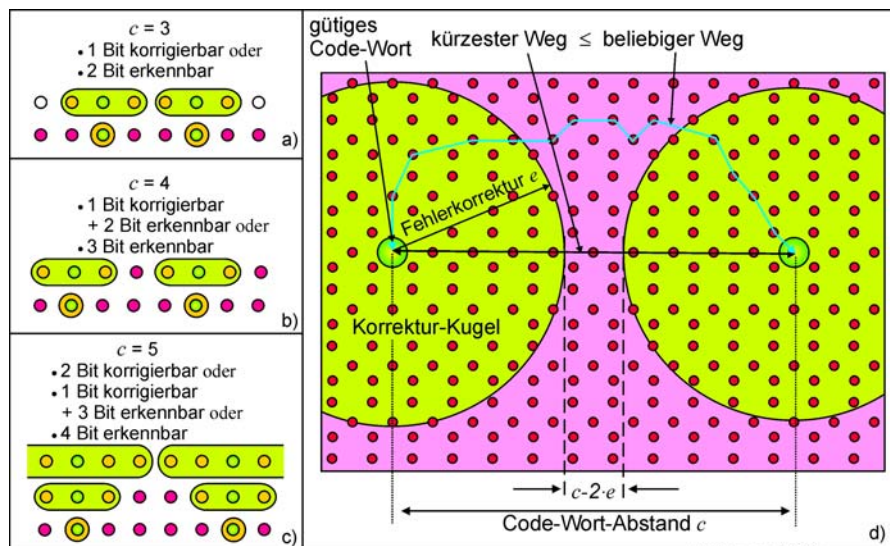
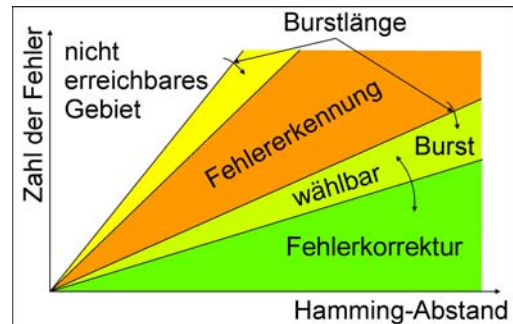
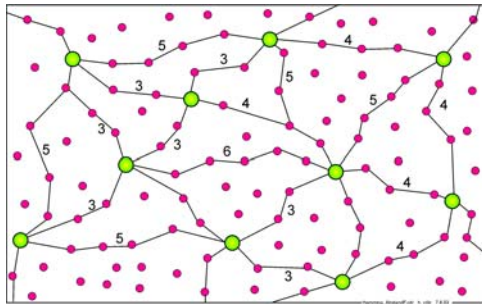
Für einzelne gültige Code-Wörter sind zwar bessere Leistungen möglich, doch es muss **worst-case** berücksichtigt werden! Dann gilt

Maximal erkennbare Fehler $e_{\max} = h - 1$

Maximal korrigierbare Fehler $c_{\max} \leq \text{Int}\left(\frac{h-1}{2}\right)$

Varianten: Werden nur $c < c_{\max}$ korrigiert, so sind noch $e(h, c) = h - c - 1$ erkennbar

Auf Formeln für **Bursts** und für die sich ergebende **förderliche Redundanz** sei hier verzichtet ($\approx$ reziprok zur Blocklänge)



Mathematische Methoden

Zwei Methoden sind wichtig, gute Einführung [Peterson]

Polynom-Methode besonders übersichtlich: Polynom gilt zugleich für Signale und Schieberegister

ermöglicht daher den leichten Übergang von Rechnung zu Hardware. Einschränkung **binäre** Signale

Matrizen-Methode benutzt GALOIS-Felder (ÉVARISTE GALOIS 1811 – 1832), auch für **mehrstufige** Signale

Bei Polynomen sind übersichtlich **Additionen**, **Subtraktionen**, **Multiplikationen** und **Divisionen** ausführbar

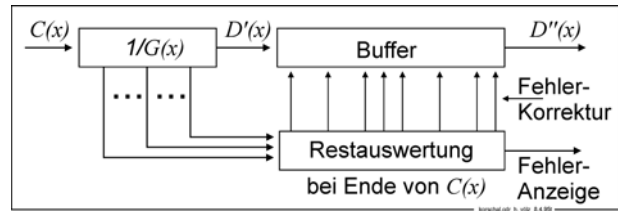
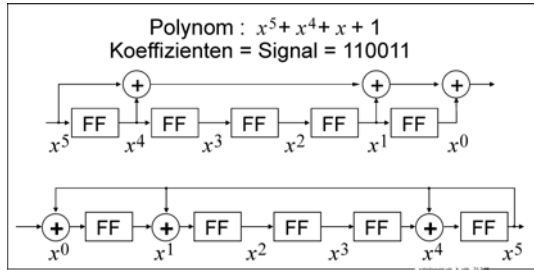
Division erfolgt ähnlich wie bei natürlichen Zahlen mit Rest, z. B. $11/4 = 2$ Rest 3

In diesem Sinne gibt es **irreduzible Polynome**, die Primzahlen ähneln

Sie sind ohne Rest nur durch $x^0 = 1$ und sich selbst teilbar. Sie sind u.a. für die CRC-Methoden (s.u.) wichtig

Die ersten sind: 11; 111; 1011; 10011; 11111; 100101; 101111; 110111 (auch rückwärts gelesen gültig)

Der **Rest** bei einer typischen Division ist dagegen für die Durchführung der **Fehlerkorrektur** wichtig

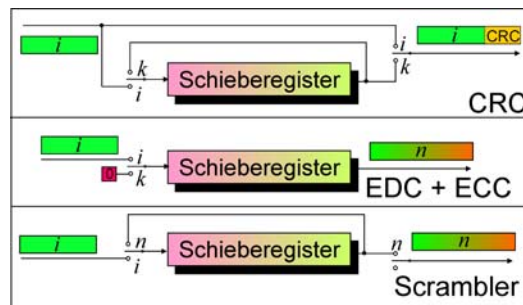


CRC- und andere Methode

Wird bei einem Schieberegister mit n FF der Ausgang mit dem Eingang verbunden, so entsteht ein Generator
 Jeder Takt verschiebt die in den FF enthaltenen Werte mit Beachtung der XOR-Operationen nach rechts → **zyklisches Signal**
 Bei einer **irreduziblen Polynom** werden alle möglichen 2^n Belegungen (zufallsähnlich) durchlaufen
 Anderen Polynomen erzeugen je nach der Anfangsbelegung mehrere unterschiedliche Zyklen → Auswahl gültige Wörter!

Irreduzible Polynome für CRC (Cyclic Redundancy Code) der Fehlererkennung, ausgewählt genormt
 Entsteht bei Wiedergabe kein Unterschied, so beträgt Sicherheit für des Schieberegisters n FF $1/(2^n - 1) \approx 10^{-0,3 \cdot n}$

Länge	Fehler	Polynom
8	$4 \cdot 10^{-3}$	$x^8 + 1$
12	$2 \cdot 10^{-4}$	$x^{12} + x^{11} + x^3 + x^2 + 1$
16	$2 \cdot 10^{-5}$	$x^{16} + x^{15} + x^2 + 1$
32	$2 \cdot 10^{-10}$	$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$



Zuverlässigkeit

Geschichte

Beachten: gewiss gab es auch bereits in der **Antike** intuitive und erfahrene Zuverlässigkeit bei Bauten usw.
 Doch es mussten zunächst die **Wahrscheinlichkeitsrechnung** und **komplexe elektronische Schaltungen** entstehen

- 1930 V. M. MONTINGER erste Arbeit; Zuverlässigkeit elektrischer Maschinen, 8°-Regel
- 1942 BUSSING theoretische Zusammenhänge zur Zuverlässigkeit: Archiv für Elektrotechnik: „Beiträge zum Lebensdauergesetz elektrischer Maschinen“
- 1950 Beginn experimenteller Untersuchungen zur Zuverlässigkeit von Bauelementen
- 1951 Zuverlässigkeit bei Radar-/Feuerleitergeräten USA entscheidend im Koreakrieg
- 1955 Zuverlässigkeit von Bauelementen mindert bei Rechnern erheblich den Betrieb
- 1958 Störungen im Funkverkehr durch Atomexplosionen, EMP noch nicht erkannt
- 1960 Zuverlässigkeit wird bei Raumfahrt wichtig
- 1962 8.7. Atomversuch 350 km Höhe, Stromversorgung Honolulu bricht zusammen, EMP
- 1978 Entdeckung der Alpha-Strahlungsfehler in DRAM, Tim May

Wichtige Begriffe

Ausfall: ≈ **Versagen:** Ende der Funktionsfähigkeit des Systems, wird erst bemerkt, wenn die Teilfunktion verlangt wird

Ausfallrate: λ Anzahl der Ausfälle, Fehler je Stunde, z. T. auch für 10 Jahre

Beanspruchung: u.a. durch Spannung, Leistung, Temperatur, Mechanik, Feuchte, Strahlung, Gewitter, EMP usw.

Fehler: unzulässige Abweichung des Systems von den Forderungen ≈ Ausfall ≈ Versagen

Einzelfehlerannahme: Vereinfachung: es tritt jeweils nur ein Fehler auf, Rest des Systems arbeitet fehlerfrei

Fehlertoleranz: Fähigkeit eines Systems, mit einer begrenzten Zahl fehlerhafter Subsysteme seine Funktion zu erfüllen
 Es werden Redundanzen benutzt, um die Zuverlässigkeit und Verfügbarkeit eines Systems zu erhöhen

Funktionsdauer: Bzgl. zuverlässiger Funktion u.a. unterscheiden: wirkliche, zu erwartende Betriebszeit und Garantie
Redundanz: Mehrfach vorhandene von Subsysteme für eine Funktion (anders bei Fehlerkorrektur s. u.)
Selbsttestbarkeit: Fähigkeit des Systems selbst Teilfehler zu erkennen
Verlässlichkeit ≈ Vertrauen in die Leistung eines Systems
Verfügbarkeit: Wahrscheinlichkeit die Systemanforderungen zu erfüllen
Versagen = nicht vorbestimmtes, unerwünschtes Reagieren des Systems
Voralterung: Verfahren zur Vermeidung von Frühausfällen (s. Badewannenkurve)
Wartbarkeit: Wie gut das System zu pflegen, warten, in Betrieb zu halten ist
Zeitraffende Verfahren: Verkürzung der Messzeit, durch erhöhte Belastung, z. B. Temperatur (8°-Regel)
Zuverlässigkeit: (umfassender als Qualität) Fähigkeit zur Erfüllung der vorgegebenen Funktion unter Betriebsbedingungen
 Auch gemitteltetes Verhältnis Dauer der Betriebsbereitschaft zur Zykluszeit aus Betriebs- und Instandsetzungsdauer
 Zusätzliche Subsysteme übernehmen im Fehlerfall die Aufgabe des primären Subsystems
 Zuweilen werden redundante Subsysteme auch mit benutzt → Erhöhung der Leistung des Systems
 Durch Abschaltung eines fehlerhaften Subsystems wird dann lediglich die Leistung reduziert

Statistik der Ausfälle

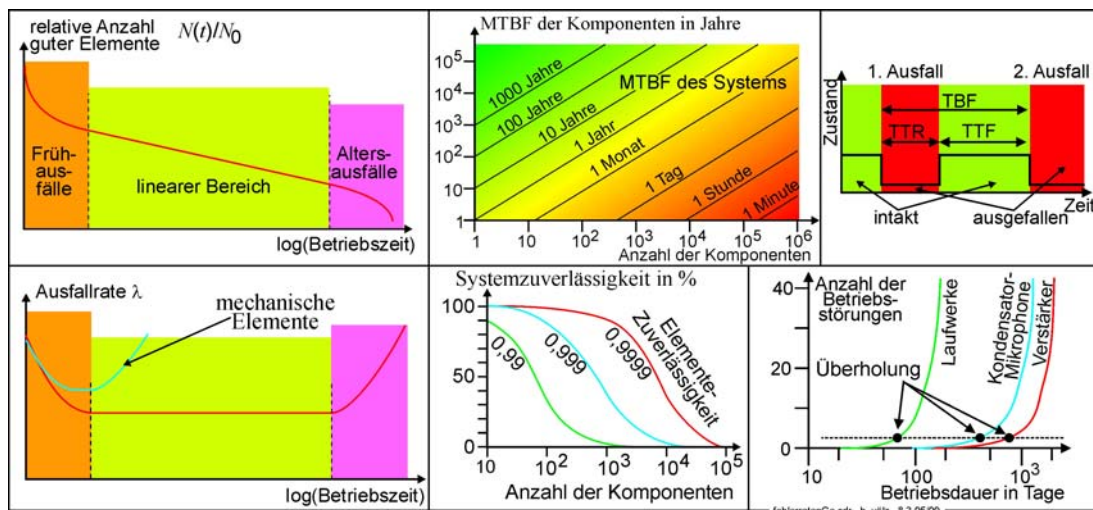
Es gibt N_0 Elemente, z.B. Geräte, Bauelemente usw. einer Serienproduktion
 Ferner eine statistische Ausfall- = Fehlerrate λ in 1/h (z.T. auch für 10 Jahre); für sie beträgt $F_W = 1 - 1/e \approx 63\%$

$$\lambda(t) = \frac{1}{N(t)} \cdot \frac{dN(t)}{dt}$$

Nach der Zeit t sind dann noch $N(t) = N_0 \cdot e^{-t \cdot \lambda}$ brauchbar

In der Praxis treten zunächst **Frühausfälle** auf, beruhen weitgehend auf mangelhafte Herstellung oder Materialfehler
 Ihnen folgt nach $\approx 100 - 2000$ h ein relativ flacher Verlauf mit **konstanter Neigung**, Er ist **optimal** für Anwendungen
 Gelangen die Elemente an die statistische Grenze „Nutzbarkeit“ wird der Abfall wieder steiler = **Spät-, Verschleißphase**
 So ergibt sich die so genannte **Badewannenkurve**

Mechanische Bauelemente besitzen einen extrem kurze Zeit für die konstante Ausfallrate
 Als häufige Fehlerquellen gelten mangelhafte Kontakte, Lötverbindungen usw. Integrierten Schaltkreise großen Fortschritt



Reserve-Elemente

Kalte Reserve

n Elemente benötigt ein System zum Betrieb
 m Elemente werden in Reserve zum Austausch bereit gehalten
 p ist die Wahrscheinlichkeit für den Ausfall eines Elements
 Die gesuchte Wahrscheinlichkeit für einen Ausfall beträgt dann

$$p_{tot} = \sum_{i=1}^m \binom{m+n}{n+1} \cdot p^{n+i} \cdot (1-p)^{m-i}$$

Beispiel: $p = 0,9$; $n = 4$; ohne Reserve gilt dann nur noch $p^4 = 0,6561$

1 Reserve-Element $\rightarrow 0,9185$, 2 Reserveelemente $\rightarrow 0,9841$

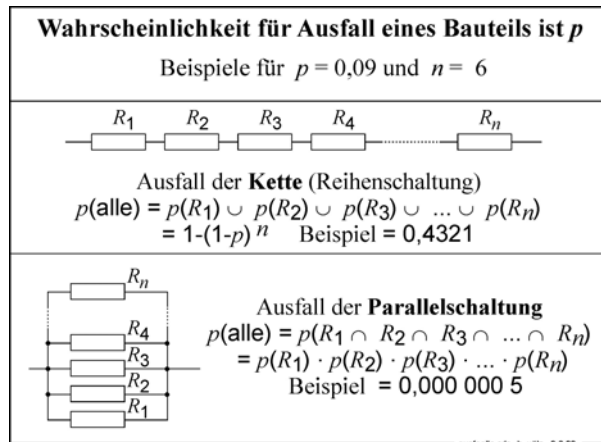
Im Beispiel bringt 1 Reserve-Element den besten Gewinn, wichtig ausgefallenes Element schnell ersetzen

Für Reservereifen bei Auto liegt p noch viel näher an 1

Wichtig ist immer die **schnelle Reparatur**. Bedingt zusätzliche Verlustzeiten

Heiße Reserve

Alle Elemente zugleich im Betrieb, Erhöht eventuell die Leistung
Sinkt dann bei einem Ausfall nur ab, Z. B. Doppelreifen
Besonderheit RAID-System mit hot logging = Austausch im laufenden Betrieb



Wichtige Begriffe

MTTF mean-time to failure, gilt ohne Reparatur bis ein Fehler eintritt ($= 1/\lambda$)
MTBF meantime between failure, gilt mit Reparatur bis zum Eintreten des nächsten Fehlers
MTTR meantime to repair, Zeit für die Reparatur eines Fehlers

In einigen Fällen entfällt das M

Ein Beispiel für günstige Zeitpunkte der Überholung wurde in den 50er in der Studientechnik des Rundfunks erfolgreich eingesetzt. Es ist auf 1 % Ausfall während der Sendung ausgelegt

8°-Regel (= 8-K)

Für gute Bauelemente dauert eine statische Untersuchung so lange, dass ihre Produktion ausgelaufen ist
Deshalb sind zeitraffende Untersuchungen unerlässlich

Erste Untersuchungen hierzu führte 1930 V. M. MONTINGER an isolierenden Kabelpapieren durch
Danach gibt es einen exponentiellen Einfluss der Temperatur T auf die Lebensdauer t_L gemäß

$$t_L = t_0 \cdot e^{-A \cdot T}$$

t_0 und A sind experimentelle, materialabhängige Konstanten

Das Ergebnis war, dass sich die Lebensdauer etwa auf die Hälfte verkürzt, wenn die Temperatur um 8° erhöht wird
Regel wird immer noch als Näherung benutzt

Prüfmethoden, Aufwand und Kosten

Bei der **Herstellung von Hard- und Software** wird die Qualität durch mehrere Einflüsse bestimmt

Fehlerverhütung bewirkt unterschiedliche Vollkommenheit und Produktqualität

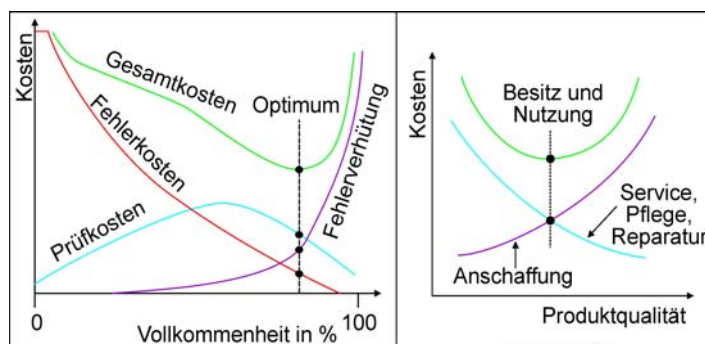
Die Kosten betreffen getrennt Herstellers und Anwender

Für **Hersteller** hauptsächlich: **Herstellungsaufwand** + **Prüfmethoden** + **Garantieleistungen**

Summe besitzt bei mittlerer Vollkommenheit ein Minimum

Für **Anwender** gilt ein anderes Minimum

Je höher Kosten der Anschaffung, desto geringer meist Kosten für Pflege, Service und Reparatur



Einflüsse

Bei der Entwicklung wichtig, **möglichst früh Fehler erkennen**, je später, desto höher der Aufwand

Beispiel: defekter Transistor schlecht eingelötet → sofortige Reparatur ca. 1 €

Erst bei der fertiger Platine → Fehlersuche + Reparatur ca. 10 €

Im fertigen Gerät → Aufwand 100 €

Beim Einsatz mit Folgeschaden vielleicht 1000 €

Beispiel: Montage elektronischer Geräte: 100 000 Operationen **mittlere Fehlerzahlen**

Vergessen, eine Komponente zu montieren 3

Zwei Drähte vertauscht 60

Eine Komponente mit Lötkolben beschädigt 100

Richtwerte für Optimum bei

50 % für Fehlerverhütung (Planung)

40 % für Prüfung, Testung

10 % für Fehlersuche und -behebung

Software fast immer weitaus komplexer als **Hardwaresysteme**

⇒ erweiterte Funktionalität und erhöhte Komplexität → geringere Zuverlässigkeit.

Fehlertoleranz

Wichtige Gründe, u. a.:

- **Extreme Bedingungen:** z. B. schlechte Belüftung, hohe Temperatur und Feuchtigkeit, instabile Stromversorgung, elektromagnetische Interferenzen

- **Unerfahrene Benutzer:** Immer mehr Menschen haben Umgang mit neuen und komplexen Systemen

- **Große Folgekosten:** Ausfall der Netzstromes, Nachrichtenverbindungen, Versorgung jeglicher Art

Beginn fehlertoleranter (Computer-) Systeme: Militär, Raumfahrt, Industrie, Kommunikation, Bankwesen

Notwendigkeit: Atomkraftwerke, Raketen, Medizinische Anlagen, Flugleit- und Steuerungssysteme, Raumfahrt

Aber auch Telefon, Transport, Logistik, Buchungs- und Transaktionssysteme

Anforderungen: Datenintegrität, Funktionskontinuität, Zuverlässigkeit, Verfügbarkeit, Lebensdauer, Instandhaltung

Nachteile: Immense Kosten bei der Anschaffung und Wartung solcher Systeme

Beispieldaten

Verfügbarkeits-Ebene	Verfügbarkeit	Ungefähre Ausfallzeit
Durchgehende Verfügbarkeit	99,999% bis 99,9995%	2,628 bis 5,256 min. / Jahr.
Kommerzielle Fehlertoleranz	99,99% bis 99,995%	26,28 bis 52,56 min. / Jahr.
Fehler beständig	99,99%	52,56 min. / Jahr.
Hoch-Verfügbarkeit	99,9%	8,76 St. / Jahr.

Nuklear-Elektromagnetischer Puls (NEMP bzw. EMP)

Geschichte des EMP

1958 zündeten USA zwei Kernexplosionen in 40 bzw. 75 km Höhe

→ beträchtliche Störungen im Funkverkehr und bei der Radarortung auf

8.7.1962 zündete USA relativ kleinen Wasserstoffbombe (1,4 Mt TNT) in 350 km Höhe

→ in Honolulu (1500 km vom Zentrum entfernt) brach Stromversorgung zusammen, Sicherungen brannten durch usw.

Anfang der 60er Jahre fielen sieben USA-Nachrichtensatelliten durch Kernexplosionen aus

Dies waren erste Beobachtung des NEMP = EMP

Ursprünglich Sicherheit, weil überall noch Röhren benutzt = 1000mal strahlungssicherer

Wirkungsmechanismen

Kernexplosion → γ -Quanten hoher Energie geradlinig in alle Richtungen

Oberhalb der Atmosphäre gezündet, treffen die Teilchen in 50 bis 20 km km auf Lufthülle

Lösen dort an den Atomen COMPTON-Elektronen aus, bewirken starke und kurze Impulse an Erdoberfläche

Feldstärken ca. 50 kV/cm, Dauer 100 ns, Anstiegflanke ca. 1 ns

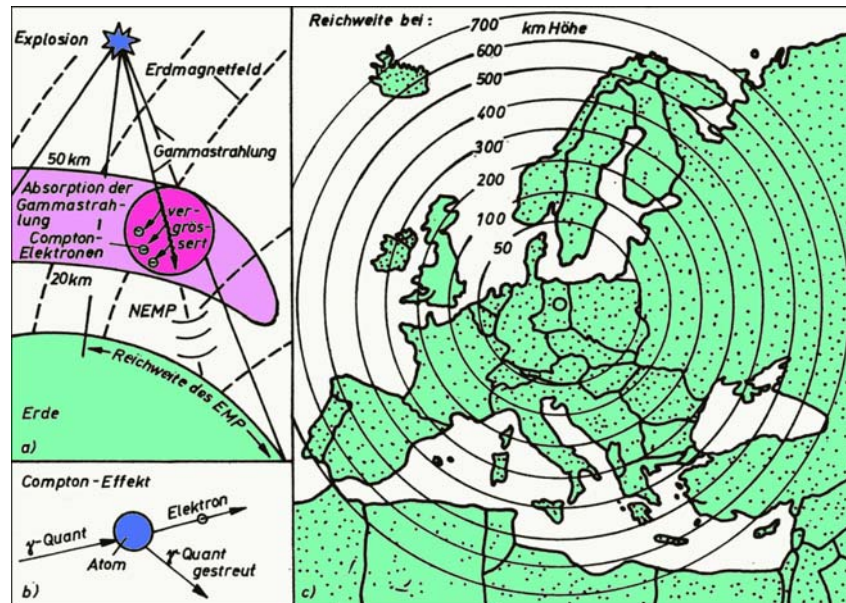
Höhe in km 50 100 200 500

EMP-Radius in km 800 1100 1600 2500

Gewitter: Feldstärke ähnlich aber lokal, Dauer 100 μ s, Flanke 1 μ s

Daher waren Sondermaßnahmen für Labore, Prüfungen und zum Schutz zu entwickeln

[Völz 89, S. 992]



Risiko: subjektiv - objektiv

Herkunft nicht eindeutig: vulgär-lateinisch **risicare/resecare** Gefahr laufen, wagen oder lateinisch **rixari** streiten
Oder griechisch **rhiza** Wurzel, Klippe bzw. widerstreben; arabisch **rizq** Lebensunterhalt, tägliches Brot

Umgangssprachlich ≈ „gefühlte“ Gefahr

Wichtig: Furcht kann durch Presse-Kampagnen usw. stark geschürt werden kann (BSE, Vogelgrippe usw.)
Bei einer Handlung kann man bewusst ein Risiko in Kauf nehmen oder gar absichtlich herausfordern
„Nur wer wagt, gewinnt – außer er verliert“

Subjektive Möglichkeit mit Risiko und Glück umzugehen ≈ Glaube an's **Schicksal**
alt-niederländisch **schicksel** Fakt, auch schicken machen, dass etwas geschieht
Verwandt damit Faktum und Los (lateinisch **fatum**, griechisch **moira**. Althochdeutsch (**h**)lôz Omen, Orakel)
Unausweichlicher **Schicksalsschlag** in der Mythologie ≈ personifizierte Macht = Schicksalsgottheiten
⇒ Wir müssen mit **Risiko** leben.

Für **subjektiv empfundenes Risiko** werden zuweilen zwei Parameter angenommen

1. wie stark „furchtauslösend“
2. Bekanntheit der möglichen Folgen

Wissenschaftlich:

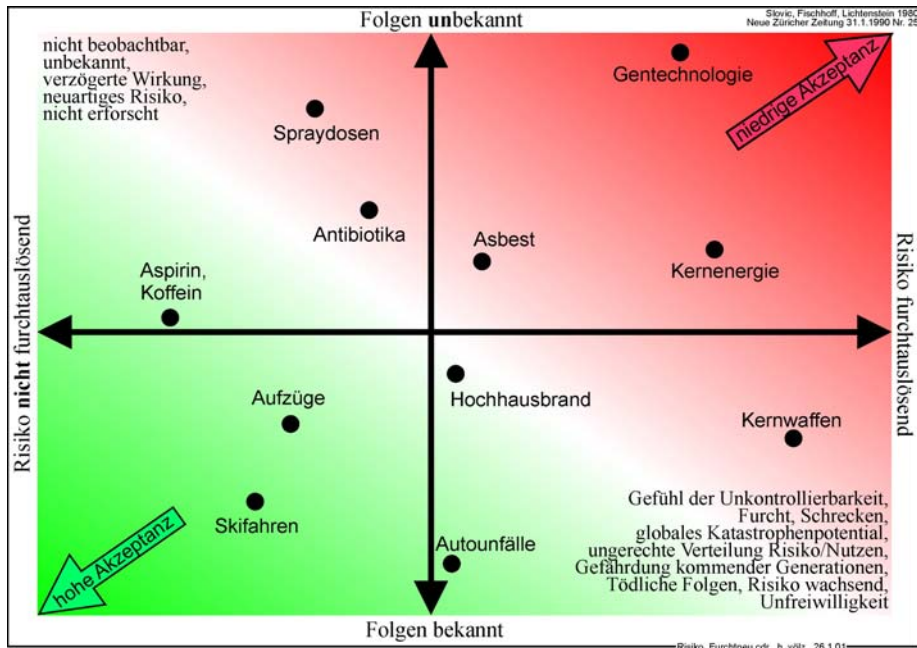
1. Wahrscheinlichkeit des Eintretens eines negativen Ereignisses × finanziellen Ausmaß des Schadens
2. Sicherheit, bzw. Gefahr und Chance (Gewinn), risikohaftes Geschehen (Gefahr) möglichst gut vermeiden
3. Optimieren: Oft gibt es gleichzeitig mehrere Risiken, dann ist das größere zu vermeiden

Fehleinschätzungen von Risiken können durchaus zu sehr negativen Folgen, ja **Katastrophen** führen
Ein sehr kleines und unvermeidliches Risiko wird oft **Restrisiko** genannt

Beispiel für Klassen von Risiko [Bonß]

- 1 KKW, Staudämme, Meteoriten, Überschwemmungen
- 2 Erdbeben, Vulkanismus, Aids, Resistenzen
- 3 BSE, Klimaänderung, neue Seuchen, Gentechnik
- 4 FCKW, Ozonloch, neue Chemikalien, Verlust der Biodiversität
- 5 Mutagene Wirkungen, langfristige Folgen; Klima-Veränderungen
- 6 Ionisierende Strahlung, „Elektrosmog“, Krebs erregende Stoffe

	Schadensausmaß	Mobilisierung gesellschaftl. Konflikte	Beständigkeit, Allgegenwärtigkeit	Reversibilität	Abschätzungssicherheit	Verzögerungswirkung	Wahrscheinlichkeit des Schadens
1	gegen unendlich	eher hoch	eher hoch		groß		gegen Null
2	groß und bekannt		eher hoch	niedrig	klein		ungewiss
3	ungewiss, evtl. hoch		hoch		sehr klein		ungewiss
4	nur Vermutungen		sehr hoch	sehr niedrig	klein		ungewiss
5	eher hoch				mittel	sehr groß	eher hoch
6	eher gering	sehr hoch	mittel bis gering	mittel bis gering	mittel bis gering	mittel bis gering	ungewiss

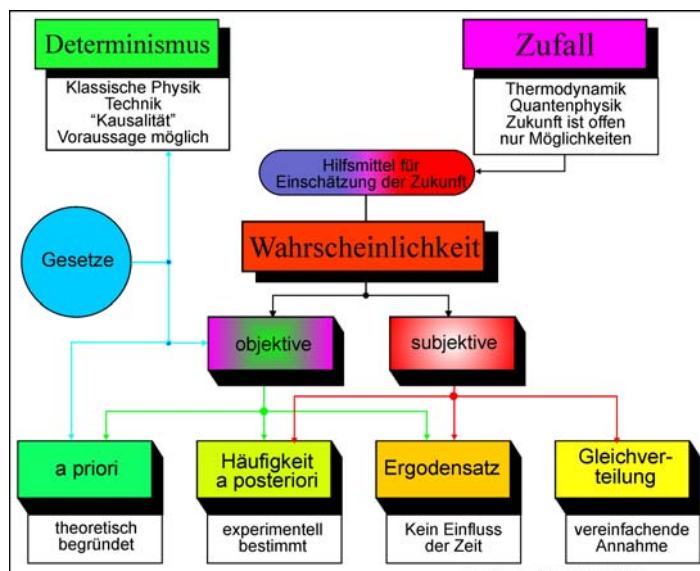


Mutationen usw.

Bei dominant vererbten Mutationen liegt die Mutationsrate bei 10^{-4} bis 10^{-6}
 Mutationsrate eines 40-jährigen ist etwa 10 Mal so hoch eine 20-jährigen Vaters
 Alkoholismus der Mutter führt zu erhöhter Missbildung der Kinder, Bei Nikotin wird es vermutet
 1990 sind ca. 60 000 chemische Substanzen im Handel, die jährlich um etwa 100 vermehren
 Größte Gefahr nicht in genetische Veränderungen sondern krebserregend
 Rot-Grün-Schwäche an X-Chromosom gekoppelt, daher nur für Männer wirksam

Zufall umdenken: **Sterne am Himmel** zufällig → Mensch konstruiert Sternbilder konstruiert

Glaube und Aberglaube: Katzen von rechts und Spinnen am Morgen Sorgen, Pech; Klopfen auf Holz, Freitag, der 13.
Weissagungen und Prophetien: Orakel, Horoskope, Handlesen, Wahrsagen



Literatur

- Ackermann, W.: Zuverlässigkeit elektronischer Bauelemente. Hüthig Verlag, Heidelberg 1976
 Arnoult, M. D.: Familiarity and Recognition of Nonsense Shapes. J. experimental Psychology, 51(1956), 4, 269-276
 Ash, D., Hewitt, P.: Wissenschaft der Götter - zur Physik des Übernatürlichen. Zweitausendundeins, Frankfurt/Main 1991
 Attnave, F.: Application of Information Theory to Psychology. Henry Holt & Co. New York 1959
 Barrow, J. D.: Ein Himmel voller Zahlen - Auf den Spuren mathematischer Wahrheit Spektrum-Verlag. Heidelberg - Berlin - Oxford, 1994

- Barrow, J. D.: Theorien für alles. Spektrum-Verlag; Heidelberg - Berlin - Oxford, 1992
- Beyer, O, Hackel, H., Pieper, V. u. Tiedte, J.: Wahrscheinlichkeitsrechnung und mathematische Statistik. B. G. Teubner, Leipzig 1976
- Böhme, J. F.: Stochastische Signale. Teubner, Stuttgart 1998
- Bongard, M. M.: Über den Begriff der nützlichen Information. Probleme der Kybernetik 6(1966), 91-130 (russ. Original 1963)
- Bonß, W.: Vom Risiko - Unsicherheit und Ungewissheit in der Moderne. Hamburger edition, , Hamburg, 1995
- Bosch, K.: Lotto und andere Zufälle. Vieweg, Braunschweig - Wiesbaden 1995.
- Bronstein, I.N., Semendjajew, K.A. u.a.: Taschenbuch der Mathematik. Verlag Hary Deutsch 1997
- Büttenmeyer, W.: Wissenschaftstheorie für Informatiker. Spektrum-Verlag. Heidelberg - Berlin - Oxford, 1995
- Creutz, G., Ehlers, R.: Statistische Formelsammlung. Deutzsch-Verlag Thun - Frankfurt 1976
- Dreszer, J.: Mathematik Handbuch. Fachbuchverlag Leipzig 1975
- Eberlein, G. L.: Schulwissenschaft - Parawissenschaft - Pseudowissenschaft. S. Hirzel, Stuttgart 1991
- Estoup, J. B.: Gammes sténographiques, Paris 1916
- Eusemann, E.: Naturkatastrophen - Das Lexikon zu ihren Ursachen und Folgen. Meyers Lexikon-Verlag. Mannheim - Leipzig - Wien - Zürich 1999
- Faletta, N.: Paradoxien. Fischer Taschenbuch-Verlag 1990
- Fano, R. M.: Informationsübertragung, Oldenbourg, München, Wien 1966 (engl. Original 1961)
- Feyerabend, P.: Wissenschaft als Kunst, Suhrkamp, es 1231, Frankfurt/M 1984
- Frank, H.: Kybernetische Grundlagen der Pädagogik. Agis-Verlag, Baden-Baden, 1969
- Friedrichs, B.: Kanalcodierung; Springer 1996, wohl bestes und gut verständliches Buch für digitale Signale, Codierung, Fehlerbehandlung, auch neueste Codes
- Fritzsche, G.: Theoretische Grundlagen der Nachrichtentechnik. 2. Aufl. 1972. brauchbare und verständliche Einführung in (vor allem analoge) Systeme, Signale und Informationstheorie
- Fucks, W.: Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen. Westdeutscher Verlag, Köln - Opladen 1955
- Gell-Mann, M.: Das Quark und der Jaguar. Piper München 1994.
- Gerdson, P. u. Kröger, P.: Digitale Signalverarbeitung in der Nachrichtenübertragung. Springer 1993. Etwas stärker hardwarebezogene (mit formalen Schaltbildern und Programmbeispielen) Betrachtung
- Gerthsen, C., Kneser, H. O., Vogel, H.: Physik - Ein Lehrbuch. Springer Verlag Berlin u.a. 1989
- Gnedenko; Chintchin: Elementare Einführung in die Wahrscheinlichkeitsrechnung, S. 158
- Gränicher, W. H. H.: Messung beendet - was nun? B. G. Teubner, Stuttgart 1996
- Grimm, H., Recknagel, R.-D.: Grundkurs Biostatistik. G. Fischer-Verlag Jena 1985
- Härtler, G.: Statistische Methoden für die Zuverlässigkeitsanalyse. Verlag Technik, Berlin 1983
- Hell, W., Fiedler, K., Gigenzer, G. (Hrsg.): Kognitive Täuschungen. Spektrum -Akademischer Verlag, Heidelberg - Berlin - Oxford 1993
- Hersch, J.: Das Philosophische Staunen. Benzinger, Zürich 1981. ISBN 3-545-37092-5
- Hesse, H.: Das Glasperlenspiel. 1943
- Hilberg, W.: Das Netzwerk der menschlichen Sprache und Grundzüge einer entsprechend gebauten Sprachmaschine, ntz Archiv, 10 (1988) H. 6. S. 133-146
- Hilberg, W.: Die texturale Sprachmaschine als Gegenpol zum Computer. Sprache und Technik, Großbieberau 1990
- Höfer, G., Klein, W.: Wahrscheinlich ganz einfach. Urania-Verlag, Leipzig - Jena - Berlin 1986
- Hoffrage, U.: Die Illusion der Sicherheit bei Entscheidungen unter Unsicherheit in Hell, W., Fiedler, K., Gigenzer, G. (Hrsg.): Kognitive Täuschungen. Spektrum -Akademischer Verlag, Heidelberg - Berlin - Oxford 1993 ; S. 73- 97
- Hyvärinen, L. P.: Information Theory for Systems Engineers. Springer, Berlin- Heidelberg - New York 1970
- Ineichen, R.: Würfel und Wahrscheinlichkeit, Stochastisches Denken in der Antike; Spektrum Akademischer Verlag, Heidelberg - Berlin - Oxford, 1996
- Klix, F.: Erwachendes Denken. 2. Aufl. Deutscher Verlag der Wissenschaften, Berlin 1983, Spektrum Akademischer Verlag, Heidelberg - Berlin - Oxford 1993
- Lichtenstein, S. Fischhoff, B. & Phillips, L. D.,(1982). Calibration of probabilities: The state of art to 1980. In D. Kahneman, P. Slovic & A. Tversky (Hrsg.), Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.
- Mandelbrot, B. B.: On the theory of word frequencies and on related Markovian models of discourse. In R. Jacobson (Hrsg.): Structures of language and its mathematical aspects (American Mathematical Society, New York, 1961)
- Mandelbrot, B.: Linguistique statistique macroscopique I, In Piaget, J. (Hrsg.): Language et Théorie de l'information. Presses universitaires de France, Paris 1957
- Mandelbrot, B.: structure formelle des textes et communication. Word 10(1954) 1-27
- Mann, Thomas: Doktor Faustus – Das Leben des deutschen Tonsetzers Adrian Leverkühn erzählt von einem Freunde, Aufbau Verlag Berlin 1952; S. 261/2
- Manteuffel, K, Stumpe, D.: Spieltheorie. B.G. Teubner, Leipzig 1977
- Marquard, O.: Abschied vom Prinzipiellen. Reclam, Stuttgart 1981
- Marquard, O.: Apologie des Zufälligen. Philipp Reclam jun. Stuttgart 1986
- Marquard, O.: Philosophie des Stattdessen. Reclam, Stuttgart, 2000
- Marquard, O.: Skepsis und Zustimmung. Reclam, Stuttgart 1994
- May, R. S.:(1987). Realismus von subjektiven Wahrscheinlichkeiten: Eine kognitionspsychologische Analyse inferentieller Prozesse beim Overconfidence-Phänomen. Frankfurt/M.: Lang.
- Mehlmann, A.: Wer gewinnt das Spiel? - Spieltheorie und Fabeln und Paradoxa. Vieweg, Braunschweig - Wiesbaden, 1997
- Meyer, J.: Die Verwendung hierarchisch strukturierter Sprachnetzwerke zur redundanzarmen Codierung von Texten. Dissertation TH Darmstadt 1960

- Meyer, J.: Gilt das Zipfsche Gesetz auch im Chinesischen, NTZ Archiv 11(1989), 1; 13-16
- Meyer-Eppler, W.: Grundlagen und Anwendungen Informationstheorie. Springer Verlag, Berlin - Heidelberg - New York 1969
- Milde, T.: Videokompressionsverfahren im Vergleich, dpunkt 1995. Für JPEG, MPEG Wavelets und Fraktale wichtig
- Miller, G. A.: Streifzüge durch die Psycholinguistik, Zweitausendeins 1993
- Modis, T.: Die Berechenbarkeit der Zukunft; Birkhäuser; Basel - Boston - Berlin, 1994
- Moles, A.: Information Theory and Esthetic Perception, University of Illinois Press, Urbana usw. 1968
- Morgenstern, D.: Einführung in die Wahrscheinlichkeitsrechnung und mathematische Statistik. Springer Verlag, Berlin - Heidelberg - New York 1968
- N.N.: Kleine Enzyklopädie Mathematik. Bibliographisches Institut Leipzig 1977
- N.N.: Wir wiederholen Wahrscheinlichkeitsrechnung. Fachbuchverlag Berlin 1982
- Neidhardt, P.: Informationstheorie und automatische Informationsverarbeitung. Verlag Technik Berlin 1964
- Neuman, J. v.: Zur Theorie der Gesellschaftsspiele. Mathematische Analen 100(1928), 295-300
- Neumann, J. v.; Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton 1944
- Nicolis, G., Prigogine, I.: Die Erforschung des Komplexen, Piper, München - Zürich 1987
- Noble, C. E.: The familiarity-frequency Relationship. J. experimental Psychology, 47(1954), 13-16
- Ohm, J.-R.: Digitale Bildcodierung. Springer, 1995. Vor allem für Codierung und Komprimierung
- Paulus, J. A.: Zahlenblind - mathematisches Analphabetentum und seine Folgen. W. Heyne - Verlag, München, 1990
- Peterson, W. W.: Prüfbare und korrigierbare Codes. Oldenbourg 1967, grundlegendes Werk, beste Einführung in das Gebiet, neue Codes fehlen allerdings
- Piatelli-Palmarini, M.: Die Illusion zu wissen, rororo, Reinbeck b. Hamburg 1997
- Planck, M.: Das Weltbild der neuen Physik. Barth, Leipzig 1929
- Planck, M.: Der Kausalitätsbegriff in der Physik. Barth, Leipzig 1965
- Planck, M.: Determinismus oder Indeterminismus. Barth, Leipzig 1938
- Planck, M.: Die Physik im Kampf um die Weltanschauung. Barth, Leipzig 1935
- Planck, M.: Scheinprobleme der Wissenschaft. Barth, Leipzig 1947
- Planck, M.: Sinn und Grenzen der exakten Wissenschaft. Barth, Leipzig 1942
- Planck, M.: Vom Wesen der Willensfreiheit. Barth, Leipzig 1936
- Prigogine, I., Stengers, I.: Dialog mit der Natur; Piper; München - Zürich 1983
- Rényi, A.: Wahrscheinlichkeitsrechnung mit einem Anhang über Informationstheorie. Deutscher Verlag der Wissenschaften, Berlin 1962
- Rescher, N.: Glück -- Die Chancen des Zufalls. Berlin-Verlag, Berlin, 1996. Original: Luck: the brilliant randomness of everyday life, 1995
- Riedel, R.: Wahrheit und Wahrscheinlichkeit. Parrey, Berlin - Hamburg 1992
- Rumpf, K.-H.: Trommeln - Telefone - Transistoren. Verlag Technik, Berlin 1971
- Schreier, W.: (Hrg.) Geschichte der Physik - Ein Abriß. Deutscher Verlag der Wissenschaften, Berlin 1988
- Schuster, H.G.: Deterministisches Chaos, Verlag Chemie, Weinheim usw. 1994
- Sedgewick, R.: Algorithmen. Addison-Wesley, 1992
- Shannon, Cl. u. Weaver, W.: Mathematische Grundlagen der Informationstheorie. Oldenbourg 1976. Meines Erachtens unverzichtbares Buch. Weaver sollte vergessen werden
- Sheldrake, R.: Das Gedächtnis der Natur, Scherz, Bern - München - Wien 1993
- Siegmund, G.: Grundlagen der Vermittlungstechnik 2. Aufl. R. v. Decker's Verlag, G. Schenk, Heidelberg 1993
- Sobol, I.M.: Die Monte-Carlo-Methode. Deutscher Verlag der Wissenschaften, Berlin 1983
- Stadler, E.: Modulationsverfahren, Vogel-Verlag 7. Aufl. 1993, Besonders für die analogen Verfahren gut geeignet
- Stegmüller, W.: (1973). Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie. Bd. IV - Personelle und Statistische Wahrscheinlichkeit. Teil E. Berlin: Springer.
- Stoyan: Stochastik für Ingenieure und Naturwissenschaftler
- Struik, D. J.: Abriß der Geschichte der Mathematik. Deutscher Verlag der Wissenschaften Berlin 1972
- Tarassow, L.: Wie der Zufall will? Spektrum Akademischer Verlag, Heidelberg – Berlin 1993
- Täubert, P.: Abschätzung der Genauigkeit von Messergebnissen. Verlag Technik. Berlin 1985
- Unbehauen, R.: System-Theorie. Oldenbourg 1980. Wichtige Einführung für die Signal-Betrachtungen
- Vollmer, G.: Auf der Suche nach Ordnung, S. Hirzel, Stuttgart 1995
- Völz, H.: Elektronik - Grundlagen - Prinzipien - Zusammenhänge. 5. Aufl. Akademie Verlag, Berlin 1989
- Völz, H.: Information I: Akademie Verlag Berlin, 1982
- Völz, H.: Wissen - Erkennen - Information. Allgemeine Grundlagen für Naturwissenschaft, Technik und Medizin. Shaker Verlag, Aachen 2001. Zusätzlich weitere Bände enthält die CD: H. Völz: Wissen - Erkennen - Information. Datenspeicher von der Steinzeit bis ins 21. Jahrhundert. Digitale Bibliothek Bd. 159, Berlin 2007
- Voss, Kl.: Statistische Theorie komplexer Systeme, EIK 5(1969) 4/5, 234-254 u. 6, 319-330
- Wallsten, T., S. & Budescu, D. V. (1983) State of Art. Encoding subjective probabilities: A psychological and psychometric review. Management Science, 29, 151 - 173.
- Weizsäcker, C. F.: Zeit und Wissen, dtv - Wissenschaft, Hanser-Verlag, München - Wien 1992
- Wygodski, M. J.: höhere Mathematik griffbereit. Akademie Verlag, Berlin, 1976
- Young, J.F.: Einführung in die Informationstheorie. R. Oldenbourg, München - Wien 1975
- Zipf, G. K.: Human behaviour and the principle of heart effort. Addison-Wesly, Cambridge, Mass. 1949
- Zipf, G. K.: The Psycho-biology of language. Mifflin, Houghten 1935