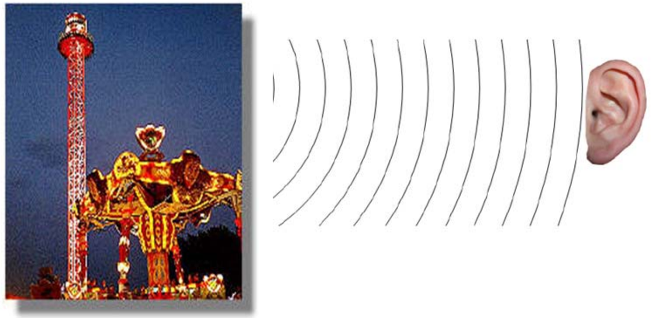
Stochastikunterricht

mit Technologieeinsatz



Streckenlängen mit Google-Earth:

Blau: 165 m

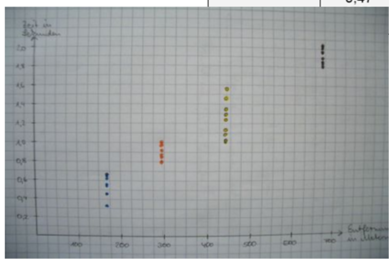
Rot: 295 m

Gelb: 440 m

Grau: 670 m

Einführung

Wann erreicht der Schall der Starterklappe unseren Standort?				
Entfernung	Gruppe 1 165 m	Gruppe 2 295 m	Gruppe 3 440 m	Gruppe 4 670 m
Zeiten [sec]	0,66	0,99	1,35	2,03
	0,68	0,91	1,57	1,96
	0,54	1,01	1,29	1,84
	0,56	0,80	1,46	1,97
	0,46	0,86	1,09	1,90
	0,47	0,85	1,13	2,04
		0,97	1,03	1,86
		0,88	1,25	1,82



„Das geht irgendwie so gerade nach oben."

Wie?

Gliederung

1. Einstiegsbeispiel
2. Wieso, weshalb, warum?
3. Aufgaben – aber wie (mit/ohne Rechner)?
   - Visualisieren und Rechnen: Strukturen suchen und Datenmasse bewältigen
   - Simulieren: Experimentieren, Testen, Schätzen, Verstehen
   - Modellieren: Fragen beantworten statt Algorithmen abarbeiten, Strukturen erfassen und abbilden
4. Ideen hinter den Aufgaben und hinter dem Rechner
5. Werbeblock

2. Wieso, weshalb, warum?

Leitidee Daten und Zufall (Sek. I): S. u. S.

- werten graphische Darstellungen und Tabellen von statistischen Erhebungen aus
- planen statistische Erhebungen entsprechend der zu untersuchenden Fragestellung
- sammeln systematisch Daten, erfassen sie in Tabellen und stellen sie graphisch dar, auch unter Verwendung geeigneter Hilfsmittel (wie Software)
- interpretieren Daten unter Verwendung von Kenngrößen reflektieren und bewerten Argumente, die auf einer Datenanalyse basieren
- beschreiben Zufallserscheinungen in alltäglichen Situationen
- bestimmen Wahrscheinlichkeiten bei Zufallsexperimenten

Die Leitidee Daten und Zufall kann prozessbezogene Kompetenzen fördern

- Modellieren (mit dem Rechner)
- Visualisieren (mit dem Rechner)
- Simulieren (mit dem Rechner)

2. Wieso, weshalb, warum?

Die Leitidee Daten und Zufall ist dafür geeignet, dass Schülerinnen und Schüler erfahren,

Fragen an alltägliche empirische Phänomene zu stellen und mit den elementaren Methoden der Sekundarstufe I (und dem Rechner) zu beantworten

Und Sek II ?

Leitidee Daten und Zufall (Sek. II): S. u. S.

Hoffentlich: Fragen an alltägliche empirische Phänomene zu stellen und mit den elementaren Methoden der Sekundarstufe II (und dem Rechner) zu beantworten

- ???
- ??? (hoffentlich nicht nur Binomialverteilungsalgorithmen)

Gliederung

1. Einstiegsbeispiel
2. Wieso, weshalb, warum?
3. Aufgaben – aber wie (mit/ohne Rechner)?
   - Visualisieren und Rechnen: Strukturen suchen und Datenmasse bewältigen
   - Simulieren: Experimentieren, Testen, Schätzen, Verstehen
   - Modellieren: Fragen beantworten statt Algorithmen abarbeiten, Strukturen erfassen und abbilden
4. Ideen hinter den Aufgaben und hinter dem Rechner
5. Werbeblock

3. Aufgaben – aber wie?

Strukturen suchen und Masse bewältigen

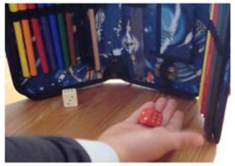
Aufgabe:
Welche Eigenschaften haben Studierende/ SchülerInnen

Der Rechner
nach händischen Bearbeitungen
als Rechenknecht und
für eine grafisch gesteuerte Analyse

Gliederung

3. Aufgaben – aber wie?

Experimentieren



Modell für die Würfel

Augenzahl	1	2	3	4	5	6
Wahrscheinlichkeit	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
Wahrscheinlichkeit	0,05	0,1	0,35	0,35	0,1	0,05

0. $P(N) = P(Q) = 0{,}5$

1. Es fällt die Augenzahl x

2. $P(N|x) = \dfrac{P(x|N) \cdot P(N)}{P(x|N) \cdot P(N) + P(x|Q) \cdot P(Q)}$

Verarbeiten der Information (x) und Anwendung der Formel von Bayes führt zu Neubewertung

Aufgabe:
Die Spielleitung wählt verdeckt den normalen oder den quaderförmigen Würfel aus.

Welcher ist es?

Der Rechner

nach händischen Bearbeitungen
für die unmittelbare Auswertung und Vertiefung des Verständnisses

3. Aufgaben – aber wie?

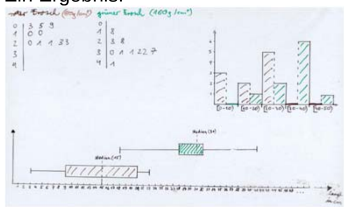
Testen

80 g/m²-Papier

160 g/m²-Papier

Wer springt weiter?

Ein Ergebnis:



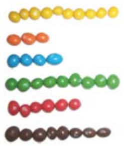
„Ist das jetzt immer so?"

Rechnergestützte Simulationsmöglichkeiten:

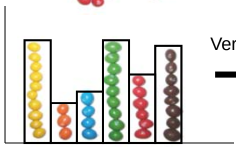
3. Aufgaben – aber wie?

Ordnen

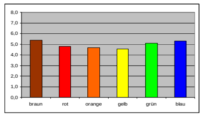
Darstellen



Abstrahieren



Verallgemeinern



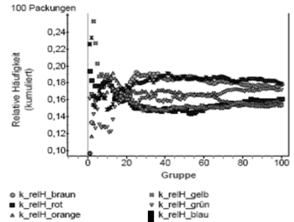
- Erhebung (Beobachtung) planen
- Daten grafisch darstellen
- Mittelwerte
- Einfluss der Stichprobengröße
- …

3. Aufgaben – aber wie?

Prognostizieren

8 Personen öffneten ihr Säckchen und zählten:

	br	rot	orange	gelb	grün	blau
z.B.	5	0	4	4	4	0
	5	3	2	6	2	0

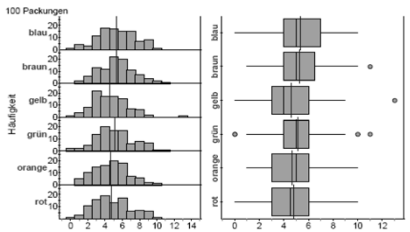


Aufgabe:

Stellt anhand der Untersuchung einiger Schokolinsen-Packungen eine Prognose über die Häufigkeit der einzelnen Farben – insbesondere der roten Linsen – auf und überprüft Eure Prognose.

(Übergang zum Zufall)

3. Aufgaben – aber wie?

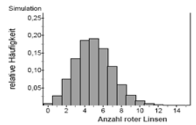


Modell:
In einer Packung sind durchschnittlich 30 Schokolinsen und im Durchschnitt je 5 Linsen einer Farbe, die Farbverteilung ist unabhängig.

3. Aufgaben – aber wie?

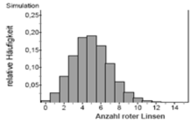
Hieb- und Stichaufgaben:
Wenn das Modell stimmt, wie groß wäre dann die Wahrscheinlichkeit, in einer Packung
- genau 3 rote Kugeln zu erhalten
- mindestens eine rote Kugel zu erhalten
- Baum, $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$

Zurück zur Realität:
Wenn das Modell stimmt, welche Wahrscheinlichkeit haben dann die verschiedenen Anzahlen roter Kugeln in einer Packung?



3. Aufgaben – aber wie?

Zurück zur Realität:
Wenn das Modell stimmt, welche Wahrscheinlichkeit haben dann die verschiedenen Anzahlen roter Kugeln in einer Packung?



Informeller Hypothesentest:
Ab welcher Anzahl von roten Kugeln in einer Packung könnte/sollte man an dem Modell der Gleichbefüllung zweifeln?
Welchen Fehler könnte man begehen?

3. Aufgaben – aber wie?

1. Datenanalyse
Farbverteilung beschreiben

2. Modell (“ist das immer so?“)
Statistische Wahrscheinlichkeit für die Farben

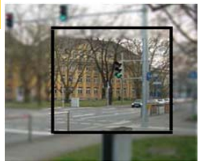
3. Wahrscheinlichkeitsrechnung
Was wäre, wenn das Modell stimmen würde, welche Prognose könnten man für … treffen (Arbeiten im Modell)

4. „Informelle“ Inferenz
Simulation des Modells. Welche Ereignisse sprechen gegen das Modell? Durchführen eines einzelnen Versuchs

5. Formelle Inferenz
Tests

3. Aufgaben – aber wie?

Schätzen

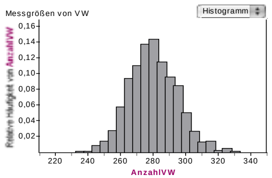


Aufgabe:
Verkehrszählung

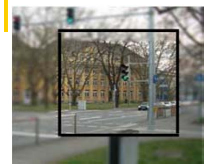
Anzahl der VW

Zufall:
- h(VW) ≈ P(VW) = 0,33; **385** bei 1166 gezählten PKW in Braunschweig
- Sind das genau 33 %?

- Wäre P(VW) = 0,25 auch möglich?
- Hieb- und Stichaufgabe
- Simulation oder Berechnung mit p = 0,25 und n = 1166



3. Aufgaben – aber wie?



Zufall:
Konfidenzintervall (Approximation durch die Normalverteilung, 95%-Niveau):
$0{,}303 < p < 0{,}357$
Interpretation (!)

Wann (Wo) würde man an diesem Modell zweifeln?
- Niedersachsen: p ≈ 0,35
- Deutschland: p ≈ 0,20
- Wie würde eine solche Zählung in Stuttgart/München/Köln … enden?

Weitere Simulationen als Schlüssel zum Verständnis

Gliederung

1. Einstiegsbeispiel
2. Wieso, weshalb, warum?
3. Aufgaben – aber wie (mit/ohne Rechner)?
   - Visualisieren und Rechenmaschine: Strukturen suchen und Datenmasse bewältigen
   - Simulieren: Experimentieren, Testen, Prognostizieren, Schätzen, Verstehen
   - Modellieren: Fragen beantworten statt Algorithmen abarbeiten, Strukturen erfassen und abbilden
4. Ideen hinter den Aufgaben und hinter dem Rechner
5. Werbeblock

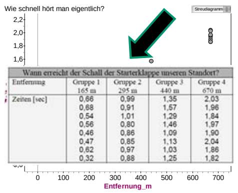
3. Aufgaben – aber wie?

Zeitgemäßes Modellieren mit der Binomialverteilung

Produktion:
Bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla **im Durchschnitt 2%** bla bla bla bla bla bla bla bla bla bla bla bla. Bla bla bla bla bla bla. Wie groß ist die Wahrscheinlichkeit dafür, dass **von den 20** bla bla bla bla bla bla **genau 3** bla bla bla bla bla bla ?“

Text: irrelevant
Sachsituation: irrelevant
Daten: irrelevant
Modell Binomialverteilung, Zahlen: relevant
Rechner: eher irrelevant, Rechenknecht

3. Aufgaben – aber wie?

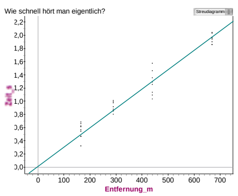


Wann erreicht der Schall der Starterklappe unseren Standort?				
Entfernung	Gruppe 1 165 m	Gruppe 2 295 m	Gruppe 3 440 m	Gruppe 4 670 m
Zeiten [sec]	0,66	0,99	1,35	2,03
	0,68	0,91	1,57	1,96
	0,54	1,01	1,29	1,84
	0,56	0,80	1,46	1,97
	0,46	0,86	1,09	1,90
	0,47	0,85	1,13	2,04
	0,62	0,97	1,03	1,86
	0,32	0,88	1,25	1,82

Daten:

vermitteln Phänomene der natürlichen, technischen und sozialen Umwelt

Reale Modellebene

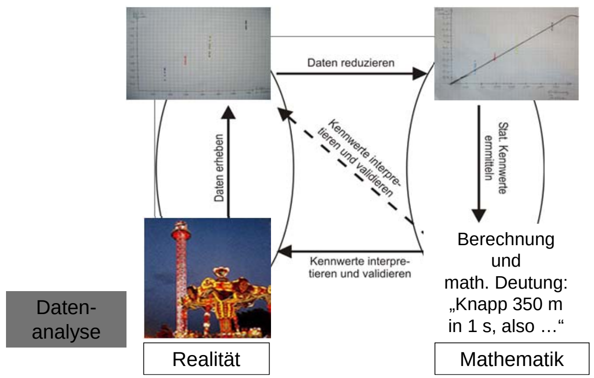


Funktion:

bildet als „deterministisches Modellierungskonzentrat“ (nur) den Datentrend ab.

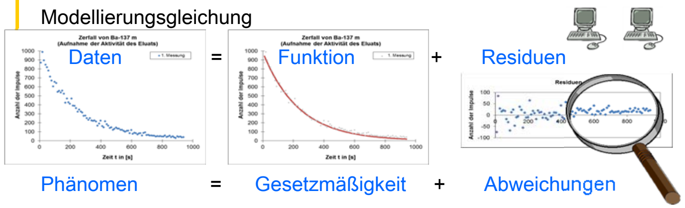
Mathematische Modellebene

3. Aufgaben – aber wie?



3. Aufgaben – aber wie?

Modellierungsgleichung



Daten = Funktion + Residuen

Phänomen = Gesetzmäßigkeit + Abweichungen

Residuen nicht nur als „Fehler“ gering schätzen, sondern als Modellierungshilfe betrachten. Sie sollen möglichst

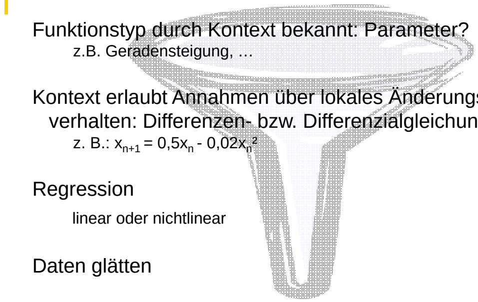
• klein, zufällig (im Sinne von trendfrei) sein

• sowie sich nach oben und unten ausgleichen

Funktionen als Werkzeug zur Umwelterschließung

3. Aufgaben – aber wie?

$(x_1, y_1), \ldots, (x_n, y_n)$



Funktionstyp durch Kontext bekannt: Parameter?

z.B. Geradensteigung, …

Kontext erlaubt Annahmen über lokales Änderungsverhalten: Differenzen- bzw. Differenzialgleichungen

z. B.: $x_{n+1} = 0{,}5x_n - 0{,}02x_n^2$

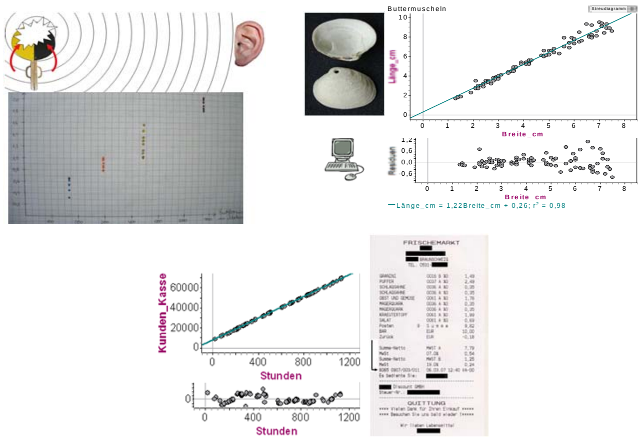
Regression

linear oder nichtlinear

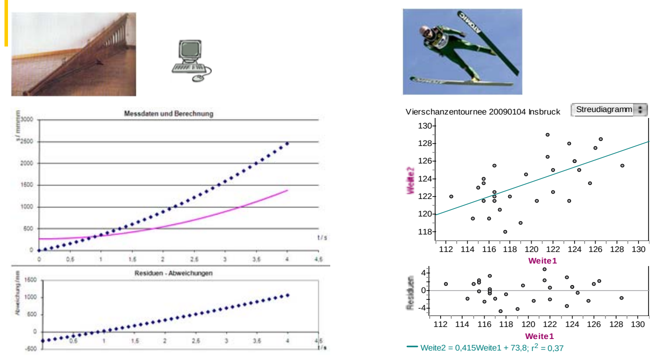
Daten glätten

Funktion

3. Aufgaben – aber wie?



3. Aufgaben – aber wie?



Fragen beantworten statt Algorithmen abarbeiten - Elementarisierung

1. Der standardisierte Weg

Datensammlung

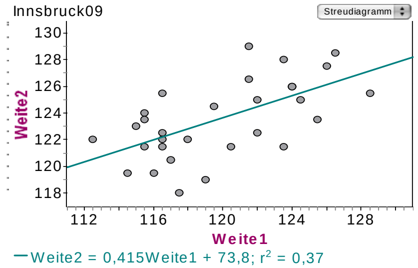
Platz	Name	Weite 1	Weite 2	Platz	Name	Weite 1	Weite 2
1	Schmitt	128,5	125,5	16	Evensen	119	119
2	Loitzl	126,5	128,5	17	Watase	118	122
3	Schlierenzauer	126	127,5	18	Eggenhofer	117,5	118
4	Amman	125,5	123,5	19	Larinto	117	120,5
5	Morgenstern	124,5	125	20	Uhrmann	116,5	125,5
6	Kasai	124	126	21	Hilde	116,5	122,5
7	Neumayer	124	126	22	Ito	116,5	121,5
8	Hautamaeki	123,5	128	23	Schoft	116,5	122
9	Rosliakow	123,5	121,5	24	Stoch	116	119,5
10	Olli	122	125	25	Hocke	115,5	124
11	Koch	122	122,5	26	Koudelka	115,5	123,5
12	Vassilev	121,5	129	27	Lackner	115,5	121,5
13	Jacobsen	121,5	126,5	28	Yumoto	115	123
14	Malysz	120,5	121,5	29	Kofler	114,5	119,5
15	Kuettel	119,5	124,5	30	Tochimoto	112,5	122

Aufgabe:
Gibt es eigentlich einen Zusammenhang zwischen erster und zweiter Sprungweite?

Fragen beantworten statt Algorithmen abarbeiten - Elementarisierung

1. Der standardisierte Weg

Regressionsgerade + Korrelationskoeffizient



Fragen beantworten statt Algorithmen abarbeiten - Elementarisierung

1. Der standardisierte Weg

Regressionsgerade + Korrelationskoeffizient

$$a = \frac{\sum_{i=1}^{n}((x_i - \overline{x}) \cdot (y_i - \overline{y}))}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{Cov(X,Y)}{s_X^2}$$

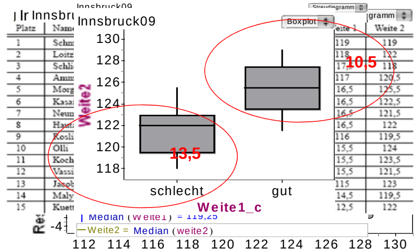
$$b = \overline{y} - a \cdot \overline{x}$$

$$r = \frac{Cov(X,Y)}{s_X \cdot s_Y}$$

?

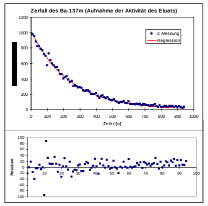
Fragen beantworten statt Algorithmen abarbeiten - Elementarisierung

2. Alternativer Weg

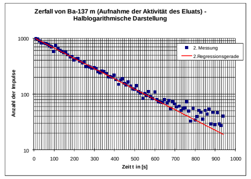


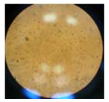
$$r_z = \frac{n^+ - (n - n^+)}{n} = \frac{24 - 6}{30} = 0{,}6 \quad (r \approx 0{,}61)$$

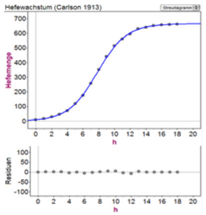
3. Aufgaben – aber wie?



Hilfsüberlegung: Logarithmieren







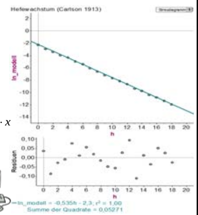
Die logistische Funktion

$$f(x)=\frac{G}{(G-y_0)\cdot e^{-G\cdot c\cdot x}+y_0}$$

umformen nach:

$$\ln\left(\frac{1}{f(x)}-\frac{1}{G}\right)=\ln\left(\frac{G-y_0}{G\cdot y_0}\right)-c\cdot G\cdot x$$

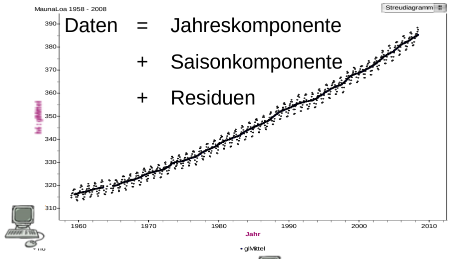
Wenn das logistische Modell passt, sollte sich eine lineare Struktur ergeben können.



3. Aufgaben – aber wie?

Atmosphärischer CO_2-Gehalt:

Analyse: Daten = Jahreskomponente + Saisonkomponente + Residuen



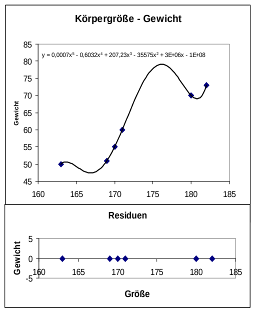
Glätten

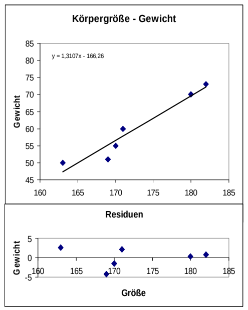
Modellieren mit param. Funktionen

Nichtlineare Regression: Minimum von $g(\theta)=\sum_{i=1}^{n}\left(f(x_i,\theta)-y_i\right)^2$

3. Aufgaben – aber wie?

Modellierungsgüte: Keine Residuen – das beste Modell?



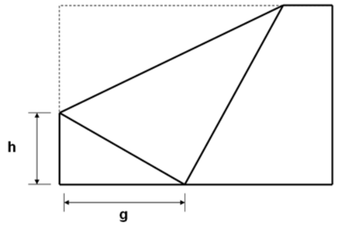


"Essentially, all models are wrong, but some are useful." (Box & Draper, 1987)

3. Aufgaben – aber wie?

Papierfaltproblem:

Schülerauftrag:
Nehmt Euch ein DIN A4-Blatt. Faltet die linke obere Ecke und die untere Seite so, wie es im Bild oben zu sehen ist. Es entsteht ein rechtwinkliges Faltdreieck mit den Seiten g und h sowie der sich daraus ergebenden Fläche F.



Führt zwanzig Messungen durch und untersucht den Zusammenhang zwischen einer der beiden Seitenlängen und dem Flächeninhalt.

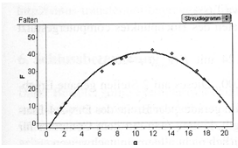
Papierfaltproblem:

Falten

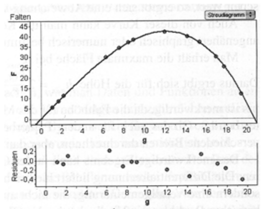
	g	h	F
=			$\frac{g \cdot h}{2}$
1	14,2	5,6	39,76
2	1,7	10,2	8,67
3	6,3	9,5	29,925
4	9,0	8,5	38,25
5	17,0	3,5	29,75
6	1,1	10,4	5,72
7	2,2	10,3	11,33
8	7,6	9,0	34,2
9	15,4	4,8	36,96
10	17,8	2,8	24,92
11	19,5	1,2	11,7
12	8,5	8,7	36,975
13	12,0	7,0	42

Abb. 2: Datentabelle

Quadratische Anpassung:



Kubische Anpassung:



Papierfaltproblem:

Überlegungen zur kubischen Anpassung:

$h + x = 21, x^2 = h^2 + g^2$

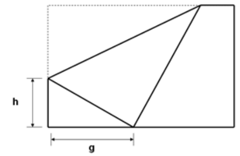
$(h-21)^2 = h^2 + g^2$

$h^2 - 2 \cdot 21h + 21^2 = h^2 + g^2$

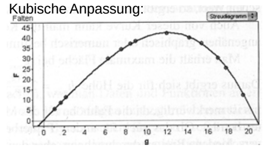
$h = -\frac{g^2}{2 \cdot 21} + \frac{21^2}{2 \cdot 21}$

$$F(g,h) = \frac{1}{2} g \cdot h = \frac{1}{2} g \cdot \left(-\frac{g^2}{2 \cdot 21} + \frac{21^2}{2 \cdot 21} \right)$$

$$F(g) = -\frac{1}{4 \cdot 21} g^3 + \frac{21^2}{4 \cdot 21} g = -\frac{1}{4 \cdot 21} g(g^2 - 21^2)$$
$$= -\frac{1}{84} g(g-21)(g+21)$$



Kubische Anpassung:



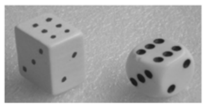
Wer sucht, der findet

Aufgaben zur realen Realität

Aufgaben zu konstruierten realen Situationen

BUNDES LIGA

Konstruierte Situationen



1. Einstiegsbeispiel
2. Wieso, weshalb, warum?
3. Aufgaben – aber wie (mit/ohne Rechner)?
   - Visualisieren und Rechenmaschine: Struktursuche und Datenmassen-Bewältigung
   - Simulieren: Experimentieren, Testen, Schätzen, Verstehen
   - Modellieren: Fragen beantworten statt Algorithmen abarbeiten, Strukturen erfassen – „Reste“ nicht vergessen
4. Ideen hinter den Aufgaben und hinter dem Rechner
5. Werbeblock

Aspekte des statistischen Denkens (Pfannkuch/Wild, 1999)

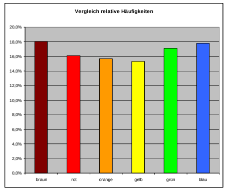
1. Notwendigkeit von Daten

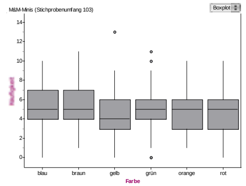
„Ich habe schon immer gewusst, dass Ärzte zu viel verdienen!“

„Die armen Ärzte, das ist ein Skandal“

Daten als Grundlage für einen “guten” Erkenntnisgewinn

2. Flexible Datendarstellungen

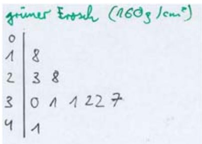


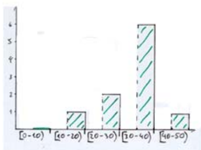


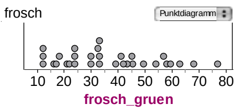
Unterschiedliche Darstellungen der Daten eröffnen unterschiedliche Perspektiven!

	braun	rot	orange	gelb	grün	blau
absolut	553	494	481	469	525	545
relativ	18,0%	16,1%	15,7%	15,3%	17,1%	17,8%

3. Datenstreuung oder Variabilität!







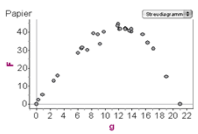
Messungen von Objekten unterscheiden sich! Nicht Uniformität, sondern Variabilität ist Gegenstand stochastischen Denkens.

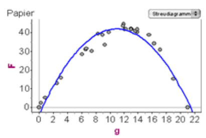
4. Struktursuche, Mustererkennung, Musterbeschreibung

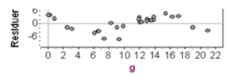
Daten = Trend + Zufall

ist ein Konstrukt, um mit der Variabilität der Daten (erklärte und nicht-erklärte) fertig zu werden.
z. B.:

Daten = Funktion + Residuen



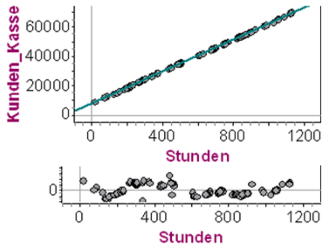




5. Zusammenhang von Zahl und Kontext

$\bar{x} = 54$

Im Durchschnitt passieren 54 Kunden pro Stunde die Kasse 1 in dem beobachten Supermarkt. Das sind … pro Jahr, woraus sich ein Umsatz von … ergibt.



Stochastisches Denken und der Rechner

- ➢Notwendigkeit von Daten – ???
- ➢Flexible Darstellung – mit dem Rechner unmittelbar möglich

- ➢Erkennen der Variabilität – Simulieren zeigt Variabilität
- ➢Muster entdecken – Residuenanalyse setzt den Rechner voraus

- ➢Integration von Daten und Kontext – reale Datenmengen benötigen den Rechner

Rechner und Leitidee Daten und Zufall

Rechner ersetzt nicht das Verständnis, aber hilft bei

- ➢Bewältigen großer Datenmengen
- ➢visuell gesteuerter Datenanalyse
- ➢Simulation
- ➢Elementarisierung konventioneller Methoden

- ➢Vertieftem Verstehen von Zusammenhängen (Interaktive Grafik)

Vieweg+Teubner
ISBN 3-8348-0681-1
ISBN 978-3-8348-0681-9

Vielen Dank für Ihre Aufmerksamkeit!

Rückmeldung erwünscht:
vogel@ph-heidelberg.de