Mit dem Rechner sicher durch den Zufall und die Daten in der Sekundarstufe II

Andreas Eichler, IMBF Markus Vogel, Luckenwalde, 24.09.2011


Eigener Herd ...
Umfragen ergaben, den Deutschen
schmeckt es zu Hause
immer noch am besten.
Ich koche oder
bereite Essen zu
Anteil der Befragten
in Prozent
mehrmals täglich 20,3
mehrmals wöchentlich 34,5
einmal wöchentlich 5,0
mehrmals monatlich 6,4
einmal monatlich 2,9
seltener 12,3
nie 16,5
Quelle: TdW Intermedia 2000/2001 von 100 Prozent Gesamtbevölkerung 2.1 Prozent o. A.

1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. Kompetenzen – aber wie (mit/ohne Rechner)

    (K1): Bewerten von Daten

    (K2): Bivariate Daten – Elementarisierung und Masse bewältigen

    (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente

    (K4): (Binomial-)Verteilungen – Modell/Realität

    (K5): Simulation – (K6) Große Zahlen

    (K8): Testen – Modelle validieren/simulieren

    (K7): Schätzen – Verfeinern von Modellen

4. Allgemeine Ideen

Gliederung

1. Einstiegsbeispiel
2. **Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?**
3. Kompetenzen – aber wie (mit/ohne Rechner)

   (K1): Bewerten von Daten

   (K2): Bivariate Daten – Elementarisierung und Masse bewältigen

   (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente

   (K4): (Binomial-)Verteilungen – Modell/Realität

   (K5): Simulation – (K6) Große Zahlen

   (K8): Testen – Modelle validieren/simulieren

   (K7): Schätzen – Verfeinern von Modellen
4. Allgemeine Ideen

2. Wieso – weshalb – warum

Beschluss für Bildungsstandards Sek. II im Oktober 2007

Konstituierung der Kommissionen für die Formulierung der Standards sowie der Aufgabenentwicklung 2010 (IQB)

Entwurf für eine Leitidee Daten und Zufall

- Formulierung
- Kritik, Ergänzungen, …
- Abgabe im April 2010
- …

Leitidee Daten und Zufall (in beiden Sekundarstufen)

Schülerinnen und Schüler sollen erfahren, dass

Fragen an alltägliche empirische Phänomene gestellt und mit den ***elementaren*** **Methoden der Sekundarstufen (und dem Rechner) beantwort werden können**

und: Ideen statt Algorithmen zählen.

ger

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 der Binomialverteilung, 3 Zahlen: relevant
Rechner: (eher) irrelevant

1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. **Kompetenzen – aber wie (mit/ohne Rechner)**
    - **(K1): Bewerten von Daten**
    - (K2): Bivariate Daten – Elementarisierung und Masse bewältigen
    - (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente
    - (K4): (Binomial-)Verteilungen – Modell/Realität
    - (K5): Simulation – (K6) Große Zahlen
    - (K8): Testen – Modelle validieren/simulieren
    - (K7): Schätzen – Verfeinern von Modellen
4. Allgemeine Ideen

Schülerinnen und Schüler

(K1) planen exemplarisch zu für sie bedeutsamen Fragestellungen statistische Erhebungen.

Sie kennen notwendige Bedingungen für die Anwendbarkeit von Methoden der beurteilenden Statistik (Randomisierung, Zufallsstichproben)

Sie können dieses Wissen in eigenen exemplarischen Datenerhebungen (ein Experiment, eine Beobachtung, eine Befragung) und für die Beurteilung fremder Datenerhebungen umsetzen.

Aufgaben:

1. Variieren von *Ort, Zeit* und *Sache* der Stichprobe
2. Prüfen der Definition (*Operationalisierung*)
3. Wie sollte man Personen auswählen? (*Repräsentativität*)
4. Wie viele Personen sollte man auswählen? (*Schätzen*)

Eigener **Herd ...**

Umfragen ergaben, den Deutschen schmeckt es zu Hause immer noch am besten:

Ich koche oder bereite Essen zu	Anteil der Befragten in Prozent
mehrmals täglich	20,3
mehrmals wöchentlich	34,5
einmal wöchentlich	5,0
mehrmals monatlich	6,4
einmal monatlich	2,9
seltener	12,3
nie	16,5

Quelle: TdW Intermedia 2000/2001 von 100 Prozent Gesamtbevölkerung 2,1 Prozent o. A.

Gliederung

1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. **Kompetenzen – aber wie (mit/ohne Rechner)**
   - (K1): Bewerten von Daten
   - **(K2): Bivariate Daten – Elementarisierung und Masse bewältigen**
   - (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente
   - (K4): (Binomial-)Verteilungen – Modell/Realität
   - (K5): Simulation – (K6) Große Zahlen
   - (K8): Testen – Modelle validieren/simulieren
   - (K7): Schätzen – Verfeinern von Modellen
4. Allgemeine Ideen

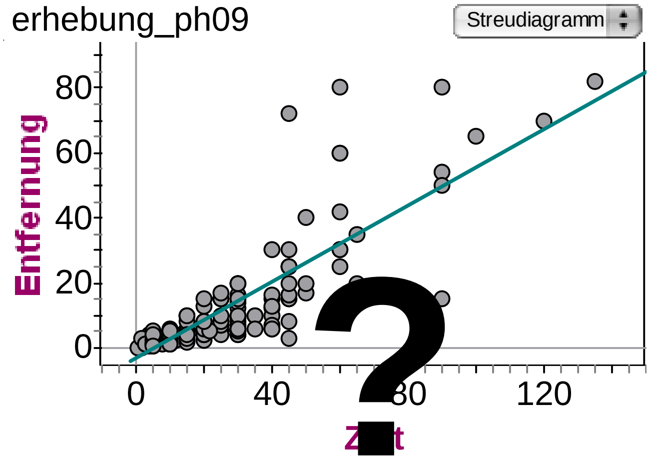
Schülerinnen und Schüler

(K2) modellieren statistische Trends und Zusammenhänge zweier Merkmale mit Hilfe von Funktionen.

Sie interpretieren die im Modell auftauchenden Parameter der Funktion und kennen geeignete Methoden wie die Residuenanalyse, um Abweichungen von Modell und Daten zu untersuchen.

Aufgabe:
Untersucht Zusammenhänge der Eigenschaften von Studierende/SchülerInnen

- lineare Abhängigkeit im Modell, im Sachkontext, in der Empirie
- Interpretation von Residuen
- Idee vor Algorithmus (Komplexität der Methoden)



$$f(x) = \frac{Cov(X,Y)}{s_X^{\,2}}x + \overline{y} - \frac{Cov(X,Y)}{s_X^{\,2}}\overline{x}$$

Aufgabe:
Untersucht Zusammenhänge der Eigenschaften von Studierende/SchülerInnen

Nach dem händischen Arbeiten

- Elementarisierung
- Visualisierung
- Idee vor Algorithmus
- Bewältigung von Masse

1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. **Kompetenzen – aber wie (mit/ohne Rechner)**
    - (K1): Bewerten von Daten
    - (K2): Bivariate Daten – Elementarisierung und Masse bewältigen
    - **(K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente**
    - (K4): (Binomial-)Verteilungen – Modell/Realität
    - (K5): Simulation – (K6) Große Zahlen
    - (K8): Testen – Modelle validieren/simulieren
    - (K7): Schätzen – Verfeinern von Modellen
4. Allgemeine Ideen

Schülerinnen und Schüler

(K3) modellieren mehrstufige zufällige Vorgänge.

Schülerinnen und Schüler sollen

- Vorgängen, deren Teilvorgänge stochastisch abhängig bzw. stochastisch unabhängig sind, unterscheiden können,
- bedingte Wahrscheinlichkeiten visualisieren können,
- den Satz von Bayes mit Hilfe einer Visualisierung wiedergeben, erklären und in relevanten Situationen anwenden können.

Uelzen:Gesundheitsamt rät von HIV-Schnelltests ab

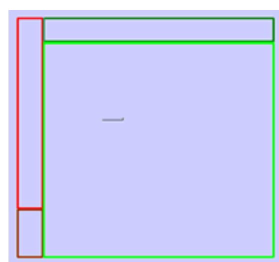
Pressemeldung vom 24. Juni, 2009, 10:40 am

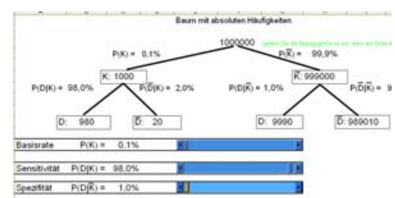
Aufgabe: HIV

- Warum ist der Schnelltest schlecht?

Basisrate	Krank: $P(K) = 0{,}001$, $P(\overline{K}) = 0{,}999$
Sensitivität	Wenn krank, dann Test+ (D): $P(D\|K) = 0{,}995$; $P(\overline{D}\|K) = 0{,}005$
Spezifität	Wenn gesund, dann Test+ (D): $P(D\|\overline{K}) = 0{,}01$; $P(D\|\overline{K}) = 0{,}99$

- Könnte man nicht bei einem positiven Ergebnis gleich noch einen machen?





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

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

Es fällt die Augenzahl x

Aufgabe:

Die Spielleitung wählt verdeckt den normalen oder den quaderförmigen Würfel aus. Welcher ist es?

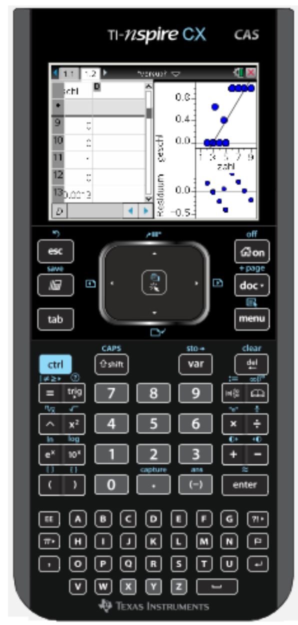
$$P(N|x) = \frac{P(x|N) \cdot P(N)}{P(x|N) \cdot P(N) + P(x|Q) \cdot P(Q)}$$

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

Modell: Würfe des Würfels sind unabhängig!

Nach dem händischen Arbeiten

- Begriffe bilden
- Lösung *und* Struktur
- Simultan: Experiment – Phänomen – Auswertung
- Visualisierung



1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. **Kompetenzen – aber wie (mit/ohne Rechner)**
   - (K1): Bewerten von Daten
   - (K2): Bivariate Daten – Elementarisierung und Masse bewältigen
   - (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente
   - **(K4): (Binomial-)Verteilungen – Modell/Realität**
   - (K5): Simulation – (K6) Große Zahlen
   - (K8): Testen – Modelle validieren/simulieren
   - (K7): Schätzen – Verfeinern von Modellen
4. Allgemeine Ideen

Schülerinnen und Schüler

(K4) modellieren zufällige Vorgänge mit Hilfe von Wahrscheinlichkeitsverteilungen und ihrer charakteristischen Kennzahlen.

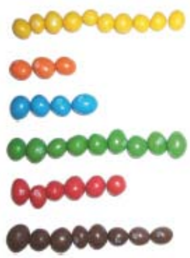
Dabei beachten sie die Voraussetzungen, die bei der Modellierung einer stochastischen Situationen mit einer spezifischen Wahrscheinlichkeitsverteilung gegeben sein müssen.

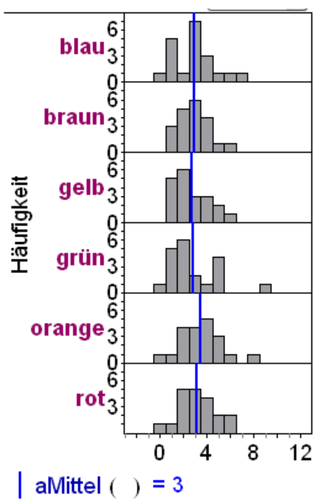
Sie beachten zudem den Zusammenhang zwischen den Parametern der Binomialverteilung und der Form der Verteilung sowie den charakteristischen Kenngrößen.

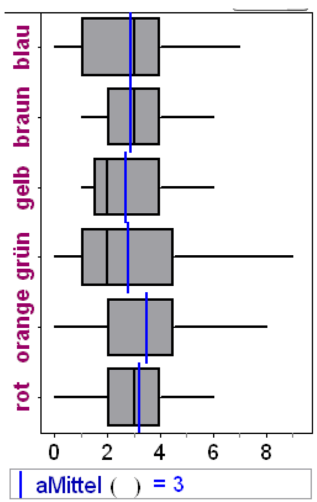
M&M

Aufgabe:

Untersucht die Inhalte dieser Tüten.







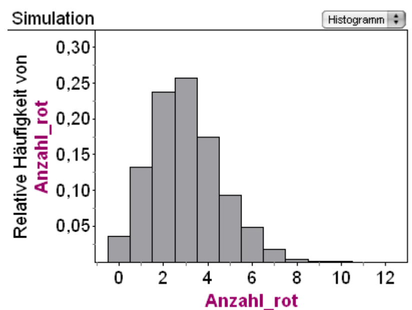
Modell:

In einer Packung sind durchschnittlich 18 Schokolinsen und im Durchschnitt je 3 Linsen einer Farbe

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



Informelle Inferenz

Bei welcher Anzahl roter Kugeln sollte man an dem Modell zweifeln

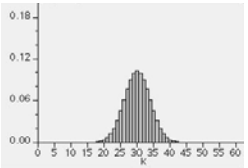
Empirische Welt der Daten

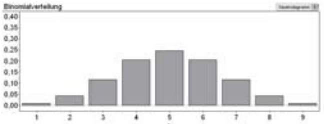
Modell-Welt der Zufalls

K4 – (Binomial-)Verteilungen

Aufgaben:

- Durcharbeiten des Modells, grafisch wie algorithmisch





- Modellierung in einer realistischen Situation

- Was wäre wenn: Prognose auf der Basis eines Modells + Prüfung (mit neuen Daten; Arbeitsphase)

Nach dem händischen Arbeiten

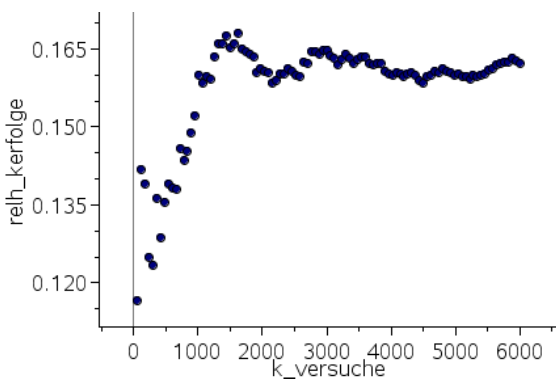
- Visualisierung
- Rechenknecht

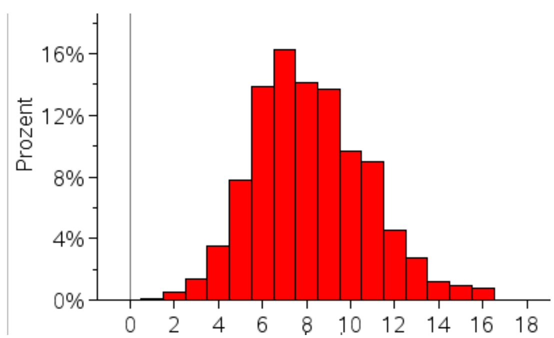
1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. **Kompetenzen – aber wie (mit/ohne Rechner)**
    - (K1): Bewerten von Daten
    - (K2): Bivariate Daten – Elementarisierung und Masse bewältigen
    - (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente
    - **(K4): (Binomial-)Verteilungen – Modell/Realität**
    - **(K5): Simulation – (K6) Große Zahlen**
    - (K8): Testen – Modelle validieren/simulieren
    - (K7): Schätzen – Verfeinern von Modellen
4. Allgemeine Ideen

Schülerinnen und Schüler

(K5) nutzen Simulationen, um mit stochastischen Situationen zu experimentieren und Näherungslösungen in komplexeren Situationen zu gewinnen.

Sie erstellen zu einer gegebenen einfachen Problemsituation ein Simulationsmodell und führen Simulationen mit dem Rechner aus.





Schülerinnen und Schüler

(K5) nutzen Simulationen, um mit stochastischen Situationen zu experimentieren und Näherungslösungen in komplexeren Situationen zu gewinnen.

Sie erstellen zu einer gegebenen einfachen Problemsituation ein Simulationsmodell und führen Simulationen mit dem Rechner aus.

Nach dem händischen Simulieren

- Simulation
- Begriffsbildung

1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. **Kompetenzen – aber wie (mit/ohne Rechner)**

   (K1): Bewerten von Daten

   (K2): Bivariate Daten – Elementarisierung und Masse bewältigen

   (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente

   (K4): (Binomial-)Verteilungen – Modell/Realität

   (K5): Simulation – (K6) Große Zahlen

   (K8): Testen – Modelle validieren/simulieren

   (K7): Schätzen – Verfeinern von Modellen
4. Allgemeine Ideen

Schülerinnen und Schüler

(K8) verwenden Methoden der beurteilenden Statistik (insbesondere Konfidenzintervalle und in Erweiterung auch Hypothesentests)

Für Schülerinnen und Schüler soll das Verständnis für **Grundideen** des Testens Vorrang haben gegenüber einer rein algorithmischen und formalen Abarbeitung von Problemstellungen.
Als wesentliches Hilfsmittel für das Vermitteln dieser Grundidee dient die **Simulation**.

Aufgaben:

- Idee des Testens durch Simulation

- Idee, solche Daten zu nehmen, die man erneut sammeln kann

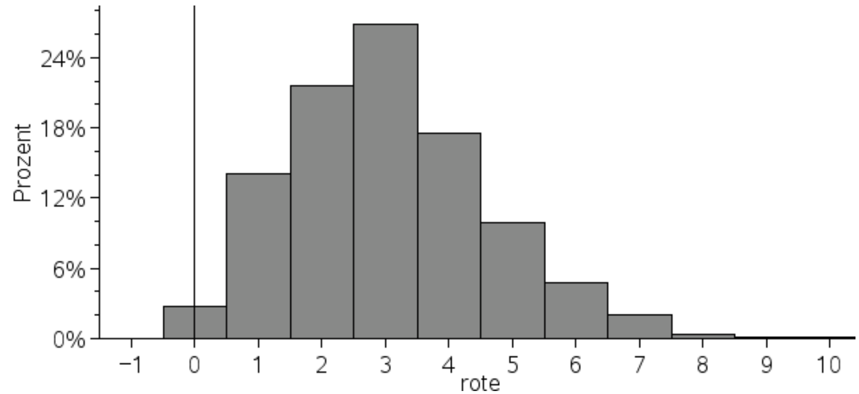
Aufgabe:

Überprüfen Sie Ihr Modell zur Verteilung der M&M-Kugeln

Modell:

In einer Packung sind durchschnittlich 18 Schokolinsen und im Durchschnitt je 3 Linsen einer Farbe.

Mit p = 1/6 und n = 18 lässt sich die Anzahl der (z.B.) roten Kugeln in einer Tüte simulieren…



… und damit ein Annahme bzw. Ablehnungebereich für das Modell konstruieren:

$$h(1 \leq R \leq 6) \approx 0{,}95$$

Nun öffne man die nächste Tüte.

Exakt (Binomialverteilung):

$$P(1 \leq R \leq 6) \approx 0{,}942$$

Approximation (Normalverteilung)

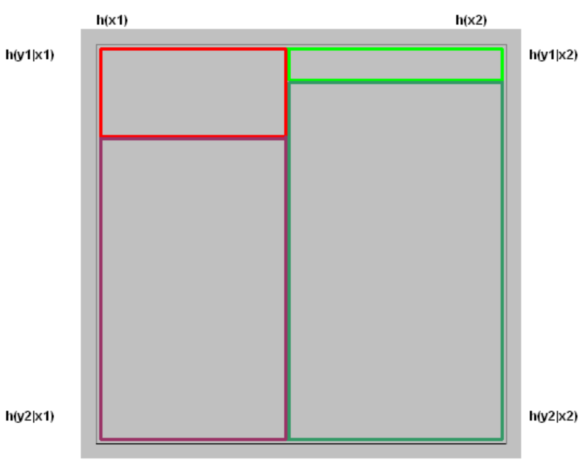
???

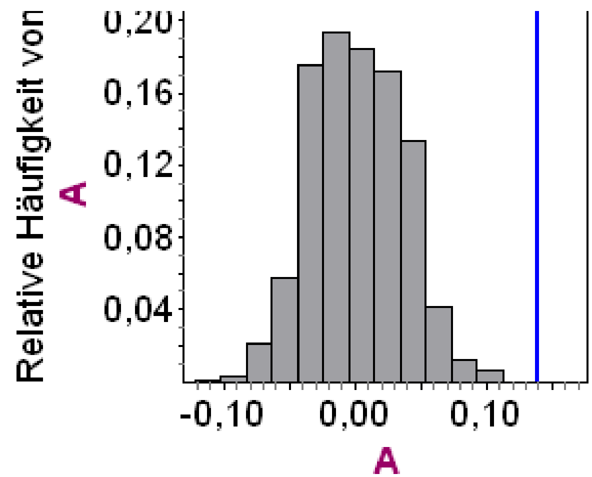
Aufgaben:

- Ausweiten der Idee auf Nicht-Schul-Standard-Tests
  - Unabhängigkeitstest, Mittelwerttest, …

Wohnverhalten an den Hochschulen in Münster und Freiburg

Vierfeldertafel (absolut)			
	PH	Uni	Summe
E	49	21	70
NE	169	239	408
Summe	218	260	478



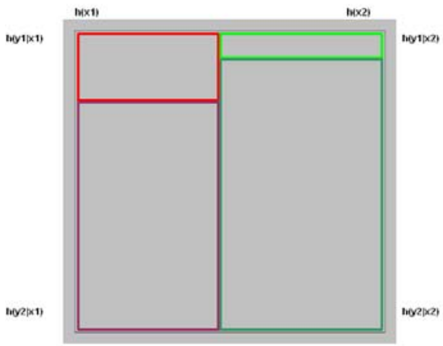


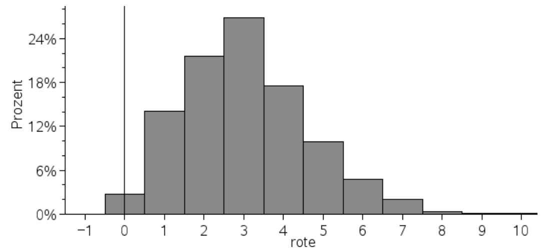
$$P(-0{,}14 \le A \le 0{,}14) < 0{,}05$$

$$A = h_{478}(E \mid PH) - h_{478}(E \mid Uni) = 0{,}14$$

K8 – Modelle testen

Vierfeldertafel (absolut)			
	PH	Uni	Summe
E	49	21	70
NE	169	239	408
Summe	218	260	478





Vor Algorithmus Klären der Idee des Testens

- Simulation
- Elementarisierung

1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. **Kompetenzen – aber wie (mit/ohne Rechner)**

    (K1): Bewerten von Daten

    (K2): Bivariate Daten – Elementarisierung und Masse bewältigen

    (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente

    (K4): (Binomial-)Verteilungen – Modell/Realität

    (K5): Simulation – (K6) Große Zahlen

    (K8): Testen – Modelle validieren/simulieren

    (K7): Schätzen – Verfeinern von Modellen

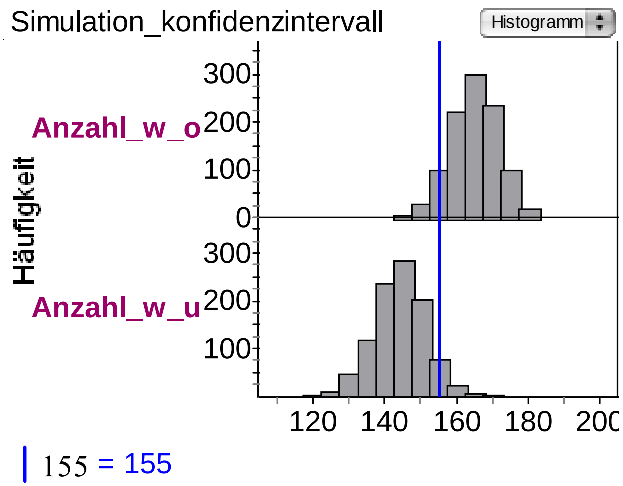
4. Allgemeine Ideen

Schülerinnen und Schüler

(K7) verwenden Methoden der beurteilenden Statistik (insbesondere Konfidenzintervalle und in Erweiterung auch Hypothesentests)

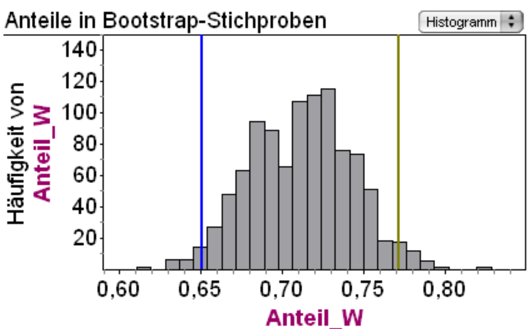
Für Schülerinnen und Schüler soll das Verständnis für **Grundideen** des Schätzens Vorrang haben gegenüber einer rein algorithmischen und formalen Abarbeitung von Problemstellungen.
Als wesentliches Hilfsmittel für das Vermitteln dieser Grundidee dient die **Simulation**.

- h (W) ≈ P(W) = 0,71 (PH Freiburg; **155** bei **218** erhobenen Studierenden)
- Sind das genau 71 %?
- Wäre P(W) = 0,76 (offizielle Angabe) auch möglich?
- Hieb- und Stichaufgabe
- Simulation oder Berechnung mit p = 0,76 und n = 218
- Welches P(W) mit P(W) < 0,71 könnte zum Stichprobenergebnis passen?



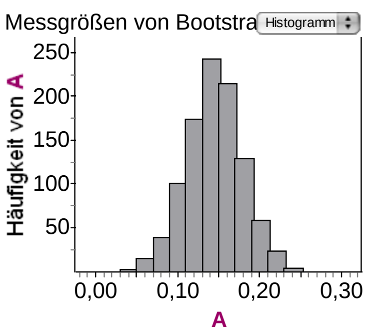
Begriffsbildung Konfidenzintervall

- Die Augenzahl 6 hat die Wahrscheinlichkeit von 1/6
- Man wirft den Würfel 1000 Mal, zählt die Anzahl der 6en und bildet das Konfidenzintervall (Näherung, Sigma-Intervalle, exakt …) zum Niveau 95%
- „Das Konfidenzintervall überdeckt den wahren Parameter von 1/6 mit einer Wahrscheinlichkeit von 95%“ **(?)**
- „95% der so gebildeten Intervalle überdecken den wahren Parameter von 1/6“ **(?)**



PH-Uni

Wohnung



- h (W) ≈ P(W) = 0,71 (PH Freiburg; **155** bei **218** erhobenen Studierenden)
- Simulation I_K = [0,650;0,772]
- Berechnung I_K = [0,651;0,772]
- Approximation (Normalverteilung) I_K = [0,651;0,771]
- Bootstrap: Neues Ziehen einer n=218-Stichprobe ohne Zurücklegen I_K = [0,651;0,771]
- Erweiterung auf weitere Konfidenzintervalle (z.B. Abhängigkeit/Assoziation)

1. Einstiegsbeispiel
2. Vorgeschichte: Bildungsstandards, wieso-weshalb-warum?
3. Kompetenzen – aber wie (mit/ohne Rechner)
   - (K1): Bewerten von Daten
   - (K2): Bivariate Daten – Elementarisierung und Masse bewältigen
   - (K3): (Un-)Abhängigkeit – Begriffsbildung/Experimente
   - (K4): (Binomial-)Verteilungen – Modell/Realität
   - (K5): Simulation – (K6) Große Zahlen
   - (K8): Testen – Modelle validieren/simulieren
   - (K7): Schätzen – Verfeinern von Modellen
4. **Allgemeine Ideen**

Die **Leitidee Daten und Zufall in der Sek. II** kann zeigen, dass

- reale Fragestellungen beantwortet nicht Algorithmen abgearbeitet werden
- mit Daten und Zufall Modelle der Realität gebildet werden können,
- auf Ideen fokussiert wird,
  - zu denen das Durcharbeiten eines Modells Mittel zum Zweck ist,
- Realität und Modell differenziert werden,
- Offenheit für Erweiterungen besteht,
  - aber auch Kürzungsmöglichkeiten in der Behandlungstiefe bestehen.

Wer sucht der findet

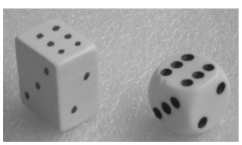




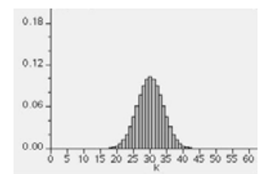
**Aufgaben zur realen Realität
(Nachempfinden/Lernen)**



**Aufgaben zu konstruierten realen Situationen
(Lernen und Testen, offene Aufgaben)**



**Konstruierte Situationen
(Üben/Testen von Aspekten des Modellierens)**



Rechner und Leitidee Daten und Zufall

Rechner ersetzt nicht das Verständnis, aber hilft bei

- Bewältigen großer Datenmengen
- visuell gesteuerter Datenanalyse
- Simulation
- Unterscheidung Modell/Realität
- Elementarisierung konventioneller Methoden
- Erweiterung von Fragestellungen

Daten und Zufall

Danke für die Aufmerksamkeit