



Wolfgang Riemer

Statistik unterrichten

Eine handlungsorientierte Didaktik
der Stochastik

3 Beurteilende Statistik im Federmäppchen

Darum geht es – das Wichtigste in Stichworten

- Anders als bei teilsymmetrischen Quadern verbindet man mit symmetrischen Zufallsobjekten wie Münzen, Würfeln oder Glücksrädern die Vorstellung eindeutiger objektiv existierender (Laplace-)Wahrscheinlichkeiten.
- Dass auch diese nur Modelle sind, wird beim Rollen rotationssymmetrischer Sechskantstifte – absolut überraschend – mehr als deutlich.
- Abweichungsmaße zwischen Erwartung (Modell) und beobachteter Häufigkeit (Realität) werden erfunden, erprobt und interpretiert.
- Durch simulationsgestütztes Aushandeln von Grenzen werden sie Testgrößen, deren Bedeutung sich durch Vergleich mit Primärintuitionen erschließt: Sie modellieren „Bauchgefühl“.

Experimente

- Rollen von Bleistiften
- Simulationen ermöglichen die Interpretation von Testgrößen.

3.1 Perspektivwechsel

Beim **Würfeln mit Quadern** steht die Begriffsentwicklung (Wahrscheinlichkeit auf der Modellebene versus relative Häufigkeit auf der Realitätsebene) im Zentrum. Man sucht nach Modellen, die gut zur Realität passen, also brauchbare Vorhersagen machen. Letztendlich bereitet man damit auf ganz elementarem Niveau die Idee der **Konfidenzbereiche** vor, in denen man Hypothesen sammelt, die gut zu einem experimentellen Ergebnis passen, denen man „vertraut“. Und man erlebt beim „Quadern“ intuitiv, dass der Bereich vertrauenswürdiger Hypothesen umso kleiner wird, je mehr Versuche man macht – je genauer man hinschaut (vgl. Abb. 2-3).

Beim **Würfeln mit Sechskantstiften** nimmt man die hierzu duale Perspektive ein: Wegen der Symmetrie industriell gefertigter Stifte steht nämlich einzig die Laplace-Hypothese im Raum. Und es gibt kaum ein reizvolles („erschütterndes“) Experiment als diese sicher geglaubte Hypothese im Klassenverband tatsächlich experimentell zu überprüfen. Wenn man dabei gefühlte Zweifel rechnerisch sortiert, entdeckt man, wie Testgrößen funktionieren, vor allem aber, wie sie – insbesondere auch unter Berücksichtigung des Versuchsumfangs – zu deuten sind.

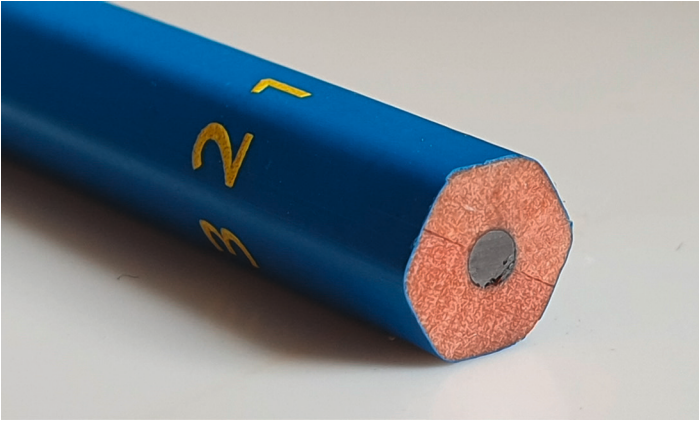


Abb. 3-1: Mit dem bloßen Auge erkennt man minimale Unsymmetrien nicht. Beim Rollen werden sie nicht nur statistisch signifikant, sondern auch relevant!

3.2 Laplace-Hypothese intuitiv bewerten

Wenn man Sechskantstifte mit den Augenzahlen 1, 2, ..., 6 so beschriftet, dass einander gegenüberliegende Seiten die Summe 7 besitzen, dann kann man mit ihnen „würfeln“, indem man sie durch leichtes Schnippen über den Tisch rollt. Aufgrund der industriellen Fertigung scheinen die Stifte absolut rotationssymmetrisch und es wäre abwegig, Zweifel an der Güte des Laplace-Modells zu hegen. Genau diese Erwartungshaltung ist eine perfekte Grundlage für den folgenden Arbeitsauftrag.

A

Arbeitsauftrag 1

- Rollt den Stift 120-mal in eine Richtung, anschließend 120-mal in die Gegenrichtung. Haltet die Häufigkeiten der sechs Seiten jeweils in einer Strichliste fest.
- Bewertet auf einer Viererskala intuitiv („nach Bauchgefühl“), wie gut das Laplace-Modell zu eurem Stift passt.

Nach wenigen Minuten geht ein Raunen durch die Lerngruppe. Yusof: „Verdammt! Das gibt's doch nicht! Ich bekomme nur Zweier und kaum Sechser!“ Antonia: „Deine Zweier kann ich brauchen, die tausche ich gegen meine Fünfer, die Sechser passen bei mir ...“

Die Stifte entfalten eine erstaunliche Individualität (siehe Abb. 3-2) und die Laplace-Hypothesen brechen vielfach wie Kartenhäuser in sich zusammen, manchmal schneller, manchmal langsamer. Das ursprünglich mit einem ho-

hen Vertrauensvorschuss favorisierte Laplace-Modell ist nach 120 Versuchen bei vielen Stiften „für die Tonne“ (genau wie das Proportionalitätsmodell bei den Quaderwürfeln). Und das ist der Hypothesentest der beurteilenden Statistik auf präformalem Niveau.

P2 =Summe((C2:H2 - 20)/20)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1		cm	1	2	3	4	5	6		gut	noch ok	schlecht	sehr schlecht		s: $\sum(h_i - 20)$	t chr2	u: $\max(h_i) - \min(h_i)$	g: $\max(h_i - 20)$	a Ausreisser	#s	#t	#u	#g	#a
2	Yusof re	16.5	37	42	15	9	12	5	120						78	60.4	37	22	4	21	22	25	21	22
3	Yusof li	16.5	23	20	18	21	15	23	120	1			1		14	2.4	8	5	0	2	2	7	2	2
4	Felix re	16.7	18	21	26	17	11	27	120		1				28	9	16	9	0	7	7	7	7	7
5	Felix li	16.7	22	13	22	23	15	25	120	1					24	5.8	12	7	0	5	6	7	5	6
6	Julia re	14.2	21	22	10	9	31	27	120			1			42	19.8	22	11	3	9	8	16	9	8
7	Julia li	14.2	18	23	24	18	17	20	120	1					14	2.1	7	4	0	1	1	7	1	1
8	Antonia re	12.7	4	18	45	1	29	23	120				1		74	66.6	44	25	3	26	24	16	26	24
9	Anna li	12.7	33	9	22	7	11	38	120						66	43.4	31	18	4	16	17	25	16	17
10	Lea re	11.8	0	33	24	9	19	35	120			1			64	46.6	35	20	4	19	20	25	19	20
11	Lea li	11.8	28	59	38	0	0	4	120				1		112	117.2	59	30	5	27	27	28	27	27
12	Catherin re	18.8	15	17	8	20	17	43	120			1			46	35.8	35	23	2	19	23	12	19	23
13	Catherin li	18.8	14	7	16	7	30	46	120						72	58.3	39	26	4	23	25	23	25	25
14	Paul re	16.5	32	26	14	36	5	7	120			1			68	43.3	31	16	4	16	14	25	16	14
15	Paul li	16.5	30	21	17	26	19	7	120			1			34	15.8	23	13	2	10	11	12	10	11
16	Anne li	16.5	24	15	13	19	27	22	120		1				26	7.2	14	7	0	6	6	7	6	6
17	Anne re	16.5	33	30	20	13	13	11	120			1			46	22.4	22	13	2	9	11	12	9	11
18	Dominik re	17.1	1	26	39	7	22	25	120						64	47.8	38	19	3	22	18	16	22	18
19	Dominik li	17.1	16	22	21	49	6	6	120			1			64	62.7	43	29	3	25	26	18	25	26
20	Stefanie re	17	37	35	23	13	5	7	120			1			70	48.3	32	17	4	17	16	25	17	16
21	Stefanie li	17	5	1	2	88	15	9	120				1		136	284	87	68	5	28	28	28	28	28
22	Lukas re	15.6	16	27	20	20	17	19	120		1				15	3.75	11	7	0	4	6	7	4	6
23	Lukas li	15.6	22	14	23	18	25	18	120			1			20	4.1	11	6	0	4	3	7	4	3
24	Luise re	16.2	14	40	35	5	4	22	120			1			74	57.3	36	20	4	20	20	25	20	20
25	Luise li	16.2	30	19	14	42	14	1	120			1			64	50.9	41	22	3	24	22	18	24	22
26	Antonia re	13.5	7	5	16	30	31	31	120				1		64	37.6	26	15	5	12	12	28	12	12
27	Antonia li	13.5	37	9	13	27	10	24	120			1			56	31.2	28	17	3	14	16	18	14	16
28	Anna Z re	18.1	16	24	31	18	4	27	120						44	23.1	27	16	2	13	14	12	13	14
29	Anna Z li	18.1	18	23	20	17	9	33	120			1			32	15.6	24	13	2	11	11	12	11	11

Abb. 3-2: Bewertung der Laplace-Hypothese nach Bauchgefühl. Überraschend ist der Einfluss der Rollrichtung re/li (vgl. z. B. Lea).

Abb. 3-3: Spalten O-S: Werte erfundener Sortiergrößen. Spalten T-X: Rangposition gemäß zugehöriger Sortiergröße (zweifel-sortieren.ggb [📄](#))

3.3 Den Zweifel sortieren

Entscheidend für die intuitive Einschätzung der Güte des Laplace-Modells sind die Abweichungen der absoluten Häufigkeiten vom Erwartungswert 20. Bei den mit „gut“ oder „noch gut“ bewerteten Verteilungen liegen alle Häufigkeiten zwischen 11 und 29. Falls Häufigkeiten besonders klein (≤ 10) oder besonders groß (≥ 30) sind, gelten sie als Ausreißer und das Laplace-Modell passt nach dem Gefühl der Schülerinnen und Schüler schlecht oder sehr schlecht. Luise linksherum hat drei, Antonia rechtsherum fünf Ausreißer. Aber Luises Ausreißer (42 und 1) sind extremer. Bei wem passt das Laplace-Modell schlechter? Gute Frage!

25	Luise li	30	19	14	42	14	1
26	Antonia re	7	5	16	30	31	31

Sie ist Anlass für folgenden Arbeitsauftrag, der genauso präzise wie offen ist.

A

Arbeitsauftrag 2

Erfindet eine Sortiergröße – besser noch: viele Sortiergrößen – mit der Eigenschaft:
Je größer die Sortiergröße, desto größer der Zweifel an der Güte des Laplace-Modells!

Wie Abb. 3-4 zeigt, überraschen Schülerinnen und Schüler ab Klassenstufe 8/9 mit konstruktiven Ideen, wenn man ihnen etwas Zeit zum Nachdenken lässt. Falls quadratische Funktionen „schon dran waren“, wird neben der Summe s der betraglichen Abweichungen vom erwarteten Wert 20 oft auch die Sortiergröße q genannt, die durch das Quadrieren große Abweichungen stärker „bestraft“ als s . Die Division $t = \frac{q}{20}$ macht q numerisch handlicher. Dass es sich hierbei um Chiquadrat (χ^2) handelt, wird man beiläufig erwähnen.

	Luise linksherum						Antonia rechtsherum					
	30	19	14	42	14	1	7	5	16	30	31	31
s: Summe aller Abweichungen vom Sollwert	64						64					
q: Summe aller Abweichungsquadrate	1018						752					
t: $q/20$ (χ^2)	50.9						37.6					
u: max - min	41						26					
g: größte Abweichung vom Sollwert 20	22						15					
a: Anzahl der "Ausreißer"	3						5					

Abb. 3-4: Vorschläge für geeignete Sortiergrößen

Vielleicht wird man auch ergänzen, dass das (analog zu g gebildete) Abweichungsmaß

$$ks = \max(|H(X \leq 1) - 20| + |H(X \leq 2) - 40| + \dots + |H(X \leq 6) - 120|)$$

von den berühmten Statistikern Kolmogoroff/Smirnof erfunden wurde.

Die erfundenen Größen sortieren den Zweifel ähnlich. Das zeigen die letzten fünf Spalten T ... X von Abb. 3-3, die für jede Häufigkeitsverteilung die Position angeben, auf der sie landet, wenn man nach der jeweiligen Größe sortiert. Damit sind alle Sortiergrößen sinnvoll. Die Schülerinnen und Schüler nehmen im Sinne von Abb. 1.1 mit: „Mathe hat etwas mit uns zu tun“. Das gilt auch für den folgenden Arbeitsauftrag 3.

3.4 Die Sortiergrößen werden Testgrößen

A

Arbeitsauftrag 3:

Betrachtet eine Sortiergröße, etwa g , s oder t , und einigt euch in eurer Gruppe auf eine Obergrenze, also einen „Benchmark“, ab dem ihr „starke Zweifel“ an der Fairness des Bleistift-Würfels für gerechtfertigt haltet. Ihr könnt dazu die in Abb. 3-2 vorgeschlagenen „Bauchgefühl-Einschätzungen“ zurate ziehen und ggf. in eurer Gruppe modifizieren.

Naturgemäß differieren die innerhalb der Gruppen ausgehandelten Benchmarks von Gruppe zu Gruppe ein wenig. Diese Unterschiede sind Anlass für den folgenden Arbeitsauftrag.

A

Arbeitsauftrag 4:

Würfelt mit echten, dann mit computersimulierten Spielwürfeln 120-mal und sucht für jede der Sortiergrößen g , s und t nach einer Obergrenze, die nur selten (in 5 % aller Fälle) überschritten wird.

Die Verlangsamung durch das händische Würfeln und das Berechnen der verschiedenen Sortiergrößen und die dabei ablaufenden Gespräche in einer Atmosphäre emotionaler Eingebundenheit sind sinnvoll investierte Unterrichtszeit. Computersimulationen, bei denen die Schwankungen der Sortiergrößen durch Boxplots (mit Ausreißern) veranschaulicht werden, ergänzen (notfalls ersetzen) diese Experimente.

Die Experimente (Abb. 3-5) zeigen: Bei 120 Würfen eines Laplace-Würfels gilt fast immer (in ca. 95 % aller Fälle) $g \leq 11$, $s \leq 30$ und $t \leq 11$. Dadurch, dass in der Computersimulation neben den Sortiergrößen auch die zugehörige Häufigkeitsverteilung zu sehen ist, entwickelt sich ein Gefühl dafür, wie „krass“ die Verteilungen aussehen, die mit $s > 30$, $g > 11$, $t > 11$ die kritischen Grenzen überschreiten. Durch folgende – auch in der „scientific community“ nur ausgehandelte – Vereinbarung werden die Sortiergrößen zu Testgrößen:

i

Offiziell bezweifeln Statistiker die Fairness des Würfels erst dann, wenn die Testgröße die kritische Grenze überschreitet. Man spricht dann von einer statistisch signifikanten Abweichung (der Wirklichkeit vom Modell).

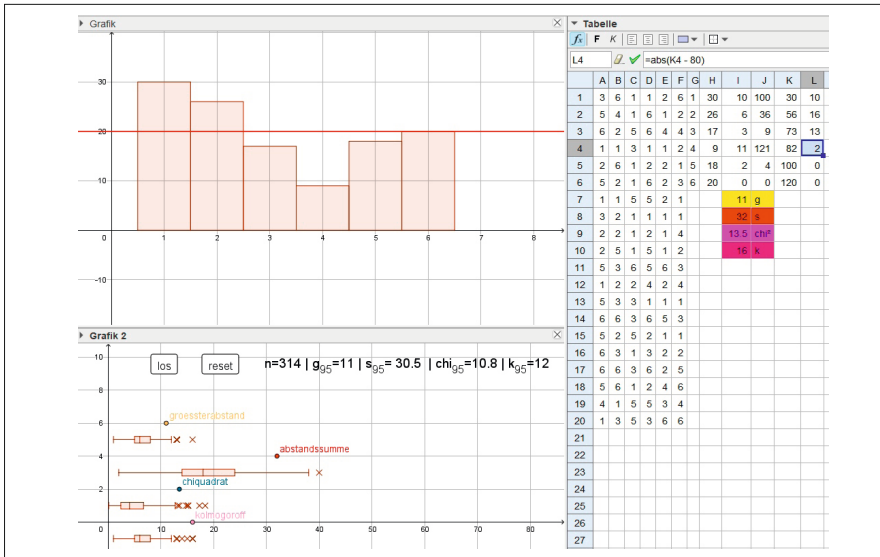


Abb. 3-5: Die abgebildete Häufigkeitsverteilung signalisiert starke Zweifel an der Güte des Laplace-Modells. Die Window-Shuttle-Technik, durch welche (rechts) die Urliste, (links oben) die zugehörige Verteilung, (rechts unten farbig markiert) die aktuellen Sortiergrößen, über die Boxplots (links unten) auch im Vergleich zu allen bisherigen Sortiergrößen gleichzeitig sichtbar sind und miteinander vernetzt werden, ist lernpsychologisch genial (03-bleistift-simulation.ggb [↓](#)).

Insgeheim zweifelt man natürlich auch schon dann, wenn die Testwerte erhöht sind. Im Schülerslang: „je höher der Testwert, desto größer der Zweifel“, und wenn man sensibel mit der Sprache umgeht und den beim Bleistiftrollen „gefühlten Zweifel“ als Begriff pflegt, kommt niemand auf die (absurde) Idee, nach der Wahrscheinlichkeit zu fragen, mit der ein verwendeter Würfel „lplacesch“ ist (vgl. auch 3.7). Es geht einzig und allein um die Einschätzung, wie gut „für den Hausgebrauch“ das Laplace-Modell für das „Würfelobjekt“ taugt.

Und das wäre die Antwort, die das Würfeln mit Bleistiften auf die von Amrhein u. a. aufgezeigten Interpretationsprobleme statistischer Aussagen liefert, nicht erst kurz vor dem Abitur, sondern im Sinne einer langfristigen Prävention – schon in Klasse 8. Gleichzeitig ist dies ein grandioser Brückenschlag zwischen beschreibender und beurteilender Statistik, bei dem sich Boxplots zur Beurteilung des Verhaltens von Sortiergrößen als sehr hilfreich erweisen.

Im Normalfall kann man sich in der Sekundarstufe I mit dem festen Stichprobenumfang $n = 120$ begnügen. Wie reizvoll eine Erhöhung sein kann, zeigt Abschnitt 3.5.

Tipp: Wer ohne experimentellen Mehraufwand im Sinne „realistischen Mathematikunterrichts“ erforschen möchte, ob bzw. wie sich die Rollrich-

tungen gegenseitig kompensieren, zerlegt die vorliegenden Ergebnisse der 120 Rechts- und Links-Rollversuche in je zwei 60er-Serien, um sie dann zu 120er-Serien mit gleichviel Rechts- und Linksdrehungen zu kombinieren. Die Testgrößen werden dadurch i. d. R. tendenziell etwas kleiner, es wäre aber weit verfehlt, von einem Ausgleich zu sprechen. Die Rollrichtungen kompensieren sich bei realen Stiften nicht.

3.5 Stichprobenumfang und Hauptsatz der beurteilenden Statistik

Wir blicken nochmals zurück auf die Bewertungen der Bleistift-Daten: Wegen der „etwas erhöhten“ Testwerte waren Felix' leichte Zweifel an der Güte des Laplace-Modells bei seinen rechtsherum Ergebnissen sicher gerechtfertigt.

Was wäre, wenn Felix seinen Stift sehr viel öfter (statt 120 z. B. 480-mal) rollen würde? Würden sich die leichten Zweifel an der Brauchbarkeit des Laplace-Modells zu massiveren Zweifeln verfestigen? Wieder eine gute Frage!

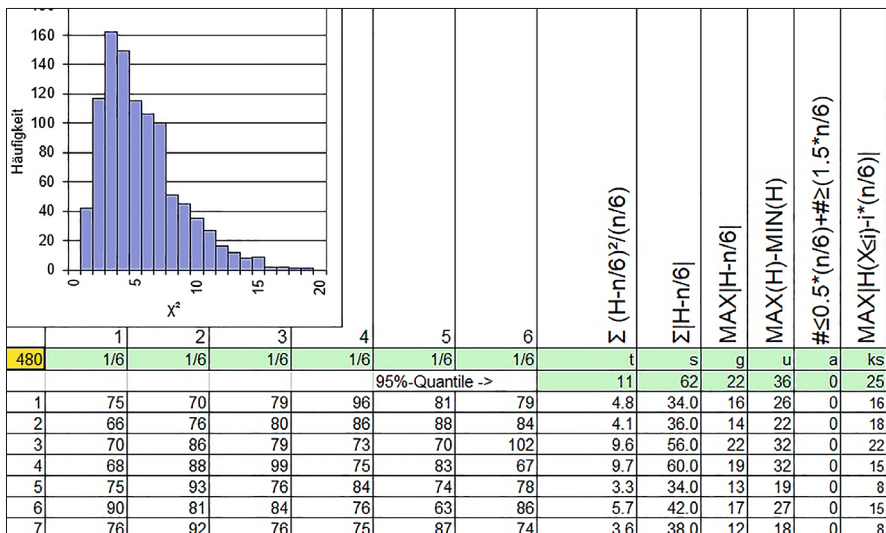


Abb. 3-6: 03-bleistift-simulation.xlm liefert in Zeile 4 die 95 % Quantile der Sortiergrößen für beliebige Stichprobenumfänge (markierte Zelle).

Aber bevor Felix weiterrollt, werden die Sortier- bzw. Testgrößen an den neuen Versuchsumfang 480 angepasst: Bei s, t, und u wird der Sollwert 20 durch 80 ersetzt und bei a erhöhen Schülerinnen und Schüler die Ausreißer-

Grenzen 10 bzw. 30 ohne Zögern proportional zum vierfachen Versuchsumfang auf 40 bzw. 120.

Erneute Simulationen zeigen: Der kritische 95 %-Wert zu t bleibt unverändert bei 11, bei s erhöht er sich von 30 auf 62, bei g von 11 auf 22 und bei u von 18 auf 36. Aber dass es nun keinen einzigen Ausreißer (das sind jetzt Werte unter 80 oder über 120) gibt, erweckt großes Staunen – und öffnet die Perspektive auf Kap. 4 ($\frac{1}{\sqrt{n}}$ -Gesetz).

Natürlich werden nun die Ergebnisse realer Rollversuche mit vierfachem Versuchsumfang mit Spannung erwartet.

n	1	2	3	4	5	6	t	s	g	u	a	ks
120	18	21	26	17	11	27	9.0	28.0	9	16	0	7
120	19	17	25	19	12	28	8.2	26.0	8	16	0	8
120	30	15	28	14	12	21	14.5	38.0	10	18	0	13
120	16	31	20	19	11	23	11.4	28.0	11	20	0	7
480	83	84	99	69	46	99	25.3	90.0	34	53	0	26

Und tatsächlich ergeben sich für Felix (re) auch in den drei folgenden 120er-Serien erhöhte Sortiergrößen, die den Zweifel schüren. Fasst man die Häufigkeiten aber zu einem 480er-Versuch zusammen, signalisieren alle Testgrößen ganz massive Zweifel. Und damit hat man den Einfluss des Stichprobenumfangs wirklich erlebt:

Ein hoher Stichprobenumfang wirkt wie ein Vergrößerungsglas, das auch kleine, vielleicht für den Alltagsgebrauch gar nicht so entscheidende Abweichungen zum Modell hervortreten lässt.

Wenn jetzt auch Anne und Lukas das Bedürfnis verspüren, weiter zu würfeln, ihre Stifte also durch das stochastische Vergrößerungsglas ... vielleicht sogar durch das statistische Mikroskop mit $n=800$... zu analysieren, ist der Funke „realistischen“ Mathematikunterrichts im Sinne Freudenthals übergesprungen ... Und man formuliert den „Hauptsatz der beurteilenden Statistik“.

„Hauptsatz der beurteilenden Statistik“

Wenn man den Stichprobenumfang groß genug macht (also genau genug hinsieht), dann kann man jede Hypothese bezweifeln.

3.6 Kontrapunkt

Was dabei herauskommen kann, wenn man das Bleistiftwürfeln nicht als authentisches Fallbeispiel zur Begriffsbildung nutzt, sondern in den Dienst standardisierter Signifikanztestaufgaben stellt, bestätigen die Bedenken, die zur Petition (Amrhein u. a. 2019) führten:

A

Zweiseitiger Signifikanztest

Aufgabe: Ein Bleistift wird auf dem Tisch gerollt. Man vermutet, dass die beschriftete Seite mit der Wahrscheinlichkeit $p = \frac{1}{6}$ oben liegen bleibt.

Jana testet dies, indem sie ihren Bleistift 300-mal rollt. Dabei zählte sie 35 Versuche mit der beschrifteten Seite nach oben. Spricht dieser Test gegen die obige Vermutung?

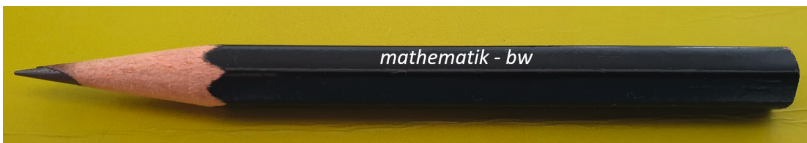


Abb. 3-7: https://www.schule-bw.de/faecher-und-schularten/mathematisch-naturwissenschaftliche-faecher/mathematik/unterrichtsmaterialien/sekundarstufe2/stochastik/testen/einf_signitest.jpg (16.04.2022)

Lösung: Eine Rechnung ergibt: 36 liegt außerhalb des 2σ Intervalls [38; 62]. Damit wird die Vermutung $p = \frac{1}{6}$ unwahrscheinlich. $p = \frac{1}{6}$ wird verworfen. Wir können aber nicht sicher sein, dass die Vermutung falsch ist. Denn in etwa 5 % aller Fälle können durchaus weniger als 38 oder mehr als 62 Treffer auftreten.

Wenn wir also behaupten, aufgrund unseres Tests wäre die Vermutung $p = \frac{1}{6}$ widerlegt, so irren wir uns mit einer Wahrscheinlichkeit von etwa 5 %. Genauer gilt: $P(\text{Irrtum}) = P(X < 38) + P(X > 62) = 0.052$.

Die gut gemeinte Warnung „Wir können nicht sicher sein, dass $p = \frac{1}{6}$ falsch ist“ führt in die Sackgasse. Tatsächlich kann man sogar absolut sicher sein, dass $p = \frac{1}{6}$ nicht gilt. Um das sagen zu können, hätte Jana ihren Bleistift nach dem Hauptsatz nicht einmal rollen müssen. Und die verkürzte Aussage: Wenn wir $p = \frac{1}{6}$ verwerfen, irren wir mit 5.2 % wird dahingehend gedeutet, dass wir uns mit der Wahrscheinlichkeit von 94.8 % nicht irren, also die Entscheidung richtig ist. Und diese Deutung ist fatal, denn Tests machen keine quantitativen Aussagen über subjektive Sicherheiten, mit denen Hypothesen gelten, sondern nur über Wahrscheinlichkeiten für (kritische) Ereignisse

bei unterstellter Gültigkeit von Hypothesen, die eigentlich keinen Anwender interessieren. Und wenn man Signifikanztests nur in Form von Routineaufgaben durchrechnen lässt, wird man Buth (2003, S. 30) zustimmen müssen:

„Also gehören Signifikanztests in den Papierkorb, und zwar auf allen drei Ebenen: in der Mathematik, in den anwendenden Disziplinen und im Mathematikunterricht.“

3.7 Resümee

Das Bleistiftrollen eignet sich hervorragend als authentisches Fallbeispiel zum Sammeln von Grunderfahrungen beurteilender Statistik – und zwar schon in der Sekundarstufe I.

- Es zeigt, dass auch höchst plausible Laplace-Wahrscheinlichkeiten nur Hypothesen sind. Damit ist das Bleistiftrollen nicht nur kognitiv, sondern auch emotional aktivierend.
- Der Auftrag, intuitiv gefühlte Zweifel mathematisch zu sortieren, verknüpft Intuition mit Präzision. Er führt zum selbstständigen (!) Entdecken und Erproben von Sortiergrößen, die durch das Vereinbaren kritischer Schranken allesamt sinnvolle Testgrößen werden. Kinder erleben sich als „Mathematik schaffende“ im Sinne von Abb. 1-1.
- Das Bleistiftrollen hat gegenüber dem Münzenwerfen den Vorteil, dass es eine Vielzahl sinnvoller Sortier- (und damit Test-)Größen zu entdecken gibt. Und es gibt keine Münzen, bei denen das Laplace-Modell ähnlich eindrucksvoll versagt wie bei den handelsüblichen Sechskantbleistiften.⁶
- Wenn eine Testgröße Zweifel an der Modellgüte signalisiert, weisen i. d. R. auch alle anderen erhöhte Werte auf, die aber nicht notwendig über den jeweiligen kritischen Grenzen liegen müssen. Man erlebt somit, dass Tests keine ja/nein Urteile über Hypothesen fällen, sondern ausschließlich behutsam zu interpretierende Anhaltspunkte zur Bewertung der Modellgüte liefern.
- Und man erlebt: Es spielt eine erhebliche Rolle, ob man das Modell mit Augenmaß ($n=120$), durch eine Lupe ($n=480$) oder durch ein Mikroskop ($n=1920$) – bei dem auch irrelevante kleine Abweichungen der Realität vom Modell signifikant werden – prüft.

⁶ Wenn man Münzen senkrecht um die eigene Achse kreiseln lässt statt sie zu werfen, werden aber auch hier kleinste Asymmetrien deutlich sichtbar. Es gibt US-Pennys, die mit > 80%iger Wahrscheinlichkeit auf der Seite Zahl landen (vgl. Riemer u. a. 2024).

So gelingt es: Stochastik mit Geogebra



WOLFGANG RIEMER | REIMUND VEHLING

Stochastik erkunden Ideenreiche Arbeitsblätter mit GeoGebra

Bestellnr.: 1840008, € 27,90

Alle sammeln Daten - jetzt auch im Matheunterricht. GeoGebra hat eine Datenanalyse - und viele Befehle, um Auswertungen von Datensätzen oder Simulationen „per Hand“ zu programmieren. Parallel zu Inhalten der Leitidee „Daten und Zufall“ erkunden Lernende diese Tools anhand konkreter Arbeitsblätter.

In einem zeitgemäßen Stochastikunterricht steht die fachliche Systematik am Ende eines Lern- und Erkenntnisprozesses, in dessen Mittelpunkt spannende Fragen und das Bewältigen kognitiver Konflikte stehen. Die Arbeitsblätter bieten Aufgaben, an deren Erforschung und Bearbeitung die Schülerinnen und Schüler zentrales Wissen und tragfähige Grundvorstellungen handelnd und reflektierend aufbauen können. Mit den exakt auf die Aufträge zugeschnittenen Dateien kann (vielfach auch ohne GeoGebra-Vorkenntnisse) inhaltlich zielgerichtet gearbeitet werden.

Wichtige Befehle und allgemein hilfreiche Dateien wie das Signifikanztest-Tool, das Binomialverteilungs-Tool, das Universalauswertungs-Tool oder die Stoppuhr bieten einen Nutzen weit über die 63 Arbeitsblätter hinaus.

Alle Preise zzgl. Versandkosten, Stand 2023.

Fachbuch



Kallmeyer



Unser Leserservice berät Sie gern:
Telefon: 05 11/400 04 -150
Fax: 05 11/400 04 -170
leserservice@friedrich-verlag.de



Die Downloadmaterialien enthalten
zahlreiche Arbeitsblätter und vertiefende
Materialien für Ihren Unterricht.

www.klett-kallmeyer.de

Wichtige GeoGebra-Befehle aus dem Bereich der Stochastik

Befehl	Erläuterung
nCr[n,k] n: Anzahl der Versuche; k: Anzahl der Treffer	Berechnung von $\binom{n}{k}$
Binomial [n, p] (p: Erfolgswahrscheinlichkeit)	erzeugt ein Balkendiagramm einer Binominalverteilung
Binomial [n, p, <Wahrheitswert Verteilungsfunktion>]	erzeugt ein Balkendiagramm einer Binomialverteilung, wenn der Wahrheitswert false ist; erzeugt ein Balkendiagramm einer kumulativen Binomialverteilung, wenn der Wahrheitswert true ist
Binomial [n, p, k, <Wahrheitswert Verteilungsfunktion>]	Sei X eine $B(n; p)$ -Zufallsvariable. $P(X = k)$, wenn der Wahrheitswert false ist $P(X \leq k)$, wenn der Wahrheitswert true ist
Normal [Mittelwert, Standardabweichung, x]	erzeugt die Wahrscheinlichkeitsdichtefunktion der Normalverteilung
Normal [Mittelwert, Standardabweichung, x, Wahrheitswert Verteilungsfunktion]	Ist der Wahrheitswert true, dann wird die kumulative Verteilungsfunktion erzeugt, ansonsten die Wahrscheinlichkeitsdichtefunktion der Normalverteilung.
InversNormal [<Mittelwert>, <Standardabweichung>, <Wahrscheinlichkeit>]	Es wird $\Phi^{-1}(p) \cdot \sigma + \mu$ berechnet. Φ ist die Verteilungsfunktion der Standardnormalverteilung. Mit diesem Befehl wird der Wert der $N(\mu; \sigma)$ -verteilten Zufallsgröße X ermittelt, welche zu dem gegebenen p-Wert gehört.
Zufallszahl [<Minimalwert>, <Maximalwert>]	erzeugt eine ganzzahlige Zufallszahl zwischen Minimalwert und Maximalwert (einschließlich) Beispiel: Mit Zufallszahl [1,6] kann der Wurf mit einem normalen Würfel simuliert werden.
ZufallszahlBinomialverteilt [n, p]	Realisation (Anzahl der Treffer) einer binomialverteilten Zufallsgröße X n: Wiederholungen, p: Trefferwahrscheinlichkeit
ZufallszahlBinomialverteilt [n, p] / n	Realisation (Stichprobenanteil h) einer binomialverteilten Zufallsgröße n: Wiederholungen, p: Trefferwahrscheinlichkeit
ZufallszahlGleichverteilt [<Min>, <Max>]	erzeugt die Realisation einer gleichverteilten Zufallsgröße X in einem Bereich [Min, Max]
ZufallszahlNormalverteilt [<Mittelwert μ >, <Standardabweichung σ >]	erzeugt die Realisation einer normalverteilten Zufallsgröße $N(\mu, \sigma)$
Stichprobe [<Liste L>, <Größe n>, <Wahrheitswert Wiederholung möglich>]	Erzeugt eine Liste mit n beliebig gewählten Elementen aus der Liste L. Wahrheitswert true: Die Elemente können mehrfach vorkommen (Ziehung mit Zurücklegen) Wahrheitswert false: Die Elemente können nur einfach vorkommen (Ziehung ohne Zurücklegen). Beispiel: Stichprobe [{1, 2, 3}, 3, false] erzeugt eine zufällige Permutation von {1, 2, 3}
Mittelwert [<Liste von Zahlen>]	Berechnung des Mittelwertes
Standardabweichung [<Liste von Zahlen>]	Berechnung der Standardabweichung
StichprobenStandardabweichung [<Liste von Zahlen>]	Berechnung der Stichprobenstandardabweichung
Stichprobenvarianz [<Liste von Zahlen>]	Stichprobenvarianz



Wolfgang Riemer ist Archimedes-Preisträger, Schulbuchautor, Fachleiter, Lehrbeauftragter, Erfinder der „Riemerwürfel“, der „Würfelbleistifte“ und des „Glücksrads auf der schiefen Ebene“.

Seit vielen Jahren baut er Brücken zwischen Schulalltag, Hochschuldidaktik, Lehrerbildung und Lehrplangestaltung.

Mit seinen Ideen hat er den Stochastikunterricht substantiell bereichert und nachhaltig zum Besseren verändert.

Statistik unterrichten ist eine erfrischend innovative Didaktik der Stochastik. Funktionierende Schulpraxis steht im Vordergrund, solide reflektierte Theorie dahinter. Auf der Grundlage eines umfassenden Wahrscheinlichkeitsbegriffs werden beschreibende Statistik, Wahrscheinlichkeitsrechnung und Kerngedanken beurteilender Statistik von Anfang an spiralcurricular miteinander vernetzt. Dies gelingt – handlungsorientiert – durch spannende und schulalltagstaugliche Fallbeispiele, in deren Zentrum Kinder und Jugendliche mit ihren Alltagsintuitionen und ihrem Interesse an realistischen Fragen stehen. Ziel ist ein nachhaltiger, kognitiv aktivierender Unterricht: Begriffe werden über konkrete Inhalte gebildet, als sinnstiftend erlebt und Zusammenhänge entdeckt.

Der Band ist modular aufgebaut, die Kapitel lassen sich unabhängig voneinander lesen und werden durch wenige Paradigmen zusammengehalten:

- Pflege einen passenden Wahrscheinlichkeitsbegriff.
- Trenne Modell und Realität messerscharf und konsequent.
- Untersuche Zufallsschwankungen statt sie wegzuwünschen.
- Stelle authentische Probleme ins Zentrum.
- Nutze den „didaktischen Dreisatz“ Spekulieren – Experimentieren – Reflektieren.

Der Band richtet sich an Referendarinnen und Referendare sowie an Mathematik-Lehrkräfte beider Sekundarstufen, die spannende Unterrichtsstunden gestalten möchten, an die sich die Lernenden noch lange nach der Schulzeit mit Vergnügen erinnern.

