Bleistiftrollen: Beurteilende Statistik im Federmäppchen

Abb. 1: Beschriftete Bleistift -„Würfel“

Im hektischen Schulalltag kommt das Experimentieren mitunter etwas kurz, obwohl jeder weiß, dass es - insbesondere in der der Stochastik - für eine nachhaltige Entwicklung von Grundvorstellungen unerlässlich ist. Wunderbare Experimente sind sensorische Tests (Cola- oder Schokoladen-Tests mit geraspelten Schokoladensorten [1, 2] oder Hörtests [3] mit CD/MP3 - Musik verschiedener Qualitätsstufen. Wer klebrige Finger oder den Gang in den Musikraum scheut, der findet mit dem hier vorgestellten Bleistiftexperiment eine höchst lohnende Alternative, die praktisch keiner organisatorischen Vorbereitung bedarf.

1 Mit Bleistiften „würfeln“

Wir beschriften die Seiten eines Bleistifts mit den Augenzahlen 1 bis 6. Das Logo bekommt die 6, und wenn es dann mit 5, 4, 1, 2, 3 in einer Richtung weiter geht, haben die Gegenseiten die Augensumme 7, wie bei einem richtigen Würfel. Nun kann man mit den Stiften „würfeln“, indem man sie über den Tisch rollen lässt.

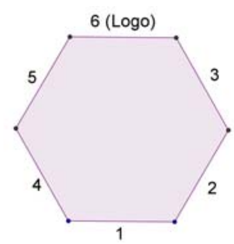


Abb. 2: Seiten-Nummerierung

Wer die Versuchsbedingen genauer festlegen möchte, nimmt das Mathebuch als schiefe Ebene, legt das Logo beim Start nach oben und lässt die Stifte abrollen, bis sie liegen bleiben.

2 Die Individualität der Stifte

Das Hypothesentesten ist bei normalen Spielwürfeln „langweilig“, denn kein Schüler, der es bis in die „Qualifikationsphase“ eines Gymnasiums geschafft hat, bezweifelt ernsthaft, dass die „6“ mit der Wahrscheinlichkeit 1/6 kommt. Bei Bleistiften, die eine oft erheblich unterschätzte Individualität haben, sieht das anders aus. Auch wenn man frisch verpackte Neuware aus Originalschachteln verwendet. Zwecks Wiedererkennung in verschiedenen Kursen bekommt jeder Bleistift einen Namen - oder wenigstens eine Nummer. Für spannende Qualitätsvergleiche empfiehlt es sich, für jeden Schüler zwei Marken bereitzuhalten. Die Stifte heißen dann z. B. Johann Faber - Johann Herlitz, Andrea Faber - Andrea Herlitz. Doch bevor es ans Bleistiftrollen geht, werden Fragestellungen zusammengetragen, eine Erwartungshaltung wird aufgebaut.

3 Nahe liegende Fragen

(1) Kommt bei meinem Stift die 6 mit der Wahrscheinlichkeit 1/6? Wie viele 6er wären dann bei 120 Rollversuchen „normal“?

(2) Kann ich meinen Bleistift als „Laplace-Bleistift“ bezeichnen, bei dem alle Seiten mit der Wahrscheinlichkeit 1/6 kommen?

(3) Gibt es unter den „Bleistiften höchster Qualität“ mehr Laplace-Bleistifte als unter den preiswerteren oder gar No Name-Produkten?

4 Erst simulieren - Erwartungshaltung aufbauen

Es ist sehr lohnend, (… für die Entwicklung tragfähiger Grundvorstellungen zur beurteilenden Statistik unerlässlich …), dass man zunächst Entscheidungskriterien für die genannten Fragen auf intuitiver Grundlage im Kursverband aushandelt und durch Simulationen überprüft und absichert:

Man lässt jeden Schüler 120 mal mit einem normalen Würfel würfeln und präsentiert die Häufigkeitsverteilungen auf einer gemeinsamen Folie oder an der Tafel. Es entsteht ein abgesichertes Gefühl dafür, wie sich ein Laplace-Bleistift verhalten müsste - und welche „Abweichungen vom „Normalen“ nicht vorkommen: Für Frage 2 misst man die Abweichungen zwischen den gewürfelten Augenzahlen $n_1,\dots,\ n_6$ zu der erwarteten Augenzahl 20 durch die Abstandsquadrat-Summe $t = \frac{(n_1-20)^2}{20} + \frac{(n_2-20)^2}{20} + \dots + \frac{(n_6-20)^2}{20}$, die um so kleiner ist, je näher die Häufigkeiten bei einer Gleichverteilung liegen. Die Simulationen geben dann Auskunft, bis zu welcher Obergrenze dieser Testwert t bei Laplace-Würfeln normalerweise schwankt. 120-mal Würfeln dauert mit Auszählen und Zusammentragen der Ergebnisse im Kursverband ca. 25 Minuten. Das ist eine gute und durch die anschließende Bewertung nachhaltig wirkende Investition. Die Erwartungshaltung vor dem realen Bleistiftrollen steigt dadurch enorm.

t bezeichnet man als Chi-Quadrat-Testwert. Die Verteilung von t hat die Dichte
f(t)= $0.133 \cdot t^{1.5} \cdot e^{-t/2}$

Als Ergebnis solcher Simulationen zeigt sich, dass bei Laplace-Würfeln/Bleistiften

- die Anzahl der Sechser mit ca. 95%iger Sicherheit zwischen 12 und 28 liegt
- Der Testwert t mit ca. 95% Wahrscheinlichkeit unter 11 liegt

Durch eine Computersimulation (n=1000 Versuche mit je 120 Laplace-Bleistiften, Abb. 3) wird das abgesichert .

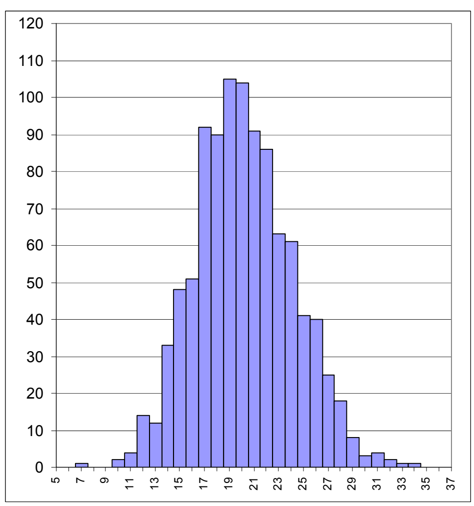
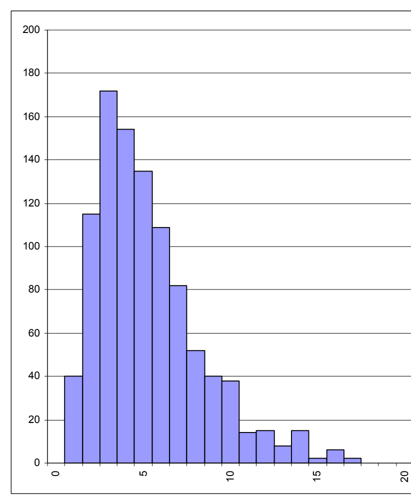


Abb. 3: Häufigkeitsverteilung der Sechser



Verteilung der Testwerte t

Stochastisch gesehen ist ein Bleistift um so besser, je „gleichverteilter“ (oder „lapacescher“) die Augenzahlen sind. Um die „stochastische Qualität“ verschiedener Marken zu vergleichen, verwendet man den Vorzeichentest. Man bildet Stiftepaare. Jeder Schüler notiert ein „+“, wenn die Anzahl der 6er bei seinem teuren Stiften näher an 20 liegt (oder der Testwert t kleiner ist), sonst ein „-“. Wenn 30 Schüler im Kurs sind, entspricht das 30 Münzwürfen, und bei gleicher Qualität erwartet man mit 95% Sicherheit zwischen 10 und 20 positive Vorzeichen (Simulation von Münzwürfen oder Binomialverteilung nutzen). Erst bei mehr als 20-mal „+“ würde man die teuren Stifte, bei weniger als 10-mal „+“ die billigen Stifte als „stochastisch hochwertiger“ bezeichnen.

Wenn man das Experiment im Rahmen beschreibender Statistik als Übung zum Umgang mit Boxplots nutzen möchte, misst man bei Verwendung einer schiefen Ebene (Mathebuch auf Becher gestützt) jeweils auch die Rollweite - oder man notiert die Zeiten, die Schüler für das Würfeln samt Auswertung benötigen (erst schätzen lassen, dann bei laufender Uhr messen)

5 Dann experimentieren

Nach dem Diskutieren und Simulieren gehört die Stunde, in der die Bleistifte tatsächlich gerollt werden, zu den schönsten -und „emotional geladensten", die man im Mathematikunterricht erleben kann. Ohne Unterlass werden (während des Rollens) neue Hypothesen generiert und sicher geglaubte verworfen. Jubeln und Fluchen, ungeduldiges Warten und überraschende Freude erfüllen den Raum: Beurteilende Statistik lebt.

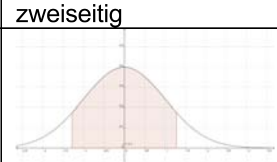
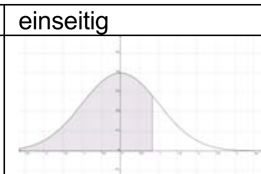
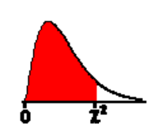
Einen Eindruck von den Ergebnissen vermittelt Abb. 4: Echte Laplace-Bleistifte sind aller Erwartung zum Trotz selten, Markenprodukte schneiden nicht besser ab als No-Name Produkte. Zur Protokollierung der Rollversuche empfiehlt sich die beiliegende Auswertungsvorlage. Wenn man parallel zur Strichliste jede Augenzahl auf den Karos notiert, entfällt die dauernde Kontrolle, ob die 120 schon voll sind.

	Herlitz							Faber							
	1	2	3	4	5	6	t	1	2	3	4	5	6	t	
Mario	18	17	17	25	22	21	2.60	37	42	15	9	12	5	60.4	+
Frauke	20	30	20	20	12	18	8.40	23	20	18	21	15	23	2.4	-
Julia	21	20	27	12	25	15	8.20	12	23	18	11	30	26	14.7	+
Michael	18	29	31	6	3	33	43.00	31	4	4	26	33	22	42.1	-
Henning	16	31	22	23	15	13	11.20	22	21	29	13	12	23	10.4	-
Daniel	18	32	39	11	13	7	40.40	16	18	22	27	17	20	4.1	-
Jaquline	10	22	20	22	20	26	7.20	10	24	26	12	24	24	12.4	+
Wolfgang	6	13	23	27	24	27	18.40	20	17	22	24	20	17	1.9	-
Simone	21	13	26	25	23	12	9.20	33	19	12	21	22	13	14.4	+
Simome	23	38	7	27	20	5	38.80	4	13	28	16	29	30	28.3	-
Nils	44	33	11	18	0	14	63.30	13	17	0	51	29	10	80	+

Abb. 4: Versuchsergebnisse für 11 Bleistiftpaare - Signifikante Abweichungen sind unterlegt

6 Vertiefende Aufgaben

Wenn man (vor oder nach dem Bleistiftexperiment) gelernt hat, mit der Binomialverteilung und der Sigmaregel (Abb. 5) zu arbeiten, wenn man also über die „Werkzeugkiste" der beurteilenden Statistik verfügt, bieten die Bleistifte ein schönes Übungsfeld für vertiefende Übungen zu sinnvollem Hypothesentests und zum Parameterschätzen. Man vergleiche das folgen Aufgabenblatt:

Niveau	zweiseitig	einseitig	Chiquadrat 5df
			
99%	μ ± 2.576σ: [9.5;30.5]	μ+2.326σ=29.7	15.09
95%	μ ± 1.960σ: [12.0;28.0]	μ+1.645σ=26.7	11.07
90%	μ ± 1.645σ: [13.3;26.7]	μ+1.282σ=25.3	9.24
80%	μ ± 1.282σ: [14.7;25.3]	μ+0.842σ=23.5	7.29

$$\mu = n\cdot p = 120\cdot\frac{1}{6} = 20 \quad \sigma = \sqrt{n\cdot p(1-p)} = \sqrt{120\cdot\frac{1}{6}\cdot\frac{5}{6}} = 4.08 \quad \chi^2 = \frac{(n_1-20)^2}{20} + \ldots + \frac{(n_6-20)^2}{20}$$

Abb. 5: Werkzeugkiste der beurteilenden Statistik

Beurteilende Statistik - Aufgaben zum Bleistiftrollen

Zweiseitiges Testen

- Kommt die „6“ bei meinem Stift mit der Wahrscheinlichkeit 1/6?

  ($H_0: p = \frac{1}{6}$ gegen $H_1: p \neq \frac{1}{6}$)
- Kommen „6“ und gegenüberliegende „1“ zusammen mit Wahrscheinlichkeit 1/3?
  (was einer Seite möglicherweise fehlt, kommt der Gegenseite zugute)

  ($H_0: p = \frac{1}{3}$ gegen $H_1: p \neq \frac{1}{3}$)
- Kann ich meinen Bleistift so in zwei Seiten zerlegen, dass Ober- und Unterseite mit gleicher Wahrscheinlichkeit ½ kommen? (z. B. Oberseite: 1-4-5, Unterseite 6-3-2)

  ($H_0: p = \frac{1}{2}$ gegen $H_1: p \neq \frac{1}{2}$)

Einseitiges Testen

- Wolfgang sammelt Stifte bei denen 6 besonders häufig oder 1 besonders selten kommt.

  Er zahlt 1€, wenn man $H_0: p(6) = \frac{1}{6}$ gegen $H_1: p(6) > \frac{1}{6}$ oder $H_0: p(1) = \frac{1}{6}$ gegen $H_1: p < \frac{1}{6}$ verwerfen kann. Kann ich meinen Stift in Zahlung geben?

Fehlerwahrscheinlichkeiten

- Herbert zahlt für Stifte, bei denen die „6“ mindestens mit Wahrscheinlichkeit 1/3 kommt. Entscheidungsregel: Bei 120 Versuchen mindestens 40mal die „6“. Formulieren Sie zwei Alternativ-Hypothesen und bestimmen Sie die Wahrscheinlichkeiten der Fehler erster und zweiter Art.
- Rainer zahlt für Stifte, bei denen die 1 höchstens mit Wahrscheinlichkeit 5% kommt. Formulieren Sie zwei Alternativhypothesen, eine sinnvolle Entscheidungsregel und plottten Sie die zugehörige Operationscharakteristik des Tests.

Schätzen

- Schätzen Sie die Wahrscheinlichkeit der „6“ bei Ihren Stiften durch ein 80%, 95%, 99% Konfidenzintervall.

Anpassungstest ($t = \frac{(n_1 - 20)^2}{20} + \frac{(n_2 - 20)^2}{20} + \ldots + \frac{(n_6 - 20)^2}{20}$)

- Testen Sie, ob Sie einen Laplace-Bleistift erwischt haben, bei dem alle Seiten mit der Wahrscheinlichkeit 1/6 kommen.
- Sie brauchen einen Laplace-Bleistift. Der Anpassungstest hat die Hypothese einer Gleichverteilung nicht zurückgewiesen (so sagen die Angloamerikaneer treffend!)
  a) bei Jan für n=120 auf dem 5% Niveau
  b) bei Bianca für n=1200 auf dem 4% Niveau
  c) bei David für n=120 auf dem 1% Niveau
  d) bei Simone für n=1200 auf dem 1%-Niveau
  Welchen der 4 Bleistifte kaufen Sie?

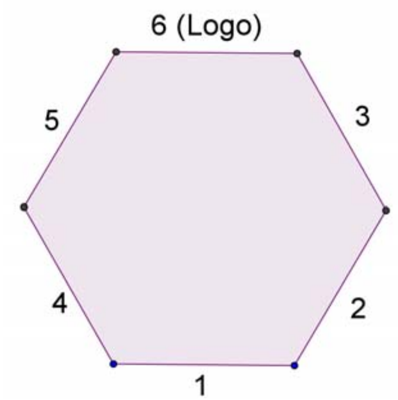
Vorzeichentest

- Sie haben einen teuren und einen billigen Bleistift je 120-mal gerollt. Wenn die Häufigkeiten des hochwertigen Bleistifts näher an der Gleichverteilung liegen (kleinerer Testwert t) als diejenigen des minderwertigen Stiftes, geben Sie ein „+“, sonst ein „–“
  Zählen Sie in Ihrem Kurs die „+“ Ergebnisse und prüfen Sie die Hypothese, ob p(+)=0.5 zurückgewiesen werden muss. Testen Sie einseitig oder zweiseitig?

Auswertungsvorlage Bleistiftrollen

Rollen Sie den Bleistift 120-mal mit ein wenig Schwung über Ihren Holztisch. Notieren Sie die „gewürfelten“ Augenzahlen in den Gitterkaros. Zählen Sie danach aus, wie oft die einzelnen Augenzahlen kamen und notiere Sie die Häufigkeiten an den entsprechenden Bleistiftseiten der Abbildung. (Sie können Strichlisten parallel zum Ausfüllen der Karos auch an den zugehörigen Kanten der Grafik führen, es muss als Summe 120 herauskommen).

Bleistiftnummer/Name

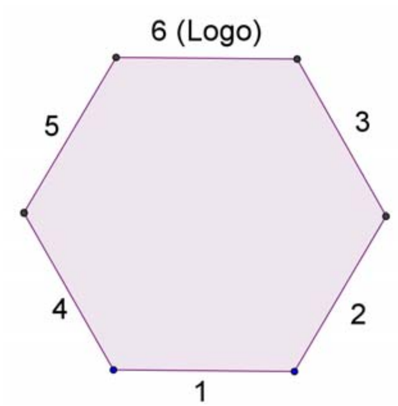


Marke

	1	2	3	4	5	6	7	8	9	10
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										

$$t = \frac{1}{20} \cdot [(\_\_\_ - 20)^2 + (\_\_\_ - 20)^2 + (\_\_\_ - 20)^2 + (\_\_\_ - 20)^2 + (\_\_\_ - 20)^2 + (\_\_\_ - 20)^2] = \_\_\_\_\_$$

Bleistiftnummer/Name



Marke

	1	2	3	4	5	6	7	8	9	10
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										

$$t = \frac{1}{20} \cdot [(\_\_\_ - 20)^2 + (\_\_\_ - 20)^2 + (\_\_\_ - 20)^2 + (\_\_\_ - 20)^2 + (\_\_\_ - 20)^2 + (\_\_\_ - 20)^2] = \_\_\_\_\_$$

Die Anzahl der Sechser liegt näher an 20

bei dem teuren Stift **(+)**

bei dem billigen Stift **(-)**

Der Testwert t ist kleiner

bei dem teuren Stift **(+)**

bei dem billigen Stift **(-)**

Zum Weiterlesen
[1] Riemer, W., Petzold, W.: Geschmackstests - Spannende und verbindende Experimente. mathematiklehren Sammelband Wege in die Stochastik (2008) und 85 (1997)
[2] Riemer, W.: Schmeckt Linde-Schokolade besser als Alpin? Sensorische Experimente im Mathematikunterricht. mathematiklehren 62 (1994)
[3] Riemer, W. Soundcheck: CD contra MP3 Ein Hörtest als Einstieg in die Stochastik. mathematiklehren 153 (2009)

Dr. Wolfgang Riemer,
August-Bebel-Str. 80
50259 Pulheim
w.riemer@arcor.de
www.riemer-koeln.de