Stochastik in der Sek I

In der Stochastik geht es um Daten. Diese werden übersichtlich dargestellt oder beschrieben (das ist mit Informationsverlust verbunden), oder in ihnen wird eine Struktur erkannt (was zum Begriff der Wahrscheinlichkeit führen kann) bzw. vermutet.
Aus diesem Grunde ist es sinnvoll, auch den schulischen Stochastik-Unterricht mit Daten zu beginnen.
Im Doppeljahrgang 5/6 werden daher Daten erhoben, und es werden Daten dargestellt.
Erst dann wird im Doppeljahrgang 7/8 der Begriff der Wahrscheinlichkeit eingeführt und anschließend für Prognosen fruchtbar gemacht.
Im Doppeljahrgang 9/10 werden dann neben den Baumdiagrammen auch Vierfeldertafeln betrachtet, um bei zweistufigen Zufallsexperimenten unbekannte Wahrscheinlichkeiten auf einfache Weise ermitteln zu können.

Klasse 5: Planung und Durchführung statistischer Erhebungen

Der Umgang mit Daten ist grundlegend für den Stochastik-Unterricht. Woher kommen die Daten? Am besten ist es, eigene Daten zu verwenden!
Daher steht im Zentrum des Stochastik-Unterrichts in Klasse 5, eine Befragung oder eine Beobachtung zu planen und durchzuführen. Die Art und Weise der Aufbereitung von Daten wird bewusst erst in der folgenden Klasse 6 thematisiert, um in Klasse 5 die Zeit zur Erhebung statistischer Daten zu haben.

Beispiele sind:

- In welcher Jahreszeit haben die Schülerinnen und Schüler Geburtstag?
  Hier kann man ggf. darüber sich verständigen, wie man mit Zwillingen umgeht: Zählen sie einzeln oder doppelt? Ist die Vorab-Hypothese der Gleichverteilung (einigermaßen) zutreffend? Was wäre, wenn man alle Schülerinnen und Schüler der Schule befragte?
- Anzahl der Geschwister
- Haarfarbe
  Hier muss man über die Klassifikation reden!
- Verkehrszählung
  Was will man alles zählen? Fahräder auch?
- Länge der Wörter im Vorwort des Mathematikbuches
  Was ist ein Wort?
- Anzahl der Stunden im Internet pro Woche
  Hier wird die Schätzproblematik tatsächlich erfahren!

Stets muss man die zu ermittelnden Merkmale identifizieren und die ggf. vorliegende Nichteindeutigkeit der Merkmale diskutieren!
Die fast zwangsläufig sich ergebenden Schwierigkeiten sollten nicht als Probleme, sondern als Chancen zur Erfahrung verstanden werden.

Man soll auch ein *Experiment* planen und durchführen. Hier kann nur ein absolutes Merkmal gezählt werden, also noch keine funktionale Abhängigkeit eines Merkmals von einem anderen.
Beispiele:

- Wie schnell können die Schülerinnen und Schüler ihre Schuhe zubinden?
- Wie schnell können die Schülerinnen und Schüler drei kleine Aufgaben fehlerfrei lösen?
- Können Mädchen schneller rechnen?
  Wie muss ein Experiment gestaltet werden, das diese Frage beantwortet?
- Bringt das Üben (etwa in Bruchrechnung) etwas?
  Wie muss ein Experiment gestaltet werden, das diese Frage beantwortet?

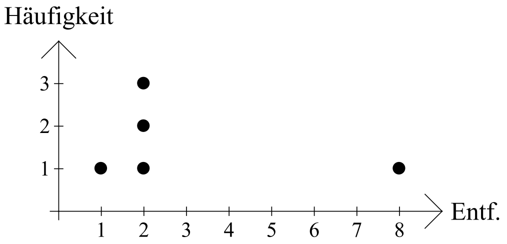
Klasse 6: Maßzahlen statistischer Erhebungen

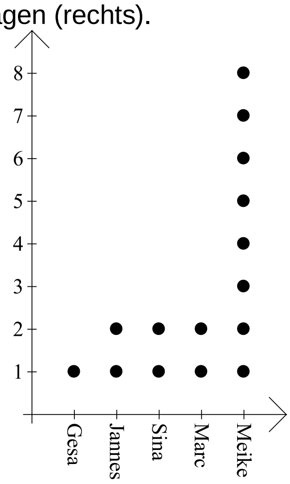
In dieser Jahrgangsstufe steht die Aufbereitung und Darstellung von Daten im Vordergrund.

Beispiel: Entfernungen vom Schulort (man wird mit allen 30 Schülerinnen und Schüler anfangen):

Gesa	1 km
Jannes	2 km
Sina	2 km
Marc	2 km
Meike	8 km

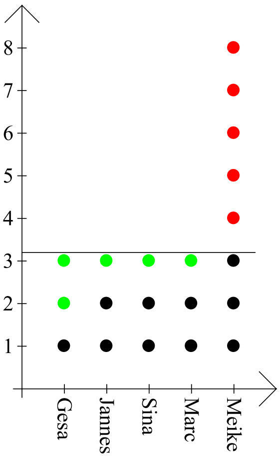
Man hat verschiedene Darstellungsmöglichkeiten: Man kann auf der Rechtsachse Entfernungen und auf der Hochachse die Häufigkeiten auftragen (im folgenden Bild links). Oder: Man kann auf der Rechtsachse die Namen und auf der Hochachse die Entfernungen auftragen (rechts).





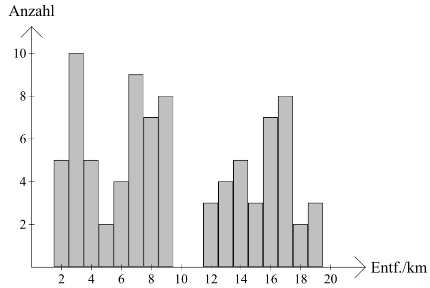
Das arithmetische Mittel ist 3. Wie kann man das veranschaulichen?
Dies geht bei der rechten Darstellung wesentlich einfacher:

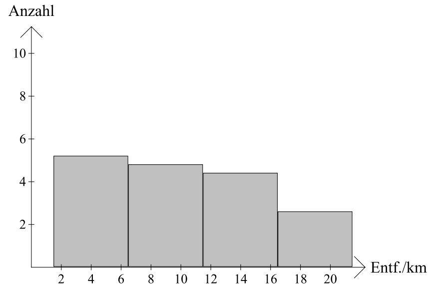
Die Anzahl der Überschüssigen oben
ist so groß wie die Anzahl der Unterschüssigen unten.



Gleichwohl: Analog zu Strichlisten wird man sich oft dafür entscheiden, auf der Rechtsachse die Entfernungen und auf der Hochachse die Häufigkeiten anzutragen.

Balkendiagramm bekommt man in GeoGebra mit dem Befehl Balkendiagramm. Je nach Breite der Balken lassen sich unterschiedliche Schlüsse ziehen:





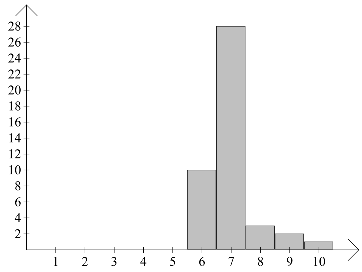
Links hat man eine zweigipflige Verteilung; in der Entfernung von ca. 10 km wohnt niemand.

Hier ist eine monotone Abnahme zu erkennen, die man links nicht sieht.

Der *Modalwert* ist der Wert mit der größten Häufigkeit. Er lässt sich inhaltlich bedeutend leichter als der Median vom arithmetischen Mittel abgrenzen. Aus diesem Grunde wird der Median erst im Sekundarbereich II behandelt.

Im obigen Doppelbild links ist 3 der Modalwert, rechts 4.
Auch die arithmetischen Mittel unterscheiden sich: links etwa 9,75 und rechts etwa 11,3.

Arithmetisches Mittel und Modalwert können auch bei asymmetrischen Verteilungen zusammenfallen, wie man rechts sieht.



Zur Problematik der unterschiedlichen Aussagekraft unterschiedlicher Parameter empfiehlt sich auch das folgende Beispiel:
Schüler Cedric hat in den drei ersten Mathematikarbeiten die Noten 3, 6 und 2 geschrieben. Bei Schülerin Lia waren es die Noten 4, 3 und 3.
Wer ist besser?
Cedric hat das arithmetische Mittel 3,7 und Lia hat das arithmetische Mittel 3,3. Also ist Lia besser.
Aber: Meistens ist Cedric besser!

Die Umfrage „Wen halten Sie für besser, A oder B?“ ist verschieden von der Umfrage „Benoten Sie A und B jeweils mit den Schulnoten 1 bis 6!“ und kann zu einem ganz anderen Ergebnis führen!

Rückgabe einer Klassenarbeit: Die Durchschnittsnote war „3“. Die häufigste Note war „2“. Kann das sein? Kann die häufigste Note auch eine „1“ sein?

Die Aussagekraft von Modal- und Mittelwert kann diskutiert werden anhand folgender Fragen:
- In welcher Jahreszeit Geburtstag?
- Anzahl der Geschwister?
- Supermarkt fragt Kunden nach Postleitzahl
- Lebensalter der Schülerinnen und Schüler (auf‘s Jahr gerundet)
- Ausfall zweier aufeinander folgender Klassenarbeiten
- Ausfall zweier Tests: (etwa einer vor einem Gruppenturnier, einer danach)

Als einfaches Streumaß bietet sich die *Spannweite* an.
Die Streumaß-Berechnung simpel, daher sollte man die Problematik besser umdrehen:
Erstelle fiktiven Datensatz mit vorgegebenen Werten für arithmetisches Mittel, Modalwert und Spannweite!
Oder:
Der Durchschnitt der Höchsttemperaturen in der letzten Woche war 8 °C. Der Modalwert war 2 °C. Welche Höchsttemperaturen passen dazu?

Klasse 7: Wahrscheinlichkeit

Nun kommt endlich der Begriff der Wahrscheinlichkeit.

Wahrscheinlichkeiten modellieren relative Häufigkeiten. Der Unterschied zwischen realer Welt und Modell muss erfahrbar sein.
Beginnt man mit asymmetrischen Objekten wie der Reißzwecke, so ist der begriffliche Unterschied zwischen Wahrscheinlichkeit und relativer Häufigkeit kaum zu vermitteln.
Beginnt man mit vollsymmetrischen Objekten, hat man für die Betrachtung relativer Häufigkeiten kaum einen Anlass (außer beim Gesetz der großen Zahlen).

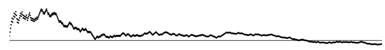
Daher haben sich zur Einführung *teilsymmetrische Objekte* (wie Riemer-Quader oder Legosteine) bewährt.
Hier sind manche Seiten schon aus geometrischen Gründen gleichwahrscheinlich, führen aber i.a. nicht zu übereinstimmenden relativen Häufigkeiten.
Der Unterschied zwischen den beobachteten relativen Häufigkeiten verringert sich bei zunehmender Wurfzahl. Gleichwohl handelt es sich *nicht* um analytische Konvergenz!
Verschiedene Schüler schätzen die Wahrscheinlichkeiten für verschiedene Seitenflächen ähnlich, aber dennoch unterschiedlich, und es ergeben sich Diskussionen, wie sie für die Genese eines tragfähigen Wahrscheinlichkeitsbegriffs fruchtbarer nicht sein könnten. Durch Experimente werden die Schüler gezwungen, ursprüngliche Hypothesen zu verwerfen, wie z. B. die der Proportionalität der Wahrscheinlichkeiten zu den Seitenflächen.

Es wird nicht einseitig der frequentistische Wahrscheinlichkeitsbegriff angesteuert (Wahrscheinlichkeit als Grenzwert relativer Häufigkeiten), sondern auch der subjektivistische (Wahrscheinlichkeiten haben mit zu revidierenden Hypothesen zu tun).

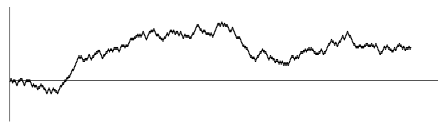
Wenn man es nicht mit vollsymmetrischen Objekten zu tun hat, gilt:
Wahrscheinlichkeiten sind Größen, denen man durch Messung näher kommt, sie aber auch durch Messung nie wird exakt bestimmen lassen.
Die Welt der Daten setzt sich zusammen aus einem Muster plus einer Variabilität:

Daten = Muster + Variabilität
= signal + noise

Nach Klärung des begrifflichen Unterschieds zwischen Wahrscheinlichkeit und relativen Häufigkeiten stellt sich die Frage: Wie steht es mit vollsymmetrischem Objekt wie einer Münze?
Betrachtet man die relativen Häufigkeiten, hat man das nebenstehende Bild; die waagerechte Linie hat die Höhe 1/2.



Was ist mit den absoluten Häufigkeiten? Trägt man bei 100 Würfen deren jeweilige Unterschiede gegen n/2 auf, ergibt sich ein ganz anderes Bild; von Konvergenz kann keinerlei Rede mehr sein!

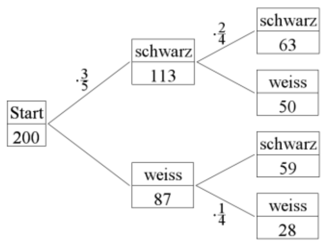


Klasse 8: Ein- und zweistufige Zufallsversuche

In Kl. 7 ging es um den modellierenden Übergang von relativen Häufigkeiten zu Wahrscheinlichkeiten. Man fragt sich: Wozu nützt das Modell? Antwort: Das Modell wird zur Prognose relativer Häufigkeiten verwendet.
Es geht daher in Klasse 8 um den modellverwendenden Übergang von Wahrscheinlichkeiten zu relativen Häufigkeiten.
Insbesondere lassen sich so auch falsche Vorstellungen zu Wahrscheinlichkeiten anhand relativer Häufigkeiten falsifizieren.

Eine Münze (Laplace-Objekt, damit die Wahrscheinlichkeit unstrittig ist) wird 100-mal geworfen. Wie häufig erscheint die nationale Seite?
Die Schülerinnen und Schüler bekommen ein Gefühl dafür, wie viele „Erfolge" beim 100-maligen Werfen einer fairen Münze zu erwarten sind: Es können 40 oder 53 „Erfolge" sein, aber nur in ganz seltenen Fällen 10. Auch die 50 wird nicht häufig auftreten.
Man bekommt auf diese Weise binomialverteilte Zufallszahlen (ohne sie so zu nennen).
Das geht auch direkt mit GeoGebra [ZufallszahlBinomialverteilt[n, p]].
Der Zufallszahlengenerator wirkt dabei jeweils als Black Box, deren Ergebnisse im Großen und Ganzen den Erfahrungen mit Münzwürfen entsprechen (die Funktionsweise wird erst in der Sek II geklärt).
Damit kann die Variablilität der Einträge (absolute Häufigkeiten!) in einem Baumdiagramm oder in einer Vierfeldertafel erfahren werden!

Auch die Pfadregeln erschließen sich mithilfe fiktiver absoluter Häufigkeiten leicht auf diese Weise: Ich habe eine Urne mit 3 scharzen und 2 weißen Kugeln und ziehe 2-mal OHNE Zurücklegen. Ich spiele dieses Experiment 200-mal:
In *etwa* 120 Fällen ziehe ich beim 1. Mal „schwarz".
Von diesen *etwa* 120 Fällen ziehe ich *etwa* 60-mal beim 2. Mal „schwarz".



Wahrscheinlichkeit für s / s?
In *etwa* 60 von 200 Fällen ziehe ich s / s.
Die Wahrscheinlichkeit für s/s ist somit $\frac{60}{200}=\frac{120}{200}\cdot\frac{60}{120}=\frac{3}{5}\cdot\frac{2}{4}$.

Klasse 9: Baumdiagramme und Vierfeldertafeln

Hier geht es um Zufallsexperimente mit zwei Merkmalen. Dabei sollte man mit merkwürdigen Phänomenen beginnen:

In der 9a tragen die meisten Mädchen eine Brille.
In der 9a sind die meisten Brillenträger männlich.

Folgt daraus, dass die meisten Mädchen männlich sind?
Die beiden Aussagen sind durchaus miteinander verträglich:

	M	J	
Br	10	15	25
¬Br	5		
	15		

Man sieht: Es kommt gar nicht darauf an, wie viele Jungen keine Brille tragen. Finden Schülerinnen und Schüler analoge Beispiele?

Bei einer Vorsorgeuntersuchung wird auf eine bestimmte Krankheit getestet. Der Prozentsatz der tatsächlich Kranken betrage 10 % (woher weiß man das?). Bei einer in Wahrheit kranken Person rekenne der Test mit Wahrscheinlichkeit 85 % auf „krank“; bei einer in Wahrheit gesunden Person erkenne der Test mit Wahrscheinlichkeit 95 % auf „gesund“.
Alle Zahlenangaben sind Schätzwerte. Man sollte also die Bayes-Aufgaben nicht zu quantitativ lösen. Der hier deutlich werdende Modellierungsaspekt ist wesentlich.
Fachbegriffe wie Prävalenz, Sensitivität, Spezifizität tragen hier überhaupt nicht zur Klärung bei.

Die Frage, mit welcher Wahrscheinlichkeit eine als „krank“ getestete Person in Wahrheit krank ist, wird viel weniger fehlerträchtig gelöst, wenn man fiktive absolute Häufigkeiten verwendet.
Nehmen wir an, dass der Test bei $n = 1000$ Personen durchgeführt wird (repräsentative Stichprobe). Dann werden etwa 100 Personen tatsächlich krank sein. Allerdings sind das sicherlich nicht genau 100 Leute. In welchem Bereich diese Anzahl liegen wird: Das beantwortet der im Schuljahr davor kennen gelernte Zufallszahlengenerator. Das mit den fiktiven absoluten Zahlen verbundenene Gefühl für Variabilität sollte nicht verloren gehen.
Von den etwa 100 kranken Personen werden etwa 85 als „krank“ angesehen usw.
Nun kann man p(k | „k“) als Quotient ermitteln. Man sieht: Dieser Wert ist eine Zufallsgröße. Er hängt ab von der 1000-er Stichprobe sowie vom Test.

	"krank"	"gesund"	
krank	99	12	111
gesund	40	849	889
			1000

Verblüffender ist: Was passiert, wenn der Anteil der tatsächlich Kranken in der Bevölkerung auf 1 % sinkt?

<table>
<tr><td></td><td>"krank"</td><td>"gesund"</td><td></td><td></td><td>pk =</td><td>0,01</td></tr>
<tr><td>krank</td><td>87</td><td>15</td><td>102</td><td></td><td>qk = p("k" | k) =</td><td>0,85</td></tr>
<tr><td>gesund</td><td>500</td><td>9398</td><td>9898</td><td></td><td>qg = p("g" | g) =</td><td>0,95</td></tr>
<tr><td></td><td>587</td><td>9413</td><td>10000</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>0,1482</td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Der Unterricht sollte auf jeden Fall Vierfeldertafel *und* Baumdiagramm pflegen, um die Schülerinnen und Schüler nicht auf eine einzige Lösungsmöglichkeit zu „trimmen“.
Im Unterricht zeigte sich nämlich immer wieder, dass manche Lernenden lieber mit dem einen Format arbeiten und andere mit dem anderen Format.