Schwierigkeiten mit dem Stabilwerden der relativen Häufigkeiten - Das $\frac{1}{\sqrt{n}}$ - Gesetz

von **Ödön Vancsó**, Budapest, und **Elke Warmuth**, Berlin

Zusammenfassung: *Auf eine kurze Einführung über Interpretationen von Wahrscheinlichkeit folgen im 2. Teil heuristische Untersuchungen des Verhaltens der relativen Häufigkeit sowie didaktische Überlegungen zum schwachen Gesetz der großen Zahlen. Im 3. Teil berichtet Ö. Vancsó über Schwierigkeiten im Umgang mit diesem Gesetz. Schließlich werden im 4. Teil Grundzüge einer möglichen Unterrichtssequenz für einen Leistungskurs in der S II zur Behandlung des Gesetzes der großen Zahlen im Rahmen des Schätzens einer unbekannten Wahrscheinlichkeit vorgestellt.*

1. Modelle für Vorgänge mit zufälligem Ergebnis und Interpretationen von Wahrscheinlichkeit

Im Stochastikunterricht entwerfen und untersuchen wir mathematische Modelle für Vorgänge mit zufälligem Ergebnis. Wir nennen einen Vorgang „Vorgang mit zufälligem Ergebnis", wenn er mehrere mögliche Ausgänge (Ergebnisse) hat und vor Ablauf des Vorgangs das Ergebnis nicht sicher vorhergesagt werden kann. Die Ergebnismenge Ω als Modellbestandteil enthält Repräsentanten für die möglichen Ergebnisse. Mit der Wahl der Ergebnismenge legen wir fest, auf welche Merkmale des Vorgangs sich unser Interesse richtet.

Beispiel: Ein Würfel wird fünfmal geworfen.

a) Es interessieren die einzelnen erhaltenen Augenzahlen. Eine mögliche Wahl ist $\Omega = \{(a, b, c, d, e)\text{: } a, b, c, d, e \in \{1, 2, ..., 6\}\}$.

b) Es interessiert die Anzahl der Würfe vor der ersten Sechs. Eine mögliche Wahl ist $\Omega = \{6, 06, 006, 0006, 00006, 00000\}$.

Wir wollen uns auf den Fall $\Omega = \{\omega_1, \omega_2, ..., \omega_r\}$, d. h. auf Vorgänge mit endlich vielen möglichen Ergebnissen beschränken. Im Modell werden den möglichen Ergebnissen Wahrscheinlichkeiten zugeordnet:

Ergebnis	ω_1	ω_2		ω_r
Wahrscheinlichkeit	p_1	p_2	...	p_r

Die einzigen Bedingungen an diese Zuordnung sind:

$$p_k \geq 0 \text{ für } k = 1, 2, ..., r \quad \text{und} \quad p_1 + p_2 + ... + p_r = 1.$$

Für das Rechnen ist es gleichgültig, was man sich unter Wahrscheinlichkeit vorstellt. Für die Anwendung und den Lernprozeß muß die Verbindung von Modellebene und Sachebene durch Interpretationen von Wahrscheinlichkeit hergestellt werden. Wir halten es für wichtig, im Mathematikunterricht die Sachebene und die Modellebene zu unterscheiden. Wir fassen Mathematik als eine **Tätigkeit** auf, bei der die Menschen (auch die Schüler) die Umwelt mit mathematischen Mitteln zu erfassen suchen. Der Wirklichkeitsbezug des Mathematikunterrichts entsteht nicht, indem den Schülern nur fertige Modelle vorgesetzt werden, in denen sie etwas "ausrechnen" sollen.

Wenn wir die Modellwahrscheinlichkeiten p_k auswählen, verarbeiten wir unser Wissen, unsere Beobachtungen, unsere Intuitionen über den realen Vorgang. Das Modell verkörpert unsere Sicht auf diesen Vorgang. Intuitiv ist die Wahrscheinlichkeit zunächst ein Maß für die Gewißheit über das Eintreten eines Ergebnisses in einer Skala von 0 bis 100% bzw. von 0 bis 1. Dieses Maß kann zahlenmäßig fixiert werden durch

- Gleichverteilungsannahmen (klassische Wahrscheinlichkeit)
- Häufigkeitsinterpretationen
- das Konstruieren von Wettsituationen
- subjektives Expertenwissen.

Der fixierte Zahlenwert muß nicht endgültig sein. Neue Informationen, weitere Beobachtungen, nicht plausible Folgerungen aus den Modellrechnungen können uns veranlassen, das Modell zu verändern.

Die meisten Rahmenpläne für den Stochastikunterricht und im Gefolge die Lehrbücher bevorzugen einseitig die Interpretation von Wahrscheinlichkeit als stabilem Wert der relativen Häufigkeit. Allgemeinheit und Anwendbarkeit der Theorie werden dadurch unnötig eingeschränkt. Ein Indiz dafür ist z. B. die Forderung nach Wiederholbarkeit als Grundvoraussetzung an die betrachteten Vorgänge. Stochastik erscheint als Mathematik der Massenerscheinungen. Viele Beispiele und Aufgaben passen aber nicht dazu.

Beispiel: "In einem Restaurant essen 60% der Gäste keine Vorspeise und 50% keinen Nachtisch. 30% bestellen weder Vorspeise noch Nachspeise. Wie groß ist die Wahrscheinlichkeit, daß ein Gast, der keine Nachspeise hatte, auch keine Vorspeise hatte?"

Man wird in diesem Beispiel ohne Mühe 0,6 für die gesuchte Wahrscheinlichkeit ausrechnen. Wie soll eine Häufigkeitsinterpretation aussehen, wenn es um einen konkreten Gast geht? Eine Deutung des Zahlenwertes 0,6 könnte darin bestehen, einen "fairen" Wetteinsatz zu beziffern. Jemand bietet 10 DM als Wetteinsatz darauf, daß der besagte Gast eine Vorspeise hatte. Wieviel ist uns dann die Wette wert, daß er keine Vorspeise hatte? Eine faire Wette darf uns 15 DM wert sein!

Im täglichen Leben begegnen uns viele solcher "Wettsituationen": beim Abschluß einer Versicherung, bei der Vorbereitung einer Reise, beim Anlegen von Geld, bei Sport und Spiel, ...

Im Stochastikunterricht sollten wir intuitive Vorstellungen und Vorerfahrungen der Schüler über Wahrscheinlichkeit aufgreifen und dabei verschiedene Deutungskonzepte miteinander verbinden.

2. Das schwache Gesetz der großen Zahlen - Die Theorie

Zunächst wird (im wesentlichen Jacob Bernoulli folgend) ein Modell für Bernoulli-Versuche konstruiert.

Das Ergebnis einer Versuchsserie der Länge n kann als "Wort" der Länge n, bestehend aus den Buchstaben **E**(rfolg) und **M**(ißerfolg), aufgeschrieben werden. Die Ergebnismenge Ω besteht aus den 2^n möglichen „Wörtern". In einem ersten Schritt wird die Bedingung ausgenutzt: "Bei einem Versuch sind der Erfolg und der Mißerfolg gleichwahrscheinlich wie z. B. beim Werfen einer regelmäßigen Münze". Das entspricht einem wichtigen didaktischen Prinzip, wonach wir von den einfacheren und greifbaren Fakten ausgehend auf einem langen Weg die abstrakteren und tieferen Tatsachen erreichen und begreifen müssen. In diesem Fall gibt es 2^n verschiedene **gleichwahrscheinliche** Ergebnisse. Uns interessiert, wie viele solcher Ergebnisse existieren, bei denen die Buchstaben E und M nahezu gleich häufig sind (wobei die Differenz der Anzahl von E und M verglichen mit n klein ist). Entgegengesetzt dazu besteht die ebenso wichtige Frage, wie oft E oder M häufiger in einem Wort vorkommt. Bekanntlich führen diese Fragen direkt zur Binomialverteilung, in diesem Fall wegen p=0,5 zu einer symmetrischen Verteilung. Wenn X_n die Anzahl der Buchstaben E in einem Wort der Länge n bezeichnet, so ist die Anzahl der Wörter, in denen genau k-mal E auftritt, gleich $\binom{n}{k}$. Mit Hilfe der klassischen Berechnungsvorschrift der Wahrscheinlichkeit (nach Laplace) erhalten wir: $P(X_n = k) = \binom{n}{k}\frac{1}{2^n}$. Wenn man mit Hilfe eines Computers diese Wahrscheinlichkeiten graphisch darstellt, so sieht man, daß der Unterschied der Anzahl der Buchstaben E und M in den meisten Wörtern "klein verglichen mit n" ist, die Verteilung also stark in der Mitte konzentriert ist. Die sehr großen Schwankungen, z. B. alle Buchstaben sind E, kommen sehr selten vor, im Mittel genau einmal unter 2^n Fällen. Ist n groß genug, wird diese Chance sehr gering. Wenn wir von den (absoluten) Häufigkeiten des Buchstaben E zu den relativen Häufigkeiten übergehen und beachten, daß die Erfolgswahrscheinlichkeit p=0,5 ist, so können wir unsere Beobachtung so formulieren: Die relative Häufigkeit des Buchstaben E liegt "viel lieber" in der Nähe von 0,5 als weit davon entfernt. Und dies wiederum umso

lieber, je größer n ist. (Die Verteilung der relativen Häufigkeit im Fall p=0,5 ist in Abb. 4 in Abschnitt 4 gezeigt.)

Hier tritt ein wesentliches Element auf: die **Unsicherheit** beim Vergleichen der relativen Häufigkeit mit der Wahrscheinlichkeit eines Ereignisses. Unterrichten wir in einer normalen Klasse, reicht es aus, diese Tatsache **qualitativ** zu verstehen. Abhängig vom Interesse der Schüler (und dem Rahmenplan) ist es möglich (notwendig), die Mathematisierung dieser Erkenntnis auf verschiedenen Stufen zu verwirklichen (Die zwei Endpunkte der Skala sind eine Formel ohne irgendeine Herleitung, die anwendbar ist bei vielen Aufgaben, und eine mathematisch präzise Herleitung.)

Die Mathematik tritt dann in Aktion, wenn wir die Aussage " klein verglichen mit n" bzw. „viel lieber" genauer analysieren möchten. Die Antwort auf die Frage:

Wie groß kann die Abweichung vom Wert p=0,5 sein?

ist: Beliebig, aber doch ist sie „viel öfter" klein als groß. Wenn wir die Formel von Stirling für n! anwenden, können wir für gerades n ausrechnen, wie groß die Wahrscheinlichkeit dafür ist, daß die relative Häufigkeit von E gleich 0,5 ist. Dies ist unter allen möglichen der wahrscheinlichste Wert. Für $n=2k$ beträgt seine Wahrscheinlichkeit ungefähr $\frac{1}{\sqrt{\pi \cdot k}}$ (vgl. Anhang). Also ist die Wahrscheinlichkeit des Ereignisses "Die relative Häufigkeit von E ist gleich 0,5" von der **Größenordnung** $\frac{1}{\sqrt{n}}$. Mit Hilfe der Tschebyschewschen Ungleichung können wir ein Intervall um 0,5 bestimmen, in das die relative Häufigkeit mit einer großen Sicherheit („sehr gern") fällt. Die Tschebyschewsche Ungleichung besagt hier

$$P\left(\left|\frac{X_n}{n} - 0{,}5\right| \geq \varepsilon\right) \leq \frac{1}{4n\varepsilon^2}.$$

Wenn wir für ε den Wert $\frac{1}{\sqrt{n}}$ wählen, dann erhalten wir als Abschätzung 0,25. Aber 0,75 ist noch keine große Sicherheit, deshalb erhöhen wir auf $\frac{2}{\sqrt{n}}$. Für dieses ε erhält das Intervall $[0{,}5-\frac{2}{\sqrt{n}}, 0{,}5+\frac{2}{\sqrt{n}}]$ eine Sicherheit von mindestens $1-\frac{1}{16} \approx 0{,}94$. Wir sehen also, daß mit großer Wahrscheinlichkeit die Abweichungen der relativen Häufigkeit von 0,5 die Größenordnung $\frac{1}{\sqrt{n}}$ besitzen. Das ist die Aussage des $\frac{1}{\sqrt{n}}$-Gesetzes.

Bis jetzt haben wir uns nur mit dem Fall p=0,5 beschäftigt. Der zweite Schritt ist: p sei eine beliebige Zahl zwischen 0 und 1.

Wir können wieder den Computer als Hilfsmittel verwenden. Die Binomialverteilung mit p<0,5 oder p>0,5 ist nicht symmetrisch und auch das Rechnen wird komplizierter. *Die Beobachtung bleibt aber dieselbe:* Für große n liegt mit hoher Wahrscheinlichkeit die relative Häufigkeit in der Nähe von p. Wieder mit Hilfe der Stirling-Formel erhalten wir: die Wahrscheinlichkeit des wahrscheinlichsten Wertes liegt in der Größenordnung $\frac{1}{\sqrt{n}}$ (vgl. Anhang). In analoger Weise findet man mit Hilfe der Tschebyschewschen Ungleichung ein Intervall der Länge $c \cdot \frac{1}{\sqrt{n}}$, in dem die relative Häufigkeit mit hoher Wahrscheinlichkeit liegt.

Das schwache Gesetz der großen Zahlen folgt durch Grenzübergang aus der Tschebyschewschen Ungleichung:

$$\text{Für alle } \varepsilon > 0 \text{ gilt } \lim_{n\to\infty} P\left(\left|\frac{X_n}{n} - p\right| \le \varepsilon\right) = 1.$$

Analysieren wir die Aussage dieses Gesetzes, so können wir folgendermaßen formulieren: Seien ein $\varepsilon > 0$ und ein **Sicherheitsniveau** (z. B. 0,95) gegeben, dann gibt es ein n_0, so daß für alle $n \ge n_0$ gilt $P\left(\left|\frac{X_n}{n} - p\right| \le \varepsilon\right) \ge 0{,}95$.

Unser Satz sichert die Existenz einer solchen, manchmal sehr großen Zahl für n_0. Die Tschebyschewsche Ungleichung ermöglicht es darüber hinaus, ein ausreichend großes n_0 konkret zu bestimmen. Dieser Satz behauptet aber **nicht**, daß **immer**, wenn n größer als diese Grenze ist, die relative Häufigkeit in der vorgeschriebenen Umgebung von p bleibt. Die hier beschriebene Konvergenz der Folge von Zufallsgrößen $\left(\frac{X_n}{n}\right)$ ist deshalb keine Konvergenz von Folgen im Sinne der klassischen Analysis.

Es gibt in der Theorie einen „stärkeren" Satz - das starke Gesetz der großen Zahlen. Es besagt, daß mit Wahrscheinlichkeit 1 die Folge $\left(\frac{X_n}{n}\right)$ gegen p konvergiert: Für fast alle ω gilt $\lim_{n\to\infty} \frac{X_n(\omega)}{n} = p$. Es ist also fast sicher, daß die Konvergenz stattfindet, aber es ist auch (für eine Ausnahmemenge mit der Wahrscheinlichkeit 0) möglich, daß sie nicht stattfindet.

3. Typische Schwierigkeiten von Lehrern im Umgang mit dem schwachen Gesetz der großen Zahlen

Zunächst muß erwähnt werden, daß die Stochastik in der ungarischen Schulpraxis sehr vernachlässigt wird. Die Probleme, über die im folgenden berichtet werden soll, kommen vor allem an der Universität vor. Unter Lehramtsstudenten gibt es mehrere Unklarheiten, die im wesentlichen in drei Gruppen eingeteilt werden können:

a) Was bedeutet das Gesetz? Gehört es in die Sachebene oder in die Modellebene? Was bedeutet die Wahrscheinlichkeit und welche Beziehung gibt es zwischen der relativen Häufigkeit und der Wahrscheinlichkeit? Sei diese Gruppe als Sinnfrage bezeichnet.
b) Welche "Konvergenz" gilt hier? Was verstehen wir unter stochastischer Konvergenz? Sei diese Gruppe als Konvergenzfrage bezeichnet.
c) Was kann man über die Geschwindigkeit der Konvergenz sagen? Wie viele Experimente müssen nacheinander wiederholt werden, damit wir mit der relativen Häufigkeit die Wahrscheinlichkeit schätzen können?

Die Erfahrungen stammen aus der seit zehn Jahren regelmäßigen Teilnahme an den Abschlußprüfungen für die Mathematiklehrerstudenten an der Eötvös-Loránd-Universität von Budapest, und aus vielen Gespräche mit Studenten und Lehrern.

Zur Sinnfrage: Eine starke Tradition ist die Betonung des axiomatischen Aufbaus. Also ist der Umgang mit diesem Problem sehr mathematisch und rein formal ohne Konkretisierung. Die Studenten bewegen sich fast nur in der Modellebene. Die Frage, woher die "Grundwahrscheinlichkeiten" stammen, aus denen wir "sekundäre Wahrscheinlichkeiten" ausrechnen, wird vermieden, als "Nicht-Mathematisches" verdrängt. Unseres Erachtens ist diese Haltung selbst für Mathematiker gefährlich, aber unzulässig für werdende Mathematiklehrer. Es muß erklärt werden, daß eben dieses Gesetz eine Brücke zwischen unseren Erfahrungen (der Sachebene) und dem Modell, das im ersten Abschnitt unseres Artikels vorgestellt worden ist, schlägt. Nach unseren Erfahrungen gibt es einen stabilisierenden Charakter der relativen Häufigkeit in einer längeren Versuchsserie. Das Bernoullische Gesetz zeigt die klare Anpassung unseres Modells an die zu beschreibende Situation. Das Gesetz der großen Zahlen ist ein wesentlicher Bestandteil des mathematischen Modells, ein mathematisches Theorem.
Zusammenfassend, wenn ein Student das Gesetz überhaupt kennt, denkt er nur an den mathematischen Inhalt ohne Reflexion über die Anwendbarkeit und die Beziehung zur realen Situation. Das ist typisch für die Schulmathematik in Ungarn: die Mathematisierung wird in den Hintergrund gedrängt gegenüber der "reinen Mathematik".

Unter der Wahrscheinlichkeit verstehen die Gefragten meistens eine objektive Zahl, oft mit einem kombinatorischen Hintergrund. Von den im ersten Abschnitt erwähnten vier Möglichkeiten der Deutung von Wahrscheinlichkeit kommen nur die ersten zwei unter Studenten vor. Das zeigt den zu eng aufgefaßten Begriff von Wahrscheinlichkeit. Weniger als 20% der gefragten Studenten (unter etwa 200) haben gute Vorstellungen über die Beziehung zwischen der relativen Häufigkeit und der Wahrscheinlichkeit. Aber mehr als 30% haben nicht einmal die Frage verstanden. Die Sinnfragen kommen so selten während ihrer Ausbildung vor, daß sie große Schwierigkeiten haben, diese zu beantworten.

Zur Konvergenzfrage: Die meisten Studenten meinen, beim schwachen Gesetz konvergiert die relative Häufigkeit gegen die Wahrscheinlichkeit im Sinne der Analysis. Auf die danach gestellte Frage, warum der Satz so kompliziert ist (mit zwei P), können sie keine Antwort geben. Die Unsicherheit der Konvergenz spielt hier eine wichtige störende Rolle, deren wesentlicher Charakter von den Studenten nicht erfaßt wird.

Der dritte Punkt - siehe z.B. Freudenthal 1972 - steht in der Schulpraxis im Mittelpunkt. Wie lange müssen wir ein Experiment nacheinander wiederholen, damit ein "zuverlässiges Ergebnis" herauskommt. Im allgemeinen gibt es wenigstens eine Größenordnung Unterschied zwischen den getippten Anzahlen und den aus dem $\frac{1}{\sqrt{n}}$-Gesetz gefolgerten Anzahlen. Die meisten Gefragten wissen nicht, wie sie eine solche Frage beantworten sollen, also können sie den Satz nicht wirklich anwenden. Die häufigste Antwort auf die Frage, eine wie lange Versuchsserie ausreichend ist, wenn wir die Wahrscheinlichkeit z. B. mit der Genauigkeit von zwei Hundertstel bestimmen möchten, ist Hundert oder einige Hundert. Die Wahrheit ist: mindestens zwei-drei Tausend, was in einer Klasse schwer realisierbar ist. (Siehe die Tabelle im 4. Abschnitt nach Abb. 5.)

4. Das schwache Gesetz der großen Zahlen im Stochastikunterricht

Ausgangspunkt

Wir betrachten eine Bernoulli-Kette der Länge n. Die Wahrscheinlichkeit des Erfolges A sei $p=P(A)$ und die Zufallsgröße X_n beschreibe die Anzahl der Erfolge in dieser Bernoulli-Kette. Den Quotienten $\frac{X_n}{n}$ nennen wir die relative Häufigkeit der Erfolge. Das Gesetz der großen Zahlen können wir nun so formulieren:

$$\text{Für alle } \varepsilon > 0 \text{ gilt } \lim_{n\to\infty} P\left(\frac{X_n}{n} \in [p-\varepsilon, p+\varepsilon]\right) = 1.$$

Das Gesetz beschreibt in der Modellebene das asymptotische Verhalten der relativen Häufigkeit $\frac{X_n}{n}$. Schon deshalb kann es kein Beweis dafür sein, daß sich in der Realität die empirisch beobachteten relativen Häufigkeiten stabilisieren. Es schlägt aber eine Brücke zur Sachebene und stützt die Häufigkeitsinterpretation einer Wahrscheinlichkeit, wenn wir folgendes annehmen:

a) Die wiederholten Beobachtungen eines gegebenen Vorgangs erfolgen so, daß eine Bernoulli-Kette als Modell für diese Beobachtungsserie geeignet erscheint.
b) Ereignisse mit sehr kleiner Wahrscheinlichkeit treten im Einzelversuch praktisch nicht ein.

Die Annahme a) führt dazu, daß die Quotienten $\frac{X_n}{n}$ mitsamt ihren Eigenschaften als Beschreibung der empirisch gewonnenen relativen Häufigkeiten herangezogen werden. Mit der Annahme b) bezogen auf die Ereignisse $\left\{\frac{X_n}{n} \notin [p-\varepsilon, p+\varepsilon]\right\}$ rechtfertigen

wir es, in der Praxis $\frac{X_n}{n} \approx P(A)$ zu setzen. (Wir betonen aber noch einmal, daß die Häufigkeitsinterpretation keine zwingende Interpretation ist.)

Die Behandlung des Bernoullischen Gesetzes der großen Zahlen wird in vielen Rahmenlehrplänen für den Stochastikunterricht gefordert. Die Ausführungen in Abschnitt 3 lassen erwarten, daß diese Forderung Lehrer und Schüler vor erhebliche Probleme stellt. Es wird nicht ausreichen, den Satz (dessen Formel für den Schüler ein Monstrum ist) zu nennen und womöglich zu beweisen. Er würde toter Kalkül bleiben. Auch Rechtfertigungsargumente von der Art „Das Gesetz zeigt, daß die Mathematik die Wirklichkeit adäquat beschreibt" werden im Verständnis der Schüler keine Spuren hinterlassen. Wird doch das Verhältnis (die Distanz) zwischen Modell und Wirklichkeit im Mathematikunterricht viel zu wenig thematisiert, als daß die Schüler über ein entsprechendes Problembewußtsein verfügen könnten.

Was also könnte ein Sinn stiftender Kontext für Betrachtungen zum Gesetz der großen Zahlen sein? Wir schlagen vor, das Schätzen einer unbekannten Wahrscheinlichkeit als Rahmen und unmittelbar praktisch relevantes Problem zu wählen. Es soll darum gehen, die unscharfe Beziehung $\frac{X_n}{n} \approx P(A)$ genauer dahingehend zu untersuchen, wie die Güte von Schätzwerten zu beurteilen ist und wie man „gute" Schätzwerte bekommen kann.

Wir stellen uns vor, daß die im folgenden skizzierte Unterrichtssequenz in einem Leistungskurs stattfindet. Die Binomialverteilung sollte bereits besprochen sein, ebenso Eigenschaften der Binomialverteilung für große n. Die dabei auftretenden $k\sigma$-Regeln müssen u. E. nicht zwingend über die Normalverteilung gewonnen werden. Den Grenzwertsatz von Moivre-Laplace wird man im Unterricht ohnehin nicht beweisen. Insofern bleibt die Approximation der Binomialverteilung durch die Normalverteilung auf einem heuristischen, durch Beispiele und entsprechende Bilder abgesicherten Niveau. Das ist durchaus legitim, geht es doch in der Schule allenfalls um die Beschreibung des Sachverhalts und nicht um eine mathematisch exakte Absicherung. Folgt man diesen Gedanken, dann kann man aber auch auf die Normalverteilung ganz verzichten und statt dessen die $k\sigma$-Regeln auf demselben Niveau plausibel machen.

Ziel dieser Unterrichtssequenz soll es sein, Verständnis für die Probleme beim Schätzen einer unbekannten Wahrscheinlichkeit aus Beobachtungsserien zu wecken, den Begriff der Güte einer Schätzgröße zu motivieren und die Schüler zu befähigen, auf einem durch die $k\sigma$-Regeln geprägten quantitativen Niveau folgende Fragen zu beantworten:

1. Wie „gut" ist ein Schätzwert, der auf n Beobachtungen beruht?
2. Wie viele Beobachtungen reichen etwa aus, um eine vorgegebene Güte der Schätzgröße zu garantieren?

Wir meinen, daß diese Fragen von Bedeutung für die Allgemeinbildung eines Abiturienten sind in einer Zeit, in der wir mit statistischen Daten überschüttet werden und gefordert sind, den Gehalt solcher Zahlen kritisch zu werten.

Im Laufe der Unterrichtssequenz wird das Verhalten der Verteilung der Zufallsgröße $\frac{X_n}{n}$ untersucht. Diese recht ausführliche Diskussion soll es den Schülern ermöglichen, ein inhaltliches Verständnis für die Aussage des schwachen Gesetzes der großen Zahlen zu erlangen. Diese Herangehensweise ist bewußt fernab von einer Beschränkung auf einen eleganten Beweis via Tschebyschewsche Ungleichung. Sie gestattet in natürlicher Weise Fortsetzungen in Richtung Konfidenzintervalle oder Testen von Hypothesen über eine unbekannte Wahrscheinlichkeit.

Wir skizzieren nun unsere Vorstellungen über die Behandlung des Themas im Unterricht. Für eine konkrete Unterrichtsvorbereitung müßten natürlich viele Details situationsabhängig ausgearbeitet werden, insbesondere Beispiele und Aufgaben ergänzt werden.

Das Problem

Beispiel: Ausschußanteil

Der Ausschußanteil p einer Massenproduktion von Bauelementen soll geschätzt werden. Man nimmt an, daß die Bauelemente bei der Produktion unabhängig voneinander mit der Wahrscheinlichkeit p defekt anfallen. In einer zufälligen Stichprobe von 500 Bauelementen fand man 27 defekte. Die Zahl $\tilde{p} = \frac{27}{500} = 0{,}054$ dient als Schätzwert für die unbekannte Wahrscheinlichkeit p. Eine andere Stichprobe hätte vermutlich zu einem anderen Schätzwert geführt. Welchen "Wert" hat also die Zahl 0,054? Wie gut ist ein Schätzverfahren, dessen Ergebnisse vom Zufall abhängen?

Die Schüler sollen erkennen: Die relative Häufigkeit $\frac{X_n}{n}$ ist eine Zufallsgröße. Wir nennen sie **Schätzgröße** für die unbekannte Wahrscheinlichkeit $p=P(A)$. Die jeweils beobachteten Werte der Schätzgröße nennen wir **Schätzwerte**. Die Schätzwerte schwanken zufällig! Um die Brauchbarkeit unseres Schätzverfahrens zu beurteilen, untersuchen wir die Verteilung der Schätzgröße $\frac{X_n}{n}$. Unser Ziel ist es, Aussagen über Größenordnung und Wahrscheinlichkeiten der Schwankungen zu gewinnen.

Aufgabe: Ermitteln Sie die (theoretische) Verteilung der relativen Häufigkeit von Wappen bei 5 bzw. 10 Würfen mit einer guten Münze. Berechnen Sie Erwartungswerte und Varianzen! Stellen Sie die Verteilungen als Säulendiagramm dar. Berechnen Sie $P\left(\left|\frac{X_n}{n} - p\right| > 0{,}2\right)$.

Ziel solcher Aufgaben ist es, die Zufälligkeit von $\frac{X_n}{n}$ deutlich werden zu lassen, Kenngrößen und deren Einfluß auf Gestalt der Verteilung zu untersuchen sowie als Vorbereitung auf den allgemeinen Fall, Wahrscheinlichkeiten von konkreten Abweichungen berechnen.

Lösung: $E\left(\frac{X_5}{5}\right) = E\left(\frac{X_{10}}{10}\right) = 0{,}5$, $Var\left(\frac{X_5}{5}\right) = 0{,}05$, $Var\left(\frac{X_{10}}{10}\right) = 0{,}025$,

$$P\left(\left|\frac{X_5}{5} - 0{,}5\right| > 0{,}2\right) = 0{,}3752, \quad P\left(\left|\frac{X_{10}}{10} - 0{,}5\right| > 0{,}2\right) = 0{,}1094.$$

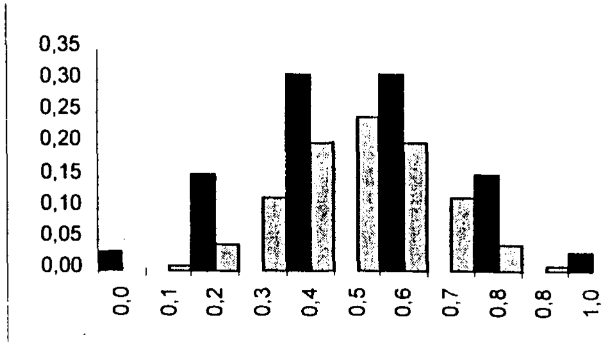


Abb. 1.

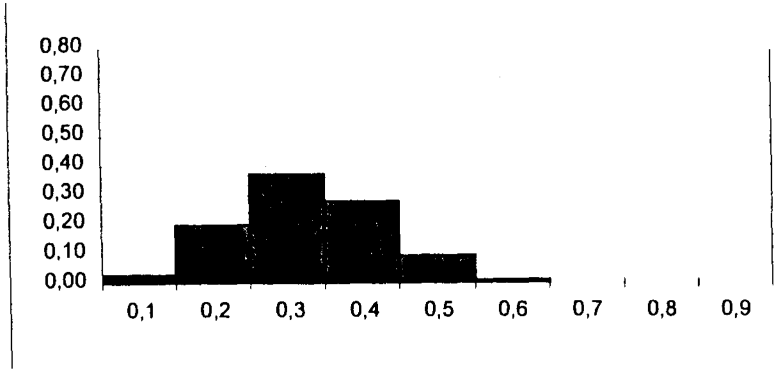
In der Praxis treten beim Schätzen von p große n auf. Die Berechnung wird aufwendig. Wir müssen uns näher mit der Verteilung von $\frac{X_n}{n}$ beschäftigen.

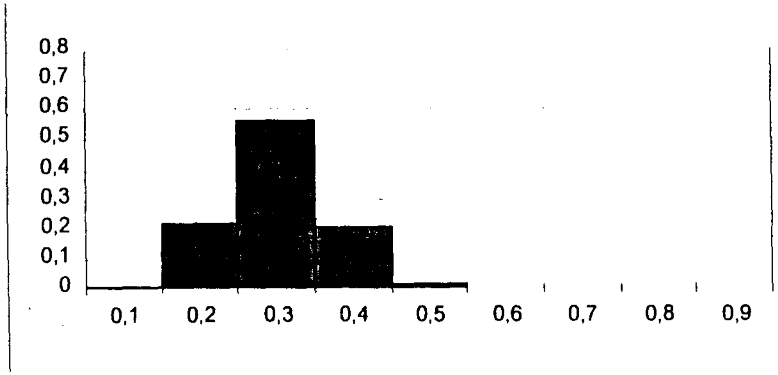
Die Verteilung der Schätzgröße $\frac{X_n}{n}$

Die Anzahl X_n der Erfolge in einer Bernoulli-Kette ist binomialverteilt: $P(X_n = k) = p_k = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0,1,\dots,n$. Daraus gewinnen wir die Verteilung von $\frac{X_n}{n}$. Da die Ereignisse $\{X_n = k\}$ und $\left\{\frac{X_n}{n} = \frac{k}{n}\right\}$ identisch sind, nimmt die Schätzgröße $\frac{X_n}{n}$ ihre Werte $\frac{0}{n}, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}$ mit denselben Wahrscheinlichkeiten

$P(\frac{X_n}{n}=\frac{k}{n})=\binom{n}{k}p^k(1-p)^{n-k}$, $k=0,1,\dots,n$, an. Wir untersuchen die Gestalt dieser Verteilung: Dazu setzen wir p zunächst als bekannt voraus. Die Histogramme in Abb. 2 zeigen die Verteilung von $\frac{X_n}{n}$ für $p=0{,}3$ und $n=20$, 50 und 100. Die möglichen Werte der Zufallsgröße $\frac{X_n}{n}$ wurden in Klassen eingeteilt: [0; 0,05), [0,05; 0,15), [0,15; 0,25), ... , [0,85; 0,95), [0,95; 1,0]

Die Klassenmitten sind 0,025, 0,1, 0,2, ..., 0,9, 0,975. In der graphischen Darstellung wurden die erste und die letzte Klasse, die hier keine Werte enthält, weggelassen. Die Abszissenachse wurde mit den Klassenmitten beschriftet. Die Säulenhöhen sind gleich den Wahrscheinlichkeiten der Klassen.





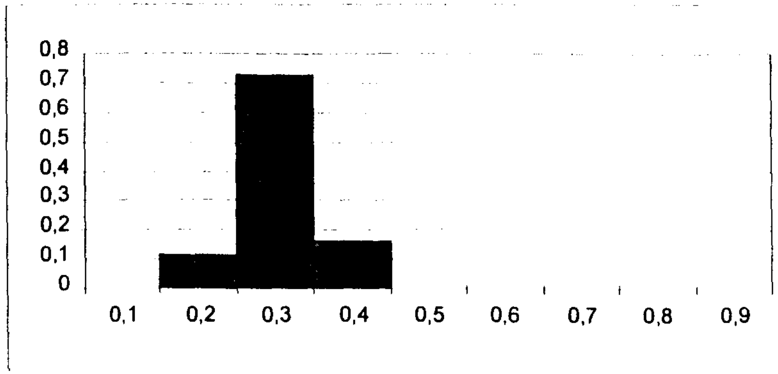


Abb. 2

Die Notwendigkeit einer Klasseneinteilung zu diskutieren, heißt, wesentliche Eigenschaften der Verteilung zu erkennen. Für große n, und um diese geht es uns ja, werden die Einzelwahrscheinlichkeiten immer kleiner, sinnvoll werden also Intervallwahrscheinlichkeiten und diese bedeuten eine Klasseneinteilung. Die Verteilung "zieht sich" mit wachsendem n immer mehr um den Wert 0,3 "zusammen". Große Abweichungen (z. B. um ±0,2) der Werte von $\frac{X_n}{n}$ von der Wahrscheinlichkeit p=0,3 sind zwar immer möglich, die Chancen dafür werden aber mit wachsendem n immer kleiner. Jetzt wird auch klar, warum wir als Klasseneinteilung nicht die vielleicht naheliegende Einteilung [0; 0,1], (0,1; 0,2], ... in gleichbreite Klassen gewählt haben. Dann wäre die Erfolgswahrscheinlichkeit 0,3 als Rand einer Klasse erschienen und der Effekt des Zusammenziehens um den Wert 0,3 wäre "verwischt" worden.

Kenngrößen der Verteilung von $\frac{X_n}{n}$ - Erwartungswert und Varianz

Satz 1: Wenn das Ereignis A die Wahrscheinlichkeit p besitzt, dann gilt für die relative Häufigkeit $\frac{X_n}{n}$ von A in einer Bernoulli-Kette der Länge n

$$E\left(\frac{X_n}{n}\right)=p,\quad Var\left(\frac{X_n}{n}\right)=\frac{p(1-p)}{n} \text{ und Standardabweichung } \sigma=\frac{\sqrt{p(1-p)}}{\sqrt{n}}.$$

Beweis: $E\left(\frac{X_n}{n}\right)=0\cdot p_0+\frac{1}{n}p_1+\frac{2}{n}p_2+\dots+\frac{n}{n}p_n$

$$=\frac{1}{n}(0\cdot p_0+1\cdot p_1+2\cdot p_2+\dots+n\cdot p_n)=\frac{1}{n}E(X_n)=\frac{1}{n}np=p.$$

Aufgabe: Beweisen Sie die Formel für $Var\left(\frac{X_n}{n}\right)$.

Erste Aussagen zur Qualität des Schätzverfahrens

Den Erwartungswert und die Varianz deuten wir als stabile Werte der entsprechenden empirischen Kenngrößen "arithmetisches Mittel" und "mittlere quadratische Abweichung". Diese Interpretation führt uns zu folgenden Schlußfolgerungen:

Q1. Das arithmetische Mittel sehr vieler Schätzwerte liegt in der Nähe von p.

Q2. Die mittlere quadratische Abweichung sehr vieler Schätzwerte aus Stichproben vom Umfang n konvergiert gegen 0 für n gegen unendlich.

Es gibt Probleme mit der Deutung: Wir bestimmen in der Regel nicht viele Schätzwerte, sondern nur einen. Die Aussagen Q1 und Q2 "ziehen" nicht. Deshalb bemühen wir uns um Wahrscheinlichkeitsaussagen über mögliche Abweichungen.

Uns interessiert der Unterschied zwischen $\frac{X_n}{n}$ und p. Warum:

- Ein bekanntes p liefert eine Vorhersage für $\frac{X_n}{n}$. Wie gut ist sie? Wie groß ist der Unterschied?
- Ein unbekanntes p wird durch $\frac{X_n}{n}$ geschätzt. Wie gut ist die Schätzung? Wie groß ist der Unterschied?

Der zufällige Unterschied zwischen $\frac{X_n}{n}$ und p

Der Unterschied ist eine Zufallsgröße. Wir werden uns mit Wahrscheinlichkeitsaussagen begnügen müssen. Wir greifen auf die $k\sigma$-Regeln für die Binomialverteilung bei genügend großem n zurück: Wenn $n \cdot p \cdot (1-p) > 9$ (Faustregel)[4] erfüllt ist, dann gilt näherungsweise

$$P(np - \sqrt{n \cdot p \cdot (1-p)} \le X_n \le np + \sqrt{n \cdot p \cdot (1-p)}\,) \approx 0{,}683,$$

$$P(np - 2\sqrt{n \cdot p \cdot (1-p)} \le X_n \le np + 2\sqrt{n \cdot p \cdot (1-p)}\,) \approx 0{,}954,$$

$$P(np - 3\sqrt{n \cdot p \cdot (1-p)} \le X_n \le np + 3\sqrt{n \cdot p \cdot (1-p)}\,) \approx 0{,}997.$$

Die Ungleichung

[4] Es gilt $p\cdot(1-p) \le 0{,}25$. Je weiter p von 0,5 entfernt ist, d. h. je asymmetrischer die Binomialverteilung ist, desto größer muß n sein, damit die Faustregel erfüllt ist. Bei $n \le 36$ gilt die Faustregel für überhaupt kein p.

$$np - \sqrt{n \cdot p \cdot (1-p)} \le X_n \le np + \sqrt{n \cdot p \cdot (1-p)}$$

geht bei Division durch n und Subtraktion von p in die äquivalente Ungleichung

$$-\frac{\sqrt{p(1-p)}}{\sqrt{n}} \le \frac{X_n}{n} - p \le \frac{\sqrt{p(1-p)}}{\sqrt{n}}$$

über. Diese Ungleichung beschreibt dasselbe Ereignis wie die ursprüngliche Ungleichung und hat folglich auch dieselbe Wahrscheinlichkeit:

$$P(-\frac{\sqrt{p(1-p)}}{\sqrt{n}} \le \frac{X_n}{n} - p \le \frac{\sqrt{p(1-p)}}{\sqrt{n}}\,) \approx 0{,}683.$$

Entsprechend werden die anderen Ungleichungen umgeformt.

Es ist eine schwierige Symbolik! Man kann sie vereinfachen und anschaulich machen durch Arbeiten mit dem Umgebungsbegriff. Es ist p der Erwartungswert und $\sigma = \frac{\sqrt{p(1-p)}}{\sqrt{n}}$ ist die Standardabweichung von $\frac{X_n}{n}$. Wir nennen das Intervall $[p - \sigma, p + \sigma]$ die σ-Umgebung von p. Entsprechend sind $[p - 2\sigma, p + 2\sigma]$ und $[p - 3\sigma, p + 3\sigma]$ die 2σ-Umgebung bzw. die 3σ-Umgebung von p:

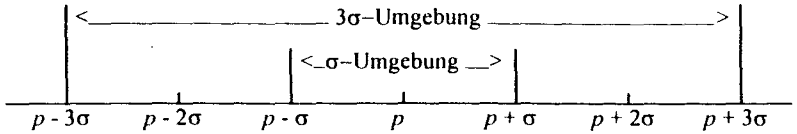


Abb. 3

Die σ-Umgebungen sind fest, $\frac{X_n}{n}$ ist zufällig. Das Ereignis "$-2\sigma \le \frac{X_n}{n} - p \le 2\sigma$" bedeutet, daß $\frac{X_n}{n}$ in der 2σ-Umgebung von p liegt, denn der Abstand von $\frac{X_n}{n}$ zu p beträgt höchstens 2σ. Folglich

$$P\left(\frac{X_n}{n} \in [p - 2\sigma, p + 2\sigma]\right) = P\left(-2\sigma \le \frac{X_n}{n} - p \le 2\sigma\right) \approx 0{,}954.$$

Analog werden die anderen Aussagen umgeformt und wir erhalten die

Merkregel 1
Wenn $n \cdot p \cdot (1-p) > 9$ (Faustregel) erfüllt ist, dann gilt für die zufällige Abweichung der relativen Häufigkeit $\frac{X_n}{n}$ von der Wahrscheinlichkeit p näherungsweise

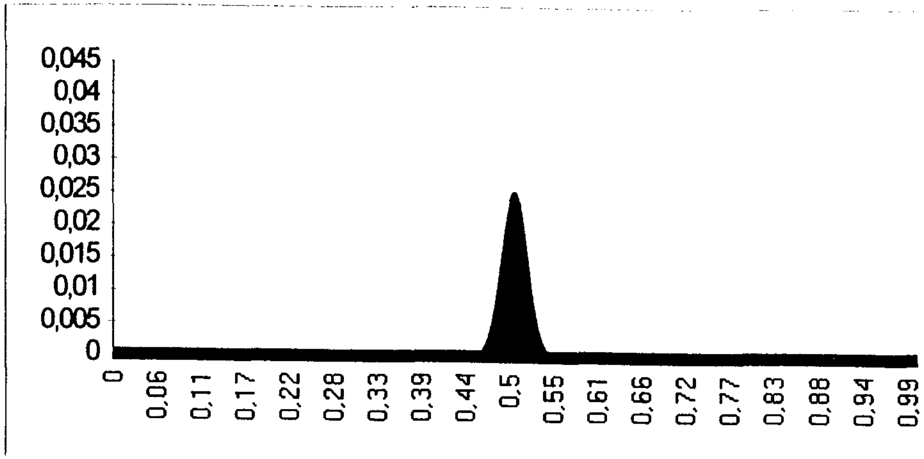
$$P\left(\frac{X_n}{n} \in [p-\sigma, p+\sigma]\right) = P\left(-\sigma \le \frac{X_n}{n} - p \le \sigma]\right) = P\left(\left|\frac{X_n}{n} - p\right| \le \sigma\right) \approx 0{,}683,$$

$$P\left(\frac{X_n}{n} \in [p-2\sigma, p+2\sigma]\right) = P\left(-2\sigma \le \frac{X_n}{n} - p \le 2\sigma]\right) = P\left(\left|\frac{X_n}{n} - p\right| \le 2\sigma\right) \approx 0{,}954,$$

$$P\left(\frac{X_n}{n} \in [p-3\sigma, p+3\sigma]\right) = P\left(-3\sigma \le \frac{X_n}{n} - p \le 3\sigma]\right) = P\left(\left|\frac{X_n}{n} - p\right| \le 3\sigma\right) \approx 0{,}997.$$

Die Aussagen in der Merkregel 1 sind Spezialfälle des sogenannten $1/\sqrt{n}$ -Gesetzes. Dieser Name kommt daher, daß in die Standardabweichung $\sigma = \frac{\sqrt{p(1-p)}}{\sqrt{n}}$ die Größenordnung $1/\sqrt{n}$ besitzt. Das bedeutet: die Schwankungen der relativen Häufigkeit lassen sich dem Betrage nach durch eine Zahl der Gestalt $c \cdot 1/\sqrt{n}$ mit einer gewissen Sicherheit nach oben abschätzen. Die Konstante c hängt nicht von n, wohl aber von er gewünschten Sicherheit ab.

Beispiel: Wir veranschaulichen die Aussage der Merkregel 1 an den Beispielen n=1000, p=0,5 und n=1000, p=0,1. Zunächst schauen wir auf die gesamte Verteilung von $\frac{X_n}{n}$ für p=0,5 bzw. p=0,1. (Abb. 4)



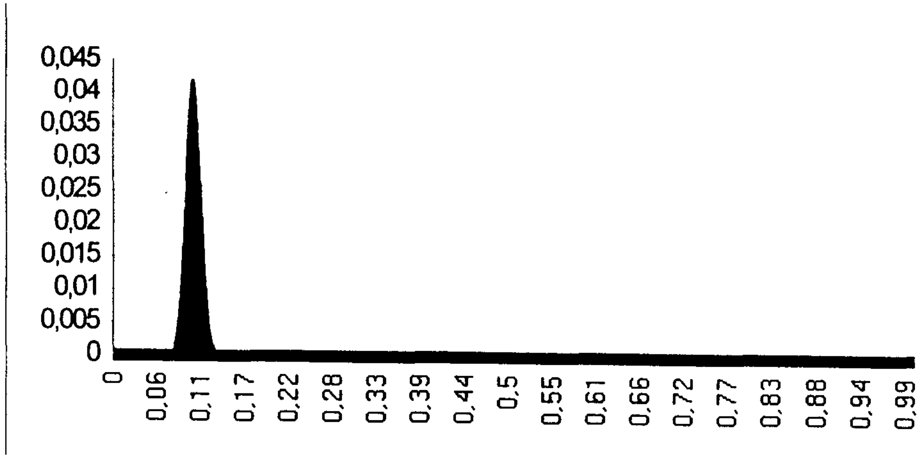


Abb. 4

Die Streuung um p ist bei 0,5 größer als bei 0,1. Die Glocke ist bei p=0,5 breiter und niedriger. Die entsprechenden σ-Umgebungen sind deshalb bei p=0,5 größer als bei p=0,1

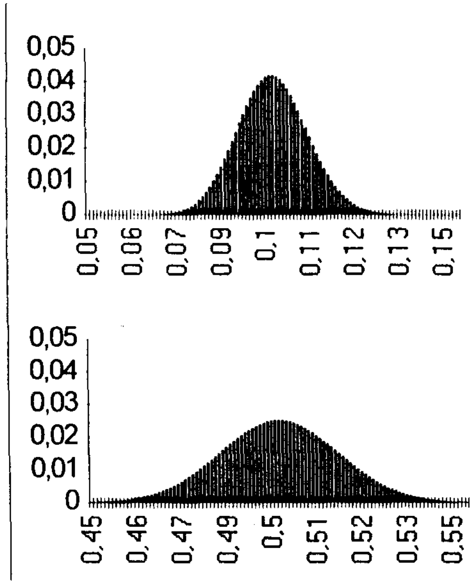


Abb. 5

Die Abb. 5 zeigt die "interessanten" Ausschnitte der Verteilungen von $\frac{X_{1000}}{1000}$. Nach der Merkregel fallen die Werte von $\frac{X_{1000}}{1000}$ praktisch sicher, nämlich mit einer Wahrscheinlichkeit von etwa 99,7%, in die 3σ-Umgebung von p. Dies sind in unserem Beispiel die Intervalle [0,45, 0,55] bzw. [0,07, 0,13]. Beobachtungswerte außerhalb dieser Intervalle kommen nur mit einer sehr kleinen Wahrscheinlichkeit vor. Die 2σ-Umgebung und die σ-Umgebung sind kleiner und haben deshalb eine geringere Sicherheit zu beanspruchen.

Aufgabe: Innerhalb welcher Grenzen liegt mit etwa 95%iger Wahrscheinlichkeit die relative Häufigkeit der Sechs bei 100, 1000 bzw. 10000 Würfen mit einem guten Würfel? Welche Schranken können Sie für die relative Häufigkeit angeben, wenn Sie nur eine Sicherheit von etwa 68% beanspruchen?

Lösung:

n	σ	σ-Umgebung	2σ-Umgebung
100	0,04	[0,12; 0,21]	[0,08; 0,25]
1000	0,01	[0,15; 0,18]	[0,14; 0,19]
10 000	0,004	[0,16; 0,17]	[0,15; 0,18]

Die Güte der Schätzgröße $\frac{X_n}{n}$ und Konsequenzen für den Stichprobenumfang

Das Problem: Der (maximale) Fehler der Schätzgröße wird für die drei Sicherheitsniveaus durch die Werte $\sigma = \frac{\sqrt{p(1-p)}}{\sqrt{n}}$, 2σ bzw. 3σ erfaßt. Die Standardabweichung σ stellt also ein Maß für die zufällige Abweichung vom Erwartungswert p dar. Sie hängt ihrerseits von p ab, wie die Abb. 4 und 5 deutlich illustrieren. Da p unbekannt ist - wir wollen es ja schätzen - nützen uns die Fehlerschranken σ, 2σ und 3σ in dieser Form nichts. Wir müssen p eliminieren. Nun gilt aber $\frac{\sqrt{p(1-p)}}{\sqrt{n}} < \frac{1}{2\sqrt{n}}$.

Wenn wir in der obigen Faustregel den Ausdruck $\frac{\sqrt{p(1-p)}}{\sqrt{n}}$ durch den größeren Ausdruck $\frac{1}{2\sqrt{n}}$ ersetzen, dann werden die Wahrscheinlichkeiten der zugehörigen Ereignisse (Intervalle) höchstens größer. Die Abbildung verdeutlicht den Sachverhalt am Beispiel der 2σ-Umgebung:

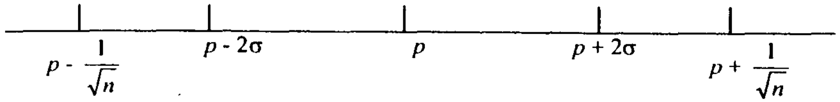


Abb. 6

Somit können wir näherungsweise die Qualität der Schätzgröße beurteilen, wenn wir stillschweigend annehmen, daß die Faustregel $n \cdot p \cdot (1-p) > 9$ erfüllt ist. Dies nachzuprüfen ist nicht möglich, da das unbekannte p eingeht. Meist hat man allerdings eine Vorstellung über die Größenordnung von p und weiß demnach, welcher Stichprobenumfang n die Gültigkeit der Faustregel sicherstellt.

Merkregel 2:

Für die zufällige Abweichung $\frac{X_n}{n} - p$ der relativen Häufigkeit von der Wahrscheinlichkeit p gilt bei genügend großem n näherungsweise

$$P\left(-\frac{0,5}{\sqrt{n}} \le \frac{X_n}{n} - p \le \frac{0,5}{\sqrt{n}}\right) \ge 0,683, \quad P\left(-\frac{1}{\sqrt{n}} \le \frac{X_n}{n} - p \le \frac{1}{\sqrt{n}}\right) \ge 0,954,$$

$$P\left(-\frac{1,5}{\sqrt{n}} \le \frac{X_n}{n} - p \le \frac{1,5}{\sqrt{n}}\right) \ge 0,997.$$

Beispiel: Man möchte den Anteil der A-Wähler in einer großen Stadt auf 3/100 genau schätzen. Die Schätzgröße soll diese Genauigkeit mit einer Sicherheit von mindestens 99,7 % aufweisen. Um zu wissen, welcher Stichprobenumfang dafür ausreicht, benutzen wir die dritte Beziehung der Merkregel 2: Es soll gelten $1{,}5 \cdot \frac{1}{\sqrt{n}} \leq 0{,}03$. Wir formen äquivalent um und erhalten $n \approx 2500$. Ungefähr 2500 auf gut Glück ausgewählte Wähler sichern die gewünschte Genauigkeit bei der vorgegebenen Sicherheit.

Aufgabe: Eine Zeitung berichtet, daß mehr als die Hälfte der Einwohner des Landes Brandenburg regelmäßig den Sender "Otto" hört. Der Berichterstatter stützt sich auf das Ergebnis einer zufälligen Stichprobe, in der 51,4% der 2250 Befragten angaben, regelmäßige "Otto-Hörer" zu sein. Diskutieren Sie die Schlußfolgerung des Reporters.

Lösung: Es ist $\frac{1}{\sqrt{n}} \approx 0{,}021$ und $1{,}5 \cdot \frac{1}{\sqrt{n}} \approx 0{,}032$. Man sollte also durchaus damit rechnen, daß es weniger als die Hälfte sind.

Die Güte der Schätzgröße $\frac{X_n}{n}$ wird durch die Fehlerschranke ε (auch Genauigkeit genannt) und die Sicherheit $1\text{-}\alpha$ beschrieben. Dabei sind ε und α vorgegebene, in der Regel kleine Zahlen aus dem Intervall (0; 1). Wir sagen, daß die Schätzgröße die Genauigkeit ε mit der Sicherheit $1\text{-}\alpha$ besitzt, wenn sich $\frac{X_n}{n}$ von p mit einer Wahrscheinlichkeit von mindestens $1\text{-}\alpha$ um höchstens den Betrag ε unterscheidet, wenn also gilt $P(|\frac{X_n}{n} - p| \leq \varepsilon) \geq 1 - \alpha$. Wenn wir das Betragszeichen "auflösen", dann lautet diese Beziehung

$$P\left(-\varepsilon \leq \frac{X_n}{n} - p \leq \varepsilon\right) \geq 1 - \alpha\,.$$

Mit anderen Worten: $\frac{X_n}{n}$ besitzt die Genauigkeit ε mit der Sicherheit $1\text{-}\alpha$, wenn $\frac{X_n}{n}$ mit einer Wahrscheinlichkeit von mindestens $1 - \alpha$ in die ε-Umgebung von p fällt. Anhand schematischer Darstellungen der Verteilung von $\frac{X_n}{n}$ kann man die Zusammenhänge zwischen der Genauigkeit ε, der Sicherheit $1\text{-}\alpha$ und dem Stichprobenumfang n verdeutlichen: Die Sicherheit $1\text{-}\alpha$, die zum Intervall $[p\text{-}\varepsilon, p\text{+}\varepsilon]$ gehört, wird durch den Inhalt der markierten Fläche dargestellt. Der Gesamtinhalt unter der Kurve ist 1:

1. Fall (Abb. 7): ε festgehalten: Je größer n, desto größer $1\text{-}\alpha$. Anschaulich: Die Verteilung von $\frac{X_n}{n}$ zieht sich mit wachsendem n um p zusammen. In eine feste ε-Umgebung von p fällt mehr "Masse".

Die anderen beiden Fälle möge sich der Leser selbständig veranschaulichen. Wir formulieren lediglich die Erkenntnisse:
2. Fall: $1\text{-}\alpha$ festgehalten: Je größer n, desto kleiner ε. Anschaulich: Um eine vorgegebene Masse "einzufangen", reicht eine kleinere ε-Umgebung.
3. Fall: n festgehalten: Je kleiner ε, desto kleiner $1\text{-}\alpha$. Anschaulich: Je kleiner die ε-Umgebung, desto kleiner die Masse, die hineinfällt.

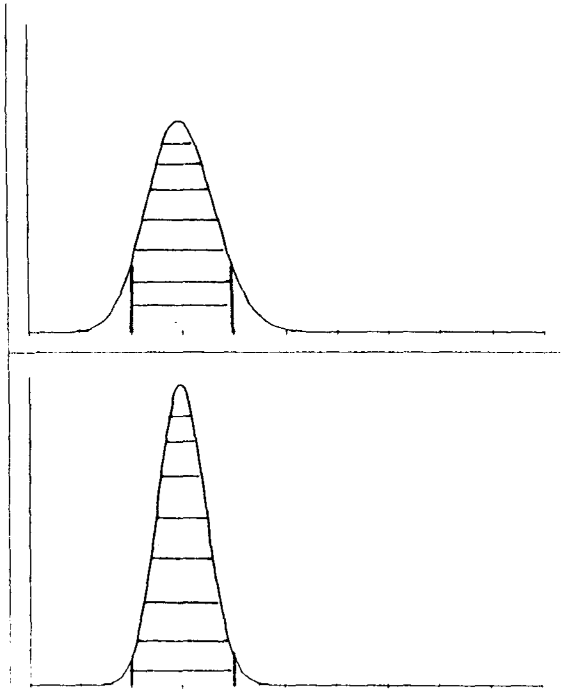


Abb. 7

Für die praktische Bestimmung von n bei gegebenem ε und α bewährt sich folgende Merkregel, bei deren Anwendung die Faustregel $np(1\text{-}p)>9$ erfüllt sein soll.

Merkregel 3:

Wenn die Schätzgröße $\frac{X_n}{n}$ die Genauigkeit ε mit einer Sicherheit

$$\left.\begin{array}{l}\text{von etwa } 86{,}3\% \\ \text{von etwa } 95{,}4\% \\ \text{von etwa } 99{,}7\%\end{array}\right\} \begin{array}{c}\text{besitzen soll,} \\ \text{dann wähle} \\ \text{das kleinste } n \text{ mit}\end{array} \left\{\begin{array}{l} n > \left(\frac{0{,}5}{\varepsilon}\right)^2 \\ n > \frac{1}{\varepsilon^2} \\ n > \left(\frac{1{,}5}{\varepsilon}\right)^2 \end{array}\right.$$

als Stichprobenumfang.

Aufgabe: Machen Sie sich klar: Wenn der Stichprobenumfang gänzlich unbekannt ist, dann läßt sich nichts über die Qualität der Schätzung sagen! Vorbildlich handelt also ein große Tageszeitung, wenn sie ihren Lesern nicht nur mitteilt: "86% der Berliner sind nur wenig vom Wahlkampf angetan", sondern auch informiert, daß durch eine Zufallsauswahl 1330 Wahlberechtigte befragt wurden.

Aufgabe: Begründen Sie mit Hilfe der Merkregel 3: Bei einer gegebenen Sicherheit wird die Fehlerschranke halbiert, wenn man den Stichprobenumfang vervierfacht.

Anhang

Die Stirlingsche Formel und die größte Einzelwahrscheinlichkeit der Binomialverteilung

Die Stirlingsche Formel gibt eine Näherung für $n!$ für große n. Sei nämlich $a_n = n^n e^{-n}\sqrt{2\pi n}$. Dann gilt $\lim\limits_{n\to\infty}\frac{n!}{a_n} = 1$. Wir schreiben dafür $n! \sim a_n$.

Für eine Bernoulli-Kette mit $p = \frac{1}{2}$ und $n = 2k$ ist es am wahrscheinlichsten, daß genau k Erfolge eintreten. Wir geben eine Näherung für diese Wahrscheinlichkeit für große n mit Hilfe der Stirlingschen Formel. Es ist

$P(X_n = k) = \binom{2k}{k}\left(\frac{1}{2}\right)^{2k} = \frac{(2k)!}{k!k!}\frac{1}{2^{2k}}$. Nähern wir die auftretenden Fakultäten mit der Stirlingschen Formel, so erhalten wir $P(X_n = k) \sim \frac{(2k)^{2k}e^{-2k}\sqrt{2\pi 2k}}{\left(k^k e^{-k}\sqrt{2\pi k}\right)^2}\frac{1}{2^{2k}}$. Nach elementaren Umformungen ergibt sich $P(X_n = k) \sim \frac{1}{\sqrt{\pi k}}$. Für die wahrscheinlichste Anzahl k_0 von Erfolgen in einer Bernoulli-Kette mit beliebigem p gilt $np-(1-p) \le k_0 \le np+p$. Es ist also $k_0 \sim np$. Um im Bereich der natürlichen Zahlen zu bleiben, ersetzen wir bei Bedarf np durch den ganzen Anteil $[np]$, und es gilt auch $k_0 \sim [np]$. Wir erhalten nun

$$P(X_n = [np]) = \binom{n}{[np]} p^{[np]}(1-p)^{n-[np]} = \frac{n!}{(n-[np])!([np])!} p^{[np]}(1-p)^{n-[np]}$$

$$\sim \frac{n^n e^{-n}\sqrt{2\pi n}}{(n(1-p))^{n(1-p)}e^{-n(1-p)}\sqrt{2\pi n(1-p)}(np)^{np}e^{-np}\sqrt{2\pi np}} p^{np}(1-p)^{n(1-p)}$$

$$= \frac{1}{\sqrt{2\pi p(1-p)n}}.$$

Literatur:

Borovcnik, M.: Stochastik im Wechselspiel von Intuitionen und Mathematik. BI Wissenschaftsverlag, Mannheim, 1992.

Dinges, H./Rost, H.: Prinzipien der Stochastik. B.G. Teubner, Stuttgart, 1982.

Freudenthal, H.: "The Empirical Law of Large Numbers" or "The Stability of Frequencies". In: Educational Studies in Mathematics 4(1972), 484-490.

Reichel, H. Ch. (Hrsg.): Wahrscheinlichkeitsrechnung und Statistik, Hölder-Pichler-Tempsky, Wien, 1989.

Riemer, W.: „Das „Eins durch Wurzel aus n"-Gesetz. Einführung in statistisches Denken auf der Sekundarstufe I". In: Stochastik in der Schule 11(1991), Heft 3, 23-36.

Adressen der Autoren:

Ödön Vancsó
Eötvös-Lorand-Universität
Mathematisch-fachdidaktische Gruppe
Rákóczi u. 5
H 1088 Budapest

Elke Warmuth
Humboldt-Universität zu Berlin
Institut für Mathematik
Unter den Linden 6
10099 Berlin