

Augensummen bis zwölf - eine Wahrscheinlichkeitstheoretische Meditation  
von Hans G. Schönwald, Siegen

Ausgehend von der Augensumme beim Werfen mit zwei Würfeln werden wir durch naheliegende Überlegungen zu einigen weiteren Wahrscheinlichkeitsexperimenten geführt und fragen schließlich nach einem Überblick über alle Mehrfachexperimente mit Augensummen bis zwölf. Durch diese Meditation sollen wesentliche, wahrscheinlichkeitstheoretische Begriffe und Techniken angesprochen und benutzt und damit vertiefend wiederholt werden. Darüber hinaus kann das Erforschen eines mathematischen Teilgebietes beispielhaft vorerlebt werden.

Wir würfeln mit zwei Würfeln (Hexaeder), einem roten und einem grünen. Die Verschiedenfarbigkeit der beiden Würfel ermöglicht eine bequeme Formulierung und eine anschauliche Vorstellung der wahrscheinlichkeitstheoretischen Situation. Sowohl der rote Würfel als auch der grüne werden mit je  $\frac{1}{6}$  Wahrscheinlichkeit jede der sechs Augenzahlen zeigen. Da die Würfel nichts voneinander wissen, unabhängig voneinander fallen, wird, wenn wir sehr oft würfeln, jedes mögliche rot/grüne Augenzahlpaar gleich oft fallen. Da es 36 Möglichkeiten gibt, steht jeder die Wahrscheinlichkeit  $\frac{1}{36}$  zu. Wir notieren die gleichwahrscheinlichen Möglichkeiten, welche man auch Laplace'sche Elementarereignisse nennt, und die zugehörigen Augensummen in einem quadratischen Schema:

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

Demnach gilt für die Verteilung der Augensumme:

Augensumme		3	4	5	6	7	8	9	10	11	12
W'keit		$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Stochastik in der Schule, Heft 2, 5, 6 (1985)

Wir können diese Verteilung als Dachverteilung bezeichnen. Dieselbe Verteilung wird sich ergeben, wenn wir im Dämmerlicht spielen, wo alle Würfel grau und nicht mehr unterscheidbar sind.

Manche Menschen vermuten zunächst eine Gleichverteilung anstelle der Dachverteilung. Eine solche entsteht, wenn wir mit einem einzigen regelmäßigen Zwölfflach, einem Pentagon-Dodekaeder würfeln. Dessen Augenzahlen fallen gleichverteilt, jede mit  $\frac{1}{12} = \frac{3}{36}$  Wahrscheinlichkeit. Übrigens fehlt bei der oben betrachteten Augensumme die Augenzahl 1, hier ist sie möglich. Die maximal mögliche Augenzahl ist jeweils zwölf. Dieser Vergleich mit dem Zwölfflachwürfel ist m.E. naheliegend. Denn es ist eine Anwendung eines üblichen Heurismus, d.h. Findeverfahrens, nämlich daß man eine mathematische Situation ein wenig variiert.

Ähnlich naheliegend ist nun, andere Mehrfach-Experimente mit Maximalaugensummenzahl 12 zu suchen, den Zerlegungsmöglichkeiten von 12 entsprechend. Es gilt  $12 = 1 \cdot 12 = 2 \cdot 6 = 3 \cdot 4 = 4 \cdot 3 = 6 \cdot 2 = 12 \cdot 1$ . Die bisher betrachteten Experimente entsprechen den ersten beiden Zerlegungsmöglichkeiten. Läßt sich zu jeder dieser Zerlegungsmöglichkeiten ein Experiment konstruieren? Wie ist es verteilt? Welche Eigenschaften besitzen die Verteilungen? Und worin stimmen sie überein bzw. unterscheiden sie sich? Und können wir in der obigen Reihenfolge der Zerlegungsmöglichkeiten Entwicklungslinien der Eigenschaften finden? Diese Fragen sollen unser weiteres Vorgehen bestimmen.

Wenn wir in drei Würfeln maximal 12 Augen sammeln können wollen, müssen wir ein regelmäßiges Vierflach benutzen. Es gibt ein solches, man nennt es auch Tetraeder. Man muß allerdings kräftig werfen, weil es schlecht rollt. Wie können wir die Verteilung errechnen? Es gibt  $4^3 = 64$  Elementarereignisse. Wir können alle Dreifach-Ziffern-Kombinationen hinschreiben. Das dauert nur zwei Minuten, und die Schüler werden daran erinnert, regelmäßige Zahlenfolgen entsprechend schematisch darzustellen. Man schreibt mit bunter Kreide die Augensummen dazu und zählt aus. Dies ist ein leicht verständliches Vorgehen. Man erhält:

Augensumme	3	4	5	6	7	8	9	10	11	12
W'keit	$\frac{1}{64}$	$\frac{3}{64}$	$\frac{6}{64}$	$\frac{10}{64}$	$\frac{12}{64}$	$\frac{12}{64}$	$\frac{10}{64}$	$\frac{6}{64}$	$\frac{3}{64}$	$\frac{1}{64}$

Etwas pfiffiger ist ein nach der Mehrfachheit voranschreitendes Vorgehen. Sich vom Einfachen schrittweise zum Mehrfachen voranzutasten, ist ein anderer mathematischer Heurismus, ein Trick, der manches Mal hilft. Beim Werfen mit zwei Tetraedern würde man entsprechend unseren allerersten Überlegungen

Augensumme	2	3	4	5	6	7	8
W'keit	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

als Verteilung erhalten. Nun kann jede dieser Möglichkeiten mit der Augenzahl 1,2,3 oder 4 beim dritten Würfel zusammentreffen, und zwar jeweils in einem Viertel der Fälle. Das folgende Schema stellt uns dies explizit vor Augen.

		$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$				
W'keit (2-fach)		$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$				
Augensumme		2	3	4	5	6	7	8	9	10	11	12
1 beim dritten W.		$\frac{1}{64}$	$\frac{2}{64}$	$\frac{3}{64}$	$\frac{4}{64}$	$\frac{3}{64}$	$\frac{2}{64}$	$\frac{1}{64}$				
2 "			$\frac{1}{64}$	$\frac{2}{64}$	$\frac{3}{64}$	$\frac{4}{64}$	$\frac{3}{64}$	$\frac{2}{64}$	$\frac{1}{64}$			
3 "				$\frac{1}{64}$	$\frac{2}{64}$	$\frac{3}{64}$	$\frac{4}{64}$	$\frac{3}{64}$	$\frac{2}{64}$	$\frac{1}{64}$		
4 "					$\frac{1}{64}$	$\frac{2}{64}$	$\frac{3}{64}$	$\frac{4}{64}$	$\frac{3}{64}$	$\frac{2}{64}$	$\frac{1}{64}$	
W'keit (3-fach)			$\frac{1}{64}$	$\frac{3}{64}$	$\frac{6}{64}$	$\frac{10}{64}$	$\frac{12}{64}$	$\frac{12}{64}$	$\frac{10}{64}$	$\frac{6}{64}$	$\frac{3}{64}$	$\frac{1}{64}$

Als nächstes möchten wir viermal mit einem Dreiflach werfen. Leider gibt es keinen solchen. Wir können uns aber eine andere Realisierung eines gleichverteilten Experiments mit drei Ausgängen erfinden. Wir können bei einem Hexaeder gegenüberliegende Seiten gleich beschriftet bzw. die Augenzahlen modulo 3 benutzen, oder bei einem Dodekaeder modulo 3. Überhaupt, jedes 3m-flach würde, modulo 3 benutzt, uns hier helfen; aber es gibt ja nur diese beiden regelmäßigen 3m-flache. Auf entsprechende Art könnten wir doch alle unsere Vorhaben verwirklichen, die sich ja alle als n-fach zusammengesetztes Werfen mit bis zu  $\frac{12}{n}$  Augen verstehen lassen. Wir werfen das Dodekaeder n mal und nehmen die Augenzahl mod  $\frac{12}{n}$ . Das könnte interessant sein. Wir wollen es später weiterbedenken. Zunächst suchen wir die Verteilung für n = 4. Wir können sie wiederum auf zwei Arten ermitteln. Leicht verständlich ist, ein Schema der Augenzahlenkom-

binations aufzuschreiben bzw. das vorige abzuändern. Es müssen  $3^4 = 81$  Kombinationen berücksichtigt werden. Pfiffiger ist, drei Augenzahlen viermal miteinander zu kombinieren. Dies läßt sich entsprechend dem obigen Vorgehen durchführen.

Oder wir gehen der Zerlegung  $81 = (3^2)^2$  gemäß vor und notieren in nochmals abgekürzter, aber doch selbstredender Schreibweise:

Augenzahl bzw. Summe	1	2	3	4	5	6	7	8	9	10	11	12
1-fach (in 3-tel)	1	1	1									
		1	1	1								
			1	1	1							
				1	1	1						
2-fach (in 9-tel)	1	2	3	2	1							
mal 1			1	2	3	2	1					
mal 2				2	4	6	4	2				
mal 3					3	6	9	6	3			
mal 2						2	4	6	4	2		
mal 1							1	2	3	2	1	
4-fach (in 81-tel)			1	4	10	16	19	16	10	4	1	

Nun werfen wir sechs Groschen, die Vorderseite bedeutet "1", die Rückseite "2". Wir wissen, daß die Summengröße binomialverteilt ist. Es gibt  $2^6 = 64$  Laplace-Ereignisse. Die Augensumme ist gleich 6 plus Anzahl der 2en.

Augensumme	6	7	8	9	10	11	12
W'keit	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

Es bleibt, zwölfmal mit einem Einflach zu werfen. Nun gibt es zwar keinen Einflach, aber man weiß, daß er immer auf seine eine Seite fallen würde. Es gibt somit auch nur  $1^{12} = 1$  Laplace-Ereignis und die Verteilung ist:

Augensumme	12
W'keit	1

Man spricht von einem sicheren Ereignis.

Nachdem wir alle sechs Zufallsexperimente untersucht haben, wollen wir unsere Resultate zusammenstellen. Wir berechnen noch Mittelwerte und Varianzen.

n-fach zusammengesetzte Experimente (Augenzahl  $a \in \frac{12}{n}$ , Augensumme  $k = a_1 + \dots + a_n$ )

Wahrscheinlichkeitsverteilungen:

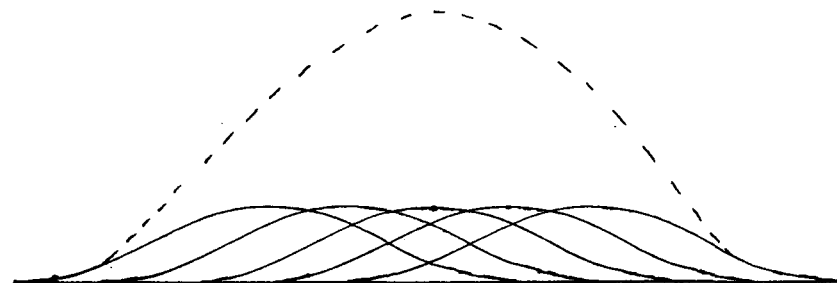
	Augensummen											
n	1	2	3	4	5	6	7	8	9	10	11	12
1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$
2		$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
3			$\frac{1}{64}$	$\frac{3}{64}$	$\frac{6}{64}$	$\frac{10}{64}$	$\frac{12}{64}$	$\frac{12}{64}$	$\frac{10}{64}$	$\frac{6}{64}$	$\frac{3}{64}$	$\frac{1}{64}$
4				$\frac{1}{81}$	$\frac{4}{81}$	$\frac{10}{81}$	$\frac{16}{81}$	$\frac{19}{81}$	$\frac{16}{81}$	$\frac{10}{81}$	$\frac{4}{81}$	$\frac{1}{81}$
6						$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$
12												1

n	$\frac{12}{n}$	Verteilungs- -formel	Verteilungs- -name	$\mu$	$s^2$	$\left(\frac{12}{n}\right)^n =$ # Laplace-E.	
1	12	$\frac{1}{12}$		Gleichv.	6,5	$11\frac{11}{12}$	12
2	6	$\frac{k}{36}$ bzw. $\frac{12-k}{36}$		Dachv.	7	$5\frac{5}{6}$	36
3	4	?		Haufenv.	7,5	$3\frac{3}{4}$	64
4	3	?		"	8	$2\frac{2}{3}$	81
6	2	$\binom{6}{k} \cdot \frac{1}{64}$		Binomialv.	9	$1\frac{1}{2}$	64
12	1	1		sicheres Ereignis	12	0	1

Siehe Abbildungen am Ende dieses Artikels.

Die graphischen Darstellungen dieser sechs Verteilungen machen offensichtlich, daß sie in dieser Reihenfolge, von der Gleich- zur Punktverteilung, zunehmend der Normalverteilung ähnlicher werden, wobei die letzte Verteilung zwar entartet ist, aber diese Regel nicht verletzt. Außerdem fällt uns die Symmetrie bei allen Graphen ins Auge. Diese muß uns nicht verwundern, denn wenn wir bei den Würfeln die höchste Augenzahl mit 1, die zweithöchste mit 2, die dritthöchste mit 3 usw. bewerten, muß sich jeweils dieselbe Verteilung ergeben und diese muß, da die Würfel nichts von der Ummumerierung wissen können, auch der am Mittelwert gespiegelten früheren gleichen.

In unserem Überblick über die sechs Verteilungen und ihre Eigenschaften, fällt uns schon bei wenig Nachdenken auf, daß für die Varianz offenbar  $s^2(n) = \frac{12}{n} - \frac{n}{12}$  gilt, auch für den Extremfall  $n = 12$ . Bewiesen haben wir diese Formel, da wir sie für jedes mögliche  $n$  bestätigt haben. Aber verstanden haben wir sie nicht. Wir haben einzeln uns sehr verschieden anmutende Verteilungen betrachtet, und nun lassen sich sich - so vermuten wir - alle über einen Kamm (eine Formel) scheren. Wenn wir geschickt vorgehen würden, könnten wir doch wohl alles auf einmal beweisen. Unser Ansatzpunkt kann zunächst nur das von uns gewollte Gemeinsame sein; nämlich das  $n$ -fache Ausführen eines Experimentes mit  $\frac{12}{n}$  gleichwahrscheinlichen Ausgängen. Weshalb sehen wir diese sechs Experimente als zusammengehörig an? Wie können wir eine (einzige, gemeinsame) Regel finden, wie wir uns die sechs Experimente aufgebaut denken? Eine erste Antwort ist naheliegend. Der schichtweise Aufbau, wie wir ihn bei  $n = 3$  und  $n = 4$  verfolgten, läßt sich in jedem Fall hineininterpretieren: Wir führen ein Experiment mit  $\frac{12}{n}$  gleichwahrscheinlichen Ausgängen  $n$ -mal nacheinander durch. Bei jeder weiteren Durchführung ist die bisherige Verteilung mit der neuen, gleichverteilten Durchführung zu kombinieren. Wenn wir die Varianz in eben diesen Schritten mitberechnen könnten, würden wir dies vielleicht als Einblick und Verstehen empfinden. Rechentechnisches Kernstück wird der Übergang der Varianz von einer Verteilung zu ihrer Kombination mit einer Gleichverteilung sein. Dies bedeutet graphisch eine Überüberlegung äquidistant gegeneinander verschobener, identischer Verteilungen (kontinuierlich, da anschaulicher):



Oder algebraisch (diskret):

$$\sum_{i,k} f(i-k) \cdot (i-\mu)^2,$$

wobei  $i$  die Augenzahlen und  $\mu$  den alten Mittelwert bedeuten und  $k$  die Verschiebungen nummeriert. Wollen wir diese mit der Varianz  $\sum f(i-k)(i-k-\mu)^2$  der bisherigen Verteilung in Beziehung setzen, so müssen wir mittels Binomischer Formeln den Summanden  $k$  aus dem Quadrat heraussezieren. Diese Formel sieht reichlich allgemein und entsprechend undurchsichtig aus. Beschränken wir uns deshalb zunächst auf einen Spezialfall! Diesem entsprechend können wir dann den allgemeinen formulieren und dabei auch unnötige Verallgemeinerungen vermeiden. Wir haben also z.B. bei  $n = 3$  schon zweimal gewürfelt und für die zugehörige Verteilung  $f$  erkannt:

	$i$	2	3	4	5	6	7	8	9	10	11	12
$f(i)$		$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$				

Während für "jede der bisherigen" Verteilungen

$$s_{\text{bisher}}^2 = \frac{1}{16} (1 \cdot 3^2 + 2 \cdot 2^2 + 3 \cdot 1^2 + 4 \cdot 0^2 + 3 \cdot 1^2 + 2 \cdot 2^2 + 1 \cdot 3^2)$$

zu berechnen wäre, weil sich "jede der bisherigen" Verteilungen auf ihren (!) Mittelwert bezieht, gilt für die neue Verteilung nun dasselbe (!). Mittelwert:

$$s_{\text{neu}}^2 = \frac{1}{64} \left( 1 \cdot 4 \frac{1}{2}^2 + 2 \cdot 3 \frac{1}{2}^2 + 3 \cdot 2 \frac{1}{2}^2 + 4 \cdot 1 \frac{1}{2}^2 + 3 \cdot \frac{1}{2}^2 + 2 \cdot \frac{1}{2}^2 + 1 \cdot \frac{1}{2}^2 + 1 \cdot 3 \frac{1}{2}^2 + 2 \cdot 2 \frac{1}{2}^2 + 3 \cdot 1 \frac{1}{2}^2 + 4 \cdot \frac{1}{2}^2 + 3 \cdot \frac{1}{2}^2 + 2 \cdot \frac{1}{2}^2 + 1 \cdot \frac{1}{2}^2 + 1 \cdot 2 \frac{1}{2}^2 + 1 \cdot 2 \frac{1}{2}^2 + 2 \cdot 1 \frac{1}{2}^2 + 3 \cdot \frac{1}{2}^2 + 4 \cdot \frac{1}{2}^2 + 3 \cdot \frac{1}{2}^2 + 2 \cdot \frac{1}{2}^2 + 1 \cdot \frac{1}{2}^2 + 1 \cdot 3 \frac{1}{2}^2 + 1 \cdot 1 \frac{1}{2}^2 + 2 \cdot \frac{1}{2}^2 + 3 \cdot \frac{1}{2}^2 + 4 \cdot \frac{1}{2}^2 + 3 \cdot \frac{1}{2}^2 + 2 \cdot \frac{1}{2}^2 + 1 \cdot \frac{1}{2}^2 \right).$$

Unser Bestreben ist ja, dies durch möglichst eine einzige Formel zu erfassen, um dann schrittweise die Entwicklung algebraisch verfolgen zu können. Unser Zahlenschema ist punktsymmetrisch, dies könnte zu einer Vereinfachung der Formel beitragen. Jeder Abschnitt in der Klammer läßt sich schreiben als

$$1 \cdot (3+k)^2 + 2 \cdot (2+k)^2 + 3 \cdot (1+k)^2 + 4 \cdot (0+k)^2 + 3 \cdot (-1+k)^2 + 2 \cdot (-2+k)^2 + 1 \cdot (-3+k)^2$$

mit  $k = 1\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -1\frac{1}{2}$ , also die gesamte Klammer

$$\text{als } \sum_{k=-1\frac{1}{2}}^{+1\frac{1}{2}} \sum_{i=-3}^{+3} f(\mu_{\text{bisher}} + i) \cdot (i+k)^2.$$

Die Binomischen Formeln ermöglichen uns nun folgende Umformungen:

$$\begin{aligned} s_{\text{neu}}^2 &= \frac{1}{64} \sum_{k=-1\frac{1}{2}}^{+1\frac{1}{2}} \sum_{i=-3}^{+3} f(\mu_{\text{bisher}} + i) \cdot (i+k)^2 \\ &= \frac{1}{64} \sum \sum f(\dots) \cdot (i^2 + 2ik + k^2) \\ &= \frac{1}{64} \sum \sum f(\dots) \cdot i^2 \quad \text{:unabhängig von } k \\ &+ \frac{2}{64} \sum \sum f(\dots) \cdot ik \quad \text{:ergibt Null} \\ &+ \frac{1}{64} \sum \sum f(\dots) \cdot k^2 \quad \text{:es ist } \sum_i f(\dots) = 16 \\ &= \frac{4}{64} \sum_i f(\dots) \cdot i^2 + \frac{16}{64} \sum_k k^2 \\ &= s_{\text{bisher}}^2 + s_{\text{zusätzlich}}^2 \end{aligned}$$

Nun haben wir eine viel allgemeiner anmutende Formel (für unseren Spezialfall) erhalten, als wir erwarteten. Gilt sie allgemein? Ohne große Mühe können wir einsehen, daß diese Rechnung allgemein für die Summe ("Faltung") einer symmetrischen mit einer gleichverteilten Zufallsgröße gilt.

Wenn wir dies Ergebnis auf unsere Situation mehrmals nacheinander anwenden, wissen wir, daß die Varianz bei einer n-fachen Summe mit einer Gleichverteilung n-mal so groß wie die Varianz dieser Gleichverteilung ist. Wir haben also als nächstes die Varianz einer Gleichverteilung mit den Elementarereignissen  $1/2/\dots/12/n$  zu suchen ( $n = 1, 2, 3, 4, 6, 12$ ). Der Mittelwert beträgt  $\frac{6}{n} + \frac{1}{2}$  und die äußeren Werte 1 und  $\frac{12}{n}$  sind um  $\frac{6}{n} - \frac{1}{2}$  von diesem entfernt.

Gemäß Definition gilt

$$s^2 = \frac{n}{12} \left( \left(\frac{6}{n} - \frac{1}{2}\right)^2 + \left(\frac{6}{n} - 1\frac{1}{2}\right)^2 + \left(\frac{6}{n} - 2\frac{1}{2}\right)^2 + \dots + \left(\frac{6}{n} - \left(\frac{12}{n} - \frac{1}{2}\right)\right)^2 \right).$$

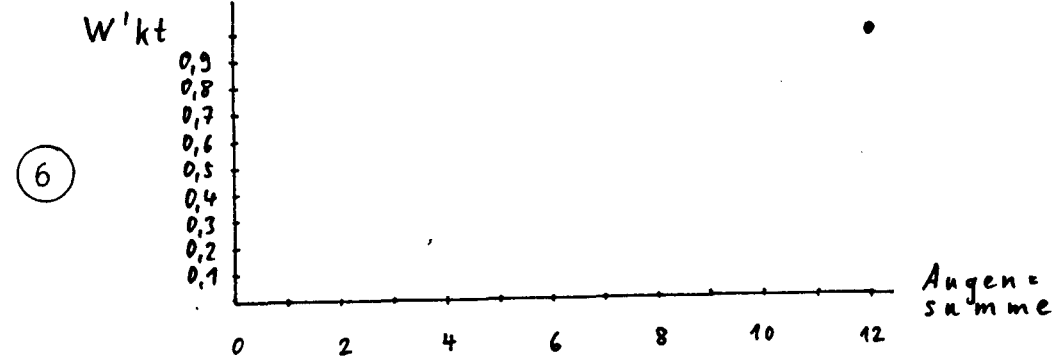
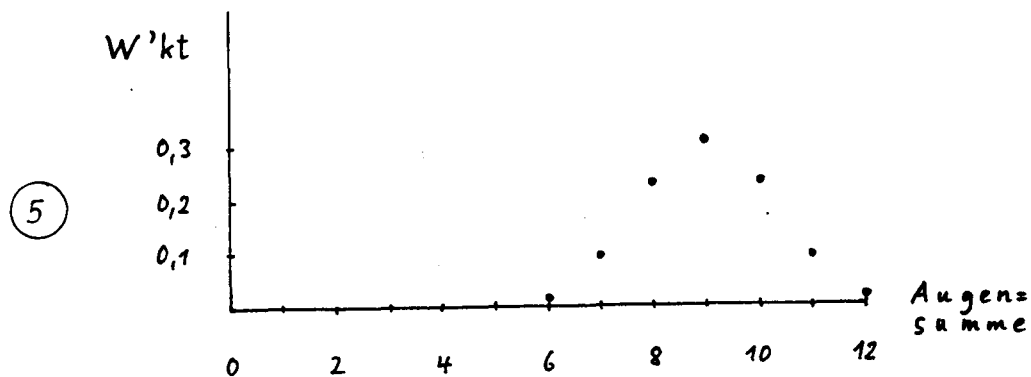
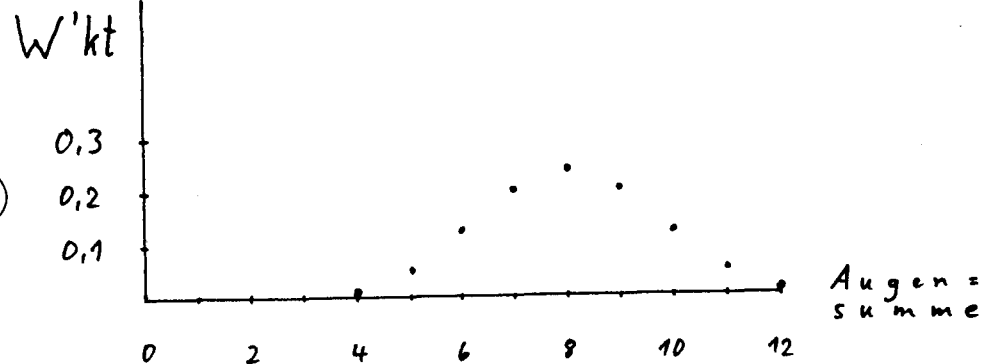
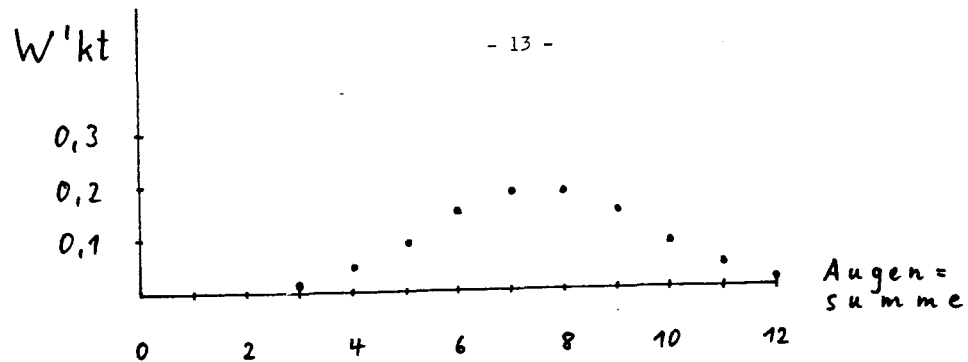
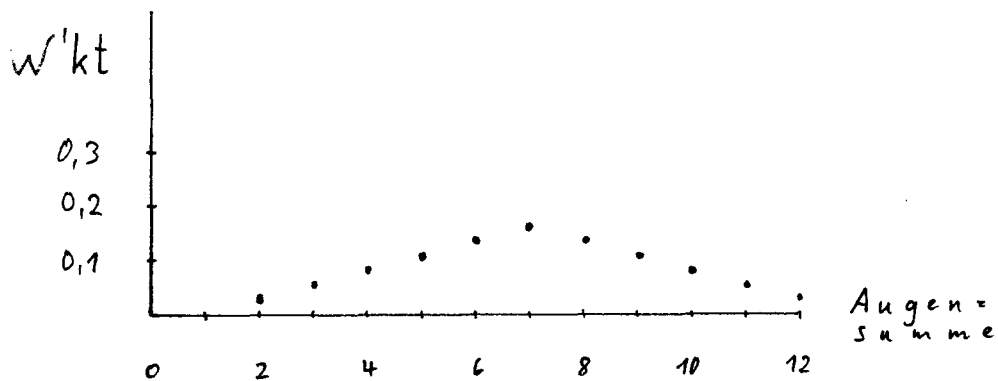
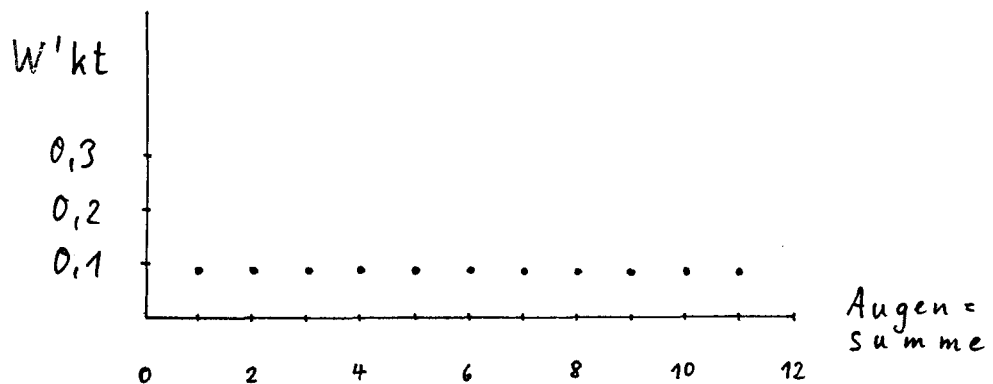
Die Klammer umfaßt  $\frac{12}{n}$  Summanden, jeden wollen wir, weil wir uns einen ähnlichen Effekt wie bei der letzten Rechnung erhoffen, gemäß der Binomischen Formel aufdröseln.

$$\begin{aligned} s^2 &= \frac{n}{12} \sum_{i=1}^{12/n} \left(\frac{6}{n} - \left(i - \frac{1}{2}\right)\right)^2 \\ &= \frac{n}{12} \sum \left(\frac{6}{n} - i + \frac{1}{2}\right)^2 \\ &= \frac{n}{12} \sum \left(\frac{36}{n^2} + i^2 + \frac{1}{4} - 12 \frac{i}{n} + \frac{6}{n} - i\right) \\ &= \frac{n}{12} \cdot \left(\frac{36}{n^2} + \frac{1}{4} + \frac{6}{n}\right) \cdot \frac{12}{n} + \frac{n}{12} \sum i^2 - \frac{n}{12} \cdot \left(\frac{12}{n} + 1\right) \sum i \\ &= \frac{36}{n^2} + \frac{1}{4} + \frac{6}{n} + \frac{n}{12} \cdot \frac{12}{n} \cdot \left(\frac{12}{n} + 1\right) \cdot \frac{(24}{n} + 1) - \frac{n}{12} \cdot \left(\frac{12}{n} + 1\right) \cdot \frac{1}{2} \cdot \frac{12}{n} \cdot \left(\frac{12}{n} + 1\right) \\ &\vdots \\ &= \frac{12}{n^2} - \frac{1}{12}. \end{aligned}$$

Multiplizieren wir dies mit n, so erhalten wir  $\frac{12}{n} - \frac{n}{12}$  und freuen uns über unsere Einsicht.

Wir hatten uns oben eine weitere Idee aufbewahrt: Da wir jede gleichwahrscheinlich verteilte Zufallsgröße mit  $\frac{12}{n}$  Elementar-Ereignissen

durch ein Dodekaeder verwirklichen können, indem wir je  $n$  Seiten zu einem einzigen Elementarereignis zusammenfassen, vermuten wir, daß wir wohl nur die Wahrscheinlichkeitstheoretische Situation für das Dodekaeder berechnen müssen und dann aus diesen Ergebnissen die Ergebnisse für unsere anderen Experimente durch Reduktion modulo  $\frac{12}{n}$  gewinnen können. Leider zeigen uns schon einfache Zahlenbeispiele, daß diese Idee sich nicht ausführen läßt, weil Summieren und Restklassenbildung nicht vertauschbar sind. (Ich habe diesen Gedanken nicht aus meinem Manuskript wieder gestrichen, weil mir als Mathematiklehrer dies zu unaufrichtig erschien. Beim Mathematik-treiben läuft man zum Überwiegenden Teil in Sackgassen, und nur manchmal erlebt man Erfolg. Den Schülern wird aber Mathematik andersartig dargestellt. Freilich kann man sie nicht selbst alles entdecken lassen, aber man sollte ihnen mathematische Mißerfolgs-erlebnisse größeren Stils auch nicht gänzlich vorenthalten).



## Probleme eines Statistikuterrichtes nach STRICK-Muster

von Raphael Diepgen, Bochum

Während ein Großteil mathematischer Verfahren bislang vor allem nur von physikorientierter Naturwissenschaft und Technik benötigt wird, verlangen nach statistischen Kenntnissen heute nahezu alle Disziplinen mit empirischem Selbstverständnis, also auch Disziplinen wie Pädagogik, Psychologie, Soziologie, Ökonomie, Medizin, Biologie. Gerade dieses breite Anwendungsspektrum hat der beurteilenden Statistik einen Platz im gymnasialen Mathematikcurriculum verschafft. Nach der zunächst euphorischen Übernahme inferenzstatistischer Verfahren in den Methodenkanon vor allem humanwissenschaftlicher Disziplinen hat indessen die wissenschaftstheoretische und methodenkritische Diskussion der letzten Jahre, etwa die sog. Signifikanztestdebatte in Soziologie und Psychologie - siehe zum Überblick etwa BREDEKAMP (1972) oder WITTE (1980) -, deutlich gemacht, daß die inferenzstatistischen Standardverfahren häufig ohne jedes Verständnis ihrer Logik, sondern lediglich in Orientierung an unverstandener Konvention angewandt und in ihren Ergebnissen miß- und in der Regel überinterpretiert werden. Aufgabe eines Statistikuterrichtes auf der Schule sollte es daher meiner Meinung nach sein, einer irrationalen statistischen Praxis im Forschungsalltag vorzubeugen, indem Logik und zugleich Problematik inferenzstatistischer Verfahren ausführlich thematisiert werden; die Vermittlung rechen-technischer Fertigkeiten im Umgang mit inferenzstatistischen Standardverfahren allein kann daher meines Erachtens nicht Ziel eines wissenschaftspropädeutischen Statistikuterrichtes sein, zumal diese Fertigkeiten durch die Verbreitung entsprechender Computerprogramme zunehmend überflüssig werden. Vielmehr sollte der Unterricht die argumentativen Grundlagen für einen aufgeklärten und kritischen Umgang mit dem statistischen Instrumentarium vermitteln.

Nun leidet aber der Unterricht über schließende Statistik unter einem gleich zweifachen Ausbildungsdefizit der Mathematiklehrer: Die meisten von ihnen haben sich während ihres Studiums nicht mit der Theorie mathematischer Statistik beschäftigt. Und noch viel weniger vertraut sind sie mit der inferenzstatistischen Praxis außerhalb der Mathematik, mit jenen Anwendungen also, auf die hin mathematische Statistik überhaupt formu-

liert ist. In einer solchen Situation bekommen Schulbücher hervorragende Bedeutung für den Unterricht. Denn es lernen aus ihnen die Theorie - und wozu sie gut sein soll - nicht nur die Schüler, sondern zunächst einmal die Lehrer. Erklärlich daher, daß ein Schulbuch Verbreitung findet besonders dann, wenn es beurteilende Statistik leicht verständlich und zugleich eingebettet in ihre Anwendungen darzustellen verspricht. Dies Versprechen bietet von der Aufmachung her insbesondere das Buch "Einführung in die Beurteilende Statistik" von STRICK (1980), dessen didaktische Konzeption vom Autor mehrfach dargestellt wurde (STRICK 1978, 1979, 1981a, 1981b). Es sei daher - stellvertretend auch für andere Schulbücher zur Statistik - näher betrachtet.

Intention von STRICK ist es, in einem einsemestrigen Stochastikkurs durch die radikale Beschränkung auf einfachste wahrscheinlichkeitstheoretische Voraussetzungen breiten Raum zu lassen für die beurteilende Statistik, d.h. für Konfidenzintervalle und Hypothesentesten. Dies gelingt STRICK, indem er Inferenzstatistik reduziert auf das Grundmuster der Bereichsschätzung für Binomialverteilungen bei großen Stichproben. Dabei nutzt er die Approximation der Binomialverteilung an die Normalverteilung - natürlich ohne dies auszuführen, sondern nur an Beispielen illustrierend - zur Gewinnung der zentralen Aussage, daß beim  $n$ -stufigen BERNOULLI-Versuch mit hinreichend großem  $n$  die relative Häufigkeit der Erfolge mit einer Wahrscheinlichkeit von ca. 95,5 %, der sog. "Sicherheitswahrscheinlichkeit", in der  $2\sigma/n$ -Umgebung um die Erfolgswahrscheinlichkeit  $p$  liegt (S. 57, 60). Mit dieser einfachen Aussage werden nun beim sog. "Schluß von der Gesamtheit auf die Stichprobe für eine bekannte Erfolgswahrscheinlichkeit  $p$ " alle Stichprobenergebnisse innerhalb der  $2\sigma/n$ -Umgebung von  $p$  als "verträglich mit der Erfolgswahrscheinlichkeit  $p$ " bezeichnet (S. 60). So weit, so gut: Denn offensichtlich hat die Rede von der Verträglichkeit eines Stichprobenergebnisses mit einer bekannten Erfolgswahrscheinlichkeit einen Sinn; ich kann auf lange Sicht bei Ziehungen aus der Gesamtheit in 95,5 % der Fälle mit solchen Stichprobenergebnissen rechnen, d.h. mein Verhalten darauf einstellen, ich kann auf das Auftreten solcher Ergebnisse mit bestimmten Einsätzen wetten usw.. Der Begriff der Sicherheitswahrscheinlichkeit ist hier ohne Problem interpretierbar.

Der Einstieg in die schließende Statistik liest sich dann aber folgendermaßen: "In den meisten Fällen kennt man jedoch die Erfolgswahrscheinlich-

keit nicht, sondern kennt das Ergebnis einer Stichprobe. Aufgaben von diesem Typ heißen: Schluß von der Stichprobe auf die Gesamtheit. Die Grundfrage dieser Aufgaben lautet: Welche Erfolgswahrscheinlichkeit liegt dem Zufallsversuch zugrunde? Wir suchen alle Erfolgswahrscheinlichkeiten  $p$ , die mit dem Stichprobenergebnis verträglich sind. In unserem Lösungsansatz gehen wir also davon aus, daß das Stichprobenergebnis innerhalb der  $2\sigma/n$ -Umgebung um  $p$  liegt. Dies gilt mit einer Sicherheitswahrscheinlichkeit von 95,5 %." (S. 64). Die Menge aller dieser mit dem Stichprobenergebnis verträglichen Parameter  $p$  bekommt dann auf der folgenden Seite den Namen Konfidenzintervall. Nur der nachdenkliche Leser wird bei diesen glatten Formulierungen stutzig: Denn bislang kennt er nur den Begriff der Verträglichkeit eines Stichprobenergebnisses mit einer (bekannten) Erfolgswahrscheinlichkeit  $p$ . Was heißt aber die Verträglichkeit einer (unbekannten) Erfolgswahrscheinlichkeit  $p$  mit einem beobachteten Stichprobenergebnis, von dem hier plötzlich geredet wird? Kann ich, in Umkehrung des Schlusses von der Gesamtheit auf die Stichprobe, nun bei Vorliegen eines Stichprobenergebnisses mit bestimmten Einsätzen darauf wetten, daß die Erfolgswahrscheinlichkeit im Konfidenzintervall liegt? Oder kann etwa ein Forscher damit rechnen, daß auf lange Sicht - etwa über sein wissenschaftliches Leben - die von ihm ermittelten Konfidenzintervalle zu 95,5 % jeweils die wahren Parameter enthalten? Bedeutet die Sicherheitswahrscheinlichkeit von 95,5 % dafür, daß "das Stichprobenergebnis innerhalb der  $2\sigma/n$ -Umgebung um  $p$  liegt", wenn ich gar nicht weiß, um welches  $p$  es sich handelt, nicht dasselbe wie die Wahrscheinlichkeit, daß  $p$  innerhalb der  $2\sigma/n$ -Umgebung um das beobachtete Stichprobenergebnis liegt? Wie gesagt: Nur der aufmerksame Leser wird sich diese Fragen stellen, aber er wird sie aufgrund seiner Lektüre kaum beantworten können und daher schließlich nicht wissen, was denn der Sinn von Konfidenzintervallen ist, warum man ihnen vertrauen sollte. Der flüchtige und weniger nachdenkliche Leser dürfte es hier besser haben: Er wird einem Konfidenzintervall vertrauen, weil er glaubt, es enthalte mit 95,5 %-iger Wahrscheinlichkeit den wahren Parameter. Daß zu dieser zweiten Lesergruppe durchaus auch Mathematiklehrer gehören, die schon jahrelang mit diesem Schulbuch arbeiten, zeigt die Fruchtlosigkeit der Warnung vor dieser Fehlinterpretation im Lehrerbegleitheft zum Schulbuch (S. 49); denn diese Warnung ist nicht verbunden mit einer positiven Interpretation, die das Vertrauen in ein Konfidenzintervall

rechtfertigen würde. STRICKs suggestiver Übergang vom wohldefinierten und wohlinterpretierten Begriff der Verträglichkeit eines Stichprobenergebnisses mit einer Erfolgswahrscheinlichkeit  $p$  zum lediglich rein formal rekonstruierbaren, die ganze inferenzstatistische Problematik aber berührenden Begriff der Verträglichkeit einer Erfolgswahrscheinlichkeit  $p$  mit einem Stichprobenergebnis legt das skizzierte Mißverständnis natürlich nahe. Wer beim Inferenzschluß von einer Sicherheitswahrscheinlichkeit redet, ohne ausdrücklich hervorzuheben, daß diese Wahrscheinlichkeit unter der Bedingung steht, daß in Wahrheit  $p$  eine bestimmte Größe hat - was man leider nicht weiß -, der braucht sich nicht darüber zu wundern, daß der Leser zwar rechentechnische Fertigkeit in der Bestimmung von Konfidenzintervallen erwirbt, im Grunde aber nicht versteht, welchen Sinn und welche Begründung dieses Tun hat.

Ähnliche Unklarheiten begleiten die Darstellung des Hypothesentestens bei STRICK (S. 71). Der Test einer Nullhypothese  $H_0$  erscheint als Regel,  $H_0$  zu verwerfen, wenn das Stichprobenergebnis nicht mit der hypothetischen Erfolgswahrscheinlichkeit unter  $H_0$  verträglich ist. Warum indessen die fehlende Verträglichkeit die Verwerfung von  $H_0$  rechtfertigt, dies bleibt im Dunkeln. Mißverständlich auch hier (S. 70) die Bezeichnung der "Wahrscheinlichkeit für eine falsche Entscheidung (hier 4,5 %) als Irrtumswahrscheinlichkeit", so als stände diese Wahrscheinlichkeit nicht unter der Bedingung der Geltung von  $H_0$  und als gäbe es nur falsche Entscheidungen in Form des Fehlers 1. Art, nicht aber in Form des Fehlers 2. Art. Rein formal werden zwar später (S. 74) beide Fehlerarten unterschieden, dies aber ohne irgendwelche Folgen. Weder wird jemals von der (bedingten) Wahrscheinlichkeit für einen Fehler 2. Art gesprochen - es fehlt also gänzlich das Konzept der Testgüte -, noch wird die gegenläufige Beziehung zwischen beiden Fehlerwahrscheinlichkeiten und damit zwischen Signifikanzniveau und Testgüte erörtert, eine Beziehung, die erst ein sinnvolles Abwägen bei der Wahl des Signifikanzniveaus erlauben würde. Der Charakter des Hypothesentestens als eines Entscheidungsverfahrens, dessen sinnvoller Einsatz ein Abwägen zwischen beiden Fehlentscheidungsmöglichkeiten erfordert, bleibt undeutlich. Dem Leser wird die Logik vorenthalten, die dem Hypothesentesten überhaupt eine Begründung als sinnvolles Entscheidungsverfahren verleiht. So dürfte er kaum zu einem begründeten Umgang mit dem Testinstrumentarium geführt werden, eher zu den vielen Un-



sinnigkeiten falschverstandener statistischer Alltagspraxis: Er wird die Wahl eines Signifikanzniveaus nicht legitimieren können, denn er weiß ja gar nicht, was sie ihm bringt und was sie ihn kostet. Er wird wahrscheinlich der Unsitte verfallen, erst nach der Inspektion der Daten das Signifikanzniveau festzulegen, denn er hat den Charakter des Tests als eines im Vorhinein definierten Entscheidungsverfahrens kaum verstanden. Er wird mit Tests, bei denen die Nullhypothese mit der Forschungshypothese übereinstimmt, mit Anpassungstests also, nichts anfangen können, oder er wird sie wie normale Tests behandeln. Er wird signifikante Ergebnisse für wichtig halten und hochsignifikante Ergebnisse für noch wichtiger, seien die Stichproben auch noch so groß und die Abweichungen von der Nullhypothese auch noch so klein und wissenschaftlich wie praktisch völlig irrelevant. Er wird nicht die Gefahr verstehen, daß Signifikanz in der Forschungspraxis degenerieren kann zu einem Maß für den empirischen Aufwand, den der Forscher bereit ist zu betreiben, der ja so gut wie sicher sein kann, daß seine Nullhypothese exakt ohnehin nicht gelten wird. Denn dem Leser fehlt das entscheidende Konzept der Fehlerwahrscheinlichkeit 2. Art und Testgüte.

Fazit: Die didaktisch zunächst reizvolle Idee von STRICK, die formale und rechnerische Äquivalenz zwischen Konfidenzintervallen und Hypothesentests einerseits, der Bereichsschätzung bei bekanntem Parameter andererseits zu nutzen in Form eines einfachen Lösungsschemas für alle statistischen Fragestellungen, hat den Nachteil, daß die spezifische Logik der Inferenzstatistik - und damit zugleich auch ihre Problematik - verschleiert wird. Der Schüler lernt die rezeptartige Berechnung von Konfidenzintervallen und Hypothesentests, hier ja immer nur die Ermittlung von  $\sigma$ -Umgebungen, er wird aber meines Erachtens kaum dazu in die Lage versetzt, dieses Tun überzeugend zu begründen. Er erlernt sein simples "Strick"-Muster, bestenfalls vielleicht eine gewisse Skepsis gegenüber veröffentlichten Untersuchungsergebnissen, eine Skepsis, die er kaum in klare Worte fassen können. Allein, diese sicher sinnvolle Skepsis könnte der Unterricht - so würde ich vermuten - viel einfacher erzeugen durch einige Demonstrationsbeispiele mithilfe von Simulation.

Stellvertretend wird hier das Buch von STRICK kritisiert für all jene Schulbücher, die auf das Konzept der Fehlerwahrscheinlichkeit 2. Art oder Testgüte verzichten. Diese Schulbücher verharren damit auf

dem allenfalls heuristischen Argumentationsniveau des ursprünglichen Signifikanzgedankens, klassisch formuliert etwa in FISHERS "Problem of the lady tasting tea", der Idee nämlich, eine isolierte einfache Hypothese als wenig plausibel zu verwerfen allein deshalb, weil in einem Experiment Ereignisse auftreten, die bei Geltung der Hypothese unwahrscheinlich sind. Das argumentative Ungenügen dieser allzu simplen Idee wurde erst behoben durch die Entwicklung der Testtheorie, etwa von NEYMAN und PEARSON, also durch die Betrachtung der Fehlerwahrscheinlichkeit 2. Art. Wenn Schulbücher darauf verzichten, so bleiben sie dem Leser eine ausreichende Begründung schuldig, warum denn bloß das als "Hypothesentesten" vorgeschlagene und eintrainierte Entscheidungsverhalten vernünftig sein soll. Ohne Kenntnis des Konzepts der Testgüte scheint mir schließlich eine aufgeklärte und kritische statistische Praxis schlechterdings unmöglich. Auf eine solche Praxis aber sollte der schulische Statistikerunterricht propädeutisch abzielen.

Warum ich - abgesehen von der Möglichkeit zum Wortspiel - unter vergleichbaren Schulbüchern gerade jenes von STRICK ausgewählt habe? Nun, das Einzigartige beim Konzept von STRICK ist, daß hier der Lehrer auch nicht mehr im Nachhinein die Möglichkeit hat, durch zusätzliche Betrachtung von Fehlerwahrscheinlichkeiten 2. Art die Argumentationslücken zu schließen. Denn dank der Reduzierung der Wahrscheinlichkeitsrechnung einzig auf die Berechnung von  $\sigma$ -Umgebungen hin ist es dem Schüler im Statistikerunterricht nach STRICK gänzlich unmöglich, die Wahrscheinlichkeit zu berechnen, daß ein Stichprobenergebnis in die  $2\sigma/n$ -Umgebung um die hypothetische Erfolgswahrscheinlichkeit  $p_0$  fällt, wenn in Wahrheit eine andere Erfolgswahrscheinlichkeit  $p_1$  (mit  $p_1 \neq p_0$ ) vorliegt. Letztlich erweist es sich an dieser Stelle als folgenreiche Fehlentscheidung von STRICK (1981a, S. 10), dem Konfidenzintervall bloß seiner rechnerischen und formalen Universalität wegen didaktische Priorität vor dem Hypothesentesten zugesprochen zu haben. Ich würde die Prioritäten andersherum setzen, im Hinblick sowohl auf die Interpretierbarkeit, als auch auf die Bedeutung im Wissenschaftsalltag.

Ich weiß nicht, ob es in einem nur einsemestrigen Stochastikgrundkurs möglich ist, angemessen beurteilende Statistik darzustellen. Sollte dazu die Zeit fehlen, dann erschiene es mir allerdings sinnvoller, die Wahrscheinlichkeitsrechnung wenigstens so ausführlich zu betreiben, daß

darauf aufbauend die Universität Statistik angemessen darstellen könnte.

#### LITERATUR

- BREDENKAMP, J., 1972: Der Signifikanztest in der psychologischen Forschung. Frankfurt
- STRICK, H. K., 1978: "Mathematische Statistik. Vorschlag für einen einsemestrigen Grundkurs." Lernzielorientierter Unterricht 4, 28-37
- STRICK, H. K., 1979: "Parameterschätzung und Hypothesentesten im einsemestrigen Grundkurs Wahrscheinlichkeitsrechnung/Statistik. In: Beiträge zum Mathematikunterricht 1979. Vorträge. Hannover, S. 356-359
- STRICK, H. K., 1980: Einführung in die Beurteilende Statistik. Hannover
- STRICK, H. K., 1981a: "Die Bestimmung von Konfidenzintervallen im Grundkurs Wahrscheinlichkeitsrechnung/Statistik." MNU 34, 7-11
- STRICK, H. K., 1981b: "Methoden der beurteilenden Statistik im Grundkurs Stochastik." In: DÖRFLER, W., FISCHER, R. (Hrsg.): Stochastik im Schulunterricht. Wien, S. 245-248
- WITTE, E. H., 1980: Signifikanztest und statistische Inferenz. Analysen, Probleme, Alternativen. Stuttgart

Anmerkung zu R. DIEPGEN: Probleme eines Statistikerunterrichtes nach STRICK-Muster

von H. K. Strick, Leverkusen

Schon vor einigen Heften hatten wir dazu aufgerufen, in kritischer Form zu unseren Beiträgen Stellung zu nehmen. Dies ist bisher leider nicht geschehen. Der vorstehend abgedruckte Aufsatz setzt sich in kritischer Weise mit der Konzeption von Stochastikkursen auseinander - er weist auf gefährliche Stellen und Schwachpunkte insbesondere meines Grundkursbuches hin. Ich sehe diese gefährlichen Stellen in meiner Konzeption ebenfalls; wenn ich auch meine, daß bei konsequenter Unterrichtsführung Mißverständnisse nicht auftreten sollten. Daß in einem einsemestrigen Grundkurs nicht alle Fragen der Beurteilenden Statistik erarbeitet werden können, ist offensichtlich. Ob man aus diesem Grunde auf eine "Einführung" (und mehr biete ich nicht an) verzichten sollte, darüber könnte man streiten. Bei der Konzeption meines Kurses erschien es mir als wichtigste Idee, die "Gesetzmäßigkeiten" des Zufalls - bei BERNOULLI-Versuchen - zu vermitteln, um damit dem Schüler einen Einblick in Schätz- und Testverfahren zu gewähren.