Österreichische Mathematische Gesellschaft

MATHE-BRIEF

Juni 2014 — Nr. 48

Herausgegeben von der Österreichischen Mathematischen Gesellschaft
http: // www.oemg.ac.at / Mathe–Brief ——— mathe–brief@oemg.ac.at

Liebe Kolleginnen und Kollegen,

das Schuljahr 2013/14 geht zu Ende — Zeit für einige unterhaltsame Resultate der Mathematik!

Wer kennt sie nicht, die „langweiligen“ Resultate der Mathematik. Es ist zum Beispiel gar nicht so einfach, zu beweisen, dass man auf einer Schiffsreise von der Nord- auf die Südhalbkugel der Erde (mindestens) einmal den Äquator passieren muss. Aber der Zwischenwertsatz ist halt ein „No-na-Ergebnis“. Wir wollen uns diesmal einige Ergebnisse ansehen, die unserer Intuition völlig widersprechen.

Frauen haben immer Recht

Kann es sein, dass ein Mann immer bessere Argumente hat, aber seine Frau/Freundin am Ende des Tages doch „gewinnt“? Ja! Sehen wir uns folgende Situation an. Einige Frauen und Männer kommen zur Fahrprüfung. Hier sind die Ergebnisse von 2 Tagen:

	Frauen	Frauen	Frauen	Männer	Männer	Männer
	Anzahl	bestanden	Anteil	Anzahl	bestanden	Anteil
1. Tag	8	7	87.5%	1	1	100%
2. Tag	2	1	50%	3	2	66.7%

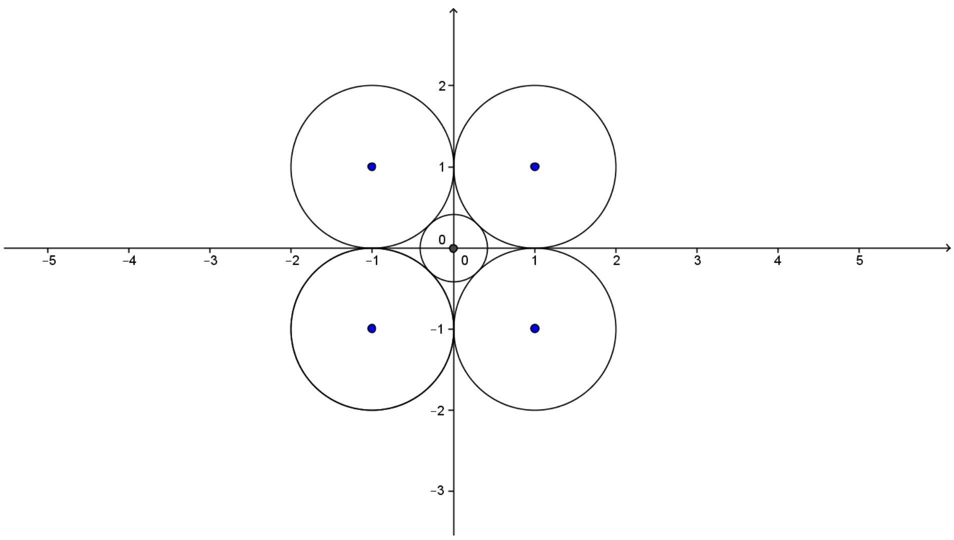
An beiden Tagen waren also die Männer erfolgreicher. Wie sieht die Gesamtbilanz aus?

	Frauen	Frauen	Frauen	Männer	Männer	Männer
	Anzahl	bestanden	Anteil	Anzahl	bestanden	Anteil
1. Tag	8	7	87.5%	1	1	100%
2. Tag	2	1	50%	3	2	66.7%
gesamt	*10*	*8*	*80%*	*4*	*3*	*75%*

Obwohl an beiden Tagen die Männer „gewonnen“ haben, gewinnen insgesamt die Frauen! Wie ist das möglich? (*Hinweis*: Vergleichen Sie die Brüche $\frac{a}{b} + \frac{c}{d}$ und $\frac{a+b}{c+d}$!) Dieses Phänomen ist als das Simpson-Paradox bekannt.

Ab Dimension 10 wird alles anders

Stellen wir uns vier Kreise, jeweils mit Radius 1 und den Mittelpunkten in $(1,1)$, $(1,-1)$, $(-1,1)$ bzw. $(-1,-1)$ vor:



Welchen Radius r hat der kleine Kreis in der Mitte? Der Satz des Pythagoras liefert uns

$$(1+r)^2 = 1^2 + 1^2, \text{ also } r = \sqrt{2} - 1, \text{ also ca. } r \approx 0.41$$

Wie sieht die entsprechende Fragestellung im 3-dimensionalen Raum aus? Da hat man 8 Kugeln mit Radius 1, die eine kleine Kugel in der Mitte „einzwicken". Deren Radius ergibt sich aus dem „3-dimensionalen Pythagoras" als $(1+r)^2 = 1^2 + 1^2 + 1^2$, diesmal also als $r = \sqrt{3} - 1$ mit einem Wert von ca. 0.73. Im 4-dimensionalen Raum hätte man analog $r = \sqrt{4} - 1 = 1$. Und wie weiter?

Die innere „Kugel" scheint also immer mehr zu wachsen. Bei Dimension 9 hat man schon unglaubliche $r = \sqrt{9} - 1 = 2$, die innere „Kugel" reicht also schon bis zum Rand unseres Würfels, und ab der Dimension 10 reicht sie also schon aus dem Würfel hinaus. Ab Dimension 10 stimmen diese Rechnungen also nicht mehr mit unseren naiven Vorstellungen von Würfeln und Kugeln überein.

Warum kann die „kleine" Kugel so groß werden? Hat ein Aufsichtsrat oder eine andere Kontrollinstanz versagt? Wie groß ist das Volumen einer n-dimensionalen Kugel und wie definiert man so etwas überhaupt? Am besten, man zählt nach, wie viele kleine „Würfelchen" der Seitenlänge s in diese Kugel hineinschlichten kann (jedes dieser kleinen Würfel hat das Volumen s^n). Sind es w Stück, so hat die Kugel ein Volumen, das etwas größer als ws^n ist.

Eine kleine Herausforderung für die Leserinnen und Leser: Versuchen Sie, auf diese Weise selbst die Formel für den ungefähren Flächeninhalt eines Kreises zu berechnen und vergleichen Sie Ihr Ergebnis mit der „offiziellen" Formel $r^2\pi$!

Dann lässt man s immer kleiner werden und sucht den Grenzwert. Die folgende Tabelle gibt die Ergebnisse, darunter auch die Inhalte von Kreis (Dimension $n = 2$) und der gewöhnlichen Kugel (Dimension $n = 3$).

Dimension	2	3	4	5	6	10
Volumen	$r^2\pi$	$\frac{4r^3\pi}{3}$	$\frac{r^4\pi^2}{2}$	$\frac{8r^5\pi^2}{15}$	$\frac{r^6\pi^3}{6}$	$\frac{r^{10}\pi^5}{120}$
Volumen bei Radius 1	3.14	4.19	4.93	5.26	5.17	2.55

Schon wieder etwas Unglaubliches: Zuerst wächst das Volumen der Kugeln, erreicht das Maximum ca. bei „Dimension $5\frac{1}{2}$", wird dann immer kleiner und geht gegen 0. Daher passt bei immer größerer Dimension eine immer größere kleine Kugel dazwischen, während das Volumen des Gesamtwürfels, in dem sich alles abspielt, mit 2^n immer größer wird. Sehr seltsam.

Wer es noch genauer wissen möchte: die allgemeine Formel für das Volumen $V(n)$ einer n-dimensionalen Kugel erhält man über die sogenannte Gammafunktion Γ (Für Genaueres siehe z.B. *http://de.wikipedia.org/wiki/Gammafunktion*). Sie genügt den Gleichungen $\Gamma(x+1) = x\Gamma(x)$ $(x > 0)$, $\Gamma(1) = 1$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, und sie interpoliert die Faktoriellen in dem Sinn, dass für alle natürlichen Zahlen n die Formel

$$\Gamma(n+1) = n! = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$$

gilt. Es gilt

$$V(n) = r^n \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}.$$

Bei geradem $n = 2k$ ergibt sich die Formel

$$V(n) = r^n \frac{\pi^{n/2}}{(n/2)!} = r^{2k}\frac{\pi^k}{k!},$$

bei ungeradem $n = 2k+1$ dagegen

$$V(n) = r^n \frac{2^{\frac{n+1}{2}}\pi^{\frac{n-1}{2}}}{n(n-2)(n-4)\cdots 1} = r^{2k+1}\frac{2^{2k+1}k!\pi^k}{(2k+1)!} = r^{2k+1}\frac{2^{k+1}\pi^k}{(2k+1)(2k-1)\cdots 1}.$$

Der Fluch der Unehrlichkeit

Der folgende Fall ist leider tatsächlich passiert. Wissenschafter wollten etwas „beweisen" und verwendeten dazu frei erfundene Daten. Hier sind 120 davon:

0,997	0,200	0,392	0,423	0,773	0,468	0,508	0,506	0,910	0,236	0,945	0,868	0,631	0,641	0,482
0,692	0,383	0,046	0,206	0,441	0,151	0,954	0,970	0,400	0,632	0,628	0,919	0,483	0,638	0,405
0,847	0,262	0,966	0,202	0,935	0,271	0,967	0,494	0,108	0,099	0,683	0,585	0,489	0,602	0,854
0,269	0,222	0,573	0,684	0,130	0,694	0,315	0,841	0,553	0,478	0,232	0,637	0,986	0,242	0,598
0,693	0,259	0,673	0,705	0,352	0,171	0,101	0,252	0,601	0,021	0,856	0,245	0,305	0,277	0,280
0,237	0,699	0,230	0,850	0,272	0,191	0,318	0,584	0,913	0,094	0,681	0,898	0,876	0,111	0,473
0,897	0,100	0,917	0,938	0,280	0,773	0,827	0,601	0,629	0,105	0,164	0,953	0,020	0,952	0,632
0,249	0,853	0,033	0,294	0,235	0,803	0,182	0,116	0,347	0,256	0,719	0,589	0,390	0,934	0,358

Ein anderer Wissenschafter (ein Biologe, der gut in Statistik geschult ist) schaute eine Zeitlang auf diese Tabelle und sagte dann: „Die Daten sind gefälscht!" *Warum hat er das behaupten können?*

Die Antwort liegt in den letzten Kommastellen. Deren relative Häufigkeit sollte jeweils um die 10% sein. Tatsächlich haben die Endziffern folgende Häufigkeiten:

0	1	2	3	4	5	6	7	8	9
12	14	16	15	10	9	11	11	12	10
10,00%	11,67%	13,33%	12,50%	8,33%	7,50%	9,17%	9,17%	10,00%	8,33%

Aus ungeklärten psychologischen Gründen verwendet man bei frei erfundenen Kommazahlen häufiger die Endziffern 1,2,3. Natürlich könnte die größere Häufigkeit der Endziffern 1,2,3, die hier auftritt, auch Zufall sein. Mit Hilfe der Statistik kann man jedoch zeigen, dass es mit mehr als 95% Sicherheit *kein* Zufall ist. Dazu braucht man kein biologisches Fachwissen, man muss nicht einmal wissen, was gemessen wurde.

Man kann dies etwa so berechnen: Die Wahrscheinlichkeit, durch bloßen Zufall eine der drei Zahlen 1,2,3 zu bekommen, ist gleich $p = \frac{3}{10} = 0.3$; die Gegenwahrscheinlichkeit beträgt $q = 0.7$. Die Wahrscheinlichkeit, bei n zufällig ausgewählten Endziffern genau k mal eine der Zahlen 1,2,3, zu bekommen, ist nach der Formel für die „Binomialverteilung“ gegeben durch $B(n,k) = \binom{n}{k} p^k q^{n-k}$. Wir haben $n = 120$ und für 1,2,3, haben wir $k = 14+16+15 = 45$. Die Wahrscheinlichkeit, höchstens 44mal die Zahlen 1,2,3 zu bekommen, ist also gegeben durch $B(120,0)+B(120,1)+B(120,2)+\ldots+B(120,44)$ (für unsere Werte $p = 0,3$, $q = 0,7$). Das kann man entweder selbst an regnerischen Sommertagen ausrechnen, oder man sieht in einer Tabelle nach und findet den Wert 0,9527. Man bekommt also bei „reinem Zufall“ mit über 95% Wahrscheinlichkeit Anzahlen für 1,2,3 unter 44. Daher ist 45 zu hoch.

Wenn man weiß, was gemessen wurde (es waren sogenannte Stopp-Zeiten), dann ist die Verteilung ebenfalls ganz falsch (es sollte eine „negative Exponentialverteilung“ sein). Kann man weiters noch einiges in Biologie, dann weiß man, dass die Varianz für diese Messungen viel zu groß ist. Die Fälscher brachen unter der Last dieser Indizien zusammen, gaben die Erfindung der Daten zu und reagierten „österreichisch“: sie kündigten die Schwächste — eine Laborantin. Man sieht: Datenfälschung will gelernt sein! Überall lauern Mathematiker...

LITERATUR

• Zum Simpson-Paradox: Julian Havil: *Impossible?*, Princeton Univ. Press, 2011, pp. 11–20.

• Zu den n-dimensionalen Kugeln: Barry Chipra: Disproving the Obvious in Higher Dimensions, in: *What's Happening in the Mathematical Sciences* No. 1 (1993) pp. 21–26. Siehe auch *http://de.wikipedia.org/wiki/Kugel*.

• Zu den Datenfälschungen: Mündliche Überlieferung.

Günter Pilz