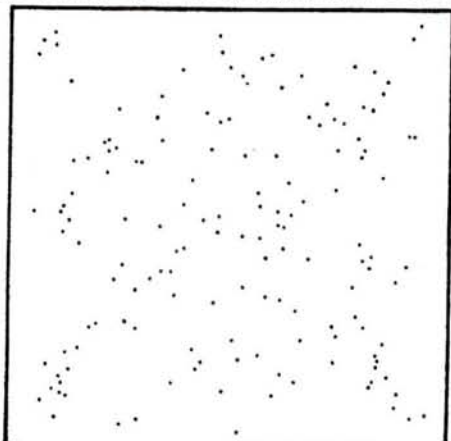


von B. Dudley

Bei der Untersuchung der Ökologie eines Waldes, einer Heide oder einer Wiese ist es üblich, das Gebiet in quadratische Flächenstücke zu unterteilen und in diesen Quadraten die entsprechenden Daten zu sammeln. Stichproben aus so einem Gebiet müssen natürlich zufällig verteilt sein, und Generationen von Biologen warfen zu diesem Zweck mit geschlossenen Augen quadratische Rahmen in das zu untersuchende Gebiet oder warfen einen Stein, um dann ein Quadrat mit dem Auftreffpunkt als Mittelpunkt festzulegen. In jüngerer Zeit werden manchmal Zufallszahlen-Tabellen benutzt, um die Quadrate auszusuchen, und es gibt weitere Zufallsverfahren (cf. z.B. Galloway, Gent, Hutton und Payne, S. 145). Das folgende Experiment beschäftigt sich mit dem Problem der Zufälligkeit bei solchen Stichproben.

Schüler/Studenten werden aufgefordert, auf einem Blatt Papier (DIN A4), in dessen Mittel ein Quadrat mit 9 cm Kantenlänge gezeichnet ist, zehn Kreuze so in das Quadrat einzuzeichnen, daß sie über das ganze Quadrat "zufällig" verteilt sind. Das Experiment wurde vom Autor mit Gruppen von graduierten Biologielehrern, graduierten Mathematiklehrern und Biologie-Studenten gemacht - mit im wesentlichen gleichen Ergebnissen. Der Versuch wurde bei fünf verschiedenen Gelegenheiten durchgeführt, insgesamt mit über 150 Personen. Die hier angegebenen Ergebnisse stammen aus einem Durchführung des Versuchs, und sie sind typisch.

Figur 1. Gesammelte Ergebnisse eines Einzelversuchs mit 15 Studenten. Jeder Punkt ist Kreuzungspunkt eines Kreuzes.



Die Versuchsergebnisse werden zusammengefaßt, indem die Kreuzungspunkte aller Kreuze in einziges Quadrat eingetragen werden (Figur 1). Durch Linien im Abstand 1 cm wird das Quadrat dann in Reihen und Spalten zerlegt, und die Anzahl der Punkte in jeder Reihe und Spalte wird tabellarisch erfaßt. Punkte, die auf eine Linie fallen, werden der Reihe unterhalb bzw. der Spalte rechts von der Linie zugezählt.

Mit dem χ^2 -Test wird für die Verteilung auf Reihen (Figur 2) und Spalten (Figur 3) die Hypothese getestet, daß die Punkte in Figur 1 zufällig über das Quadrat verteilt sind.

Figur 2. χ^2 -Test für die Verteilung auf die Reihen (Daten von 15 Studenten)

Reihennummer	1	2	3	4	5	6	7	8	9
gezählte Punktanz. (f_o)	8	26	14	18	24	15	21	14	10
erwartete Punktanz. (f_e)	16,67	16,67	16,67	16,67	16,67	16,67	16,67	16,67	16,67
$(f_o - f_e)^2 / f_e$	4,51	5,22	0,43	0,11	3,22	0,17	1,13	0,43	2,67

Figur 3. χ^2 -Test für die Verteilung auf die Spalten (Daten von 15 Studenten)

Spaltennummer	1	2	3	4	5	6	7	8	9
gezählte Punktanz. (f_o)	7	25	12	24	19	19	14	21	9
erwartete Punktanz. (f_e)	16,67	16,67	16,67	16,67	16,67	16,67	16,67	16,67	16,67
$(f_o - f_e)^2 / f_e$	5,61	4,16	1,31	3,22	0,33	0,33	0,43	1,13	3,53

Die ermittelten Werte für χ^2 ($\chi^2 = 17,89$ bzw. $\chi^2 = 20,05$) sind größer als 15,51; die Wahrscheinlichkeit für ein solches Ergebnis ist kleiner als 0,05; die Nullhypothese wird verworfen.

Sozialistik in der Schule, Heft 1, Band 3 (1983)

Interpretation dieser Ergebnisse: die Kreuze sind nicht zufällig über das Quadrat verteilt, obwohl es so zu sein scheint. Wenn also das Quadrat ein Wiesengebiet und die Kreuze quadratische Teilgebiete darstellen, so könnte man aus den Daten, die man in diesen Teilgebieten gesammelt hat, mit statistischen Methoden keine gültigen Schlüsse ziehen; denn diese Tests setzen voraus, daß die Stichproben zufällig entnommen worden sind.

Einen großen Beitrag zu den beiden χ^2 -Werten liefern die Reihen bzw. Spalten 1 und 9, und dies in beiden Fällen deshalb, weil die Häufigkeit der Punkte geringer ist als erwartet. Diese Reihen und Spalten bilden den 1 cm breiten Randstreifen des Quadrats, und ein weiterer χ^2 -Test zeigt, daß die Nullhypothese verworfen wird, daß die Punkte auf dem Randstreifen die gleiche relative Häufigkeit haben wie im Inneren (Figur 4). Es ist nicht nur so, daß die Punkte nicht zufällig verteilt sind: der Randstreifen des Quadrats wird gemieden.

Figur 4. χ^2 -Test für die Verteilung auf dem Randstreifen und im Inneren des Quadrats (Daten von 15 Studenten)

Teilgebiet des Quadrats	Randstreifen	Inneres
Flächeninhalt des Teilgebiets	32 cm ²	49 cm ²
gezählte Punktzahl (f_o)	24	126
erwartete Punktzahl (f_e)	$\frac{32}{81} \cdot 150 = 59,26$	$\frac{49}{81} \cdot 150 = 90,74$
$\frac{(f_o - f_e)^2}{f_e}$	20,98	13,70

$$\chi^2 = 44,68$$

Der ermittelte Wert für χ^2 ist größer als 10,8; die Wahrscheinlichkeit für ein solches Ergebnis ist kleiner als 0,001; die Nullhypothese wird verworfen.

In dem inneren Gebiet mit Flächeninhalt 49 cm² befinden sich 126 Punkte. Die Nullhypothese, daß es keine signifikanten Unterschiede in der Häufigkeit gibt, mit der diese Punkte über die 1 cm

breiten Reihen und Spalten im inneren Gebiet verteilt sind, wird nicht verworfen ($\chi^2 = 6,68$ für die Reihen und $\chi^2 = 7,67$ für die Spalten sind kleiner als der kritische Wert von 12,59 - auf 5 %-Niveau mit sechs Freiheitsgraden), obwohl die obere linke und die obere rechte Ecke des inneren Gebiets signifikant mehr Punkte aufweisen, als man bei zufälliger Verteilung erwarten würde ($\chi^2 = 23,6$)

Das Ausgangsproblem war, die Kreuze so in dem Quadrat zu plazieren, daß sie zufällig verteilt sind. Jeder Student war davon überzeugt, dies getan zu haben, und als der Gruppe die gesammelten Ergebnisse gezeigt wurden, waren alle überzeugt, daß diese zufällig verteilt sind. So war das Ergebnis der Untersuchung überraschend, zeigte es doch, daß die Gruppe nicht zufällig funktioniert hatte, obwohl alle es glaubten, und daß man zufällige Stichproben nicht so einfach auf diese Art erhält.

Das Experiment deutet auf die Notwendigkeit hin, bei ökologischen Untersuchungen genau das Gebiet zu definieren, aus dem Stichproben genommen werden; denn sonst ist fraglich, ob der Randstreifen einbezogen wurde, und dies zieht die Gültigkeit jeder statistischen Analyse der Daten in Zweifel. Wo es keine natürliche Grenze gibt, etwa bei Untersuchung eines Gebiets innerhalb eines Waldes, einer Heide oder einer Wiese, muß man das Gebiet abstecken, bevor man beginnt, Proben zu nehmen. Gibt es natürliche Grenzen - z.B. eine Hecke, einen Bach oder einen Weg -, so ist es wichtig, den Randstreifen mit einzubeziehen, indem man bis zum Rand hin Proben entnimmt, allerdings nicht häufiger als es mit der Entnahme einer zufälligen Stichprobe aus dem ganzen Gebiet vereinbar ist; denn natürliche Grenzen haben einen besonderen Einfluß auf die Verteilung von Pflanzenarten, bilden oft einen Blickfang und könnten daher bevorzugt oder gemieden werden. Häufig findet man entlang einer Hecke oder an einem Bach Arten, die es weiter entfernt davon nicht gibt, und oft wachsen unmittelbar links und rechts von einem Weg durch eine Wiese Arten, die man auf der Wiese selbst nicht findet. Sollten diese weder übersehen noch bevorzugt werden, soll also die Stichprobe repräsentativ sein, so müssen aus dem ganzen Gebiet zufällig Proben entnommen werden.

Menschen funktionieren nicht zufällig, selbst wenn sie es versuchen; die Entnahme einer Stichprobe muß hohen Ansprüchen genügen. Deshalb sollte man das Gebiet abstecken und dann Zufallszahlen benutzen, um die Koordinaten jeder Probe festzulegen.

Literatur

D.Galloway, A.M.Gent, J.Hutton and V.Payne, Maths for scientists. Mills and Boon 1980.

Originaltitel in 'TEACHING STATISTICS' (1982) Vol.4 / Nr. 2:
Exploring a Random Sampling Method used in Biology

Übersetzung und Bearbeitung: R.Oselies