

SCHÜLEREXPERIMENTE ZUM VORZEICHENTEST

von Mary Rouncefield und Peter Holmes

Originaltitel in Teaching Statistics Vol 11 (1989) No 1:

Practical Hypothesis Testing using the Sign Test

Bearbeitung: Manfred Borovcnik, Klagenfurt

Kurzfassung:

Die Frage "Werden Leute mit zunehmender Übung besser?" kann anhand erfrischender Experimente im Unterricht behandelt werden. An Voraussetzung ist nur die Binomialverteilung nötig. Der Vorzeichentest kann wegen der geringeren Schwierigkeiten im Hinblick auf die erforderlichen Verteilungen auch als Einstieg in die indirekte Schlußweise des Signifikanztests benützt werden.

1. Experimente

Reaktionszeit

Ziel dieses Experiments ist es, die Reaktionszeit von Schülerinnen und Schülern vor und nach einer Übungsphase wie folgt zu testen: Ein Schüler hält ein Lineal von 30cm Länge senkrecht mit Daumen und Zeigefinger an der 30cm-Marke. Ein zweiter Schüler ist mit seinem Daumen und Zeigefinger in Lauerstellung über der 0cm-Marke, er berührt das Lineal jedoch nicht. Innerhalb einer Zeit von 20 Sekunden vom Start kann der erste Schüler den Zeitpunkt, in dem er das Lineal losläßt, frei wählen. Sowie er das Lineal fallen gelassen hat, versucht der zweite, es aufzufangen. Die Distanz, die das Lineal frei gefallen ist, kann vom Lineal abgelesen werden. Daraus erhält man entweder durch Umrechnung nach folgender Tabelle oder durch Ablesen an einer Vergleichsskala die benötigte Reaktionszeit.

Zeit in 1/100 sec	5	6	7	8	9	10	11	12	13
Distanz in mm	12.3	17.6	24	31.4	40	49	59	71	83
	14	15	16	17	18	19	20	21	22
	96	110	125	142	159	177	196	216	237
								259	282

Diese Tabelle kann man sich auch auf einem Vergleichslineal einzeichnen und dieses zur Schnellablesung der Reaktionszeit verwenden.

Sortieren von Münzen

Folgende Münzen werden auf einem Tablett durcheinander gemischt: 4 Pfennigstücke, 4 Zwei-Pfennig-Stücke, 4 Groschen, 4 Knöpfe, 4 Automatenmarken. Dem "Prüfling" werden die Augen verbunden. Notiert wird die Zeit, die er/sie benötigt, bis alle Münzen korrekt in fünf Haufen getrennt sind. Das Experiment wird wiederholt. Im zweiten Versuch sollten die notierten Zeiten kürzer sein. Die Schüler sollten keine Gelegenheit haben, bei den Versuchen der anderen zuzuschauen, sie würden dabei zuviel an Strategie lernen, ehe sie selbst an die Reihe kommen.

Figuren nachzeichnen

Eine Figur mit einfachem Umriß wird sechsmal auf ein Blatt Papier gedruckt. Die Probanden müssen alle sechs Figuren mit einem Stift nachfahren. Dabei müssen sie ihre nicht-dominante Hand benutzen (Rechtshänder benutzen die linke Hand). Gemessen wird die Zeit, die sie für ihre erste bzw. für ihre letzte Figur benötigt haben. Der sechste Versuch sollte schneller als der erste sein.

2. Hypothesen

Für jedes der angeführten Experimente ist die Nullhypothese, H_0 , Übung hat keinerlei Auswirkung auf das Abschneiden einer Person, in dem Fall, auf die Zeit, die sie benötigt, bis die Aufgabe bewältigt ist. Die Zahl derer, die besser abschneiden nach einiger Übung, sollte sich mit der, die sich verschlechtern, in etwa aufheben. Daher ist, im Falle des Zutreffens der Nullhypothese, die Wahrscheinlichkeit für eine Verbesserung (Verschlechterung) gleich $\frac{1}{2}$.

Die Alternativhypothese, H_1 , ist, daß Übung eine Verbesserung zur Folge hat. Diese zwei Hypothesen werden so ange-

schrieben:

$H_0: p = \frac{1}{2}$
 $H_1: p > \frac{1}{2}$,

wobei p gleich der Wahrscheinlichkeit der Verbesserung ist. Die folgende Tabelle enthält die Ergebnisse von acht Probanden im Figuren-Nachzeichnen.

Zeit vor Übung (1. Figur)	9	7	20	16	9	15	11	16
Zeit nach Übung (2. Figur) in Sekunden	7	5	15	13	16	10	9	16

Die tatsächliche Veränderung der Reaktionszeit wird vergrößert in folgende Kategorien:

Verbesserung	6
Verschlechterung	1
Keine Änderung	1

Unter der Nullhypothese sollten sich die Fälle von Verbesserung bzw. Verschlechterung gegenseitig in der Zahl aufheben; die Kategorie "Keine Änderung" verändert diese Balance nicht und wird daher in der weiteren Analyse weggelassen. Die Zahl der Leute, die eine Veränderung der benötigten Zeit aufweisen, beträgt 7, unser Hypothesentest wird auf deren Ergebnissen basieren.

Unter der Nullhypothese folgt die Verteilung positiver (bzw. negativer) Differenzen einer Binomialverteilung $B(7, \frac{1}{2})$, das ist, als ob man mit einem siebenfachen Münzwurf die Richtung der Unterschiede der Leute ermittelt hätte. Falls man vom bestehenden Einfluß der Wirksamkeit der Übung spricht, sollte man sich gegen diesen Fall absichern.

Die Teststatistik ist die kleinere Zahl der positiven bzw. negativen Unterschiede. In diesem Fall haben wir einen negativen Unterschied, und wir möchten die Wahrscheinlichkeit berechnen, daß eine solche oder noch kleinere Zahl beobachtet wird unter den Bedingungen der H_0 .

3. Der Signifikanztest

Das mathematische Modell für die Situation unter der Nullhypothese ist die Binomialverteilung $B(7, \frac{1}{2})$. Die Anzahl negativer (wie positiver) Differenzen folgt unter H_0 gerade dieser Verteilung. Damit können wir für jede Anzahl negativer Differenzen ihre Wahrscheinlichkeit ausrechnen.

Anzahl negativer Differenzen	Wahrscheinlichkeit
0	0.008
1	0.055
2	0.164
3	0.273
4	0.273
5	0.164
6	0.055
7	0.008

Aus dieser Tabelle ergibt sich die Wahrscheinlichkeit von 0.063 dafür, daß die Zahl negativer Differenzen kleiner oder gleich 1 ist

Ist diese Wahrscheinlichkeit ausreichend klein, damit wir die Nullhypothese verwerfen können? Sollten wir die Behauptung, daß $p = \frac{1}{2}$ ist, aufrecht erhalten? Immerhin, falls wir das Experiment mit 100 verschiedenen Gruppen von je sieben Leuten wiederholen, dann würden wir ähnlich extreme Ergebnisse in lediglich 6 Fällen (in etwa) erhalten. Oder sollten wir das Resultat für ausreichend ungewöhnlich befinden, sodaß wir uns für die Möglichkeit entscheiden, daß p größer als $\frac{1}{2}$ sein muß? Wenn wir uns an die konventionelle Ablehnregel mit 5% oder weniger halten, dann dürften wir im Experiment keine einzige Verschlechterung haben, ehe wir H_0 ablehnen können.

4. Der Vorzeichentest

Der Vorzeichentest, wie das beschriebene Verfahren heißt, ist anwendbar, wenn die Ergebnisse gepaart sind, wie hier im Ergebnis ein und derselben Person vor und nach der Übung. Die tatsächliche Differenz spielt keine Rolle, sie wird lediglich mit positiv ("Verbesserung") oder negativ ("Verschlechterung") eingestuft. Das Verfahren gestattet daher auch keine Beurteilung, um wieviel man besser oder schlechter wird. Es prüft lediglich, ob es sinnvoll ist, davon auszugehen, daß Unterschiede bestehen. Weil man die Situation stark vergrößert hat, verwendet man auch kaum Annahmen, so z.B. ist die Verteilung der Merkmale (in den Beispielen die Reaktionszeiten) völlig egal. Es muß lediglich gesichert sein, daß sie in den beiden zu vergleichenden Situationen dieselbe Gestalt hat. Verfahren, die nicht auf die spezielle Gestalt der Verteilung der Variablen zurückgreifen, heißen nicht-parametrische Verfahren.

Weil man weniger Annahmen verwendet, kann man dabei auch weniger Gefahr laufen, Fehler zu machen. Darüberhinaus ist der Vorzeichentest auch begrifflich einfacher, weil neben die Schwierigkeiten mit der indirekten Schlußweise des Signifikanztests keine weiteren mit der Art der Verteilung

treten. Die Situation unter der Nullhypothese ist wie gewöhnliches Münzwerfen. Ein weiterer Vorteil liegt darin, daß in den praktischen Experimenten die Formulierung von Null- und Alternativhypothese direkt verständlich ist.

LITERATUR

ROUNCEFIELD, M. und HOLMES, P.: *Practical Statistics*. Macmillan, 1989
 INEICHEN, R.: *Zufällig oder nicht zufällig?* In: *Der Mathematik-Unterricht* 29 (1983), S. 24-40
 LORENZ, R. J.: *Grundbegriffe der Biometrie*. G. Fischer, Stuttgart, 1988

Anmerkung: Im Abschnitt 4 hat der Bearbeiter einige allgemeine Hinweise zum Vorzeichentest eingefügt.