

UNEHELICH GEBORENE IN EXETER

nach F. LEWES, University of Exeter

Originaltitel in 'Teaching Statistics' Vol. 6 (1984), Nr. 2:

Exeter Bastards

Übersetzung: H. Gundel

Zusammenfassung

Anhand der Auswertung von Daten zur Unehelichkeit von Neugeborenen wird die Argumentation und Logik beim Testen von Hypothesen dargestellt.

Kürzlich betonte Ian Birnbaum (Birnbaum, 1982), wie wichtig es ist, daß Schüler die Logik des Hypothesentestens verstehen und nicht bloß erfahren, wie man zu Signifikanzniveaus kommt. Im folgenden wird der Versuch unternommen, diese Notwendigkeit in einem neuen Zusammenhang herauszustellen und von einem anderen Standpunkt aus zu begründen.

Als Beispiel benutze ich einige Daten, die ich zufällig beim Stöbern in den "Local Authority Vital Statistics" gefunden habe, dabei ist die Auswahl von Exeter gezielt und nicht zufällig, denn ich lebe dort.

Geburten in Exeter nach Geschlecht und Ehelichkeit, 1980

	männlich	weiblich	Summe
ehelich	566	547	1113
unehelich	54	84	138
			1251

Wir betrachten zuerst den Anteil unehelicher Geburten. Im Jahre 1980 waren in England und Wales 13,47 % aller lebend Geborenen unehelich. Wenn man in Exeter wohnt, erwartet man dort einen geringeren Anteil. Also liegt ein einseitiger Test nahe, wenn man die Nullhypothese prüfen will, daß der Anteil in Exeter derselbe ist wie im Land insgesamt. Tatsächlich war der Anteil in Exeter 11,03 %; das entspricht

Stochastik in der Schule Heft 1, 5. Jg. (1983)

unter der Nullhypothese einer Abweichung um 2,53 Standardfehler¹⁾ in der erwarteten Richtung. Selbstzufrieden lehnen wir die Nullhypothese ab und halten für wahr, daß der geringere Unehelichenanteil in Exeter keine zufallsbedingte Begebenheit ist. Eine Reihe von Erklärungen drängen sich auf: Tugendhaftigkeit, Vorsicht, Weisheit, die soziale Bevölkerungsstruktur oder der Verheiratetenanteil. Aber wollten wir uns für eine der Ursachen entscheiden, so müßten wir raten. Der entscheidende Punkt ist unsere Reaktion auf solche Zahlen. Eine so große Abweichung wie diese ist unter der Nullhypothese selten, deshalb lehnen wir die Nullhypothese ab.

Fairerweise müssen wir nun das Geschlechterverhältnis bei diesen Kindern betrachten. Im Jahr 1980 betrug in England und Wales der Prozentsatz männlicher Neugeborener 51,17 %. Dagegen gab es in Exeter nur 54 Jungen unter 138 unehelich Geborenen, das sind 39,13 %. Unter der Nullhypothese bedeutet dies eine Abweichung von 2,83 Standardfehlern, die sogar größer als beim Unehelichenanteil ist. Folgt man der Logik des ersten Tests, so muß man die Nullhypothese, daß das Geschlechterverhältnis bei unehelich Geborenen in Exeter genauso ist wie im ganzen Land, ablehnen. Um Himmels willen! Irgendjemand oder irgendetwas verfälscht in Exeter das Geschlechterverhältnis bei den unehelich Geborenen.

Nun ist die Unveränderlichkeit des Geschlechterverhältnisses bei Neugeborenen in der ganzen Welt seit langem bekannt. Die Konstanz ist derart, daß Florence Nightingale ihre Erhaltung der "Göttlichen Hand" zuschrieb; und, soweit ich weiß, hat bisher niemand eine bessere Erklärung angeboten. Angesichts der möglichen Schlußfolgerung aus unserer Entdeckung wird man zögernd. Aus diesem Grund würden die meisten von uns die Nullhypothese in der Praxis nicht ablehnen - es sei denn, man würde etwa nach einem vermuteten Zusammenbruch des Meldesystems suchen-, sondern man würde eher annehmen, daß etwas Unwahrscheinliches, aber Mögliches eingetreten ist. Was bedeu-

¹⁾ Mit Standardfehler ist die Standardabweichung σ_p der relativen Häufigkeit (des prozentualen Anteils) bezeichnet, und zwar gilt $\sigma_p = \sqrt{\frac{p(1-p)}{n}}$, wobei n der Stichprobenumfang (hier n = 1251) und p der Anteil (hier 13,47 % unter der Nullhypothese) ist.

tet es nun für den ersten Fall, d.h. für die Gesamtzahl unehelich geborener Kinder, wenn man so schließt? Sind Signifikanztests willkürlich, wird nach Gutdünken auf Niveaus, die unserer Laune entspringen, angenommen oder abgelehnt? Aus welchem Grund ist man berechtigt anzunehmen, daß nichts geschehen ist, was den Jungenanteil vermindert, während man gleichzeitig die Ehre der Mütter aus Exeter verteidigt?

Nun ist der interessante Aspekt eines Tests dieser Art nicht die dahinterstehende Mathematik, die ja trivial ist, sondern die sehr komplexe Argumentation bei der Anwendung des Tests. Entsprechend den Anwendungssituationen wählen wir zu Recht unterschiedliche Signifikanzniveaus. Wenn die Alternativhypothese vernünftig und wahrscheinlich ist, wie im ersten Fall, akzeptieren wir gerne ihre Bestätigung. Dies ist einem bayesianischen Schluß ähnlich, obwohl Grade des Vertrauens keine Wahrscheinlichkeiten sind, worauf Birnbaum hinweist. Aber auch wenn Signifikanztests nicht mit a-priori-Wahrscheinlichkeiten oder Glauben verbunden werden können, um zu a-posteriori-Wahrscheinlichkeiten zu gelangen, so wird doch nach einem ähnlichen Prinzip vorgegangen. Jeder Jurist wird bestätigen, daß an sich Unwahrscheinliches ein höheres Beweisniveau erfordert. Man könnte folgendes hinzufügen: Wenn man die Nullhypothese ablehnt, ohne eine Alternativhypothese zu spezifizieren, dann muß man wenigstens sicher sein, daß es eine solche geben könnte.

Statistik ist nicht nur das Rechnen mit Wahrscheinlichkeiten. Sie umfaßt auch das logische System, das in Situationen mit Unsicherheiten benötigt wird. Die Bücher über Elementare Statistik haben sich weiterentwickelt, und es ist üblicher geworden, Fehler erster Art und Fehler zweiter Art anzusprechen, obwohl Fehler zweiter Art eher knapp behandelt werden. Aber es gibt auch noch Bücher, die andeuten, daß jede Abweichung, die auf dem 5 %-Niveau signifikant ist, als real anzusehen ist - trotz der Anzahl falscher Aussagen, die dies zur Folge haben muß.

Wie schon Benjamin Franklin betonte, ist nichts auf der Welt sicher außer dem Tod und Steuern. Dies trifft ganz besonders und unabweisbar auf Ergebnisse statistischer Tests zu. Dies mag unbequem sein,

vielleicht besonders für diejenigen von uns, deren mathematischer Hintergrund sie mit einzelnen richtigen Antworten, die mit einem roten Stift abgehakt werden können, glücklicher werden läßt. Aber man darf der Logik der Statistik nicht dadurch ausweichen, daß man sich in das Dickicht seiner Mathematik zurückzieht. Manchmal gibt es viel zu diskutieren. Mehr noch, in vielen Fällen beginnt es interessant zu werden, wenn die Mathematik am Ende ist. Dies gilt sicherlich für das dargestellte Beispiel.

Literatur

BIRNBAUM, I. (1982). Interpreting statistical significance.
Teaching Statistics, 4 (1)