Glossar: Testtheorie / Hypothesentest

Testtheorie, Testen von Hypothesen [Stochastik, Wahrscheinlichkeitsrechnung, Beurteilende Statistik]

Es geht darum, eine Hypothese (die sogenannte Nullhypothese H_0) statistisch zu wiederlegen.
Das klingt seltsam, aber genau zu diesem Zweck hat man in der Regel die Nullhypothese aufgestellt.

Der Test ist also nur erfolgreich (und wird leider häufig auch nur dann veröffentlicht), wenn die Nullhypothese H_0 auf Grund des Testergebnisses abgelehnt werden kann.

Bei dieser Ablehnung soll eine im Vorhinein gesetzte Grenze für die Irrtumswahrscheinlichkeit α nicht überschritten werden.
Die gewählte Grenze heißt Signifikanzniveau.
Der entsprechende Fehler (fälschliches Ablehnen der Nullhypothese) heißt Fehler 1. Art (oder α-Fehler).

In der Regel sind (in der Aufgabenstellung) zwei Entscheidungen vorgegeben:
die *Anzahl der Einzelversuche* n (Stichprobengröße) und das Signifikanzniveau.

Zur Anzahl der Versuche: Ist die Anzahl der Versuche zu klein, kommt unter Umständen „nichts raus", denn wenn es nicht gelingt, H_0 statistisch zu widerlegen, hat man nichts in der Hand (und darf sich auf gar keinen Fall einbilden, nun hätte man H_0 stattdessen statistisch bestätigt.)

Zum Signifikanzniveau: Dies entspricht der Entscheidung, was „dann doch kein Zufall mehr sein kann". Man legt die Obergrenze für den genannten Fehler 1. Art fest.
(Häufig wird die Irrtumswahrscheinlichkeit 1. Art α mit dem Signifikanzniveau 0,05 gedeckelt, aber es werden auch andere Signifikanzniveaus verwendet, wie z.B. $\propto \leq 0{,}01$)

Durch ein Experiment – also indem man den Würfel immer wieder wirft – kann man dann (statistisch) zeigen, dass der Würfel gefälscht ist.

Es handelt sich um Anwendungen der Wahrscheinlichkeits-rechnung – daher kann man Irrtümer nicht ganz ausschließen, sondern das Entscheidungsverfahren bestenfalls so anlegen, dass sie sehr unwahrscheinlich sind.

Schritt 1: Überlegen, was man bezüglich p zeigen will – und auf dieser Basis H_1 festlegen.
Dann legt man als Gegenbehauptung dazu die Nullhypothese H_0 festlegen. .

Beispiel: Bei einem Würfelspiel hat ein Zocker seine Millionen-Dollar-Yacht verloren. Nun vermutet er, dass der Würfel seines Gegenspielers gefälscht war.
In einem Gerichtsprozess wird ein Gutachter beauftragt, den Würfel zu überprüfen. Materialtechnische Untersuchungen ergeben keine Unregelmäßigkeiten.
Der Gutachter plant nun, eine Versuchsreihe mit dem Würfel durchzuführen.
X: Anzahl der Sechsen im Versuch.
Was muss er entscheiden und wann wird er in seinem Gutachten den Würfel für gefälscht erklären?

LORE LORENTZ SCHULE

Wenn man also z.B. zeigen will, dass ein Würfel gefälscht ist und häufiger eine Sechs würfelt also normal,

so ist das die Vermutung $H_1: p > \frac{1}{6}$

Man will zeigen, dass die Sechs häufig fällt (X nimmt große Werte an, der Ablehnungsbereich liegt dann rechts), es handelt sich dann um einen rechtsseitigen Test.

Die Nullhypothese ist dann, die Verneinung dessen (das Gegenereignis) $p \leq \frac{1}{6}$. Man rechnet dann mit dem Grenzfall $p \ = \frac{1}{6}$ (auch das wird häufig statt $p \leq \frac{1}{6}$ als Nullhypothese bezeichnet).

Der Gutachter will den Betrug durch 1000-faches Würfeln auf dem Signifikanzniveau 0,01 zeigen:

n= 1000, $\propto \leq 0{,}01$

Schritt 2 ist die eigentliche mathematische Arbeit:

Aus H_0, n und dem Signifikanzniveau wird nun der Ablehnungsbereich bestimmt.

Da man die Einzelversuche unabhängig voneinander immer mit der gleichen Trefferwahrscheinlichkeit p durchführt, benutzt man dazu die Binomialverteilung. Da die wahre Trefferwahrscheinlichkeit p unbekannt ist, rechnet man mit dem p aus der Nullhypothese (hier: $p = \frac{1}{6}$).

(Bei der Nullhypothese $p \ = \frac{1}{6}$, rechnet man mit $p \ = \frac{1}{6}$ (klar!),

Bei der Nullhypothese $p \geq \frac{1}{6}$ und bei der $p \leq \frac{1}{6}$ aber *ebenfalls* mit $p \ = \frac{1}{6}$, da dies der Grenzfall ist. Womit soll man dann auch sonst rechnen?)

Bei einem rechtsseitigen Test ist der Ablehnungsbereich

$$A = \{k;\ \dots;\ n\}$$

Gesucht ist die kleinste Zahl k für die $P(X \geq k)$ kleiner gleich dem Signifikanzniveau ist.
Im Beispiel ist also ein k gesucht, so dass $P(X \geq k) \leq 0{,}01$.

Man findet es z.B., indem an mit dem Taschenrechner oder CAS Zahlen zwischen dem Erwartungswert 166,7 und 1000 für k einsetzt und die Wahrscheinlichkeit berechnet (z.B. mit binomcdf(1000,1/6, 200, 1000))

Erst dann sollte in **Schritt 3** das Experiment durchgeführt und ausgewertet werden.
Ist das Ergebnis im Ablehnungsbereich, so ist H_0 statistisch widerlegt (auf dem entsprechenden Signifikanzniveau).
Andernfalls ist H_0 aber keineswegs bestätigt. Das Ergebnis ist dann zwar mit H_0 vereinbar, aber in der Regel genauso mit H_1.

Fehler:
Es können zwei verschiedene Fehler auftreten:
Fehler 1. Art: Die Nullhypothese wird fälschlicherweise abgelehnt. D.h., das Ergebnis lag zufällig im Ablehnungsbereich, und man hat einen „statistischen Nachweis" geführt für etwas, was eigentlich falsch ist.
Die Wahrscheinlichkeit, dass dies eintritt, wird durch das Signifikanzniveau gedeckelt.
Fehler 2. Art: Die Nullhypothese wird fälschlicherweise nicht abgelehnt. D.h., das, was man zeigen wollte, stimmt, aber der statistische Nachweis ist nicht gelungen.

Links: Checkliste
umfangreiches Lernmaterial: Strobl: Grundlagen Übungen Lsg
http://raschweb.de/Q12-M-Hypothesentest-Abituraufgaben.pdf