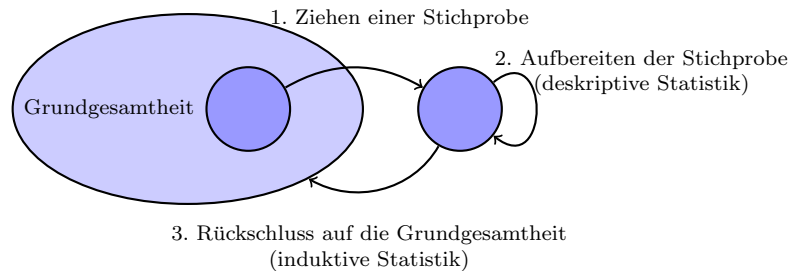


Hypothesentests für die Maturaarbeit

Das Gesamtbild (Ineichen & Stocker, 1996)



Beispiel

Aline vermutet, dass 16-jährige Mitschülerinnen und Mittelschüler an deutschschweizer Gymnasien lieber Englisch als Französisch lernen.

Dies führt zu zwei komplementären Hypothesen (Vermutungen):

- Die *Alternativhypothese* H_1 : Der Anteil der 16-jährigen Gymnasiast(inn)en, die lieber Englisch als Französisch lernen, ist grösser als 0.5.
- Die *Nullhypothese* H_0 : Der Anteil der 16-jährigen Gymnasiast(inn)en, die die lieber Englisch als Französisch lernen, ist kleiner oder gleich 0.5.

Die Alternativhypothese

Die Alternativhypothese H_1 stellt (oft) eine „innovative Behauptung“ dar. H_1 beschreibt einen Zustand, den die Wissenschaftlerin oder der Wissenschaftler nachweisen möchte.

Im Beispiel fragt sich Aline, ob die 16-jährigen Mittelschüler lieber Englisch als Französisch lernen? Dies führt zu einem weiteren Begriffspaar:

- Eine Alternativhypothese ist *gerichtet*, wenn die Richtung eines Unterschieds vorgegeben wird.
- Eine Alternativhypothese ist *ungerichtet*, wenn nur ein Unterschied aber nicht seine Richtung von Interesse ist.

Die Nullhypothese

- Die Nullhypothese H_0 stellt den aktuellen Wissensstand dar.
- Die Nullhypothese behauptet, dass der in der Alternativhypothese postulierte Unterschied bzw. Zusammenhang nicht vorhanden ist.
- Die Nullhypothese ist komplementär zur Alternativhypothese.

Aufgabe 1

Formuliere die Nullhypothese zur folgenden Alternativhypothese:

H_1 : Frauen besitzen durchschnittlich mehr Paar Schuhe als Männer.

H_0 : Frauen besitzen durchschnittlich gleich viele oder weniger Paar Schuhe als Männer.

Gib an, ob H_1 gerichtet oder ungerichtet ist.

H_1 ist gerichtet

Aufgabe 2

Formuliere die Null- und die Alternativhypothese zur folgenden Fragestellung und gib an, ob H_1 gerichtet oder ungerichtet ist.

Ist Kurzsichtigkeit abhängig vom Geschlecht?

H_0 : Die Anteile der kurzsichtigen Männer und Frauen unterscheiden sich nicht.

H_1 : Die Anteile der kurzsichtigen Männer und Frauen unterscheiden sich. (ungerichtet)

Aufgabe 3

Formuliere die Null- und die Alternativhypothese zur folgenden Fragestellung und gib an, ob H_1 gerichtet oder ungerichtet ist.

Wird der Politiker X bei der kommenden Wahl mehr als 60% der Stimmen erhalten?

H_0 : Der Anteil der Wähler, die X wählen, ist $\leq 60\%$.

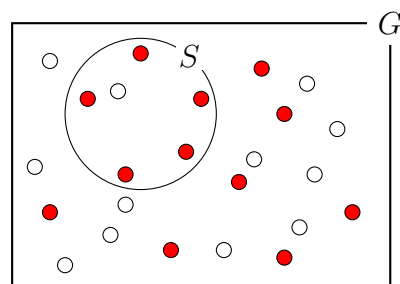
H_1 : Der Anteil der Wähler, die X wählen, ist $> 60\%$. (gerichtet)

Die Stichprobe

Da Aline nicht alle gleichaltrigen Schweizer Gymnasiasten befragen kann, muss sie sich auf eine Stichprobe beschränken.

Idealerweise sollte ihre Stichprobe die Verhältnisse in der Grundgesamtheit abbilden.

Die im folgenden Bild dargestellte Stichprobe S liefert ein verzerrtes Bild der Grundgesamtheit.



Gründe:

- Die Elemente der Stichprobe wurden *nicht unabhängig* ausgewählt. Dies ist beispielsweise dann der Fall, wenn Kollegen mit ähnlichen Interessen oder Haltungen befragt werden. Diese Verzerrung lässt sich durch eine Zufallsauswahl reduzieren.
- Die Elemente der Stichprobe wurden unabhängig ausgewählt aber dennoch erhält man ein verzerrtes Bild. Dagegen kann man zwar nichts machen aber die Statistik kann eine Antwort darauf geben, wie gross die Wahrscheinlichkeit („Chance“) ist, dass eine solche Situation eintritt.

Aline wählt also in ihrer Schule aus 6 Klassen zufällig ein Mädchen und einen Jungen aus, die sie befragt, ob sie lieber Englisch oder Französisch lernen.

lernt lieber Englisch als Französisch	Anzahl
Ja	9
Nein	3

Nun geht es darum herauszufinden, ob das Ergebnis der Stichprobe mit der Nullhypothese vereinbar ist oder nicht.

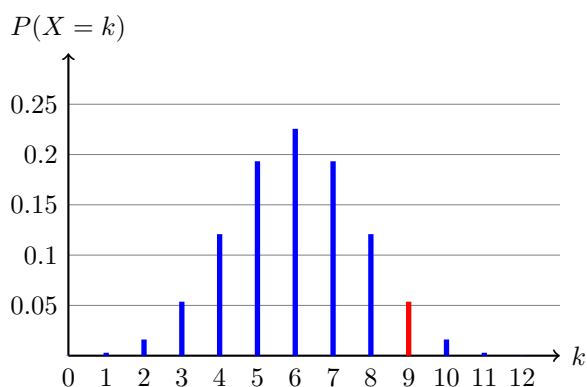
Bei kleineren Abweichungen gehen wir davon aus, dass sie zufälliger Natur sind. In diesem Fall werden wir die Nullhypothese beibehalten.

Je grösser die Abweichungen sind, desto eher werden wir dazu tendieren die Alternativhypothese als Grund für den Unterschied anzusehen.

Ist das Stichprobenergebnis zu „extrem“, sind wir bereit, die Nullhypothese zu Gunsten der Alternativhypothese zu verwerfen.

Um herauszufinden, wie (un)wahrscheinlich ein Stichprobenergebnis ist, muss man wissen, wie alle möglichen Stichprobenergebnisse verteilt sind, wenn man H_0 (vorläufig) als gültig voraussetzt.

Im Fall von Beispiel 1 mit $n = 12$ befragten Schülern könnte man beispielsweise 12 faire Münzen werfen und zählen wie oft jedes Mal Kopf erscheint; also jemand Englisch bevorzugt. Wiederholt man dieses Experiment sehr oft und bestimmt jeweils die relative Häufigkeiten für k Mal Kopf, so erhält man näherungsweise die gesuchte Verteilung (Monte Carlo-Methode). In der Regel lässt sich diese Verteilung auch rechnerisch bestimmen und man erhält dann folgendes Bild:



Der p -Wert

Nun können wir angeben, wie wahrscheinlich das Stichprobenresultat oder ein noch extremeres in Richtung der Alternativhypothese ist.

Dazu addieren wir die relative Häufigkeiten für 9–12 Mal Kopf. Diese Summe beträgt $p = 0.073$ und wird p -Wert genannt. Sie bedeutet anschaulich, dass in etwa 73 von 1000 Versuchen 9–12 Mal Kopf erscheint.

Damit erhalten wir ein objektives Mass dafür, wie wahrscheinlich das vorliegende Stichprobenresultat ist, wenn man davon ausgeht, dass die Nullhypothese gültig ist.

Das Signifikanzniveau α

Es ist bei Hypothesentests üblich, dass man die Nullhypothese verwirft, wenn das in der Stichprobe beobachtete Resultat oder ein extremeres (in Richtung von H_1) in höchstens 5% aller Fälle eintritt.

Diese willkürliche und im Voraus gewählte Grenze wird *Signifikanzniveau* genannt, mit α bezeichnet und muss bei der Diskussion der Ergebnisse jeweils angegeben werden.

Der Fehler 1. Art

Wenn eine Stichprobe einen p -Wert hat, der kleiner oder gleich 5% ist, dann verwerfen wir die Nullhypothese zu Gunsten der Alternativhypothese.

Es könnte aber dennoch sein, dass die Alternativhypothese falsch ist und das beobachtete Ergebnis durch Zufall zustande gekommen ist. In diesem Falle würden wir uns aber nur in 5 von 100 Fällen irren, weshalb α auch *Irrtumswahrscheinlichkeit* oder *Fehler 1. Art* genannt wird.

Dort wo ein Fehler 1. Art besonders weitreichende Konsequenzen hat, wählt man als Signifikanzniveau $\alpha = 1\%$ oder sogar $\alpha = 0.1\%$.

Gleichzeitig führt eine Senkung des Fehlers 1. Art dazu, dass es wahrscheinlicher wird, dass die Nullhypothese beibehalten wird. Damit steigt auch die Wahrscheinlichkeit eines *Fehlers 2. Art*, bei dem die Nullhypothese beibehalten wird, obwohl die Alternativhypothese richtig wäre.

Die Berechnung des Fehlers 2. Art geht aber über den Stoff dieses Workshops hinaus.

Zusammenfassung

Ein Hypothesentest ist eine Regel, die es erlaubt, für jedes Stichprobenergebnis eine Entscheidung zwischen der Null- und der Alternativhypothese zu treffen.

Dazu wählt man im Voraus das Signifikanzniveau (üblich ist $\alpha = 5\%$), ein geeignetes Testverfahren (dazu später mehr) und berechnet mit dem Test den p -Wert der Stichprobe.

- ist der p -Wert kleiner oder gleich α , wird die Nullhypothese *verworfen*.
- ist der p -Wert grösser als α , wird die Nullhypothese *beibehalten*.

Verboten

- Manipulieren von Zahlen
- Das Signifikanzniveau α nach dem Test „anpassen“, um ein signifikantes Resultat zu bekommen.
- So viele Male eine Stichprobe ziehen, bis ein „signifikantes“ Resultat entsteht.

Vorsicht

Signifikanztests sind kein „Beweis“ für die absolute Gültigkeit einer Hypothese:

- Wenn man die Nullhypothese verwirft bedeutet das nicht, dass die Alternativhypothese richtig sein muss. Das Testergebnis stellt bloss ein Indiz für deren Richtigkeit dar.
- Wenn man die Nullhypothese beibehält bedeutet das nicht, dass die Alternativhypothese falsch sein muss. Das Testergebnis ist nur nicht „stark genug“, damit man sich für die Alternativhypothese ausspricht.