

Hypothesentests

Allgemein: Ein Hypothesentest verfolgt das Ziel, die aufgestellte Hypothese auf ihre Gültigkeit zu testen. Für diesen Test muss eine Nullhypothese und eine Alternativhypothese aufgestellt werden, welche sich gegenseitig ausschließen. Die Nullhypothese wird nur abgelehnt, wenn ausreichend Evidenz gegen sie vorliegt.

Parametrischer Test: Falls wir beim Hypothesentest Aussagen bezüglich einem Parameter ψ eines Merkmals der Population treffen, sprechen wir von einem parametrischen Test.

- ▶ **Nullhypothese H_0 :** Diese Hypothese wird getestet und stellt in der Regel die Annahme dar, dass es keinen Effekt oder keinen Unterschied gibt. Ein Beispiel für eine Nullhypothese wäre: "Der Mittelwert ist gleich 100."
- ▶ **Alternativhypothese H_1 :** Sie widerspricht der Nullhypothese und besagt, dass es einen Effekt bzw. einen Unterschied gibt. Zum Beispiel: "Der Mittelwert ist ungleich 100."
- ▶ **Signifikanzniveau α :** Das Signifikanzniveau, oft mit 0.05 festgelegt, beschreibt die maximale Wahrscheinlichkeit, die Nullhypothese fälschlicherweise zu verwerfen. Ein Signifikanzniveau von 0.05 bedeutet, dass wir bereit sind, in maximal 5 % der Fälle eine falsche Entscheidung zu treffen.
- ▶ **p-Wert:** Dieser Wert beschreibt die Wahrscheinlichkeit, unter der Annahme von H_0 , ein Ergebnis zu erhalten, das mindestens so extrem ist, wie das beobachtete Ergebnis. Demnach wird der p -Wert durch die Stichprobe bestimmt. Ein niedriger p -Wert (z.B. kleiner als 0.05) spricht dafür, die Nullhypothese abzulehnen. Daher wird die Nullhypothese verworfen, wenn der p -Wert $\leq \alpha$ ist.

Einfache Hypothese: Wir sprechen von einer einfachen Hypothese, falls die Hypothese für den Parameter ψ aus einem Wert besteht. Zum Beispiel: $\psi = \psi_0$.

Zusammengesetzte Hypothese: Wir sprechen von einer zusammengesetzten Hypothese, falls die Hypothese für den Parameter ψ aus mehreren Werten besteht. Zum Beispiel: $\psi \leq \psi_0$, $\psi \geq \psi_0$, $\psi \neq \psi_0$.

Um den Hypothesentest durchführen zu können, brauchen wir eine Stichprobe und eine Teststatistik $T(X_1, \dots, X_n)$, welche die Stichprobe als Input nimmt. Die Teststatistik bildet die Grundlage für die Entscheidung. Falls der Wert der Teststatistik für die Alternativhypothese spricht, lehnen wir die Nullhypothese ab.

α -Fehler sowie β -Fehler:

	Test entscheidet sich für H_0	Test entscheidet sich für H_1
H_0 ist wahr	✓	α -Fehler: Fehler 1. Art
H_1 ist wahr	β -Fehler: Fehler 2. Art	✓

Den α -Fehler sowie β -Fehler beim Testen können wir mit einem Gerichtsverfahren vergleichen, das sich ebenfalls in vier Fällen darstellen lässt:

	unschuldig	schuldig
Freispruch	✓	Freispruch wegen fehlenden Beweisen
Verurteilung	Fehlverurteilung	✓

Beim Gerichtsverfahren ohne Zeugen kann der Angeklagte seine Unschuld nicht beweisen. Beim Testen von Hypothesen ist es auch nicht möglich zu beweisen, dass die Nullhypothese gültig ist.

Zweiseitiger Hypothesentest:

$$H_0 : \psi = \psi_0 \quad \text{vs.} \quad H_1 : \psi \neq \psi_0$$

Einseitiger Hypothesentest:

$$H_0 : \psi \geq \psi_0 \quad \text{vs.} \quad H_1 : \psi < \psi_0$$

$$H_0 : \psi \leq \psi_0 \quad \text{vs.} \quad H_1 : \psi > \psi_0$$

Konfidenzintervalle spielen eine zentrale Rolle bei Hypothesentests. Wenn die berechnete Teststatistik innerhalb des Konfidenzintervalls liegt, wird die Nullhypothese H_0 nicht verworfen. Befindet sich die Teststatistik jedoch außerhalb des Konfidenzintervalls, wird die Nullhypothese abgelehnt.

Anmerkung 1: Wir bezeichnen den Bereich, in dem wir H_0 ablehnen als kritischen Bereich (auch Ablehnungsbereich genannt). Sein Komplement ist der Bereich, in welchem wir H_0 beibehalten.

Anmerkung 2: Entweder wird H_0 beibehalten, da wir nicht ausreichend Beweise zum Verwerfen haben oder H_0 wird verworfen. H_0 können wir niemals beweisen.

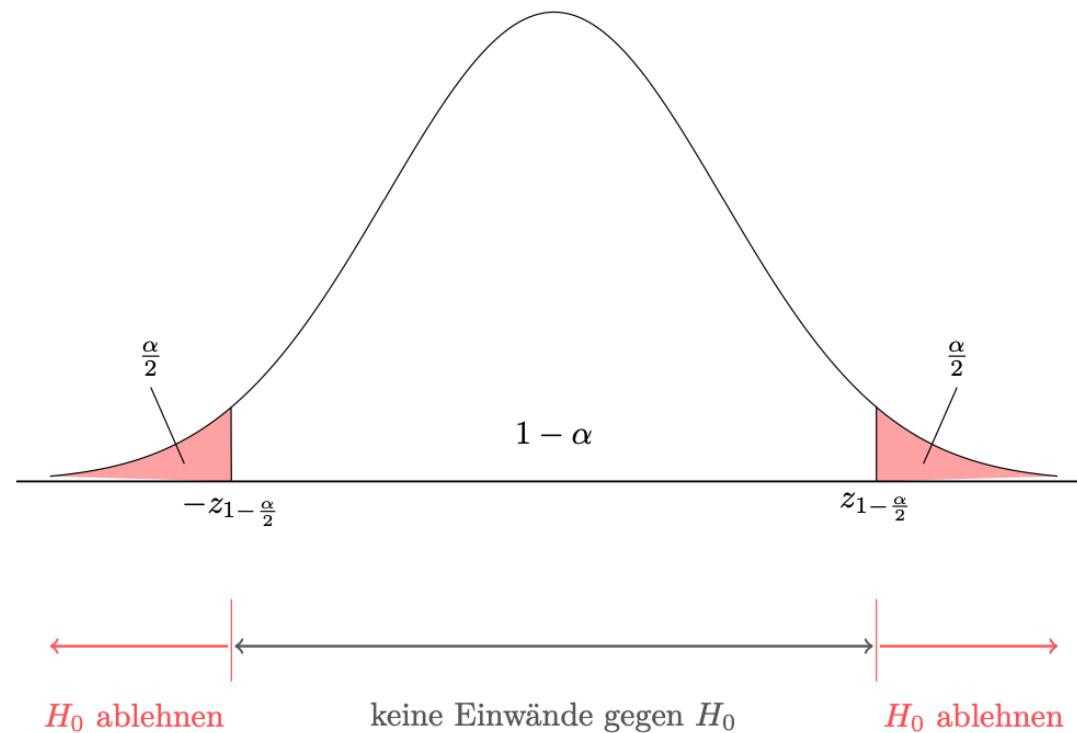


Abb.: Ablehnungsbereich und Nicht-Ablehnungsbereich der Nullhypothese.

Der **Gauß Test** wird zum Testen von Hypothesen über Mittelwerte bei bekannter Standardabweichung verwendet. Voraussetzung ist ein normalverteiltes Merkmal.

Wir wählen ein Signifikanzniveau α und eine der folgenden Hypothesen (je nach Fragestellung):

1. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$
2. $H_0 : \mu \geq \mu_0$ vs. $H_1 : \mu < \mu_0$
3. $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$.

Die Teststatistik wird wie folgt berechnet:

$$z = \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma},$$

wobei n die Stichprobengröße, σ die Standardabweichung der Grundgesamtheit und μ_0 der Erwartungswert unter H_0 beschreibt.

Wir lehnen H_0 ab, wenn

1. $|z| > z_{1-\frac{\alpha}{2}}$
2. $z < -z_{1-\alpha}$
3. $z > z_{1-\alpha}$,

wobei z_α das α -Quantil der Standardnormalverteilung ist.

Der **t-Test** wird zum Testen von Hypothesen über Mittelwerte bei unbekannter Standardabweichung verwendet. Voraussetzung ist ein normalverteiltes Merkmal.

Wir wählen ein Signifikanzniveau α und eine der folgenden Hypothesen (je nach Fragestellung):

1. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$
2. $H_0 : \mu \geq \mu_0$ vs. $H_1 : \mu < \mu_0$
3. $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$.

Die Teststatistik wird wie folgt berechnet:

$$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s},$$

wobei n die Stichprobengröße, s die empirische Standardabweichung und μ_0 der Erwartungswert unter H_0 beschreibt.

Wir lehnen H_0 ab, wenn

1. $|t| > t_{n-1; 1-\frac{\alpha}{2}}$
2. $t < -t_{n-1; 1-\alpha}$
3. $t > t_{n-1; 1-\alpha}$,

wobei $t_{n-1; \alpha}$ das α -Quantil der t -Verteilung mit $n - 1$ Freiheitsgraden ist.

Beispiel 1: Wir untersuchen das Abfüllgewicht einer 1000 Gramm Nussmischung und wollen nachweisen, dass das Abfüllgewicht kleiner als 1000 Gramm ist. Die Stichprobe ergibt folgende Werte:

$$\{980, 1000, 1010, 1000, 970, 1010, 960, 980, 1010, 960\}.$$

Wir erhalten: $\bar{x} = 988$ und $s = 20.44$. Wir setzen $\alpha = 0.05$ fest. Der einseitige Hypothesentest lautet:

$$H_0 : \mu \geq 1000 \quad \text{vs.} \quad H_1 : \mu < 1000$$

Wir erhalten aufgrund der Stichprobe folgende Teststatistik:

$$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s} = \frac{\sqrt{10}(988 - 1000)}{20.44} = -1.86.$$

Das 0.95-Quantil der t -Verteilung mit 9 Freiheitsgraden ist:

$$t_{n-1;1-\alpha} = t_{9;0.95} = 1.833.$$

Siehe https://de.wikipedia.org/wiki/Studentsche_t-Verteilung.

Somit ist der Ablehnungsbereich $A = \{t \in \mathbb{R} : t < -1.833\}$. Die Wert der Teststatistik liegt in A und demnach lehnen wir die Nullhypothese ab und haben einen signifikanten Nachweis, dass $\mu < 1000g$ ist.

Der χ^2 **Test** wird zum Testen von Hypothesen für die Varianz verwendet.
Voraussetzung ist ein normalverteiltes Merkmal.

Wir wählen ein Signifikanzniveau α und eine der folgenden Hypothesen (je nach Fragestellung):

1. $H_0 : \sigma^2 = \sigma_0^2$ vs. $H_1 : \sigma^2 \neq \sigma_0^2$
2. $H_0 : \sigma^2 \geq \sigma_0^2$ vs. $H_1 : \sigma^2 < \sigma_0^2$
3. $H_0 : \sigma^2 \leq \sigma_0^2$ vs. $H_1 : \sigma^2 > \sigma_0^2$.

Die Teststatistik wird wie folgt berechnet:

$$c = \frac{(n-1)s^2}{\sigma_0^2},$$

wobei n die Stichprobengröße und s die empirische Standardabweichung beschreibt.

Wir lehnen H_0 ab, wenn

1. $c < \chi_{n-1; \frac{\alpha}{2}}^2$ oder $c > \chi_{n-1; 1-\frac{\alpha}{2}}^2$
2. $c < \chi_{n-1; \alpha}^2$
3. $c > \chi_{n-1; 1-\alpha}^2$,

wobei $\chi_{n-1; \alpha}^2$ das $(1 - \alpha)$ -Quantil der Chi-Quadrat-Verteilung mit $n - 1$ Freiheitsgraden ist.

Beispiel 2: Die Produktion eines bestimmten Impfstoffes soll 1 Milligramm Phenol enthalten. Eine Stichprobe wurde entnommen und die gemessenen Mengen an Phenol in Milligramm sind:

0.995, 1.003, 1.001, 0.998, 1.002, 0.997, 1.000, 1.004, 0.996, 0.999.

1. Muss der Produktionsprozess neu kalibriert werden oder kann weiterhin $\mu = 1mg$ als Sollwert angenommen werden? Verwenden Sie ein Signifikanzniveau von $\alpha = 0.05$.
2. Die maximal zulässige Streuung für Phenol darf $\sigma = 0.015$ nicht übersteigen. Kann dies zu einem Signifikanzniveau von $\alpha = 0.01$ gewährleistet werden?

Lösung Frage 2.1: Wir erhalten $\bar{x} = 0.9995$ und $s = 0.00303$. Der zweiseitige Hypothesentest lautet:

$$H_0 : \mu = 1 \quad \text{vs.} \quad H_1 : \mu \neq 1.$$

Wir erhalten aufgrund der Stichprobe folgende Teststatistik:

$$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s} = -0.522.$$

Das 0.975-Quantil der t -Verteilung mit 9 Freiheitsgraden ist $t_{9;0.975} = 2.262$.

Da $|t| = |-0.522|$ nicht größer als 2.262 ist, wird die Nullhypothese nicht abgelehnt. Somit muss der Produktionsprozess nicht neu kalibriert werden, da kein signifikanter Unterschied festgestellt wurde.

Lösung Frage 2.2. Wir wollen nachweisen, dass σ den Wert 0.015 nicht übersteigt und wählen daher die Alternativhypothese entsprechend dieser Aussage. Daher gilt:

$$H_0 : \sigma^2 \geq 0.015^2 \quad \text{vs.} \quad H_1 : \sigma^2 < 0.015^2.$$

Die Teststatistik wird wie folgt berechnet:

$$c = \frac{(n-1)s^2}{\sigma_0^2} = \frac{9 \cdot 0.00303^2}{0.015^2} = 0.3672.$$

Wir lehnen H_0 ab, wenn $c < \chi_{n-1;\alpha}^2 = \chi_{9;0.01}^2$ gilt.

Laut Tabellen aus <https://datatab.de/tutorial/tabelle-chi-quadrat> sowie https://www.uibk.ac.at/econometrics/einf/tab_chisq.pdf gilt: $\chi_{9;0.01}^2 = 21.666$. Demnach wird H_0 abgelehnt und H_1 angenommen.

Vielen Dank!

Tobias Forster
tobias.forster@fhv.at