

Statistische Test-/Prüfverfahren

Wilfried Mann,
Mettmann

Mit dem statistischen Test bzw. dem Hypothesen-Prüfverfahren kann anhand vorliegender Beobachtungen entschieden werden, ob mit einer bestimmten Wahrscheinlichkeit eine zuvor aufgestellte Hypothese zutrifft oder zu verwerfen ist.

1. Beschreibung

1.1 Funktion

Ein **statistischer Test** dient dazu, anhand vorliegender Beobachtungen (Stichprobe), eine begründete Entscheidung über die Gültigkeit oder Ungültigkeit einer zuvor festgelegten Behauptung/Vermutung (sog. Hypothese) zu treffen. Kann man eine Hypothese »annehmen« oder muss man sie »ablehnen«? Wo liegen die Grenzen zwischen kleinen, rein zufallsbedingten, unvermeidlichen Abweichungen und größeren Abweichungen, die nicht mehr durch den Zufall erklärbar sind?

Diese Fragen lassen sich im Rahmen eines **Hypothesen-Prüfverfahrens** beantworten. Man spricht auch von Hypothesentest oder Signifikanztest. Eine 100 %-ige Sicherheit, dass die angenommene Hypothese auch tatsächlich wahr ist, kann der Hypothesentest naturgemäß niemals bieten, da von einer Stichprobe auf die Grundgesamtheit geschlossen wird.

1.2 Nullhypothese und Alternativhypothese

Die beiden Hypothesen, die bei einem Signifikanztest einander gegenüberstehen, sind die Nullhypothese (H_0) und die Gegen- oder Alternativhypothese (H_1 , H_A). Welche der beiden Hypothesen die Nullhypothese ist und welche die Gegenhypothese, hängt von der in der Aufgabenstellung gegebenen Ausgangssituation ab.

Die Nullhypothese stellt die Basis dar, von der aus entschieden wird, ob die Alternativhypothese akzeptiert werden kann oder nicht.

Zum Beispiel sind, bezogen auf einen Mittelwertvergleich, drei Hypothesen möglich:

$$\begin{aligned} H_1: \mu_0 > \mu_1 \quad H_0: \mu_0 \leq \mu_1 \\ H_1: \mu_0 < \mu_1 \quad H_0: \mu_0 \geq \mu_1 \\ H_1: \mu_0 \neq \mu_1 \quad H_0: \mu_0 = \mu_1 \end{aligned}$$

1.3 Fehlerarten

Neben den beiden richtigen Entscheidungen können auch zwei fehlerhafte Entscheidungen aufgrund der Stichprobenergebnisse getroffen werden. Entweder kann die an sich richtige Nullhypothese zugunsten der Alternativhypothese verworfen werden (α -Fehler oder Fehler 1. Art) oder es wird die Nullhypothese akzeptiert obwohl die Alternativhypothese richtig ist (β -Fehler oder Fehler 2. Art).

	H_0 ist wahr	H_0 ist falsch (H_1 ist wahr)
Testgröße T nimmt bei einer Stichprobe einen Wert im Annahmebereich von H_0 an	Richtige Entscheidung H_0 wird zu Recht beibehalten Wahrscheinlichkeit: $1 - \alpha$	Falsche Entscheidung H_0 wird zu Unrecht beibehalten Fehler 2. Art Wahrscheinlichkeit: β
Testgröße T nimmt bei einer Stichprobe einen Wert im Ablehnungsbereich von H_0 an	Falsche Entscheidung H_0 wird zu Unrecht verworfen Fehler 1. Art Wahrscheinlichkeit: α	Richtige Entscheidung H_0 wird zu Recht verworfen Wahrscheinlichkeit: $1 - \beta$

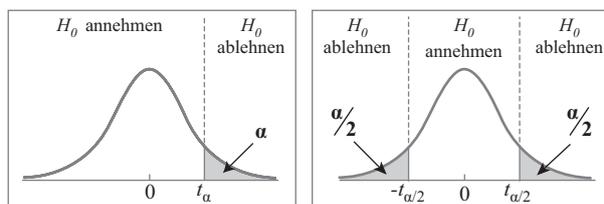
1.4 Prüfgröße und kritischer Wert

Um die Abweichung einer Stichprobe von der angenommenen Grundgesamtheit zu messen, berechnet man eine **Prüfgröße**. Der Berechnungsmodus variiert von Anwendung zu Anwendung, berücksichtigt aber immer die Abweichung zwischen den in der Nullhypothese spezifi-

zierten Parametern und den entsprechenden Stichprobenstatistiken, wie Standardabweichung und Stichprobenumfang. Um extreme oder weniger extreme Werte erkennen zu können, muss man die zur Prüfgröße gehörende Testverteilung kennen, mit der man die Wahrscheinlichkeit bestimmen kann, mit der die Prüfgröße in einen bestimmten Wertebereich fällt. Werte, die einen bestimmten **kritischen Wert** der Testverteilung überschreiten, sind dann eher »unwahrscheinliche« Werte.

1.5 Annahme- und Ablehnungsbereich

Um zu einer Entscheidung zu kommen, wird durch Vorgabe eines Signifikanzniveaus der Annahme- und Ablehnungsbereich festgelegt (Entscheidungsregel). Das Signifikanzniveau ist dabei die Komplementärwahrscheinlichkeit (Gegenwahrscheinlichkeit) zur Sicherheitswahrscheinlichkeit.



Annahme und Ablehnungsbereich bei einem einseitigen Test (für eine gerichtete Hypothese) und zweiseitigen Test (für eine ungerichtete Hypothese)

2. Anwendung

Das hier eher systematisch beschriebene Hypothesen-Prüfverfahren kennt in der Praxis eine Vielzahl von komplexen Facetten. Im Rahmen der Wertermittlung kommt z.B. die Prüfung von Mittelwerten (Mittelwertvergleich) zum Tragen. Hierbei ist zunächst abzuklären, ob die Stichproben abhängig oder unabhängig sind, wie viele Messungen bzw. Gruppen es gibt (mindestens zwei), ob es sich um normalverteilte oder nicht normalverteilte abhängige Variable handelt und ob diese metrisch oder nicht metrisch verteilt sind. Bei zwei Gruppen, metrischen und normalverteilten Variablen, wird ein sog. t-Test durchgeführt, die Testverteilung ist hier die Student-t-Verteilung*. Andererseits kommen auch nichtparametrische Testverfahren zum Einsatz.

Eine Statistik-Software berechnet den **kritischen Wert**, einen sog. **P-Wert**, als Wahrscheinlichkeit, dass die Teststatistik – bei Gültigkeit von H_0 – mindestens den in der Stichprobe berechneten Wert oder einen größeren Wert annimmt.

Das Prinzip ist auch bei komplexen Kaufpreisanalysen, wie der multiplen Regressionsanalyse, anwendbar. Mit Hilfe von geeigneten Hypothesentests lassen sich Aussagen treffen, ob eine Zielgröße von verschiedenen Einflussgrößen signifikant abhängig ist (Fisher-F-Test für das Gesamtmodell oder Student-t-Test für einzelne Regressionskoeffizienten).

* Siehe Glossar Theoretische Verteilungen, GuG 3/2019, S. 166.