

4.3 Allgemeines Vorgehen bei statistischen Tests

1. **Schritt:** Formulierung von Annahmen. Ganz ohne Annahmen kommt man meist nicht aus. Übliche Annahmen betreffen meist die Verteilung der Stichprobenvariablen und deren Unabhängigkeit.
2. **Schritt:** Formulierung der Nullhypothese.
3. **Schritt:** Auswahl des Testverfahrens.
4. **Schritt:** Durchführung des Tests und Entscheidung.

4.4 Ausgewählte statistische Tests

4.4.1 Wie findet man das richtige Testverfahren?

Statistische Tests kann man nach mehreren Kriterien in Klassen einteilen.

- **Anzahl der beteiligten Zufallsgrößen**

Sollen zwei Zufallsgrößen mit potentiell unterschiedlichen Verteilungen verglichen werden, für die jeweils eine Stichprobe erzeugt wird, (**Zwei-Stichproben-Test**) oder wird nur eine einzelne Zufallsgröße untersucht (**Ein-Stichproben-Test**)?

Bei der Fragestellung

Beträgt die mittlere Zugriffszeit auf einen Datenbankserver im Mittel höchstens 10ms?

hat man es mit einem Ein-Stichproben-Test zu tun, während die Untersuchung der Frage

Hat Datenbankserver A eine kürzere mittlere Zugriffszeit als Datenbankserver B?

auf einen Zwei-Stichproben-Test führt.

Bei mehreren beteiligten Zufallsgrößen wird zusätzlich unterschieden, ob aus voneinander unabhängigen Grundmengen Stichproben erhoben werden oder nicht. Beim vorigen Beispiel werden **unabhängige Messungen** vorgenommen, sofern die Server A und B getrennt voneinander arbeiten. Wenn man jedoch die Frage

Läuft ein Datenbankserver auf einer Menge festgelegter Testanfragen mit Query-Optimierung schneller als ohne?

untersucht, so spricht man von **verbundenen Messungen**.

Gelegentlich betrachtet man auch den Zusammenhang zwischen mehreren Zufallsgrößen. Beispielsweise könnte man sich für die Frage interessieren:

Wie stark wächst der Zeitbedarf für eine Datenbankanfrage im Mittel mit der (syntaktischen) Länge der Anfrage, d. h. führen kompliziertere Formulierungen zu proportional längeren Laufzeiten?

Mit solchen Fragenstellungen, bei denen ein funktionaler Zusammenhang zwischen Zufallsgrößen ermittelt werden soll, beschäftigt sich die **Regressionsanalyse**. Wenn überhaupt erst zu klären ist, ob ein solcher Zusammenhang besteht oder ob die Zufallsgrößen vielmehr unabhängig voneinander sind, so spricht man von **Zusammenhangsanalyse**.

- **Formulierung der Nullhypothese**

Welche Größe dient zur Definition der Nullhypothese? Hierbei werden in erster Linie Tests unterschieden, die Aussagen über verschiedene so genannte **Lageparameter** treffen, wie z.B. den **Erwartungswert** oder die **Varianz** der zugrunde liegenden Verteilungen.

Im Zwei-Stichproben-Fall könnte man beispielsweise untersuchen, ob der Erwartungswert der Zufallsgröße A größer oder kleiner als bei Zufallsgröße B ist.

Gelegentlich wird zur Formulierung der Nullhypothese auch der so genannte **Median** betrachtet: Der Median einer Verteilung entspricht dem (kleinsten) Wert x mit $F(x) = 1/2$.

Neben solchen Tests auf Lageparameter gibt es z.B. auch Tests, die auf eine **vorgegebene Verteilung** oder auf ein Maß für die Abhängigkeit verschiedener Zufallsgrößen testen.

- **Annahmen über die Zufallsgrößen**

Was ist über die Verteilung der untersuchten Größe(n) bekannt? Bei entsprechenden Annahmen könnte es sich z.B. um die Art der Verteilung, den Erwartungswert oder die Varianz handeln.

4.4.2 Ein-Stichproben-Tests für Lageparameter

Beim approximativen Binomialtest wird ausgenutzt, dass die Binomialverteilung für große n nach dem Grenzwertsatz von de Moivre (Korollar 116) gegen die Normalverteilung konvergiert. Aus diesem Grund kann man diesen Test auch als Spezialfall eines allgemeineren Testverfahrens ansehen, nämlich des **Gaußtest**, der nun dargestellt wird.

Table: Gaußtest

Annahmen:

X_1, \dots, X_n seien unabhängig und identisch verteilt mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$, wobei σ^2 bekannt ist.

Alternativ gelte $\mathbb{E}[X_i] = \mu$ und $\text{Var}[X_i] = \sigma^2$, und n sei groß genug.

Hypothesen:

- a) $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$,
- b) $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$,
- c) $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$.

Testgröße:

$$Z := \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}.$$

Ablehnungskriterium für H_0 bei Signifikanzniveau α :

- a) $|Z| > z_{1-\alpha/2}$,
- b) $Z < z_\alpha$,
- c) $Z > z_{1-\alpha}$.

Der Gaußtest hat den Nachteil, dass man die Varianz σ^2 der beteiligten Zufallsgrößen kennen muss.

Wenn diese unbekannt ist, so liegt es nahe, die Varianz durch die Stichprobenvarianz S^2 (siehe Definition 121) anzunähern. Dies führt auf den so genannten *t-Test*, der in der folgenden Übersicht dargestellt ist.

Tabelle: *t*-Test

Annahmen:

X_1, \dots, X_n seien unabhängig und identisch verteilt mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$.
Alternativ gelte $\mathbb{E}[X_i] = \mu$ und $\text{Var}[X_i] = \sigma^2$, und n sei groß genug.

Hypothesen:

- a) $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$,
- b) $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$,
- c) $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$.

Testgröße:

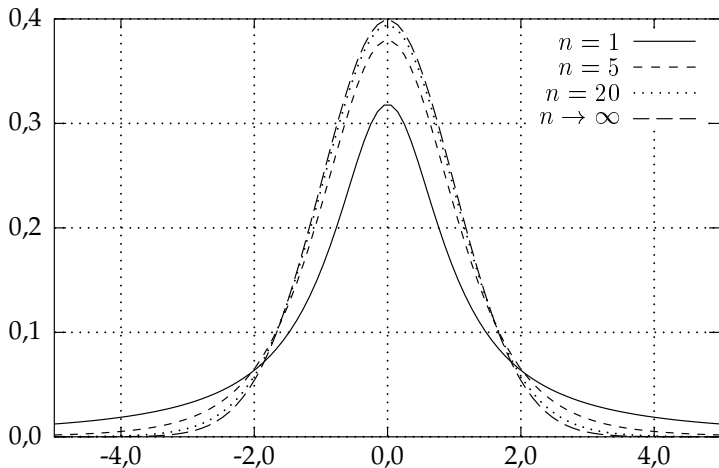
$$T := \frac{\bar{X} - \mu_0}{S} \sqrt{n}.$$

Ablehnungskriterium für H_0 bei Signifikanzniveau α :

- a) $|T| > t_{n-1, 1-\alpha/2}$,
- b) $T < t_{n-1, \alpha}$,
- c) $T > t_{n-1, 1-\alpha}$.

Hierbei gibt $t_{n-1,1-\alpha}$ das $(1 - \alpha)$ -Quantil der *t-Verteilung* mit $n - 1$ Freiheitsgraden an. Die *t-Verteilung* taucht manchmal auch unter dem Namen *Student-Verteilung* auf, da sie ursprünglich unter dem Pseudonym „Student“ publiziert wurde.

Wir gehen an dieser Stelle nicht darauf ein, wieso die Testgröße die *t-Verteilung* besitzt, sondern weisen nur darauf hin, dass die Dichte dieser Verteilung (eigentlich handelt es sich um eine ganze Familie von Verteilungen, da die Anzahl der Freiheitsgrade jeweils noch gewählt werden kann) der Dichte der Normalverteilung ähnelt. Für große n (Faustregel: $n \geq 30$) liegen die beiden Dichten so genau übereinander, dass man in der Praxis die *t-Verteilung* durch die Normalverteilung annähert.



Dichte der t -Verteilung mit n Freiheitsgraden