Definition eines Tests.

Gegeben: Stichprobenvariablen $X_1, \ldots, X_n$ bzw. Stichprobe $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Aufgabe: Entscheide, ob eine gegebene Hypothese H_0 (Nullhypothese) gilt oder die Alternative H_1.

Beispiel: Ist X Bernuolli-verteilt mit (unbekanntem) Parameter p, so könnten man setzten

$$H_0 : p \geq 1/3 \quad \text{und} \quad H_1 : p < 1/3$$

Ein echter Alternativtest liegt vor, wenn man ansetzt

$$H_0' : p \geq 1/3 \quad \text{und} \quad H_1' : p \leq 1/6.$$

(Dies ist in einigen Anwendungen durchaus sinnvoll, s.u.)

Vorgehensweise: Wir definieren eine Menge

$$K := \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \text{ führt zur Ablehnung der Hypothese}\}.$$

K nennen wir den *Ablehnungbereich* oder den *kritischen Bereich* des Tests.

BEISPIEL: (Echter Alternativtest)

Wir untersuchen eine Festplatte, von der bekannt ist, dass sie zu einer von zwei Baureihen gehört. Die mittleren Zugriffszeiten dieser Baureihen betragen 9ms bzw. 12ms. Wir möchten nun herausfinden, zu welchem Typ die betrachtete Festplatte gehört, indem wir die Zugriffszeit bei n Zugriffen bestimmen. Hier würde man dann ansetzen: $H_0 : \mu \leq 9$ und $H_1 := \mu \geq 12$, wobei μ die mittlere Zugriffszeit bezeichnet.

Fehler bei statistischen Tests

Bei jedem statistischen Test können mit einer gewissen Wahrscheinlichkeit falsche Schlüsse gezogen werden. Hierbei unterscheidet man zwischen den beiden Fällen, ob die Hypotheses in Wahrheit zutrifft oder eben nicht.

Fehler 1. Art	:	H_0 gilt und wird irrtümlich abgelehnt.
Fehler 2. Art	:	H_0 gilt nicht und wird irrtümlich angenommen.

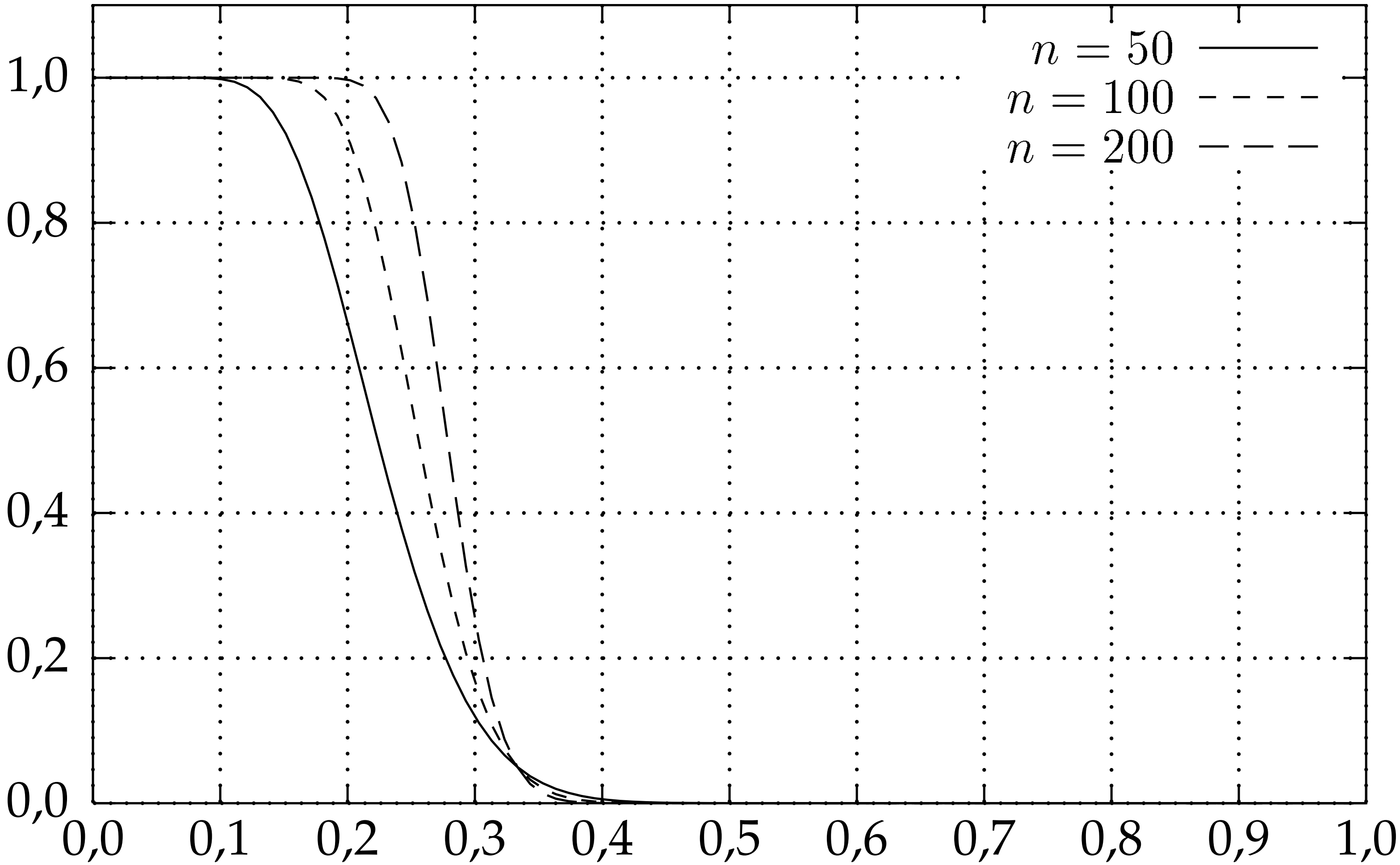
Für die Beurteilung eines Tests ist es wesentlich, mit welcher Wahrscheinlichkeit diese beiden Fehler eintreten können. Ziel ist es natürlich, diese Wahrscheinlichkeiten möglichst klein zu halten.

Achtung: Die Minimierung des Fehlers 1. Art und des Fehlers 2. Art sind gegenläufige Ziele:

Wenn man beispielsweise $K = \emptyset$ setzt, so erhält man Wahrscheinlichkeit Null für den Fehler 1. Art, da H_0 immer angenommen wird. Allerdings tritt der Fehler 2. Art dann mit Wahrscheinlichkeit Eins ein, wenn H_0 nicht gilt.

Vorgehensweise: In der Praxis gibt man sich eine Fehlerwahrscheinlichkeit α für den Fehler 1. Art vor und versucht dann den Ablehnungsbereich so einzustellen, dass der Fehler 2. Art möglichst klein wird.

α heißt daher auch *Signifikanzniveau* des Tests.



Gütefunktion $g(n, \frac{1}{3})$ für verschiedene Werte von n zum Signifikanzniveau $\alpha = 0.05$

Tabelle 1: Testbeschreibung Approximativer Binomialtest

Approximativer Binomialtest

Annahmen:

$X_1, \ldots, X_n$ seien unabhängig und identisch verteilt mit $\Pr[X_i = 1] = p$ und $\Pr[X_i = 0] = 1 - p$, wobei p unbekannt sei. n sei hinreichend groß, so dass die Approximation gemäß Zentralen Grenzwertsatzes brauchbare Ergebnisse liefert.

Hypothesen:

a) $H_0 : p = p_0$ gegen $H_1 : p \neq p_0$,

b) $H_0 : p \geq p_0$ gegen $H_1 : p < p_0$,

c) $H_0 : p \leq p_0$ gegen $H_1 : p > p_0$.

Testgröße:

$$Z := \frac{h - np_0}{\sqrt{np_0(1 - p_0)}},$$

wobei $h := X_1 + \ldots + X_n$ die Häufigkeit bezeichnet, mit der die Ereignisse $X_i = 1$ aufgetreten sind.

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

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

b) $Z < z_\alpha$,

c) $Z > z_{1-\alpha}$.

Allgemeines Vorgehen bei statistischen Tests

1. Schritt: Formulierung von Annahmen. Ganz ohne Annahmen kommt man meist nicht aus. In jedem Fall werden die Testergebnisse durch (korrekte) Annahmen aussagekräftiger. Übliche Annahmen betreffen meist die Verteilung der Stichprobenvariablen und deren Unabhängigkeit.

2. Schritt: Formulierung der Nullhypothese. Die Nullhypothese ergibt sich in in der Regel direkt aus der Problemstellung.

Man geht gewöhnlich so vor, dass man als Nullhypothese das Komplement der Aussage verwendet, die man verifizieren möchte. Wenn der Test dazu führt, dass die Nullhypothese verworfen und damit die zu überprüfende Aussage angenommen wird, so geschieht dies auf dem zuvor festgelegten Signifikanzniveau und man kann sich somit hinreichend sicher sein, keine Fehlentscheidung getroffen zu haben.

Achtung: Bei statistischen Tests mit der trivialen Alternative H_1 : “H_0 gilt nicht” gibt es keine Möglichkeit, Aussagen über die Wahrscheinlichkeit anzugeben, dass H_0 irrtümlich angenommen wird.

3. Schritt: Auswahl des Testverfahrens. Dieser Schritt hängt stark von den vorherigen ab. Man muss dazu in der Literatur ein Verfahren finden, auf das die getroffenen Annahmen zutreffen, und das auf die zu untersuchende Nullhypothese passt.

4. Schritt: Durchführung des Tests und Entscheidung. Man berechnet die zu dem gewählten Verfahren gehörenden Testgröße, wählt ein Signifikanzniveau, vergleicht die Testgröße mit dem entsprechenden Entscheidungskriterium.

Tabelle 2: Testbeschreibung Gaußtest

Gaußtest

Annahmen:

$X_1, \ldots, X_n$ 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 $\mathrm{Var}[X_i] = \sigma^2$ 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:

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

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

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

b) $Z < z_\alpha$,

c) $Z > z_{1-\alpha}$.

Tabelle 3: Testbeschreibung t-Test

t-Test

Annahmen:

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

Alternativ gelte $\mathbb{E}[X_i] = \mu$ und $\mathrm{Var}[X_i] = \sigma^2$ 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:

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

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

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

b) $T < t_{n-1,\alpha}$,

c) $T > t_{n-1,1-\alpha}$.

Tabelle 4: Testbeschreibung Zwei-Stichproben-t-Test

Zwei-Stichproben-t-Test

Annahmen:

$X_1, \ldots, X_m$ und $Y_1, \ldots, Y_n$ seien unabhängig und jeweils identisch verteilt, wobei $X_i \sim \mathcal{N}(\mu_X, \sigma_X^2)$ und $Y_i \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ gelte. Die Varianzen seien identisch, also $\sigma_X^2 = \sigma_Y^2$.

Hypothesen:

a) $H_0 : \mu_X = \mu_Y$ gegen $H_1 : \mu_X \neq \mu_Y$,

b) $H_0 : \mu_X \geq \mu_Y$ gegen $H_1 : \mu_X < \mu_Y$,

c) $H_0 : \mu_X \leq \mu_Y$ gegen $H_1 : \mu_X > \mu_Y$.

Testgröße:

$$T := \sqrt{\frac{n+m-2}{\frac{1}{m}+\frac{1}{n}}} \cdot \frac{\overline{X}-\overline{Y}}{\sqrt{(m-1)\cdot S_X^2 + (n-1)\cdot S_Y^2}}.$$

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

a) $|T| > t_{m+n-2,1-\alpha/2}$,

b) $T < t_{m+n-2,\alpha}$,

c) $T > t_{m+n-2,1-\alpha}$.