Hypothesenprüfungen

Dr László Smeller

1

Vergleich der Schätzungen und Hypothesenprüfungen

Schätzungen:

Frage: Wie groß (ist eine physikalische Größe) $\mu=?$

z.B.: Körperhöhe, Blutdruck, Blutzuckerkonzentration...

Antwort: Punktschätzung: Ein Wert

Intervallschätzung: Ein Intervall + Konfidenzniveau (Sicherheitswahrscheinlichkeit)

Hypothesenprüfungen:

Frage: Eine Entscheidungsfrage (ist es wahr oder nicht?)

zB: hat ein Medikament eine Wirkung oder nicht?

Mathematisch: ist $\mu=\mu_0$?

Antwort: Ja oder Nein + Konfidenzniveau (Sicherheitswahrsch.)

Signifikanzniveau (Irrtumswahrsch.)

Typische Aufgaben der Hypothesenprüfung

1. Hat ein Medikament/Behandlung eine Wirkung?
   1a. Verursacht es eine Änderung (zB. Blutdruckänderung, d.h.: ist der Blutdruck kleiner nach der Eingabe?)
   1b. Gibt es einen Unterschied zwischen den unbehandelten und behandelten Gruppen?

2. Gibt es eine Korrelation
   2. a. zwischen zwei Parametern (zB. Körperhöhe und Gewicht, …)
   2. b. Gibt es eine Korrelation zwischen zwei Eigenschaften (Alkoholismus, Leberschrumpfung)

Typische Fragen - gebrauchte Merkmale

1. Hat ein Medikament/Behandlung eine Wirkung?
   Änderung von einer **numerischen (kontinuierlichen) Größe**
   (zB. Blutdruck, Körpertemperatur, Blutzuckerkonzentration, …)
   1a. Änderung nach einem Einfluss an einer Stichprobe
   1b. Unterschied zwischen zwei Stichproben

2. Gibt es eine Korrelation
   2a. zwischen **zwei numerischen Größen**
   (zB. Körperhöhe und Gewicht, …)
   2b. zwischen zwei (oder mehreren) **kategorischen Merkmalen**
   (zB: Alkoholiker – Antialkoholiker, Leberschrumpfung – keine Leberschr.
   Raucher – Nichtraucher, Lungenkrebs – kein Lungenkrebs)

Grundprinzip der Hypothesenprüfungen

Die Wirkung
Die Korrelation } können viele unterschiedliche Stärke aufnehmen

⇩

unprognostizierbar

⇩

mathematisch unbehandelbar

☹

Keine Wirkung
Keine Korrelation } nur zufällige Effekte treten auf

⇩

mathematisch gut beschreibbar!

☺

⇒ **Nullhypothese**

Die Nullhypothese und die Alternativhypothese

Nullhypothese (H_0):

Es gibt **keine Wirkung** $\mu=\mu_0$

Alle **Abweichungen** von dem theoretischen Wert sind rein **zufällig**.

Alternativhypothese (H_1)

Es gibt eine Wirkung $\mu\neq\mu_0$

Die Abweichungen sind nicht zufällig, sondern systematisch!

Eine von H_0 und H_1 wird unbedingt auftreten! $p(H_0 \text{ oder } H_1)=1$

Beispiele für Nullhypothesen

beim Fiebermittel:

Keine Wirkung (H_0):

Die Temperaturänderung nach der Eingabe = 0.

Genauer: Erwartungswert der Temperaturänderungen ist $\mu_0=0$

beim Würfelspiel:

Grundprinzip der Hypothesenprüfungen

Sei es vorausgesetzt, dass wir keine Wirkung/Korrelation haben!
(D.h. H_0 ist richtig.)

Wenn unsere Ergebnisse diesen Voraussetzung nicht entsprechen, dann haben wir wahrscheinlich eine Wirkung/Korrelation.

Beispiel 1.: Fibermittel

Die Temperaturen wurden vor und nach der Eingabe der Tablette gemessen.

Die Messergebnisse (in °C):

Temperatur - änderung

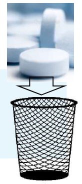
T_{vor}	T_{nach}	$x = T_{nach} - T_{vor}$
39,7	39,0	-0,7
38,8	38,4	-0,4
37,9	38,3	0,4
39,2	38,9	-0,3
38,9	38,4	-0,5
Durchschnitt $\bar{x}$		-0,3

Nullhypothese: das Fiebermittel is unwirksam.
Die Temperaturänderungen (x_i) sind zufällig.
Der Erwartungswert der Temperaturänderungen ist null.

1. Beispiel: Fibermittel

Nullhypothese: $\mu_0 = 0$

Wenn die Nullhypothese gültig ist, dann befindet sich $\bar{x}$ nicht weit von μ_0.



Ist $\bar{x} = -0{,}3\,°C$ klein genug um es als zufällig zu betrachten? d.h. die Nullhypothese anzunehmen?

oder

Ist $\bar{x} = -0{,}3\,°C$ groß genug um es als systematische Abweichung zu betrachten? d.h.Nullhypothese abzulehnen?

Aber wo ist die Grenze?
Wie groß muss der Durchschnitt sein um die Nullhypothese abzulehnen?

2. Beispiel: Kniebeugungen

Pulszahl vor und nach 10 Kniebeugungen.
Wird die Pulszahl geändert nach der Kniebeugungen?
H_0: keine Änderung $\mu=0$

p_{vor}	p_{nach}	$x=\Delta p$
65	79	14
68	77	9
72	91	19
63	70	7
74	88	14
69	84	15
	Durchsch.	13

Ist 13 $^1/_{Min}$ klein genug um die Nullhypothese anzunehmen?

oder

Ist 13 $^1/_{Min}$ groß genug um die Nullhypothese abzulehnen?

Wie große zufällige Abweichung ist erlaubt?

(Rein zufällige, statistische) Streuung des Durchschnittes beträgt: $s_{\bar{x}}$

$s_{\bar{x}}$ ist der „Maßstab“ für die Abweichung des Durchschnittes von dem Erwartungswert μ_0.

Einige $s_{\bar{x}}$ große Abweichung ist „erlaubt“,
aber merfache $s_{\bar{x}}$ Abweichung ist sehr unwahrscheinlich.
(Angenommen dass die Nullhypothese gültig ist.)

Der t-Wert

Weil die zufällige Abweichungen des Durchschnittwertes von μ können einige $s_{\bar{x}}$ betragen, vergleichen wir $\bar{x}$ mit $s_{\bar{x}}$.

Definieren wir eine neue Größe:

$$t = \frac{\bar{x}}{s_{\bar{x}}} \qquad (\bar{x} \text{ in } s_{\bar{x}} \text{ Einheiten gemessen})$$

oder mathematisch:

$$t = \frac{\bar{x}}{s_{\bar{x}}} = \frac{\frac{\sum_{i=1}^{n} x_i}{n}}{\frac{s}{\sqrt{n}}} = \frac{\frac{\sum_{i=1}^{n} x_i}{n}}{\frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}}{\sqrt{n}}} = f(x_1, x_2, \ldots, x_n)$$

t für unseren „Fibermittel“:

$t = -0{,}3/0{,}187 = -1{,}6$

für Kniebeugungen $t' = 7{,}34$

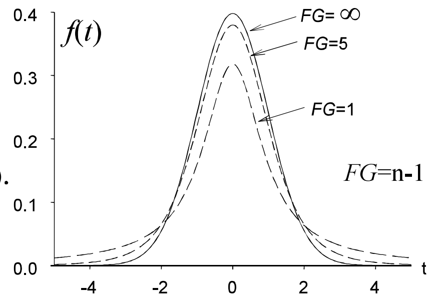
Die Verteilung des t-Wertes: die t-Verteilung

$t = f(x_1, x_2, \ldots, x_n)$

Wenn die Nullhypothese gültig ist, alle x_i Werte folgen einer t-Verteilung mit $\mu = 0$.

Wenn die $x_1, x_2, \ldots x_n$ Daten einer Normalverteilung folgen, die Verteilung von $t = \frac{\bar{x}}{s_{\bar{x}}}$ kann berechnet werden: t-Verteilung

Wenn die Nullhypothese gültig ist, der aus unserer Stichprobe ausgerechnete t-Wert folgt einer t-Verteilung (Student-Vert.).



Bedingung: x muss normalverteilt sein.

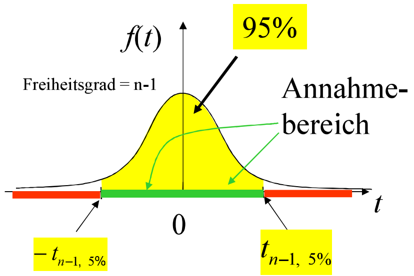
D.h.: man braucht normalverteilte Messwerte.

Die Anwendung der t-Verteilung

Die Wahrscheinlichkeit, dass bei einer richtigen Nullhypothese

$$-t_{n-1,\ 5\%} < t < +t_{n-1,\ 5\%}$$

gilt, beträgt 95%.



Bei richtiger Nullhypothese ist der aus der Stichprobe ausgerechnete t-Wert mit 95% Wahrscheinlichkeit in dem Annahmebereich. Wir können diesen kleinen t-Wert mit zufälligen Abweichungen erklären. $\Rightarrow$ Wir müssen keine Wirkung voraussetzen.

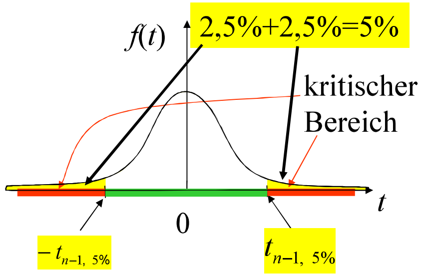
Die Nullhypothese kann nicht abgelehnt werden.
(H_0 wird angenommen)

Die Anwendung der t-Verteilung

Die Wahrscheinlichkeit, dass bei einer richtigen Nullhypothese

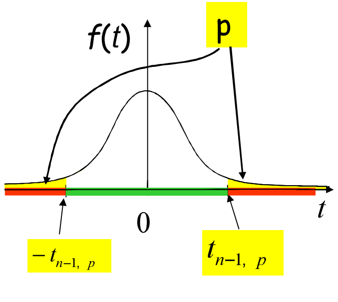
$$t < -t_{n-1,\ 5\%} \quad \text{oder} \quad t > +t_{n-1,\ 5\%}$$

gilt, beträgt 5%.

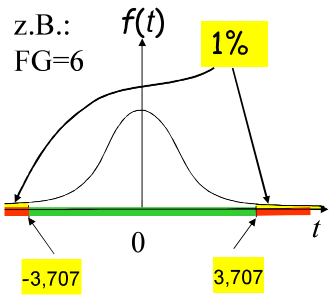


D.h.: Es ist sehr unwahrscheinlich (<5%), dass wir ***bei richtiger Nullhypothese*** einen so großen t-Wert bekommen. $\Rightarrow$ Wir haben wahrscheinlich eine Wirkung, ***die Nullhypothese kann abgelehnt werden, die Alternativhypothese wird angenommen.***

Das 5% nennt man als ***Signifikanzniveau*** oder ***Irrtumswahrscheinlichkeit.***



z.B.:
FG=6



Freiheits-grad (FG)	p (Irrtumswahrscheinlichkeit)				
	0,2	0,1	0,05	0,02	0,01
1	3,078	6,314	12,70	31,82	63,65
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807

Ablauf der Hypothesenprüfung bei einem *t*-Test

1. Fragestellung (mit der Definition der Population!)
   (Bedingung: Normalverteilung)
2. Nullhypothese - Alternativhypothese
3. Festlegung des Signifikanzniveaus (α)
4. Messung (Stichprobe mit n Messungen, Repräsentativität!)
5. Berechnung des t-Wertes
6. Vergleich von unserem t und dem Grenzwert ($t_{n-1,\alpha}$)

$|t| < t_{n-1,\alpha}$ $\qquad$ $|t| > t_{n-1,p}$

7. Die Entscheidung:

die **Nullhypothese** kann mit einem α Signifikanzniveau **nicht abgelehnt** werden.

Anhand unserer Messung kann die Alternativhypothese nicht bewiesen werden.

die **Nullhypothese** kann mit einem α Signifikanzniveau **abgelehnt** werden

Die Alternativhypothese ist angenommen **(mit einem Signifikanzniveau von α).**

1. Beispiel: Fibermittel

1. Die Frage: Ist das Fibermittel wirksam?
2. Nullhypothese: Das Fibermittel ist unwirksam.
   (Die Temperaturänderungen sind rein zufällig).
3. Sei p=5%
4. Die Messergebnisse:
   Temperaturen vor und nach der Eingabe (in °C):
   x= Temperaturänderung

Temperatur - änderung

T_{vor}	T_{nach}	$x = T_{nach} - T_{vor}$
39,7	39,0	-0,7
38,8	38,4	-0,4
37,9	38,3	0,4
39,2	38,9	-0,3
38,9	38,4	-0,5
Durchschnitt		-0,3

5.
$$\bar{x} = -0{,}3\,°C$$
$$s_{\bar{x}} = 0{,}187^\circ C$$
$$t = \frac{-0{,}3^\circ C}{0{,}187^\circ C} = -1{,}6$$

FG= n-1=4

6. Vergleich des unseren *t*-Wertes mit dem Grenzwert

p=0,05

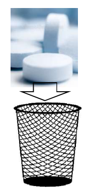
FG=4

Freiheits-grad (FG)	p (Irrtumswahrscheinlichkeit)				
	0,2	0,1	0,05	0,02	0,01
1	3,078	6,314	12,70	31,82	63,65
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055

$t_{4;5\%}$=2,776

$|t| < t_{4;\,5\%}$ die Nullhypothese kann nicht abgelehnt werden.
Die Nullhypothese wird angenommen
(mit einem *5%* Signifikanzniveau).

Das Fiebermittel ist mit 5% Sign.n unwirksam!



2. Beispiel: Kniebeugungen

Pulszahl vor und nach 10 Kniebeugungen.

p_{vor}	p_{nach}	$x = \Delta p$
65	79	14
68	77	9
72	91	19
63	70	7
74	88	14
69	84	15
	Durchsch.	13
	Stabw.	4,34
	Stfehler	1,77

Freiheits-grad (FG)	*p* (Irrtumswahrscheinlichkeit)				
	0,2	0,1	0,05	0,02	0,01
1	3,078	6,314	12,70	31,82	63,65
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898

n=6

$$t = \frac{13}{1{,}77} = 7{,}34$$

Hypothesenprüfung mit Excel

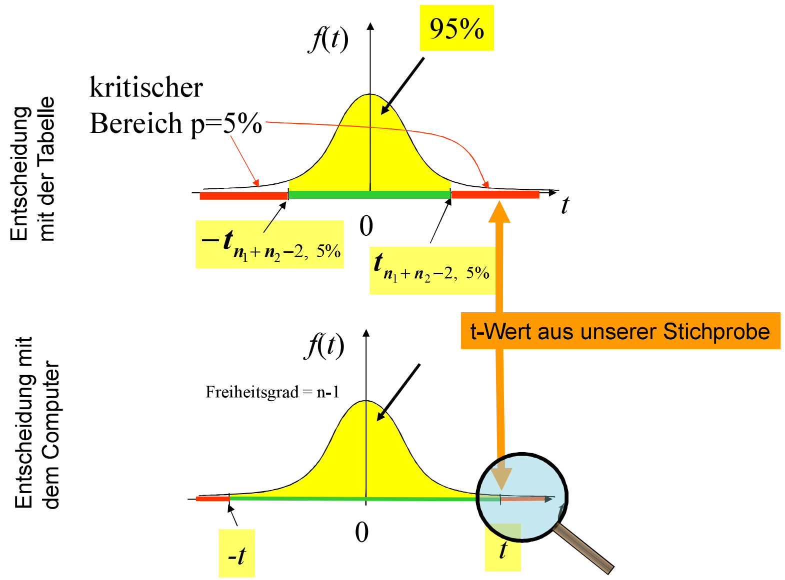
Excel Funktion für t-Teste:

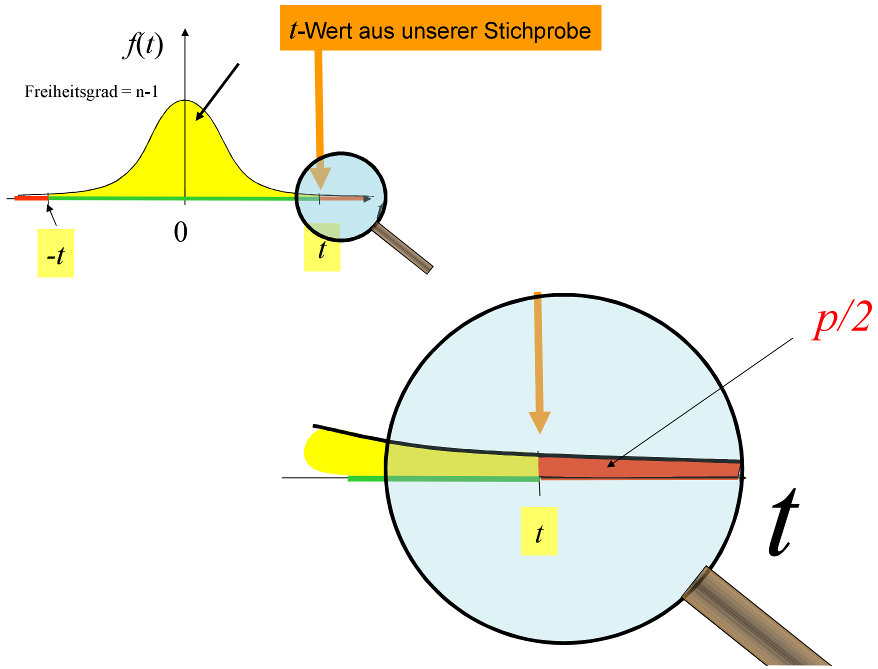
ttest(Reihe1; Reihe2; Seiten; Typ)

Typ: 1 - gepaart (Eine Stichprobe)
2 - Zwei Stichproben, gleiche Varianz
3 - Zwei Stichproben, ungleiche Varianz

Diese Funktion gibt einen p Wert an!

Die Bedeutung des p-Wertes der Excel Funktion





Entscheidung mit dem p-Wert

1. Fragestellung (mit der Definition der Population!)
   (Bedingung: Normalverteilung)
2. Nullhypothese - Alternativhypothese
3. Festlegung des Signifikanzniveaus (α)
4. Messung (Stichprobe mit n Messungen, Repräsentativität!)
5. Berechnung des p-Wertes
6. Vergleich von unserem p und dem Signifikanzniveau (α)

$p > \alpha$ | $p < \alpha$

7. Die Entscheidung:

$p > \alpha$	$p < \alpha$
die **Nullhypothese** kann mit einem α Signifikanzniveau **nicht abgelehnt** werden.	die **Nullhypothese** kann mit einem α Signifikanzniveau **abgelehnt** werden
Anhand unserer Messung kann die Alternativhypothese nicht bewiesen werden.	***Die Alternativhypothese ist angenommen*** **(mit einem Signifikanzniveau von α).**

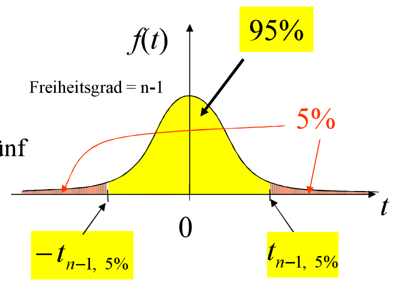
Die Bedeutung des Signifikanzniveaus

Bei einem unwirksamen Medikament beträgt die Wahrscheinlichkeit dafür, dass $|t| > t_{n-1,\alpha}$ ist, 5%.

(⇨Bei der Untersuchung von hundert unwirksamen Pillen werden zufällig fünf als wirksam gefunden!)

⇩

Fehler erster Art



Fehler von 1. und 2. Art

Fehler erster Art:

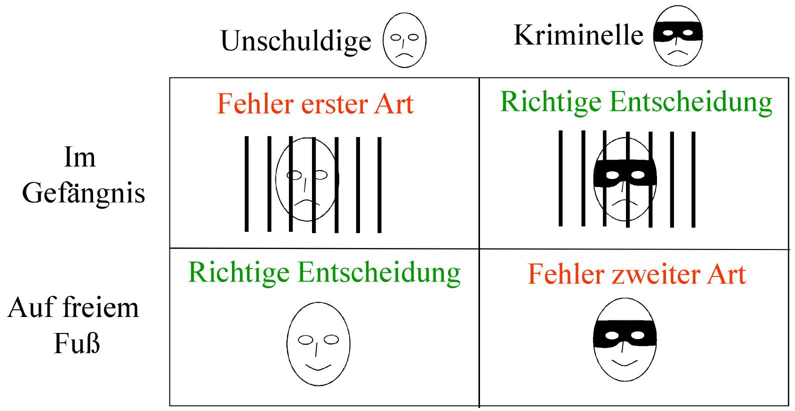
Die Nullhypothese wird zufällig abgelehnt werden, obwohl sie richtig ist!
Die Wahrscheinlichkeit des Fehlers erster Art = Signifikanzniveau
zB: Unwirksame Pille als wirksam gefunden

Fehler zweiter Art:

Die Nullhypothese wird nicht abgelehnt, obwohl sie nicht richtig ist.
Wahrscheinlichkeit =?
zB: Die Wirkung einer Pille ist so klein, dass man es aus der Messung nicht beweisen kann. ⇨ Man braucht noch mehrere Messungen.
⇨ So kleine Wirkung ist oft uninteressant

Fehler von 1. und 2. Art

Nullhypothese: unschuldig

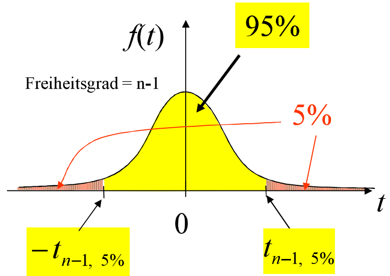


Einseitige/zweiseitige Teste

Ist es interessant wenn das Medikament die Körpertemperatur erhöht?

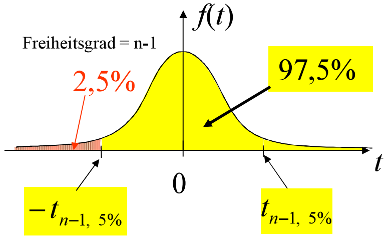
Zweiseitiger Test:

Nullhyp: das Medikament ändert die Körpertemperatur nicht.



Einseitiger Test

Nullhyp: das Medikament erniedrigt die Körpertemperatur nicht.



Verallgemeinerung: $\mu_0 \neq 0$

Beispiel:

Eine Maschine stellt Pillen mit einem nominalen Wirkstoffsgehalt von 20mg her. Man mißt 10 Tabletten und die Wirkstoffgehalte sind (in mg):

20,1 19,8 19,5 17,9 18,8 19,9 18,6 20,3 19,2 19,3

Durchschnitt 19,34 mg, Standardabweichung 0,74 mg, Standardfehler 0,24 mg

Nullhypothese: $\mu_0 = 20$ mg

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}}$$

Alle weitere Schritte sind wie früher.

$t = -2{,}80$