Hypothesenprüfungen

Dr László Smeller

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!

☺

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! $p(H_0 \text{ oder } H_1)=1$

Grundprinzip der Hypothesenprüfungen

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

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

Keine Wirkung:
zB: Fiebermittel: Wenn es keine Wirkung gibt, ist die Temperaturänderung nach der Eingabe = 0.

1. Beispiel: Fibermittel

Seien die Temperaturen vor und nach der Eingabe 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 $\overline{x}$		-0,3

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

1. Beispiel: Fibermittel

Die Nullhypothese entspricht μ = 0

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

Ist $\bar{x} = -0{,}3\,°C$ klein genug um die Nullhypothese anzunehmen?

oder

Ist $\bar{x} = -0{,}3\,°C$ groß genug um die 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 μ=0

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

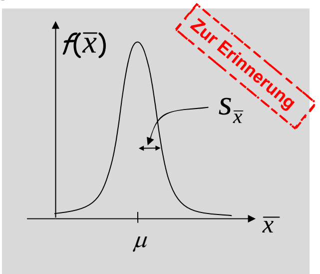
Ist 13 $^1/_{Min}$ klein genug um die Nullhypothese anzunehmen?

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 $s_{\bar{x}}$ sein, wir vergleichen $\bar{x}$ mit $s_{\bar{x}}$.

Definieren wir eine neue Größe:

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

($\bar{x}$ „in $s_{\bar{x}}$ Einheiten” gemessen)

oder mathematisch:

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

t für unseren „Fibermittel“:
$t = -0{,}3/0{,}187 = -1{,}6$
für Kniebeugungen t‘=7,34

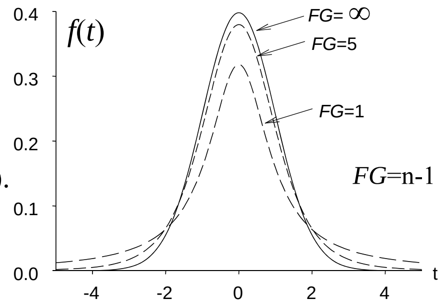
Die Verteilung des *t*-Wertes: die *t*-Verteilung

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

Angenommen, dass, die Messdaten normalverteilt sind, und die Nullhypothese gültig ist, die Verteilung von $t = \frac{\bar{x}}{s_{\bar{x}}}$ kann mathematisch 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.

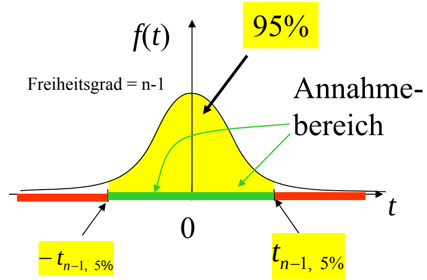


Die Anwendung der *t*-Verteilung

Die Wahrscheinlichkeit, dass bei Richtigkeit der 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. ⇒ Wir müssen keine Wirkung voraussetzen.

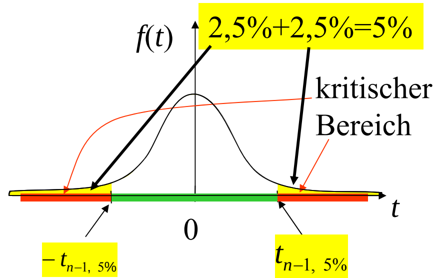
Die Nullhypothese kann nicht abgelehnt werden.

Die Anwendung der *t*-Verteilung

Die Wahrscheinlichkeit, dass bei Richtigkeit der Nullhypothese

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

gilt, beträgt 5%.



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

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

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 ($t_{n-1,\alpha}$)

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

7. Die Entscheidung:

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).**

Beispiel des Fiebermittels

α=5%

$\bar{x} = -0{,}30°C \quad s_{\bar{x}} = 0{,}187°C$

$n = 5 \qquad FG = 4$

$$t = \frac{-0{,}30}{0{,}187} = -1.6$$

$$-2{,}776 = -t_{4;\,5\%} < t < +t_{4;\,5\%} = 2{,}776$$

$\Rightarrow$ ***t*** liegt in dem Annahmebereich, die Nullhypothese kann nicht abgelehnt werden, $\Rightarrow$

$\Rightarrow$ Das Medikament ist unwirksam

(mit p=5% Irrtumswahrscheinlichkeit)

FG \ p	0.05	0.02	0.01
2	4.30266	6.96455	9.92499
3	3.18245	4.54071	5.84085
4	2.77645	3.74694	4.60408
5	2.57058	3.36493	4.03212
6	2.44691	3.14267	3.70743
7	2.36462	2.99795	3.49948
8	2.30601	2.89647	3.35538
9	2.26216	2.82143	3.24984
10	2.22814	2.76377	3.16926
11	2.20099	2.71808	3.10582
12	2.17881	2.68099	3.05454
13	2.16037	2.65030	3.01228
14	2.14479	2.62449	2.97685
15	2.13145	2.60248	2.94673
20	2.08596	2.52798	2.84534
50	2.00856	2.40327	2.67779
70	1.99444	2.38080	2.64790
100	1.98397	2.36421	2.62589
unendlich	1.95996	2.32635	2.57583

Beispiel der Kniebeugungen

$\bar{\boldsymbol{x}} = 13 \quad \boldsymbol{s}_{\bar{x}} = 1{,}77$

$n = 6 \qquad FG = 5$

$$\boldsymbol{t} = \frac{13}{1{,}77} = 7.34$$

$$t > +t_{5;\,5\%} = 2{,}57$$

$\Rightarrow$ ***t*** liegt in dem kritischen Bereich, die Nullhypothese kann abgelehnt werden, $\Rightarrow$

$\Rightarrow$ Die „Behandlung“ ist wirksam

(mit p=5% Irrtumswahrscheinlichkeit)

Auch bei 2 % und bei 1% Irrtumswahrscheinlichkeit!

FG \p	0.05	0.02	0.01
2	4.30266	6.96455	9.92499
3	3.18245	4.54071	5.84085
4	2.77645	3.74694	4.60408
5	2.57058	3.36493	4.03212
6	2.44691	3.14267	3.70743
7	2.36462	2.99795	3.49948
8	2.30601	2.89647	3.35538
9	2.26216	2.82143	3.24984
10	2.22814	2.76377	3.16926
11	2.20099	2.71808	3.10582
12	2.17881	2.68099	3.05454
13	2.16037	2.65030	3.01228
14	2.14479	2.62449	2.97685
15	2.13145	2.60248	2.94673
20	2.08596	2.52798	2.84534
50	2.00856	2.40327	2.67779
70	1.99444	2.38080	2.64790
100	1.98397	2.36421	2.62589
unendlich	1.95996	2.32635	2.57583

Hypothesenprüfung mit Excel

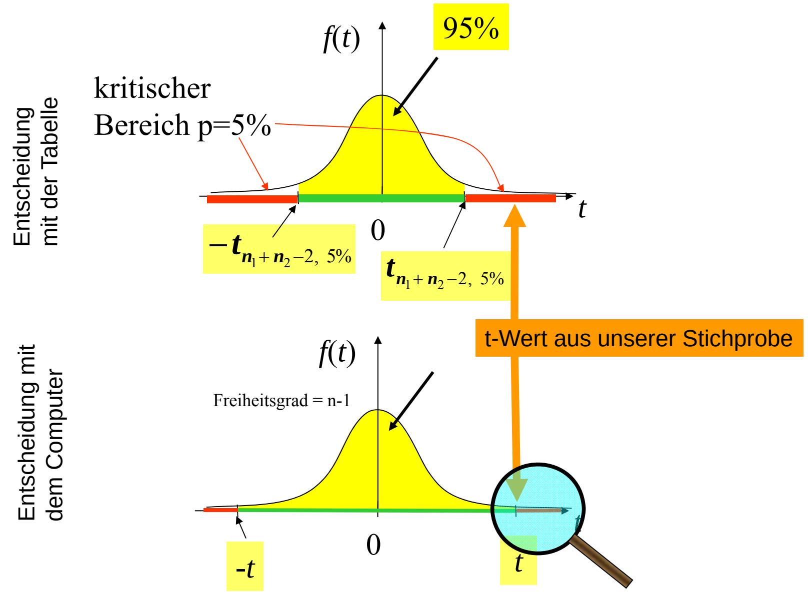
Excel Funktion für *t*-Teste:

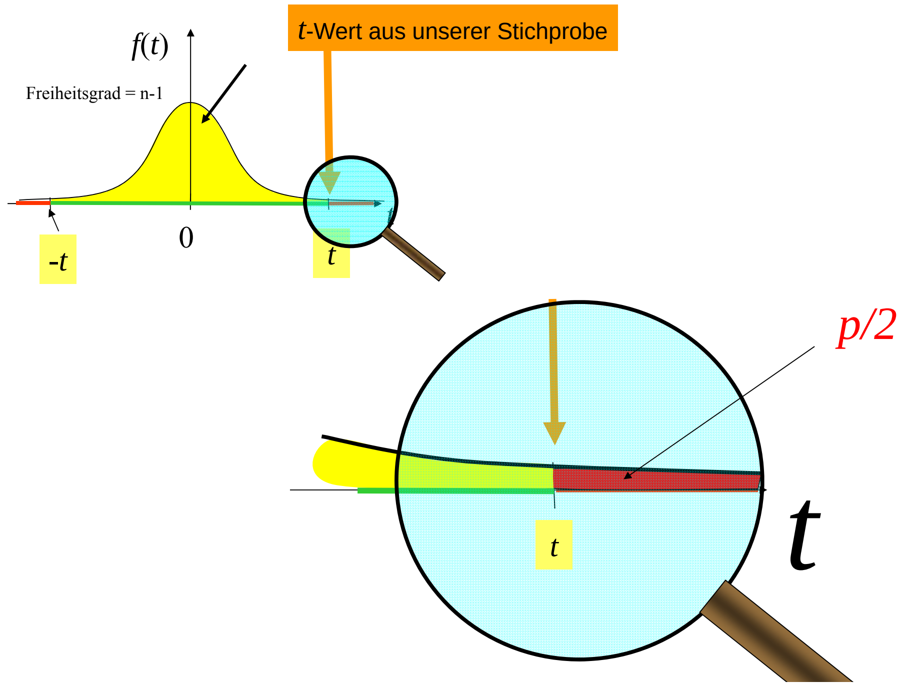
ttest(Reihe1; Reihe2; Seiten; Typ)

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:

die **Nullhypothese** wird mit einem Signifikanzniveau von α **angenommen** werden.

Anhand unserer Messung kann die Alternativhypothese nicht bewiesen werden.

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

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