Grundlagen der Biostatistik und Informatik

Hypothesenprüfungen II.

Zwei Stichproben *t*-Test, F-Test,
Bedingungen der Anwendung der *t*-Teste
Varianzanalyse

László Smeller

Widerholung: Grundprinzip der Hypothesenprüfungen

Zu entscheidende Frage
Indirekter Beweis
Nullhypothese (H_0): nur zufällige Änderungen
mathematisch behandelbar
Ein geeigneter Parameter (zB. t)
Bei Gültigkeit der H_0 t folgt einer gut bestimmten Verteilung
Zu 95% $|t|<t_{FG,5\%}$
Wenn $|t|> t_{FG,5\%}$ $\Rightarrow$ H_0 wird abgelehnt zu 5% Irrtumswahrscheinlichkeit.
$|t|< t_{FG,5\%}$ $\Rightarrow$ H_0 kann nicht abgelehnt werden (bei 5% Irrtumsw.).

Übersicht der Teste

Stichproben \ Verteilung	Normalverteilte Daten	Die Verteilung der Daten ist unbekannt
Eine Stichprobe	Einstichproben t-Test	Vorzeichnentest Wilcoxon Test
Zwei Stichproben	Zweistichproben t-test	Mann-Whittney U-Test
Mehrere Stichproben	ANOVA (Varianzanalyse)	Kruskal-Wallis Test

Zweistichproben t-Test

Vergleich von zwei Stichproben (zwei Populationen)

Warum?
- zwei wesentlich unterschiedliche Populationen
  (z.B.: Männer und Frauen)
- Vermeidung des Placeboeffektes
  (Placebo: Pille ohne Wirkstoff)
- ethische Hinsicht: kein Patient darf unbehandelt bleiben: Vergleich von alte und neue Medikamente oder Behandlungen.

Zweistichproben t-Test: Frage, Nullhypothese

Frage: Ist der zu vergleichende Parameter unterschiedlich in der zwei Populationen?

Mathematisch: Sind die Erwartungswerte in der zwei Populationen unterschiedlich? (Stammen die zwei Stichproben aus einem Population?)

$$\mu_1 \neq \mu_2$$

Nullhypothese: Es gibt kein Unterschied, die Erwartungswerte sind gleich: $\mu_1 = \mu_2$

Zweistichproben t-Test : Beispiel

Ist die Körperhöhe der Jungen höher als die Körperhöhe der Mädchen?

Zwei Populationen: Jungen u. Mädchen

Nullhypothese:

Jungen u. Mädchen sind gleich hoch.

Der Erwartungswert der Körperhöhe der Jungen ist gleich der Erwartungswert der Körperhöhe der Mädchen

$$\mu_{\text{Mädchen}} = \mu_{\text{Jungen}}$$

Zweistichproben t-Test

Auch wenn $\mu_{\text{Mädchen}}=\mu_{\text{Jungen}}$ die Durchschnittwerte können unterschiedlich sein:

$$\overline{x}_{\text{Mädchen}} \neq \overline{x}_{\text{Jungen}}$$

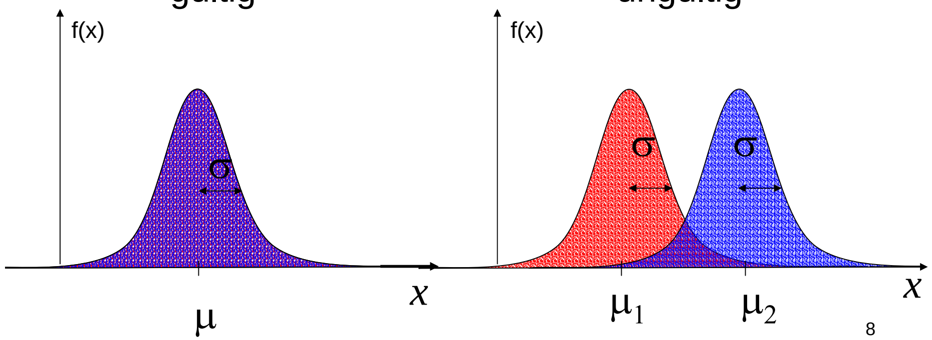
Ist dieser Unterschied zufällig (statistisch), oder ist es die Konsequenz des Unterschiedes zwischen der zwei Populationen (Mädchen u. Jungen)?

Nehmen wir an:

- Die beide Verteilungen sind Normalverteilungen,
- und die Varianzen (Streuungen) sind gleich

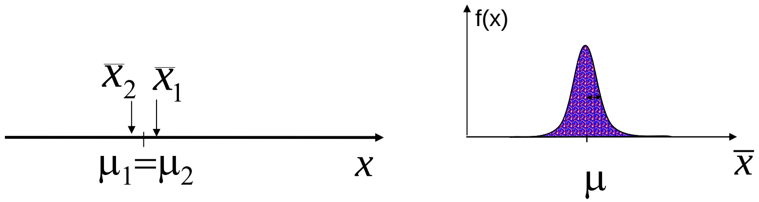
(Bedingungen des Zweistichproben t-Testes)

Nullhypothese ist gültig	Nullhypothese ist ungültig



Aus der zwei Populationen nimmt man zwei Stichproben, man kann die zwei Durchschnitts-werte vergleichen.

Angenommen dass die Nullhypothese gültig ist:



Ist $\bar{x}_1 - \bar{x}_2$ zufällig oder groß genug um die Nullhypothese abzulehnen?

Die Berechniund des Parameters *t*

Wir brauchen einen Parameter ähnlich zu *t* beim Einstichprobentest

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{Q_{x1} + Q_{x2}}{n_1 + n_2 - 2}}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

$$Q_{x1} = \sum_{i=1}^{n_1} (x_{1i} - \bar{x}_1)^2$$

$$Q_{x2} = \sum_{i=1}^{n_2} (x_{2i} - \bar{x}_2)^2$$

Ähnlichkeit zum Einstich-probentest:

$$t = \frac{\bar{x}}{s_{\bar{x}}} = \frac{\bar{x}}{s_x}\sqrt{n} = \frac{\bar{x}}{\sqrt{\frac{Q_x}{n-1}}}\sqrt{n}$$

Der Parameter *t*

Bei Gültigkeit der Nullhypothese ***t*** folgt eine *t*–Verteilung mit Freiheitsgrad von n_1+n_2-2.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{Q_{x1} + Q_{x2}}{n_1 + n_2 - 2}}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

Entscheidung:
wie bei Einstichproben *t*-Test

Die Entscheidung

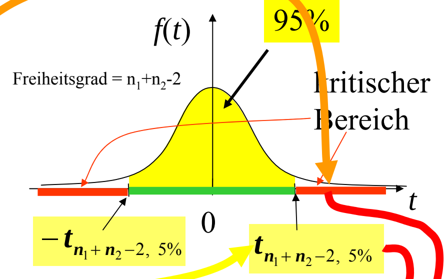
5 %

t-Tabelle

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	[illegible]	[illegible]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

zB: $n_1=7$, $n_2=9$ $t=2{,}35$

FG: $n_1+n_2-2=14$ $t_{FG,5\%}=2{,}14479$



$t > t_{FG,5\%}$

$|t| > t_{n_1+n_2-2,\ 5\%}$ Die H_0 ablehnen (die zwei Populationen sind unterschiedlich

$|t| < t_{n_1+n_2-2,\ 5\%}$ Die H_0 ist gültig (die zwei Populationen sind nicht unterschiedlich

Gepaarte –ungepaarte Teste

Einstichprobentest

Name	T_{vor}	T_{nach}
Anna	39,7	39,2
Benjamin	38,8	38,4
Christina	37,9	38,7
Daniel	39,2	38,7

Gepaarte Daten

Zweistichprobentest

Name	Höhe [cm]
Benjamin	189
Christian	175
Daniel	180
Gabriel	165
Henrik	187

Name	Höhe [cm]
Anna	175
Eva	155
Frederike	167
Judith	180

Ungepaarte Daten

Diese Daten können nicht in Paare geordnet werden

Vergleich der Effektivität der gepaarten-ungepaarten Teste

Ungepaarte Test
Zweistichproben t-Test

Kein signifikanter Unterschied

Gepaarte Test:
Einstichproben t-Test

Signifikanter Unterschied

F-test

Frage:
Sind die Varianzen in zwei Stichproben Gleich?

Nullhypothese: Die Varianzen sind gleich

Parameter: $\boxed{F = \frac{s_1^2}{s_2^2}}$ $\quad s_1 > s_2$

Bei der Gültigkeit der Nullhypothese F folgt eine F-Verteilung mit n_1-1 und n_2-1 Freiheitsgrade
Bemerkung: Tabelle zum einseitigen Test!
Wir brauchen einen zweiseitigen Test!

F-test

$$\boxed{F = \frac{s_1^2}{s_2^2}}$$

Aus der Tabelle

$F < F_{n_1-1, n_2-1; 5\%}$ Nullhypothese ist gültig zu 5% Irrtumswahrsch.
d.h. die Varianzen sind gleich

Gut für Zweistichproben t-Test

$F > F_{n_1-1, n_2-1; 5\%}$ Nullhypothese ist ungültig zu 5% Irrtumsw.
d.h. die Varianzen sind nicht gleich

Wenn die Streuungen sind ungleich

Die Daten können transformiert werden so dass der Zweistichprobentest durchgeführt werden kann. (Excel kann diese Transformation ausrechnen).

Hypothesenprüfungen mit Excel

Excel Funktion für t-Teste:
(Ein- u. Zweistichproben t-Teste)

ttest(Reihe1; Reihe2; Seiten; Typ)

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

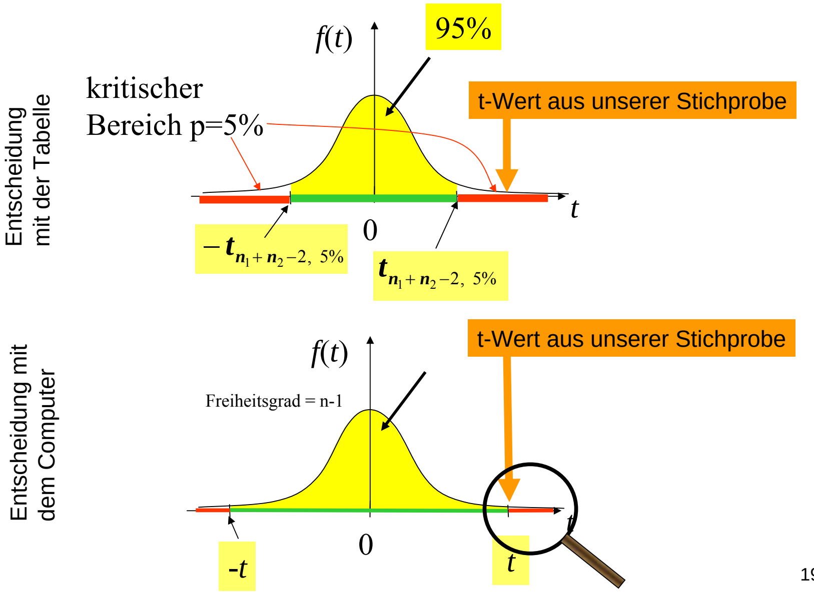
Ftest(Reihe1; Reihe2)

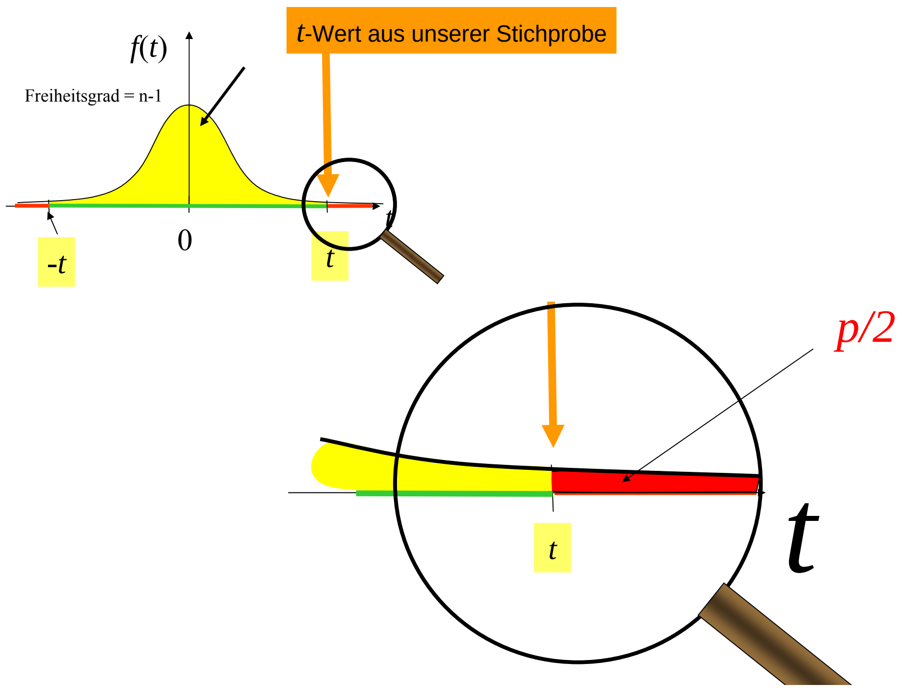
Diese Funktionen geben p an

Entscheidung: $p<5\%$ H_0 wird mit p Sing.N abgelehnt
$p>5\%$ H_0 wird nicht abgelehnt (5% S.N.)

Bemerkung: die Ftest() Funktion im Excel gibt p des zweiseitigen Tests

Die Bedeutung des p-Wertes der Excel Funktion





Zusammenfassung: Zweistichproben *t*-Test

Vergleich von zwei Populationen durch zwei Stichproben

Bedingung: Normalverteilung mit derselben Varianz

Prüfung der Varianzen: F-Test

Die Varianzen sind: gleich ungleich

↓

Transformation

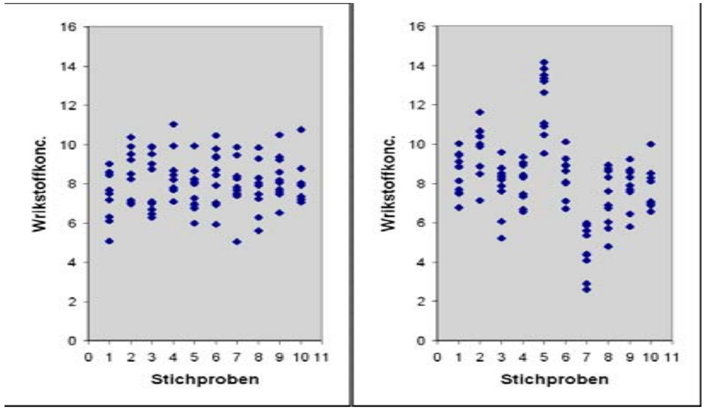
Berechnung des *t*-Wertes oder des p-Wertes

Ist $t > t_{n-2\,,\,5\%}$ oder p<5% ?

→ ja: H_0 mit 5% oder p Irrtumswahrsch. ablehnen

→ nein: H_0 kann nicht abgelehnt werden (mit 5% bzw. p Irrtumswahrsch.)

Vergleich von mehreren Stichproben



Vergleich von mehreren Stichproben

Bonferroni - Problem

Vergleich von mehreren Stichproben

Paarweise Vergleichungen:

- Hohe Wahrscheinlichkeit des Fehlers von 1. Art
- z.B.: 10 Stichproben, 45 Vergleichungen
  alle mit 5% Irrtumswahrscheinlichkeit
  Gesamtirrtumsw.:→ $1-(1-0{,}05)^{45}=90{,}0\%$

Parametrische Methode: ANOVA
(**AN**alysis **O**f **VA**riance)

ANOVA

Vorbedingungen:

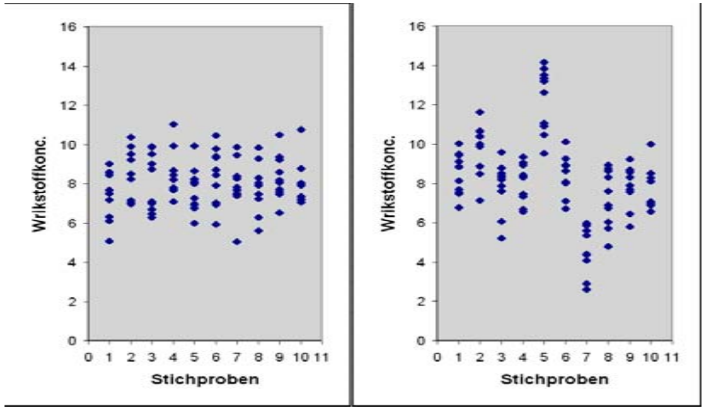
- Unabhängigkeit der Stichproben
- Normalverteilung
- Gleiche Streuungen

H_0: Alle Stichproben stammen aus der selben Grundgesamtheit

H_1: Mindestens *eine* Stichprobe stammt aus einer anderen Grundgesamtheit

ANOVA

Wenn H_0 gültig ist, sollen die Streuungen *zwischen* den Stichproben und *innerhalb* der Stichproben dieselbe sein.



ANOVA

h Stichproben

$$H_0: \mu_1 = \mu_2 = \dots \mu_h$$

Zwei unabhängige Varianzschätzungen:

Varianz innerhalb der Stichproben: S_i^2

Varianz zwischen den Stichproben: S_g^2

Wenn $S_i^2 << S_g^2 \rightarrow$ Varianzen sind unterschiedlich $\rightarrow$ H_0 ablehnen

Wenn $S_i^2 \approx S_g^2 \rightarrow$ Varianzen sind die Schätzungen derselben Varianz $\rightarrow$ H_0 annehmen

$$F = \frac{S_g^2}{S_i^2}$$ *F* – Test; Einseitig, Freiheitsgrad: *h-1; N-h*

ANOVA

Varianz zwischen den Stichproben:

$$s_g^2 = \frac{\sum_{j=1}^{h} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2}{h-1} = \frac{Q_g}{h-1}$$

h: Anzahl der Stichproben

n_j: Anzahl der Elementen in der *j*-ten Stichprobe

$\bar{\bar{x}}$: Durchschnitt von allen Elementen

$\bar{x}_j$: Durchschnitt in der *j*-ten Stichprobe

ANOVA

Varianz innerhalb der Stichproben:

$$s_i^2 = \frac{\sum_{j=1}^{h} Q_j}{N-h} = \frac{\sum_{j=1}^{h}\sum_{i=1}^{n_j}(x_{ij}-x_j)^2}{N-h} = \frac{Q_i}{N-h}$$

h: Anzahl der Stichproben

n_j: Anzahl der Elementen in der *j*-ten Stichprobe

x_{ij}: *i*-ten Element der *j*-ten Stichprobe

$\bar{x}_j$: Durchschnitt in der *j*-ten Stichprobe

N: Gesamte Anzahl der Stichprobenelementen

ANOVA

$$F = \frac{S_g^2}{S_i^2}$$

$\downarrow$

F – Test

FG: h-1; N-h

Nicht Signifikant,
keine Unterschiede

H_0: OK

Signifikant,
Es gibt Unterschiede

$\cancel{H_0}$

$\downarrow$

Post-hoc Test
(Paarweise Vergleichungen
der Stichproben, z.B. Min-Max)

(Paarweise Vergleichungen,
z.B: gegen Kontrolle)

Ende

Ich habe es jetzt statistisch beweist, daß alle diese Kristallkugeln sind gleich gut für Wahsagung!