

Grundlagen der Biostatistik und Informatik

Hypothesenprüfungen II.

Fehlern von erste u. zweite Art,
Zwei Stichproben t -Test, F-Test,
Bedingungen der Anwendung der t -Teste
Varianzanalyse

László Smeller

Widerholung: Grundprinzip der Hypothesenprüfungen

Zu entscheidende Frage

Indirekter Beweis

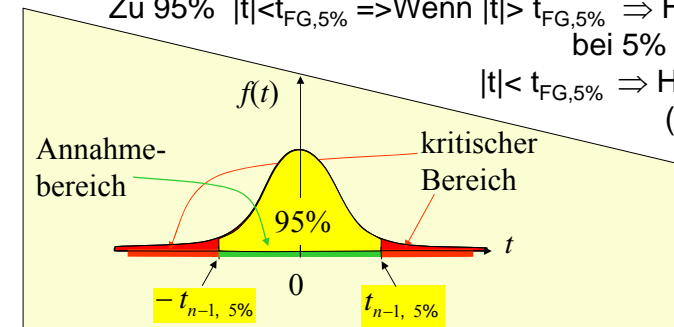
Nullhypothese (H_0): nur zufällige Änderungen
mathematisch behandelbar

Ein geeigneter Parameter (zB. t)

Bei Gültigkeit der H_0 t folgt einer gut bestimmten Verteilung
(t -Verteilung)

Zu 95% $|t| < t_{FG,5\%} \Rightarrow$ Wenn $|t| > t_{FG,5\%} \Rightarrow H_0$ wird abgelehnt
bei 5% Irrtumswahrsch.

$|t| < t_{FG,5\%} \Rightarrow H_0$ wird beibehalten
(bei 5% Irrtumsw.).



2

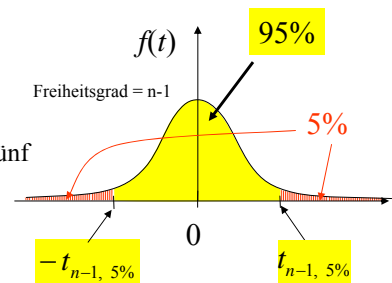
Die Bedeutung des Signifikanzniveaus

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

(\Rightarrow 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

Auch als α -Fehler genannt.

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. \Rightarrow Man braucht noch mehrere Messungen.*

\Rightarrow So kleine Wirkung ist oft uninteressant

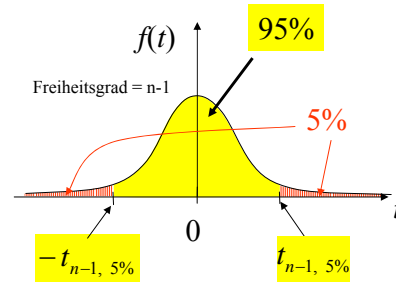
Auch als β -Fehler genannt.

Einseitige/zweiseitige Teste

Ist es interessant wenn das Medikament die Körpertemperatur erhöht?

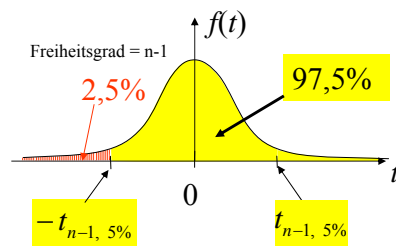
Zweiseitiger Test:

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 Wirkstoffgehalt 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

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

$t = -2,80$

Übersicht der Teste

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

7

Zweistichproben t-Test

Vergleich von zwei Stichproben (zwei Populationen)

Warum?

- zwei wesentlich unterschiedliche Populationen (z.B.: Männer und Frauen)
- Vermeidung des Placeboeffektes mit Anwendung einer Kontrollgruppe. (Doppelblindstudie) (Placebo: Pille ohne Wirkstoff)
 - Randomisierung ist wichtig! (wenn möglich)
 - ethische Hinsicht: kein Patient darf unbehandelt bleiben: Vergleich von alte und neue Medikamente oder Behandlungen.

8

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?
(oder stammen die zwei Stichproben aus einer Population?)



$$\mu_1 \neq \mu_2$$

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

9

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

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

10

Zweistichproben t-Test

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

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

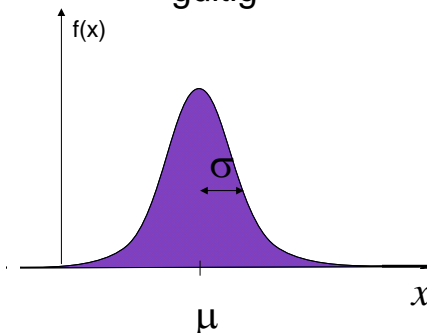
Ist dieser Unterschied zufällig (statistisch), oder ist es die Konsequenz des Unterschiedes zwischen der zwei Populationen (Mädchen u. Jungen)?

11

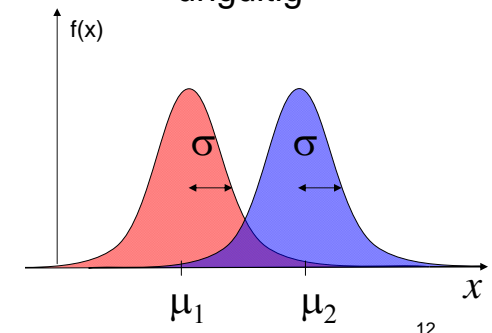
Nehmen wir an:

- Die beide Verteilungen sind Normalverteilungen,
- und die Varianzen (Streuungen) sind gleich (Bedingungen des Zweistichproben t-Testes)

Nullhypothese ist gültig



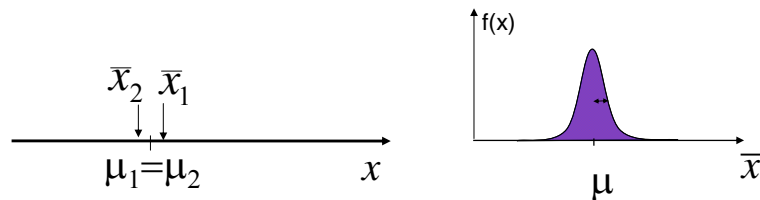
Nullhypothese ist ungültig



12

Aus der zwei Populationen nimmt man zwei Stichproben, man kann die zwei Durchschnittswerte vergleichen.

Angenommen dass die Nullhypothese gültig ist:



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

13

Die Berechnung 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 Einstichprobentest:

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

14

Der Parameter t

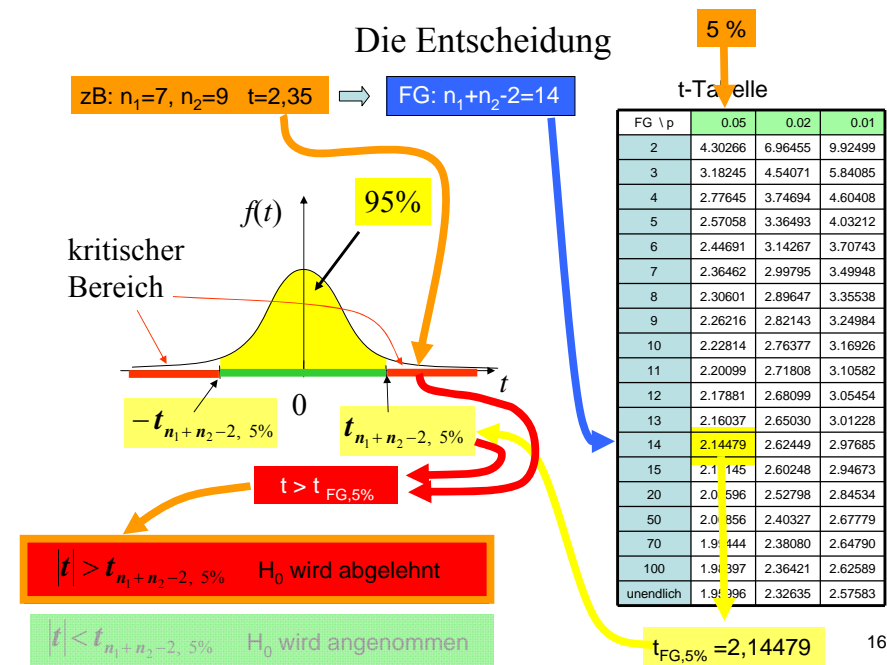
Bei Gültigkeit der Nullhypothese t folgt eine t -Verteilung mit Freiheitsgrad von $n_1 + n_2 - 2$.

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

Entscheidung:
wie bei Einstichproben t -Test

15

Die Entscheidung



16

F-test

Frage:

Sind die Varianzen in zwei Stichproben Gleich?

Nullhypothese: Die Varianzen sind gleich

Parameter:

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

$s_1 > s_2$

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!

17

F-test

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

Aus der Tabelle

$$F < F_{n_1-1, n_2-1; 5\%}$$

Nullhypothese ist gültig zu 5% Irrtumswahrsch.
d.h. die Varianzen sind gleich

Erfüllt die Bedingungen des
Zweistichproben t-Testes

$$F > F_{n_1-1, n_2-1; 5\%}$$

Nullhypothese ist ungültig zu 5% Irrtumsw.
d.h. die Varianzen sind nicht gleich



18

Wenn die Streuungen sind ungleich

Die Daten können transformiert werden so dass der Zweistichprobentest durchgeführt werden kann.
Auch als Welch-Test bekannt.

(Excel kann diese Transformation ausrechnen).

TTEST(Daten1;Daten2;Seiten;3)



19

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

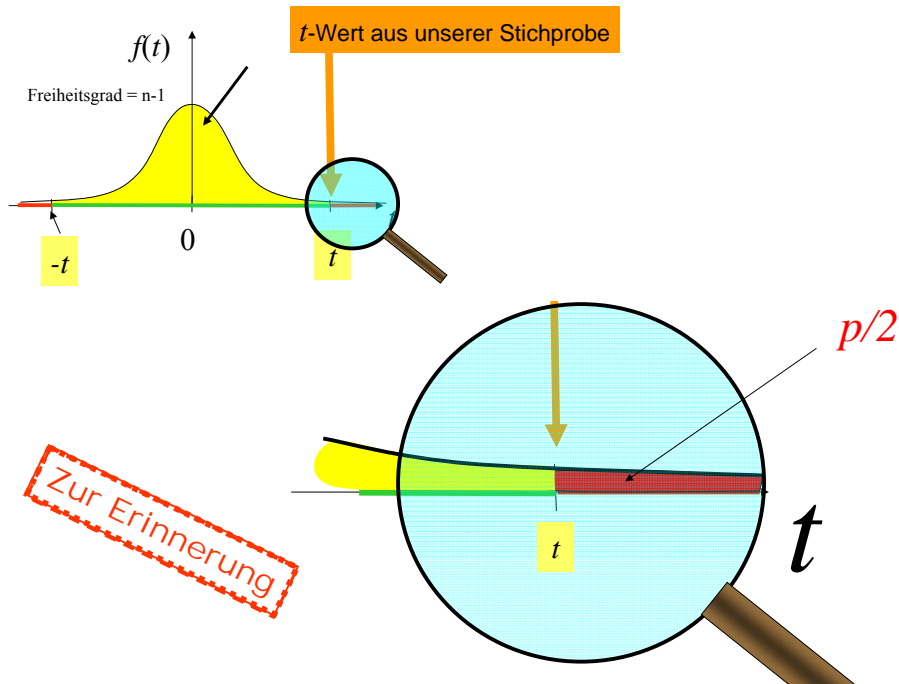
Ftest(Reihe1; Reihe2)

Diese Funktionen geben **p** an

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

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

20



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 (oder Typ 3 in Excel)

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.)

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

Zweistichprobentest

Name	Höhe [cm]	Name	Höhe [cm]
Benjamin	189	Anna	175
Christian	175	Eva	155
Daniel	180	Frederike	167
Gabriel	165	Judith	180
Henrik	187		

Gepaarte Daten

Ungepaarte Daten

Diese Daten können nicht in Paare geordnet werden

23

Vergleich der Effektivität der gepaarten-ungepaarten Teste

Ungepaarte Test
Zweistichproben t -Test

Kein signifikanter Unterschied



Gepaarte Test:
Einstichproben t -Test

Signifikanter Unterschied



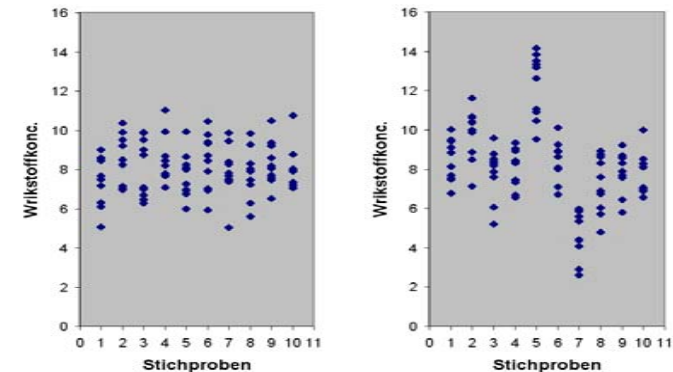
24

Übersicht der Teste

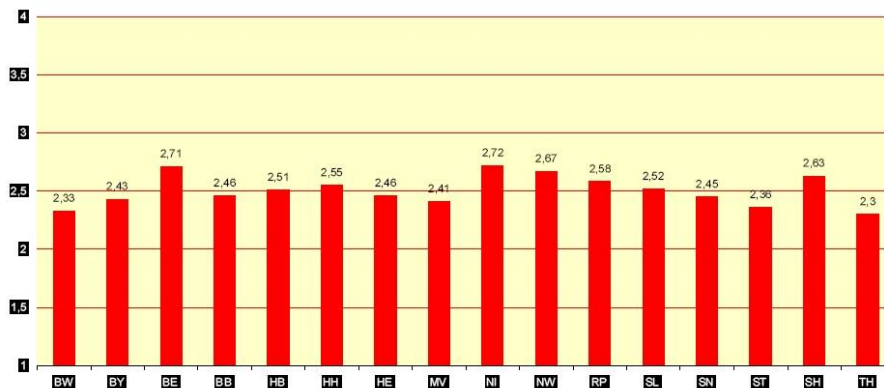
Verteilung Stichproben	Normalverteilte Daten	Die Verteilung der Daten ist unbekannt
Eine Stichprobe	Einstichproben t-Test ✓	Vorzeichentest Wilcoxon Test
Zwei Stichproben	Zweistichproben t-test ✓	Mann-Whittney U-Test
Mehrere Stichproben	ANOVA (Varianzanalyse)	Kruskal-Wallis Test

25

Vergleich von mehreren Stichproben ANOVA



Vergleich von mehreren Stichproben



Bonferroni - Problem

Vergleich von mehreren Stichproben

Paarweise Vergleichen:

- Hohe Wahrscheinlichkeit des Fehlers von 1. Art
- z.B.: 10 Stichproben, 45 Vergleichen
alle mit 5% Irrtumswahrscheinlichkeit

Gesamtirrtumsw.: $\rightarrow 1 - (1 - 0,05)^{45} = 90,0\%$

Lösung (für normalverteilte Daten): **ANOVA**
(ANalysis Of VAriance)

ANOVA

Vorbedingungen:

- Unabhängigkeit der Stichproben
- Normalverteilung
- Gleiche Streuungen

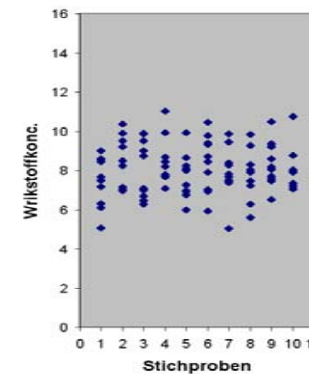
H_0 : Alle Stichproben stammen aus der selben Grundgesamtheit

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

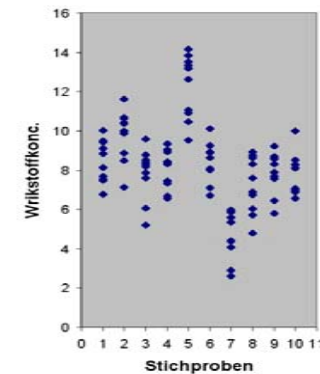
ANOVA

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

Beispiel 1.



Beispiel 2.



ANOVA

h : Anzahl der Stichproben

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

Varianz innerhalb der Stichproben: S_i^2

Varianz zwischen den Stichproben: S_g^2

Wenn $S_i^2 \ll S_g^2 \rightarrow$ 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

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

ANOVA

Varianz zwischen den Stichproben:

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

h : Anzahl der Stichproben

n_j : Anzahl der Elementen in der j -ten Stichprobe

\bar{x} : Durchschnitt von allen Elementen

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

ANOVA

Varianz innerhalb der Stichproben:

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

h : Anzahl der Stichproben

n_j : Anzahl der Elementen in der j -ten Stichprobe

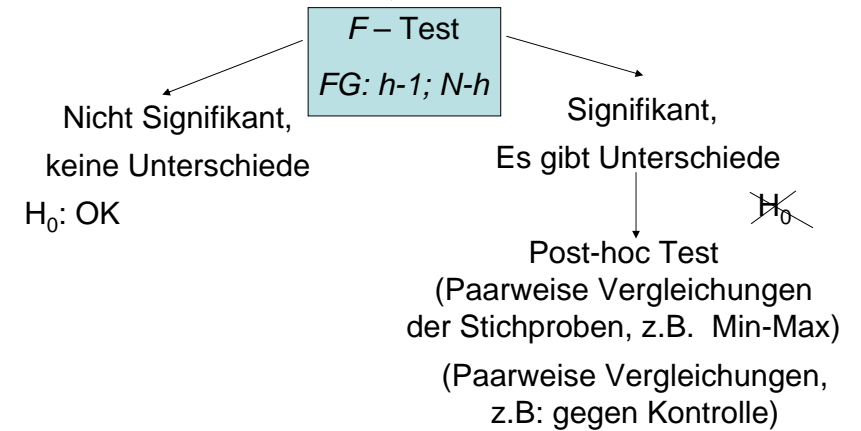
x_{ij} : i -ten Element der j -ten Stichprobe

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

N : Gesamte Anzahl der Stichprobenelementen

ANOVA

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



Ende

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

