Varianzanalyse (linke Gruppe)

4.1 Wie lässt sich die Varianzanalyse im Vergleich zum t- Test charakterisieren? Wieso werden hier Mittelwertsunterschiede mit dem F- Test, einem Test für Varianzunterschiede, geprüft? Was versteht man in der Varianzanalyse unter einem Faktor, was unter einer Faktorenstufe?

Wie lässt sich die Varianzanalyse im Vergleich zum t- Test charakterisieren?
t- Test:

- Signifikanztest für den Vergleich zweier unabhängiger Stichproben (einer abhängigen Variablen in zwei unabhängigen Stichproben)
- dabei wird angenommen, daß es sich dabei um in der Population normalverteilte Variablen handelt die Berechnung von Mittelwert und Varianz sind also dabei sinnvolle und zulässige Operatoren
- Verfahren, die diese Voraussetzung erfüllen, werden parametrisch genannt.
  t- Test für den Vergleich zweier Stichprobenmittelwerte

⇨ Im Grunde ist der t-Test ein Spezialfall (oder wie Glaser ihn bezeichnet: ein Grenzfall) der VA. Vergleicht man nämlich die Mittelwerte von nur 2 Gruppen, so stimmen die Ergebnisse von VA und t-Test überein (Vgl. Bortz S. 228).

Varianzananlyse:

- kann als verallgemeinerter t- Test für mehr als zwei Stichproben aufgefaßt werden (t- Test: Grenzfall der Varianzanalyse für nur 2 Gruppen)
- parametrischer Signifikanztest für den Mittelwertsvergleich einer abhängigen Variablen in mehr als zwei Stichproben
- mit dem Varianztest bei mindesten 2 UV Darstellung von Wechselwirkungen (bei t- Test nicht möglich)

⇨ Prinzipiell könnte man auch mit entsprechend vielen t-Tests alle Paare von Mittelwerten vergleichen. Problematisch dabei ist allerdings, dass sich durch das mehrmalige Testen die Gefahr eines α–Fehlers erhöht. Um dies zu verhindern, müsste die Irrtumswahrscheinlichkeit (bzw. das Signifikanzniveau) modifiziert werden.

einfaktorielle Varianzanalyse: 1 UV mit mehreren Stufen, 1 AV, einfachste mögliche Varainzanalyse
Zweifaktorielle Varianzanalyse: 2 UV mit mehreren Stufen, 1 AV, Hinzunahme einer zweiten unabhängigen Variablen; Wechselwirkung taucht erstmals auf (2 Haupteffekte A, B und Wecheselwirkung A * B)
Dreifaktorielle Varianzanalyse: 3 UV mit mehreren Stufen, Hinzunahme einer dritten unabhängigen Variable, drei Wechselwirkungen zwischen je einer UV (A * B, A * C, B * C) und eine Dreifachwechselwirkung (A * B * C)
Mehrfaktorielle Varianzanalyse: mehrere UV mit mehreren Stufen,

Wieso werden hier Mittelwertsunterschiede mit dem F- Test, einem Test für Varianzunterschiede, geprüft?
F- Test:
Mit dem F-Test wird geprüft, weil eine Voraussetzung der VA die Varianzhomogenität ist, d.h. die Fehler in den verschiedenen Gruppen sollten die gleiche Varianz haben.
Die grundsätzliche Idee der VA geht davon aus, dass zur Überprüfung von Mittelswertsunterschieden die Varianz der Daten analysiert wird. Der F-Test prüft, ob sich zwei Varianzen signifikant voneinander unterscheiden. Sollten sie nicht signifikant sein ,so spricht man von homogenen Varianzen. Bei einem signifikanten F-Test stammen die Stichproben, die den Varianzschätzungen zugrunde liegen, nicht aus der gleichen Population. Nun kann die Ho zugunsten der Alternativhypothese verworfen werden.

- Obwohl die ANOVA Mittelwertsunterschiede prüft, erhielt sie ihren Namen aus der Tatsache, daß (geschätzte) Varianzen Größen sind, die einen quantitativen Vergleich von Mittelwertsunterschieden ermöglichen

- s^2_{IG}: gewogenes arithmetisches Mittel der einzelnen Varianzschätzungen innerhalb der Gruppen
- s^2_{ZG}: Varianzschätzung auf der Basis der einzelnen Gruppenmittelwerte
- Die Varianzschätzung ZG wird verglichen mit der Varianz, die wir erwarten würden
  (Varianz bzw. Standarabweichung sollte größer sein als Standardfehler des Mittelwertes
  wenn Mittelwerte stärker streuen, als es dem Stichprobenfehler des Mittelwertes
  entspricht: H_1
- Varianzunterschied wird einseitig durchgeführt (wenn $\hat{\sigma}^2_{Fehler} > \hat{\sigma}^2_{treat}$ Durchführung keine Durchführung des F – Test, da Treatmentunterschiede unbedeutend)
- Tabelle E Werte F > 1 tabelliert
- Diese einseitige Varianzüberprüfung entspricht jedoch die Überprüfung einer ungerichteten Mittelwertshypothese $H_0 = \mu_1 = \mu_2 \dots \mu_p$ $H_1 = \mu_i \neq \mu_{i'}$ (mindestens 2 Mittelwerte sind ungleich)
- Der einseitige F- Test überprüft eine ungerichtete Alternativhypothese bezüglich der Mittelwerte (Varianz der Verteilung der Mittelwerte verglichen mit der Varianz, die wir erwarten würden)

Kürzere Antwort (=Glaser):

Der Test für Varianzunterschiede (der F-Test) prüft hier Mittelwertsunterschiede, weil diese in einer Varianzschätzung, eben der Varianzschätzung zwischen den Gruppen, ihren Ausdruck finden.

Weitere Frage: warum finden Mittelwertsunterschiede ihren Ausdruck in einer Varianzschätzung?

Weil die Varianzschätzung zwischen den Gruppen genau so definiert ist: Die Mittelwerte werden dabei wie Maßzahlen behandelt, für die eine Varianzschätzung berechnet wird.

Kommentar [a1]:

$F = s^2_{ZG} / s^2_{IG}$

$F = (S_{ZG} / df_{ZG}) / (S_{IG} / df_{IG})$

$F = \dfrac{(\sum 1/m_j (\sum x_{ij})^2 - (\sum\sum x_{ij})^2 / N) / n-1}{\sum\sum x_{ij}^2 - (\sum 1/m_j (\sum x_{ij})^2 / N - n}$

Varianzanalyse:

1. F- Test:

Signifikanztest zur Überprüfung der der generellen H_0

$H_{0\,generell}$: $\mu_1 = \mu_2 = \dots = \mu_p$ und $\sigma^2_1 = \sigma^2_2 = \dots \sigma^2_n$

sämtliche Effekte α_j in der Population gleich Null

Gesamtmittelwert bester Schätzer der Population (Mittelwert der Mittelwertsverteilung)

H_1: : $\mu_1 \neq \mu_2$ aber $\sigma^2_1 = \sigma^2_2 = \dots \sigma^2_n$

Stichprobenmittelwert bester Schätzer für Population

H_0 und H_1: Erwartungswert von $s^2_{IG} = \sigma^2_e$;

H_0 :Erwartungswert von $s^2_{ZG} = \sigma^2_e$ (sämtliche Effekte sind gleich Null);

H_1 :Erwartungswert von $s^2_{ZG} > \sigma^2_e$ (falls mindestens ein Effekt ungleich Null ist)

Erwartungswert s^2_{ZG} = Erwartungswert s^2_{IG} sämtliche Effekte gleich Null

Erwartungswert s^2_{ZG} > Erwartungswert s^2_{IG} falls mindestens ein Effekte ungleich Null ist

Aufgrund dieser Erwartungswerte für die mittleren Abweichungsquadrate S^2_{ZG} und S^2_{IG} werden diese in folgendes Verhältnis gesetzt:

F - Test: $F = S^2_{ZG} / S^2_{IG}$ $F = \hat{\sigma}^2_{ZG} = \hat{\sigma}^2_{IG}$ $(\hat{\sigma}^2_{treat} / \hat{\sigma}^2_{Fehler})$

mit den Zählerfreiheitsgraden **$df_{ZG} = n - 1$ $df_{treat}; = n - 1$**

und Nennerfreiheitsgraden **$df_{IG} = N - n$ $df_{Fehler} = N - 1$**

$F_{emp} = S^2_{treat} / S^2_{Fehler}$

$F_{tab} = (df_{treat};\ df_{Fehler,\ \alpha\ einseitig})$

2. Post hoc Test (S- N- K) zur Überprüfung der speziellen H_0`s

3. F- Test: F_{max}, Bartlett, Leven (SPSS) prüft, ob Voraussetzung der Varianzhomogenität gegeben ist

Voraussetzungen für die Varianzanalyse:

1. Population normalverteilt (Voraussetzung um Mittelwert und Varianz zu berechnen)
2. Varianzhomogenität (Varianzschätzungen „zwischen“ und „innerhalb“ der Gruppen sind voneinander unabhängig)

Varianzhomogenität:

- F- Test vorausgeschaltet, testet Teil der Hypothese H_1 und H_0: homogene Varianzen
- Prüfen, ob Varianzhomogenität, also Gleichheit der Populationsvarainzen σ^2 angenommen werden kann, wenn sich die Varianzschätzungen in den einzelnen Gruppen unterscheiden (Varianztest innerhalb der Gruppen Schätzung für den allgemeinen Fehler)

Varianzhomogenität: Unabhängigkeit des Meßfehlers von den Bedingungen

- weichen Varianzschätzungen innerhalb der einzelnen Gruppen zu sehr voneinander ab, so ist wahrscheinlich, daß die Voraussetzung der Unabhängigkeit von Fehlern und Effekten verletzt ist

- nur, wenn Varianzhomogenität erfüllt ist, können wir von unter allen Bedingung gleichen Fehlervarianz ausgehen

Was versteht man in der Varianzanalyse unter einem Faktor, was unter einer Faktorenstufe?

In der VA bezeichnet man eine unabhängige Variable als Faktor, die verschiedenen Ausprägungen werden als Faktorstufe bezeichnet. Ist nur eine UV vorhanden, so spricht man von einer einfaktoriellen VA, sind mehrere Faktoren vorhanden, so spricht man von einer mehrfaktoriellen VA.

Abhängige Variablen: quantitative intervallskallierte Variablen
Univariate Varianzanalyse: eine AV
Multivariate Varianzanalyse: mehr als eine AV

4.2 Was versteht man unter Quadratsumme, Zahl der Freiheitsgrade und Varianzschätzung? Welcher Zusammenhang besteht zwischen diesen Größen? Warum ist die Zahl der Freiheitsgrade bei Varianzschätzungen n-1? Was versteht man bei der Quadratsumme unter Korrekturglied?

Was versteht man unter Quadratsumme, Zahl der Freiheitsgrade und Varianzschätzung?

• Quadratsumme S:

- ist die Summe der quadrierten Abweichungen der Maßzahlen von ihrem Mittelwert.
- bildet den Zähler in der Formel zur Berechnung der Varianz. Über die Quadratsummen wird somit eine Varianzschätzung über die Population vorgenommen.
- Die Quadratsumme total setzt sich aus der Quadratsumme „innerhalb" und der Quadratsumme „zwischen" zusammen.

- jede Quadratsumme = Summe x^2 – Korrekturglied

Freiheitsgrade df:

$df_{ZG} = n-1$ $df_{IG} = N-n$ $df_T = N-1$ $df_T = df_{ZG} + df_{IG}$

⇨ $df_T = df_{IG} + df_{ZG}$; bewiesen wird diese Beziehung durch N-1 = N-n + n-1

- bezeichnet die Anzahl der Werte, die frei variieren können, abzüglich ihrer Bestimmungsstücke
- $df_{IG} = N-n$ für die Berechnung von S_{IG} werden n Mittelwerte benötigt
- $df_{ZG} = n-1$ da sich bei Kenntnis des einen Gruppenmittelwertes und des Gesamtmittelwertes sich der andere berechnen läßt (und damit nicht mehr frei variieren kann)

→ Freiheitsgrade werden festgelegt, indem eine der n Abweichungen von n Messungen festgelegt ist, die Varianz hat somit n-1 Freiheitsgrade. Die Anzahl der Bestimmungsstücke (wie z.B. die Populationsparameter sie darstellen), die frei variieren können, ist ebenfalls auf „n-1" begrenzt.
→ Die Bezeichnung Freiheitsgrade rührt daher, dass nach Berechnung des Mittelwertes für eine Stichprobe nur noch m-1 Maßzahlen frei gewählt werden können.

• Varianzschätzung $\hat{\sigma}^2$:

- wird von Glaser als korrekter Begriff für die Varianz (s^2) angeführt.
- Ist als s^2 die erwartungstreue Schätzung der Populationsvarianz.
- Generell ist die totale Varianzschätzung zerlegbar in die Varianzschätzung innerhalb von Gruppen (s^2_{IG}) und in die Varianzschätzung zwischen Gruppen (s^2_{ZG}).
- s^2_{ZG} ist die Varianz der Einzelmaßzahlen um den Gruppenmittelwert: stellt eine erwartungstreue Schätzung der Populationsvarianz dar, sofern man bei den einzelnen Gruppen Normalverteilung annehmen kann.
- s^2_{ZG} : jede der gegebenen Einzelstichproben ist nur mit dem Mittelwert in der Schätzung repräsentiert.

Varianzschätzung = Quadratsumme/ Freiheitsgrade
= mittlere Abweichungsquadrate
$\hat{\sigma}^2 = s^2 = S/ df$
$\hat{\sigma}^2_{ZG} = s^2_{ZG} = S_{ZG}/ df$ Varianzschätzung zwischen den Gruppen auf der Basis der einzelnen Gruppenmittelwerte (Mittelwerte, als ob sie Maßzahlen wären) (Between group)
$\hat{\sigma}^2_{IG} = s^2_{IG} = S_{IG}/ df$

gemittelte Varianzschätzung auf der Basis der Abweichungen der einzelnen Maßzahlen von ihren zugehörigen Gruppenmittelwerten: durchschnittliche Varianz der Einzelverteilung (Within group) gewogenes arithmetisches Mittel der einzelnen Varianzschätzungen innerhalb der Gruppen

$\hat{\sigma}^2_{Total} = s^2_{Total} = S_{Total}/ df$

- In der ANOVA werden Varianzkomponenten gebildet, die unter der H_0 (erwartet Gleichheit der Erwartungswerte in den Gruppen) identische Größen aufweisen müssten

Welcher Zusammenhang besteht zwischen diesen Größen?

Bei jeder Varianzberechnung wird als erstes Glied der Quadratsumme (S) die Summe der quadrierten Maßzahlen berechnet.
Die totale Quadratsumme setzt sich additiv aus der Quadratsumme innerhalb und der Quadratsumme zwischen den Gruppen zusammen. Somit ist eine Zerlegung der Gesamtvarianz in die Varianzschätzungen „innerhalb“ und „zwischen“ möglich.
Die Freiheitsgrade der totalen Varianzschätzung teilen sich ebenso additiv in die s^2_{IG} und s^2_{ZG} auf.
Glaser dazu: aus jeder Quadratsumme wird durch Division durch die Zahl der zugehörigen Freiheitsgrade eine Varianzschätzung. Die Bezeichnung „Varianzanalyse“ rührt daher.
⇨ Varianzschätzung = Quadratsumme/ Freiheitsgrade

Warum ist die Zahl der Freiheitsgrade bei Varianzschätzungen n-1?

- Werden Varianzen von Zufallsstichproben des Umfangs n aus einer Grundgesamtheit gemittelt, erhalten wir eine Durchschnittsvarianz, die die Populationsvarianz um den Faktor n-1/ n unterschätzt
- Damit Stichprobenvarianzen erwartungstreue Schätzer der Populationsvarianz werden, muß man sie mit dem Faktor n/ n-1 multiplizieren

$\hat{\sigma}^2 = \sum (x_i - M)^2/ n * n/ n-1 = \sum (x_i - M)^2/ n-1$ **n- 1 Freiheitsgrade der Varianz**

Was versteht man bei der Quadratsumme unter Korrekturglied?

Korrekturglied:
$\sum (x_i - M)^2 = \sum x_i^2 - 1/n*(\sum x_i)^2$
Quadratsumme= Summenquadrat x * Korrekturglied

Jede Quadratsumme (S) ist eine Summe x^2 – Korrekturglied
(versteh ich nicht ganz, weil bei Varianzanalyse nur bei S_T so, ansonsten kommt das totale Korrekturglied nur noch bei S_{ZG} vor)
Wenn M = 0 wird Korrekturglied 0
Totales Korrekturglied bei der Varianzanalyse: **$1/N * (\sum\sum x_{ij})$**

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m_j} x_{ij} \right)^2 / m_j$$

Gewinnung des Korrekturgliedes bei der VA:

- Summation der unquadrierten Elemente innerhalb jeder Spalte über alle Zeilen hinweg.
- Quadrierung der Spaltensummen
- Division der quadrierten Spaltensummen durch die Anzahl der Maßzahlen in den Spalten.

Quadratsummen bestehen also zumeist aus der Summe der quadrierten Maßzahlen minus das durch m dividierte Quadrat der Summe der nichtquadrierten Maßzahlen (Korrekturglied).
Bedeutung:
Es gibt den Einfluss des Stichprobenmittelwertes auf die Quadratsumme S wieder.

4.3 Was versteht man bei der einfaktoriellen Varianzanalyse unter der Quadratsumme zwischen den Gruppen, der Quadratsumme innerhalb der Gruppen und der Quadratsumme total? Wie hängen diese drei Quadratsummen miteinander zusammen? Wie groß ist die jeweils zugehörige Zahl der Freiheitsgrade? Wie hängen die Freiheitsgradezahlen der drei Quadratsummen miteinander zusammen?

Was versteht man bei der einfaktoriellen Varianzanalyse unter der Quadratsumme zwischen den Gruppen, der Quadratsumme innerhalb der Gruppen und der Quadratsumme total?

⇨ Bei der einfaktoriellen Varianzanalyse gibt es eine generelle H0; hier wird eine UV (Merkmal) anhand mehreren Stichroben (Gruppen) in mehreren Faktoren (Abstufungen des Merkmals) untersucht. i = Spalte, j = Zeile

• **Quersumme zwischen Gruppen**: Erst werden die Zeilen der Faktorgruppen summiert, dann quadriert und durch die Anzahl der jeweiligen Gruppenmitglieder dividiert, hiervon wird die quadrierte Summe aller Maßzahlen dividiert durch alle Vpn abgezogen. Bei Glaser SZG

$$SZG = \sum_{j=1}^{N} \frac{1}{mj} * \left(\sum_{i=1}^{mj} xij\right)^2 - \frac{1}{N}\left(\sum_{j=1}^{n}\sum_{i=1}^{mj} xij\right)^2$$

Zahl der Freiheitsgrade df= n-1 (n= Anzahl d. Gruppen)

$$SZG^2 = \frac{SZG}{dfZG}$$ (auf Basis der Mittelwerte)

⇨ Hier werden die Mittelwerte der Gruppen mit dem Gruppengesamtmittelwert zueinander in Beziehung gebracht. Hier wird untersucht, inwieweit die Mittelwerte vom Gesamtmittelwert abweichen.

• **Quersumme innerhalb** der Gruppen: Ausgangspunkt ist die Summe aller quadrierten Maßzahlen und hiervon wird die Differenz abgezogen, die aus der Summierung der Zeilen und des Quadrates hiervon entsteht, dividiert durch die Anzahl der jew. Gruppenmitglieder.
Bei Glaser SIG

$$SIG = \sum_{j=1}^{n}\sum_{i=1}^{mj} xij^2 - \sum_{j=1}^{n} \frac{1}{mj} * \left(\sum xij\right)^2$$

Zahl der Freiheitsgrade df= N-n

$$SIG^2 = \frac{SIG}{dfIG}$$ (auf Basis der Abweichungen der einzelnen Maßzahlen zum Mittelwert der einzelnen Gruppen)

• Bei der **totalen Quadratsumme** werden alle quadrierten Maßzahlen summiert und hiervon die Summe abgezogen, die sich aus der Quadrierung aller summierten Maßzahlen dividiert durch alle Vpn ergibt. Wird bei Glaser ST genannt.

$$ST = \sum_{j=1}^{n}\sum_{i=1}^{mj} xij^2 - \frac{1}{N}\left(\sum\sum xij\right)^2$$

Zahl der Freiheitsgrade: df= N – 1 (N – alle Vpn)

$$ST^2 = \frac{St}{dfT}$$ (totale Verteilung)

⇨ Hier wird über alle Maßzahlen hinweg gerechnet. Es wird die generelle H0 getestet, da untersucht wird, ob sich die Unterschiedlichkeit des Faktors zwischen den einzelnen Gruppen und innerhalb der jeweiligen Gruppen mehr als zufällig auswirkt.

Wie hängen diese drei Quadratsummen miteinander zusammen?

Beziehung der Quadratsummen miteinander:
ST= SIG + SZG

Wie groß ist die jeweils zugehörige Zahl der Freiheitsgrade?

Die Freiheitsgrade benötigen wir bei der Feststellung, ob die Ergebnisse der Varianzanalyse signifikant sind. Es genügen die der IG und ZG

- $df_T = N-1$
- $df_{IG} = N-n$
- $df_{ZG} = n-1$

➔ wobei N = Zahl der insgesamt vorhandenen Maßzahlen

➔ n= Zahl der Gruppen

⇨

Tabelle F (dfZG, dfIG, 5 %)= .... $Femp = \frac{SZG^2}{SIG^2}$

Wie hängen die Freiheitsgradezahlen der drei Quadratsummen miteinander zusammen?
$df_T = df_{IG} + df_{ZG}$

4.4 Was versteht man bei einer wenigstens zweifaktoriellen Varianzanalyse unter Haupteffekt, Einfacheffekt und Wechselwirkung ? Wie kann man das Resultat einer zweifaktoriellen Varianzanalyse graphisch darstellen, so dass die Wechselwirkung veranschaulicht wird?

Was versteht man bei einer wenigstens zweifaktoriellen Varianzanalyse unter Haupteffekt, Einfacheffekt und Wechselwirkung ?
➔ betrachtet an dem einfachsten Fall einer zweifaktoriellen Varianzanalyse mit nur zwei Stufen für Faktor A und Faktor B. Zum Beispiel Placebo/Beruhigungsmittel und männlich/weiblich. → AV = Subjektive Befindlichkeit

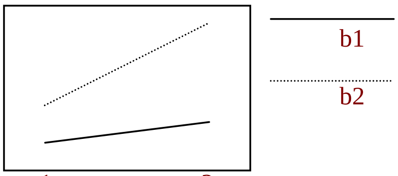
a) Als **Haupteffekt** bezeichnet man den Unterschied zwischen den Mittelwerten der Stufen einer unabhängigen Variablen.
⇨ Haupteffekthypothesen beziehen sich auf die gesamte Zielpopulation, die also aussagt, dass die durch verschieden Treatments bewirkten Unterschiede für die gesamte untersuchte Population gelten.
➔ Die Wirkung eines Faktors alleine
(➔ Wenn Interaktionen vorliegen kann das Konsequenzen für die Bedeutung vorhandener Haupteffekte haben. Bei bestimmten Interaktionen dürfen Haupteffekte nicht mehr interpretiert werden, selbst wenn sie signifikant sind)
b) Im Gegensatz dazu nennt man den Unterschied zwischen den Zellen einer Zeile oder Spalte **Einfacheffekt**. Eine zweifaktorielle Varianzanalyse hat also zwei Haupteffekte und so viele einfache Effekte, wie sie Zeilen und Spalten hat.
c) Eine **Wechselwirkung** (Interaktion) bedeutet, dass die Wirkung eines Faktors auf den einzelnen Stufen des anderen Faktors verschieden ist.
⇨ Interaktionshypothesen beziehen sich auf die differentielle Wirkung der einzelnen Treatments, d.h. auf die Treatmentwirkungen, die von der Art der untersuchten Population abhängen bzw. auf die Wirkung von Treatmentkombinationen.
➔ wenn sich zwei oder mehrere unabhängige Variablen in ihrem Einfluss auf eine abhängige Variable gegenseitig beeinflussen, wenn also die Wirkung einer unabhängigen Variable auf die abhängige Variable bei einer Veränderung in einer oder mehrerer weiteren unabhängigen Variablen wechselt. Es können drei Interaktionen unterschieden werden
a) ordinale Interaktion:
→ Graphen laufen in beiden Interaktionsdiagrammen gleichsinnig (aber nicht parallel), beide aufsteigend oder beide abfallend
b) Hybriden Interaktion:
→ Graphen laufen nur in einem der Interaktionsdiagramme gleichsinnig, im anderen nicht
c) Disordinale Interaktion
→ Graphen laufen in beiden Interaktionsdiagrammen nicht gleichsinnig
(⇨ Gleichsinnigkeit sagt nichts darüber aus, ob sie sich durchkreuzen)

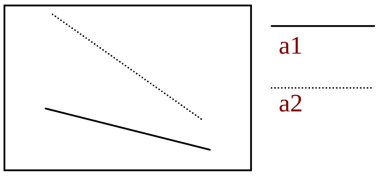
Wie kann man das Resultat einer zweifaktoriellen Varianzanalyse graphisch darstellen, so dass die Wechselwirkung veranschaulicht wird?
Graphisch kann eine Wechselwirkung dargestellt werden, indem auf der X-Achse

(= Abszisse) die Stufen eines Faktors (=UV) und auf der Y-Achse (= Ordinate) die Mittelwerte der abhängigen Variable eingetragen werden.

C) Wechselwirkung

<u>a) ordinale Interaktion:</u>

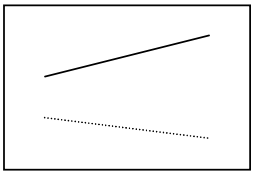
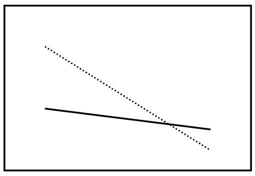




→ a1 a2 b1 b2

⇨ Beide Linien im Interaktionsdiagramm weise immer den gleichen Trend auf.
⇨ Beide Haupteffekte (von Faktor A und B) sind interpretierbar
⇨ Die Interaktion zeigt sich in unterschiedlicher Steilheit der Linien

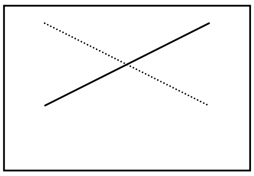
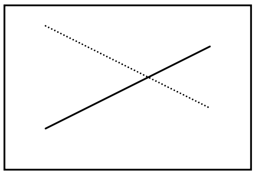
<u>b) Hybriden Interaktion:</u>

→ a1 a2 b1 b2

⇨ Bei einem Faktor gibt es einen gegenläufigen Trend, beim anderen einen gleichen Trend
⇨ In diesem Falle sollte Faktor A nicht interpretiert werden

<u>c) Disordinale Interaktion</u>

→ a1 a2 b1 b2

⇨ Bei beiden Faktoren gibt es einen gegenläufigen Trend
⇨ Beide Faktoren sollten daher nicht isoliert interpretiert werden (= kein Haupteffekt interpretiert!)

4.5 Was versteht man bei einem varianzanalytischen Versuchsplan unter Kreuzen, was unter Schachteln von Faktoren ? Welche Konsequenzen haben diese Kombinationsweisen von Faktoren für die varianzanalytische Auswertung ? Was ist ein lateinisches Quadrat?

Was versteht man bei einem varianzanalytischen Versuchsplan unter Kreuzen, was unter Schachteln von Faktoren ?

⇨ Kreuzen und Schachteln von Faktoren sind zwei prinzipielle Weisen, in der Faktoren, also UV, in der Varianzanalyse miteinander in Verbindung gebracht werden können.

• Kreuzen von Faktoren:
→ 2-faktorielle Varianzanalyse
→ jede Stufe A_i des Faktors A ist mit jeder Stufe B_j des Faktors B kombiniert (die Faktoren A und B sind gekreuzt)

• Schachteln von Faktoren:

Dies nennt man auch einen unvollständigen Versuchsplan
→ jede Stufe A_i des Faktors A kommt genau einmal vor, und zwar kombiniert mit einer bestimmten Stufe B_j des Faktors B
→ diese Untersuchung kombiniert zwei Faktoren derart, daß jede Faktorstufe des einen Faktors nur mit bestimmten Faktorstufen des anderen Faktors auftritt (Faktor A ist in Faktor B geschachtelt; deshalb auch „hierarchischer Versuchsplan")
Z.B. untersucht man die Kombinationen $a_1b_1b_2$, a_2 b_3 b_4, a_3 b_5 b_6 usw. Auf diese Weise entsteht eine Hierarchie von Faktoren wobei darauf geachtet werden muss dass jede Stufe des Faktors A mit der gleichen Anzahl von Stufen des Faktors B kombiniert wird.

Welche Konsequenzen haben diese Kombinationsweisen von Faktoren für die varianzanalytische Auswertung ?
• Kreuzen von Faktoren:
➔ Wechselwirkungen kommen nur zwischen gekreuzten, niemals zwischen geschachtelten Faktoren in der Auswertung vor.
• Schachteln von Faktoren:
→ führt zu einer Einsparung von Vpn, da nicht alle möglichen Kombinationen untersucht werden..
➔ Allerdings können mit diesem Verfahren keine Interaktionen zwischen A und B untersucht werden nur Haupteffekte sind prüfbar

⇨ Bei vielen Untersuchungen ist es organisatorisch unmöglich, die Mitglieder vorgefundener Gruppen (z.B. Schulklassen, Abteilungen etc.) individuell nach Zufall aufzuteilen. Vielmehr wird jede vorgefundene Personengruppe als ganze nach Zufall jeweils einer Versuchsbedingung zugeordnet. Daten werden auch oft für alle Personen gleichzeitig in einer gemeinsamen Sitzung (Gruppenuntersuchung) erhoben.

<u>Was ist ein lateinisches Quadrat?</u>
Ein lateinisches Quadrat ist ebenfalls eine Variante eines unvollständigen Versuchsplans.
→ kommt dann zum Einsatz, wenn man aufgrund von Voruntersuchungen zur Überzeugung gelangt, dass Interaktionen unwahrscheinlich sind, und man nur an den Haupteffekten interessiert ist.
➔ Anwendung nur dann, wenn Interaktionen zu vernachlässigen sind, denn sonst sind Haupteffekte nicht eindeutig interpretierbar
⇨ Es lassen sich dabei drei Haupteffekte untersuchen, wobei darauf geachtet werden muss, dass alle drei Faktoren die gleiche Anzahl an Faktorstufen besitzen. Das lateinische Quadrat ist so aufgebaut, dass jede Faktorstufe pro Zeile bzw. Spalte nur einmal vorkommt.

	B_1	B_2	B_3
A_1	$\mathbf{C_1}$	C_2	C_3
A_2	C_2	C_3	C_1
A_3	C_3	C_1	C_2

⇨ Jede Stufe des Faktors C ist mit jeder Stufe des Faktors A und mit jeder Stufe des Faktors B ein Mal kombiniert d.h. das lateinische Quadrat ist in Bezug auf die Haupteffekte vollständig ausbalanciert.

4.6 Was ist bei einer Varianzanalyse eine Kovariate, was ist eine Kovarianzanalyse?

Üblicherweise haben die unabhängigen Variablen bei der Varianzanalyse nur Nominalskalenniveau. Die Ausprägungen dieser UV bilden die Faktorstufen.
Hat man aber in einer Untersuchung eine intervallskalierte Variable, von der ausgegangen wird, dass sie ebenfalls Einfluss auf die abhängige Variable hat, mit erhoben, so kann dieser Einfluss mit der **Kovarianzanalyse** berücksichtigt werden (ohne die Gesamtzahl der Vpn wie in mehrfaktoriellen Varianzanalysen erhöhen zu müssen).
➔ Eine solche intervallskalierte unabhängige Variable wird als **Kovariate**
(= Kontrollvariable) bezeichnet.
Im Therapiebeispiel könnte z.B. das Alter der Patienten eine solche Kovariate sein.

⇨ Die Idee bei der Kovarianzanalyse ist, den Einfluss dieser Kovariate mittels Regressionsrechnung aus der abhängigen Variable herauszurechnen. Dadurch wird die Fehlervarianz in den Daten gesenkt und ein signifikanter Einfluss der anderen Faktoren leichter nachweisbar.

4.7 Was versteht man bei einer Varianzanalyse unter Messwiederholungen? Welche Vorteile, welche Nachteile haben Messwiederholungen? Was ist ein „between-subjects", was ein „within-subjects" Faktor?

Was versteht man bei einer Varianzanalyse unter Messwiederholungen?
Die Vpn der einzelnen Stichproben werden mehrfach unter verschiedenen Bedingungen (Faktorstufen) beobachtet.

Welche Vorteile, welche Nachteile haben Messwiederholungen?
Vorteile: man benötigt weniger Vpn
Nachteile: Möglichkeit, dass die Vpn durch wiederholte Untersuchungen zu sehr beansprucht werden, was zu Motivations- und Aufmerksamkeitsabnahme bzw. allgemein zu Sequenzeffekten führen kann, wodurch die Interpretation einer Untersuchung erschwert wird.

Was ist ein „between-subjects", was ein „within-subjects" Faktor?
• **„between - subjects" Faktor:** einzelne Stufen eines Faktors werden an verschiedenen Personen getestet, d.h. ohne Messwiederholung
• **„within - subjects" Faktor:** einzelne Stufen eines Faktors werden an der gleichen Person getestet,, d.h. mit Messwiederholung

4.8) Was ist eine multivariate Varianzanalyse?

Mit multivariaten Methoden werden Hypothesen geprüft, die sich auf das Zusammenwirken vieler abhängiger und unabhängiger Variablen beziehen.

Beispiel: Geschulte Psychiater werden gebeten, zufällig ausgewählte Patienten der 4 Kategorien (Depressiv, schizophren, paranoid und manisch) auf einer Ratingskala danach einzustufen, wie stark 16 bestimmte Merkmale ausgeprägt sind. Das komplexe Merkmal „Krankheitssymptomatik" wird jedoch nicht nur durch ein, sondern durch 16 Merkmale erfasst, d.h. wir müssten 16 einfaktorielle Varianzanalysen durchführen, um die 4 Patientengruppen hinsichtlich der gesamten Krankheitssymptomatik differenzieren zu können. ← Hier würde man nun eine multivariate Varianzanalyse durchführen.

4.9 Was versteht man unter dem Datenmodell einer Varianzanalyse, was unter festen, was unter zufälligen Faktoren? Woher weiß man bei den Faktoren einer Varianzanalyse, ob sie "fest" oder "zufällig" sind? Welche Auswirkungen hat das auf die Auswertungen?

Was versteht man unter dem Datenmodell einer Varianzanalyse,
• **Datenmodell einer einfaktoriellen Varianzanalyse**
- Datenwahl:

$xij = \mu + \beta j + \varepsilon ij$ (Population) → GL (61)

μ = Grundmittelwert
βj = Moment der j-ten Gruppe
εij = individuelles Moment (Abweichung der i-ten Person vom Mittelwert):

- Momente:

$\varepsilon ij = xij - \mu j$ → Gl. (62)
$bj == \mu j - \mu$ → Gl. (63)

oder

- Datenwahl

  xij = M**..** + bj + eij (Stichprobe) → Gl. (64)

- Momente:

  eij = xij – M**.**j → Gl. (65)

  bj = M**.**j –M**..** → Gl. (66)

$$M.. = \sum_{j=1}^{n}\sum_{i=1}^{mj} x_{ij} \times \frac{1}{N}$$

<u>was unter festen, was unter zufälligen Faktoren?</u>

(Allgemein: Faktor entspricht der Stufung einer unabhängigen Variable)

a) "Zufällige" Faktoren/ "randomized factors"

Faktoren, deren Stufen aus der Population möglicher Faktorenstufen **zufällig** ausgewählt werden, bezeichnet man als Faktoren mit zufälligen Effekten („random factors“). Die konkrete Auswahl der Faktorenstufen ist im Grunde also beliebig (z.B. Untersuchung zur Überprüfung des Einflusses des Therapeuten auf den Therapieerfolg: hier geht es nicht darum, z.B. Unterschiede zwischen bestimmten Therapeuten festzustellen, sondern um die Frage, ob Therapeuten überhaupt die abhängige Variable beeinflussen. Die Auswahl der Therapeuten ist daher beliebig). Will man nun die Ergebnisse auf die Population aller Therapeuten generalisieren, wird man als Stufen des Faktors „Therapeuten“ eine Stichprobe zufällig ausgewählter Therapeuten einsetzen (Ausmaß der Generalisierbarkeit hängt natürlich von der Repräsentativität und Größe der Stichprobe ab.

b) "Feste" Faktoren/ "fixed factors"

Wählt man jedoch systematisch nur diejenigen Faktorenstufen aus, über die man letztlich Aussagen formulieren will, sprechen wir von einem Faktor mit **festen** Effekten („fixed factors“). Dies gilt auch für Faktoren, die alle möglichen Abstufungen einer unabhängigen Variablen umfassen (z.B. männlich-weiblich, Unterschicht-Mittelschicht-Oberschicht).

<u>Woher weiß man bei den Faktoren einer Varianzanalyse, ob sie "fest" oder "zufällig" sind?</u>

Müsste sich im Zusammenhang zeigen, sprich aus dem Kontext ersichtlich werden, oder auch nicht, und dann...? Wahrscheinlich lieber von einem festen Faktor ausgehen, da enger gefasst.

<u>Welche Auswirkungen hat das auf die Auswertungen?</u>

Für die rechnerische Durchführung einer einfaktoriellen Varianzanalyse ist es unerheblich, ob der untersuchte Faktor zufällig oder fest ist. Unterschiede ergeben sich lediglich in der Interpretation.

(zu a) Bei **zufälligen Faktoren**, besagt ein signifikanter F-Test, dass die Wirkung aller möglichen Faktorenstufen nicht gleich sind.

→ Die Interpretation eines signifikanten zufälligen Faktors ist damit weitergehend als die Interpretation eines signifikanten festen Faktors. Vorrausetzung ist aber, dass bei den zufälligen Faktoren, die mit allen möglichen Faktorenstufen verbundenen Treatmenteffekte normalverteilt sind.

(zu b) Die einfaktorielle Varianzanalyse über einen **festen Faktor** überprüft die $H0: \mu1 = \mu2 = = \mu p$, d.h. bei einem signifikanten F-Test können wir behaupten, dass sich mindestens zwei der tatsächlich untersuchten Faktorenstufen unterscheiden.

4.10 Bei einer varianzanalytischen Auswertung kann man auch spezielle Nullhypothesen bezüglich einzelner Mittelwertsunterschiede aufstellen und prüfen. Welche Probleme stellen sich dabei? Was versteht man unter apriorischen und aposteriorischen, was unter orthogonalen und nicht-orthogonalen Tests? Was ist die Bonferoni-Korrektur? Bitte erläutern Sie den Newman-Keuls-Test.

<u>Bei einer varianzanalytischen Auswertung kann man auch spezielle Nullhypothesen bezüglich einzelner Mittelwertsunterschiede aufstellen und prüfen. Welche Probleme stellen sich dabei?</u>

Das Problem lässt sich daraus ableiten, daß zunächst der t-Test das angemessene Instrument zur Prüfung eines Unterschiedes zwischen zwei Mittelwerten ist.

Im vorliegenden Falle führt es aber in die **Problematik** der **Alpha-Erniedrigung** durch wiederholte Anwendung von t-Tests am gleichen Datensatz oder **an systematisch ausgewählten Teilmengen** eines Datensatzes.

• **Problematik** der **Alpha-Erniedrigung:**

Die Gültigkeit eines einzigen speziellen H_1 verwirft die generelle H_0. Dies erzeugt in der praktischen Anwendung ein Problem, da sich der **∇-Fehler kumuliert** (siehe Multiplikationstheorem für voneinander unabhängige Ereignisse). Das Risiko, die varianzanalytische H0 über einen Einzelvergleichstest zu verwerfen, ist also gegenüber dem ursprünglichen ∇-Niveau um ein Vielfaches höher.

Deshalb muss das **∇-Niveau korrigiert** werden: Soll eine globale Nullhypothese über verschiedene Einzeltests auf einem zuvor spezifizierten ∇-Niveau verworfen werden, muss mindestens ein Einzeltest die folgende Irrtumswahrscheinlichkeit ∇‘ erreichen oder unterschreiten: **∇‘ = $\mathbf{1-(1-\nabla)^{1/m}}$** (m = n-1 bei n Gruppen). Bei wachsendem n lässt sich ∇‘ durch die **Bonferoni-Korrektur** approximieren: **∇‘ = ∇ / n**

• **Problematik systematisch ausgewählten Teilmengen:**

??

Einzelvergleiche werden oft auch Kontraste genannt. Es gibt ***3 grundsätzliche Probleme:***

<u>1.) Wie viele Einzelvergleiche gibt es bei n Gruppen?</u>

a) $C = \binom{n}{2} = n*(n-1) / 2$

b) nur die orthogonalen Einzelvergleiche: n-1

c) eigentlich gibt es unendlich viele, wenn man die gebrochenen Einzelvergleiche mit einbezieht: $H_{0\ spez.}$ $⇐_1 = (⇐_2+⇐_3) / 2$; $⇐_1 = (⇐_2+⇐_3+⇐_4) / 2$ -> die Brüche werden immer kleiner, es gibt immer mehr Einzelvergleiche.

[Unter ***orthogonalen Tests*** versteht man Tests, die keine gemeinsamen Informationen aus den Daten ausnutzen, es gibt also keine Mehrfachverwendung der Daten -> Korrelation ***r = 0***; ***nicht-orthogonales*** Tests sind also ***abhängige Tests***]

<u>2.) Wann sollen die Einzelvergleiche geplant werden?</u>

a) ***a priori:*** Interesse an einem Einzelvergleich ***vor der Untersuchung*** bekannt, dies muss sachlich (aus der Literatur) begründet sein (Erkenntnisgewinn höher)

b) ***a posteriori:*** aufgestellt, nachdem die Ergebnisse bekannt sind, erst dann nach einer plausiblen Erklärung gesucht -> ***kein wissenschaftliches Hypothesenprüfen*** -> durch ***Bonferoni-Korrektur*** (siehe später) muss der ∇-Fehler korrigiert werden

<u>3.) Welchen Zusammenhang gibt es zwischen der spezifischen und der generellen H_0?</u>

Die generelle H_0 impliziert, dass auch alle speziellen H_0 gelten, sie verbietet also auch alle spezifischen Alternativhypothesen. Die Gültigkeit eines einzigen speziellen H_1 verwirft die generelle H_0. Dies erzeugt in der praktischen Anwendung ein Problem, da sich der ***∇-Fehler kumuliert*** (siehe Multiplikationstheorem für voneinander unabhängige Ereignisse). Das Risiko, die varianzanalytische H0 über einen Einzelvergleichstest zu verwerfen, ist also gegenüber dem ursprünglichen ∇-Niveau um ein Vielfaches höher.

Deshalb muss das ***∇-Niveau korrigiert*** werden: Soll eine globale Nullhypothese über verschiedene Einzeltests auf einem zuvor spezifizierten ∇-Niveau verworfen werden, muss mindestens ein Einzeltest die folgende Irrtumswahrscheinlichkeit ∇‘ erreichen oder unterschreiten: ***∇‘ = $1-(1-\nabla)^{1/m}$*** (m = n-1 bei n Gruppen). Bei wachsendem n lässt sich ∇‘ durch die ***Bonferoni-Korrektur*** approximieren: ***∇‘ = ∇ / n***

<u>Was versteht man unter apriorischen und aposteriorischen, was unter orthogonalen und nicht-orthogonalen Tests?</u>

Apriorische und aposteriorische Tests bzw. Hypothesen

A) a priori: Interesse an einem Einzelvergleich **vor der Untersuchung** bekannt, dies muss sachlich (aus der Literatur) begründet sein (Erkenntnisgewinn höher)

Ergänzung: Auffassung im Kontext der Varianzanalyse bzw. mit deren α-Fehler-Korrektur: A priori formulierte Einzelvergleichshypothesen, die theoretisch gut begründet sind, oder auf Grund von Vorversuchen aufgestellt werden konnten, machen keine α-Fehler-Korrektur erforderlich.

⇨ In der Regel sind es nur eine oder zwei Einzelvergleichshypothesen, die man im Rahmen einer einfaktoriellen varianzanalytischen Untersuchung den Status einer "echten" A-priori-Hypothese zubilligen kann, und die deshalb- jede für sich- mit dem unkorrigierten, nominellen α-Niveau getestet werden können.

Typischerweise sind A-priori-Hypothesen gerichtet, so daß statt des F-Tests nach Gl. (7.44) oder (7.51) ein einseitiger t-Test nach Gl. (7.45) gerechtfertigt ist. Der t-Wert lässt sich einfach als Wurzel des F-Werts ermitteln (s. Gl. 2.60).

Man geht davon aus, dass **a priori** formulierte Einzelvergleichshypothesen, die theoretisch gut begründet sind, **keine ∇-Fehler-Korrektur** benötigen. Im Normalfall handelt es sich um sehr wenige, gerichtete Hypothesen, die einen **einseitigen t-Test** rechtfertigen.

B) a posteriori: aufgestellt, nachdem die Ergebnisse bekannt sind, erst dann nach einer plausiblen Erklärung gesucht -> **kein wissenschaftliches Hypothesenprüfen** -> durch **Bonferoni-Korrektur** (siehe später) muss der ∇-Fehler korrigiert werden (Fragestellung ergibt sich hinterher)

⇨ A-posteriori- Einzelvergleiche hingegen können jederzeit durchgeführt werden, wenn man nach einer „Overall-„ Signifikanz feststellen möchte, welche Einzelvergleiche maßgeblich dafür verantwortlich sind, daß die globale Ho zu verwerfen ist. In diesem Falle muß der α-Fehler gemäß Gl. (7.62) oder (7.63) korrigiert werden.

Wenn ein Einzelvergleich mit dem korrigierten α'-Niveau signifikant wird, so ist dies zwar ein wichtiger Hinweis für die Interpretation der Overall-Signifikanz, aber noch keine Bestätigung der entsprechenden Einzelvergleichshypothese. Diese kann nur in einer neuen Untersuchung erbracht werden, der diese Einzelvergleichshypothese als A-priori-Hypothese vorangestellt wird.

Orthogonale und nicht-orthogonale Tests

C) Unter **orthogonalen Tests** versteht man Tests, die keine gemeinsamen Informationen aus den Daten ausnutzen, es gibt also keine Mehrfachverwendung der Daten -> Korrelation **r = 0**;

D) **nicht-orthogonales** Tests sind also **abhängige Tests**

⇨ In Abhebung von Varianzanalysen mit gleichgroßen Stichproben und damit unkorrelierten (orthogonalen) Effekten bezeichnet man zwei- oder mehrfaktorielle Varianzanalysen mit ungleichgroßen Stichproben als nicht-orthogonale Varianzanalysen

<u>Was ist die Bonferoni-Korrektur?</u>

⇨ Bonferoni-Korrektur ist eine α-Fehler-Korrektur

Wird in einer einfaktoriellen Varianzanalyse die $H_o : \mu = \mu = ... = \mu = ... = \mu$ zugunsten der $H : \mu \neq \mu$ mit $\alpha = 0{,}05$ verworfen, beträgt die Wahrscheinlichkeit einer irrtümlichen Entscheidung 5%. Die Vorgehensweise, die H der einfaktoriellen Varianzanalyse über p-1 orthogonale Einzelvergleiche mit $\alpha = 0{,}05$ zu prüfen ist dennoch nicht korrekt. Es muß eine α-Fehler-Korrektur erfolgen.

Soll eine globale Nullhypothese über m verschiedene Einzeltests auf einem zuvor spezifizierten α-Niveau verworfen werden, muß mindestens ein Einzeltest die Irrtumswahrscheinlichkeit α' erreichen oder unterschreiten:

$\alpha' = 1- (1-\alpha)^{1/m}$ → GL (7.62).

Der nach Gl. (7.62) ermittelte α'-Wert lässt sich mit wachsendem m durch eine sehr viel einfachere Gleichung approximieren, die in der Literatur als Bonferoni-Korrektur bekannt ist:

→ Gl. (7.63): $\alpha' = \alpha/m$

<u>SPSS-Skript S.65:</u>

Im Bonferoni-Test werden alle Mittelwertsdifferenzen berechnet und auf Signifikanz geprüft. Die Basis ist der t-Test mit der sog. Bonferoni-Korrektur für die Tatsache, daß die einzelnen Mittelwertsvergleiche zu einer Grundmenge von Mittelwerten gehören und daher nicht voneinander unabhängig sind (er ist wie der S-N-K auch ein Post-Hoc-Test für Einzelvergleiche).

<u>Bitte erläutern Sie den Newman-Keuls-Test (SNK).</u>

= Post-Hoc-Test ein wichtiger Signifikanztest aller geordneten (der Größe nach) Einzelmittelwerte (nicht-orthogonalen a-posteriori-Hypothesen).
Er prüft alle kombinatorisch möglichen ganzen (also nicht gebrochenen) Mittelwertsdifferenzen auf Signifikanz und bildet dabei Teilmengen voneinander nicht verschiedenen Mittelwerten.
Voraussetzung ist die gleich Stichprobengröße. Wird eine kritische Differenz zweier Mittelwerte überschritten, so ist die Differenz signifikant. Diese Prüfgröße berechnet sich wie folgt: D = q* sqr (s2IG / m) (sqr = Wurzel aus; m = Stichprobengröße; q ist der Tabelle „studentized range statistics“ zu entnehmen, es hängt ab vom Signifikanzniveau, der Zahl der Freiheitsgrade und der Distanz der zu vergleichenden Mittelwerte). Man beginnt den SNK mit der Prüfung der größten Mittelwerts-Differenz und bearbeitet dann jeweils alle Differenzen bis zur ersten nicht-signifikanten Differenz.
(Vergleiche tabellarische Mittelwertsunterschiede in SPSS: zeilenweise durcharbeiten)