Erwin Grüner
FB Psychologie

Versuchsplanung

WS 2006/07
29.01.2007

Veranstaltungsbeschreibung

Inhalte der Lehrveranstaltung

Inhalte der Lehrveranstaltung

Der Vorlesungsstoff gliedert sich in drei Teile:

- **Wissenschaftstheoretische Grundlagen**: Prinzipien der empirischen und experimentellen Forschung, Logik des Experiments, Konstruktion von Versuchsplänen
- **Stichprobentechniken**: uneingeschränkte Zufallsauswahl, geschichtete Stichproben, u.a.
- **Varianzanalytische Modelle und Auswertungsverfahren**: ein- und mehrfaktorielle Versuchspläne mit unabhängigen Gruppen, Versuchspläne mit abhängigen Gruppen, hierarchische Versuchspläne, lateinisches Quadrat, Kovarianzanalyse, allgemeines lineares Modell

Lernziel

Lernziel

- Die Teilnehmer sollen in die Lage versetzt werden, selbständig Versuche zu planen, anzuwenden und die Daten selbständig auszuwerten.
- Die den Verfahren zugrunde liegenden theoretischen Grundlagen sollen zumindest auf einem einfachen, intuitiven Niveau verstanden werden.
- Die Teilnehmer sollen anschließend auch in der Lage sein, irgendwelche Behauptungen und Aussagen, z.B. empirische Befunde, kritisch zu bewerten.

Ablauf

Ablauf

- Vorlesung (mit der Möglichkeit, Fragen zu stellen)
  - Vermittlung der neuen Inhalte

  - Übersicht und Strukturierung
  - Anwendungsbeispiele
  - Demonstrationen am PC
- Parallel zur Vorlesung wird eine Übung angeboten.

Literaturhinweise

Literaturhinweise

Wissenschaftstheoretische Grundlagen, Prinzipien der Versuchsplanung

- Huber, O. (2002). *Das psychologische Experiment* (3. Aufl.). Bern: Huber.
- Bortz, J., & Döring, N. (2002). *Forschungsmethoden und Evaluation* (3. Aufl.). Berlin: Springer.

Speziell: Versuchspläne

- Campbell, D.T., & Stanley, J.C. (1963). Experimental and Quasi-Experimental Designs for Research on Teaching. In: Gage, N.L. (Ed.). *Handbook of Research on Teaching* (S. 171-246). Chicago: Rand McNally.

Stichprobentechniken

- Kellerer, H. (1997). *Theorie und Technik des Stichprobenverfahrens.* Göttingen: Vandenhoek & Ruprecht.

Varianzanalytische Methoden

- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5. Aufl.). Berlin: Springer.
- Kirk, R.E. (1982). *Experimental Design: Procedures for the Behavioral Sciences* (2nd ed.). Monterey: Brooks/Cole Pub. Co.

Veranstaltungsskript

I. Stelzl / E. Grüner (2006). *Versuchsplanung WS 2006/07.*(Erhältlich bei Frau Skitschak, Raum 176)

Organisatorische Hinweise

Organisatorische Hinweise

- Übung (Termine, Beginn, Teilnehmerzahl, ...)
- Elektronischer Semesterapparat (Zugangsdaten[1], Zweck, ...)
- Klausur (Termin, Schein, ...)
- ...

Wissenschaftstheoretische Grundlagen

Wozu Versuchsplanung?

Wozu Versuchsplanung?

- Versuchplanung ist eine Forschungsmethode, die in der Psychologie und in anderen Wissenschaften eingesetzt wird.
- Ihr Ziel ist es, auf empirischem Wege zu wissenschaftlichen Aussagen zu kommen bzw. wissenschaftliche Aussagen zu überprüfen.
- Sie befasst sich im einzelnen mit drei Problembereichen:
  1. Fragen des Versuchsdesigns
  2. Erhebung von Stichproben
  3. Auswertung der gewonnenen Daten

Psychologie als Wissenschaft

Alltagspsychologie vs. wissenschaftliche Psychologie

Wissenschaftliche Aussagen müssen sein:

- logisch widerspruchsfrei
- nachvollziehbar und intersubjektiv pürüfbar
- werturteilsfrei

Alltagspsychologie

[1] `https://lernen.uni-marburg.de`, Username/Passwort

- Die Alltagspsychologie beruht auf Intuition, Spekulation, naiven Analogieschlüssen.
- Die Kriterien für wissenschaftliche Aussagen sind nicht oder nur teilweise erfüllt.

Wissenschaftliche Psychologie

- Die Psychologie ist eine Erfahrungswissenschaft.
- Sie gründet ihre Aussagen auf Beobachtbares, selbst dann, wenn sich ihre Theorien auf innere, der direkten Beobachtung nicht zugängliche Prozesse beziehen.
- Die wissenschaftliche Psychologie verwendet ein Methodeninstrumentarium.

Methodeninstrumentarium der Psychologie

- Operationalisierung psychologischer Konstrukte, indem man sie erfahrbar, bzw. sogar messbar macht.
- Operationalisierung von Untersuchungsstrategien
- Induktion: aus Daten werden Hypothesen und Theorien abgeleitet
- Deduktion: aus theoretischen Annahmen und Modellen werden Vorhersagen für Daten abgeleitet

---

Beispiel: Aggressions-Frustrations-Hypothese

- Aus der Theorie werden konkrete Verhaltensvorhersagen abgeleitet ...
- Operationalisierung der betrachteten Variablen ...
- Versuchsanordnung ...
- Gewinnung und Auswertung der Daten

Wissenschaftliche Hypothesen

Wissenschaftliche Hypothesen

- Im Erkundungsstadium werden wissenschaftliche Aussagen als Hypothesen formuliert.
- “Eine *wissenschaftliche Hypothese* behauptet eine mehr oder weniger präzise Beziehung zwischen zwei oder mehreren Variablen, die für eine bestimmte Population vergleichbarer Objekte oder Ereignisse gelten soll.” (Bortz & Döring, 2002)

- Wissenschaftliche Hypothesen können im allgemeinen als "Wenn-Dann"-Sätze formuliert werden.

Überprüfung wissenschaftlicher Hypothesen

- Wissenschaftliche Hypothesen oder Theorien müssen empirisch überprüfbar sein.
- Sie können durch Erfahrungsdaten überprüft, d.h. verifiziert (gestützt) oder falsifiziert (widerlegt) werden.
- In der Praxis ist der Nachweis oder die Widerlegung der Gültigkeit einer Hypothese oder Theorie nie mit letzter Sicherheit zu führen; es sind lediglich Wahrscheinlichkeits- und Plausibilitätsaussagen möglich.

Klassifizierung von Variablen

Klassifizierung von Variablen

- Die Variable, die im Wenn-Satz einer Hypothese steht, nennt man die **unabhängige** Variable (UV).
- Die Variable, die im "Dann"-Satz steht, nennt man die **abhängige** Variable (AV).
- Eine Variable, welche die Wirkung der UV modifiziert, nennt man **Moderatorvariable**.
- Neben der UV hat man es in der Praxis immer auch mit **Störvariablen** zu tun.
- In manchen Theorien tauchen neben den unabhängigen und den abhängigen Variablen noch **intervenierende** Variablen auf.

Unabhängige vs. abhängige Variablen

- Die Bezeichnung "unabhängig" und "abhängig" sind keine Prädikate, die den Variablen selbst zukommen, sie hängen vielmehr von der jeweiligen Hypothese ab.
- Mit diesen Bezeichnungen wird auch nicht unbedingt eine Kausalitätsbeziehung ausgedrückt.
- Auch was eine Störvariable ist, hängt von dem jeweiligen Kontext ab: ein und dieselbe Variable kann in einem Experiment UV und in einem anderen eine Störvariable sein.
- Übrigens können die UV und AV auch multivariat sein.

Klassifizierung von unabhängigen Variablen

- **Experimentelle Variablen (Treatmentvariablen)**
  Beispiele: Dosierung eines Präparats, Lehrmethode, Art der Rückmeldung
- **Präexperimentelle Variablen (Organismusvariablen)**
  Beispiele: Geschlecht, Alter, Intelligenz, Änstlichkeit, Einstellungen, Schulbildung, Schicht

---

Intervenierende Variablen

- Intervenierende Variablen spielen in manchen Theorien die Rolle von hypothetischen Konstrukten.
- Intervenierende Variablen sind selbst nicht beobachtbar.
- Sie vermitteln in einer Theorie zwischen den unabhängigen und den abhängigen Variablen, sie erklären in der Theorie den Zusammenhang zwischen UV und AV.
- Beispiele für intervenierende Variablen: Persönlichkeitskonstrukte oder z.B. der Trieb ("drive") in der Motivationstheorie von Hull.

Korrelation und Kausalität

Kausalität und Korrelation

Kausalität

- Viele wissenschaftlichen Aussagen haben die Form von "Wenn-Dann"Aussagen.
- Die Aussage "Wenn A, dann B" impliziert nicht notwendigerweise eine Kausalitätsbeziehung zwischen A und B.
- Bedingungen für eine Ursache-Wirkungs-Beziehung zwischen A und B:
  - Räumliche Kontiguität von A und B
  - Zeitliche Kontiguität von A und B (genauer: A **vor** B)
  - Ist A eine notwendige Bedingung für B?
    (Wenn nicht-A, dann immer nicht-B?)
  - Ist A eine hinreichende Bedingung für B?
    (Wenn A, dann immer B?)
- Übrigens: Kausalität kann auch probabilistisch sein.

Korrelation

- Mit Korrelation bezeichnet man lediglich man den Zusammenhang (genauer: den linearen Zusammenhang) zwischen zwei Variablen.
- Damit ist noch keine Aussage über eine Kausalbeziehung gemacht.
- Eine Korrelation zwischen A und B kann durch folgende **kausalen** Beziehungen zustande kommen:
  - A $\rightarrow$ B (A bewirkt B)
  - A $\leftarrow$ B (B bewirkt A)
  - A $\leftrightarrow$ B (wechselseitige Wirkung)
  - A $\leftarrow$ C $\rightarrow$ B (eine oder mehrere dritte Variablen C bewirken A und B, eventuell unterschiedlich stark)
  - Neben diesen Formen der Beeinflussung treten oft auch Mischformen auf (z.B. aus den letzten beiden).

Interpretation von Korrelationen

- Abnahme der Geburten und der Anzahl der Störche in Ostpreußen
- Anzahl der männlichen und der weiblichen Studentierenden einer Universität
- Anzahl der Fehler und Punktezahl bei einem Test
- Geschlecht und räumliches Vorstellungsvermögen
- Körpergröße und Beinlänge
- Intelligenz und Deutschnote

Probleme mit Korrelationen

- Abhängigkeit von der Homogenität der Population/Stichprobe:
  1. hohe Homogenität[2] $\rightarrow$ Korrelationen verschwinden
  2. niedrige Homogenität[3] $\rightarrow$ Korrelationen werden dem Betrag nach größer
- Ein Korrelationskoeffizient ist ein Maß für einen <u>linearen</u> Zusammenhang, nicht für einen <u>funktionalen</u> Zusammenhang
  Beispiel: $y = x^2$ ist ein perfekter funktionaler Zusammenhang, aber die Korrelation zwischen x und y ist nicht 1.

[2] z.B. durch Streichung der beiden Extremgruppen
[3] z.B. durch Auswahl der beiden Extremgruppen

Arten von Studien

Arten von Studien

Man unterscheidet

- Experimentelle Studien
- Quasiexperimentelle Studien
- Feldexperimente
- Korrelationsstudien

Es gibt Fragestellungen, für die eine experimentelle Untersuchung nicht durchführbar ist ...

Experiment

- Ein wesentliches Merkmal des Experiments ist das willkürliche Herstellen und Manipulieren von Versuchsbedingungen.
- Mit Hilfe von Experimenten können Kausalaussagen abgeleitet bzw. überprüft werden.
- Bei präexperimentelle Variablen sind Kausalaussagen meist problematisch.

Quasi-Experiment

- Das willkürliche Herstellen und Manipulieren von Versuchsbedingungen ist noch möglich.
- Im Unterschied zum Experiment sind aber nicht alle relevanten Störvariablen kontrolliert, da das Randomisierungsprinzip nicht anwendbar ist.
- Quasiexperimentelle Versuchspläne sollte man verwenden, wenn keine besseren einsetzbar sind.
- Vorteil von Quasi-Experimenten: höhere ökologische Validität

---

Feldexperiment

- Ein Feldexperiment findet nicht im Labor, sondern in der natürlichen Umgebung statt.
- Die Versuchspersonen wissen i.d.R. nicht, dass sie an einem Experiment teilnehmen, sie verhalten sich daher vollkommen natürlich.

- Vorteil eines Feldexperiment: höhere ökologische Validität.
- Nachteil: ungenügende Kontrolle von Störvariablen, daher interne Validität gering.

Korrelationsstudie

- Bei der Korrelationsstudie werden Variablen nur beobachtet und registriert.
- Das willkürliche Herstellen und Manipulieren der Versuchsbedingungen ist nicht möglich.

Wissenschaftliche Fragestellungen

Wissenschaftliche Fragestellungen

Experimentelle oder Korrelationsstudien?

- Sprachentwicklung und sozioökonomische Schicht
- Fernsehen schädlich für Kinder?
- Alkoholgenuß und Reaktionszeit
- Anzahl der Belohnungen und Leistungsmotivation
- Depression als Reaktion auf Umweltfaktoren
- Selbstwertgefühl bei Gesunden und bei Schizophrenen

Interpretation wissenschaftlicher Befunde

(„Beispiele zum Nachdenken“[4])

- Vorlieben für Bier
- Sicherheitsgurte
- Selbstmordrate in Schweden
- Hausgeburten
- Lungentuberkulose
- Lebenserwartung von berühmten Dirigenten
- Rauchen und Lebenserwartung
- Über die Vorteile des Gassiführen von Hunden

[4] siehe Handapparat

- Über den Nutzen von Läusen

Weitere Beispiele

- Probleme mit bedingten Wahrscheinlichkeiten, z.B.:
  - **Version 1**: “Ich habe zwei Kinder. Eines davon ist ein Mädchen. Wie groß ist die Wahrscheinlichkeit, daß das andere ebenfalls ein Mädchen ist?”
  - **Version 2**: “Ich habe zwei Kinder. Das ältere Kind ist ein Mädchen. Wie groß ist die Wahrscheinlichkeit, daß das andere ebenfalls ein Mädchen ist?”
- Ziegenproblem
- Simpsons Paradox
- ...

Literaturhinweise

- Beck-Bornholdt, H.-P., & Dubben, H.-H. (2001). *Der Hund, der Eier legt.* Rowohlt.
- Krämer, W. (2000). *Denkste! (3. Aufl.).* Piper.
- Dubben, H.-H., & Beck-Bornholdt, H.-P. (2005). *Mit an Wahrscheinlichkeit grenzender Sicherheit.* rororo.

Prinzipien der Versuchsplanung

Das Max-Kon-Min Prinzip

Das Max-Kon-Min Prinzip

Es interessiert die **Varianz der AV**.

Diese ist die Summe von drei Varianzkomponenten:

- **Primärvarianz**: Varianz aufgrund der Wirkung der UVs auf die AV
- **Sekundärvarianz**: systematisch verzerrender Einfluss (“bias”) von Störvariablen
- **Fehlervarianz**: unsystematischer, zufälliger Einfluss von nicht identifizierbaren Störvariablen

Prinzip und Ziel der Versuchsplanung

Max-Kon-Min-Regel:

- Maximierung der Primärvarianz
- Kontrolle der Sekundärvarianz
- Minimierung der Fehlervarianz

Maximierung der Primärvarianz

- Variablen geeignet operationalisieren
- Varianz der UV maximieren, Wertebereich ausschöpfen (z.B. bei linearer Beziehung zwischen UV und AV auch Extrembereiche verwenden)
- Optimale Werte der UV wählen (vor allem bei nichtlinearer Beziehung zwischen UV und AV)
- Möglichst viele Stufen der UV wählen (falls ein funktionaler Zusammenhang zwischen UV und AV entdeckt werden soll)

Minimierung der Fehlervarianz

- Wenn große individuelle Differenzen vorhanden sind, einen großen Gruppenumfang verwenden (siehe Standardfehler des Mittelwerts)
- Wenn Messungen ungenau, Reliabilität der Messungen erhöhen (siehe Testtheorie)
- Wenn Auswertungsfehler vorhanden, durch Standardisierung etc. Auswertungsobjektivität erhöhen

Kontrolle der Sekundärvarianz

- Es gibt verschiedene Techniken, die Sekundärvarianz zu kontrollieren.
- Es ist nicht möglich, **alle** denkbaren störenden Einflussgrößen zu kontrollieren.

Kontrolltechniken

Kontrolltechniken

Die wichtigsten Techniken zur Kontrolle der Sekundärvarianz sind:

- Störvariablen eliminieren
- Störvariablen konstanthalten

- Gruppen randomisieren
- Kontrollgruppen verwenden
- Gruppen parallelisieren
- Verwendung von abhängigen Gruppen durch Messwiederholung
- Bedingungen ausbalancieren
- Störvariable als zusätzliche explizite UV einführen
- Störvariable als Kovariable einführen

Eliminieren

Eliminieren

1. Beispiel: schalltoter Raum, falls störender Lärm von der Straße kommt

Eliminieren in der Psychologie:

- oft nicht möglich, da Umgebung immer da (z.B. Beleuchtung, Alter von Personen ...),
- auch nicht immer anzustreben, daher Forderung nach "Feldexperimenten", die in der natürlichen Umgebung stattfinden (im Gegensatz zu Laborexperimenten).

---

2. Beispiel: Versuchsleitereffekt ("Rosenthal-Effekt"): Die Erwartungen des Vl bezüglich der Ergebnisse seiner Untersuchung können zu einer systematischen Verfälschung in Richtung Erwartung führen.

Ausschalten mit dem "Doppelblindversuch":

- Vl weiß nichts über die Hypothesen.
- Vl weiß nicht, welches Treatment gerade läuft.
- Vp weiß nicht, welches Treatment gerade läuft.

Konstanthalten

Konstanthalten

Beispiele:

- Konstante Umgebung schaffen (wenn sie sich nicht eliminieren lässt), gleiches Alter, gleiches Geschlecht, usw.
- "Hawthorne-Effekt" ("Versuchskanincheneffekt", Mayo und Mitarbeiter): Wirkung der Arbeitsplatzbedingungen auf die Arbeitsleistung von Arbeiterinnen
  - Experimentalgruppe: bessere Beleuchtung
  - Kontrollgruppe: normale Beleuchtung

  → nicht konstant-gehaltene Störvariable: das Wissen, Teilnehmer an einem Versuch zu sein

Parallelisieren

Parallelisieren

Beispiel: Einfluss der Lehrmethode auf die Leseleistung Mögliche Störvariable: Intelligenz? Vorgehen:

1. Zunächst wird ein Vortest für Intelligenz durchgeführt.
2. Hinsichtlich der Intelligenz werden nun homogene Blöcke[5] von Vpn gebildet.
3. Die Vpn innerhalb der Blöcke werden nach dem Zufall oder systematisch auf die Gruppen verteilt.

Hinweis: Parallelisieren ist eine spezielle Form der Blockbildung.

Wiederholungsmessung

Wiederholungsmessung

- Die Wiederholungsmessung ist ebenfalls eine spezielle Variante der Blockbildung.
- Durch sie können viele organismische Störvariablen ausgeschaltet werden.

Vorteile

- Man braucht die Störvariablen nicht zu kennen.
- Man braucht in der Regel weniger Vpn.

Nachteile

[5] Paare, Tripel, Quadrupel oder . . . : abhängig von der Anzahl der zu bildenden parallelisierten Gruppen

- Die Methode ist oft nicht anwendbar (z.B. bei Lehrmethoden).
- Reihenfolgeeffekte: Menschen lernen, d.h. vorherige Messungen haben einen Einfluss.
- Allgemein: Übertragungseffekte, d.h. Nachwirkungen von vorangegangenen Versuchsbedingungen auf nachfolgende ("carry-over-effects").
- Empfehlung: mehrere Gruppen verwenden und die Bedingungen ausbalancieren, um Reihenfolge- bzw. Übertragungseffekte zu kontrollieren.

Umwandlung in eine UV

Umwandlung der Störvariable in eine UV

Beispiel: Geschlecht als Störvariable → Geschlecht wird zusätzlich als UV verwendet

Hinweis: Mehrfaktorielle Versuchspläne

→ Neben Haupteffekten sind auch Wechselwirkungen wichtig!

"**Moderatorvariable**": eine Variable, die die Wirkung einer anderen Variable modifiziert, d.h. je nach Ausprägung der Moderatorvariable ist die Wirkung der UV auf die AV unterschiedlich.

Randomisieren

Randomisieren

- Randomisierung ist die am häufigsten angewandte Kontrolltechnik.
- Vpn werden per Zufall den verschiedenen Versuchsgruppen zugeordnet.

Vorteil

Man kann davon ausgehen, dass sich alle möglichen (auch unbekannten) Störvariablen bezüglich ihrer Wirkung zufällig auf die Versuchsbedingungen verteilen. Dies ist umso wahrscheinlicher, je größer die Gruppen sind.

Voraussetzungen

Das Verfahren zur Ziehung der Zufallsstichprobe muss

- gleiche Chancen für alle Mitglieder der Population und
- Unabhängigkeit der Ziehungen gewährleisten.

Empfohlene Strategie

Kontrolltechniken: Empfohlene Strategie

- Einzelne Kontrolltechniken schließen sich nicht aus.
- Man sollte im Experiment möglichst viele Kontrolltechniken anwenden, d.h. sie kombinieren.
- **Vor allem sollte die Randomisierung immer angewandt werden!**

Beispiel: Lernexperiment

- Lärm von draußen → Schild "Bitte nicht stören. Versuch" (Eliminieren)
- Beleuchtung, Temperatur, Vl, usw. → geicher Versuchsraum, gleicher Vl usw. (Konstanthalten)
- Geschlecht der Vp → Gruppen bilden nach Geschlecht (Umwandeln in UV)
- Individuelle Merkfähigkeit → Merkfähigkeit vorher erfassen (Kovariable verwenden)
- Sonstige individuelle Differenzen → Vpn zufällig auf Gruppen verteilen (Randomisieren)

Prinzip der isolierenden Variation

Prinzip der isolierenden Variation

Ein Grundprinzip der Versuchsplanung ist das "**Prinzip der isolierenden Variation**":

1. Manipuliere stets nur <u>eine</u> Variable X "zu einem Zeitpunkt", d.h.
2. Kontrolliere die möglichen Störfaktoren.
3. Wenn dann eine Veränderung in der abhängigen Variable Y eintritt, dann kann
4. X als Ursache dafür angesehen werden.

Übrigens: die Variablen können auch multivariat sein.

Unter diesem Prinzip lassen sich letztlich alle Kontrolltechniken subsumieren.

Konfundierung

- Konfundierung: zwei (oder mehrere) UV variieren gleichzeitig und lassen sich nicht isoliert voneinander untersuchen.
- Dies stellt einen Verstoß gegen das Prinzip der isolierenden Variation dar.

Gütekriterien von Versuchsplänen

Interne Validität

Interne Validität

Die interne Validität[6] eines Versuchsplans betrifft die Interpretierbarkeit der Untersuchungsergebnisse.

Definition:

Eine Untersuchung ist dann intern valide, wenn die gefundenen Effekte eindeutig auf die Treatmentvariable, d.h. auf die verschiedenen Versuchsbedingungen, zurückgeführt werden können.

Faktoren, welche die interne Validität gefährden

- History (Zeitgeschehen)
- Maturation (Reifung)
- Testing (mehrfache Testung)
- Instrumentation (Instrumentierung)
- Statistic Regression (statistische Regression)
- Selection (Auswahlverzerrung)
- Differential Mortality (Ausfalleffekte)
- Interactions (interaktive Effekte)

Die genannten Faktoren bieten jeweils alternative Erklärungsmöglichkeiten für aufgetretene Effekte.

---

[6] Campbell, D.T., & Stanley, J.C. (1963).

History (Zeitgeschehen)

Die völlige experimentelle Isolation ist meist nicht realisierbar.

Es gibt fast immer externe unkontrollierte Ereignisse, die außerhalb der Untersuchung zur gleichen Zeit oder zwischen zwei Untersuchungszeitpunkten auftreten.

Maturation (Reifung)

Hierunter sind spontane zeitliche Veränderungen bei den AVs zu verstehen: z.B. organismische Veränderungen bei der Vp wie Reifung, Ermüdung, Übung.

Testing (Testung)

Beim wiederholten Testen kann es einen Effekt des Vortests auf den Nachtest geben: z.B. Sensibilisierung der Vp als Wirkung des Vortests.

Instrumentation (Instrumentierung)

Darunter sind alle Veränderungen in der Messprozedur zu verstehen: z.B. anderer Versuchsraum, anderer Versuchsleiter, Versuchsleiter anders gelaunt.

Es sind solche Faktoren, die die Objektivität und Reliabilität der Messung herabsetzen.

Statistic Regression (statistische Regression)

Aufgrund unvollkommener Messgenauigkeit tritt bei einer Messwiederholung gegenüber der ersten Messung eine Regression zur Mitte auf.

Dieser Effekt ist umso stärker, je extremer der erste Messwert war.

Vor allem bei Verwendung von Extremgruppen ist mit diesem Effekt immer zu rechnen.

Selection (Auswahlverzerrung)

Das Ergebnis kann von der Stichprobenauswahl abhängen: bei nicht zufälliger Stichprobenauswahl könnten sich die Gruppen von vornherein, ohne Wirkung der Treatments, unterscheiden.

Differential Mortality (Ausfalleffekte)

Dieser Faktor betrifft den unterschiedlichen Ausfall von Vpn in den einzelnen Gruppen.

Interactions (Interaktionseffekte)

Zwischen den bisher genannten Faktoren können auch Interaktionen auftreten:

- **Interaktion zwischen Selection und Maturation**: unterschiedliche spontane Reifungsprozesse bei den verschiedenen Gruppen (z.B. Versuchs- und Kontrollgruppen, die selegiert sind)
- **Interaktion zwischen Selection und Mortality**: aufgrund der Auswahl unterschiedlicher Ausfall von Vpn in den ausgewählten Gruppen
- ...

Externe Validität

Externe Validität

Die externe Validität[7] eines Versuchsplans betrifft die Frage, inwieweit experimentelle Ergebnisse generalisiert werden können.

Definition:

Eine Untersuchung gilt dann als extern valide, wenn man von den Stichproben von Individuen, experimentellen Variablen und Umgebungen auf ebensolche Populationen schließen kann.

Voraussetzung für die Generalisierbarkeit der Ergebnisse ist

1. die **Repräsentativität der Stichprobe**, d.h. der Grad, in dem die Stichprobe ein getreues Abbild der Population ist, und
2. die **Repräsentativität der Untersuchungssituation**, d.h. der Grad, in dem die Untersuchungssituation der "natürlichen" Umgebung ähnlich ist, auf die man generalisieren möchte[8].

Faktoren, welche die externe Validität gefährden

Generell handelt es sich hier um Interaktionen der bisher genannten Einflussfaktoren mit der Treatmentvariable (X). Beispiele:

- Interaction of Testing and X

[7] Campbell, D.T., & Stanley, J.C. (1963).
[8] Brunswik spricht von der "ökologischen Validität".

- Interaktion zwischen Selection und X
- Reactive Arrangements (Reaktive Effekte der Versuchsanordnung)
- Multiple-X Interference (Interferenzeffekte bei mehrfachen Treatments)

Interaktion zwischen Testing und X

Es können durch einen Vortest Sensibilisierungsprozesse auftreten, die die Wirkung des Treatments verstärken oder abschwächen.

Interaktion zwischen Selection und X

Das Treatment könnte sich unterschiedlich auf die ausgewählten Gruppen auswirken.Die

Auswahl der Gruppen könnte nicht repräsentativ sein, z.B. wenn manche Personen die Teilnahme an der Untersuchung abgelehnt haben.

Reaktive Effekte der Versuchsanordnung

Solche Effekte sind z.B. unterschiedliche äußere Bedingungen für die verschiedenen Gruppen oder die Wirkung der experimentellen Situation (z.B. Laboratmosphäre).

Multiple-X Interference (Interferenzeffekte)

Bei Anwendung von mehreren Treatments nacheinander kann es zu Wechselwirkungen zwischen den verschiedenen Treatments kommen.

Versuchspläne: Beispiele

Beispiele von Versuchsplänen

Wir wollen im Folgenden einige ausgewählte Versuchspläne und deren Validität betrachten[9]. Für die einzelnen Komponenten eines Versuchsplanschemas werden folgende Symbole verwendet:

- X: Treatment (UV)
- O: Beobachtung der AV
- R: Randomisierung
- Eine einzelne Zeile: eine Versuchsgruppe

[9] Lit.: Campbell, D.T., & Stanley, J.C. (1963).

- Trennlinie: Gruppen wurden nicht durch Randomisierung vergleichbar gemacht

---

One-shot Case Study

$$X \qquad O$$

- Ein einzelnes Ereignis wird in seinen Auswirkungen beobachtet.
- Es wird kein Vergleich gezogen.
- Was wäre gewesen, wenn X nicht eingetreten wäre?
- Die interne Validität ist nicht gegeben. Nicht kontrolliert sind beispielsweise:
  - History
  - Maturation
  - Selection
  - Mortality

One Group Pretest-Posttest Design

$$O_1 \qquad X \qquad O_2$$

- Unterschiede zwischen den beiden Erhebungen werden X zugeschrieben.
- Die interne Validität ist nicht gewährleistet. Nicht kontrolliert sind:
  - History
  - Maturation
  - Testing
  - Instrumentation
  - Statistical Regression
- Beispiel: Untersuchung von Collier (1940) über die Wirkung nazistischer Propaganda (aber: zwischen der ersten und der zweiten Erhebung lag die Kapitulation Frankreichs)

Static-Group Comparison

$$\begin{array}{cc} X & O_1 \\ \hline & O_2 \end{array}$$

- Die zweite Gruppe soll die Kontrollgruppe darstellen.
- Nicht kontrolliert sind bei diesem Versuchsplan:
  - Selection (die Gruppen könnten sich bereits durch ihre Zusammensetzung unterscheiden)
  - Differential Mortality (Beispiel: “Mädchen ohne Examen sind hübscher als fertige Akademikerinnen” → macht Examen hässlich?)

Pretest-Posttest Control Group Design

R	O_1	X	O_2
R	O_3		O_4

- Dies ist ein echter experimenteller Versuchsplan, d.h. die interne Validität ist gegeben.
- Die externe Validität ist jedoch durch folgende Faktoren eventuell gefährdet:
  - Interaktion zwischen Testing und X
  - Interaktion zwischen Selection und X
  - Andere Interaktionen mit X

Solomon Four-Group Design

R	O_1	X	O_2
R	O_3		O_4
R		X	O_5
R			O_6

- Die interne Validität ist gegeben.
- Zusätzlich ist die Interaktion zwischen Testing und X kontrolliert.
- Weitere Interaktionen mit X sind jedoch möglich.
- Der Einfluss des Testing kann gemessen werden.
- Die Wirkung von X kann auf verschiedene Arten festgestellt werden.

Posttest-Only Control Group Design

R	X	O_1
R		O_2

- Dieser Versuchsplan ist praktisch die zweite Hälfte des Solomon Versuchplans[10].
- Es wird kein Prätest durchgeführt (vielleicht ist er aus irgendwelchen Gründen nicht möglich?).
- Die interne Validität ist bei diesem Versuchsplan ebenfalls gegeben.

Time Series Experiment

$$\boxed{O_1 \quad O_2 \quad O_3 \quad O_4 \quad X \quad O_5 \quad O_6 \quad O_7 \quad O_8}$$

- Experimente dieser Art werden oft in der Physik und Biologie durchgeführt.
- Eine eventuelle Wirkung des Treatments zeigt sich in einer Diskontinuität der Zeitreihe nach Eintreten von X.
- Nicht kontrolliert ist der Faktor History sowie möglicherweise auch Instrumentation.
- Externe Validität:
  - Nicht kontrolliert ist die Interaktion von Testing mit X.
  - Weiterhin ist eine Interaktion zwischen Selection und X sowie ein Effekt der Reactive Arrangements möglich.

Counterbalanced Designs

X_1O	X_2O	X_3O	X_4O
X_2O	X_4O	X_1O	X_3O
X_3O	X_1O	X_4O	X_2O
X_4O	X_3O	X_2O	X_1O

- Ein typisches Beispiel für ein solches Design ist das „lateinische Quadrat“.
- Mit einer solchen Anordnung möchte man Reihenfolgeeffekte eliminieren bzw. kontrollieren.
- Die interne Validität ist gegeben außer einer möglichen Interaktion zwischen verschiedenen Faktoren.
- Hinsichtlich der externen Validität ist neben Interferenz- und reaktiven Effekten auch mit einer Interaktion zwischen X und Testing sowie X und Selection zu rechnen.

[10] Der Effekt des Testing kann allerdings nicht gemessen werden

Stichprobenverfahren I

Population und Stichprobe

Population und Stichprobe

Wissenschaftliche Aussagen: Aussagen über Populationen

Aussagen:

- deskriptiv (z.B. Bevölkerungsstatistiken, Umfragen, Testeichung)
- inferenzstatistisch (z.B. Wahlprognosen, Medikamentenforschung)

Populationen

Populationen (Grundgesamtheiten) können sein:

- Personen, Gruppen, Situationen, ...
- real existierend oder denkbar (durch Bedingungen definiert)
- endlich oder unendlich

Die Population muss genau abgegrenzt werden:

- zeitlich
- räumlich
- personell
- sachlich

Stichproben

Um etwas über eine Population zu erfahren, führt man eine Erhebung durch.

Man unterscheidet

- **Vollerhebungen**: die gesamte Zielpopulation wird befragt
- **Teilerhebungen**: nur eine Stichprobe aus der Zielpopulation wird befragt

---

Vorteile der Teilerhebung

- Geringere Kosten

- Schnellere Ergebnisse
- Gründlichere Durchführung

Aber:

Es ist prinzipiell nicht möglich, mittels Stichproben eine Aussage über eine Grundgesamtheit zu verifizieren (bzw. falsifizieren). Es sind nur Wahrscheinlichkeitsaussagen möglich.

Die wichtigsten Stichprobenverfahren

- Uneingeschränkte Zufallsauswahl
- Schichtenstichprobe
- Klumpenstichprobe
- Mehrstufige Auswahlverfahren
- Quotenstichprobe

Uneingeschränkte Zufallsauswahl

Uneingeschränkte Zufallsauswahl

Dies ist das einfachste Stichprobenverfahren.

- Gegeben ist eine Grundgesamtheit.
- Wir ziehen daraus eine Stichprobe vom Umfang n.

Vorgehen bei der Stichprobenziehung

1. Gesamtliste der Grundgesamtheit verwenden
2. Echtes Zufallsverfahren anwenden, d.h.
   a) gleiche Wahrscheinlichkeit für alle
   b) Unabhängigkeit der Ziehungen
   c) Ziehen: im allg. ohne Zurücklegen

Schätzung der Populationsparameter

Zufallsauswahl: Schätzung von Populationsparametern

Folgende Parameter sollen im Folgenden geschätzt werden:

- Mittelwerte
- Varianzen
- Wahrscheinlichkeiten

Dabei sollen jeweils sowohl Punkt- als auch Intervallschätzungen angegeben werden.

Schätzung des Mittelwerts μ

Schätzfunktion:

$$\hat{\mu} = \bar{X} = \frac{\sum X_i}{n}$$

Für den Erwartungswert dieser Schätzfunktion gilt:

$$E(\bar{X}) = \mu$$

Ihre Varianz ist:

$$Var(\bar{X}) = \frac{\sigma^2}{n}$$

Schätzung der Varianz σ^2

Schätzfunktion bei bekanntem μ:

$$\hat{\sigma}^2_{(n)} = s^2_{(n)} = \frac{\sum (X_i - \mu)^2}{n}$$

Schätzfunktion bei unbekanntem μ:

$$\hat{\sigma}^2_{(n-1)} = s^2_{(n-1)} = \frac{\sum (X_i - \bar{X})^2}{n-1}$$

Für den Erwartungswert gilt in beiden Fällen:

$$E(s^2_{(.)}) = \sigma^2$$

Schätzung der Varianz des Mittelwerts

Voraussetzung:

X_1, X_2, …, X_n unabhängige identisch verteilte Zufallsvariablen mit Varianz σ^2.

Dann gilt für die Varianz der Summe der X_i:

$$Var(\sum X_i) = n\sigma^2$$

und für die Varianz bzw. die Standardabweichung (den "Standardfehler") des Mittelwerts:

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} \qquad \text{bzw.} \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Schätzung von μ: Konfidenzintervall

Nun soll für μ ein Konfidenzintervall angegeben werden. Darin spiegelt sich auch die Genauigkeit der Schätzung wider.

Voraussetzung:

$\bar{X}$ ist eine normalverteilte Zufallsvariable.

Aus der Beziehung

$$Pr\Big\{\mu - z_{1-\frac{\alpha}{2}}\sigma_{\bar{X}} \leq \bar{X} \leq \mu + z_{1-\frac{\alpha}{2}}\sigma_{\bar{X}}\Big\} = 1 - \alpha$$

folgt …

$$\boxed{Pr\Big\{\bar{X} - z_{1-\frac{\alpha}{2}}\sigma_{\bar{X}} \leq \mu \leq \bar{X} + z_{1-\frac{\alpha}{2}}\sigma_{\bar{X}}\Big\} = 1 - \alpha}$$

mit

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Es wird davon ausgegangen, dass die Populationsvarianz bekannt.

Konfidenzintervall bei unbekannter Varianz

Ist die Varianz nicht bekannt, muss sie ebenfalls geschätzt werden:

$$\hat{\sigma}^2 = \frac{\sum(X - \bar{X})^2}{n - 1}$$

In diesem Fall ist nicht die Standardnormalverteilung, sondern die t-Verteilung zu verwenden ...

$$\boxed{Pr\Big\{\bar{X} - t_{(1-\frac{\alpha}{2}),(n-1)}\hat{\sigma}_{\bar{X}} \leq \mu \leq \bar{X} + t_{(1-\frac{\alpha}{2}),(n-1)}\hat{\sigma}_{\bar{X}}\Big\} = 1 - \alpha}$$

mit

$$\hat{\sigma}_{\bar{X}} = \frac{\hat{\sigma}}{\sqrt{n}}$$

Schätzung einer Wahrscheinlichkeit π

Wir betrachten ein Ereignis A, das mit Wahrscheinlichkeit π eintritt und führen n Bernoulli-Versuche durch.

Wenn dabei das Ereignis A r mal eingetreten ist, gilt für die relative Häufigkeit p:

$$p = \frac{r}{n}$$

p wird als Schätzfunktion für π verwendet:

$$p = \hat{\pi}$$

Wir definieren nun eine Zufallsvariable X in folgender Weise:

$$X = \begin{cases} 1 & \text{A ist eingetreten} \\ 0 & \text{sonst} \end{cases}$$

Dann ergibt sich für die relative Häufigkeit:

$$\hat{\pi} = p = \bar{X} = \frac{\sum X_i}{n}$$

Auch hier gilt: $E(p) = \pi$.

Für die Varianz von X gilt übrigens:

$$\sigma_X^2 = \pi(1 - \pi)$$

Sie wird geschätzt mit:

$$\hat{\sigma}_X^2 = p(1 - p)$$

Also ergibt sich für die Varianz bzw. die Standardabweichung (den "Standardfehler") von p:

$$\sigma_p^2 = \frac{\pi(1-\pi)}{n} \qquad \text{bzw.} \qquad \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

Sie werden geschätzt mit:

$$\hat{\sigma}_p^2 = \frac{p(1-p)}{n} \qquad \text{bzw.} \qquad \hat{\sigma}_p = \sqrt{\frac{p(1-p)}{n}}$$

Konfidenzintervall für eine Prozentzahl

Damit können wir für das Konfidenzintervall einer Prozentzahl dieselbe Formel verwenden wie für den Mittelwert.

Wir ersetzen $\bar{X}$ durch p und $\hat{\sigma}^2$ durch $p(1-p)$ und erhalten die Näherungsformel:

$$\boxed{Pr\Big\{p - z_{1-\frac{\alpha}{2}}\sqrt{\tfrac{p(1-p)}{n}} \le \pi \le p + z_{1-\frac{\alpha}{2}}\sqrt{\tfrac{p(1-p)}{n}}\Big\} = 1-\alpha}$$

Zentraler Grenzwertsatz

Exkurs: Der zentrale Grenzwertsatz

Die Summe von unabhängigen, identisch verteilten Zufallsvariablen ist asymptotisch normalverteilt.

X_1, X_2, X_3, ... seien unabhängige, identisch verteilte („iid") Zufallsvariablen mit Erwartungswert μ und Varianz σ^2 und ansonsten beliebiger Verteilung.

Wir betrachten nun die n-te Teilsumme dieser Zufallsvariablen

$$S_n = X_1 + X_2 + \cdots + X_n$$

Der Erwartungswert von S_n ist $n\mu$ und die Varianz ist $n\sigma^2$.

Der zentrale Grenzwertsatz besagt, dass die Verteilung von S_n für $n \to \infty$ gegen die Normalverteilung $\mathcal{N}(n\mu, n\sigma^2)$

bzw. die Größe

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

gegen die Standardnormalverteilung geht.

Stichprobenverfahren II

Schichtenstichprobe

Schichtenstichprobe

Vorgehen bei der Stichprobenziehung

- Die Population wird in Teilpopulationen eingeteilt.
- Die Einteilung muss disjunkt und vollständig sein.
- Aus jeder Teilpopulation wird getrennt eine Stichprobe gezogen (jeweils: uneingeschränkte Zufallsauswahl!).

Vorteil der Schichtenstichprobe

- Bei gleichem Stichprobenumfang ist eine höhere Genauigkeit möglich als bei der uneingeschränkten Zufallsauswahl.
- Oder: für die gleiche Genauigkeit ist ein kleinerer Stichprobenumfang nötig.

Nachteil der Schichtenstichprobe

- Das Schichtmerkmal muss mit erhoben werden.

Voraussetzungen für die Anwendung

- Die Zusammensetzung der Schichten in der Population muss bekannt sein.
- Das Schichtmerkmal sollte mit dem zu untersuchenden Merkmal zusammenhängen.

- Die Schichtenstichprobe ist von Vorteil bei möglichst großer Heterogenität zwischen den Schichten und möglichst großer Homogenität innerhalb der Schichten.

Notation

Mit k bezeichnen wir die Anzahl der Schichten.

Wir betrachten erst den Fall **endlicher Populationen**:

- N: Gesamtumfang der Population
- N_j: Umfang der j-ten Schicht ($\sum N_j = N$)
- g_j: Anteil der j-ten Schicht ($g_j = \frac{N_j}{N}$) an der Gesamtpopulation
- μ_j: Mittelwert des zu untersuchenden Merkmals in der j-ten Schicht
- σ_j^2: Varianz des zu untersuchenden Merkmals in der j-ten Schicht

---

Für den Fall **unendlicher Populationen** haben wir:

- g_j: Anteil der j-ten Schicht (es gilt: $\sum g_j = 1$) an der Gesamtpopulation
- μ_j: Mittelwert des zu untersuchenden Merkmals in der j-ten Schicht
- σ_j^2: Varianz des zu untersuchenden Merkmals in der j-ten Schicht

---

Für die Stichprobe verwenden wir die folgende Notation:

- n: Gesamtstichprobenumfang
- n_j: Stichprobenumfang der j-ten Schicht
- $\bar{X}_j$: Mittelwert des zu untersuchenden Merkmals in der j-ten Schicht
- s_j^2: Varianz des zu untersuchenden Merkmals in der j-ten Schicht

---

Gesamtmittelwert μ der Population

$$\mu = \sum g_j \mu_j$$

μ ist also ein "gewichteter Mittelwert".

Schätzung des Mittelwerts μ_j der j-ten Schicht

$$\hat{\mu_j} = \bar{X}_j = \frac{\sum_{i=1}^{n_j} X_i^{(j)}}{n_j}$$

Schätzung des Varianz σ_j^2 der j-ten Schicht

$$\hat{\sigma_j^2} = s_j^2 = \frac{\sum_{i=1}^{n_j} (X_i^{(j)} - \bar{X}_j)^2}{n_j - 1}$$

Schätzung des Gesamtmittelwerts μ

$$\hat{\mu} = \bar{X} = \sum g_j \bar{X}_j$$

Standardfehler von $\bar{X}$

$$\sigma_{\bar{X}}^2 = \sum g_j^2 \frac{\sigma_j^2}{n_j}$$

Also:

$$\sigma_{\bar{X}} = \sqrt{\sum g_j^2 \frac{\sigma_j^2}{n_j}}$$

Schätzung des Standardfehlers von $\bar{X}$

$$\hat{\sigma}_{\bar{X}} = \sqrt{\sum g_j^2 \frac{s_j^2}{n_j}}$$

Konfidenzintervall für den Mittelwert

$$\boxed{Pr\{\bar{X} - z_{1-\frac{\alpha}{2}} \sigma_{\bar{X}} \leq \mu \leq \bar{X} + z_{1-\frac{\alpha}{2}} \sigma_{\bar{X}}\} = 1 - \alpha}$$

bzw.

$$\boxed{Pr\{\bar{X} - t_{(1-\frac{\alpha}{2}),(n-k)} \hat{\sigma}_{\bar{X}} \leq \mu \leq \bar{X} + t_{(1-\frac{\alpha}{2}),(n-k)} \hat{\sigma}_{\bar{X}}\} = 1 - \alpha}$$

Schätzung eines Prozentwerts π

Das Vorgehen ist völlig analog zum Vorgehen bei der uneingeschränkten Zufallsauswahl, d.h. wir substituieren

- μ durch π
- $\bar{X}$ durch p

und verwenden die obigen Formeln für geschichtete Stichproben.

Auch hier gilt wiederum:

$$\sigma_j^2 = \pi_j(1-\pi_j)$$

und

$$\hat{\sigma}_j^2 = p_j(1-p_j)$$

Wir haben also:

$$\hat{\sigma}_p^2 = \sum g_j^2 \frac{p_j(1-p_j)}{n_j}$$

und

$$\hat{\sigma}_p = \sqrt{\sum g_j^2 \frac{p_j(1-p_j)}{n_j}}$$

Wir erhalten also als Näherungsformel:

$$\boxed{Pr\{p - z_{1-\frac{\alpha}{2}}\hat{\sigma}_p \leq \pi \leq p + z_{1-\frac{\alpha}{2}}\hat{\sigma}_p\} = 1-\alpha}$$

Aufteilung des Stichprobenumfangs auf die Schichten

Im Prinzip gibt es viele mögliche Aufteilungen des Gesamtstichprobenumfangs.

Zwei Spezialfälle sind hervorzuheben:

- Die proportionale Aufteilung
- Die bestmögliche Aufteilung

Unabhängig von der Art der Aufteilung bleiben die Formeln für die Schätzung des Mittelwerts und das Konfidenzintervall natürlich allesamt gültig[11].

Die proportionale Aufteilung

$$n_j = ng_j$$

Die Zahlenverhältnisse in der Stichprobe entsprechen jenen in der Population.

Die bestmögliche Aufteilung

$$n_j = n\frac{g_j\sigma_j}{\sum g_i\sigma_i}$$

- Hier wird neben den Proportionen auch noch die Varianz des Merkmals innerhalb der Schichten berücksichtigt, d.h. Schichten mit größerer Heterogenität gehen mit einem größeren Stichprobenumfang ein.
- Die bestmögliche Aufteilung erbringt die höchste Schätzgenauigkeit.
- Sind die Varianzen in den Schichten alle gleich, so ergibt sich die proportionale Aufteilung.

0.1 Quotenstichprobe

Quotenstichprobe

- In der Markt- und Meinungsforschung ist dieses Stichprobenverfahren stark verbreitet.
- Ähnlich wie bei der Schichtenstichprobe wird die Population aufgeteilt, allerdings normalerweise nach mehreren Merkmalen[12] gleichzeitig.

Eigenschaften der Quotenstichprobe

- Die Auswahl der Personen ist nicht zufällig, sondern bleibt dem Interviewer überlassen.

[11]Durch Einsetzen und Umformen vereinfachen sie sich allenfalls.
[12]Z.B. Geschlecht, Alter, Bildung

- Bei mehr als einem Auswahlmerkmal werden nur die zu erreichenden Randhäufigkeiten und nicht die verbundenen Häufigkeiten vorgegeben [13].

Vorteile der Quotenstichprobe

- Kostenvorteile
- Organisatorische Vorteile (Verfügbarkeit, Freiwilligkeit)

Nachteile der Quotenstichprobe

- Statistische Aussagen (Schätzwerte und deren Genauigkeit) sind fraglich.
- Mangelnde Repräsentativität

0.2 Mehrstufige Auswahlverfahren und Klumpenstichprobe

Mehrstufige Auswahlverfahren

Bei einem mehrstufigen Auswahlverfahren wird folgendermaßen vorgegangen:

- Die Population wird in “Primäreinheiten” eingeteilt (z.B. Länder, Betriebe).
- Einige Primäreinheiten werden nach dem Zufall gezogen.
- Die Primäreinheiten können weiter aufgeteilt werden in “Sekundäreinheiten” (z.B. Städte, Abteilungen).
- Von diesen werden wiederum einige nach dem Zufall gezogen.
- usw.

Klumpenstichprobe

- Bei einem mehrstufigen Auswahlverfahren können auf der untersten Stufe alle Mitglieder untersucht werden oder nur eine Zufallsstichprobe.
- Werden auf der untersten Stufe alle Mitglieder untersucht (Vollerhebung), so spricht man von einer **Klumpenstichprobe**.

Voraussetzungen für die Anwendung des Verfahrens

[13] Dadurch kann es theoretisch vorkommen, dass bestimmte Kombinationen der Auswahlmerkmale in der Stichprobe nicht vertreten oder unterrepräsentiert sind.

- Alle möglichen Einheiten bzw. Klumpen müssen zugänglich sein und für die Untersuchung prinzipiell zur Verfügung stehen.
- Das Merkmal, nach dem die Aufteilung der Populationen erfolgt, sollte mit dem zu untersuchenden Merkmal **nicht** zusammenhängen.
- Die Einheiten bzw. Klumpen sollten in sich möglichst heterogen sein.
- Die Klumpenstichprobe ist von Vorteil bei möglichst großer Homogenität zwischen den Schichten und möglichst großer Heterogenität innerhalb der Teilpopulationen.

Vorteile

- Organisatorische Vorteile
- Kostenvorteil
- Zuweilen die einzige Möglichkeit [14]

Nachteile

- Mangelnde Repräsentativität?
- Gesamtstichprobenumfang zufallsabhängig

t-Tests

Übersicht über die t-Tests

- Vergleich einer Stichprobe mit einer Population (Ein-Stichproben t-Test)
- Vergleich von zwei unabhängigen Stichproben
  - bei homogenen Varianzen (t_{hom})
  - bei heterogenen Varianzen (t_{het})
- Vergleich von zwei abhängigen Stichproben (t_{corr})

Ein-Stichproben t-Test

Ein-Stichproben t-Test

[14] Eine uneingeschränkte Zufallsauswahl bzw. Schichtenstichprobe ist oft nicht einsetzbar bzw. auch nicht nicht sinnvoll, beispielsweise bei Schulklassen oder Betrieben.

Voraussetzungen

- Gegeben ist eine normalverteilte Population mit unbekannten Parametern μ und σ^2.
- Daraus wird eine Zufallsstichprobe vom Umfang n gezogen.
- Wit betrachten also unabhängige, normalverteilte Zufallsvariablen $X_1 \dots X_n$ mit Mittelwert μ und Varianz σ^2.

Hypothese

- Nullhypothese: $\mu = \mu_0$
- Alternativhypothese:
  - Zweiseitige Fragestellung: $\mu \neq \mu_0$
  - Einseitige Fragestellung: $\mu < \mu_0$ bzw. $\mu > \mu_0$

Teststatistik

$$\boxed{t = \frac{\bar{X} - \mu_0}{\hat{\sigma}_{\bar{X}}}}$$

mit

$$\hat{\sigma}_{\bar{X}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{s}{\sqrt{n}}$$

Sind die obigen Vorausetzungen erfüllt, so gilt unter H_0:

Die Teststatistik t ist t-verteilt mit $n-1$ Freiheitsgraden.

t-Tests für zwei unabhängige Stichproben

Zwei-Stichprobent t-Test bei homogenen Varianzen

Voraussetzungen

- Gegeben sind zwei normalverteilte Populationen mit unbekannten Parametern μ_1, μ_2, σ_1, σ_2.

- $\sigma_1 = \sigma_2 = \sigma$.
- Aus jeder Population wurde eine Zufallsstichprobe gezogen (Stichprobenumfänge: n_1 bzw. n_2).

Hypothesen

- Nullhypothese: $\mu_1 = \mu_2 = \mu$
- Alternativhypothese:
  - Zweiseitige Fragestellung: $\mu_1 \neq \mu_2$
  - Einseitige Fragestellung: $\mu_1 < \mu_2$ bzw. $\mu_1 > \mu_2$

Teststatistik

$$\boxed{t = \frac{\bar{X}_1 - \bar{X}_2}{\hat{\sigma}_{\bar{X}_1 - \bar{X}_2}}}$$

mit

$$\hat{\sigma}_{\bar{X}_1 - \bar{X}_2} = \hat{\sigma}\sqrt{\tfrac{1}{n_1} + \tfrac{1}{n_2}}$$

und

$$\hat{\sigma} = \sqrt{\tfrac{s_1^2(n_1-1)+s_2^2(n_2-1)}{(n_1-1)+(n_2-1)}}$$

Sind die Stichprobenumfänge gleich, so vereinfacht sich die Formel zu

$$\hat{\sigma} = \sqrt{\tfrac{s_1^2+s_2^2}{2}}$$

Sind die obigen Vorausetzungen erfüllt, so gilt unter H_0:

Die Teststatistik t ist t-verteilt mit $n_1 + n_2 - 2$ Freiheitsgraden.

t-Test bei heterogenen Varianzen

Voraussetzungen

- Gegeben sind zwei normalverteilte Populationen mit unbekannten Parametern μ_1, μ_2, σ_1, σ_2.
- Varianzengleichheit wird nicht vorausgesetzt.
- Aus jeder Population wurde eine Zufallsstichprobe gezogen (Stichprobenumfänge: n_1 bzw. n_2).

Hypothesen

- Nullhypothese: $\mu_1 = \mu_2 = \mu$
- Alternativhypothese:
  - Zweiseitige Fragestellung: $\mu_1 \neq \mu_2$
  - Einseitige Fragestellung: $\mu_1 < \mu_2$ bzw. $\mu_1 > \mu_2$

Teststatistik

$$\boxed{t = \frac{\bar{X}_1 - \bar{X}_2}{\hat{\sigma}_{\bar{X}_1 - \bar{X}_2}}}$$

mit

$$\hat{\sigma}_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Die Freiheitsgrade sind

$$df = \frac{(n_1-1)(n_2-1)}{(n_2-1)c^2+(n_1-1)(1-c^2)}$$

mit

$$c = \frac{s_1^2/n_1}{s_1^2/n_1 + s_2^2/n_2} = \frac{n_2 s_1^2}{n_2 s_1^2 + n_1 s_2^2}$$

Sind die Stichprobenumfänge gleich ($n_1 = n_2 = n$), so gilt

$$df = n - 1$$

Sind die obigen Vorausetzungen erfüllt, so gilt unter H_0:

Die Teststatistik t ist t-verteilt mit df Freiheitsgraden.

t-Test für verbundene Stichproben

t-Test für verbundene Stichproben

Voraussetzungen

- Die Zufallsvariablen

$$\begin{aligned} D_1 &= X_1 - Y_1, \\ &\dots \\ D_n &= X_n - Y_n \end{aligned}$$

  sind unabhängig und stammen aus einer Normalverteilung mit Mittelwert μ_D und Standardabweichung σ_D.

Hypothesen

- Nullhypothese: $\mu_D = 0$
- Alternativhypothese:
  - Zweiseitige Fragestellung: $\mu_D \neq 0$
  - Einseitige Fragestellung: $\mu_D < 0$ bzw. $\mu_D > 0$

Teststatistik

$$\boxed{t = \frac{\bar{D}}{\hat{\sigma}_D}}$$

mit

$$\hat{\sigma}_D = s_D = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2}$$

Sind die obigen Vorausetzungen erfüllt, so gilt unter H_0:

Die Teststatistik t ist t-verteilt mit $n-1$ Freiheitsgraden.

Einfache Varianzanalyse

Zweck

Vergleich der Mittelwerte mehrerer unabhängiger Gruppen.

Beispielsweise soll die Wirkung verschiedener Stufen einer unabhängigen Variable, d.h. verschiedener Bedingungen, Behandlungen und dgl. auf die abhängige Variable getestet werden.

Modell

Modell

Voraussetzungen

1. Gegeben seien p normalverteilte Populationen mit unbekannten Parametern $\mu_1, \mu_2, \ldots, \mu_p$ und $\sigma_1, \sigma_2, \ldots, \sigma_p$.
2. $\sigma_1 = \sigma_2 = \cdots = \sigma_p = \sigma$.
3. Aus jeder Population i wird eine Zufallsstichprobe mit dem Stichprobenumfang n_i gezogen.

Modell

Für den Wert der abhängigen Variable X von Person m aus der Stichprobe i nehmen wir das folgende Modell an:

1. $X_{im} = \mu_i + \varepsilon_{im}$
2. $\varepsilon_{im} \overset{\text{iid}}{\hookrightarrow} \mathcal{N}(0, \sigma_\varepsilon^2)$.

   In Worten: die Fehler sind unabhängige normalverteilte Zufallsvariablen mit Erwartungswert 0 und homogenen Varianzen[15].

Strukturmodell

Da wir weniger an den Mittelwerten μ_i selbst, sondern an den sog. <u>Effekten</u> interessiert sind, nehmen wir eine Reparametrisierung des Modells vor:

- Sei μ der Gesamtmittelwert der p Populationen.

[15]Wir bezeichnen die Fehlervarianz mit σ_ε^2. Es gilt $\sigma_\varepsilon^2 = \sigma^2$ (siehe Voraussetzungen).

- Wir definieren nun die Effekte α_i:

$$\alpha_i = \mu_i - \mu$$

- Trivialerweise gilt:

$$\sum_{i=1}^{p} \alpha_i = 0$$

Wir schreiben nun das Modell mit Hilfe der neuen Parameter und erhalten das sog. Strukturmodell der einfaktoriellen Varianzanalyse:

$$X_{im} = \mu + \alpha_i + \varepsilon_{im}$$

Der Messwert einer Person m in der Gruppe i setzt sich also additiv zusammen aus dem Gesamtmittelwert, dem Effekt der Bedingung i sowie einer Zufallskomponente, dem "Fehler".

Im Gegensatz zu den ε_{im} sind die Parameter μ und α_i keine Zufallsvariablen, sondern fixe Größen.

Für den Erwartungswert und die Varianz von X_{im} gilt:

$$\begin{aligned} E(X_{im}) &= \mu + \alpha_i \\ Var(X_{im}) &= \sigma_\varepsilon^2 \end{aligned}$$

Die eingangs genannten Voraussetzungen der Varianzanalyse könnten also, bezogen auf das Strukturmodell, auch so formuliert werden: die ε_{im} sind unabhängige, identisch normalverteilte Zufallsvariablen mit Mittelwert 0 und Varianz σ_ε^2.

Hypothesen

Hypothesen

Nullhypothese

$$\mu_1 = \cdots = \mu_p = \mu$$

bzw.

$$\alpha_1 = \cdots = \alpha_p = 0$$

Eine äquivalente Formulierung wäre:

$$\sum_{i=1}^{p} \alpha_i^2 = 0 \qquad \text{bzw.} \qquad \sigma_\alpha^2 = 0$$

Alternativhypothese

Für mindestens ein i und j gilt

$$\mu_i \neq \mu_j$$

bzw. für mindestens ein i gilt

$$\alpha_i \neq 0$$

Eine äquivalente Formulierung wäre:

$$\sum_{i=1}^{p} \alpha_i^2 > 0 \qquad \text{bzw.} \qquad \sigma_\alpha^2 > 0$$

Teststatistik

Teststatistik

Für die Teststatistik müssen Quadratsummen, Freiheitsgrade berechnet werden:

$$\begin{aligned}
QS_{zw} &= \sum_{i=1}^{p} n_i(\bar{X}_{i.} - \bar{X}_{..})^2 \\
QS_{inn} &= \sum_{i=1}^{p} \sum_{m=1}^{n_i} (X_{im} - \bar{X}_{i.})^2 \\
df_{zw} &= p - 1 \\
df_{inn} &= \sum_{i=1}^{p} (n_i - 1) = \sum_{i=1}^{p} n_i - p
\end{aligned}$$

Sind die Stichprobenumfänge gleich ($n_1 = n_2 = \cdots = n$)[16], so vereinfachen sich die Formeln etwas:

[16] Davon wollen wir bis auf weiteres immer ausgehen.

$$
\begin{aligned}
QS_{zw} &= n\sum_{i=1}^{p}(\bar{X}_{i.} - \bar{X}_{..})^2 \\
QS_{inn} &= \sum_{i=1}^{p}\sum_{m=1}^{n}(X_{im} - \bar{X}_{i.})^2
\end{aligned}
$$

$$
\begin{aligned}
df_{zw} &= p - 1 \\
df_{inn} &= p(n-1)
\end{aligned}
$$

Daraus erhält man die mittleren Quadrate:

$$
\begin{aligned}
MQ_{zw} &= \frac{QS_{zw}}{df_{zw}} \\
MQ_{inn} &= \frac{QS_{inn}}{df_{inn}}
\end{aligned}
$$

Die Teststatistik lautet schließlich:

$$F = \frac{MQ_{zw}}{MQ_{inn}}$$

Sind die Voraussetzungen erfüllt, so gilt unter H_0:

> Die Teststatistik F ist F-verteilt mit $df_1 = df_{zw}$ und $df_2 = df_{inn}$ Freiheitsgraden.

Begründung für den F-Test

Man kann zeigen, dass unter der Nullhypothese MQ_{inn} und MQ_{zw} zwei unabhängige erwartungstreue Schätzungen derselben Varianz σ_ε^2 sind.

Die Erwartungswerte der mittleren Quadrate sind:

$$
\begin{aligned}
E(MQ_{inn}) &= \sigma_\varepsilon^2 \\
E(MQ_{zw}) &= n\sigma_\alpha^2 + \sigma_\varepsilon^2
\end{aligned}
$$

Unter H_0 ist $\sigma_\alpha^2 = 0$, so dass man erhält:

$$E(MQ_{zw}) = \sigma_\varepsilon^2$$

Herleitung der Erwartungswerte

Wenn die Voraussetzungen erfüllt sind, gilt unter H_0 für die Varianz der Mittelwerte:

$$\sigma^2_{\bar{X}_{i.}} = \frac{\sigma^2_\varepsilon}{n} \quad \text{bzw.} \quad n\sigma^2_{\bar{X}_{i.}} = \sigma^2_\varepsilon$$

$\sigma^2_{\bar{X}_{i.}}$ kann durch die Varianz der Stichprobenmittelwerte geschätzt werden:

$$\hat{\sigma}^2_{\bar{X}_{i.}} = \frac{\sum(\bar{X}_{i.} - \bar{X}_{..})^2}{p-1}$$

Nun gilt aber

$$n\hat{\sigma}^2_{\bar{X}_{i.}} = \frac{n\sum(\bar{X}_{i.} - \bar{X}_{..})^2}{p-1} = MQ_{zw}$$

Also ist MQ_{zw} eine Schätzung von σ^2_ε.

Unabhängig davon kann σ^2_ε aus den Daten jeder Stichprobe geschätzt werden:

$$\hat{\sigma}^2_{\varepsilon(i)} = \frac{\sum_{m=1}^n (X_{im} - \bar{X}_{i.})}{n-1}$$

Der Mittelwert dieser Schätzungen (die „gepoolte Varianz innerhalb“) stellt eine gemeinsame Schätzung von σ^2_ε dar. Dieser Mittelwert entspricht exakt MQ_{inn}.

Quadratsummenzerlegung

Quadratsummenzerlegung

Der Ausdruck $\sum\sum(X_{im} - \bar{X}_{..})^2$ wird als totale Quadratsumme QS_{total} bezeichnet. Man kann leicht zeigen, dass für sie die folgende additive Beziehung gilt:

$$QS_{total} = QS_{zw} + QS_{inn}$$
$$\sum_i \sum_m (X_{im} - \bar{X}_{..})^2 = n\sum_i (\bar{X}_{i.} - \bar{X}_{..})^2 + \sum_i \sum_m (X_{im} - \bar{X}_{i.})^2$$

Für die Freiheitsgrade gilt die analoge Beziehung

$$df_{total} = df_{zw} + df_{inn}$$

Die Quadratsummenzerlegung kann das folgende Baumdiagramm veranschaulicht werden:

$$QS_{total}$$

$$QS_{zw} \qquad\qquad QS_{inn}$$

Rechnerische Durchführung der Varianzanalyse

Für die rechnerische Durchführung der Varianzanalyse müssen die Größen QS_{zw}, QS_{inn}, df_{zw}, df_{inn}, MQ_{zw}, MQ_{inn} und F berechnet werden.

Der F-Wert ist mit dem kritischen F-Wert

$$F_{krit} = F_{1-\alpha;df_{zw};df_{inn}}$$

zu vergleichen. Übersteigt er diesen, so ist H_0 zu verwerfen.

Die Ergebnisse werden normalerweise in einer Tabelle zusammengefasst:

Quelle der Variation	**QS**	**df**	**MQ**	**F**
Zwischen				
Innerhalb				——
Total			——	——

Teststärke der Varianzanalyse

Teststärke der Varianzanalyse

Effektstärke

Als Maß für die Stärke des Einflusses der UV auf die AV ist folgende Größe definiert:

$$\eta^2 = \frac{QS_{zw}}{QS_{total}}$$

Teststärke

Bei Gültigkeit der Voraussetzungen ist die Teststärke der Varianzanalyse umso höher

- je größer das Risiko α,
- je größer der Stichprobenumfang n,
- je größer die Effektstärke,
- je kleiner die Varianz innerhalb der Gruppen σ_ε^2.

Konsequenzen bei Verletzung der Voraussetzungen

Konsequenzen bei Verletzung der Voraussetzungen

Generell wurde nachgewiesen, dass der F-Test robust ist gegen die Verletzung einiger seiner Voraussetzungen.

Im Detail hat sich gezeigt:

- Verletzung der Unabhängigkeitsvoraussetzung: der F-Test ist verfälscht.
- Verletzung der Normalverteilungsvoraussetzung:
  - Schiefwinklige Verteilung: F-Test ist kaum verfälscht.
  - Schmalgipflige Verteilung: F-Test ist konservativ ($\alpha_{nominell} > \alpha_{real}$), die Teststärke ist kleiner
  - Breitgipflige Verteilung: F-Test ist progressiv ($\alpha_{nominell} < \alpha_{real}$), die Teststärke ist größer

---

- Verletzung der Homogenitätsvoraussetzung:
  - Gleiche Stichprobenumfänge: F-Test ist robust
  - Ungleiche Stichprobenumfänge: Ergebnisse verfälscht, vor allem bei kleinen Stichprobenumfängen (vor allem, wenn die kleineren Stichproben die größeren Varianzen haben)

Beispiel

Numerisches Beispiel[17]

Phobiker wurden mit einer Konfrontationstherapie (F), einer kognitiven Therapie (K) oder beiden (F+K) behandelt, die Kontrollgruppe wartet ohne Behandlung.

Nach der Behandlung wurde die abhängige Variable erhoben.

Dazu wurde die Vp mit einem angstauslösenden Reiz konfrontiert.

Gemessen wurde die Zeit, wie lange die Vp diese Testsituationausgehalten hat.

Daten ($p = 4$, $n = 3$):

Kontrollgruppe	8,7,9
F	11,9,10
K	9,10,11
F+K	13,11,12

Auswertung

...

Die Prüfung spezieller Hypothesen

Problem

- Wir nehmen an, wir hätten eine Varianzanalyse durchgeführt und ein signifikantes Ergebnis erhalten, d.h. wir hätten mit einer Irrtumswahrscheinlichkeit α die Nullhypothese abgelehnt.
- Dies bedeutet, dass mindestens zwei Populationsmittelwerte verschieden sind.
- Leider sind die Ergebnisse der Varianzanalyse unspezifisch, denn wir wissen damit nicht, welche Mittelwerte zu dem signifikanten Ergebnis beigetragen haben.
- Wir wollen also versuchen, spezielle Mittelwertsunterschiede näher zu untersuchen.

Lösungsansätze

Es gibt zwei Möglichkeiten:

- Man prüft jeden Mittelwert gegen jeden (paarweise Mittelwertsvergleiche, multiple Vergleiche).
- Man testet spezielle Hypothesen: in diesem Fall betrachtet man nicht jeweils zwei Mittelwerte isoliert, sondern mehrere bzw. alle Mittelwerte simultan.

  Hypothesen dieser Art können entweder

  - a priori (vorher) oder
  - a posteriori (erst im Nachhinein)

  aufgestellt werden.

Multiple Vergleiche

Multiple Vergleiche

Wir stellen mehrere Vergleiche zwischen Mittelwerten an. Dabei unterscheiden wir ein Risiko α^* für den Fehler 1. Art bei einem einzelnen Vergleich und ein Gesamtrisiko α dafür, den Fehler 1. Art bei mindestens einem Vergleich zu machen.

Ziel der multiplen Vergleiche:

Das Gesamtrisiko α soll kontrolliert werden, d.h. einen bestimmten Wert nicht übersteigen.

Nicht nur in diesem Kontext, sondern allgemein gilt, wenn mehrere Tests (Vergleiche) durchgeführt werden, und die einzelnen Vergleiche unabhängig sind:

Wenn bei einem einzelnen Test die Wahrscheinlichkeit, einen Fehler 1. Art zu begehen, gleich α^* ist, so ist die Wahrscheinlichkeit α, bei k unabhängigen Vergleichen mindestens einen Fehler 1. Art zu begehen:

$$\alpha = 1 - (1 - \alpha^*)^k$$

Also habe ich für den einzelnen Vergleich zu wählen:

$$\alpha^* = 1 - (1 - \alpha)^{\frac{1}{k}}$$

Dann wird das Gesamtrisiko α eingehalten.

Im allgemeinen gilt aber die Unabhängigkeit nicht, und so gibt mehrere Verfahren, die das Gesamtrisiko α kontrollieren, z.B.:

- Scheffé Test
- Tukey HSD
- Bonferroni t-Test.

Lineare Kontraste

Lineare Kontraste

Ein Kontrast von irgendwelchen Größen $V_1, V_2, \ldots, V_p$ (z.B. Variablen, Parametern) ist eine Linearkombination dieser Größen mit Koeffizienten c_i, die nicht alle gleich 0 sind und deren Summe gleich 0 ist:

$$\sum c_i V_i$$

mit $\quad \sum c_i = 0 \quad$ und $\quad \sum c_i^2 > 0.$

Normierte Kontrastkoeffizienten c_i^* erhält man durch folgende Umrechnung:

$$c_i^* = \frac{c_i}{\sqrt{\sum c_j^2}}$$

Insbesondere betrachten wir nun Kontraste zwischen Populationsmittelwerten.

Ein Kontrast ψ zwischen Populationsmittelwerten μ_i wird mit Hilfe von Koeffizienten c_i folgendermaßen definiert:

$$\psi = \sum_{i=1}^{p} c_i \mu_i$$

Unter Verwendung von normierten Kontrastkoeffizienten c_i^* erhält man einen normierten Kontrast:

$$\psi^* = \sum_{i=1}^{p} c_i^* \mu_i$$

Kontraste zwischen Populationsmittelwerten werden durch entsprechende Kontraste zwischen Stichprobenmittelwerten geschätzt:

$$\hat{\psi} = \sum_{i=1}^{p} c_i \bar{X}_{i.}$$

bzw.

$$\hat{\psi}^* = \sum_{i=1}^{p} c_i^* \bar{X}_{i.}$$

Spezielle Hypothesen sind jeweils durch eine geeignete Wahl der Koeffizienten c_i auszudrücken.

Beispiele

...

Test von linearen Kontrasten

Die Nullhypothese lautet:

$H_0 : \psi = 0$

Als Alternativhypothese hat man:

$H_1 : \psi \neq 0$ (zweiseitige Fragestellung)

bzw.

$H_1 : \psi > 0$ bzw. $H_1 : \psi < 0$ (einseitige Fragestellung)

Quadratsumme eines Kontrasts (für gleiche n_i)

$$QS_\psi = n \frac{\hat{\psi}^2}{\sum c_i^2}$$

Liegt ein normierter Kontrast ψ^* vor, so gilt:

$$QS_{\psi^*} = n\hat{\psi^*}^2$$

Ein einzelner Kontrast hat immer 1 Freiheitsgrad.

Skaleninvarianz

Die Größenordnung der Koeffizienten spielt beim Test eines Kontrasts keine Rolle: multipliziert man jeden Kontrastkoeffizienten mit demselben Faktor, so ändert sich zwar der Wert des Kontrasts um diesen Faktor, aber die Quadratsumme des Kontrasts bleibt unverändert.

Test eines einzelnen a priori Kontrastes

$$F = \frac{QS_\psi}{MQ_{inn}}$$

mit $\quad df_1 = 1 \quad$ und $\quad df_2 = df_{inn}$.

Test mehrerer a priori Kontraste

Die Prüfgröße ist dieselbe wie im vorigen Fall, nämlich wieder:

$$F = \frac{QS_\psi}{MQ_{inn}}$$

mit $df_1 = 1$ und $df_2 = df_{inn}$.

Allerdings muss das α nach Bonferroni korrigiert, d.h durch α^* ersetzt werden:

$$\alpha^* = \alpha/k \qquad (k\text{: Anzahl der Kontraste})$$

Der Nachteil ist, dass das Risiko für den Fehler zweiter Art ansteigt.

Test mehrerer a posteriori Kontraste nach Scheffé

Voraussetzung: der F-Wert der einfachen Varianzanalyse ist signifikant.

Die Prüfgröße ist

$$F = \frac{QS_\psi/(p-1)}{MQ_{inn}}$$

Die Freiheitsgrade sind hier:

$df_1 = p - 1$ und $df_2 = df_{inn}$.

Orthogonale Kontraste

Orthogonale Kontraste

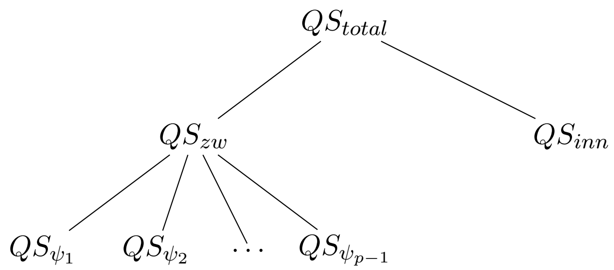
Zwei lineare Kontraste mit den Koeffizienten $c_1, \ldots, c_p$ und $c'_1, \ldots, c'_p$ sind zueinander orthogonal, wenn gilt:

$$\sum_i^p c_i c'_i = 0$$

- Orthogonale Kontraste spielen eine besondere Rolle innerhalb der Varianzanalyse: sie erfassen jeweils verschiedene, unabhängige Aspekte der UV.
- Die Summe der Quadratsummen von paarweise orthogonalen Kontrasten ergibt selbst wiederum eine Quadratsumme. Die Anzahl der Freiheitsgrade dieser Quadratsumme ist gleich der Anzahl der beteiligten Kontraste. Wie üblich kann auch hier ein F-Test (für die kombinierten Hypothesen) durchgeführt werden.

- Bei p Mittelwerten gibt es maximal p–1 paarweise orthogonale Kontraste, nämlich so viele, wie die UV Freiheitsgrade hat. Deren Quadratummen ergeben insgesamt die QS_{zw} für die Mittelwerte.
- D.h. umgekehrt, dass QS_{zw} in $p-1$ paarweise orthogonale Komponenten QS_{ψ_1}, QS_{ψ_2}, …, $QS_{\psi_{p-1}}$ zerlegt werden kann.

Diese Quadratsummenzerlegung kann das folgende Baumdiagramm veranschaulicht werden:



Orthogonale Kontraste und das Strukturmodell

Wenn wir für eine p-stufige UV $p-1$ orthogonale Kontraste ψ_1, ψ_2, …, ψ_{p-1} definiert haben, so können wir das Strukturmodell mit Hilfe der normierten Koeffizienten und normierten Kontraste auch in folgender Weise schreiben:

$$X_{im} = \mu + c^*_{1i}\psi^*_1 + c^*_{2i}\psi^*_2 + \cdots + c^*_{p-1,i}\psi^*_{p-1} + \varepsilon_{im}$$

Für die μ_i gilt:

$$\mu_i = \mu + c^*_{1i}\psi^*_1 + c^*_{2i}\psi^*_2 + \cdots + c^*_{p-1,i}\psi^*_{p-1}$$

Approximation der Gruppenmittelwerte mit Hilfe der Kontraste

Für die Gruppenmittelwerte erhält man mit Hilfe der Kontrasteffekte die folgende Zerlegung:

$$\bar{X}_{i.} = \bar{X}_{..} + c^*_{1i}\hat{\psi}^*_1 + c^*_{2i}\hat{\psi}^*_2 + \cdots + c^*_{p-1,i}\hat{\psi}^*_{p-1}$$

Verwendet man dabei nicht alle p–1 Kontrasteffekte, sondern nur einen oder einige ausgewählte, so werden die Gruppenmittelwerte nicht exakt reproduziert, sondern lediglich approximiert.

Trendtests

Trendtests

- Die UV sei eine quantitative Variable, auf Intervallskalenniveau gemessen.
- Als eine Form spezieller Hypothesen interessiert dann vielleicht die Frage, welcher Art die Beziehung zwischen der UV und der AV ist.
- Eine mögliche Art des Zusammenhangs wäre ein funktionaler Zusammenhang in der Form einer linearen, quadratischen, kubischen etc. Funktion, kurzum in Form eines Polynoms n-ten Grades.

- Ein Polynom n-tes Grades enthält einen linearen, einen quadratischen, einen kubischen Term usw. bis hin zu einem Term n-ter Ordnung[18].
- Für den Test einer solchen Hypothese ist es üblich, orthogonale polynomiale Kontraste zu verwenden, um damit die lineare, quadratische, kubische usw. Komponente des Zusammenhangs zu testen.
- Die Testgröße ist jeweils genau dieselbe wie bei jedem linearen Kontrast.

Koeffizienten für orthogonale Polynome[19]

Stufen	Trend	Koeffizienten						$\sum c_i^2$
3	linear	-1	0	1				2
	quadratisch	1	-2	1				6
4	linear	-3	-1	1	3			20
	quadratisch	1	-1	-1	1			4
	kubisch	-1	3	-3	1			20
5	linear	-2	-1	0	1	2		10
	quadratisch	2	-1	-2	-1	2		14
	kubisch	-1	2	0	-2	1		10
	quartisch	1	-4	6	-4	1		70
6	linear	-5	-3	-1	1	3	5	70
	quadratisch	5	-1	-4	-4	-1	5	84
	kubisch	-5	7	4	-4	-7	5	180
	quartisch	1	-3	2	2	-3	1	28
	quintisch	1	-5	10	-10	5	-1	252

Kurvenanpassung mit orthogonalen polynomialen Kontrasten

[18] Bei p Stufen kann das Polynom höchstens vom Grade $p-1$ sein.

Die Gruppenmittelwerte können mit Hilfe der polynomialen Trendkomponenten approximiert werden:

$$\hat{\bar{X}}_{i.} = \bar{X}_{..} + c^*_{lin,i}\hat{\psi}^*_{lin} + c^*_{quad,i}\hat{\psi}^*_{quad} + \dots$$

Beispiel

Schlafdeprivation und Tremor[20]

Mittelwertstabelle (4 unabhängige Gruppen à 8 Vpn):

Schlafentzug (Stunden)	12	24	36	48
Tremorstärke	2.75	3.50	6.25	9.00

Außerdem sei $QS_{inn} = 41.0$.

Es soll der lineare Trend getestet werden, d.h. ob zwischen Schlafentzug und Tremorstärke ein linearer Zusammenhang besteht (a priori Hypothese).

Als Kontrastkoeffizienten finden wir:

$c_{lin,1}$	$c_{lin,2}$	$c_{lin,3}$	$c_{lin,4}$
-3	-1	1	3

$$\begin{aligned}
\hat{\psi}_{lin} &= (-3) \times 2.75 + (-1) \times 3.50 + 1 \times 6.25 + 3 \times 9.00 \\
&= 21.5 \\
QS_{lin} &= 8 \times \frac{21.5^2}{20} = 184.9 \\
F &= \frac{184.9}{41/28} = 126.3 \qquad df_1 = 1 \qquad df_2 = 28
\end{aligned}$$

Der kritische F-Wert ist $F_{.95;1,28} = 4.20$.

Wir lehnen H_0 ab.

Lineare Approximation der Mittelwerte:

$$\hat{\bar{X}}_{i.} = \bar{X}_{..} + c^*_{lin,i}\hat{\psi}^*_{lin}$$

Also:

[20]Beispiel nach Kirk

$c^*_{lin,i}$	-0.671	-0.224	0.224	0.671
$\hat{\psi}^*_{lin}$	4.808	4.808	4.808	4.808
$c^*_{lin,i}\hat{\psi}^*_{lin}$	-3.225	-1.075	1.075	3.225
$\bar{X}_{..}$	5.375	5.375	5.375	5.375
$\hat{\bar{X}}_{i.}$	**2.15**	**4.30**	**6.45**	**8.60**
$\bar{X}_{i.}$	2.75	3.50	6.25	9.00

Zweifaktorielle Varianzanalyse

Zweck und Voraussetzungen

- Gegeben sind die Daten eines Versuchsplans mit zwei unabhängigen Variablen (Faktoren).
- Alle Bedingungskombinationen wurden realisiert.
- Es wurden jeweils unabhängige Stichproben verwendet.

Zweck der zweifaktoriellen Varianzanlyse

Test der Wirkung der verschiedenen Bedingungen und Bedingungskombinationen auf die AV.

Voraussetzungen

- Gegeben sind $p \times q$ normalverteilte Populationen mit unbekannten Parametern μ_{ij} und σ_{ij}.
- Alle Populationen haben die gleiche Standardabweichung: $\sigma_{11} = \cdots = \sigma_{ij} = \cdots = \sigma_{pq} = \sigma$.
- Aus jeder Population wurde eine Zufallsstichprobe mit Stichprobenumfang n gezogen[21].

Definitionen, Notation

- Die beiden unabhängigen Variablen seien Faktor A mit p Stufen und Faktor B mit q Stufen.
- Jede der $p \times q$ Bedingungskombinationen sei im Versuchsplan realisiert worden.

[21] Der Einfachheit halber nehmen wir im folgenden wieder an, dass alle Stichprobenumfänge gleich sind.

- Wir betrachten nun folgende Mittelwerte:

μ_{ij}	Mittelwert der Population mit A-Stufe i und B-Stufe j $(i = 1, \ldots, p;\ j = 1, \ldots, q)$
μ_i	Mittelwert aller Populationen mit A-Stufe i $(i = 1, \ldots, p)$
μ_j	Mittelwert aller Populationen mit B-Stufe j $(j = 1, \ldots, q)$
μ	Gesamtmittelwert der $p \times q$ Populationen

Effekte

Dann lassen sich folgende Effekte definieren:

- Haupteffekte des Faktors A: $\alpha_i = \mu_i - \mu$
- Haupteffekte des Faktors B: $\beta_j = \mu_j - \mu$
- Wechselwirkungseffekte: $\alpha\beta_{ij} = (\mu_{ij} - \mu) - \alpha_i - \beta_j$

Wir betrachten im Folgenden alle Effekte als fixe Größen.

Trivialerweise gilt für die Haupteffekte:

$$\sum \alpha_i = 0 \quad \text{und} \quad \sum \beta_j = 0$$

und für die Wechselwirkungseffekte:

$$\sum_i \alpha\beta_{ij} = 0 \quad \text{für alle j}$$

$$\sum_j \alpha\beta_{ij} = 0 \quad \text{für alle i}$$

Modell

Strukturmodell

Der Messwert einer Vp m in der Gruppe mit der Bedingungskombination A-Stufe i und B-Stufe j ist eine Zufallsvariable und setzt sich nun folgendermaßen zusammen:

$$X_{ijm} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijm}$$

In Worten: der Messwert einer Person m in der Gruppe i setzt sich additiv zusammen aus

- dem Gesamtmittelwert,
- dem Haupteffekt der A-Stufe i,
- dem Haupteffekt der B-Stufe j,
- dem Wechselwirkungseffekt der Kombination der A-Stufe i und B-Stufe j
- sowie einer Zufallskomponente, dem "Fehler".

Für den Erwartungswert und die Varianz von X_{ijm} gilt:

$$\begin{aligned} E(X_{ijm}) &= \mu + \alpha_i + \beta_j + \alpha\beta_{ij} \\ Var(X_{ijm}) &= \sigma_\varepsilon^2 \end{aligned}$$

Die eingangs genannten Voraussetzungen der Varianzanalyse könnten also, bezogen auf das Strukturmodell, auch so formuliert werden:

Die ε_{ijm} sind unabhängige, identisch normalverteilte Zufallsvariablen mit Mittelwert 0 und Varianz σ_ε^2.

Hypothesen

Hypothesen

Haupteffekt A

- Nullhypothese:

$$\mu_i = \mu \quad \text{bzw.} \quad \alpha_i = 0 \qquad \text{für alle } i = 1, \ldots, p$$

  Eine äquivalente Formulierung ist:

$$\sum \alpha_i^2 = 0 \quad \text{bzw.} \quad \sigma_\alpha^2 = 0$$

- Alternativhypothese:

  ...

Haupteffekt B

- Nullhypothese:

$$\mu_j = \mu \quad \text{bzw.} \quad \beta_j = 0 \qquad \text{für alle } j = 1, \ldots, q$$

Eine äquivalente Formulierung ist:

$$\sum \beta_j^2 = 0 \quad \text{bzw.} \quad \sigma_\beta^2 = 0$$

- Alternativhypothese:

  ...

Wechselwirkung AxB

- Nullhypothese:

$$\mu_{ij} = \mu + \alpha_i + \beta_j \quad \text{bzw.} \quad \alpha\beta_{ij} = 0 \qquad \text{für alle } i \text{ und } j$$

  Eine äquivalente Formulierung ist:

$$\sum \sum \alpha\beta_{ij}^2 = 0 \quad \text{bzw.} \quad \sigma_{\alpha\beta}^2 = 0$$

- Alternativhypothese:

  ...

Quadratsummenzerlegung

Mittelwerte

Für die Quadratsummenzerlegung müssen zunächst folgende Mittelwerte berechnet werden:

$\bar{X}_{ij\cdot}$	Mittelwert der Stichprobe mit A-Stufe i und B-Stufe j (Zellenmittelwerte, Gruppenmittelwerte)
$\bar{X}_{i\cdot\cdot}$	Mittelwert aller Stichproben mit A-Stufe i
$\bar{X}_{\cdot j\cdot}$	Mittelwert aller Stichproben mit B-Stufe j
$\bar{X}_{\cdots}$	Gesamtstichprobenmittelwert

Quadratsummenzerlegung

- Quadratsumme total

$$QS_{tot} = \sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{m=1}^{n} (X_{ijm} - \bar{X}_{\cdots})^2$$

- Quadratsumme zwischen

$$QS_{zw} = n\sum_{i=1}^{p}\sum_{j=1}^{q}(\bar{X}_{ij.} - \bar{X}_{...})^2$$

- Quadratsumme innerhalb

$$QS_{inn} = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{m=1}^{n}(X_{ijm} - \bar{X}_{ij.})^2$$

- Quadratsumme A

$$QS_A = nq\sum_{i=1}^{p}(\bar{X}_{i..} - \bar{X}_{...})^2$$

- Quadratsumme B

$$QS_B = np\sum_{j=1}^{q}(\bar{X}_{.j.} - \bar{X}_{...})^2$$

- Quadratsumme AxB[22]

$$QS_{AB} = QS_{zw} - QS_A - QS_B$$

Die Quadratsummen der Haupteffekte und Wechselwirkung lassen sich bequem auch aus den Schätzungen der Effekte berechnen:

- Quadratsumme A

$$QS_A = nq\sum_{i=1}^{p}\hat{\alpha_i}^2$$

- Quadratsumme B

$$QS_B = np\sum_{j=1}^{q}\hat{\beta_j}^2$$

- Quadratsumme AxB

$$QS_{AB} = n\sum_{i=1}^{p}\sum_{j=1}^{q}\widehat{\alpha\beta}_{ij}^2$$

Es gilt:

$$QS_{tot} = QS_{zw} + QS_{inn}$$

und

$$QS_{zw} = QS_A + QS_B + QS_{AB}$$

[22]Anstatt AxB wird häufig auch die Kurzbezeichung AB verwendet.

Freiheitsgrade

- Freiheitsgrade total
$$df_{tot} = npq - 1$$
- Freiheitsgrade zwischen
$$df_{zw} = pq - 1$$
- Freiheitsgrade innerhalb
$$df_{inn} = pq(n - 1)$$

- Freiheitsgrade A
$$df_A = p - 1$$
- Freiheitsgrade B
$$df_B = q - 1$$
- Freiheitsgrade AxB
$$df_{AB} = (p - 1)(q - 1)$$

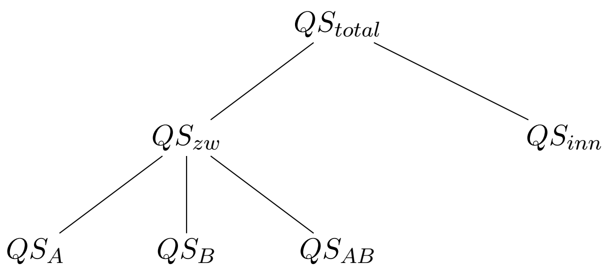
Für die Freiheitsgrade gilt analog zu den Quadratsummen:

$$df_{tot} = df_{zw} + df_{inn}$$

und

$$df_{zw} = df_A + df_B + df_{AB}$$

Diese Quadratsummenzerlegung kann das folgende Baumdiagramm veranschaulicht werden:



Erwartungswerte der mittleren Quadrate

Erwartungswerte der mittleren Quadrate

$$\begin{aligned} E(MQ_A) &= \sigma_\varepsilon^2 + nq\sigma_\alpha^2 \\ E(MQ_B) &= \sigma_\varepsilon^2 + np\sigma_\beta^2 \\ E(MQ_{AB}) &= \sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 \\ E(MQ_{inn}) &= \sigma_\varepsilon^2 \end{aligned}$$

Teststatistiken

Teststatistiken

- Haupteffekt A

  $F_A = \frac{MQ_A}{MQ_{inn}}$ mit $MQ_A = \frac{QS_A}{df_A}$

- Haupteffekt B

  $F_B = \frac{MQ_B}{MQ_{inn}}$ mit $MQ_B = \frac{QS_B}{df_B}$

- Wechselwirkung AB

  $F_{AB} = \frac{MQ_{AB}}{MQ_{inn}}$ mit $MQ_{AB} = \frac{QS_{AB}}{df_{AB}}$

Sind die Vorausetzungen erfüllt, so gilt jeweils unter H_0:

Die Teststatistiken $F_A/F_B/F_{AB}$ sind jeweils F-verteilt mit $df_1 = df_A/df_B/df_{AB}$ und $df_2 = df_{inn}$ Freiheitsgraden.

Rechnerische Durchführung

Rechnerische Durchführung

Für die rechnerische Durchführung der Varianzanalyse müssen folgende Größen berechnet werden:

- Quadratsummen QS_A, QS_B, QS_{AB}, QS_{inn}
- Freiheitsgrade df_A, df_B, df_{AB}, df_{inn}
- Mittlere Quadrate MQ_A, MQ_B, MQ_{AB}, MQ_{inn}
- F-Werte F_A, F_B, F_{AB}

Der F-Wert ist jeweils mit dem kritischen F-Wert zu vergleichen. Übersteigt er diesen, so ist H_0 zu verwerfen.

Die Ergebnisse werden normalerweise in einer Tabelle zusammengefasst:

Quelle der Variation	**QS**	**df**	**MQ**	**F**
Zwischen				
A				
B				
AxB				
Innerhalb				—
Total			—	—

Interaktionseffekte

Interaktionseffekte

- Lassen sich die Zellenmittelwerte nicht nur durch die Haupteffekte reproduzieren, so sind Wechselwirkungseffekte (Interaktionseffekte) im Spiel.
- Wechselwirkungen sind also eine Maß für die Nichtadditivität der Haupteffekte.
- Existenz von Interaktionseffekten hängt in gewisser Weise auch von der Mess-Skala ab (z.B. bei einer logarithmierten Skala kann eine Interaktion verschwinden, die bei der unlogarithmierten Skala vorhanden ist).
- Die Verwendung der Mess-Skala hängt normalerweise vom Experimentator ab. Falls die Mess-Skala frei wählbar ist, sollte jene Skala verwendet werden, die zu einem einfacheren Modell führt, d.h. zu einem solchen ohne Wechselwirkungen.

Interaktionstabellen

- Die Ergebnisse einer zweifaktoriellen Varianzanalyse werden häufig in Form eine AxB-Interaktionstabelle dargestellt.

- Dies ist eine zweidimensionale Tabelle der Zellenmittelwerte, erweitert um die Randmittelwerte und den Gesamtmittelwert.
- Auf Basis einer solchen Tabelle lässt sich bequem eine Übersichtstabelle mit den Haupt- und Wechselwirkungseffekten erstellen.
- Ein vollständiges Fehlen einer Wechselwirkung wäre daran zu erkennen, dass sich die Zellenmittelwerte von jeweils zwei Zeilen (bzw. Spalten) um einen konstanten Wert unterscheiden (d.h. die Differenzen erster Ordnung sind konstant).

Interaktionsdiagramme

- Die verschiedenen varianzanalytischen Effekte lassen sich auch in einem Interaktionsdiagramm veranschaulichen.
- Hierbei werden die Gruppenmittelwerte eines zweifaktoriellen Versuchsplans als überlagertes Liniendiagramm in einem zweidimensionalen Koordinatensystem dargestellt.
- Die Stufen des einen der beiden Faktoren werden dabei auf der x-Achse abgetragen.
- Jede Stufe des anderen Faktors wird durch einen getrennten Linienzug repräsentiert.
- Ein vollständiges Fehlen einer Wechselwirkung wäre daran zu erkennen, dass die einzelnen Linienzüge (in Y-Richtung) parallel sind.

Klassifikation von Interaktionseffekten

Bringt man getrennt für die Stufen des einen Faktors die bedingten Mittelwerte des anderen Faktors in eine Rangreihe und umgekehrt, so kann man durch Vergleich der Rangreihen verschiedene Formen der Interaktion unterscheiden:

- Ordinale Interaktion: die Rangreihen stimmen überein (die einzelnen Linienzüge im Interaktionsdiagramm kreuzen sich nicht).
- Disordinale Interaktion: die Rangreihen stimmen in beiden Fällen nicht überein (die einzelnen Linienzüge im Interaktionsdiagramm kreuzen sich immer).
- Hybride Interaktion: die Rangreihen stimmen in einem Fall überein, im anderen nicht (in einem Fall kreuzen sich die einzelnen Linienzüge, im anderen Fall nicht).

□

Wenn bei einer Varianzanalyse Wechselwirkungseffekte signifikant werden, ist bei der Interpretation Vorsicht geboten:

- Beim ordinalen Interaktionstyp ist die Interaktion möglicherweise ein Skalenartefakt.
- Beim hybriden Interaktionstyp sollte nur jener Haupteffekt interpretiert werden (falls er signifikant ist), bei dem die Rangreihen der bedingten Mittelwerte übereinstimmen.

Die Prüfung spezieller Hypothesen

Prüfung spezieller Hypothesen

Haupteffekte

- Für jeden der beiden Faktoren lassen sich spezielle Hypothesen mit Hilfe linearer Kontraste prüfen.
- Das Vorgehen incl. F-Test ist analog zu jenem bei der einfachen Varianzanalyse.
- Zu berücksichtigen ist dabei, dass für die Mittelwerte jeweils nq bzw. np Einzelwerte gemittelt wurden.
- Bei einem vollständigen faktoriellen Design mit gleichen Stichprobenumfängen sind die linearen Kontraste des einen Faktors orthogonal zu jenen des anderen Faktors.

Wechselwirkungen

- Auch für die Wechselwirkung lassen sich spezielle Hypothesen prüfen.
- Hat man für die Haupteffekte einen Satz von orthogonalen Kontrasten aufgestellt, so lassen sich daraus Kontrastkoeffizienten für die Wechselwirkung ableiten: man nimmt jeweils einen Kontrast des einen Faktors und einen Kontrast des anderen Faktors und multiplizert paarweise die Kontrastkoeffizienten. Diese ergeben einen Kontrast für die Wechselwirkung.
- Für die Wechselwirkung gibt es wiederum maximal so viele orthogonale Kontraste, wie sie Freiheitsgrade hat.
- Getestet wird jeweils gegen MQ_{inn}.

Bedingte Haupteffekte und bedingte Vergleiche

Bedingte Haupteffekte

Um die Wechselwirkung bei einem zweifaktoriellen Versuchsplan näher zu untersuchen, können auch bedingte Haupteffekte (auch „simple main effects“ genannt) berechnet werden.

Dabei betrachtet man die Haupteffekte eines der Faktoren für eine bestimmte Stufe des anderen Faktors, d.h. beispielsweise die α_i für die Stufe j des Faktors B oder die β_j für die Stufe i des Faktors A:

$$\begin{aligned}
\alpha_{i|b_j} &= \mu_{ij} - \mu_j \\
\beta_{j|a_i} &= \mu_{ij} - \mu_i
\end{aligned}$$

Für die bedingten Haupteffekte lassen sich auch Quadratsummen berechnen, d.h. die Quadratsumme von A für Stufe j des Faktors B bzw. die Quadratsumme von B für Stufe i des Faktors A:

$$\begin{aligned}
QS_{A|b_j} &= n\sum_{i=1}^{p}(\bar{X}_{ij\cdot} - \bar{X}_{\cdot j\cdot})^2 \\
QS_{B|a_i} &= n\sum_{j=1}^{q}(\bar{X}_{ij\cdot} - \bar{X}_{i\cdot\cdot})^2
\end{aligned}$$

Die Freiheitsgrade sind wie gewohnt $p-1$ bzw. $q-1$.

Für die Summe der Quadratsummen der bedingten Haupteffekte eines Faktors gilt übrigens:

$$\begin{aligned}
\sum_{j=1}^{q} QS_{A|b_j} &= QS_{AB} + QS_A \\
\sum_{i=1}^{p} QS_{B|a_i} &= QS_{AB} + QS_B
\end{aligned}$$

Für deren Freiheitsgrade gilt die analoge Beziehung:

$$\begin{aligned}
df_{A|b_{j=1,\ldots,q}} &= df_{AB} + df_A \\
df_{B|a_{i=1,\ldots,p}} &= df_{AB} + df_B
\end{aligned}$$

Daraus ergibt sich als Prüfgröße:

$$F = \frac{(QS_A + QS_{AB})/(df_A + df_{AB})}{MQ_{inn}}$$

bzw.

$$F = \frac{(QS_B + QS_{AB})/(df_B + df_{AB})}{MQ_{inn}}$$

mit $df_1 = df_A + df_{AB}$ bzw. $df_1 = df_B + df_{AB}$

und $df_2 = df_{inn}$.

Beispiel für ein 3×4-Design

μ_{11}	μ_{12}	μ_{13}	μ_{14}	$\mu_{1\cdot}$
μ_{21}	μ_{22}	μ_{23}	μ_{24}	$\mu_{2\cdot}$
μ_{31}	μ_{32}	μ_{33}	μ_{34}	$\mu_{3\cdot}$
$\mu_{\cdot 1}$	$\mu_{\cdot 2}$	$\mu_{\cdot 3}$	$\mu_{\cdot 4}$	$\mu_{\cdot\cdot}$

$\bar{X}_{11\cdot}$	$\bar{X}_{12\cdot}$	$\bar{X}_{13\cdot}$	$\bar{X}_{14\cdot}$	$\bar{X}_{1\cdot\cdot}$
$\bar{X}_{21\cdot}$	$\bar{X}_{22\cdot}$	$\bar{X}_{23\cdot}$	$\bar{X}_{24\cdot}$	$\bar{X}_{2\cdot\cdot}$
$\bar{X}_{31\cdot}$	$\bar{X}_{32\cdot}$	$\bar{X}_{33\cdot}$	$\bar{X}_{34\cdot}$	$\bar{X}_{3\cdot\cdot}$
$\bar{X}_{\cdot 1\cdot}$	$\bar{X}_{\cdot 2\cdot}$	$\bar{X}_{\cdot 3\cdot}$	$\bar{X}_{\cdot 4\cdot}$	$\bar{X}_{\cdot\cdot\cdot}$

Bedingte Einzelvergleiche

Basierend auf bedingten Haupteffekten (d.h. basierend auf einer einzelnen Zeile oder Spalte der Interaktionstabelle) können auch bedingte Einzelvergleiche aufgestellt und geprüft werden.

Das Vorgehen ist dabei völlig analog zu jenem bei linearen Kontrasten für Haupteffekte, d.h. die Kontrastkoeffizienten, der Wert des Kontrasts, seine Quadratsumme und der F-Test sind entsprechend zu bestimmen.

Wenn wir beispielsweise mit $\psi_{A_k|b_j}$ einen linearen Kontrast für Faktor A für die Stufe j des Faktors B bezeichnen, so ergibt sich als Quadratsumme:

$$QS_{\psi_{A_k|b_j}} = \frac{n\hat{\psi}^2_{A_k|b_j}}{\sum c^2_{i;A_k|b_j}}$$

(Für weitergehende Informationen zu solchen Einzelvergleichen sei auf die Literatur, wie z.B. Bortz oder Kirk, verwiesen).

Zweifaktorielle Varianzanalyse

Zweck und Voraussetzungen

- Gegeben sind die Daten eines Versuchsplans mit zwei unabhängigen Variablen (Faktoren).
- Alle Bedingungskombinationen wurden realisiert.
- Es wurden jeweils unabhängige Stichproben verwendet.

Zweck der zweifaktoriellen Varianzanlyse

Test der Wirkung der verschiedenen Bedingungen und Bedingungskombinationen auf die AV.

Voraussetzungen

- Gegeben sind $p \times q$ normalverteilte Populationen mit unbekannten Parametern μ_{ij} und σ_{ij}.
- Alle Populationen haben die gleiche Standardabweichung: $\sigma_{11} = \cdots = \sigma_{ij} = \cdots = \sigma_{pq} = \sigma$.
- Aus jeder Population wurde eine Zufallsstichprobe mit Stichprobenumfang n gezogen[23].

Definitionen, Notation

- Die beiden unabhängigen Variablen seien Faktor A mit p Stufen und Faktor B mit q Stufen.
- Jede der $p \times q$ Bedingungskombinationen sei im Versuchsplan realisiert worden.
- Wir betrachten nun folgende Mittelwerte:

μ_{ij}	Mittelwert der Population mit A-Stufe i und B-Stufe j $(i = 1, \ldots, p;\ j = 1, \ldots, q)$
μ_i	Mittelwert aller Populationen mit A-Stufe i $(i = 1, \ldots, p)$
μ_j	Mittelwert aller Populationen mit B-Stufe j $(j = 1, \ldots, q)$
μ	Gesamtmittelwert der $p \times q$ Populationen

[23] Der Einfachheit halber nehmen wir im folgenden wieder an, dass alle Stichprobenumfänge gleich sind.

Effekte

Dann lassen sich folgende Effekte definieren:

- Haupteffekte des Faktors A: $\alpha_i = \mu_i - \mu$
- Haupteffekte des Faktors B: $\beta_j = \mu_j - \mu$
- Wechselwirkungseffekte: $\alpha\beta_{ij} = (\mu_{ij} - \mu) - \alpha_i - \beta_j$

Wir betrachten im Folgenden alle Effekte als fixe Größen.

Trivialerweise gilt für die Haupteffekte:

$$\sum \alpha_i = 0 \quad \text{und} \quad \sum \beta_j = 0$$

und für die Wechselwirkungseffekte:

$$\sum_i \alpha\beta_{ij} = 0 \quad \text{für alle j}$$

$$\sum_j \alpha\beta_{ij} = 0 \quad \text{für alle i}$$

Modell

Strukturmodell

Der Messwert einer Vp m in der Gruppe mit der Bedingungskombination A-Stufe i und B-Stufe j ist eine Zufallsvariable und setzt sich nun folgendermaßen zusammen:

$$X_{ijm} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijm}$$

In Worten: der Messwert einer Person m in der Gruppe i setzt sich additiv zusammen aus

- dem Gesamtmittelwert,
- dem Haupteffekt der A-Stufe i,
- dem Haupteffekt der B-Stufe j,
- dem Wechselwirkungseffekt der Kombination der A-Stufe i und B-Stufe j
- sowie einer Zufallskomponente, dem “Fehler”.

Für den Erwartungswert und die Varianz von X_{ijm} gilt:

$$\begin{aligned} E(X_{ijm}) &= \mu + \alpha_i + \beta_j + \alpha\beta_{ij} \\ Var(X_{ijm}) &= \sigma_\varepsilon^2 \end{aligned}$$

Die eingangs genannten Voraussetzungen der Varianzanalyse könnten also, bezogen auf das Strukturmodell, auch so formuliert werden:

Die ε_{ijm} sind unabhängige, identisch normalverteilte Zufallsvariablen mit Mittelwert 0 und Varianz σ_ε^2.

Hypothesen

Hypothesen

Haupteffekt A

- Nullhypothese:

$$\mu_i = \mu \quad \text{bzw.} \quad \alpha_i = 0 \qquad \text{für alle } i = 1, \ldots, p$$

Eine äquivalente Formulierung ist:

$$\sum \alpha_i^2 = 0 \quad \text{bzw.} \quad \sigma_\alpha^2 = 0$$

- Alternativhypothese:

  ...

Haupteffekt B

- Nullhypothese:

$$\mu_j = \mu \quad \text{bzw.} \quad \beta_j = 0 \qquad \text{für alle } j = 1, \ldots, q$$

Eine äquivalente Formulierung ist:

$$\sum \beta_j^2 = 0 \quad \text{bzw.} \quad \sigma_\beta^2 = 0$$

- Alternativhypothese:

  ...

Wechselwirkung AxB

- Nullhypothese:

$$\mu_{ij} = \mu + \alpha_i + \beta_j \quad \text{bzw.} \quad \alpha\beta_{ij} = 0 \qquad \text{für alle } i \text{ und } j$$

  Eine äquivalente Formulierung ist:

$$\sum\sum \alpha\beta_{ij}^2 = 0 \quad \text{bzw.} \quad \sigma_{\alpha\beta}^2 = 0$$

- Alternativhypothese:

  ...

Quadratsummenzerlegung

Mittelwerte

Für die Quadratsummenzerlegung müssen zunächst folgende Mittelwerte berechnet werden:

$\bar{X}_{ij\cdot}$	Mittelwert der Stichprobe mit A-Stufe i und B-Stufe j (Zellenmittelwerte, Gruppenmittelwerte)
$\bar{X}_{i\cdot\cdot}$	Mittelwert aller Stichproben mit A-Stufe i
$\bar{X}_{\cdot j\cdot}$	Mittelwert aller Stichproben mit B-Stufe j
$\bar{X}_{\cdots}$	Gesamtstichprobenmittelwert

Quadratsummenzerlegung

- Quadratsumme total

$$QS_{tot} = \sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{m=1}^{n} (X_{ijm} - \bar{X}_{\cdots})^2$$

- Quadratsumme zwischen

$$QS_{zw} = n \sum_{i=1}^{p} \sum_{j=1}^{q} (\bar{X}_{ij\cdot} - \bar{X}_{\cdots})^2$$

- Quadratsumme innerhalb

$$QS_{inn} = \sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{m=1}^{n} (X_{ijm} - \bar{X}_{ij.})^2$$

- Quadratsumme A

$$QS_A = nq \sum_{i=1}^{p} (\bar{X}_{i..} - \bar{X}_{...})^2$$

- Quadratsumme B

$$QS_B = np \sum_{j=1}^{q} (\bar{X}_{.j.} - \bar{X}_{...})^2$$

- Quadratsumme AxB[24]

$$QS_{AB} = QS_{zw} - QS_A - QS_B$$

Die Quadratsummen der Haupteffekte und Wechselwirkung lassen sich bequem auch aus den Schätzungen der Effekte berechnen:

- Quadratsumme A

$$QS_A = nq \sum_{i=1}^{p} \hat{\alpha_i}^2$$

- Quadratsumme B

$$QS_B = np \sum_{j=1}^{q} \hat{\beta_j}^2$$

- Quadratsumme AxB

$$QS_{AB} = n \sum_{i=1}^{p} \sum_{j=1}^{q} \widehat{\alpha\beta}_{ij}^2$$

Es gilt:

$$QS_{tot} = QS_{zw} + QS_{inn}$$

und

$$QS_{zw} = QS_A + QS_B + QS_{AB}$$

Freiheitsgrade

[24] Anstatt AxB wird häufig auch die Kurzbezeichung AB verwendet.

- Freiheitsgrade total

$$df_{tot} = npq - 1$$

- Freiheitsgrade zwischen

$$df_{zw} = pq - 1$$

- Freiheitsgrade innerhalb

$$df_{inn} = pq(n - 1)$$

- Freiheitsgrade A

$$df_A = p - 1$$

- Freiheitsgrade B

$$df_B = q - 1$$

- Freiheitsgrade AxB

$$df_{AB} = (p - 1)(q - 1)$$

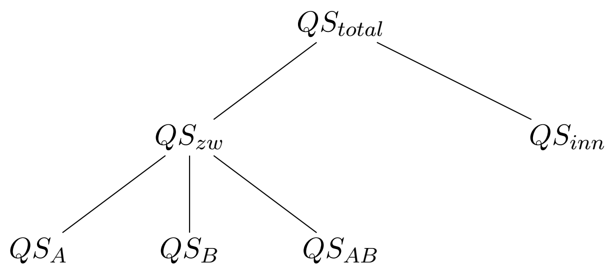
Für die Freiheitsgrade gilt analog zu den Quadratsummen:

$$df_{tot} = df_{zw} + df_{inn}$$

und

$$df_{zw} = df_A + df_B + df_{AB}$$

Diese Quadratsummenzerlegung kann das folgende Baumdiagramm veranschaulicht werden:



Erwartungswerte der mittleren Quadrate

Erwartungswerte der mittleren Quadrate

$$\begin{aligned} E(MQ_A) &= \sigma_\varepsilon^2 + nq\sigma_\alpha^2 \\ E(MQ_B) &= \sigma_\varepsilon^2 + np\sigma_\beta^2 \\ E(MQ_{AB}) &= \sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 \\ E(MQ_{inn}) &= \sigma_\varepsilon^2 \end{aligned}$$

Teststatistiken

Teststatistiken

- Haupteffekt A

  $F_A = \frac{MQ_A}{MQ_{inn}}$ mit $MQ_A = \frac{QS_A}{df_A}$

- Haupteffekt B

  $F_B = \frac{MQ_B}{MQ_{inn}}$ mit $MQ_B = \frac{QS_B}{df_B}$

- Wechselwirkung AB

  $F_{AB} = \frac{MQ_{AB}}{MQ_{inn}}$ mit $MQ_{AB} = \frac{QS_{AB}}{df_{AB}}$

Sind die Vorausetzungen erfüllt, so gilt jeweils unter H_0:

Die Teststatistiken $F_A/F_B/F_{AB}$ sind jeweils F-verteilt mit $df_1 = df_A/df_B/df_{AB}$ und $df_2 = df_{inn}$ Freiheitsgraden.

Rechnerische Durchführung

Rechnerische Durchführung

Für die rechnerische Durchführung der Varianzanalyse müssen folgende Größen berechnet werden:

- Quadratsummen QS_A, QS_B, QS_{AB}, QS_{inn}
- Freiheitsgrade df_A, df_B, df_{AB}, df_{inn}
- Mittlere Quadrate MQ_A, MQ_B, MQ_{AB}, MQ_{inn}
- F-Werte F_A, F_B, F_{AB}

Der F-Wert ist jeweils mit dem kritischen F-Wert zu vergleichen. Übersteigt er diesen, so ist H_0 zu verwerfen.

Die Ergebnisse werden normalerweise in einer Tabelle zusammengefasst:

Quelle der Variation	**QS**	**df**	**MQ**	**F**
Zwischen				
A				
B				
AxB				
Innerhalb				—
Total			—	—

Interaktionseffekte

Interaktionseffekte

- Lassen sich die Zellenmittelwerte nicht nur durch die Haupteffekte reproduzieren, so sind Wechselwirkungseffekte (Interaktionseffekte) im Spiel.
- Wechselwirkungen sind also eine Maß für die Nichtadditivität der Haupteffekte.
- Existenz von Interaktionseffekten hängt in gewisser Weise auch von der Mess-Skala ab (z.B. bei einer logarithmierten Skala kann eine Interaktion verschwinden, die bei der unlogarithmierten Skala vorhanden ist).
- Die Verwendung der Mess-Skala hängt normalerweise vom Experimentator ab. Falls die Mess-Skala frei wählbar ist, sollte jene Skala verwendet werden, die zu einem einfacheren Modell führt, d.h. zu einem solchen ohne Wechselwirkungen.

Interaktionstabellen

- Die Ergebnisse einer zweifaktoriellen Varianzanalyse werden häufig in Form eine AxB-Interaktionstabelle dargestellt.

- Dies ist eine zweidimensionale Tabelle der Zellenmittelwerte, erweitert um die Randmittelwerte und den Gesamtmittelwert.
- Auf Basis einer solchen Tabelle lässt sich bequem eine Übersichtstabelle mit den Haupt- und Wechselwirkungseffekten erstellen.
- Ein vollständiges Fehlen einer Wechselwirkung wäre daran zu erkennen, dass sich die Zellenmittelwerte von jeweils zwei Zeilen (bzw. Spalten) um einen konstanten Wert unterscheiden (d.h. die Differenzen erster Ordnung sind konstant).

Interaktionsdiagramme

- Die verschiedenen varianzanalytischen Effekte lassen sich auch in einem Interaktionsdiagramm veranschaulichen.
- Hierbei werden die Gruppenmittelwerte eines zweifaktoriellen Versuchsplans als überlagertes Liniendiagramm in einem zweidimensionalen Koordinatensystem dargestellt.
- Die Stufen des einen der beiden Faktoren werden dabei auf der x-Achse abgetragen.
- Jede Stufe des anderen Faktors wird durch einen getrennten Linienzug repräsentiert.
- Ein vollständiges Fehlen einer Wechselwirkung wäre daran zu erkennen, dass die einzelnen Linienzüge (in Y-Richtung) parallel sind.

Klassifikation von Interaktionseffekten

Bringt man getrennt für die Stufen des einen Faktors die bedingten Mittelwerte des anderen Faktors in eine Rangreihe und umgekehrt, so kann man durch Vergleich der Rangreihen verschiedene Formen der Interaktion unterscheiden:

- Ordinale Interaktion: die Rangreihen stimmen überein (die einzelnen Linienzüge im Interaktionsdiagramm kreuzen sich nicht).
- Disordinale Interaktion: die Rangreihen stimmen in beiden Fällen nicht überein (die einzelnen Linienzüge im Interaktionsdiagramm kreuzen sich immer).
- Hybride Interaktion: die Rangreihen stimmen in einem Fall überein, im anderen nicht (in einem Fall kreuzen sich die einzelnen Linienzüge, im anderen Fall nicht).

Wenn bei einer Varianzanalyse Wechselwirkungseffekte signifikant werden, ist bei der Interpretation Vorsicht geboten:

- Beim ordinalen Interaktionstyp ist die Interaktion möglicherweise ein Skalenartefakt.
- Beim hybriden Interaktionstyp sollte nur jener Haupteffekt interpretiert werden (falls er signifikant ist), bei dem die Rangreihen der bedingten Mittelwerte übereinstimmen.

Die Prüfung spezieller Hypothesen

Prüfung spezieller Hypothesen

Haupteffekte

- Für jeden der beiden Faktoren lassen sich spezielle Hypothesen mit Hilfe linearer Kontraste prüfen.
- Das Vorgehen incl. F-Test ist analog zu jenem bei der einfachen Varianzanalyse.
- Zu berücksichtigen ist dabei, dass für die Mittelwerte jeweils nq bzw. np Einzelwerte gemittelt wurden.
- Bei einem vollständigen faktoriellen Design mit gleichen Stichprobenumfängen sind die linearen Kontraste des einen Faktors orthogonal zu jenen des anderen Faktors.

Wechselwirkungen

- Auch für die Wechselwirkung lassen sich spezielle Hypothesen prüfen.
- Hat man für die Haupteffekte einen Satz von orthogonalen Kontrasten aufgestellt, so lassen sich daraus Kontrastkoeffizienten für die Wechselwirkung ableiten: man nimmt jeweils einen Kontrast des einen Faktors und einen Kontrast des anderen Faktors und multiplizert paarweise die Kontrastkoeffizienten. Diese ergeben einen Kontrast für die Wechselwirkung.
- Für die Wechselwirkung gibt es wiederum maximal so viele orthogonale Kontraste, wie sie Freiheitsgrade hat.
- Getestet wird jeweils gegen MQ_{inn}.

Bedingte Haupteffekte und bedingte Vergleiche

Bedingte Haupteffekte

Um die Wechselwirkung bei einem zweifaktoriellen Versuchsplan näher zu untersuchen, können auch bedingte Haupteffekte (auch „simple main effects“ genannt) berechnet werden.

Dabei betrachtet man die Haupteffekte eines der Faktoren für eine bestimmte Stufe des anderen Faktors, d.h. beispielsweise die α_i für die Stufe j des Faktors B oder die β_j für die Stufe i des Faktors A:

$$\begin{aligned}\alpha_{i|b_j} &= \mu_{ij} - \mu_j \\ \beta_{j|a_i} &= \mu_{ij} - \mu_i\end{aligned}$$

Für die bedingten Haupteffekte lassen sich auch Quadratsummen berechnen, d.h. die Quadratsumme von A für Stufe j des Faktors B bzw. die Quadratsumme von B für Stufe i des Faktors A:

$$\begin{aligned}QS_{A|b_j} &= n\sum_{i=1}^{p}(\bar{X}_{ij\cdot} - \bar{X}_{\cdot j\cdot})^2 \\ QS_{B|a_i} &= n\sum_{j=1}^{q}(\bar{X}_{ij\cdot} - \bar{X}_{i\cdot\cdot})^2\end{aligned}$$

Die Freiheitsgrade sind wie gewohnt $p-1$ bzw. $q-1$.

Für die Summe der Quadratsummen der bedingten Haupteffekte eines Faktors gilt übrigens:

$$\begin{aligned}\sum_{j=1}^{q} QS_{A|b_j} &= QS_{AB} + QS_A \\ \sum_{i=1}^{p} QS_{B|a_i} &= QS_{AB} + QS_B\end{aligned}$$

Für deren Freiheitsgrade gilt die analoge Beziehung:

$$\begin{aligned}df_{A|b_{j=1,\dots,q}} &= df_{AB} + df_A \\ df_{B|a_{i=1,\dots,p}} &= df_{AB} + df_B\end{aligned}$$

Daraus ergibt sich als Prüfgröße:

$$F = \frac{(QS_A + QS_{AB})/(df_A + df_{AB})}{MQ_{inn}}$$

bzw.

$$F = \frac{(QS_B + QS_{AB})/(df_B + df_{AB})}{MQ_{inn}}$$

mit $df_1 = df_A + df_{AB}$ bzw. $df_1 = df_B + df_{AB}$

und $df_2 = df_{inn}$.

Beispiel für ein 3×4-Design

μ_{11}	μ_{12}	μ_{13}	μ_{14}	$\mu_{1\cdot}$
μ_{21}	μ_{22}	μ_{23}	μ_{24}	$\mu_{2\cdot}$
μ_{31}	μ_{32}	μ_{33}	μ_{34}	$\mu_{3\cdot}$
$\mu_{\cdot 1}$	$\mu_{\cdot 2}$	$\mu_{\cdot 3}$	$\mu_{\cdot 4}$	$\mu_{\cdot\cdot}$

$\bar{X}_{11\cdot}$	$\bar{X}_{12\cdot}$	$\bar{X}_{13\cdot}$	$\bar{X}_{14\cdot}$	$\bar{X}_{1\cdot\cdot}$
$\bar{X}_{21\cdot}$	$\bar{X}_{22\cdot}$	$\bar{X}_{23\cdot}$	$\bar{X}_{24\cdot}$	$\bar{X}_{2\cdot\cdot}$
$\bar{X}_{31\cdot}$	$\bar{X}_{32\cdot}$	$\bar{X}_{33\cdot}$	$\bar{X}_{34\cdot}$	$\bar{X}_{3\cdot\cdot}$
$\bar{X}_{\cdot 1\cdot}$	$\bar{X}_{\cdot 2\cdot}$	$\bar{X}_{\cdot 3\cdot}$	$\bar{X}_{\cdot 4\cdot}$	$\bar{X}_{\cdot\cdot\cdot}$

Bedingte Einzelvergleiche

Basierend auf bedingten Haupteffekten (d.h. basierend auf einer einzelnen Zeile oder Spalte der Interaktionstabelle) können auch bedingte Einzelvergleiche aufgestellt und geprüft werden.

Das Vorgehen ist dabei völlig analog zu jenem bei linearen Kontrasten für Haupteffekte, d.h. die Kontrastkoeffizienten, der Wert des Kontrasts, seine Quadratsumme und der F-Test sind entsprechend zu bestimmen.

Wenn wir beispielsweise mit $\psi_{A_k|b_j}$ einen linearen Kontrast für Faktor A für die Stufe j des Faktors B bezeichnen, so ergibt sich als Quadratsumme:

$$QS_{\psi_{A_k|b_j}} = \frac{n\hat{\psi}^2_{A_k|b_j}}{\sum c^2_{i;A_k|b_j}}$$

(Für weitergehende Informationen zu solchen Einzelvergleichen sei auf die Literatur, wie z.B. Bortz oder Kirk, verwiesen).

Dreifaktorielle Varianzanalyse

Zweck und Voraussetzungen

- Gegeben sind die Daten eines Versuchsplans mit drei unabhängigen Variablen (Faktoren).
- Alle Bedingungskombinationen wurden realisiert.
- Es wurden jeweils unabhängige Stichproben verwendet.

Zweck der dreifaktoriellen Varianzanlyse

Test der Wirkung der verschiedenen Bedingungen und Bedingungskombinationen auf die AV.

Voraussetzungen

- Gegeben sind $p \times q \times r$ normalverteilte Populationen mit unbekannten Parametern μ_{ijk} und σ_{ijk}.
- Alle Populationen haben die gleiche Standardabweichung: $\sigma_{111} = \cdots = \sigma_{ijk} = \cdots = \sigma_{pqr} = \sigma$.
- Aus jeder Population wurde eine Zufallsstichprobe mit Stichprobenumfang n gezogen[25].

Definitionen, Notation

- Die unabhängigen Variablen seien Faktor A mit p Stufen, Faktor B mit q Stufen und Faktor C mit r Stufen.
- Jede der $p \times q \times r$ Bedingungskombinationen sei im Versuchsplan realisiert worden.

Tabelle der Mittelwerte:

[25] Der Einfachheit halber nehmen wir im folgenden wieder an, dass alle Stichprobenumfänge gleich sind.

μ_{ijk}	Mittelwert der Population mit A-Stufe i, B-Stufe j und C-Stufe k ($i = 1, \ldots, p; j = 1, \ldots, q; k = 1, \ldots, r$)
μ_{ij}	Mittelwert aller Populationen mit A-Stufe i und B-Stufe j ($i = 1, \ldots, p; j = 1, \ldots, q$)
μ_{ik}	Mittelwert aller Populationen mit A-Stufe i und C-Stufe k ($i = 1, \ldots, p; k = 1, \ldots, r$)
μ_{jk}	Mittelwert aller Populationen mit B-Stufe j und C-Stufe k ($j = 1, \ldots, q; k = 1, \ldots, r$)
μ_i	Mittelwert aller Populationen mit A-Stufe i ($i = 1, \ldots, p$)
μ_j	Mittelwert aller Populationen mit B-Stufe j ($j = 1, \ldots, q$)
μ_k	Mittelwert aller Populationen mit C-Stufe k ($k = 1, \ldots, r$)
μ	Gesamtmittelwert der $p \times q \times r$ Populationen

Effekte

Haupteffekte

$$\begin{aligned} \alpha_i &= \mu_i - \mu \\ \beta_j &= \mu_j - \mu \\ \gamma_k &= \mu_k - \mu \end{aligned}$$

Zweifache Wechselwirkungseffekte (Wechselwirkungen erster Ordung)

$$\begin{aligned} \alpha\beta_{ij} &= \mu_{ij} - \mu - \alpha_i - \beta_j \\ \alpha\gamma_{ik} &= \mu_{ik} - \mu - \alpha_i - \gamma_k \\ \beta\gamma_{jk} &= \mu_{jk} - \mu - \beta_j - \gamma_k \end{aligned}$$

Dreifache Wechselwirkungseffekte (Wechselwirkung zweiter Ordung)

$$\alpha\beta\gamma_{ijk} = \mu_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k - \alpha\beta_{ij} - \alpha\gamma_{ik} - \beta\gamma_{jk}$$

Alle Effekte sollen als fixe Größen betrachtet werden.

Trivialerweise gilt für die Haupteffekte:

$$\sum \alpha_i = 0 \quad \text{und} \quad \sum \beta_j = 0$$

und für die Wechselwirkungseffekte:

$$\sum_i \alpha\beta_{ij} = 0 \quad \text{für alle j}$$

$$\sum_j \alpha\beta_{ij} = 0 \quad \text{für alle i}$$

Modell

Strukturmodell

Der Messwert einer Vp m in der Gruppe mit der Bedingungskombination A-Stufe i, B-Stufe j und C-Stufe k ist eine Zufallsvariable und setzt sich nun folgendermaßen zusammen:

$$X_{ijkm} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \alpha\gamma_{ik} + \beta\gamma_{jk} + \alpha\beta\gamma_{ijk} + \varepsilon_{ijkm}$$

In Worten: der Messwert einer Person m in der Gruppe i setzt sich additiv zusammen aus ...

Für den Erwartungswert und die Varianz von ε_{ijkm} gilt:

$$\begin{aligned} E(\varepsilon_{ijkm}) &= 0 \\ Var(\varepsilon_{ijkm}) &= \sigma_\varepsilon^2 \end{aligned}$$

Die eingangs genannten Voraussetzungen der Varianzanalyse könnten also, bezogen auf das Strukturmodell, auch so formuliert werden:

Die ε_{ijkm} sind unabhängige, identisch normalverteilte Zufallsvariablen mit Mittelwert 0 und Varianz σ_ε^2.

Hypothesen

Hypothesen

Haupteffekte

Nullhypothesen für die Haupteffekte der drei Faktoren:

$$
\begin{aligned}
\alpha_i &= 0 \qquad (i = 1, \ldots, p) \\
\beta_j &= 0 \qquad (j = 1, \ldots, q) \\
\gamma_k &= 0 \qquad (k = 1, \ldots, r)
\end{aligned}
$$

Oder anders ausgedrückt:

$$
\begin{aligned}
\mu_i &= \mu \qquad (i = 1, \ldots, p) \\
\mu_j &= \mu \qquad (j = 1, \ldots, q) \\
\mu_k &= \mu \qquad (k = 1, \ldots, r)
\end{aligned}
$$

Zweifache Wechselwirkungen

Nullhypothesen für die zweifachen Wechselwirkungen:

$$
\begin{aligned}
\alpha\beta_{ij} &= 0 \qquad (i = 1, \ldots, p; j = 1, \ldots, q) \\
\alpha\gamma_{ik} &= 0 \qquad (i = 1, \ldots, p; k = 1, \ldots, r) \\
\beta\gamma_{jk} &= 0 \qquad (j = 1, \ldots, q; k = 1, \ldots, r)
\end{aligned}
$$

Oder anders ausgedrückt:

$$
\begin{aligned}
\mu_{ij} &= \mu + \alpha_i + \beta_j \\
\mu_{ik} &= \mu + \alpha_i + \gamma_k \\
\mu_{jk} &= \mu + \beta_j + \gamma_k
\end{aligned}
$$

Dreifache Wechselwirkungen

Nullhypothese für die dreifachen Wechselwirkungen:

$$\alpha\beta\gamma_{ijk} = 0 \qquad (i = 1, \ldots, p; j = 1, \ldots, q; k = 1, \ldots, r)$$

Oder anders ausgedrückt:

$$\mu_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \alpha\gamma_{ik} + \beta\gamma_{jk}$$

Quadratsummenzerlegung

Mittelwerte

Für die Quadratsummenzerlegung müssen folgende Mittelwerte berechnet werden:

- Zellenmittelwerte ($\bar{X}_{ijk\cdot}$)
- Interaktionstabellen AB, AC und BC (d.h. die Mittelwerte $\bar{X}_{ij\cdot\cdot}$, $\bar{X}_{i\cdot k\cdot}$ und $\bar{X}_{\cdot jk\cdot}$)
- Randmittelwerte ($\bar{X}_{i\cdots}$, $\bar{X}_{\cdot j\cdot\cdot}$ und $\bar{X}_{\cdot\cdot k\cdot}$)
- Gesamtmittelwert ($\bar{X}_{\cdots\cdot}$)

Quadratsummenzerlegung

- Quadratsumme total

$$QS_{tot} = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{m=1}^{n}(X_{ijkm} - \bar{X}_{\cdots\cdot})^2$$

- Quadratsumme zwischen

$$QS_{zw} = n\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}(\bar{X}_{ijk\cdot} - \bar{X}_{\cdots\cdot})^2$$

- Quadratsumme innerhalb

$$QS_{inn} = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{m=1}^{n}(X_{ijkm} - \bar{X}_{ijk\cdot})^2$$

- Quadratsumme A

$$QS_A = nqr\sum_{i=1}^{p}(\bar{X}_{i\cdots} - \bar{X}_{\cdots\cdot})^2$$

- Quadratsumme B

$$QS_B = npr \sum_{j=1}^{q} (\bar{X}_{\cdot j \cdot \cdot} - \bar{X}_{\cdot\cdot\cdot\cdot})^2$$

- Quadratsumme C

$$QS_C = npq \sum_{k=1}^{r} (\bar{X}_{\cdot\cdot k \cdot} - \bar{X}_{\cdot\cdot\cdot\cdot})^2$$

- Quadratsumme Zwischen AB

$$QS_{zwAB} = nr \sum_{i=1}^{p} \sum_{j=1}^{q} (\bar{X}_{ij \cdot\cdot} - \bar{X}_{\cdot\cdot\cdot\cdot})^2$$

- Quadratsumme Zwischen AC

$$QS_{zwAC} = nq \sum_{i=1}^{p} \sum_{k=1}^{r} (\bar{X}_{i \cdot k \cdot} - \bar{X}_{\cdot\cdot\cdot\cdot})^2$$

- Quadratsumme Zwischen BC

$$QS_{zwBC} = np \sum_{j=1}^{q} \sum_{k=1}^{r} (\bar{X}_{\cdot jk \cdot} - \bar{X}_{\cdot\cdot\cdot\cdot})^2$$

- Quadratsumme AB

$$QS_{AB} = QS_{zwAB} - QS_A - QS_B$$

- Quadratsumme AC

$$QS_{AC} = QS_{zwAC} - QS_A - QS_C$$

- Quadratsumme AB

$$QS_{BC} = QS_{zwBC} - QS_B - QS_C$$

- Quadratsumme ABC

$$QS_{ABC} = QS_{zw} - QS_A - QS_B - QS_C - QS_{AB} - QS_{AC} - QS_{BC}$$

Die Quadratsummen der Haupteffekte und Wechselwirkungen lassen sich auch hier wieder bequem aus den Schätzungen der Effekte berechnen.

...

Freiheitsgrade

$$
\begin{aligned}
df_{tot} &= npqr - 1 \\
df_{zw} &= pqr - 1 \\
df_{inn} &= pqr(n-1) \\
\\
df_A &= p - 1 \\
df_B &= q - 1 \\
df_C &= r - 1 \\
df_{AB} &= (p-1)(q-1) \\
df_{AC} &= (p-1)(r-1) \\
df_{BC} &= (q-1)(r-1) \\
df_{ABC} &= (p-1)(q-1)(r-1)
\end{aligned}
$$

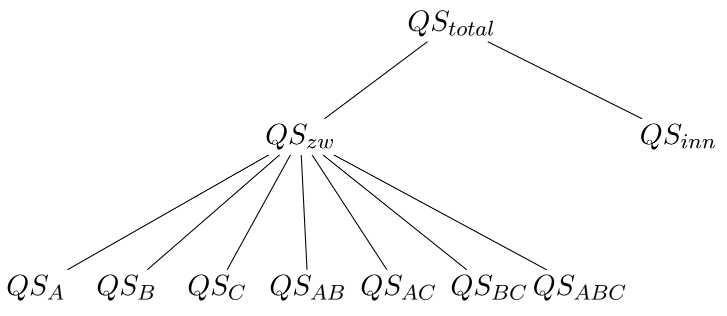
Für die Freiheitsgrade gilt analog zu den Quadratsummen:

$$df_{tot} = df_{zw} + df_{inn}$$

und

$$df_{zw} = df_A + df_B + df_C + df_{AB} + df_{AC} + df_{BC} + df_{ABC}$$

Diese Quadratsummenzerlegung kann das folgende Baumdiagramm veranschaulicht werden:



Erwartungswerte der mittleren Quadrate

Erwartungswerte der mittleren Quadrate

$$\begin{aligned}
E(MQ_A) &= \sigma_\varepsilon^2 + nqr\sigma_\alpha^2 \\
E(MQ_B) &= \sigma_\varepsilon^2 + npr\sigma_\beta^2 \\
E(MQ_C) &= \sigma_\varepsilon^2 + npq\sigma_\gamma^2 \\
\\
E(MQ_{AB}) &= \sigma_\varepsilon^2 + nr\sigma_{\alpha\beta}^2 \\
E(MQ_{AC}) &= \sigma_\varepsilon^2 + nq\sigma_{\alpha\gamma}^2 \\
E(MQ_{BC}) &= \sigma_\varepsilon^2 + np\sigma_{\beta\gamma}^2 \\
\\
E(MQ_{ABC}) &= \sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 \\
\\
E(MQ_{inn}) &= \sigma_\varepsilon^2
\end{aligned}$$

Rechnerische Durchführung

Rechnerische Durchführung

Aus den Erwartungswerten der einzelnen Mittleren Quadrate kann man ableiten, dass alle Haupteffekte und Wechselwirkung gegen MQ Innerhalb getestet werden müssen.

...

Varianzanalysetabelle

Quelle der Variation	**QS**	**df**	**MQ**	**F**
Zwischen				
A				
B				
C				
AxB				
AxC				
BxC				
AxBxC				
Innerhalb				—
Total			—	—

Interaktionseffekte

Interaktionseffekte

- Interaktionseffekte erster Ordnung (d.h. zweifache Interaktionseffekte) werden genauso behandelt wie bei einer zweifaktoriellen Varianzanalyse (Betrachtung von Interaktionsdiagrammen, Prüfung von bedingten Haupteffekte u.dgl.).
- Für Interaktionseffekte zweiter Ordnung (d.h. die dreifachen Interaktionseffekte) kann man beispielsweise die bedingten Interaktionseffekte erster Ordnung vergleichen.

Interaktionstabellen

- Für Interaktionseffekte zweiter Ordnung bildet man die bedingten zweidimensionalen Interaktionstabellen und prüft die Interaktion in der jeweiligen Tabelle (Profilvergleich). Unterscheiden sich diese Interaktionen, so liegt eine dreifache Interaktion vor.
- Letztlich betrachtet man also Differenzen zweiter Ordnung (Differenzen von Differenzen): sind sie konstant, so liegt keine dreifache Interaktion vor.

Interaktionsdiagramme

- Haupteffekte und Wechselwirkungen erster Ordnung werden wie bei einer zweifaktoriellen Varianzanalyse mit Hilfe eines Interaktionsdiagramms dargestellt.
- Eine dreifache Wechselwirkung lässt sich durch bedingte Interaktionsdiagramme veranschaulichen. Die Stufen eines der Faktoren stellen jeweils die Bedingung dar, d.h. für jede Stufe dieses Faktors wird ein zweidimensionales Interaktionsdiagramm erstellt.
- Das Fehlen einer dreifachen Wechselwirkung zeigt sich darin, dass die einzelnen Diagramme einander ähnlich sind, d.h. dass beispielsweise die gegenseitigen Abstände der Linienzüge in den verschiedenen Diagrammen einander entsprechen.

Die Prüfung spezieller Hypothesen

Prüfung spezieller Hypothesen

Die Prüfung von speziellen Hypothesen erfolgt völlig analog zur zweifachen Varianzanalyse.

...

Bedingte Haupteffekte und bedingte Vergleiche

Bedingte Haupteffekte und Wechselwirkungen

Ebenso lassen sich bei der dreifaktoriellen Varianzanalyse bedingte Haupteffekte, bedingte Wechselwirkungen (erster Ordnung) und bedingte Einzelvergleiche prüfen.

...

Rechenbeispiel

Rechenbeispiel

Einzelwerte

a_1				a_2			
b_1		b_2		b_1		b_2	
c_1	c_2	c_1	c_2	c_1	c_2	c_1	c_2
3	4	7	7	1	2	5	10
6	5	8	8	2	3	6	10
3	4	7	9	2	4	5	9
3	3	6	8	2	3	6	11

ABC-Interaktionstabelle (Zellenmittelwerte)

	b_1		b_2	
	c_1	c_2	c_1	c_2
a_1	3.75	4	7	8
a_2	1.75	3	5.5	10

AB-Interaktionstabelle

	b_1	b_2	
a_1	3.875	7.5	5.6875
a_2	2.375	7.75	5.0625
	3.125	7.625	5.375

AC-Interaktionstabelle

	c_1	c_2	
a_1	5.375	6	5.6875
a_2	3.625	6.5	5.0625
	4.5	6.25	5.375

BC-Interaktionstabelle

	c_1	c_2	
b_1	2.75	3.5	3.125
b_2	6.25	9	7.625
	4.5	6.25	5.375

Bedingte AB-Interaktionstabellen

C=1

	b_1	b_2	
a_1	3.75	7	5.375
a_2	1.75	5.5	3.625
	2.75	6.25	4.5

C=2

	b_1	b_2	
a_1	4	8	6
a_2	3	10	6.5
	3.5	9	6.25

Bedingte AC-Interaktionstabellen

B=1

	c_1	c_2	
a_1	3.75	4	5.375
a_2	1.75	3	3.625
	2.75	3.5	3.125

B=2

	c_1	c_2	
a_1	7	8	7.5
a_2	5.5	10	7.75
	6.25	9	7.625

Bedingte BC-Interaktionstabellen

A=1

	c_1	c_2	
b_1	3.75	4	3.875
b_2	7	8	7.5
	5.375	6	5.6875

A=2

	c_1	c_2	
b_1	1.75	3	2.375
b_2	5.5	10	7.75
	3.625	6.5	5.0625

Modelle mit zufälligen Effekten

Einordnung des Problems

Einteilung der varianzanalytischen Modelle

- Modell I: Modelle mit festen Effekten („fixed-effects models“)
- Modell II: Modelle mit zufälligen Effekten („random-effects models“)
- Modell III: Gemischte Modelle („mixed-effects models“)

Modelle mit festen Effekten

- Alle interessierenden Stufen des Faktors sind realisiert, d.h. es sind nur Aussagen über die realisierten Stufen beabsichtigt:
  - Der Faktor hat entweder endlich viele Stufen, und alle Stufen sind realisiert oder
  - der Faktor hat endlich oder unendlich viele Stufen, und es interessieren nur einige ausgewählte, d.h. die Gültigkeit der Ergebnisse wird auf diese eingeengt.
- Falls das Experiment wiederholt würde, wären dieselben Stufen zu verwenden
- Bezeichnung dieser Modelle: „fixed-effects models“ („Modell I“)

Beispiele

- Lehrmethoden
- Therapieformen
- Geschlecht, Alter bzw. Altersgruppe, Schicht, ...
- Instruktionen
- ...

Modelle mit zufälligen Effekten

Für alle Faktoren gilt:

- Nicht alle interessierenden Stufen sind realisiert, d.h. es sollen Aussagen auch über die Stufen gemacht werden, die nicht realisiert wurden, d.h. über den gesamten Wertebereich von möglichen Ausprägungen des Faktors.
- Realisierte Stufen sind eine Zufallsstichprobe des gesamten Variationsbereichs, d.h. ob ein bestimmtes Treatment realisiert ist oder nicht, hängt vom Zufall ab.
- Falls das Experiment wiederholt würde, kämen vermutlich andere Stufen zum Einsatz
- Zufallsfaktoren: Anzahl der Ausprägungen groß bzw. unendlich und es interessiert nicht die Wirkung einzelner, <u>bestimmter</u> Ausprägungen
- Bezeichnung dieser Modelle: „random-effects models“ („Modell II“)

Beispiele

- Einfluss der Persönlichkeit des Experimentators auf die Ergebnisse
- Lehrer
- Therapeuten
- Schulklassen
- Versuchspersonen
- ...

Gemischte Modelle

- Modelle enthalten sowohl Faktoren mit festen als auch solche mit zufälligen Effekten
- Bezeichnung dieser Modelle: „mixed-effects models“ („Modell III“)
- Beispiel: Lehrer (Random Faktor) unterrichten zwei verschiedene Unterrichtsmethoden (fixer Faktor)...

Strukturmodelle

Strukturmodell bei zufälligen oder bei gemischten Effekten

- Das Strukturmodell enthält jeweils neben konstanten Termen (z.B. μ) und neben dem zufälligen Fehler ε auch Parameter, die Ausprägungen von Zufallsvariablen sind.
- Innerhalb der Stichprobe von n Beobachtungen sind die realisierten Ausprägungen eines Random Faktors jedoch eine Konstante.

Varianzkomponenten

Varianzkomponenten

Als Kennzehl zur Beschreibung der zufälligen Effekte werden nicht die Effekte selbst angegeben (es gibt ja in den meisten Fällen unendlich viele), sondern deren Varianz.

Das Ziel der Analyse ist es damit, die Varianzkomponenten der einzelnen Faktoren zu schätzen, etwa so:

$$\sigma_y^2 = \sigma_A^2 + \sigma_B^2 + \sigma_{AB}^2 + \sigma_\varepsilon^2$$

Übrigens wird meist auch angenommen, dass die zufälligen Effekte normalverteilt sind mit Erwartungswert Null und einer bestimmten Varianz σ_{Effekt}^2.

ANOVA einfaktoriell

Einfache ANOVA mit zufälligen Effekten

- Faktor A: Faktor mit zufälligen Effekten, eine Zufallsstichprobe von p Stufen ist realisiert
- Unabhängige Stichproben
- Pro Stichprobe n Versuchspersonen

Beispiel

- Ein psychologisches Experiment soll von 5 verschiedenen Experimentatoren durchgeführt werden.
- Jedem Experimentator wird per Zufall eine Stichprobe von 8 Versuchspersonen zugeteilt.

- Es interessiert nicht der Einfluss der 5 ausgewählten Experimentatoren, sondern generell der Einfluss des Experimentators auf die Ergebnisse.
- Daher ist der Faktor Experimentator ein sog. „Random-Faktor“

Strukturmodell

$$X_{im} = \mu + a_i + \varepsilon_{im}$$

a_i sind Realisationen der Zufallsvariable A (normalverteilt mit Mittelwert 0 und Varianz σ_A^2)V

ε: wie gewohnt (d.h normalverteilt mit Mittelwert 0 und Varianz σ_ε^2)

Nullhypothese

$$\sigma_A^2 = 0$$

Alternativhypothese

$$\sigma_A^2 > 0$$

Hypothesenprüfung

Berechne wie beim Modell mit festen Effekten:

- Quadratsumme, Freiheitsgrade und Mittleres Quadrat Zwischen (QS_A, df_A, MQ_A)
- Quadratsumme, Freiheitsgrade und Mittleres Quadrat Innerhalb (QS_{inn}, df_{inn}, MQ_{inn})

Erwartungswerte

Die Erwartungswerte der mittleren Quadrate sind:

$$\begin{aligned} E(MQ_A) &= n\sigma_A^2 + \sigma_\varepsilon^2 \\ E(MQ_{inn}) &= \sigma_\varepsilon^2 \end{aligned}$$

Unter H_0 ist $\sigma_A^2 = 0$, so dass man erhält:

$$E(MQ_A) = \sigma_\varepsilon^2$$

F-Test

Also ist die Testgröße dieselbe wie beim Modell mit festen Effekten:

$$F = \frac{MQ_A}{MQ_{inn}}$$

mit derselben Anzahl von Freiheitsgraden.

ANOVA zweifaktoriell

Zweifaktorielle ANOVA mit zufälligen Effekten

Beispiel

Projektiver Test mit 10 Karten. Die Karten werden einzeln nacheinander in einer bestimmten Reihenfolge den Vpn vorgelegt. Vp muss zu jeder Karte möglichst viele freie Assoziationen nennen.AV: Gesamtzahl aller Nennungen[26].

- Faktor A: Testleiter (Random Faktor, realisiert mit p Stufen)
- Faktor B: Reihenfolge (Random Faktor, realisiert mit q Stufen)
- Jede Kombination wird realisiert.
- $p \times q$ unabhängige Stichproben vom Umfang n

Strukturmodell

$$X_{ijm} = \mu + a_i + b_j + ab_{ij} + \varepsilon_{ijm}$$

μ	Mittelwert der Gesamtpopulation (Konstante)
a_i	Realisationen der Zufallsvariable A
b_j	Realisationen der Zufallsvariable B
ab_{ij}	zufällige Wechselwirkungseffekte
ε_{ijm}	Fehler

Voraussetzungen

- a_i: normalverteilt mit Mittelwert 0 und Varianz σ_A^2
- b_j: normalverteilt mit Mittelwert 0 und Varianz σ_B^2

[26]Beispiel aus Hays: *Statistics for the Social Scientists.*

- ab_{ij}: normalverteilt mit Mittelwert 0 und Varianz σ_{AB}^2
- ε_{ijm}: normalverteilt mit Mittelwert 0 und Varianz σ_ε^2
- Die a_i, b_j, ab_{ij} und ε sind paarweise stochastisch unabhängig.

Nullhypothesen

$$\begin{aligned} \sigma_A^2 &= 0 \\ \sigma_B^2 &= 0 \\ \sigma_{AB}^2 &= 0 \end{aligned}$$

Alternativhypothesen

$$\begin{aligned} \sigma_A^2 &> 0 \\ \sigma_B^2 &> 0 \\ \sigma_{AB}^2 &> 0 \end{aligned}$$

Hypothesenprüfung

Die Quadratsummen, Freiheitsgrade und mittleren Quadrate von A, B, AB und Innerhalb (Error) werden in exakt der gleichen Weise berechnet wie bei Modell I …

Erwartungswerte der mittleren Quadrate

Wir betrachten zunächst ein Modell mit festen Effekten.

Hier sind die Summen der Wechselwirkungseffekte für jede Zeile und jede Spalte gleich Null.

Nicht so bei einem Modell mit zufälligen Effekten.

Hier ist lediglich die Summe über alle (realisierten und nicht realisierten) Wechselwirkungseffekte zeilen- und spaltenweise Null, aber nicht unbedingt bei den realisierten Stufen.

Daraus folgt, dass der Erwartungswert des mittleren Quadrats der Haupteffekte auch von der Varianz der Wechselwirkungseffekte abhängt.

Entsprechend sind die Erwartungswerte der mittleren Quadrate wie folgt:

$$\begin{aligned}E(MQ_A) &= \sigma_\varepsilon^2 + nq\sigma_A^2 + n\sigma_{AB}^2\\ E(MQ_B) &= \sigma_\varepsilon^2 + np\sigma_B^2 + n\sigma_{AB}^2\\ E(MQ_{AB}) &= \sigma_\varepsilon^2 + n\sigma_{AB}^2\\ E(MQ_{inn}) &= \sigma_\varepsilon^2\end{aligned}$$

Unter der Nullhypothese ist σ_A^2 bzw. σ_B^2 bzw. σ_{AB}^2 gleich 0, so dass man dann erhält:

$$\begin{aligned}E(MQ_A) &= \sigma_\varepsilon^2 + n\sigma_{AB}^2\\ E(MQ_B) &= \sigma_\varepsilon^2 + n\sigma_{AB}^2\\ E(MQ_{AB}) &= \sigma_\varepsilon^2\end{aligned}$$

Prüfgrößen

$$\begin{aligned}F_A &= \frac{MQ_A}{MQ_{AB}}\\ df_1 &= p-1\\ df_2 &= (p-1)(q-1)\\[1em] F_B &= \frac{MQ_B}{MQ_{AB}}\\ df_1 &= q-1\\ df_2 &= (p-1)(q-1)\\[1em] F_{AB} &= \frac{MQ_{AB}}{MQ_{Inn}}\\ df_1 &= (p-1)(q-1)\\ df_2 &= pq(n-1)\end{aligned}$$

Gepoolter Errorterm

Falls die Wechselwirkung nicht signifikant ist, kann deren Quadratsumme für eine genauere Schätzung der Fehlervarianz herangezogen werden (Paull, 1950):

Falls die mittleren Quadrate sowohl für den Fehler als auch für die Wechselwirkung mehr als 6 Freiheitsgrade haben, dann sollen sie zu einer gemeinsamen Schätzung der

Fehlervarianz gepoolt werden, falls der F-Wert für die Wechselwirkung kleiner als 2 ist:

$$MS_{error\ gepoolt} = \frac{QS_{AB} + QS_{inn}}{(p-1)(q-1) + pq(n-1)}$$

Beide Haupteffekte sind dann gegen diesen Errorterm zu testen.

ANOVA zweifaktoriell

Zweifaktorielle ANOVA mit zufälligen Effekten

- Faktor A: Beurteiler (Zufallsstichprobe)
- Faktor B: beurteilte Person (Zufallsstichprobe)
- Jede Person wird von jedem Urteiler einmal beurteilt ($n = 1$)

Strukturmodell

$$X_{ij} = \mu + a_i + b_j + ab_{ij} + \varepsilon_{ij}$$

μ	Mittelwert der Gesamtpopulation (Konstante)
a_i	Realisationen der Zufallsvariable A
b_j	Realisationen der Zufallsvariable B
ab_{ij}	zufällige Wechselwirkungseffekte
ε_{ij}	Fehler

Voraussetzungen

- a_i: normalverteilt mit Mittelwert 0 und Varianz σ^2_A
- b_j: normalverteilt mit Mittelwert 0 und Varianz σ^2_B
- ab_{ij}: normalverteilt mit Mittelwert 0 und Varianz σ^2_{AB}
- ε_{ij}: normalverteilt mit Mittelwert 0 und Varianz σ^2_ε
- Die a_i, b_j, ab_{ij} und ε sind paarweise stochastisch unabhängig.

Das Modell enthält zu viele Parameter:

Wegen $n = 1$ kann die Wechselwirkung bzw. die Fehlervarianz nicht geschätzt werden.

Daher wird die Quadratsumme Total zerlegt in

- Quadratsumme A (QS_A),

- Quadratsumme B (QS_B) und
- Quadratsumme Residuum (QS_{res}, enthält QS_{AB} und QS_{inn}).

Nullhypothesen

$$\begin{aligned} \sigma_A^2 &= 0 \\ \sigma_B^2 &= 0 \end{aligned}$$

Alternativhypothesen

$$\begin{aligned} \sigma_A^2 &> 0 \\ \sigma_B^2 &> 0 \end{aligned}$$

Hypothesenprüfung

Quadratsummenzerlegung

Berechne:

$$\begin{aligned} QS_{tot} &= \sum_i \sum_j (X_{ij} - \bar{X}_{\cdot\cdot})^2 \\ QS_A &= q \sum_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 \\ QS_B &= p \sum_j (\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot})^2 \\ QS_{res} &= QS_{tot} - QS_A - QS_B \end{aligned}$$

Freiheitsgrade

Berechne:

$$\begin{aligned} df_{tot} &= pq - 1 \\ df_A &= p - 1 \\ df_B &= q - 1 \\ df_{res} &= df_{tot} - df_A - df_B \\ &= (p-1)(q-1) \end{aligned}$$

Hypothesenprüfung

Faktor A und Faktor B werden beide gegen das Residuum getestet.

Gemischte Effekte

Zweifaktorielle ANOVA mit gemischten Effekten

Beispiel

Wie obiges Beispiel, nur soll der Faktor Reihenfolge als fixer Faktor aufgefasst werden, d.h. es soll lediglich der Effekt die vorgegebenen Reihenfolgen geprüft werden.

- Faktor A: Testleiter (p Stufen)
- Faktor B: Reihenfolge (q Stufen)
- Jede Kombination wird realisiert.
- $p \times q$ unabhängige Stichproben vom Umfang n

Strukturmodell

$$X_{ijm} = \mu + a_i + \beta_j + a\beta_{ij} + \varepsilon_{ijm}$$

μ	Mittelwert der Gesamtpopulation (Konstante)
a_i	Realisationen der Zufallsvariable A
β_j	Effekt der Reihenfolge j
$a\beta_{ij}$	zufällige Wechselwirkungseffekte
ε_{ijm}	Fehler

Voraussetzungen

- a_i: normalverteilt mit Mittelwert 0 und Varianz σ_A^2
- β_j: feste Parameter
- $a\beta_{ij}$: normalverteilt mit Mittelwert 0 und Varianz σ_{AB}^2
- ε_{ijm}: normalverteilt mit Mittelwert 0 und Varianz σ_ε^2
- Die a_i, $a\beta_{ij}$ und ε_{ijm} sind paarweise stochastisch unabhängig.

Erwartungswerte der mittleren Quadrate

Auch hier gilt, dass die Summen der Wechselwirkungseffekte pro Stufe des fixen Faktors nicht Null sein müssen.

Daraus folgt, dass der Erwartungswert des mittleren Quadrats des Haupteffekts des fixen Faktors auch von der Varianz der Wechselwirkungseffekte abhängt.

Entsprechend sind die Erwartungswerte der mittleren Quadrate wie folgt:

$$\begin{aligned} E(MQ_A) &= \sigma_\varepsilon^2 + nq\sigma_A^2 \\ E(MQ_B) &= \sigma_\varepsilon^2 + np\sigma_B^2 + n\sigma_{AB}^2 \\ E(MQ_{AB}) &= \sigma_\varepsilon^2 + n\sigma_{AB}^2 \\ E(MQ_{inn}) &= \sigma_\varepsilon^2 \end{aligned}$$

Unter der Nullhypothese ist σ_A^2 bzw. σ_B^2 bzw. σ_{AB}^2 gleich 0, so dass man dann erhält:

$$\begin{aligned} E(MQ_A) &= \sigma_\varepsilon^2 \\ E(MQ_B) &= \sigma_\varepsilon^2 + n\sigma_{AB}^2 \\ E(MQ_{AB}) &= \sigma_\varepsilon^2 \end{aligned}$$

Testgrößen

Also ist der Haupteffekt A und die Wechselwirkung gegen Innerhalb zu testen.

Der Haupteffekt B hingegen muss gegen die Wechselwirkung getestet werden.

Hierarchische Versuchspläne

Einordnung des Problems

Vollständige vs. unvollständige Versuchspläne

- Bisher haben wir faktorielle Versuchspläne betrachtet, bei denen alle Faktorstufenkombinationen realisiert waren. Solche Versuchspläne nennt man vollständig kombiniert.
- Mit Hilfe vollständiger Versuchspläne ist es möglich, alle Haupteffekte und alle Wechselwirkungen zu testen.

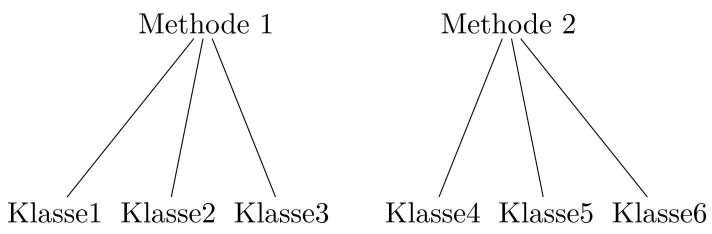
- Der Nachteil dieser Versuchsplänen ist, dass sie mit einem großen Aufwand verbunden sind.
- Sind nicht alle möglichen Faktorstufenkombinationen realisiert worden, so spricht man von einem unvollständigen Versuchsplan.
- Spezielle Formen von unvollständigen Versuchsplänen sind hierarchische bzw. teilhierarchische Versuchspläne und das lateinische bzw. lateinisch-griechische Quadrat.

Hierarchische Versuchspläne

- Zuweilen ist es auch nicht möglich, alle Kombinationen verschiedener Faktoren zu erzeugen.
- Dies ist beispielsweise dann der Fall, wenn jede Stufe eines Faktors nur mit einer bestimmten Stufe eines anderen Faktors auftritt oder auftreten kann.
- In diesem Fall spricht man von einem hierarchischen Versuchsplan. Man sagt, der eine Faktor ist unter dem anderen geschachtelt („genestet“).
- Beispiel: jede Schulklasse (Faktor B) wird nur nach einer von zwei Lehrmethoden (Faktor A) unterrichtet.

- In einem hierarchischen Versuchsplan können Wechselwirkungen zwischen genesteten Faktoren nicht geprüft werden.
- Ist Faktor B unter Faktor A genestet, so schreibt man üblicherweise B(A), und überall dort, wo der Index j auftaucht, schreibt man den Index i in Klammern dahinter: $j(i)$. Das Strukturmodell für das obige Beispiel mit den Lehrmethoden und Schulklassen schreibt man also wie folgt:

$$X_{ijm} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijm}$$



Quadratsummenzerlegung

Quadratsummenzerlegung

1. Bei einem hierarchischen Versuchsplan wird zunächst eine Quadratsummenzerlegung ohne die hierarchisch höheren Faktoren durchgeführt.
2. Die Quadratsumme und Freiheitsgrade Zwischen des genesteten Faktors (d.h. zwischen allen Zellen) werden nun zerlegt: es werden die bedingten Quadratsummen (und Freiheitsgrade) des genesteten Faktors pro Stufe des hierarchisch höheren Faktors berechnet.
3. Die Summe der bedingten Quadratsumme (und Freiheitsgrade) ergeben die gepoolte Quadratsumme (und Freiheitsgrade) des genesteten Faktors.
4. Die Differenz zwischen QS_{zw}(und df_{zw}) bei (1) und und der gepoolten Quadratsumme (und Freiheitsgrade) bei (3) ergibt die Quadratsumme (und Freiheitsgrade) des hierarchisch höheren Faktors.
5. Alternativ zu (2) und (3) können auch wie üblich die Quadratsumme und Freiheitsgrade des hierarchisch höheren Faktors berechnet werden. Die Werte für den genesteten Faktor erhält man dann als Differenz zu den Werten zwischen allen Zellen.
6. Als Fehlerterm ist (bei fixen Faktoren) jeweils MQ_{inn} zu nehmen.

Zweifaktorieller hierarchischer Versuchsplan

Zweifaktorieller hierarchischer Versuchsplan

Beispiel

- Faktor A: p Stufen (fixe Effekte)
- Faktor B: q Stufen, unter A genestet (zufällige Effekte)
- $q = q_1 + q_2 + \cdots + q_p$ unabhängige Stichproben vom Umfang n, u.zw. unter a_1 q_1 Stichproben, unter a_2 q_2 Stichproben usw.[27].

Schema

a_1			a_2			... ...
b_1	...	b_{q_1}	b_{q_1+1}	...	$b_{q_1+q_2}$	... ...
Gruppe 1	...	Gruppe q_1	Gruppe $q_1 + 1$	...	Gruppe $q_1 + q_2$	... ...

[27]Häufig wird $q_1 = q_2 = \cdots = q_p$ angenommen, eine Voraussetzung, die nicht notwendig ist.

Strukturmodell

$$X_{ijm} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijm}$$

Beispiel

a_1			a_2		
b_1	b_2	b_3	b_4	b_5	b_6
5,6,18,12	7,18,4,11	16,5,9,14	24,21,12,16	9,23,28,19	17,26,24,19

Prüfung von Effekten

Erwartungswerte von mittleren Quadraten

Erwartungswerte von mittleren Quadraten

Vorbereitung

Es wird eine aus 3 Teilen (Spaltenblöcken) bestehende Tabelle angefertigt:

1. Als erste Spalte (Teil I) werden die Parameter des Strukturmodells außer dem Parameter μ untereinander eingetragen. Die einzelnen Zeilen der Tabelle beziehen sich jeweils auf den entsprechenden Effekt. Zu jedem Effekt werden auch dessen Indizes, so wie im Strukturmodell, geschrieben. Beim Parameter ε werden alle Indizes eingeklammert.
2. Teil II besteht aus ebenso vielen Spalten, wie das Strukturmodell Indizes enthält. Die einzelnen Indizes ergeben die Spaltenüberschriften. Darunter (als zweite Überschriftszeile) wird noch notiert, wie weit der jeweilige Index läuft.
3. Teil III enthält für jeden Effekt den Erwartungswert des mittleren Quadrats.

Ausfüllen von Teil II

In die Zellen dieses Teil werden nun folgende Werte eingetragen:

1. Falls der Spaltenindex beim Parameter in Teil I vorkommt und nicht eingeklammert ist und der Spaltenfaktor ein fixer Faktor ist, wird der Wert 0 eingetragen.
2. Falls der Spaltenindex beim Parameter in Teil I nicht vorkommt, wird der Wert der zweiten Überschriftszeile übernommen.
3. In allen anderen Fällen wird der Wert 1 eingetragen.

Ausfüllen von Teil III

Eine Zeile dieses Teils wird folgendermaßen ausgefüllt:

1. Zunächst werden sämtliche Varianzanteile aufgelistet, die alle Indizes des Parameters dieser Zeile enthalten; gleichzeitig werden aus Teil II alle Spalten abgedeckt, die diesen Indizes entsprechen.
2. Jeder der aufgelisteten Varianzanteile erhält als Koeffizienten das Produkt der nicht abgedeckten Werte aus seiner Zeile in Teil II.
3. Die Summe der einzelnen Varianzanteile ergibt jeweils den Erwartungswert für das entsprechende mittlere Quadrat.

Konstruktion von Prüfgrößen

Konstruktion von Prüfgrößen in der Varianzanalyse

- Hypothesen über Effekte werden in der Varianzanalyse mit Hilfe der Prüfgröße F getestet.
- Ein Effekt wird dabei jeweils gegen einen passenden Fehlerterm getestet. Die Varianz eines Effekts ... Der Varianzanteil eines Effekts steckt als Komponente in einer mittleren Quadratsumme ...
- Der F-Wert enthält dabei im Zähler eine mittlere Quadratsumme, deren Erwartungswert außer den Varianzanteilen des Nenners den zu testenden Varianzanteil enthält
- Zähler und Nenner des F-Wertes unterscheiden sich dabei nur in dem zu testenden Varianzanteil.

- Zuweilen kann jedoch kein passender F-Quotient gebildet werden, so dass der entsprechende Effekt nicht direkt testbar ist.
- Zwei Strategien stehen dann zur Auswahl:
  1. Streichen von höheren Interaktionen und „Pooling“
  2. Quasi F-Brüche

Gepoolte Fehlerterme

Pooling

- Zunächst wird das volle Modell angenommen, d.h. jenes mit allen Haupteffekten und allen Wechselwirkungen jeder Ordnung.
- Anschließend werden z.B. die Wechselwirkungen höherer Ordnung getestet.
- Sind sie nicht signifikant, werden sie dem Fehler zugeschlagen: die Quadratsumme und die Freiheitsgrade werden zu der Quadratsumme und den Freiheitsgraden des Fehlers hinzuaddiert. Das Ergebnis wird oft auch als Residuum (QS_{res}, df_{res}) bezeichnet.
- Damit wird das urspüngliche Modell modifiziert, d.h. vereinfacht, da es weniger Parameter (Einflussgrößen) enthält.

- Der neue Fehlerterm ist dann $MQ_{res} = QS_{res}/df_{res}$.
- Dieses Verfahren lässt sich für weitere im Strukturmodell enthaltenen Terme fortsetzen.
- Natürlich kann die Pooling-Prozedur auch angewandt werden, wenn für die Prüfung von Effekten Fehlerterme für direkte F-Tests zur Verfügung stehen.

Beispiel

Quelle der Variation	**QS**	**df**	**MQ**	**F**	
A	120	3	40.00		
B	60	1	60.00		
C	40	2	20.00		
AB	96	3	32.00		
AC	72	6	12.00		
BC	18	2	9.00	0.90	$F_{.75;2,24} = 1.47$
ABC	72	6	12.00	1.20	$F_{.75;6,24} = 1.41$
Innerhalb	240	24	10.00		

Wir poolen nun BC, ABC und Innerhalb und erhalten ein Residual als neuen Fehlerterm ...

Beispiel

Quelle der Variation	**QS**	**df**	**MQ**	**F**	
A	120	3	40.00		
B	60	1	60.00		
C	40	2	20.00		
AB	96	3	32.00	3.10	$F_{.75;3,32} = 1.44$
AC	72	6	12.00	1.16	$F_{.75;6,32} = 1.39$
Residual	330	32	10.31		

AC können wir noch zum Residual schlagen, nicht aber AB. Wir erhalten als endgültige Tabelle . . .

Beispiel

Quelle der Variation	**QS**	**df**	**MQ**	**F**	
A	120	3	40.00	1.25	$F_{.95;3,38} = 9.28$
B	60	1	60.00	5.67	$F_{.95;1,38} = 4.10$
C	40	2	20.00	1.89	$F_{.95;2,38} = 3.25$
AB	96	3	32.00	3.02	$F_{.95;3,38} = 2.85$
Residual	402	38	10.58		

Quasi F-Brüche

Quasi F-Brüche

- Neben der Streichung von nicht signifikanten Wechselwirkungen aus dem Modell und dem Pooling von Quadratsummen gibt es noch die Möglichkeit der Konstruktion von Quasi F-Brüchen, um Effekte wenigstens näherungsweise zu testen, die nicht direkt testbar sind.
- Ein Quasi F-Bruch ist ein Quotient aus Summen und evtl. Differenzen von mittleren Quadraten.
- Die Summen (bzw. Differenzen) der Erwartungswerte der mittleren Quadrate im Zähler und im Nenner müssen gleich sein bis auf den einzelnen zu testenden Effekt, um den der Zähler größer ist.
- Ein Quasi F-Bruch hat nur eine angenäherte F-Verteilung[28].

[28] Deshalb wird er auch nicht mit F, sondern z.B. mit F' bezeichnet.

Freiheitsgrade eines Quasi F-Bruchs

Ist ein F-Bruch z.B. folgendermaßen definiert:

$$F' = \frac{MQ_1 + MQ_2}{MQ_3 + MQ_4}$$

mit den mittleren Quadraten MQ_1, MQ_2, MQ_3, MQ_4 mit jeweils df_1, df_2, df_3, df_4 Freiheitsgraden, so sind für den F-Bruch die folgenden aproximierten Freiheitsgrade zu verwenden (eventuell abgerundet auf die nächste ganze Zahl):

$$\begin{aligned} df_{zaehler} &= \frac{(MQ_1 + MQ_2)^2}{MQ_1^2/df_1 + MQ_2^2/df_2} \\ df_{nenner} &= \frac{MQ_3 + MQ_4}{MQ_3^2/df_3 + MQ_4^2/df_4} \end{aligned}$$

Beispiel

Quelle der Var.	E(MQ)
A	$\sigma_\varepsilon^2 + n\sigma_{abc}^2 + nq\sigma_{ac}^2 + nr\sigma_{ab}^2 + nqr\sigma_a^2$
B	$\sigma_\varepsilon^2 + n\sigma_{abc}^2 + np\sigma_{bc}^2 + nr\sigma_{ab}^2 + npr\sigma_b^2$
C	$\sigma_\varepsilon^2 + n\sigma_{abc}^2 + np\sigma_{bc}^2 + nq\sigma_{ac}^2 + npq\sigma_c^2$
AB	$\sigma_\varepsilon^2 + n\sigma_{abc}^2 + nr\sigma_{ab}^2$
AC	$\sigma_\varepsilon^2 + n\sigma_{abc}^2 + nq\sigma_{ac}^2$
BC	$\sigma_\varepsilon^2 + n\sigma_{abc}^2 + np\sigma_{bc}^2$
ABC	$\sigma_\varepsilon^2 + n\sigma_{abc}^2$
Fehler	σ_ε^2

Quasi F-Brüche für den Test von A:

$$\begin{aligned} F'_A &= \frac{MQ_A}{MQ_{AC} + MQ_{AB} - MQ_{ABC}} \\ F''_A &= \frac{MQ_A + MQ_{ABC}}{MQ_{AC} + MQ_{AB}} \end{aligned}$$

Versuchspläne mit abhängigen Stichproben

Einführung

Versuchspläne mit abhängigen Stichproben

Versuchspläne mit abhängigen Gruppen können in verschiedenen Situationen auftreten:

1. Bei Verwendung von parallelisierten Stichproben
2. Bei Messwiederholungen an denselben Versuchspersonen

Strukturmodell bei abhängigen Stichproben

- Da bei Versuchsplänen mit abhängigen Gruppen dieselbe Beobachtungseinheit (Versuchsperson bzw. Paar, Tripel usw.)[29] unter mehreren Bedingungen auftaucht, muss im Strukturmodell auch ein entsprechender Effekt vorgesehen werden.
- Der Versuchspersonenfaktor wird immer als Faktor mit zufälligen Effekten verwendet, da man auf alle Versuchspersonen („like these") verallgemeinern möchte.
- Dabei ist es allgemein üblich, den Effekt einer Versuchsperson m mit π_m zu bezeichnen[30].

Voraussetzungen

Über die Voraussetzungen der Normalverteilung und Varianzenhomogenität hinaus müssen bei einer Varianzanalyse mit abhängigen Gruppen noch weitere Voraussetzungen erfüllt sein:

- Normalverteilung der Haupt- und Wechselwirkungseffekte, an denen der Versuchspersonenfaktor beteiligt ist, mit Erwartungswert 0 und Varianz σ^2_π bzw. $\sigma^2_{\alpha\pi}$ bzw. ...
- Gleichheit der Kovarianzen (bzw. Korrelationen) der AV bei den verschiedenen Bedingungen eines Messwiederholungsfaktors, d.h. zwischen den verschiedenen Messzeitpunkten (Zirkularitätsannahme oder „compound symmetry").

Gleichheit der Kovarianzen

- Für die Kovarianzmatrix für ein einfaktorielles Design mit abhängigen Gruppen muss in der Population z.B. gelten:

$$\begin{bmatrix} \sigma^2 & \rho\sigma^2 & \cdots & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \cdots & \rho\sigma^2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho\sigma^2 & \rho\sigma^2 & \cdots & \sigma^2 \end{bmatrix}$$

[29] Im Folgenden soll immer von Versuchsperson (Vp) gesprochen werden, auch wenn es um Beobachtungseinheiten in parallelisierten Stichproben geht.

[30] Der griechische Buchstabe π für Person, entgegen der Konvention, für Effekte von Random Faktoren lateinische Buchstaben zu verwenden.

- Die Gleichheit der Kovarianzen kann mit dem Mauchly Sphärizitäts-Test überprüft werden.

Falls die Voraussetzung der Sphärizität nicht erfüllt ist, stehen verschiedene Möglichkeiten zur Verfügung, z.B.:

- Konservativer F-Test nach Greenhouse-Geisser: als kritischer F-Wert ist $F_{krit} = F_{1-\alpha;1,n-1}$ zu verwenden.
- Korrektur der Freiheisgrade: Multiplikation der Freiheitsgrade mit einem Faktor ε, berechnet nach Greenhouse-Geisser oder Huynh-Feldt.

Einfaktorieller Versuchsplan mit Messwiederholungen

Einfaktorieller Versuchsplan mit Messwiederholungen

- Faktor A (fix) mit p Stufen
- n Versuchspersonen
- Jede Versuchsperson durchläuft alle Bedingungen

Schema

a_1	a_2	...	a_p
Gruppe 1	Gruppe 1	...	Gruppe 1

Strukturmodell

$$X_{im} = \mu + \alpha_i + \pi_m + \alpha\pi_{im} + \varepsilon_{im}$$

Die Parameter $\alpha\pi_{im}$ und ε_{im} sind nicht unterscheidbar, sie bilden bei der Quadratsummenzerlegung das Residual.

Quadratsummenzerlegung

$$
\begin{aligned}
QS_{tot} &= \sum_i \sum_m (X_{im} - \bar{X}_{..})^2 \\
QS_A &= n \sum_i (\bar{X}_{i.} - \bar{X}_{..})^2 \\
QS_{Vp} &= p \sum_m (\bar{X}_{\cdot m} - \bar{X}_{..})^2 \\
QS_{res} &= QS_{tot} - QS_A - QS_{Vp}
\end{aligned}
$$

Freiheitsgrade

$$
\begin{aligned}
df_{tot} &= np - 1 \\
df_A &= p - 1 \\
df_{Vp} &= n - 1 \\
df_{res} &= (p-1)(n-1)
\end{aligned}
$$

Erwartungswerte der mittleren Quadrate

$$
\begin{aligned}
E(MQ_A) &= n\sigma_\alpha^2 + \sigma_{\alpha\pi}^2 + \sigma_\varepsilon^2 \\
E(MQ_{Vp}) &= p\sigma_\pi^2 + \sigma_\varepsilon^2 \\
E(MQ_{res}) &= \sigma_{\alpha\pi}^2 + \sigma_\varepsilon^2
\end{aligned}
$$

Prüfgröße

$$
\begin{aligned}
F &= \frac{MQ_A}{MQ_{res}} \\
df_1 &= p - 1 \\
df_2 &= (p-1)(n-1)
\end{aligned}
$$

Zweifaktorieller Versuchsplan mit Messwiederholungen auf beiden Faktoren

Zweifaktorieller Versuchsplan mit Messwiederholungen auf beiden Faktoren

- Faktor A (fix) mit p Stufen
- Faktor B (fix) mit q Stufen
- n Versuchspersonen
- Jede Versuchsperson durchläuft alle Bedingungen

Schema

	a_1	a_2	...	a_p
b_1	Gruppe 1	Gruppe 1	...	Gruppe 1
b_2	Gruppe 1	Gruppe 1	...	Gruppe 1
...	...	...	...	...
b_q	Gruppe 1	Gruppe 1	...	Gruppe 1

Strukturmodell

$$X_{im} = \mu + \alpha_i + \beta_j + \pi_m + \alpha\beta_{ij} + \alpha\pi_{im} + \beta\pi_{jm} + \alpha\beta\pi_{ijm} + \varepsilon_{ijm}$$

Die Parameter $\alpha\beta\pi_{ijm}$ und ε_{ijm} sind nicht unterscheidbar, sie bilden bei der Quadratsummenzerlegung das Residual.

Prüfgrößen

$$\begin{aligned} F_A &= \frac{MQ_A}{MQ_{AVp}} \\ df_1 &= p - 1 \\ df_2 &= (p-1)(n-1) \end{aligned}$$

$$\begin{aligned} F_B &= \frac{MQ_B}{MQ_{BVp}} \\ df_1 &= q - 1 \\ df_2 &= (q-1)(n-1) \end{aligned}$$

$$
\begin{aligned}
F_{AB} &= \frac{MQ_{AB}}{MQ_{res}} \\
df_1 &= (p-1)(q-1) \\
df_2 &= (p-1)(q-1)(n-1)
\end{aligned}
$$

Versuchspläne mit abhängigen Stichproben

Quadratsummenzerlegung bei Versuchsplänen mit Messwiederholungen

- Bei Versuchsanordnungen mit Messwiederholungen ist die Versuchsperson als Random Faktor aufzufassen.
- Damit gibt es auch keine Replikationen, d.h. zu jeder Bedingungskombination existiert nur eine einzige Beobachtung.
- Es gibt also auch keine Quadratsumme innerhalb.

- Üblicherweise wird bei diesen Versuchsplänen die totale Quadratsumme zunächst zerlegt in die Quadratsumme zwischen den Versuchspersonen und in jene innerhalb der Versuchspersonen.
- Beide Quadratsummen werden ihrerseits weiter zerlegt.
- Zu der Quadratsumme innerhalb der Versuchspersonen gehören alle Quadratsummen, bei denen Messwiederholungsfaktoren allein oder zusammen mit anderen Faktoren beteiligt sind.
- Zur Quadratsumme zwischen der Versuchspersonen zählen jene Quadratsummen, bei denen lediglich Gruppenfaktoren (und der Versuchspersonenfaktor selbst) beteiligt sind.

Einfaktorieller Versuchsplan mit Messwiederholungen

Einfaktorieller Versuchsplan mit Messwiederholungen

- Faktor A (fix) mit p Stufen
- n Versuchspersonen
- Jede Versuchsperson durchläuft alle Bedingungen

Schema

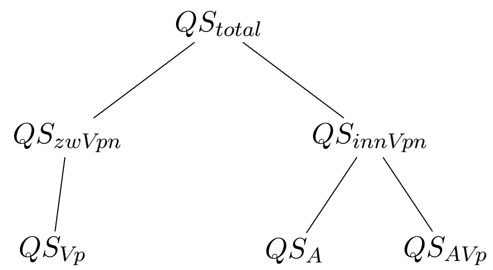
a_1	a_2	...	a_p
Gruppe 1	Gruppe 1	...	Gruppe 1

Strukturmodell

$$X_{im} = \mu + \alpha_i + \pi_m + \alpha\pi_{im} + \varepsilon_{im}$$

Die Parameter $\alpha\pi_{im}$ und ε_{im} sind nicht unterscheidbar, sie bilden bei der Quadratsummenzerlegung das Residual, bzw. die Varianz von ε_{im} ist nicht einzeln schätzbar.

Die Quadratsummenzerlegung für den einfaktoriellen Versuchsplan mit Messwiederholungen kann durch das folgende Baumdiagramm veranschaulicht werden:



$$\begin{aligned}
QS_{tot} &= \sum_i \sum_m (X_{im} - \bar{X}_{\cdot\cdot})^2 \\
QS_A &= n \sum_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 \\
QS_{Vp} &= p \sum_m (\bar{X}_{\cdot m} - \bar{X}_{\cdot\cdot})^2 \\
QS_{AVp} &= QS_{tot} - QS_A - QS_{Vp}
\end{aligned}$$

QS_{AVp} wird häufig auch mit QS_{res} bezeichnet.

Freiheitsgrade

$$\begin{aligned}
df_{tot} &= np-1 \\
df_A &= p-1 \\
df_{Vp} &= n-1 \\
df_{AVp} &= (p-1)(n-1)
\end{aligned}$$

Erwartungswerte der mittleren Quadrate

$$\begin{aligned}
E(MQ_A) &= n\sigma^2_\alpha + \sigma^2_{\alpha\pi} + \sigma^2_\varepsilon \\
E(MQ_{Vp}) &= p\sigma^2_\pi + \sigma^2_\varepsilon \\
E(MQ_{AVp}) &= \sigma^2_{\alpha\pi} + \sigma^2_\varepsilon
\end{aligned}$$

Prüfgröße

$$F = \frac{MQ_A}{MQ_{AVp}}$$

$$\begin{aligned}
df_1 &= p-1 \\
df_2 &= (p-1)(n-1)
\end{aligned}$$

Zweifaktorieller Versuchsplan mit Messwiederholungen auf beiden Faktoren

Zweifaktorieller Versuchsplan mit Messwiederholungen auf beiden Faktoren

- Faktor A (fix) mit p Stufen
- Faktor B (fix) mit q Stufen
- n Versuchspersonen
- Jede Versuchsperson durchläuft alle Bedingungen

Schema

	a_1	a_2	...	a_p
b_1	Gruppe 1	Gruppe 1	...	Gruppe 1
b_2	Gruppe 1	Gruppe 1	...	Gruppe 1
...	...	...	...	...
b_q	Gruppe 1	Gruppe 1	...	Gruppe 1

Strukturmodell

$$X_{ijm} = \mu + \alpha_i + \beta_j + \pi_m + \alpha\beta_{ij} + \alpha\pi_{im} + \beta\pi_{jm} + \alpha\beta\pi_{ijm} + \varepsilon_{ijm}$$

Die Parameter $\alpha\beta\pi_{ijm}$ und ε_{ijm} sind nicht unterscheidbar, sie bilden bei der Quadratsummenzerlegung das Residual, bzw. die Varianz von ε_{ijm} ist nicht einzeln schätzbar.

Quadratsummenzerlegung

$$\begin{aligned} QS_{tot} &= \sum_i \sum_j \sum_m (X_{ijm} - \bar{X}_{...})^2 \\ QS_{zwVpn} &= pq \sum_m (\bar{X}_{..m} - \bar{X}_{...})^2 \\ QS_{innVpn} &= QS_{tot} - QS_{zwVpn} \end{aligned}$$

$$\begin{aligned} QS_{Vp} &= QS_{zwVpn} \\ QS_A &= nq \sum_i (\bar{X}_{i..} - \bar{X}_{...})^2 \\ QS_B &= np \sum_j (\bar{X}_{.j.} - \bar{X}_{...})^2 \end{aligned}$$

$$\begin{aligned} QS_{zwAB} &= n \sum_i \sum_j (\bar{X}_{ij.} - \bar{X}_{...})^2 \\ QS_{AB} &= QS_{zwAB} - QS_A - QS_B \end{aligned}$$

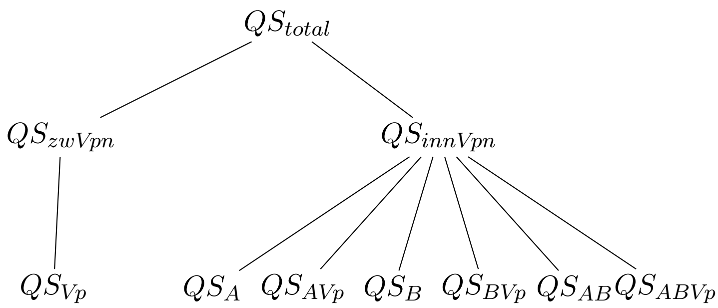
$$\begin{aligned}
QS_{zwAVp} &= q\sum_i \sum_m (\bar{X}_{i\cdot m} - \bar{X}_{\cdots})^2 \\
QS_{AVp} &= QS_{zwAVp} - QS_A - QS_{Vp}
\end{aligned}$$

$$\begin{aligned}
QS_{zwBVp} &= p\sum_j \sum_m (\bar{X}_{\cdot jm} - \bar{X}_{\cdots})^2 \\
QS_{BVp} &= QS_{zwBVp} - QS_B - QS_{Vp}
\end{aligned}$$

$$\begin{aligned}
QS_{ABVp} = {} & QS_{innVp} - QS_A - QS_B \\
& - QS_{AB} - QS_{AVp} - QS_{BVp}
\end{aligned}$$

QS_{ABVp} wird häufig auch mit QS_{res} bezeichnet.

Die Quadratsummenzerlegung kann durch das folgende Baumdiagramm veranschaulicht werden:



Prüfgrößen

$$\begin{aligned} F_A &= \frac{MQ_A}{MQ_{AVp}} \\ df_1 &= p-1 \\ df_2 &= (p-1)(n-1) \end{aligned}$$

$$\begin{aligned} F_B &= \frac{MQ_B}{MQ_{BVp}} \\ df_1 &= q-1 \\ df_2 &= (q-1)(n-1) \end{aligned}$$

$$\begin{aligned} F_{AB} &= \frac{MQ_{AB}}{MQ_{ABVp}} \\ df_1 &= (p-1)(q-1) \\ df_2 &= (p-1)(q-1)(n-1) \end{aligned}$$

Zusammenfassung

Effekt	wird getestet gegen
A	AVp
B	BVp
AB	ABVp

Zweifaktorieller Versuchsplan mit Messwiederholungen auf einem Faktor

Zweifaktorieller Versuchsplan mit Messwiederholungen auf einem Faktor

- Faktor A (fix) mit p Stufen
- Faktor B (fix) mit q Stufen
- Für jede A-Bedingungen eine unabhängige Gruppe à n Versuchspersonen

- Der Vp-Faktor ist unter A genestet.
- Jede Versuchsperson durchläuft alle B-Bedingungen

Schema

	b_1	b_2	...	b_q
a_1	Gruppe 1	Gruppe 1	...	Gruppe 1
a_2	Gruppe 2	Gruppe 2	...	Gruppe 2
...	...	...	...	...
a_p	Gruppe p	Gruppe p	...	Gruppe p

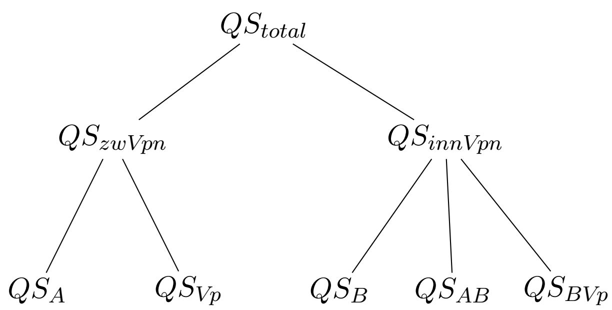
Strukturmodell

$$X_{ijm} = \mu + \alpha_i + \pi_{m(i)} + \beta_j + \alpha\beta_{ij} + \beta\pi_{jm(i)} + \varepsilon_{ijm}$$

Die Parameter $\beta\pi_{jm(i)}$ und ε_{ijm} sind nicht unterscheidbar, sie bilden bei der Quadratsummenzerlegung das Residual, bzw. die Varianz von ε_{ijm} ist nicht einzeln schätzbar.

Quadratsummenzerlegung

Die Quadratsummenzerlegung kann durch das folgende Baumdiagramm veranschaulicht werden:



$$
\begin{aligned}
QS_{tot} &= \sum_i \sum_j \sum_m (X_{ijm} - \bar{X}_{\cdots})^2 \\
QS_{zwVpn} &= q \sum_i \sum_m (\bar{X}_{i \cdot m} - \bar{X}_{\cdots})^2 \\
QS_{innVpn} &= QS_{tot} - QS_{zwVpn}
\end{aligned}
$$

$$
\begin{aligned}
QS_A &= qn \sum_i (\bar{X}_{i\cdot\cdot} - \bar{X}_{\cdots})^2 \\
QS_{Vp} &= QS_{zwVpn} - QS_A
\end{aligned}
$$

QS_{Vp} wird häufig auch mit $QS_{VpInnGr}$ (Versuchspersonen innerhalb der Gruppen) bezeichnet.

$$
QS_B = pn \sum_j (\bar{X}_{\cdot j \cdot} - \bar{X}_{\cdots})^2
$$

$$
\begin{aligned}
QS_{zwAB} &= n \sum_i \sum_j (\bar{X}_{ij\cdot} - \bar{X}_{\cdots})^2 \\
QS_{AB} &= QS_{zwAB} - QS_A - QS_B
\end{aligned}
$$

$$
QS_{BVp} = QS_{innVp} - QS_B - QS_{AB}
$$

QS_{BVp} wird häufig auch mit QS_{res} bezeichnet.

Prüfgrößen

$$
F_A = \frac{MQ_A}{MQ_{Vp}}
$$

$$
\begin{aligned}
df_1 &= p - 1 \\
df_2 &= p(n - 1)
\end{aligned}
$$

$$F_B = \frac{MQ_B}{MQ_{BVp}}$$

$$df_1 = q-1$$
$$df_2 = (p-1)(q-1)(n-1)$$

$$F_{AB} = \frac{MQ_{AB}}{MQ_{BVp}}$$

$$df_1 = (p-1)(q-1)$$
$$df_2 = (p-1)(q-1)(n-1)$$

Zusammenfassung

Effekt	wird getestet gegen
A	Vp
B	BVp
AB	BVp

Dreifaktorieller Versuchsplan mit Messwiederholungen auf einem Faktor

Dreifaktorieller Versuchsplan mit Messwiederholungen auf einem Faktor

- Faktor A (fix) mit p Stufen
- Faktor B (fix) mit q Stufen
- Faktor C (fix) mit r Stufen
- Für jede AB-Kombination eine unabhängige Gruppe à n Versuchspersonen
- Der Vp-Faktor ist unter A und B genestet.
- Jede Versuchsperson durchläuft alle C-Bedingungen

Schema

		c_1	c_2	$\dots$	c_r
a_1	b_1	Gruppe 11	Gruppe 11	$\dots$	Gruppe 11
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	b_q	Gruppe 1q	Gruppe 1q	$\dots$	Gruppe 1q
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_p	b_1	Gruppe p1	Gruppe p1	$\dots$	Gruppe p1
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	b_q	Gruppe pq	Gruppe pq	$\dots$	Gruppe pq

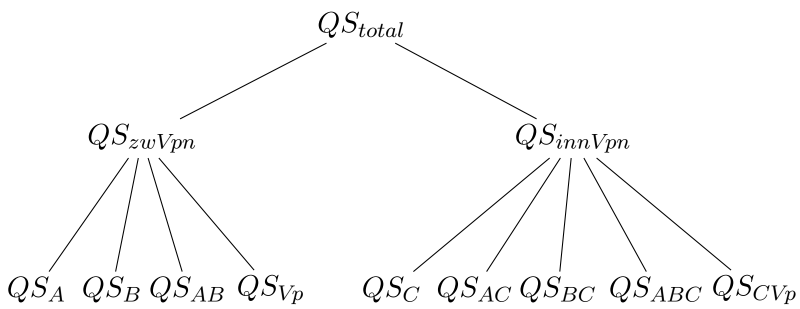
Strukturmodell

$$\begin{aligned} X_{ijkm} &= \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \pi_{m(ij)} + \gamma_k \\ &\quad + \alpha\gamma_{ik} + \beta\gamma_{jk} + \alpha\beta\gamma_{ijk} + \gamma\pi_{km(ij)} + \varepsilon_{ijkm} \end{aligned}$$

Die Parameter $\gamma\pi_{km(ij)}$ und ε_{ijkm} sind nicht unterscheidbar, sie bilden bei der Quadratsummenzerlegung das Residual.

Quadratsummenzerlegung

Die Quadratsummenzerlegung kann durch das folgende Baumdiagramm veranschaulicht werden:



QS_{CVp} wird häufig auch mit QS_{res} bezeichnet.

Testgrößen

Effekt	wird getestet gegen
A	Vp
B	Vp
AB	Vp
C	CVp
AC	CVp
BC	CVp
ABC	CVp

Dreifaktorieller Versuchsplan mit Messwiederholungen auf zwei Faktoren

Dreifaktorieller Versuchsplan mit Messwiederholungen auf zwei Faktoren

- Faktor A (fix) mit p Stufen
- Faktor B (fix) mit q Stufen
- Faktor C (fix) mit r Stufen
- Für jede A-Bedingung eine unabhängige Gruppe à n Versuchspersonen
- Der Vp-Faktor ist unter A genestet.
- Jede Versuchsperson durchläuft alle BC-Bedingungen

Schema

	b_1			$\dots$	b_q		
	c_1	$\dots$	c_r	$\dots$	c_1	$\dots$	c_r
a_1	Gr. 1	$\dots$	Gr. 1	$\dots$	Gr. 1	$\dots$	Gr. 1
a_2	Gr. 2	$\dots$	Gr. 2	$\dots$	Gr. 2	$\dots$	Gr. 2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_p	Gr. p	$\dots$	Gr. p	$\dots$	Gr. p	$\dots$	Gr. p

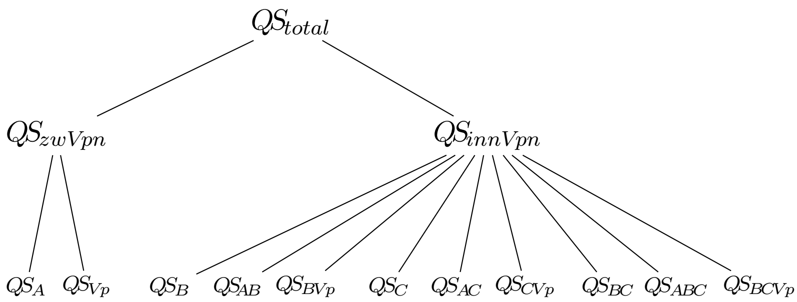
Strukturmodell

$$\begin{aligned} X_{ijkm} &= \mu + \alpha_i + \pi_{m(i)} + \beta_j + \alpha\beta_{ij} + \beta\pi_{jm(i)} + \gamma_k \\ &\quad + \alpha\gamma_{ik} + \gamma\pi_{km(i)} + \beta\gamma_{jk} + \alpha\beta\gamma_{ijk} + \beta\gamma\pi_{jkm(i)} + \varepsilon_{ijkm} \end{aligned}$$

Die Parameter $\beta\gamma\pi_{jkm(i)}$ und ε_{ijkm} sind nicht unterscheidbar, sie bilden bei der Quadratsummenzerlegung das Residual.

Quadratsummenzerlegung

Die Quadratsummenzerlegung kann durch das folgende Baumdiagramm veranschaulicht werden:



QS_{BCVp} wird häufig auch mit QS_{res} bezeichnet.

Testgrößen

Effekt	wird getestet gegen
A	Vp
B	BVp
C	CVp
AB	BVp
AC	CVp
BC	BCVp
ABC	BCVp

Testgrößen bei Versuchsplänen mit fixen Effekten

Testgrößen bei Versuchsplänen mit fixen Effekten

Die Testgröße für einen bestimmten Effekt ergibt sich auf folgende Weise:

- Schreibe den zu testenden Effekt hin.
- Hänge hinter diesen Ausdruck „Vp“.

- Streiche aus dem Ausdruck alle Gruppenfaktoren.
- Das Ausdruck, der dann übrigbleibt, ist die Testgröße für den Effekt.

Beispiel

Vierfaktorieller Versuchsplan mit A und B als Gruppenfaktoren, C und D als Messwiederholungsfaktoren.

Gegen welchen Ausdruck ist BCD zu testen?

BCD → BCDVp → **CDVp**.

Lateinisches Quadrat

Versuchsplan

Lateinisches Quadrat

- Ein lateinisches Quadrat ist ein dreifaktorieller, unvollständiger, ausbalancierter Versuchsplan.
- Die Anzahl der Stufen muss bei allen drei Faktoren gleich sein (p).
- Es wird eine quadratische Anordnung erstellt mit den Stufen des ersten Faktors (z.B. A) als Zeilen und den Stufen des zweiten Faktors (z.B. B) als Spalten.
- Die Stufen des dritten Faktors (z.B. C) werden nun als Zellen in das Schema so eingetragen, dass jede Stufe dieses Faktors in jeder Zeile und in jeder Spalte genau einmal vorkommt.

Beispiel ($p = 3$)

	a_1	a_2	a_3
b_1	c_1	c_2	c_3
b_2	c_2	c_3	c_1
b_3	c_3	c_1	c_2

In der folgenden Tabelle sieht man, dass es sich um einen unvollständigen Versuchplan handelt[31].

[31] realisierte Kombinationen sind mit einem „+“ gekennzeichnet

	c_1			c_2			c_3		
	a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3
b_1	+				+				+
b_2			+	+				+	
b_3		+				+	+		

Aufbau eines lateinischen Quadrats

- Die sogenannte Standardform eines lateinischen Quadrats ist ein solches, bei dem die erste Zeile und die erste Spalte jeweils durch die Reihe c_1, c_2, $\ldots$, c_p gebildet wird.
- Allgemein gilt: werden in einem lateinischen Quadrat Zeilen und/oder Spalten vertauscht, so ergibt sich ein neues lateinisches Quadrat.
- Die Anzahl der möglichen lateinischen Quadrate wird mit wachsendem p schnell sehr groß (bei 3 Stufen sind es 12, bei 4 bereits 576 und z.B. bei 6 Stufen 812851200).
- Idealerweise sollte der Experimentator für seinen Versuchsplan aus der Menge der möglichen lateinischen Quadrate eines zufällig ziehen.

Vor- und Nachteile

Vor- und Nachteile des lateinischen Quadrats

Vorteile

- Ein lateinisches Quadrat ist ein ökonomischer Versuchsplan, er erfordert einen geringeren Aufwand als ein vollständiger Versuchsplan: anstatt p^3 sind nur p^2 Zellen nötig.
- Ein lateinisches Quadrat eignet sich für die Kontrolle von zwei Störvariablen (in unserem Beispiel die Faktoren A und B) bei der Prüfung von Treatmenteffekten (Faktor C).
- Bevorzugt wird das lateinische Quadrat bei Messwiederholungsdesigns zur Kontrolle von Reihenfolgeeffekten eingesetzt.

Nachteile

- Ein lateinisches Quadrat kann nur eingesetzt werden, wenn Wechselwirkungen vernachlässigt werden können und nur Hauteffekte geprüft werden sollen.
- Bezüglich der Haupteffekte ist das lateinische Quadrat balanciert, aber nicht bezüglich der Interaktionen.

D.h. konkret wird jede Stufe eines der Faktoren wird mit allen Stufen eines anderen Faktors kombiniert, d.h. die drei zweidimensionalen Interaktionstabellen sind jeweils vollständig.

Hinsichtlich der Wechselwirkung ist das lateinische Quadrat jedoch nicht balanciert, z.B. war im obigen Beispiel die AB-Interaktionstabelle für c_1 wie folgt:

	a_1	a_2	a_3
b_1	+		
b_2			+
b_3		+	

Nähme man das volle Modell incl. aller Wechselwirkungen an, so ergäbe sich als Erwartungswert für den Mittelwert des Treatments c_1 ...

Erwartungswert für den Mittelwert des Treatments c_1:

$$\begin{aligned} E(\bar{X}_{..1.}) &= \mu + \gamma_1 + (\alpha\beta_{11} + \alpha\beta_{23} + \alpha\beta_{32})/3 \\ & \quad +(\alpha\beta\gamma_{111} + \alpha\beta\gamma_{231} + \alpha\beta\gamma_{321}/3 \end{aligned}$$

d.h. die Schätzungen der C-Haupteffekte wären mit AB-Wechselwirkungseffekten und dreifachen Wechselwirkungseffekten überlagert.

Also muss man die Wechselwirkungseffekte ausschließen.

Strukturmodell

Strukturmodell beim lateinischen Quadrat

Das volle dreifaktorielle Modell kann nicht geschätzt werden, da die Anzahl der Parameter zu groß ist:

- Dieses enthält $3p$ Haupteffekte, $3p^2$ zweifache Wechselwirkungseffekte und p^3 dreifache Wechselwirkungseffekte, d.h. unter Berücksichtigung der Restriktionen p^3 Parameter.
- Da aber nur p^2 Zellenmittelwerte vorliegen, muss das Strukturmodell vereinfacht werden, z.B. indem es neben μ und ε nur Haupteffekte enthält.

Quadratsummenzerlegung

Quadratsummenzerlegung

- Die Anzahl der Zellenmittelwerte ist p^2, d.h. die Anzahl der Freiheitsgrade Zwischen ist $p^2 - 1$.
- Die Anzahl der Freiheitsgrade für die 3 Haupteffekte beträgt $3(p-1)$.
- Es verbleiben also von den Freiheitsgraden Zwischen als Rest für alle Interaktionen zusammengenommen

$$df_{res} = (p^2 - 1) - 3(p-1) = p^2 - 3p + 2 = p(p-3) + 2$$

- Neben den Quadratsummen für die Haupteffekte ergibt sich also eine Residual-Quadratsumme, die entweder für einen globalen Test auf Interaktionen (falls eine Quadratsumme Innerhalb zur Verfügung steht) oder als Fehlerterm verwendet werden kann.

Beispiel

Beispiel (Winer)

Untersucht wurde die Wirksamkeit von drei Präparaten (Faktor B) auf drei Kategorien von Patienten (Faktor C) in drei verschiedenen Kliniken (Faktor A).

Versuchsplan

Die zufällige Ziehung aus der Menge der möglichen lateinischen Quadrate ergab das folgende Schema:

	b_1	b_3	b_2
a_2	c_3	c_2	c_1
a_1	c_2	c_1	c_3
a_3	c_1	c_3	c_2

Für jede der 9 Bedingungskombinationen wurde jeweils eine Gruppe von 4 Patienten untersucht.

Daten

	b_1	b_3	b_2
a_2	6,8,12,7	0,0,1,4	0,2,2,5
a_1	2,5,3,1	2,2,4,6	9,10,12,12
a_3	0,1,1,4	2,1,1,5	0,1,1,4

Mittelwerte

	b_1	b_3	b_2	
a_2	8.25	1.25	2.25	**3.92**
a_1	2.75	3.50	10.75	**5.67**
a_3	1.50	2.25	1.50	**1.75**
	4.17	**2.33**	**4.83**	**3.78**

Mittelwerte für Faktor C:

c_1	c_2	c_3
2.42	**1.83**	**7.08**

Varianzanalysetabelle

Quelle der Variation	**QS**	**df**	**MQ**	**F**
Zwischen	364.72	8		
A (Kliniken)	92.39	2	46.20	12.52
B (Präparate)	40.22	2	20.11	5.45
C (Kategorien)	198.72	2	99.36	26.93
Residuum	33.39	2	16.70	4.53
Innerhalb	99.50	27	3.69	
Total	464.22	35		

$F_{.95;2,27} = 3.35$

Lateinisches Quadrat bei Messwiederholungen

Lateinisches Quadrat bei Messwiederholungen

- Das lateinische Quadrat wird bevorzugt auch bei Messwiederholungsplänen eingesetzt.
- Damit sollen Reihenfolgeeffekte kontrolliert werden.
- Die Stufen des interessierenden Treatmentfaktors werden verschiedenen Versuchspersonen (Versuchspersonengruppen) in unterschiedlicher Reihenfolge dargeboten und zwar so, dass jede Stufe genau einmal an erster Stelle, genau einmal an zweiter Stelle usw. auftaucht.
- Die drei Faktoren sind also:
  1. Darbietungsreihenfolge (Gruppenfaktor)
  2. zeitliche Position (Messwiederholungsfaktor)
  3. Treatment (Messwiederholungsfaktor)

Strukturmodell

Bezeichnen wir den Gruppenfaktor mit A, den Positionsfaktor mit B und den Treatmentfaktor mit C, so ergibt sich das folgende Strukturmodell:

$$\begin{aligned} X_{ijkm} &= \mu + \alpha_i + \pi_{m(i)} + \beta_j + \gamma_k + (\beta\gamma_{jk}) \\ &\quad + \beta\pi_{jm(i)} + \gamma\pi_{km(i)} + \beta\gamma\pi_{jkm(i)} + \varepsilon_{ijkm} \end{aligned}$$

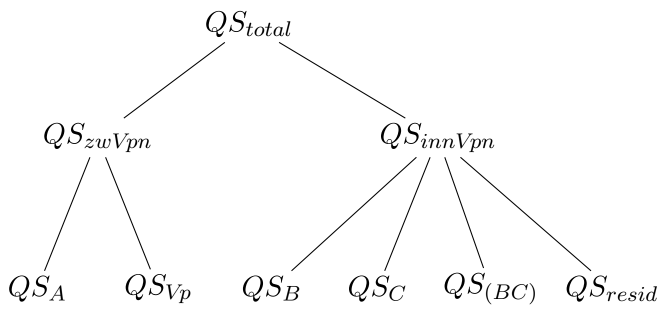
Der Parameter $\beta\gamma_{jk}$ ist eingeklammert, da über ihn nur partielle Information vorliegt.

Die Parameter $\beta\pi_{jm(i)}$, $\gamma\pi_{km(i)}$ und $\beta\gamma\pi_{jkm(i)}$ (Wechselwirkungen der Messwiederholungsfaktoren mit dem Versuchspersonenfaktor) bilden zusammen das Residuum.

Die Varianz von ε lässt sich nicht schätzen, da keine Replikationen vorliegen.

Quadratsummenzerlegung

Die Quadratsummenzerlegung sieht dann wie folgt aus:



Testgrößen

Effekt	wird getestet gegen
A	Vp
B	resid
C	resid
BC	resid

Wenn das Modell gilt, sollten nur die Haupteffekte B und C signifikant werden.

Verallgemeinerungen

Verallgemeinerungen

- Über das einfache lateinische Quadrat hinaus gibt es noch eine Reihe von Möglichkeiten, in einem Experiment mehrere lateinische Quadrate zu kombinieren.
- Dabei verwendet man vorzugsweise sog. orthogonale lateinische Quadrate, das sind solche, bei denen jede Zellenkombination nur einmal auftritt.
- Werden zwei orthogonale lateinische Quadrate mit den Faktoren A,B,C und A,B,D (mit C und D als Treatmentvariablen und A und B als Kontrollvariablen) kombiniert, so spricht man von einem griechisch-lateinischen Quadrat.

Beispiel für orthogonale lateinische Quadrate

I

1	2	3
3	1	2
2	3	1

II

2	3	1
1	2	3
3	1	2

III

1	2	3
2	3	1
3	1	2

I und II sind nicht orthogonal zueinander:

12	23	31
31	12	23
23	31	12

I und III jedoch sind zueinander orthogonal:

11	22	33
32	13	21
23	31	12

Ungleiche Stichprobenumfänge

Einführung

- In der Varianzanalyse sind gleiche Stichprobenumfänge unbedingt empfehlenswert.
- Obschon man bei der Planung eines Experiments gleiche Stichprobenumfänge vorgesehen hat, ist dies nach Beendigung des Experiments aber nicht immer gegeben.

- Gründe dafür sind z.B.:
  - Ausfall von Versuchspersonen (Mortality)
  - Für manche Bedingungskombinationen sind Versuchspersonen schlecht oder überhaupt nicht verfügbar.
  - u.a.
- Manchmal ist gleicher Stichprobenumfang auch von vorneherein nicht gegeben, z.B. bei der Verwendung von intakten, vorgegebenen Gruppen oder bei Verwendung von nicht-experimentellen unabhängigen Variablen („Organismusvariablen").

Einfache Varianzanalyse

Einfache Varianzanalyse

Liegen ungleiche Stichprobenumfänge vor, so müssen Quadratsummen und Freiheitsgrade wie folgt berechnet werden:

$$
\begin{aligned}
QS_{zw} &= \sum_{i=1}^{p} n_i(\bar{X}_{i.} - \bar{X}_{..})^2 \\
QS_{inn} &= \sum_{i=1}^{p}\sum_{m=1}^{n_i}(X_{im} - \bar{X}_{i.})^2 \\
df_{zw} &= p - 1 \\
df_{inn} &= \sum_{i=1}^{p}(n_i - 1) = \sum_{i=1}^{p} n_i - p
\end{aligned}
$$

Auch bei der Prüfung spezieller Hypothesen (lineare Kontraste, multiple Mittelwertsvergleiche) sind zum Teil spezielle Formeln zu verwenden[32].

Ungewichtete Mittelwerte

Methode der ungewichteten Mittelwerte

- Die Abweichungen der Gruppenmittelwerte vom Gesamtmittelwert wurden in den bisher angegebenen Formeln mit dem Stichprobenumfang gewichtet.
- Falls die Fehlervarianzen in den Zellen gleich sind (siehe Voraussetzungen für den F-Test), wird mit dieser Methode jede quadratische Abweichung proportional zu ihrem Fehler gewichtet, d.h. mit σ_ε^2/n_i.

[32] Für nähere Informationen siehe Winer (1962, S. 96f.)

- Falls andererseits die (ungleichen) Stichprobenumfänge rein zufällig sind und in keinerlei Beziehung stehen mit der zu testenden Hypothese, wäre es angemessen, in die Berechnung der Quadratsumme Zwischen jeden Treatmentmittelwert mit gleichen Gewicht eingehen zu lassen.
- Dies ist die Methode der <u>ungewichteten Mittelwerte</u> (unweighted means).

Bei der Methode der ungewichteten Mittelwerte sind folgende Formeln zu verwenden:

$$\begin{aligned}\bar{X}_{..} &= \frac{\sum \bar{X}_{i.}}{p} \\ QS_{zw} &= \bar{n}_h \sum_{i=1}^{p} (\bar{X}_{i.} - \bar{X}_{..})^2\end{aligned}$$

In diesem Fall gilt dann aber im allgemeinen nicht mehr

$$QS_{zw} + QS_{inn} = QS_{total}$$

$\bar{n}_h$ ist der harmonische Mittelwert, der wie folgt berechnet wird:

$$\bar{n}_h = \frac{p}{1/n_1 + 1/n_2 + \cdots + 1/n_p}$$

Beispiel (Winer)

Treatment 1	3,2,4,3,1,5	**3.00**
Treatment 2	7,8,4,10,6	**7.00**
Treatment 3	3,2,1,2,4,2,3,1	**2.25**
Treatment 4	10,12,8,5,12,10,9	**9.43**

Gesamtmittelwert

$$\begin{aligned}\text{gewichtet:} \quad \bar{X}_{..} &= \frac{6 \times 3.00 + \cdots + 7 \times 9.43}{26} = 5.35 \\ \text{ungewichtet:} \quad \bar{X}_{..} &= \frac{3.00 + 7.00 + 2.25 + 9.43}{4} = 5.42\end{aligned}$$

Quadratsummenzerlegung

QS Total	313.12
QS Innerhalb	73.21
QS Zwischen	239.95

Quadratsumme Zwischen nach der Methode der ungewichteten Mittelwerte:

$$\bar{n}_h = \frac{4}{1/6 + 1/5 + 1/8 + 1/7} = 6.3$$

$$\begin{aligned} QS_{zw} &= 6.3[(3.00 - 5.42)^2 + \cdots + (9.43 - 5.42)^2] \\ &= 217.24 \end{aligned}$$

Mehrfaktorielle Varianzanalyse

Mehrfaktorielle Varianzanalyse

- Auch bei einer mehrfaktoriellen Varianzanalyse kann man zwischen der Methode der gewichteten und ungewichteten Mittelwerte wählen.
- Im allgemeinen bedient man sich der Methode der gewichteten Mittelwerte.
- Darüberhinaus unterscheidet man zwischen orthogonalen und nicht-orthogonalen Varianzanalysen: bei einer orthogonalen Varianzanalyse sind die Zellenhäufigkeiten proportional zu den Randhäufigkeiten der einzelnen Faktoren verteilt.

Beispiel

Den Unterschied zwischen der Auswertung mit gewichteten und einer solchen mit ungewichteten Mittelwerten verdeutlicht das folgende fiktive Zahlenbeispiel (Howell & McConaughy, 1982):

	Krankenhaus (Faktor B)	
	1	2
Entbindungs-station (a_1)	$\bar{x}_{11\cdot} = 3.0$ $n_{11} = 10$	$\bar{x}_{12\cdot} = 2.6$ $n_{12} = 5$
Geriatrische Station (a_2)	$\bar{x}_{21\cdot} = 20.5$ $n_{21} = 4$	$\bar{x}_{22\cdot} = 21.0$ $n_{22} = 12$

AV: Verweildauer des Patienten

Mittelwerte

Krankenhaus	1	2	Gesamt
gewichtet	8.00	15.59	12.161
ungewichtet	11.75	11.80	11.775

Station	1	2	Gesamt
gewichtet	2.87	20.875	12.161
ungewichtet	2.80	20.750	11.775

Frage: Gibt es zwischen den Krankenhäusern einen Unterschied?

Die Antwort hängt davon ab, welche Mittelwerte (gewichtet oder ungewichtet) man betrachtet:

- Wenn es um die Güte der Versorgung der Patienten geht, sind die ungewichteten Mittelwerte zu nehmen.
- Wenn es um den Bedarf an Fernsehgeräten geht, sind die gewicheten Mittelwerte die angemessenere Wahl.

Die F-Werte für die Effekte sind:

Modell	F_A	F_B	F_{AB}
ungewichtete Mittelwerte	2270.53	0.02	1.43
gewichtete Mittelwerte	2802.13	493.39	1.43

Orthogonale und nicht-orthogonale Varianzanalyse

- Bei einer orthogonalen Varianzanalyse sind die Korrelationen der Zellenhäufigkeiten gleich 0.
- Dies ist dann der Fall, wenn die bedingten Verteilungen der Zellenhäufigkeiten proportional der Verteilung der entsprechenden Randhäufigkeiten sind.
- Bei einer orthogonalen Varianzanalyse sind die Effekte voneinander unabhängig.

- Nicht so bei einer nicht-orthogonalen Varianzanalyse: hier sind die Haupt- und die Wechselwirkungseffekte nicht unabhängig.
- Es gibt verschiedene Strategien, die einzelnen Effekte zu testen:
  - Sequentielle Zerlegung der Effekte: zuerst die Haupteffekte der Reihe nach, dann die zweifachen Wechselwirkungen usw. (SPSS: `sstype(1)`) Hier kommt es auf die Reihenfolge an, in der die einzelnen Effekte geschätzt werden (hierarchisches Modell).
  - Partielle QS: der Effekt eines Faktors wird berechnet, nachdem alle aunderen Effekte auspartialisiert wurden (SPSS: `sstype(3)`, bzw. unique).

Beispiel

...

Kovarianzanalyse

Einsatzzweck

Fehler in der Versuchsplanung

Der Fehler, von dem bei den einzelnen varianzanalytischen Modellen jeweils die Rede ist, besteht genau genommen aus verschiedenen Komponenten:

- Messfehler, d.h. die Variabilität in der Messung selbst
- Individuelle Variabilität

Kontrolltechniken

Es gibt nun zwei Methoden, den Fehler (experimental error) zu kontrollieren:

- direkt z.B. durch Bildung homogener Blöcke, Verwendung von Messwiederholungen etc.
- statistisch z.B. durch Einsatz der Kovarianzanalyse

Kontrolle der individuellen Variabilität

Individuen unterscheiden sich hinsichtlich vieler Variablen (Intelligenz, Motivation, Vorerfahrung usw.), die mit den Treatments nichts zu tun haben, aber einen Einfluss auf ihre Scores in der abhängigen Variable haben können.

Nun kann man Störvariablen als zusätzliche unabhängige Variablen in die Analyse einbeziehen (z.B. durch Post-Blocking).

Bei intervallskalierten metrischen Störvariablen gibt es eine ähnliche, aber bessere Kontrolltechnik: die Kovarianzanalyse.

Grundidee der Kovarianzanalyse

Durch Erfassung von Störvariablen versucht man, die Scores um irrelevante Fehlerkomponenten zu bereinigen.

Bei der Kovarianzanalyse werden also eine oder mehrere Störvariablen als Kontrollvariablen erfasst und aus der abhängigen Variable auspartialisiert.

Vor- und Nachteile

Damit lässt sich die Teststärke steigern.

Als Nachteil steht dem gegenüber, dass die Kontrollvariable mit erhoben werden muss.

Weiterer Einsatzzweck

Neben der Verringerung der Fehlervarianz wird die Kovarianzanalyse auch eingesetzt, um apriori Unterschieden zwischen Gruppen auszugleichen.

Randomisierungstechniken sind nicht immer anwendbar, z.B. dann nicht, wenn man intakte Gruppen verwenden muss.

Einfache Kovarianzanalyse

Einfache Kovarianzanalyse

Modell

Im Unterschied zu den bisherigen Formeln soll die abhängige Variable (wie die Kriteriumsvariable bei der Regression) mit Y bezeichnet werden. Die Kontrollvariable sei - wie bei der Regressionsrechnung üblich - mit X bezeichnet.

Das Strukturmodell der einfachen Varianzanalyse ist wie folgt:

$$Y_{im} = \mu + \alpha_i + \varepsilon_{im}$$

Durch eine lineare Regression wollen wir nun ε mit Hilfe von X schätzen:

$$\hat{\varepsilon}_{im} = \beta(X_{im} - \bar{X}_{..})$$

β ist der Regressionskoeffizient von ε auf die X-Abweichungen.

Die additive Konstante ist 0, da die Mittelwert von ε und von $X_{im} - \bar{X}_{..}$ gleich 0 sind.

Wir zerlegen den Fehler also in zwei Komponenten:

$$\begin{aligned}\varepsilon_{im} &= \hat{\varepsilon}_{im} + \varepsilon'_{im} \\ &= \beta(X_{im} - \bar{X}_{..}) + \varepsilon'_{im}\end{aligned}$$

Die ε' sind die Residuen der Regression.

Für die Varianzen gilt:

$$\sigma^2_{\varepsilon} = \sigma^2_{\hat{\varepsilon}} + \sigma^2_{\varepsilon'}$$

Das Strukturmodell der einfachen Kovarianzanalyse ist dann:

$$Y_{im} = \mu + \alpha_i + \beta(X_{im} - \bar{X}_{..}) + \varepsilon'_{im}$$

Den Fehler ε' wollen künftig wieder mit ε bezeichnen.

Voraussetzungen der Kovarianzanalyse

Zunächst gelten für die Fehler dieselben Voraussetzungen wie bei der Varianzanalyse, nämlich

- Unabhängigkeit
- Normalverteilung mit Mittelwert 0 und homogenen Varianzen

Zusätzlich muss noch gelten:

- Linearität der Regression von Y auf X
- Homogene Regressionskoeffizienten für die verschiedenen Gruppen, d.h. $\beta_1 = \beta_2 = \cdots = \beta_p$.
- Es gibt keine Treatmenteffekte für die Kontrollvariable.

Quadratsummenzerlegung

Quadratsummenzerlegung

Bei der Kovarianzanalyse werden gleichzeitig mehrere Regressionen betrachtet:

1. Regression Total, d.h. Regression aller Y_{im} auf alle X_{im} (Regressionskoeffizient: b_{total})
2. Regression Zwischen, d.h. Regression der $\bar{Y}_{i.}$ auf die $\bar{X}_{i.}$ (Regressionskoeffizient: b_{zw})
3. Pro Gruppe eine Regression von Y auf X (Regressionskoeffizienten: $b_{inn(i)}$).
4. Regression Innerhalb, die sich durch Poolen der Gruppenregressionen ergibt (Regressionskoeffizient: b_{inn})

- Die Residuen dieser Regressionen ergeben die korrigierten Werte, und deren Quadratsummen ergeben die korrigierten Quadratsummen.
- Dabei bekommt man je eine korrigierte Quadratsumme Total, Zwischen und Innerhalb.

- Im Unterschied zu die unkorrigierten Quadratsummen bezeichnen wir die korrigierten jeweils mit QS^*.

Rechenformeln

Es müssen nun mehrere Zerlegungen durchgeführt werden, nämlich für die Quadratsumme von Y, für die Quadratsumme von X und für die Summe der Abweichungsprodukte von X und Y.

Für die Formeln wollen wir eine neue, einheitliche Notation verwenden:

- SYY entspricht den Quadratsummen von Y (bisher mit QS bezeichnet).
- SXX entspricht den Quadratsummen von X.
- SXY entspricht den Abweichungsproduktsummen von X mit Y. Dies sind Maße für die Kovariation von X und Y.

Quadratsummenzerlegung für Y

$$
\begin{aligned}
SYY_{zw} &= n\sum_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 \\
SYY_{inn(i)} &= \sum_m(Y_{im} - \bar{Y}_{i\cdot})^2 \\
SYY_{inn} &= \sum_i SYY_{inn(i)} \\
SYY_{total} &= \sum_i\sum_m(Y_{im} - \bar{Y}_{\cdot\cdot})^2 \\
&= SYY_{zw} + SYY_{inn}
\end{aligned}
$$

Quadratsummenzerlegung für X

$$
\begin{aligned}
SXX_{zw} &= n\sum_i(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 \\
SXX_{inn(i)} &= \sum_m(X_{im} - \bar{X}_{i\cdot})^2 \\
SXX_{inn} &= \sum_i SXX_{inn(i)} \\
SXX_{total} &= \sum_i\sum_m(X_{im} - \bar{X}_{\cdot\cdot})^2 \\
&= SXX_{zw} + SXX_{inn}
\end{aligned}
$$

Produktsummenzerlegung für X, Y

$$\begin{aligned}
SXY_{zw} &= n\sum_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}) \\
SXY_{inn(i)} &= \sum_m (X_{im} - \bar{X}_{i\cdot})(Y_{im} - \bar{Y}_{i\cdot}) \\
SXY_{inn} &= \sum_i SXY_{inn(i)} \\
SXY_{total} &= \sum_i \sum_m (X_{im} - \bar{X}_{\cdot\cdot})(Y_{im} - \bar{Y}_{\cdot\cdot}) \\
&= SXY_{zw} + SXY_{inn}
\end{aligned}$$

Regressionskoeffizienten

- Regressionskoeffizient Innerhalb für Gruppe i:

$$b_{inn(i)} = \frac{SXY_{inn(i)}}{SXX_{inn(i)}}$$

- Regressionskoeffizient Innerhalb gepoolt:

$$b_{inn} = \frac{SXY_{inn}}{SXX_{inn}}$$

- Regressionskoeffizient Zwischen

$$b_{zw} = \frac{SXY_{zw}}{SXX_{zw}}$$

- Regressionskoeffizient Total

$$b_{total} = \frac{SXY_{total}}{SXX_{total}}$$

Korrelationskoeffizienten

Entsprechend kann man auch verschiedene Korrelationskoeffizienten betrachten:

- Korrelationskoeffizient Innerhalb für Gruppe i:

$$r_{inn(i)} = \frac{SXY_{inn(i)}}{\sqrt{SXX_{inn(i)}}\sqrt{SYY_{inn(i)}}}$$

- Korrelationskoeffizient Innerhalb gepoolt:

$$r_{inn} = \frac{SXY_{inn}}{\sqrt{SXX_{inn}}\sqrt{SYY_{inn}}}$$

- Korrelationskoeffizient Zwischen

$$r_{zw} = \frac{SXY_{zw}}{\sqrt{SXX_{zw}}\sqrt{SYY_{zw}}}$$

- Korrelationskoeffizient Total

$$r_{total} = \frac{SXY_{total}}{\sqrt{SXX_{total}}\sqrt{SYY_{total}}}$$

Korrigierte Quadratsummen

$$\begin{aligned}
QS^*_{total} &= SYY_{total} - \frac{SXY^2_{total}}{SXX_{total}} \\
QS^*_{inn} &= SYY_{inn} - \frac{SXY^2_{inn}}{SXX_{inn}} \\
QS^*_{zw} &= QS^*_{total} - QS^*_{inn}
\end{aligned}$$

Freiheitsgrade

$$\begin{aligned}
df^*_{total} &= np - 2 \\
df^*_{zw} &= p - 1 \\
df^*_{inn} &= p(n-1) - 1
\end{aligned}$$

Mittlere Quadrate

$$\begin{aligned}
MQ^*_{zw} &= \frac{QS^*_{zw}}{df^*_{zw}} \\
MQ^*_{inn} &= \frac{QS^*_{inn}}{df^*_{inn}}
\end{aligned}$$

Prüfgröße

$$
\begin{aligned}
F &= \frac{MQ^*_{zw}}{MQ^*_{inn}} \\
df_1 &= p - 1 \\
df_2 &= p(n-1) - 1
\end{aligned}
$$

Beispiel

...

Das allgemeine lineare Modell

Einführung

Varianzanalyse und lineare Regression

- In der Literatur und in Computerprogrammen wird die Varianzanalyse normalerweise als Spezialfall des ALM (allgemeines lineares Modell) betrachtet.
- Dabei wird eine UV jeweils in einen Satz von Codiervariablen überführt.
- Diese werden dann als Prädiktorvariablen in einem linearen Regressionsmodell verwendet.
- Die Güte der Modellanpassung dient als Grundlage für die Beurteilung der Wirksamkeit der Effekte.

Dummyvariablen

- Wir geben die Gruppenzugehörigkeit einer Beobachtung mit Hilfe geeigneter Codiervariablen X_1, X_2, ... an.
- Eine Möglichkeit besteht darin, für jede Gruppe eine eigene Variable zu verwenden, die den Wert 1 hat, wenn die Beobachtung zu dieser Gruppe gehört, und sonst den Wert 0. Diese Form der Codierung heißt *Dummycodierung.*
- Normalerweise enthält ein Regressionsmodell aber auch eine additive Konstante. In diesem Fall sind für die Codierung einer unabhängigen Variablen mit p Stufen nur $k = p - 1$ Codiervariablen nötig.

- So läuft der Vergleich von zwei Gruppen (t_{hom}) auf eine einfache lineare Regression hinaus und bei einem Faktor mit mehr als zwei Stufen auf eine multiple Regression.

Codierungsstandards beim ALM

Neben der Dummycodierung gibt es verschiedene in der Literatur und Praxis verwendete Codierverfahren:

- **Abweichungskontraste**: Abweichungen jeweils einer Stufe vom Mittel der anderen werden betrachtet; die letzte Stufe entfällt, sie stellt sozusagen die Vergleichsstufe dar.
- **Differenzkontraste**: Eine Stufe wird jeweils mit dem Mittel aller vorhergehenden Stufen verglichen.
- **Helmert-Kontraste**: Eine Stufe wird jeweils mit dem Mittel aller nachfolgenden Stufen verglichen.
- **Einfache Kontraste**: Jede Stufe wird mit einer Referenzstufe verglichen.
- **Wiederholungskontraste**: Jede Stufe außer der ersten wird mit der vorhergehenden verglichen.
- **Polynomiale Kontraste**: Es werden Koeffizienten orthogonaler Polynome verwendet (siehe Trendtests).

Strukturgleichung

Beim ALM wird mit Hilfe der Codiervariablen die folgende Regressionsgleichung für die Schätzung von Y angesetzt:

$$\hat{Y}_m = \beta_0 + \beta_1 X_{1m} + \beta_2 X_{2m} + \cdots + \beta_k X_{km}$$

Der tatsächliche Y-Wert setzt sich zusammen aus dem geschätzten Y-Wert und dem Fehler („Residuum“):

$$Y_m = \hat{Y}_m + \varepsilon_m$$

Y_m	abhängige Variable bei Vp m
$\hat{Y}_m$	geschätzter Wert von Y_m
X_{im}	Codiervariable i für Vp m
β_0	additive Konstante
β_i	Regressionsgewicht von X_i
ε_m	Fehler, Residuum

Wir behandeln nun die Varianzanalyse wie eine multiple lineare Regressionsanalyse und führen die Parameterschätzung und die Hypothesentests usw. wie bei dieser durch ...

Modelltest

Quadratsummenzerlegung

Für den Test auf Modellgeltung wird eine Quadratsummenzerlegung durchgeführt, d.h. die totale Quadratsumme in die Quadratsumme der Regression und die Quadratsumme der Residuen zerlegt.

Einfache Varianzanalyse

Einfache Varianzanalyse

Für die einfache Varianzanalyse mit p Gruppen und insgesamt n Versuchspersonen werden $k = p - 1$ Codiervariablen benötigt.

Die Quadratsummenzerlegung ist dann wie folgt: :

$$\begin{array}{ccccc} QS_{Total} & = & QS_{Regression} & + & QS_{Residuum} \\ \sum(y_m - \bar{y}.)^2 & = & \sum(\hat{y}_m - \bar{y}.)^2 & + & \sum(y_m - \hat{y}_m)^2 \end{array}$$

Freiheitsgrade

Die Freiheitsgrade werden in analoger Weise zerlegt:

$$\begin{array}{ccccc} df_{Total} & = & df_{Regression} & + & df_{Residuum} \\ n-1 & & k & & n-k-1 \end{array}$$

F-Test

Der F-Wert ist dann wie folgt zu berechnen:

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{R^2}{1-R^2} \cdot \frac{n-k-1}{k}$$

$$df_1 = k \qquad df_2 = n-k-1$$

R ist die multiple Korrelation, d.i. die Korrelation zwischen Y und $\hat{Y}$.

Zweifaktorielle Varianzanalyse

Zweifaktorielle Varianzanalyse

- Bei einem mehrfaktoriellen Versuchsplan mit unabhängigen Gruppen codiert man die Haupteffekte jedes Faktors mit einer Codiervariablen weniger als der Faktor Stufen hat.
- Interaktionen zwischen den Faktoren werden durch das elementweise Produkt der Codiervariablen der beteiligten Faktoren codiert.
- Das weitere Vorgehen ist das gleiche wie bei der multiplen Regression, mit dem Unterschied, dass neben dem globalen Modelltest vor allem auch Tests für die einzelnen Haupteffekte und Wechselwirkungen durchgeführt werden.

- Ist das Design orthogonal (dies ist z.B. bei gleichem n in den Gruppen der Fall), so sind die einzelnen Tests unabhängig.
- Ist es nicht orthogonal, so hängen die Ergebnisse der Tests von der Reihenfolge ab, in der die einzelnen Effekte berücksichtigt werden, sowie von der Methode der Quadratsummenberechnung.

Modell mit metrischen und Faktorvariablen

Modell mit metrischen und Faktorvariablen

- Mit Hilfe des ALM ist es auch kein Problem, eine Kovarianzanalyse durchzuführen, d.h. ein lineares Modell anzupassen, das sowohl Faktorvariablen als auch metrische Variablen enthält.
- Zum Beispiel könnte man modellieren, wie Y von den Treatments und zusätzlich noch von intervallskalierten Variablen abhängt.
- Die Prädiktorvariablen könnten dann sein:
  - Codiervariablen für den Haupteffekt des Treatmentfaktors.
  - Die Kovariaten stellen weitere Prädiktorvariablen dar.
  - Schließlich könnte man durch das Produkt der Codiervariablen mit den Kovariaten sogar eine Interaktion zwischen Treatment und Kovariaten modellieren.
- Das weitere Vorgehen ist dann wie bei jeder Regressionsanalyse.