

Datenanalyse Statistische Methoden

PD Dr. Dr. Dipl.-Psych. Guido Strunk

Themenübersicht

- Methodenlehre.
 - Was ist Wissenschaft? ([Seite](#))
 - Forschungsprozess. ([Seite](#))
 - Theorie und Hypothese. ([Seite](#))
 - Forschungsdesigns. ([Seite](#))
 - Hypothetische Konstrukte. ([Seite](#))
 - Forschungsmethoden. ([Seite](#))
 - Gütekriterien. ([Seite](#))
 - Population und Stichprobe. ([Seite](#))
 - Messtheorie. ([Seite](#))
 - Skala und Item. ([Seite](#))

- Statistik im engeren Sinn.
 - Datenverarbeitung und Deskription. ([Seite](#))
 - Inferenzstatistik – Grundlagen. ([Seite](#))
 - Einfache Testverfahren ([Seite](#))
 - T-Test. (SPSS, GStat, Excel)
 - χ^2 -Test.
 - Varianzanalyse (ANOVA).
 - Inferenzstatistik – Zusammenhangshypothesen. ([Seite](#))
- Software ([Seite](#))
 - Excel
 - GStat
 - Arbeiten mit SPSS
- Test-Finder ([Seite](#))
- Literatur

Was ist Wissenschaft?

Beantwortung von **Forschungsfragen** ...

.... mit Hilfe von **Bildern** und **Modellen** über die Realität.

Folgt akzeptierten Regeln und Grundprinzipien, v.a.

- Transparenz
- Nachvollziehbarkeit
- Überprüfbarkeit

Bezugnahme auf vorhandenes, bereits „gesichertes“ Wissen. Wissenschaft ist kumulativ.

- Arbeitet mit Literatur.
- Baut darauf auf.
- Vermeidet Fehler der Vergangenheit.

Beispiel 1

- Welche Auswirkungen hat die Einführung des Managementinstruments XYZ?
 - Die Frage ist vielleicht noch zu offen formuliert: Mit „Auswirkungen“ kann alles Mögliche gemeint sein (z. B. Widerstand der Beschäftigten, Kostenvorteile, Flexibilität, übersichtliche Darstellung von Entscheidungen, Veränderungen in der Zahl von Krankenständen etc.).
 - Um diese offene Frage empirisch beantworten zu können müsste die Situation vor der Einführung und die nach der Einführung verglichen werden.
 - Dabei würde man auf alle auftretenden „Auswirkungen“ gleichermaßen achten müssen. Die Studie wäre daher explorativ entdeckend.
 - **Die Antwort auf die Frage hätte die Form einer „Beschreibung“.**

Beispiel 2

- Wieso kommt es durch die Einführung des Managementinstruments XYZ zu einer Benachteiligung von Frauen?
 - Die Frage enthält eine Behauptung. Es muss mit Vorstudien im Theorieteil belegt werden, dass diese Behauptung auch tatsächlich stichhaltig ist.
 - Im Gegensatz zum ersten Beispiel ist hier eine Auswirkung konkret benannt. Es geht nun darum die Gründe dafür zu klären.
 - Man könnte im theoretischen Teil der Arbeit mögliche Gründe aus Theorien ableiten.
 - In einem empirischen Teil könnten die vermuteten Gründe dann empirisch geprüft werden.
 - Die Forschungsfrage sucht also nach einer „Erklärung“ für ein bereits bekanntes Phänomen.

Beispiel 3

- Wird es durch die Einführung des Managementinstruments XYZ in Zukunft zu weniger Fehlzeiten kommen?
 - Die Frage zielt auf eine „Prognose“ ab.
 - Im Gegensatz zum ersten Beispiel ist hier eine Auswirkung konkret benannt. Es geht nun darum diese Auswirkung zu überprüfen.
 - Man könnte im theoretischen Teil der Arbeit mögliche Gründe für oder gegen die Hypothese der Fehlzeitenreduktion aus Theorien ableiten.
 - In einem empirischen Teil könnte die vermutete Reduktion der Fehlzeiten empirisch geprüft werden.
 - Die Forschungsfrage sucht nach Belegen für eine Prognose.

Beispiel 4

- Wie sollte das Managementinstrument XYZ implementiert werden, um der Benachteiligung von Frauen entgegenzuwirken?
 - Die Frage formuliert ein zu erreichendes Ziel.
 - Der Weg, um dieses Ziel zu erreichen, ist offen und soll mit wissenschaftlichen Methoden begründet werden.
 - **Es geht darum eine „Technologie“ zu entwickeln, die das Ziel erreicht.**
 - Im theoretischen Teil werden Erfahrungsberichte aus der Literatur ausgewertet und zunächst theoretische Möglichkeiten zur Zielerreichung formuliert.
 - In einem empirischen Teil könnte gezeigt werden, was von diesen Möglichkeiten funktioniert.

Beispiel 5

- Werden die positiven Auswirkungen, die in der Literatur benannt werden, mit der Einführung des Managementinstruments XYZ auch tatsächlich erzielt?
 - Die Frage stellt die versprochenen positiven Auswirkungen auf den Prüfstand.
 - **Die Antwort enthält also eine „Kritik bzw. Bewertung“.**
 - Man könnte im theoretischen Teil der Arbeit die versprochenen positiven Auswirkungen herausarbeiten. Diese gilt es ganz konkret zu benennen und klar zu definieren.
 - In einem empirischen Teil könnte geprüft werden ob die zuvor definierten Auswirkungen auch erreicht werden.
 - Die Studie stellt eine Art Qualitätsprüfung dar.

Beschreibung

- Was ist der Fall? Wie sieht die „Realität“ aus? (oder auch: Sieht die Realität wirklich so aus?)

Erklärung

- Warum ist etwas der Fall? Warum und unter welchen Bedingungen treten bestimmte Phänomene auf?

Prognose

- Was wird zukünftig der Fall sein? Wie wird etwas künftig aussehen? Welche Veränderungen werden eintreten?

Gestaltung/Technologie

- Welche Maßnahmen sind geeignet, um ein bestimmtes Ziel zu erreichen?

Kritik, Bewertung

- Wie ist ein bestimmter Zustand vor dem Hintergrund explizit genannter Kriterien zu bewerten?

(Nienhüser & Marcel, 1998, S. 6, 2003, S. 4)

- Konkrete Kennzahlen sind gefragt.
- Gegenstand lässt sich gut in Zahlen abbilden.
- Das Feld ist so bekannt, dass Kennwerte und Konstrukte soweit formuliert sind, dass man weiß, dass man sie messen, zählen kann.
- Komprimierte Darstellung durch statistische Kennwerte.
- Man hat Hypothesen und möchte diese konkret prüfen.

- Messung: Abbildung empirischer Gegebenheiten im Zahlenraum.
- Deskriptive Statistik: Zusammenfassung von Daten um diese mit wenigen Kennwerten zu beschreiben.
- Prüfende Statistik: Unterschieds- und Zusammenhangshypothesen können statistisch geprüft werden.

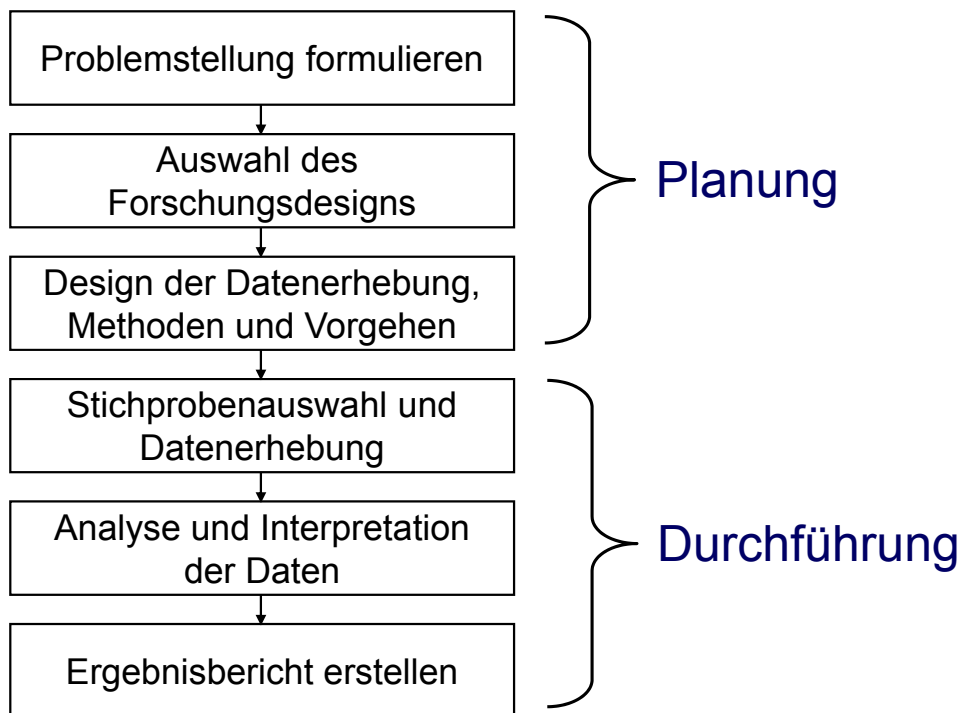
Die Art der Messung bestimmt die Methoden der deskriptiven und der prüfenden Statistik.

- Messung: Abbildung empirischer Gegebenheiten im Zahlenraum.
- Deskriptive Statistik: Zusammenfassung von Daten um diese mit wenigen Kennwerten zu beschreiben.
- Prüfende Statistik: Unterschieds- und Zusammenhangshypothesen können statistisch geprüft werden.

Die Art der Messung bestimmt die Methoden der deskriptiven und der prüfenden Statistik.

Forschungsprozess

Phasen im Forschungsprozess



Frei nach: Churchill und Iacobucci (2002, S. 56)

Fragen, für die es Antworten geben sollte I

Phase des Prozesses	Typische Fragen
<i>Problemstellung formulieren</i>	<ul style="list-style-type: none"> • Was ist das Ziel der Studie? Ein Problem zu lösen? Möglichkeiten aufzuzeigen? • Welche Informationen sind nötig, um die Ziele der Studie zu erreichen? • Werden zusätzliche, andere Informationen benötigt? • Wie werden die Ergebnisse der Studie wohl Verwendung finden?
<i>Auswahl des Forschungsdesigns</i>	<ul style="list-style-type: none"> • Was ist eigentlich bereits bekannt? • Können konkrete Hypothesen formuliert werden? • Welche Arten von Fragen sollen konkret beantwortet werden? Beschreibung, Erklärung, Prognose, Gestaltung, Kritik, Bewertung? • Welche Art von Studie (qualitativ, quantitativ, Primärstudie, Sekundäranalysen etc.) kann die Forschungsfragen am besten beantworten?

Fragen, für die es Antworten geben sollte II

Phase des Prozesses	Typische Fragen
<i>Design der Datenerhebung, Methoden und Vorgehen</i>	<ul style="list-style-type: none">• Kann bereits existierendes Datenmaterial genutzt werden?• Was soll gemessen werden? Wie? Operationalisierung?• Welche Datenquellen stehen zur Verfügung?• Sind kulturelle, legale, ethische oder andere Faktoren bei der Datenerhebung zu berücksichtigen?• Können durch Befragungen „objektive“ Daten erwartet werden?• Wie sollte man befragen (offene Fragen, Antwortkategorien, Ratings etc.)?• Können Beobachtungen mit technischen Hilfsmitteln erleichtert werden?• Bei einer Beobachtung, was genau soll beobachtet werden?• Experiment oder Quasi-Experiment?

Fragen, für die es Antworten geben sollte III

Phase des Prozesses	Typische Frage
<i>Stichprobenauswahl und Datenerhebung</i>	<ul style="list-style-type: none">• Was ist die Ziel-Population?• Gibt es ein vollständiges Verzeichnis der Ziel-Population?• Ist es notwendig eine Stichprobe zu untersuchen?• Ist eine Zufallsstichprobe möglich und nötig?• Wie groß sollte die Stichprobe sein?• Wie wird die Stichprobe ausgewählt?• Wer erhebt die Daten?• Wie lange wird die Erhebung dauern?• Wie wird die Qualität der erhobenen Daten sichergestellt?
<i>Analyse und Interpretation</i>	<ul style="list-style-type: none">• Wie werden die erhobenen Daten kodiert?• Wer kodiert und wie wird die Qualität überwacht?• Welche Software soll genutzt werden?• Welche deskriptiven Darstellungen (Tabellen, Abbildungen, Kennwerte) werden benötigt?• Mit welchen (inferenz-)statistischen Methoden soll gearbeitet werden?

Fragen, für die es Antworten geben sollte IV

Phase des Prozesses	Typische Frage
<i>Ergebnisbericht erstellen</i>	<ul style="list-style-type: none">• Wer wird den Forschungsbericht lesen?• Wie kompliziert darf der Bericht werden?• In welcher Form soll der Bericht erstellt werden?• Soll es auch eine mündliche Präsentation geben?• Wie ist die mündliche Präsentation zu strukturieren?

Theorie und Hypothese

- Eine Theorie ist ein System von Begriffen, Definitionen und Hypothesen. Dieses System sollte in sich geordnet und widerspruchsfrei sein.
- Daraus folgt, dass
 - die Konstruktion und Überprüfung von Theorien zunächst auf einer rein formalen Ebene nach den Regeln der Logik erfolgt;
 - je nach erkenntnistheoretischer Richtung bestimmte Spielregeln zur Überprüfung von Theorien einzuhalten sind (z.B. Falsifikationsprinzip des *Kritischen Rationalismus*);
 - Theorien rein logische Gebäude bleiben. Auch dann, wenn sie empirisch fundiert sind, ergeben sich Probleme der Übertragbarkeit in Bezug auf das Verhältnis zwischen der Theoriesprache einerseits und der sogenannten Realität andererseits.

Hypothesen 1

Prüfbarkeit. Eine Hypothese blickt nach vorne. Sie ist eine Behauptung, die man prüfen kann. Eine Hypothese mag dem gesunden Menschenverstand widersprechen oder mit ihm übereinstimmen. Sie kann sich als richtig oder falsch erweisen. In jedem Fall führt sie jedoch zu einer empirischen Nachprüfung.

Beantwortbarkeit. Unabhängig von dem Ergebnis ist eine Hypothese eine Frage, die so gestellt ist, dass irgendeine Antwort darauf gegeben werden kann.

Systematische Skepsis. Sie ist ein Beispiel für die systematische Skepsis der Wissenschaft, für ihre Weigerung, irgendeine Behauptung ohne empirische Bestätigung anzuerkennen.

Hypothese und Theorie. Die Verifikation oder Falsifikation von Hypothesen wird in der Regel angestrebt, um bestimmte Elemente oder eine ganze Theorie auf den Prüfstand zu stellen. An einer Hypothese hängt also mehr als nur die Überprüfung einer x-beliebigen Vorhersage.

Hypothesen 2

Hypothesen sind Schlussfolgerungen aus einer Theorie, die die Brücke zur beobachtbaren „Wirklichkeit“ schlagen. Aus dem Zutreffen der Hypothese im Rahmen ihrer empirischen Prüfung kann dann die Bewährung der Theorie gefolgert werden.

Hypothesen sind im engeren Sinne nur dann „wissenschaftlich“, wenn sie sich in der empirischen Prüfung als *falsch* oder *wahr* herausstellen können.

Wissenschaftstheoretiker, sowohl der *Logischen Empiristen*, als auch der *Kritisch Rationalisten*, nutzen die prinzipielle empirische Überprüfbarkeit als Abgrenzungskriterium, um wissenschaftliches Arbeiten von metaphysischen Spekulationen zu unterscheiden.

Nicht empirisch überprüfbar ist z.B. die Existenz eines sich nicht einmischenden Gottes, die Existenz eines freien Willens, die Existenz des Fremdseelischen.

Hypothesen 3 Verifikation vs. Falsifikation

Die *logischen Empiristen* verlangen die *Verifikation* von Hypothesen und dadurch die Verifikation von Theorien.

Für (*allgemeine*) *Existenzaussagen* ist die Verifikation die einzig logische und die *effektivste* Vorgehensweise. Die Behauptung, dass es kleine grüne Männchen auf dem Mars gibt, kann durch das Vorzeigen nur eines kleinen grünen Männchens vom Mars bewiesen werden.

Die *kritischen Rationalisten* verlangen die Falsifikation von Hypothesen und darüber die Falsifikation von Theorien. Wissenschaftliche Erkenntnis muss prinzipiell falsifizierbar sein. Widersteht eine Theorie, eine Hypothese in einer Prüfung der Falsifikation, so gilt sie als *vorläufig bewährt*.

K. R. Popper proklamiert: „Wir finden das Richtige, indem wir das Falsche ausschließen.“

Hypothesen 4

Für eine Hypothese, die zur Falsifikation taugt, wird also eine *Schwachstelle der Theorie* gesucht und diese in der Realität geprüft.

Die Logik dahinter geht davon aus, dass (*allgemeine*) *Gesetzesaussagen* zwar beliebig oft verifiziert werden können, dass dies aber kein zwingender Beweis für die Gültigkeit der Gesetzesaussagen ist.

Die Behauptung, dass alle Schwäne weiß seien, ist durch jeden beobachteten weißen Schwan verifizierbar, aber damit nicht vollständig bewiesen. Ein einziger Gegenbeweis, ein einziger schwarzer Schwan, würde jedoch genügen, um die Hypothese als nicht zutreffend festzustellen.

2, 4, 6

Hypothesen 5

Wir können also zu einer Hypothese, die aus einer Theorie folgt zwei verschiedene Aussagen formulieren.

Die Verifikationsannahme führt zur so genannten Alternativ-Hypothese (H_1).

H_1 : „Alle Schwäne sind weiß.“

Die Falsifikationsannahme führt zur so genannten Null-Hypothese (H_0).

H_0 : „Es gibt zumindest einen schwarzen Schwan.“

Häufig ist die Null-Hypothese das einfache Gegenteil der Alternativ-Hypothese. Das ist aber nicht immer der Fall. Mitunter gibt es viele verschiedene Null-Hypothesen (z.B. grüne Schwäne) oder verschieden spezifische Null-Hypothesen (z.B. nicht-weiß vs. schwarz). Es gilt die Null-Hypothese zu wählen, die bei ihrem Scheitern logisch am stringentesten die Alternativ-Hypothese als bewährt erscheinen lässt.

In einem sauberen Forschungsdesign wird alles getan, um die Falsifikationsaussage, also die H_0 , zu verifizieren.

Arten von Alternativ-Hypothesen Übersicht

	Unterschieds-Hypothesen		Zusammenhangs-Hypothesen	
2 Gruppen / Variablen / Objekte	Einfache Gruppenvergleiche		Einfache Korrelation	
Mehr als 2 Gruppen / Variablen / Objekte	Multiple Gruppenvergleiche		Multiple Korrelation / Regression	
Hypothesentyp	gerichtet (1-seitig)	ungerichtet (2-seitig)	gerichtet (1-seitig)	ungerichtet (2-seitig)

Unterschiedshypothesen

- Zwei Objekte/Gruppen/Treatments
 - H_1 : A ist anders als B (ungerichtete Hypothese / 2-seitig).
 H_0 : $A = B$.
 - H_1 : $A > B$ (gerichtete Hypothese / 1-seitig).
 $H_{0.1}$: $A = B$.
 $H_{0.2}$: $A < B$.
 - H_1 : $A < B$ (gerichtete Hypothese / 1-seitig).
 $H_{0.1}$: $A = B$.
 $H_{0.2}$: $A > B$.

Unterschiedshypothesen

- Mehr als zwei Objekte/Gruppen/Treatments
 - H_1 : Zwischen A, B, C, D bestehen Unterschiede (keine Festlegung darauf, ob sich alle von allen unterscheiden sollen oder ob ein einziger paarweiser Unterschied genügt).
 - H_0 : $A = B = C = D$.

Unterschiedshypothesen

- Mehr als zwei Objekte/Gruppen/Treatments
 - H_1 : Zwischen A, B, C, D bestehen Unterschiede (alle unterscheiden sich von allen).
 - $H_{0.1}$: A = B
 - $H_{0.2}$: A = C
 - $H_{0.3}$: A = D
 - $H_{0.4}$: B = C
 - $H_{0.5}$: B = D
 - $H_{0.6}$: C = D

Bei mehr als zwei Objekten gilt es zu überlegen, ob A, B, C, D alle in Frage kommenden Objekte sind (*fixed factors*) oder ob es z.B. noch E und F gibt (*fixed factors mit beschränkter Aussage*). Es kann aber auch sehr viele Objekte geben und A, B, C, D sind eventuell nur eine zufällige Auswahl (*random factors*). Das hat eine Bedeutung für die Verallgemeinerbarkeit (Generalisierbarkeit) der Befunde.

Zusammenhangshypothesen

- Zwei Variablen
 - H_1 : x zeigt einen Zusammenhang mit y.
x *korreliert* mit y (ungerichtete Hypothese / 2-seitig).
 - H_0 : x zeigt keinen Zusammenhang mit y.
Die *Korrelation* zwischen x und y ist Null ($r = 0$).
 - H_1 : x korreliert *positiv* mit y. Also: je mehr x, desto mehr y. Und je weniger x, desto weniger y. Und die gleichen Aussagen mit x und y vertauscht (gerichtete Hypothese/ 1-seitig).
 - $H_{0.1}$: Die Korrelation zwischen x und y ist Null ($r = 0$).
 - $H_{0.2}$: Die Korrelation zwischen x und y ist negativ ($r < 0$).

- H_1 : x korreliert *negativ* mit y . Also: je mehr x , desto weniger y .
Und je weniger x , desto mehr y . Und die gleichen Aussagen mit x und y vertauscht (gerichtete Hypothese / 1-seitig).
- $H_{0.1}$: Die Korrelation zwischen x und y ist Null ($r = 0$).
- $H_{0.2}$: Die Korrelation zwischen x und y ist positiv ($r > 0$).

Arten von Alternativ-Hypothesen Zusammenhänge 3

Zusammenhangshypothesen

– Mehr als zwei Variablen

- Unterscheidung zwischen unabhängigen (uV) und abhängigen Variablen (aV).

Die uV ist die Variable, deren Auswirkung untersucht werden soll.
Die Alternativ-Hypothese geht davon aus, dass die aV von der uV abhängig ist.



- Unterscheidung zwischen Prädiktor/en und Kriterium.
Prädiktoren sind *mehrere* uV's, die in der Regel *eine* aV vorhersagen helfen.
- H_1 : Ist in der Regel ungerichtet (2-seitig): uV korreliert mit aV
- H_0 : Die Korrelation zwischen aV und uV ist Null ($r = 0$)

Arten von Alternativ-Hypothesen

Zusammenhänge 4

Zusammenhangshypothesen

– Mehr als zwei Variablen

- Unterscheidung zwischen Haupteffekten und Interaktionseffekten.

- Bei einem Haupteffekt besteht die Beziehung einer uV auf eine aV unabhängig von anderen uV's. (Z.B.: je schwerer die Krankheit, desto geringer die Patientenzufriedenheit. Wenn dies unabhängig vom Alter, vom Geschlecht und vom Heilungserfolg etc. gilt, liegt ein Haupteffekt vor.)
- Bei einem Interaktionseffekt zwischen zwei Variablen (Interaktion 1. Ordnung) ändert sich die Beziehung zwischen einer uV und einer aV je nach Ausprägung einer anderen uV. (Z.B.: mit der schwere der Krankheit und der Größe des Heilungserfolges wächst die Patientenzufriedenheit.)

Forschungsdesigns

Wichtige Begriffe

- **Unterscheidung zwischen unabhängigen (uV) und abhängigen Variablen (aV).**

Die uV ist die Variable, deren Auswirkung untersucht werden soll. Die Alternativ-Hypothese geht davon aus, dass die aV von der uV abhängig ist.

uV \longrightarrow aV

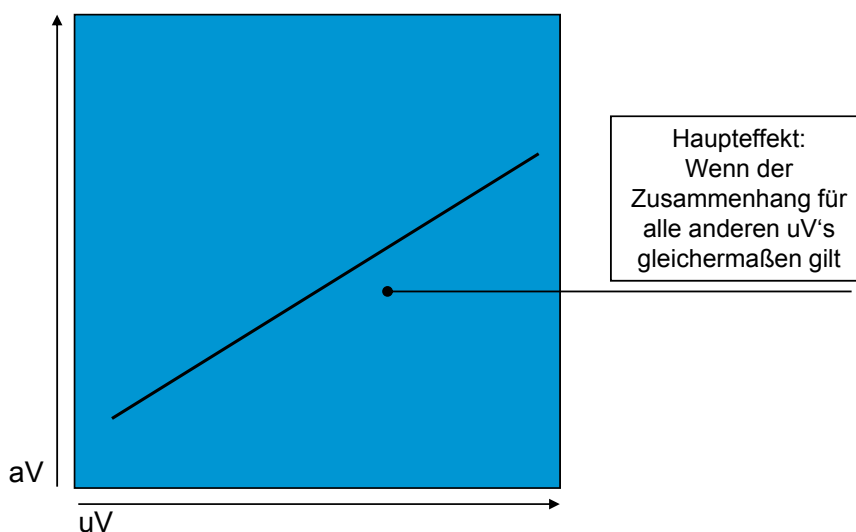
- **Unterscheidung zwischen Prädiktor/en und Kriterium.**

Prädiktoren sind **mehrere** uV's, die in der Regel **eine** aV (Kriterium) vorhersagen helfen.

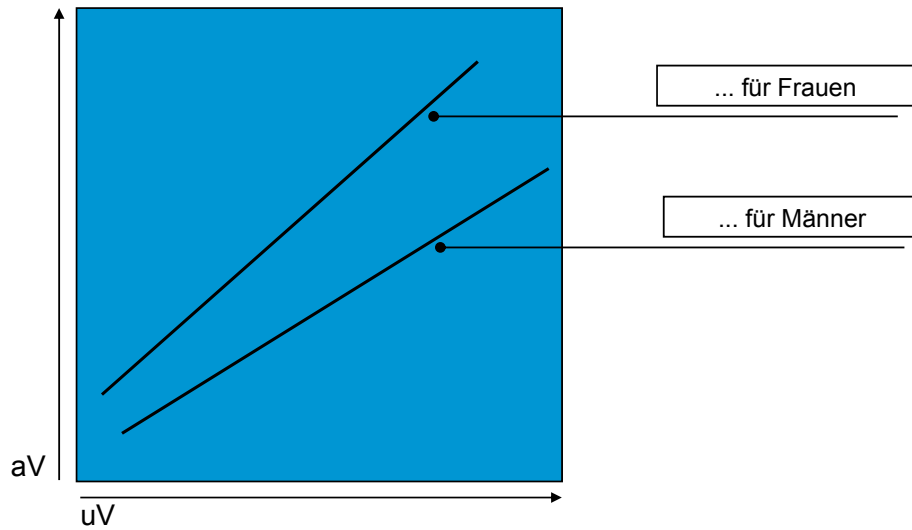
- **Unterscheidung zwischen Haupteffekten und Interaktion.**

Zeigt eine uV unabhängig von anderen uVs immer die gleiche Wirkung bei der aV, so heißt das Haupteffekt. Eine Interaktion bedeutet unterschiedliche Wirkungen bei Kombination der uVs auf die aV.

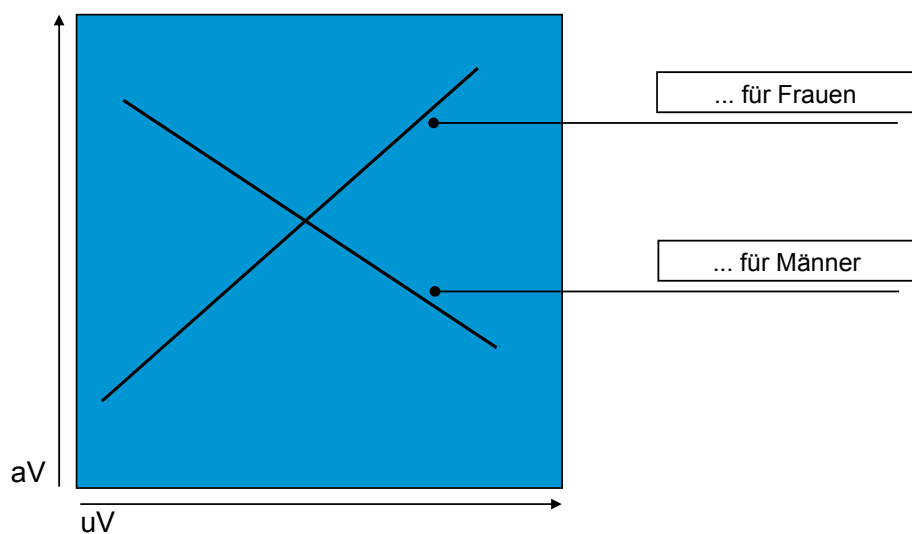
Typen von Effekten 1 Haupteffekt



Typen von Effekten 2 Ordinale Interaktion



Typen von Effekten 3 Disordinale Interaktion



Wichtige Begriffe

- **Unterscheidung zwischen ein- und mehrfaktoriellem Design.**

Liegt nur eine uV vor, die mehrere klare Abstufungen besitzt (z.B. Berufe), so handelt es sich um ein einfaktorielles Design. Mehrere uVs bilden ein mehrfaktorielles Design.

- **Unterscheidung zwischen festen und zufälligen Faktoren.**

Bei mehrstufigen uVs gilt es zu überlegen, ob alle in Frage kommenden Stufen untersucht werden (***fixed factors***) oder ob es noch andere nicht untersuchte Stufen gibt (***fixed factors mit beschränkter Aussage***). Es kann aber auch sehr viele Stufen geben und eventuell wird nur eine zufällige Auswahl (***random factors***) untersucht. Das hat eine Bedeutung für die Verallgemeinerbarkeit (Generalisierbarkeit) der Befunde.

Hypothetische Konstrukte

Echte und operationale Definitionen 1

- Definitionen sind wichtige Elemente wissenschaftlicher Theorien.
- Eine *Realdefinition* ist die Erklärung eines Begriffs, die zum Ziel hat festzustellen, wie der Begriff im Rahmen der Arbeit verwendet wird.
- Eine *Nominaldefinition* ist die explizite Einführung eines Terminus, die zum Ziel hat, für einen Begriff einen Terminus festzusetzen, ihn also durch Kopplung an ein Wort dauerhaft handhabbar zu machen.
- Definitionen sind *Identitäten*. Das *Definiendum* (das, was definiert wird) ist nach der Definition *identisch* mit dem *Definiens* (dem Definierenden). Damit gilt die Forderung der *Eliminierbarkeit*: Das Definiendum muss jederzeit durch das Definiens ersetzbar sein.

Echte und operationale Definitionen 2

- Forderung nach *Nicht-Kreativität*. Es darf nicht erst durch eine Definition eine Wahrheit erzeugt/bewiesen werden, die ohne sie unbeweisbar wäre.
- Für eine empirische Erhebung ist besonders die Definition der zu erhebenden Variablen bedeutsam. Dabei sind zwei Gruppen von Variablen zu unterscheiden:
 - *Echte Definition empirischer Variablen*. Die Definition betrifft Variablen, die direkt wahrnehmbar, zählbar, messbar sind und tatsächlich mit der beobachtbaren Variable „identisch“ sind. Wenn das Gehalt als zu versteuerndes Gehalt aus unselbständiger Arbeit laut Steuerbescheid definiert wird, ist die Definition eindeutig und bezeichnet eine Identität, die bei der Messung des Gehalts 1-zu-1 genutzt werden kann.

- **Echte operationale Definition.** Die Definition betrifft Variablen, die nicht direkt wahrnehmbar, zählbar, messbar sind. Es muss eine *Operation* durchgeführt werden, um eine interessierende Eigenschaft „hervorzulocken“. Z.B. ist die Wasserlöslichkeit eines Stoffes definiert über das Vorliegen einer vollständigen Lösung eines Stoffes, nachdem man ihn ins Wasser gegeben und lange genug umgerührt hat. Auch diese operationale Definition ist eine Identität.
- **Unechte operationale Definitionen.** In den Sozialwissenschaften werden häufig *hypothetische Konstrukte* als Elemente einer Theorie benutzt. Z.B. ist die Intelligenz ein solches hypothetisches Konstrukt. „Intelligenz“ ist eine gute „Erklärung“ für bestimmte psychische Phänomene. Wie aber kann Intelligenz definiert werden? Die Zahl der Operationen, um die Intelligenz „hervorzulocken“, ist unbegrenzt. Damit umfasst das hypothetische Konstrukt aber immer *mehr*, als die praktisch begrenzte operationale Definition. Intelligenz als hypothetisches Konstrukt ist immer mehr als der Intelligenztest misst.

Man erzählt, dass drei Schiedsrichter über die Frage des Pfeifens von Fouls uneins waren.

Der erste sagte: „Ich pfeife sie, wie sie sind.“

Der zweite sagte: „Ich pfeife sie, wie ich sie sehe.“

Der dritte und cleverste Schiedsrichter sagte:

„Es gibt sie überhaupt erst, wenn ich sie pfeife.“

(Simons, 1976, S. 29, zitiert in Weick 1985, S.9)



Forschungsmethoden

Strategien empirischer Forschung 1

- Explorative Datenanalyse:
 - Pilotstudie, Voruntersuchung; Hypothesengenerierung
- Deskriptive Datenanalyse:
 - Beschreibung, Klassifizierung, Bericht der beobachtbaren Daten; deskriptive Statistik (z.B. Verteilungsparameter, Mittelwerte, Streuungsmaße, Tabellen, Diagramme)
- Konfirmatorische Datenanalyse:
 - Hypothesentest, Suche nach Zusammenhängen, Schlüsse von der Stichprobe auf die Grundgesamtheit, Inferenzstatistik (Tests, Wahrscheinlichkeitsrechnung)
- Diagnostische Studien:
 - Erklärung von Einzelfällen; Rückgriff auf bereits bewährte Hypothesen
- Studien zur Technologieentwicklung:
 - Entwicklung und Test von Erhebungsinstrumenten, z.B. eines Fragebogens zur Patientenzufriedenheit

Strategien können unterschieden werden, ...

- nach der Zahl der berücksichtigten Ebenen:
 - Personen, Gruppen, Organisationen, Gesellschaft; Mikro-/Makroebene
- nach der Zahl der verwendeten Methoden:
 - Verfahren derselben Methode (z.B. Vorbereitung von schriftlichen Befragungen durch Interviews), Verfahren unterschiedlicher Methoden (z.B. Beobachtung und Befragung)
- nach der Zahl der Zeitpunkte:
 - Querschnittuntersuchungen, Längsschnittuntersuchungen
- nach der Originalität
 - Primäranalyse, Sekundäranalyse (zweite Analyse bereits abgeschlossener Untersuchungen), Replikationsstudie

- Methoden:
 - grundlegende Ausrichtung und Wege, wie man Modelle/Abbilder der Realität erzeugt
 - konstruieren ein Stück sozialer Realität (Selektivität)
- Verfahren:
 - Techniken der Datenerhebung, die im Rahmen einer Methode angewendet werden (konkrete „Varianten“ einer Methode)
- Instrumente:
 - Werkzeuge, standardisierte Elemente der Datenerhebung
 - Z.B. Meßinstrumente, Indizes, Skalen, Tests (Fragebögen)
- Regeln:
 - Richtlinien, die bei der konkreten Datenerhebung bzw. der Erstellung von Instrumenten zu beachten sind

Methoden-Verfahren- Instrumente-Regeln 2

<i>Methoden</i>	<i>Verfahren</i>	<i>Instrumente</i>	<i>Regeln</i>
Befragung	mündlich: strukturiert / unstrukturiert / Intensivinterview schriftlich: standardisiert nicht standardisiert / Tagebuch Delphi-Methode	Fragebogen Interview-Leitfaden Testverfahren ...	angemessene Sprache, keine Suggestivfragen, ...
Beobachtung	teilnehmend / nicht teilnehmend offen / verdeckt standardisiert / nicht standardisiert / reaktiv / non-reaktiv	Beobachtungs- leitfaden, Beobachtungs- schema, Notationssystem ...	Definition der kleinsten Beobachtungs- einheit, ...
Textanalyse	qualitativ / quantitativ interpretativ / messend Inhaltsanalyse Diskursanalyse ...	Kategorien- schema, Beispiel- sammlung ...	Definition von Bedeutungs- einheiten, Abgrenzung relevanter Text- stellen, Codierung

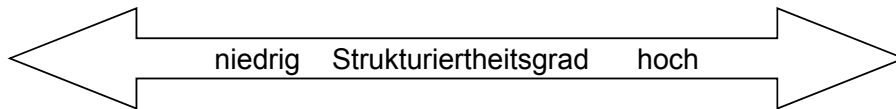
Methoden-Verfahren- Instrumente-Regeln 3

<i>Methoden</i>	<i>Verfahren</i>	<i>Instrumente</i>	<i>Regeln</i>
Experiment	Laborexperiment / Feldexperiment / Quasi-Experiment	nutzt Instrumente anderer Methoden bei gezielter Manipulation der Experimentalsituation	Isolierende Variation

- Bei **systematischen Beobachtungsverfahren** ermöglicht ein *Beobachtungsschema* die zielgerichtete Beobachtung und eine systematische Erfassung der Daten in Beobachtungsprotokollen.
- Das Beobachtungsschema umfasst:
 - die Auswahl der Beobachtungsgegenstände (Beobachtungssitem) (Ereignisse, Prozesse, Handlungen), die beobachtet werden sollen;
 - eine Zuordnung der Beobachtungssitem zu den Bedeutungen, die ihnen zugrunde liegen;
 - eine Klassifikation von Ereignissen und Handlungen mit ähnlichen Bedeutungen zu bestimmten Beobachtungskategorien. Diesen Kategorien werden im Beobachtungsprotokoll bestimmte Zeichen oder Zahlen zugeordnet (**Kodierung**).
- Ein Beobachtungsprotokoll beinhaltet Angaben zu Dauer, Ort, anwesenden Personen usw. und die Beobachtungssitem (Ereignisse, Prozesse, Verhaltensweisen), die bestimmten Beobachtungskategorien zugeordnet sind.

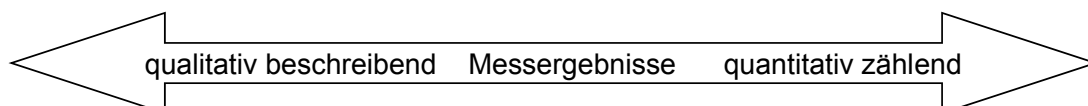
- *Manipulation* der unabhängigen Variablen
- *Messung* der abhängigen Variablen durch Instrumente anderer Methoden
- *Kontrolle* aller anderen eventuell relevanten Bedingungen durch:
 - Elimination von Störvariablen (sofern möglich)
 - Konstanthaltung der Störvariablen
 - Umwandlung von Störvariablen in weitere unabhängige Variablen
 - Parallelisierung von Experimental- und Kontrollgruppe („matching“)
 - Zufallszuweisung („Randomisierung“) der Versuchsteilnehmer zur Experimental- bzw. Kontrollgruppe

Strukturiertheitsgrad



- unstrukturiertes Interview
- nicht standardisierte Befragung
- nicht standardisierte Beobachtung
- qualitative Textanalyse
- Quasi-Experiment
- strukturiertes Interview
- standardisierte Befragung
- standardisierte Beobachtung
- quantitative Textanalyse
- Experiment

Messergebnisse



- unstrukturiertes Interview
- nicht standardisierte Befragung
- nicht standardisierte Beobachtung
- qualitative Textanalyse
- strukturiertes Interview
- standardisierte Befragung
- standardisierte Beobachtung
- quantitative Textanalyse
- Experiment / Quasi-Experiment

Vorsicht: Die Zuordnung von Verfahren ist nicht immer eindeutig.
Die von Instrumenten schon eher.

Gütekriterien

Gütekriterien 1: Intersubjektivität

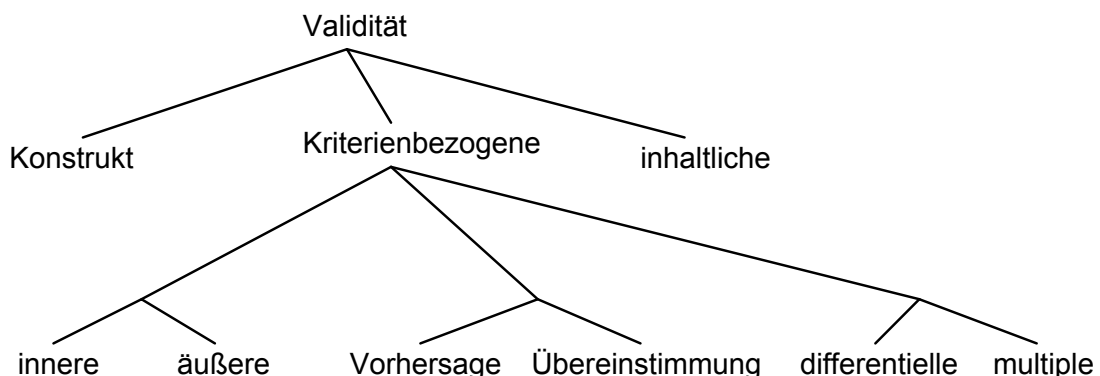
- Die Forderung nach *Intersubjektivität* entspricht im Wesentlichen der historisch älteren Forderung nach *Objektivität*.
- Theoretische Aussagen, Forschungsprozess und Ergebnisse müssen auch von anderen Personen (als den unmittelbar involvierten ForscherInnen) überprüft und nachvollzogen werden können („*Nachvollziehbarkeit*“).
- Daher sollten in der *Scientific Community* akzeptierte Methoden, Instrumente und Regeln Verwendung finden.
- Wichtig ist eine präzise Definition der zu erhebenden Variablen.

Gütekriterien 2: Reliabilität (Zuverlässigkeit)

- Das Konzept der **Reliabilität** geht davon aus, dass ein Merkmal „in Wirklichkeit“ eine bestimmte Ausprägung besitzt und dass eine Messung dieser Ausprägung verschieden genau durchgeführt werden kann. Die Genauigkeit ist die Reliabilität.
- Überprüfung der Reliabilität:
 - **Re-Test:** wiederholte Messung, gleiches Instrument, gleiche Objekte, *verschiedene Zeitpunkte*. *Erinnerungseffekte*.
 - **Parallel-Test:** wiederholte Messung, gleiche Objekte, *ähnliche Instrumente*. Eventuell *verschiedene Zeitpunkte*.
 - **Split-Half-Verfahren:** Instrument wird in zwei Hälften geteilt (z.B. Items eines Fragebogens) und die Ergebnisse der beiden Hälften werden verglichen.
 - **Innere Konsistenz (Cronbach-Alpha):** Nicht nur Hälften werden verglichen. Jedes Item wird mit jedem Item verglichen.

Gütekriterien 3: Validität (Gültigkeit) 1

- Die Validität gibt Antworten auf die folgenden Fragen: Wurde tatsächlich das gemessen, was man messen wollte? Wie groß ist die Übereinstimmung zwischen empirischer Messung und dem zu messenden Konstrukt?



Gütekriterien 3: Validität (Gültigkeit) 2

- **Konstruktvalidität**, fragt danach, wie gut bzw. passend ein Konstrukt tatsächlich erfasst wird (Überprüfung z.B. durch logische Analyse, Extremgruppenvergleiche, Experimente).
- **Inhaltliche Validität**: Beschreibt, ob ein Verfahren nach inhaltlichen Kriterien (Expertenurteile, per Augenschein bzw. face-Validität, auch logische Validität, triviale Validität genannt) erfasst, was er zu erfassen vorgibt. Dies ist bei hypothetischen Konstrukten nur schwer möglich. Daher sollte bei hypothetischen Konstrukten die Konstruktvalidität ermittelt werden. Kennwerte für die inhaltliche Validität gibt es nicht.
- **Kriterienbezogene Validität** wird durch eine Korrelation zu einem Kriterium empirisch bestimmt (empirische Validität).

Gütekriterien 3: Validität (Gültigkeit) 3

- Die **Innere Kriterienbezogene Validität** wählt als Kriterium einen schon bestehenden vergleichbaren Test. Es können aber auch multiple Validitäten aus mehreren verwandten (oder auch divergenten) Tests bestimmt werden.
- Die **Äußere Kriterienbezogene Validität** wählt als Kriterium Expertenurteile oder objektive Maße, wie Fehlzeiten, produzierte Stückzahlen.
- Die **Vorhersagevalidität (prognostische Validität)** muss ermittelt werden, wenn mit dem Verfahren Prognosen über zukünftiges Verhalten angestellt werden sollen (z.B. Eignungstests). In einigen Testmanualen finden sich sog. Erwartungstabellen, die angeben, wie viel Prozent der Probanden bei einem bestimmten Testwert z.B. eine Ausbildung gut abschließen.

Gütekriterien 3: Validität (Gültigkeit) 4

- Die **Übereinstimmungsvalidität** zielt hingegen keine Prognose an, so dass Kriterium und Test gleichzeitig erhoben werden können.
- Die **differentielle Validität** gibt verschiedene Validitätskoeffizienten für (a) verschiedene Kriterien oder (b) verschiedene Stichproben an.
- Die **Multiple Validität** belässt es nicht bei der Aufzählung einzelner differentieller Validitätskoeffizienten, sondern vereinigt verschiedene Kriterien und/oder Stichproben mittels multipler Regressionsgleichungen zu einer Validität.
- Die **ökologische Validität** beschreibt zudem die Künstlichkeit bzw. Echtheit der mit einem Verfahren gewonnenen Ergebnisse und gibt damit an, wie sehr die Ergebnisse generalisiert werden dürfen. Reaktive und experimentelle Verfahren schneiden hier schlechter ab als nicht reaktive, nicht manipulative Verfahren.

Gütekriterien 4: Weitere Kriterien

- Repräsentativität (Generalisierbarkeit)
- Bedeutsamkeit / Relevanz
- Ethischen Kriterien
 - Schutz der Menschenwürde der UntersuchungsteilnehmerInnen
 - Informationspflicht gegenüber den untersuchten Personen
 - Verantwortung der ForscherIn für alle Vorkommnisse während der Untersuchung
 - Freiwillige Teilnahme und Recht auf jederzeitigen Abbruch der Teilnahme
 - Vermeiden psychischer und körperlicher Beeinträchtigungen
 - Anonymität und Datenschutz

Population und Stichprobe

Population und Grundgesamtheit

- Eine Studie soll Aussagen über eine bestimmte Gruppe von Personen treffen. Welche Gruppe ist gemeint? Wie kann die Gruppe definiert und abgegrenzt werden?
- Die abgegrenzte und definierte Gruppe ist die **Population** bzw. **Grundgesamtheit** der Studie.
- Ist keine **Vollerhebung** möglich, muss eine Stichprobe aus der Population zusammengestellt werden, die für diese möglichst **repräsentativ** ist.

Sampling – Stichprobenauswahl 1

- Eine **Zufallsstichprobe** liegt dann vor, wenn...
... für jedes Element in der Grundgesamtheit die selbe Wahrscheinlichkeit besteht, in die Stichprobe aufgenommen zu werden.
... die Entnahme der einzelnen Elemente unabhängig voneinander erfolgt.
- **Schichtung:** Bei einer proportional geschichteten Stichprobe wird die Grundgesamtheit zunächst in Schichten mit homogenen Merkmalen unterteilt, aus denen dann Zufallsstichproben gezogen werden, deren Größenverhältnis untereinander dem Verhältnis der Teilgesamtheiten in der Grundgesamtheit entspricht.
- **Klumpenstichprobe:** Eine Klumpenstichprobe liegt dann vor, wenn mehrere zufällig ausgewählte Klumpen (natürliche Gruppen, z.B. Schulklassen) vollständig untersucht werden.
- Bei einer **mehrstufigen Auswahl** werden nach einer Klumpenauswahl, in einem zweiten Schritt, die UntersuchungsteilnehmerInnen nach einem anderen Verfahren gewählt.
- Eine bewusste, **gezielte Auswahl** der UntersuchungsteilnehmerInnen ist bei explorativen Studien sinnvoll (z.B. Fallstudien). Rückschlüsse auf eine Grundgesamtheit sind dann aber nur mit Einschränkungen möglich.

Stichprobenauswahl – nicht zufällig

Verfahren	Regel
Bequemlichkeit	Auswahl derjenigen, die für die Studie erreicht werden können.
sehr ähnliche, sehr unterschiedliche	Auswahl von Fällen, die in sich besonders ähnlich sind; oder als Alternative: Auswahl von Fällen, die möglichst unterschiedlich sind.
typische Fälle	Auswahl von Fällen, bei denen man im Vorhinein weiß, dass sie typisch sind und nicht extrem aus dem Rahmen fallen.
kritische Fälle	Auswahl von Fällen, die kritisch sind oder Schlüsselfunktionen haben bei der späteren Anwendung der Studienergebnisse.
Schneeball	Die UntersuchungsteilnehmerInnen verteilen die Fragebögen weiter.
Quotierung	Gezielte Auswahl von Personen, die zur Grundgesamtheit in Hinblick auf bestimmte Merkmale passen.

- Bedeutsam für die Festlegung der Stichprobengröße sind eine Reihe von Faktoren:
 - Forschungsansatz: Qualitative Untersuchungen beruhen auf Daten von einigen wenigen Personen. Das Abbruchkriterium rät dazu mit der Erhebung aufzuhören sobald sich keine neuen Informationen mehr ergeben.
 - Größe der Grundgesamtheit: Wenn es weltweit nur 10 Personen mit der oder der Krankheit gibt, können auch nicht mehr untersucht werden. Ist die Verfügbarkeit gegeben, sollte bei einer kleinen Grundgesamtheit grundsätzlich die gesamte Grundgesamtheit untersucht werden.
 - Verfügbarkeit: Die finanziellen, zeitlichen oder sonstigen Beschränkungen in der Verfügbarkeit einer Stichprobe spielt eine Rolle. Es macht wenig Sinn ein Stichprobe von 100 Personen zu fordern, wenn jede Untersuchung Unsummen kostet.

- Repräsentativität: Die Repräsentativität wächst bei einer echten Zufallsstichprobe mit der Größe der Stichprobe. Bei einer Zufallsstichprobe entscheidet der blinde Zufall über die Repräsentativität. Das ist gut, weil man damit keine systematischen Verzerrungen vornimmt. Das ist schlecht, weil die Stichprobe groß sein muss, damit man „alles mit drin hat“.
Repräsentativ ist eine Stichprobe dann, wenn sie hinsichtlich aller relevanter Merkmale mit der Grundgesamtheit übereinstimmt. Man könnte daher diese relevanten Merkmale in der Stichprobe mit denen in der Grundgesamtheit vergleichen (z.B. über amtliche Statistiken).
Probleme bereiten häufig Untersuchungsverweigerer (echte Zufallsstichproben kann es eigentlich nicht geben). Hier kann man die schnell Antwortenden (die ersten 20%) mit den Nachzüglern (die letzten 20%) vergleichen.

- Statistisches Verfahren: Viele statistische Verfahren erfordern die Normalverteilung der Mittelwerte der Datenstichproben. Diese ist ab 25 bis 50 Befragten Personen pro Untersuchungsgruppe ungefähr erreicht.
- Zu erwartende Effektgröße: In der Regel gilt, dass mit der Größe der Stichprobe auch die Chance wächst tatsächlich vorhandene Unterschiede als signifikant nachzuweisen. Störgrößen mitteln sich bei großen Stichproben heraus und der eigentliche Effekt wird klarer sichtbar. Daher gilt, dass die Stichprobe so groß wie möglich sein sollte.
 - Aus finanziellen oder anderen Gründen muss die Stichprobe aber dennoch meistens begrenzt bleiben. Die Frage danach, wie groß die Stichprobe denn mindestens sein muss, um einen vermuteten Effekt auch zu zeigen, wird durch die Power-Analyse beantwortet.

Power-Analyse Kleine Effekte

Alpha = 0,05 1-Beta = 0,80	Unterschieds- Hypothesen	Zusammenhangs- Hypothesen
2 Gruppen / Variablen / Objekte	Einfache Gruppenvergleiche 310 Fälle pro Gruppe	Einfache Korrelation 614 Fälle Die Angabe meint die Gesamtgröße.
Mehr als 2 Gruppen / Variablen / Objekte	Multiple Gruppenvergleiche 3 Gruppen: 969 Fälle 4 Gruppen: 1096 Fälle 5 Gruppen: 1200 Fälle Die Angabe meint jeweils die Gesamtgröße.	Multiple Korrelation 2 Prädiktoren: 485 3 Prädiktoren: 550 4 Prädiktoren: 602 Die Angabe meint jeweils die Gesamtgröße.

Power-Analyse Mittelgroße Effekte

Alpha = 0,05 1-Beta = 0,80	Unterschieds- Hypothesen	Zusammenhangs- Hypothesen
2 Gruppen / Variablen / Objekte	Einfache Gruppenvergleiche 50 Fälle pro Gruppe	Einfache Korrelation 64 Fälle Die Angabe meint die Gesamtgröße.
Mehr als 2 Gruppen / Variablen / Objekte	Multiple Gruppenvergleiche 3 Gruppen: 159 Fälle 4 Gruppen: 180 Fälle 5 Gruppen: 200 Fälle Die Angabe meint jeweils die Gesamtgröße.	Multiple Korrelation 2 Prädiktoren: 68 3 Prädiktoren: 77 4 Prädiktoren: 85 Die Angabe meint jeweils die Gesamtgröße.

Power-Analyse Große Effekte

Alpha = 0,05 1-Beta = 0,80	Unterschieds- Hypothesen	Zusammenhangs- Hypothesen
2 Gruppen / Variablen / Objekte	Einfache Gruppenvergleiche 20 Fälle pro Gruppe	Einfache Korrelation 22 Fälle Die Angabe meint die Gesamtgröße.
Mehr als 2 Gruppen / Variablen / Objekte	Multiple Gruppenvergleiche 3 Gruppen: 66 Fälle 4 Gruppen: 76 Fälle 5 Gruppen: 80 Fälle Die Angabe meint jeweils die Gesamtgröße.	Multiple Korrelation 2 Prädiktoren: 31 3 Prädiktoren: 36 4 Prädiktoren: 40 Die Angabe meint jeweils die Gesamtgröße.

- Das Programm G*Power wird seit 1992 programmiert und kann als Freeware aus dem Internet bezogen werden.
- Das Programm erlaubt die genau Abschätzung der mindestens nötigen Stichproben-Größe für verschiedene Testverfahren, Alpha- und Beta-Werte.
- Es kann z.B. bezogen werden unter:
www.psych.uni-duesseldorf.de/aap/projects/gpower/

Messtheorie

- Messen ist die homomorphe Zuordnung eines empirischen Relativs zu einem numerischen Relativ.
- Homomorphie: Strukturerhaltene Abbildung.
- Struktur?
 - Empirisches Relativ: Relationen zwischen Objekten der erfahrbaren/erschließbaren Welt.
 - Numerisches Relativ: Relationen zwischen Zahlen.

Skalenniveaus und Transformation

Skalenniveau	Das darf eine Transformation nicht verändern ...	Zulässige Interpretation
Nominal	Ein-eindeutig Zuordnung	Code, Bezeichnung, Beispiel: Berufe
Ordinal	Reihenfolge	Rangordnung Beispiel: Schulbildung
Intervall	Intervalle zwischen den Zahlen (erlaubt ist die Addition/Subtraktion von Konstanten, sowie die Multiplikation/Division mit Konstanten)	Abstände (Intervalle) zwischen den Zahlen Beispiel: Alter
Verhältnis	Verhältnisse zwischen den Zahlen (erlaubt ist die Multiplikation/Division mit Konstanten)	Verhältnisse zwischen den Zahlen Beispiel: Gehalt
Absolut	Nichts darf verändert werden.	Verhältnisse zwischen den Zahlen, Kardinalzahl Beispiel: Häufigkeiten

Skalenniveaus und Transformation

Skalenniveau	Das darf eine Transformation nicht verändern ...	Zulässige Interpretation
Nominal	Ein-eindeutig Zuordnung	Code, Bezeichnung, Beispiel: Berufe
Ordinal	Reihenfolge	Rangordnung Beispiel: Schulbildung

Metrisch

Intervall	Intervalle zwischen den Zahlen (erlaubt ist die Addition/Subtraktion von Konstanten, sowie die Multiplikation/Division mit Konstanten)	Abstände (Intervalle) zwischen den Zahlen Beispiel: Alter
Verhältnis	Verhältnisse zwischen den Zahlen (erlaubt ist die Multiplikation/Division mit Konstanten)	Verhältnisse zwischen den Zahlen Beispiel: Gehalt
Absolut	Nichts darf verändert werden.	Verhältnisse zwischen den Zahlen, Kardinalzahl Beispiel: Häufigkeiten

Skalenniveaus und Stetigkeit

- **Skalenniveaus (extrem wichtig!)**
 - Nominal
 - Ordinal
 - **Metrisch**
 - Intervall
 - Verhältnis
 - Absolut
- **Stetigkeit**
 - Diskret
 - Mehrstufig
 - Dichotom
 - **Stetig**

Skalen-Niveaus I: Nominal-Skala

Beliebige Zahlen stehen für beliebige Begriffe (z.B. Hausnummern, Geschlecht, Berufe)

Was man darf, was man nicht darf...

- ✓ **Zählen (45 Frauen und 38 Männer nahmen an der Befragung teil)**
- ✓ **Prozentangaben (55% der Befragten waren Frauen)**
- ✓ **Modalwert (die meisten Antwortenden waren Frauen)**
- ☹ **Kein Median**
- ☹ **Kein Mittelwert**
- ☹ **Keine Streuung (Standardabweichung), Varianz**

- ✓ **Häufigkeitstabellen**
- ✓ **Balkendiagramme**
- ✓ **Kreisdiagramme**

Skalen-Niveaus II: Ordinal-Skala

Die Zahlen stehen für Relationen wie größer oder kleiner. Die Ordnung der Zahlen entspricht inhaltlich der Ordnung der Dinge für die sie stehen.

z.B.

- ₁ 10-15 Jahre
- ₂ 16-25 Jahre
- ₃ 26-35 Jahre
- ₄ älter als 35 Jahre

Was man darf, was man nicht darf...

- ✓ **Zählen (45 Personen fallen in die Kategorie 3)**
- ✓ **Prozentangaben (55% der Befragten fallen in die Kategorie 3)**
- ✓ **Modalwert (die meisten Antwortenden gehören zur Kategorie 3)**
- ✓ **Median (in der Mitte der Anwohnhäufigkeiten lag die Kategorie 2)**
- ☹ **Kein Mittelwert**
- ☹ **Keine Streuung (Standardabweichung), Varianz**

Skalen-Niveaus II: Ordinal-Skala

Was man darf, was man nicht darf...

- ✓ Häufigkeitstabellen
- ✓ Balkendiagramme
- ✓ Kreisdiagramme

Skalen-Niveaus III: Intervall-Skala

Die Zahlen stehen nicht nur für Relationen wie größer oder kleiner, sondern entsprechen in ihren Abständen (Intervallen) inhaltlich und logisch den Abständen von dem was sie bezeichnen.

z.B.

Alter (wenn jemand 2 Jahre älter ist als jemand anders, stimmt der Abstand 2 unabhängig davon ob

a) die eine Person 10 und die andere 12 Jahre alt ist oder

b) die eine Person 80 und die andere 82 Jahre alt ist.

Was man darf, was man nicht darf...

- ✓ **Zählen (45 Personen waren 20 Jahre alt)**
[wenig sinnvoll!]
- ✓ **Prozentangaben (55% der Befragten waren 20 Jahre alt)**
[wenig sinnvoll!]

Skalen-Niveaus III: Intervall-Skala

Was man darf, was man nicht darf...

- ✓ **Modalwert (die meisten Antwortenden waren 25 Jahre alt)**
- ✓ **Median (die Mitte der Altersverteilung liegt bei 24 Jahren)**
- ✓ **Mittelwert (im Durchschnitt betrug das Alter 24 Jahre)**
- ✓ **Streuung (Standardabweichung), Varianz (die Standardabweichung beträgt ± 2 Jahre)**
- ✓ **Häufigkeitstabellen**
[nur sinnvoll, wenn man Gruppen bildet]
- ✓ **Balkendiagramme**
[auch Mittelwert und Streuung angeben und eventuell einzeichnen]
- ✓ **Kreisdiagramme**
[nur sinnvoll, wenn man Gruppen bildet]

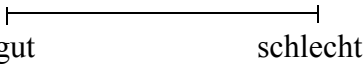
Fragebogen 1

Basisdaten	Art der Erhebung (Beispiele)	Skalenniveau
Alter	Alter: _____ Geburtsjahr: _____	Intervall-Skala
	<input type="radio"/> ₁ 10-15 Jahre <input type="radio"/> ₂ 16-25 Jahre <input type="radio"/> ₃ 26-35 Jahre <input type="radio"/> ₄ älter als 35 Jahre	Ordinal-Skala

Fragebogen 2

Basisdaten	Art der Erhebung (Beispiele)	Skalenniveau
Geschlecht	<input type="radio"/> ₀ männlich <input type="radio"/> ₁ weiblich	Nominal-Skala
Beruf	<input type="radio"/> ₁ arbeitslos <input type="radio"/> ₂ Arbeitsunfähigkeit <input type="radio"/> ₃ ArbeiterIn <input type="radio"/> ₄ Angestellte/r <input type="radio"/> ₅ Selbstständige/r <input type="radio"/> ₆ StudentIn <input type="radio"/> ₇ Ausbildung <input type="radio"/> ₈ Hausfrau / Hausmann <input type="radio"/> ₉ RentnerIn <input type="radio"/> ₁₀ Wehr- / Zivildienst Leistender <input type="radio"/> ₋₁ unbekannt	Nominal-Skala

Fragebogen 3

Ratingskalen	Art der Erhebung (Beispiele)	Skalenniveau
Gerade Anzahl an Abstufungen	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> gut schlecht	Intervall-Skala/ OrdinalSkala
Ungerade Anzahl an Abstufungen	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> gut schlecht	Intervall-Skala/ OrdinalSkala
Keine Abstufungen		Intervall-Skala
Ungerade Anzahl Abstufungen Beschriftet	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> sehr gut gut mittel schlecht sehr schlecht	Intervall-Skala/ OrdinalSkala

Gute Beispiele finden sich auch in Bortz und Döring (2002, S. 176-177)

Übungsaufgabe

- **a) Videoband cm**
 - metrisch
 - Verhältnisskala
 - diskret / (stetig)
- **b) Inflationsraten**
 - metrisch
 - Verhältnisskala
 - stetig

Übungsaufgabe

- **c) Anzahl Kinobesuche**
 - metrisch
 - Absolutskala
 - diskret
- **d) Einstellungs-Rating**
 - ordinal/(metrisch)
 - Intervallskala
 - diskret

Übungsaufgabe

- **e) Interesse, Rating**
 - ordinal / (metrisch)
 - Intervallskala
 - diskret
- **f) Einkommen**
 - metrisch
 - Verhältnisskala
 - diskret / (stetig)
- **g) Einstellungs-Rating, dreistufig**
 - ordinal / (metrisch)
 - Intervallskala
 - diskret
- **h) Temperatur**
 - Metrisch
 - Intervallskala
 - stetig / (diskret)

Übungsaufgabe

- **i) Anzahl produzierter Autos**
 - metrisch
 - Absolutskala
 - diskret
- **j) Höhe eines Berges**
 - metrisch
 - Verhältnisskala, wenn klar ist, was der Nullpunkt ist.
 - stetig / (diskret)
 - Cm mit 5 Dezimalstellen.

- Nominal und Ordinal sind immer diskret.
- Metrische Messungen können stetig oder diskret sein.
- Können metrische Merkmale auch dichotom sein?

- Die Art der Kodierung kann aus einem stetigen Merkmal ein diskretes machen.

- Eigentlich gibt es nach der Messung keine stetigen Merkmale mehr. Es ist wichtig sich zu fragen, ob das „echte“ Merkmal hinter der Messung stetig ist.

Skalenniveaus und Stetigkeit

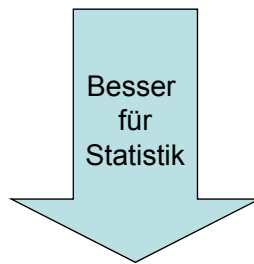
- **Skalenniveaus (extrem wichtig!)**
 - Nominal
 - Ordinal
 - **Metrisch**
 - Intervall
 - Verhältnis
 - Absolut

- **Stetigkeit**
 - Diskret
 - Mehrstufig
 - Dichotom
 - **Stetig**

Skalenniveaus und Stetigkeit

- **Skalenniveaus (extrem wichtig!)**

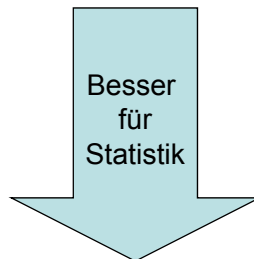
- Nominal
- Ordinal
- **Metrisch**
 - Intervall
 - Verhältnis
 - Absolut



Verbesserungsmethoden:
z.B. Nominale Objekte zählen führt zu einer Häufigkeit, die metrisch ist.

- **Stetigkeit**

- Diskret
 - Mehrstufig
 - Dichotom
- Stetig

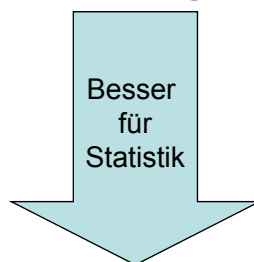


Verbesserungsmethoden:
z.B. Zusammenfassung vieler diskreter Messungen nähert sich einer stetigen Messung an.

Skalenniveaus und Stetigkeit

- **Skalenniveaus (extrem wichtig!)**

- Nominal
- Ordinal
- **Metrisch**
 - Intervall
 - Verhältnis
 - Absolut

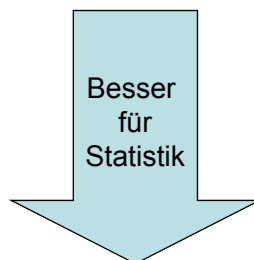


Verschlechterungsmethoden:
z.B. Metrisches Merkmal mit Kategorien abfragen.

Kann helfen:
In Bezug auf Anonymität.

- **Stetigkeit**

- Diskret
 - Mehrstufig
 - Dichotom
- Stetig



Verschlechterungsmethoden:
z.B. Gruppenbildung.

Kann helfen:
Extremgruppenvergleich.

Wofür der ganze Aufwand?

- Maße der zentralen Tendenz ...
- Dispersionsmaße ...
- Häufigkeiten / relative Häufigkeiten ...

Skala und Item

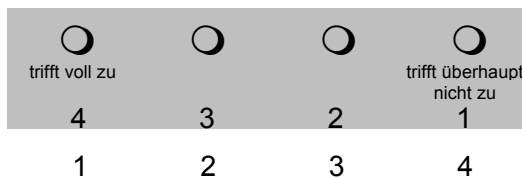
Messung hypothetischer Konstrukte – ein Beispiel

- + Diese KundenbetreuerIn hat mich freundlich behandelt.
- Ich habe gewisse Zweifel über die Kompetenz dieser KundenbetreuerIn .
- Diese KundenbetreuerIn wirkte kühl und unpersönlich.
- + Diese KundenbetreuerIn hat ihr Bestes getan, um mich nicht zu informieren.
- + Diese KundenbetreuerIn wirkte sehr kompetent und erfahren.
- + Diese KundenbetreuerIn schien ein echtes und persönliches Interesse an mir zu haben.
- Diese KundenbetreuerIn hat mich mit vielen unbeantworteten Fragen zurückgelassen.
- Diese KundenbetreuerIn verwendete Fachausdrücke, die ich nicht verstanden habe.
- + Ich habe großes Vertrauen in diese KundenbetreuerIn .
- Ich habe mich nicht getraut, dieser KundenbetreuerIn Fragen zu stellen.

(In Anlehnung an: Langewitz, Keller & Denz, 1995)

Messung hypothetischer Konstrukte

- **Skalierung: z.B. ja / nein**
 - Zählen der Zustimmungen zum Konstrukt (+ und – beachten)
- **Skalierung über Ratings**



bei positiven Items

bei negativen Items

**Skalenwert gleich Summe oder
Skalenwert gleich Mittelwert (Umgang mit fehlenden Werten ist beim
Mittelwert besser)**

Herleitung der klassischen Testtheorie

Klassische Testtheorie 1

- Klassische Testtheorie

$$x = \mu + e$$

Der Messwert x entspricht dem wahren Wert μ plus einem Fehler e .
Der Fehler kann minimiert werden, wenn viele x erhoben und gemittelt werden.

Ein Mittelwert ist gegeben durch:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

n : Anzahl der Messungen

x : Messwert

i : Laufvariable für den 1., den 2., 3. ... n -ten Messwert.

Die Varianz s^2 (durchschnittliche quadrierte Abweichung vom Mittelwert) ist gegeben durch:

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Klassische Testtheorie 2

$$S_x^2 = S_\mu^2 + S_e^2$$

Varianz der Messwerte ist gleich Varianz der wahren Werte plus Fehlervarianz

$$r = \frac{S_\mu^2}{S_x^2}$$

Die Reliabilität r ist das Verhältnis der Varianz der wahren Werte zur Varianz der gemessenen Werte. Man spricht hier auch von Varianzaufklärung.

$$r = 1 - \frac{S_e^2}{S_x^2}$$

Die Reliabilität r ist auch Eins minus das Verhältnis der Varianz der Fehler zur Varianz der gemessenen Werte. Ist die Fehlervarianz groß, so ist die Reliabilität klein.

$$\alpha = r = \frac{c}{c-1} \left[1 - \frac{\sum_j^c s_j^2}{s_x^2} \right]$$

Die Varianzen jedes einzelnen Items ($j = 1$ bis c) werden aufsummiert und durch die Varianz des Skalenwertes x geteilt.

Eins minus diese Zahl wird mit der Zahl der Items – dividiert durch die Zahl der Items minus Eins – multipliziert. Diese Form der Reliabilität heißt Cronbach Alpha

Datenverarbeitung & Deskription

Goldene Regeln Dateneingabe

- Fragebögen vor der Eingabe mit fortlaufenden eindeutigen Nummern versehen!
- Möglichst alle Fragebogendaten als Zahlen kodieren!
- Immer genau die Daten eingeben, die im Fragebogen stehen! Zusammenfassungen, Gruppenbildung etc. wird erst später mit der Software vorgenommen.
- Für dichotome Variablen (ja / nein; männlich / weiblich) 1 und 0 vergeben und nicht 1 und 2!

Vom Fragebogen zur Datendatei

Kodierung des
Fragebogens in einer
leeren Papierversion

Definition der
Variablen in Excel oder SPSS bzw. PSPP

Dateneingabe mit
laufender Nummerierung.
Fehlende Daten auslassen

Fragebogenkodierung 1

Wie zufrieden sind Sie mit Ihrer KundenbetreuerIn?
KundInnenbefragung zur Qualität der Beratung

Laufnummer

Bitte beantworten Sie folgende Fragen, die sich auf das Beratungsgespräch beziehen, welches Sie kürzlich in einem unserer Institute hatten

Bitte zutreffendes Kästchen ankreuzen

**1
2
Anzahl**

**15
Minuten**

Codeliste

<p>1. Wie oft haben Sie bereits ein Beratungsgespräch in einer Bank in Anspruch genommen?</p> <p><input type="checkbox"/> noch nie, das war mein erstes Beratungsgespräch</p> <p><input type="checkbox"/> bereits 1 mal zuvor</p> <p><input type="checkbox"/> mehr als 1 mal zuvor und zwar insgesamt ungefähr _____ mal</p>
<p>2. Wie lange dauerte das Gespräch?</p> <p><input type="checkbox"/> ca. 15 Minuten</p> <p><input type="checkbox"/> mehr als 15 Minuten und zwar ungefähr _____ Minuten</p>
<p>3. In welcher Filiale haben Sie das Gespräch in Anspruch genommen?</p> <p>_____</p>
<p>4. Auf wessen Empfehlung hin haben Sie die Beratung in Anspruch genommen?</p> <p>_____</p>

Dateneingabe im Excel

WU
EXECUTIVE
ACADEMY

Jede Zeile ist ein Fall

Jede Spalte ist eine Variable

Fehlende Angaben bleiben leer

The screenshot shows an Excel spreadsheet with the following columns: A (LNR), B (WANN), C (GEBJ), D (GESCHL), E (Item1), F (Item2), G (Item3), H (Item4), I (Item5), J (Item6), K (Item7), L (Item8), M (Item9), N (Item10), O (Item11). The data rows contain numerical values for each variable. Annotations include:

- A box pointing to a row: "Jede Zeile ist ein Fall"
- A box pointing to a column: "Jede Spalte ist eine Variable"
- A box pointing to empty cells: "Fehlende Angaben bleiben leer"
- A box pointing to the formula bar: "Nur Zahlen eingeben und nicht m / w"

Häufigkeitstabellen

Excel: Daten->PivotTabelle

Anzahl von GEBJ	GESCHL		Gesamtergebnis
1923	0	1	1
1941		2	2
1945		1	1
1946		3	3
1948		3	3
1950	1		1
1952		1	1
1954	1		1
1955		1	1
1956		1	1
1957	1		1
1958		1	1
1960	1		1
1961		1	1
1962	1		1
1963	1	4	5
1964	4	2	6
1965	3	3	6
1966		5	5
1967	2		2
1968	2	8	10
1969	1	2	3
1970	1	4	5
1971	1	5	6
1972	4	6	10
1973	4	5	9
1974	8	16	24
1975	6	6	12
1976	8	5	13
1977	5	9	14
1978	5	7	12
1979	4	1	5
1980	1	4	5
1981	2	3	5
1982	1	4	5
1983	1	4	5
1984	1		1
1985		2	2
1987		1	1
Gesamtergebnis	70	121	191

Tabelle 1: Altersverteilung und Geschlecht (absolute Häufigkeiten)

Geburtsjahr	Frauen	Männer	Summe
< 1950	1	10	11
1950 – 1960	3	4	7
1961 – 1970	15	19	44
1971 – 1980	46	64	110
> 1980	5	14	19
Summe	70	121	191

Tabelle 1: Altersverteilung und Geschlecht (relative Häufigkeiten)

Geburtsjahr	Frauen-Anteil	N (%)
< 1950	9,1%	11 (5,7%)
1950 – 1960	42,9%	7 (3,7%)
1961 – 1970	34,1%	44 (23,0%)
1971 – 1980	41,8%	110 (57,6%)
> 1980	26,3%	19 (10,0%)
Insgesamt	36,6%	191 (100,0%)

Diagramme: Häufigkeiten, Balkendiagramm

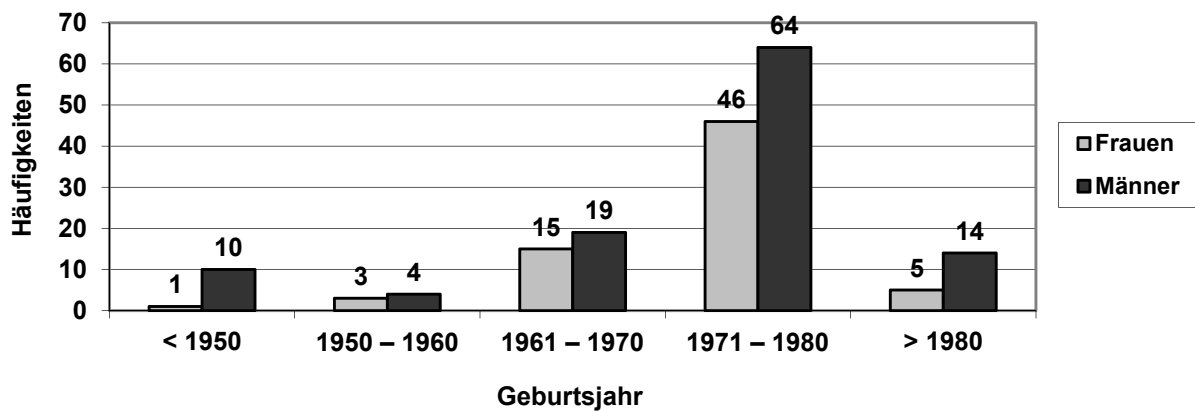


Abbildung 1: Altersverteilung und Geschlecht (absolute Häufigkeiten)

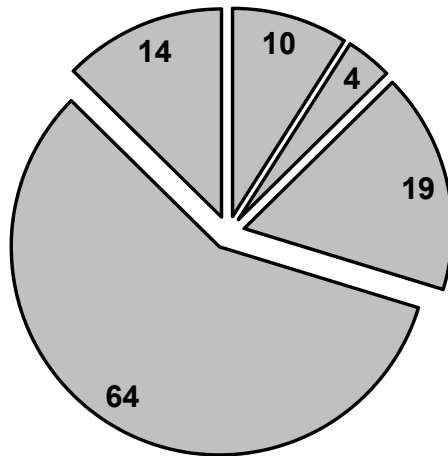


Abbildung 2: Altersverteilung der Männer (absolute Häufigkeiten)

Maße der zentralen Tendenz

- **Modalwert:** häufigster Wert
- **Median:** Ist der Wert, über dem genau so viele Fälle liegen wie unter ihm. Der Median ist die Mitte der sortierten Häufigkeits-Verteilung.
- **Arithmetisches Mittel:** Summe aller Werte, dividiert durch die Anzahl der Werte.

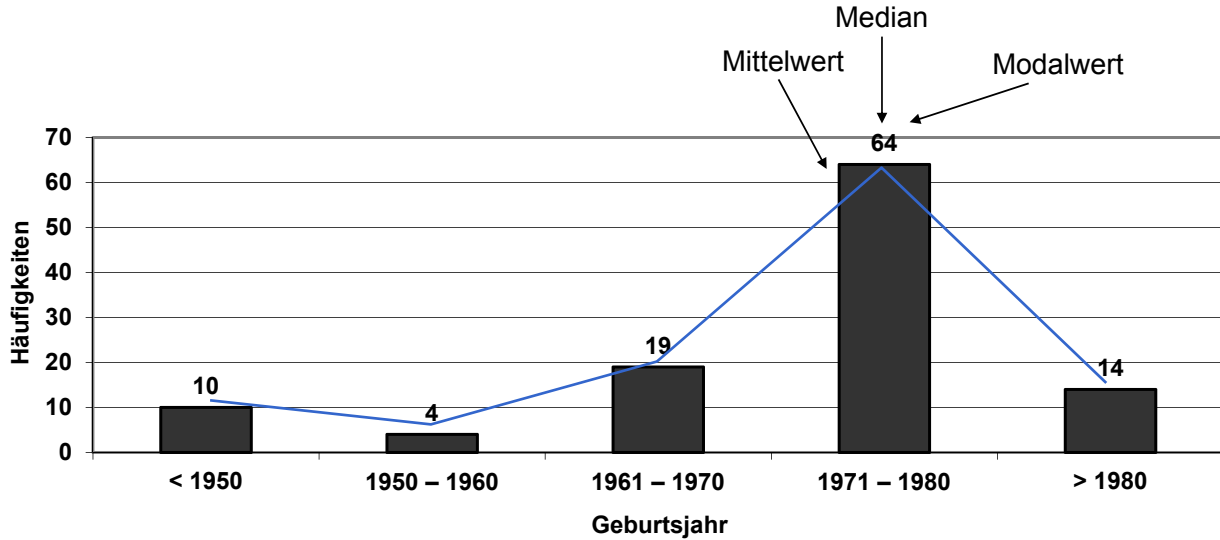
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- **Geometrisches Mittel** (Verhältnisskala – Mittelwert eines multiplikativen Faktors):

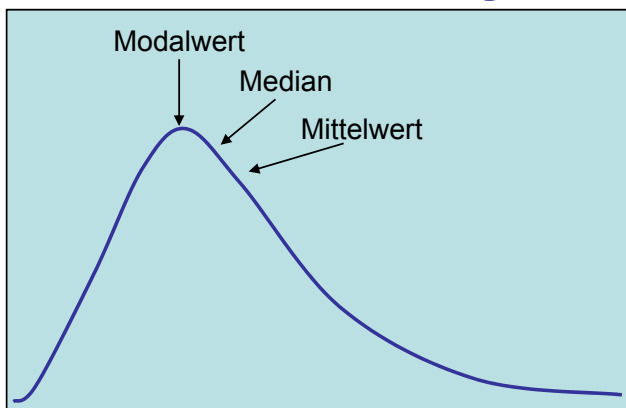
$$GM = \sqrt[n]{\prod_{i=1}^n x_i}$$

- **Harmonisches Mittel** (Verhältnisskala – Mittelwert von Brüchen mit konstanten Zähler):

$$HM = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

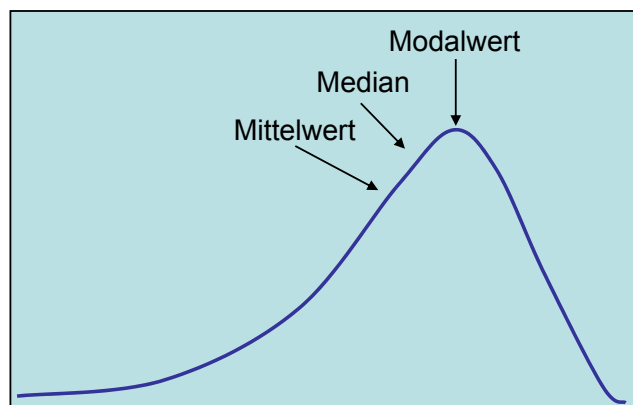


Schiefe Verteilungen



Linkssteil bzw. rechtsschief

Rechtssteil bzw. linksschief



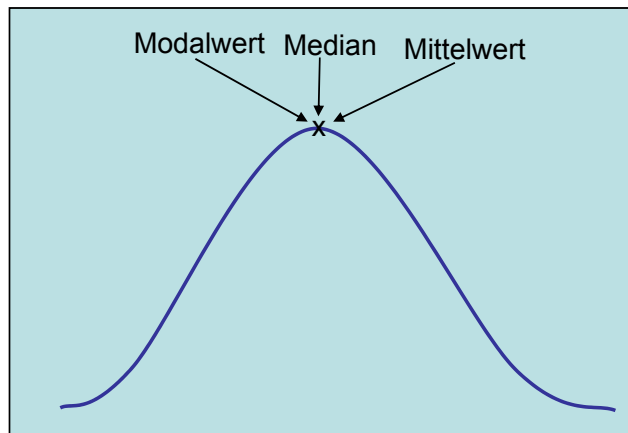
➔ Schiefe

$$Sch = \frac{\bar{x} - \text{Modalwert}}{s}$$

Sch < 0: Rechtssteil bzw. linksschief

Sch > 0: Linkssteil bzw. rechtsschief

Symmetrische Verteilungen



Streuungsparameter

- **Variationsbreite (range):** größter und kleinster Wert
- **Quartilsabstand:** Differenz zwischen dem Perzentil 25 und dem Perzentil 75. **Perzentil:** x-tes Perzentil ist diejenige Merkmalsausprägung, die x Prozent der Verteilungsfläche abschneidet (Median ist Perzentil 50)

- **Varianz**

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

- **Streuung bzw. Standardabweichung**

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Streuungsparameter Stichprobe oder Population

- **Stichproben-Varianz**

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

- **Stichproben-Streuung bzw. Stichproben-Standardabweichung**

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Ist n groß, dann gleichen sich die Ergebnisse zunehmend an und es wird unbedeutend, ob mit n oder mit n-1 gerechnet wird.

Grundidee – ein Beispiel

- Behauptung: „Ich kann zaubern, die Münze fällt immer auf Kopf!“
 - Ab wann wäre diese Behauptung empirisch gesehen statistisch signifikant abgesichert?
 - Wenn bei 20 mal Werfen 10 mal Kopf kommt?
 - Wenn bei 20 mal Werfen 11, oder 12, oder 13 mal Kopf kommt?
 - Null-Hypothese: „Kein Zaubern, sondern Zufall!“
 - Die Null-Hypothese geht von Zufall aus. Die höchste mathematische Wahrscheinlichkeit für Zufall wäre bei 10 mal Kopf (bei 20 mal Werfen) gegeben.
 - Es lässt sich ausrechnen, wie groß die Wahrscheinlichkeit für 11, 12, 13 usw. mal Kopf ist. Dabei wird davon ausgegangen, dass 10 mal normal wäre und jeder Wurf mit 50%-iger Wahrscheinlichkeit Kopf ergibt.
 - Die Berechneten Wahrscheinlichkeiten gehen also von der Null-Hypothese (50% für Kopf) aus.

Grundidee – ein Beispiel

- Der Münzwurf ist mathematisch gesehen binomial verteilt:

$$P(X=k) = n! / (k! (n-k)!) * P^k * (1-P)^{(n-k)}$$

P: Wahrscheinlichkeit für das Ereignis Kopf = 0,5

$P(x=k)$ Wahrscheinlichkeit dafür, dass das Ereignis X (Kopf) bei n Versuchen exakt k mal auftritt.

Im konkreten Beispiel ist jedoch eigentlich $P(X \leq k)$, also die Summe aller

Einzelwahrscheinlichkeiten für $k=1, k=2, \dots, k=15$ (oder bis $k=12$ oder $k=13$, je nach dem) gefragt, um die „Zauberfrage“ beantworten zu können.

- Für $n=20$ Würfe und $k=15$ mal Kopf beträgt die Wahrscheinlichkeit nur mehr 2%. Das dieses Ereignis (15 Treffer) mit einer 50-50-Münze im Einklang steht ist also sehr unwahrscheinlich.
- Weil die Nullhypothese sehr unwahrscheinlich ist, wird sie fallen gelassen. Das heißt, das Ereignis ist signifikant.

Sind die Frauen im Beispiel jünger?

Tabelle 1: Geburtsjahr nach Geschlecht

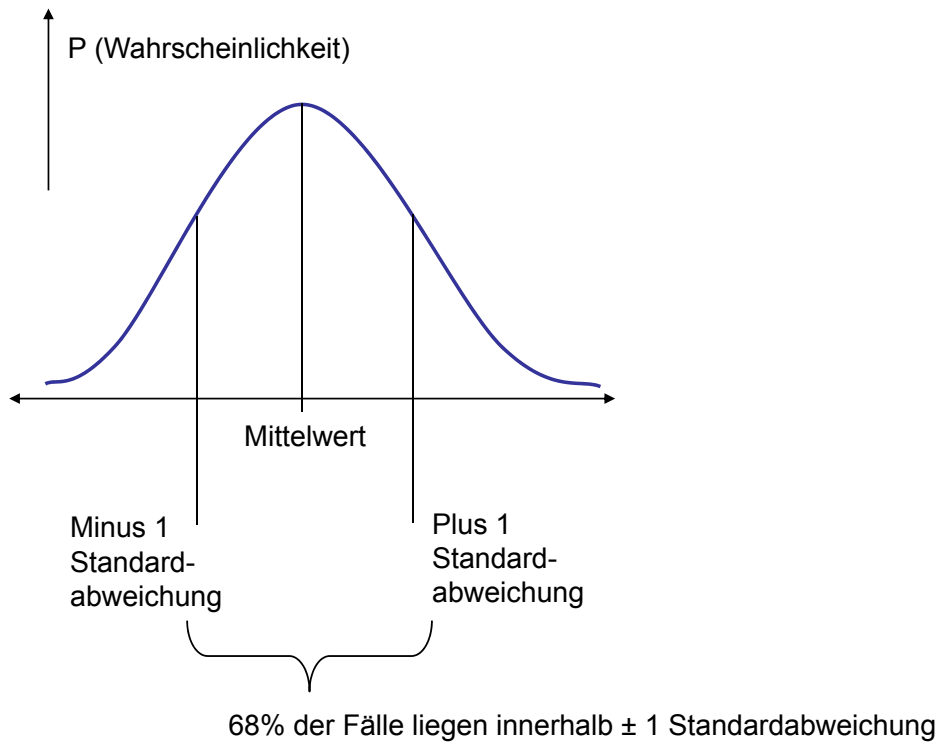
Geschlecht	Mittelwert	Streuung	N
Frauen	1972,73	6,70	121
Männer	1970,62	10,59	70
Gesamt	1971,38	9,40	191

Der Unterschied zwischen den beiden Mittelwerten beträgt rund 2 Jahre (exakt sind es 2,11 Jahre).

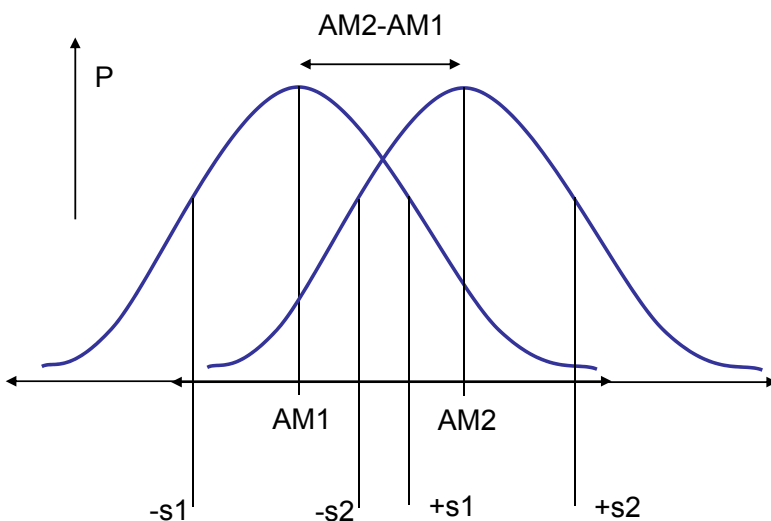
Aber es gibt Frauen, die älter sind als der Durchschnitt der Männer und Männer, die jünger sind als der Durchschnitt der Frauen.

Ob der Unterschied statistisch bedeutsam ist (also signifikant ist) kann mit Hilfe der Wahrscheinlichkeitsrechnung entschieden werden.

Normalverteilung



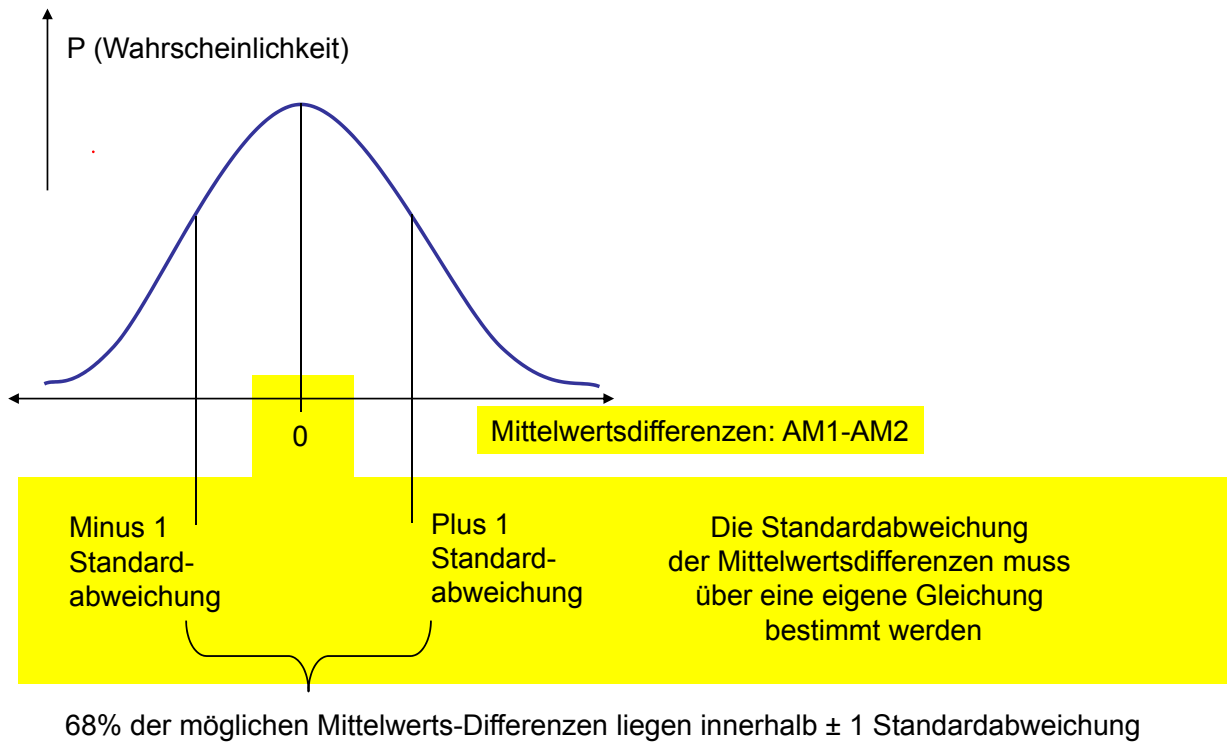
Vergleich zweier Mittelwerte



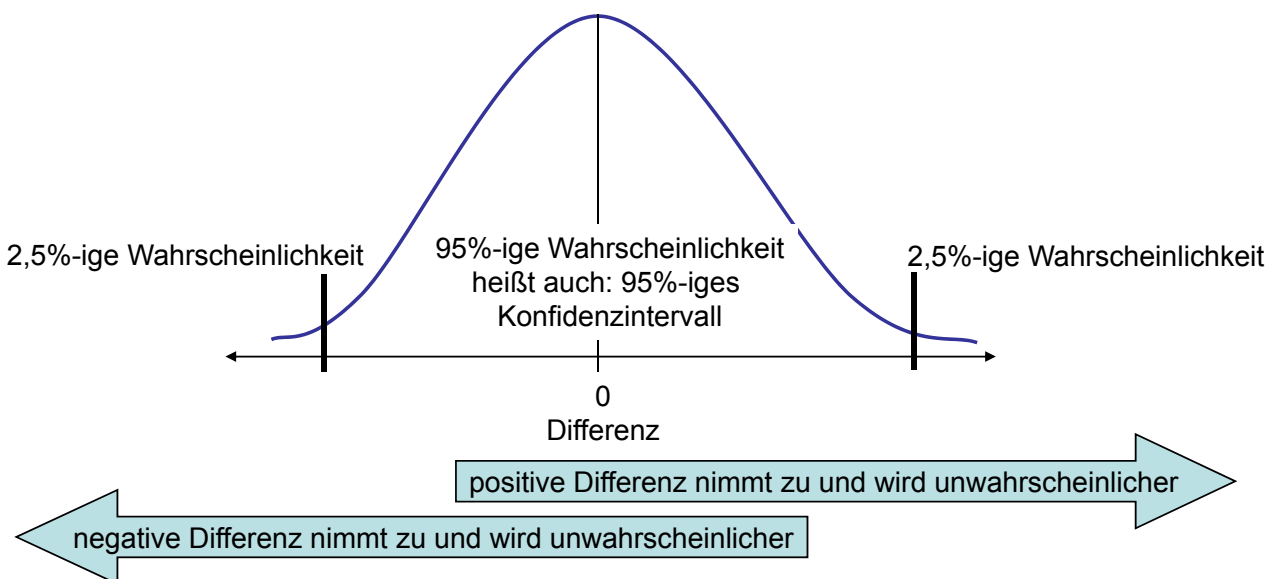
AM1: Mittelwert der Gruppe 1
AM2: Mittelwert der Gruppe 2
s: Standardabweichungen.

- Ziel wäre es die Wahrscheinlichkeit für den Unterschied zwischen AM1 und AM2 zu kennen.
- Die Verteilung müsste die Nullhypothese (kein Unterschied zwischen AM1 und AM2) als wahrscheinlichsten Fall ansehen.
- Tatsächlich sind Mittelwerts-Differenzen normalverteilt mit einer erwarteten Differenz von Null und einer Standardabweichung, die sich aus den gegebenen Standardabweichungen berechnen lässt.

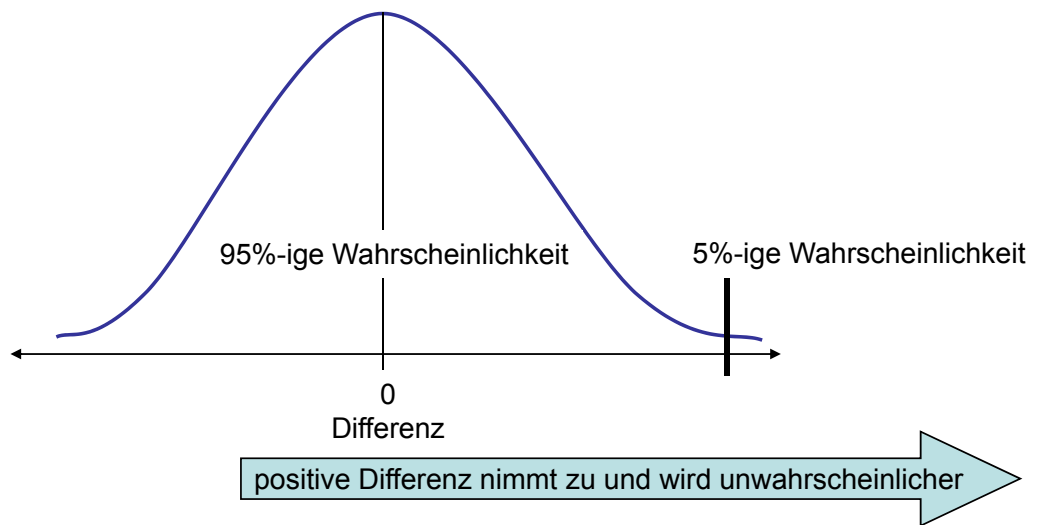
Normalverteilung von Mittelwerts-Differenzen



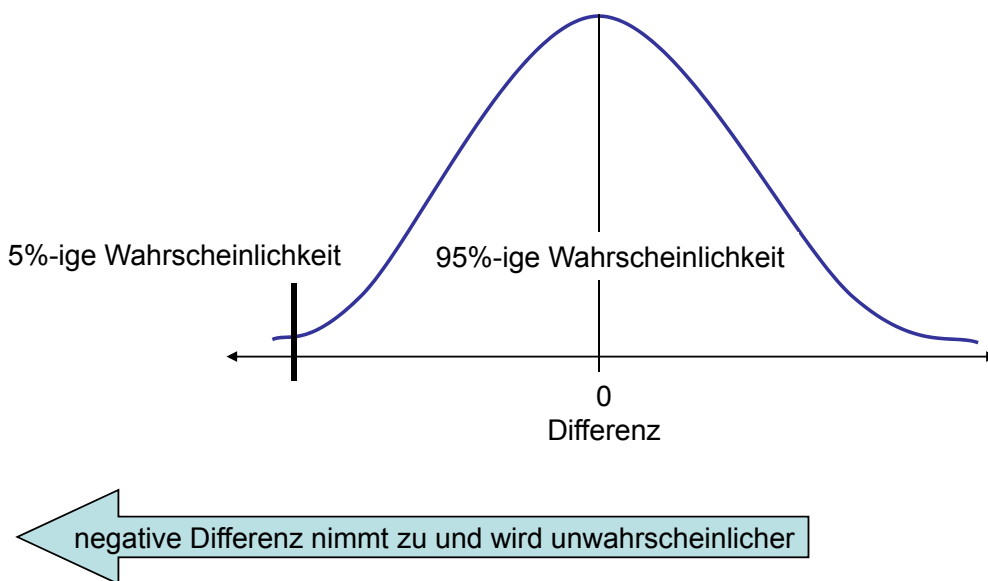
Vergleich zweier Mittelwerte 2-seitig



Vergleich zweier Mittelwerte 1-seitig



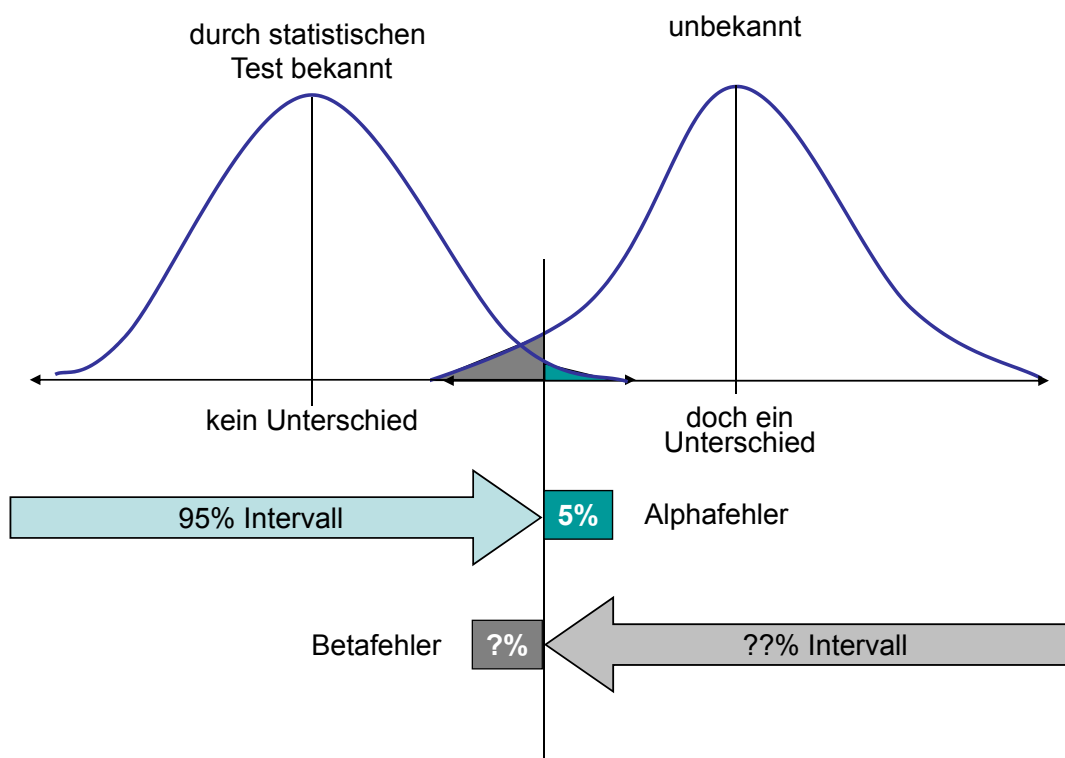
Vergleich zweier Mittelwerte 1-seitig



Zusammenfassung

- Der Vergleich zweier Mittelwerte geht davon aus, dass die **Differenz der Mittelwerte** normalverteilt ist.
- Die Nulldifferenz ist die Mitte der Verteilung. Eine Nulldifferenz ist also am wahrscheinlichsten. **Jeder statistische Test geht von der Null-Hypothese aus.**
- Weicht die tatsächliche Differenz von der Nullhypothese ab, so kann eventuell ein signifikanter Unterschied vorliegen.
- Beträgt die Wahrscheinlichkeit dafür, dass die tatsächliche Differenz Null ist nur mehr 5%, so gilt die Differenz als signifikant. Die Null-Hypothese wird verworfen und die Alternativ-Hypothese wird akzeptiert.
- Der **Alpha-Fehler** beträgt hier $\alpha = 5\%$ (Man schreibt statt $\alpha = 5\%$ auch gerne $\alpha = 0,05$).
- Bei einem Alpha-Fehler von nur mehr 1% spricht man von **sehr signifikanten** Differenzen ($\alpha = 1\%$ bzw. $\alpha = 0,01$).

Alpha- und Betafehler



Einfache Testverfahren

T-Test

t-Test

Voraussetzungen

- **t- bzw. Normalverteilung der Differenzen.** Ist bei n_1 und n_2 jeweils größer 25 (oder 30 oder 50) automatisch gegeben. Bei kleineren n müssen die Messwerte normalverteilt sein. Dies kann mittels **Kolmogorov-Smirnov-Anpassungstest** geprüft werden. Liegt keine Normalverteilung vor muss ein anderer Test gewählt werden, z.B. der **U-Test** von Mann-Whitney.
- **Varianzhomogenität.** Die Varianzen müssen in beiden Stichproben in etwa gleich groß sein. Dies kann geprüft werden mit dem **F-Test**. Ist das nicht der Fall gibt es Korrekturformeln.
- **Unabhängigkeit der Messwerte.** Die Daten müssen aus zwei voneinander unabhängigen Stichproben stammen. Zwei Messungen an der selben Stichprobe zu verschiedenen Zeiten, verletzen diese Forderung. Hierfür gibt es einen **t-Test für abhängige Stichproben**.
- Sind die Daten abhängig und ist die Normalverteilung ebenfalls verletzt, ist der **Wilcoxon-Rangsummen-Test** zu rechnen.

t-Test

- Vergleich zweier Stichprobenmittelwerte aus unabhängigen Stichproben:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)}} \quad \hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$FG = n_1 + n_2 - 2$$

t ist eine Zufallsvariable, die für kleine Stichproben mit FG t -verteilt und für größere Stichproben ($n_1 + n_2 \geq 50$) annähernd normalverteilt ist.

Rechnet man den Test von Hand, so muss in einer Tabelle nachgeschlagen werden, wie wahrscheinlich der errechnete Wert für t bei den gegebenen Freiheitsgraden ist.

Die Wahrscheinlichkeit sinkt mit wachsendem t . Ist die Wahrscheinlichkeit kleiner 0,05 so liegt ein signifikanter Unterschied vor.

t-Test Beispiel

Tabelle 1: Geburtsjahr nach Geschlecht

Geschlecht	Mittelwert	Streuung	N
Frauen	1972,73	6,70	121
Männer	1970,62	10,59	70
Gesamt	1971,38	9,40	191

$$\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)} = 1,251$$

$$t = 2,11 / 1,251 = 1,687$$

$$FG = 189$$

Tabelle der t-Verteilung (Ausschnitt)

FG	t (für $\alpha = 5\%$, 1-seitig)
80	1,664
90	1,662
100	1,660
150	1,655
200	1,653

Ist der positive Wert von $t \geq t_{\text{Tabelle}}$, dann ist der Unterschied signifikant. t_{Tabelle} wird auch als kritische Grenze bzw. t_{krit} bezeichnet.

SPSS bzw. PSPP

T-Test

T-Test – SPSS

OK: Ausführen
Einfügen: Syntax schreiben

Variable an der die beiden Gruppen unterschieden werden

Variablenauswahl für die Auswertung

T-Test – unabhängige Stichproben

Titel
Anmerkungen
Gruppenstatistiken
Test bei unabhängigen Stichp

T-Test

Gruppenstatistiken					
	GSCHL	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
ITEM16	männlich	123	1,67	,70	6,29E-02
Gesamtzufriedenheit	weiblich	248	1,75	,74	4,72E-02

Test bei unabhängigen Stichproben								
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit				
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz
ITEM16	Varianzen sind gleich	,206	,650	-1,037	369	,300	-8,33E-02	8,03E-02
Gesamtzufriedenheit	Varianzen sind nicht gleich			-1,060	257,815	,290	-8,33E-02	7,86E-02

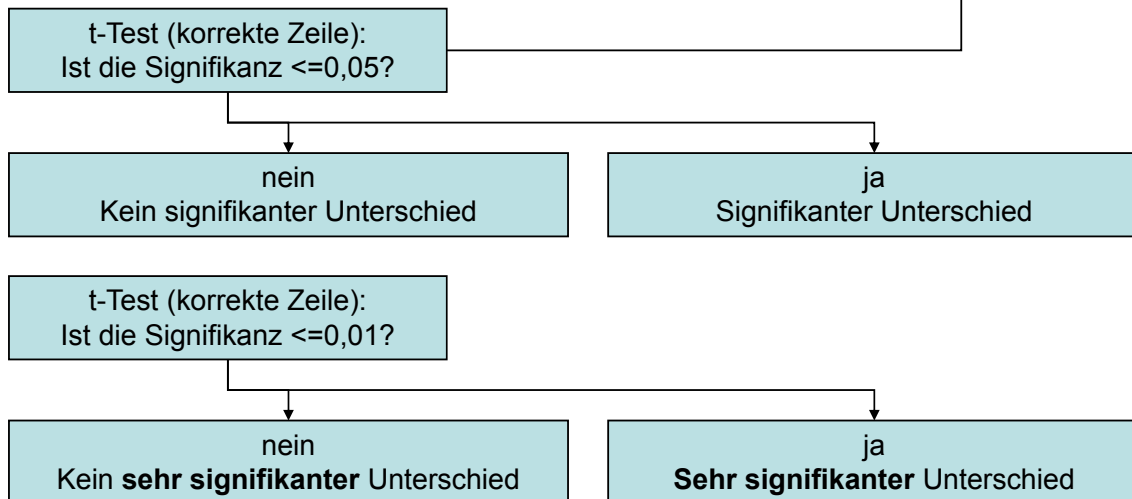
Levene-Test:
Ist die Signifikanz $\geq 0,20$?

nein
Varianzen sind nicht gleich: untere Zeile

ja
Varianzen sind gleich: obere Zeile

T-Test – unabhängige Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwerte			
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz
ITEM16	Varianzen sind gleich	,206	,650	-1,037	369	,300	-8,333
Gesamtzufriedenheit	Varianzen sind nicht gleich			-1,060	257,815	,290	-8,333



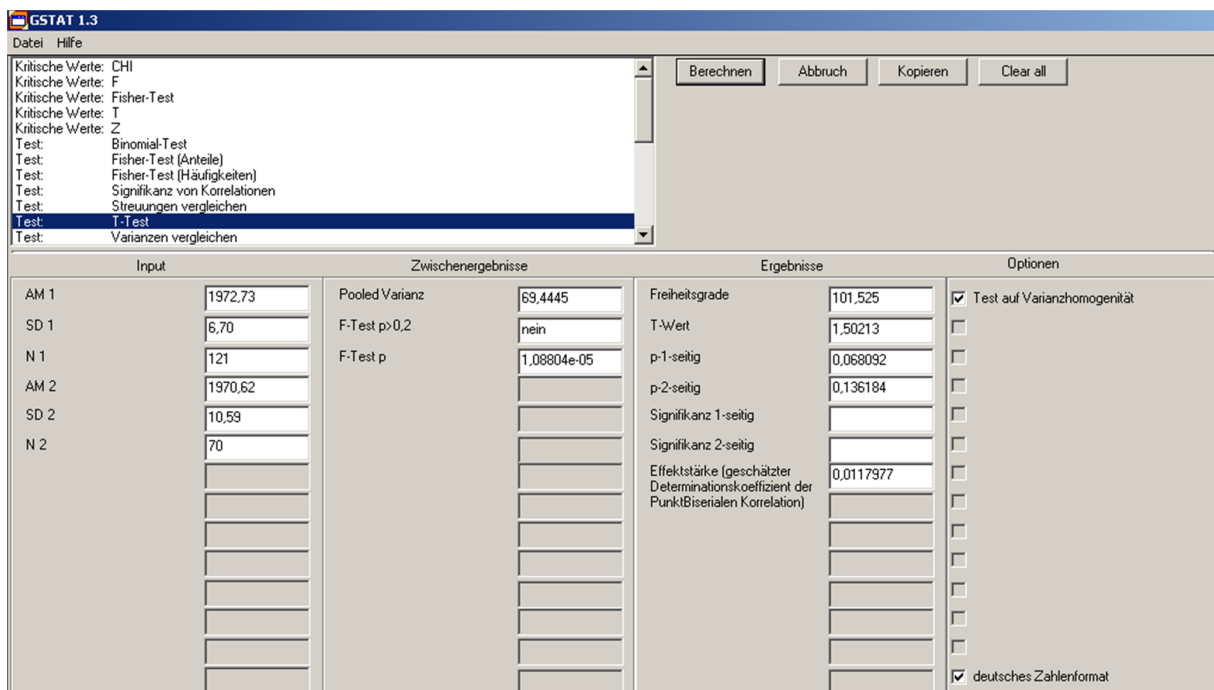
1- oder 2-seitig

- **Ungerichtete Hypothese:** Die von SPSS angegebene Wahrscheinlichkeit ist 2-seitig zu interpretieren.
- **Gerichtete Hypothese:** Für den 1-seitigen Wert gilt, dass er die Hälfte des 2-seitigen Wertes beträgt.

GStat

T-Test

T-Test GStat

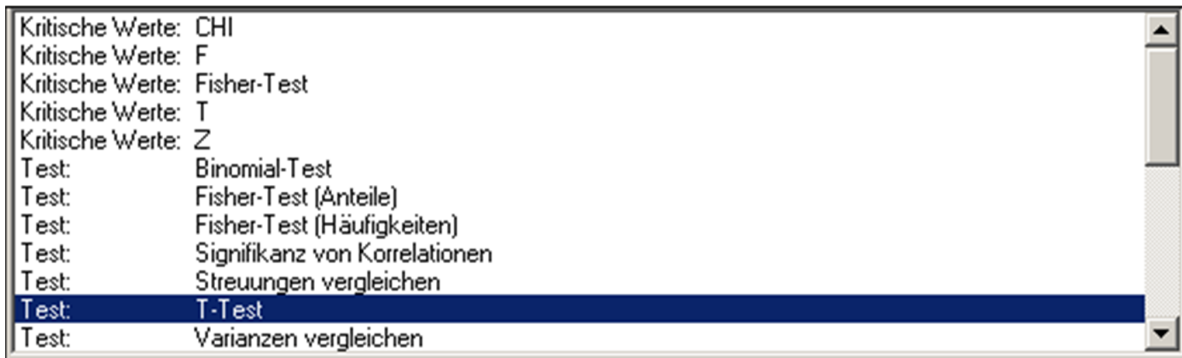


The screenshot shows the GSTAT 1.3 software interface. The 'Test' dropdown menu is set to 'T-Test'. The results table is as follows:

Input		Zwischenergebnisse		Ergebnisse		Optionen	
AM 1	1972,73	Pooled Varianz	69,4445	Freiheitsgrade	101,525	<input checked="" type="checkbox"/>	Test auf Varianzhomogenität
SD 1	6,70	F-Test p<0,2	nein	T-Wert	1,50213	<input type="checkbox"/>	
N 1	121	F-Test p	1,08804e-05	p-1-seitig	0,068092	<input type="checkbox"/>	
AM 2	1970,62			p-2-seitig	0,136184	<input type="checkbox"/>	
SD 2	10,59			Signifikanz 1-seitig		<input type="checkbox"/>	
N 2	70			Signifikanz 2-seitig		<input type="checkbox"/>	
				Effektstärke (geschätzter Determinationskoeffizient der Punktbiserialen Korrelation)	0,0117977	<input type="checkbox"/>	
						<input type="checkbox"/>	
						<input type="checkbox"/>	
						<input type="checkbox"/>	
						<input type="checkbox"/>	
						<input type="checkbox"/>	
						<input checked="" type="checkbox"/>	deutsches Zahlenformat

T-Test GStat

Verfahren auswählen



T-Test GStat

Werte eintragen (vorher ausrechnen)

Input	
AM 1	1972,73
SD 1	6,70
N 1	121
AM 2	1970,62
SD 2	10,59
N 2	70

Tabelle 1: Geburtsjahr nach Geschlecht

Geschlecht	Mittelwert	Streuung	N
Frauen	1972,73	6,70	121
Männer	1970,62	10,59	70
Gesamt	1971,38	9,40	191

T-Test GStat Ergebnisse ablesen

Ergebnisse	
Freiheitsgrade	101,525
T-Wert	1,50213
p-1-seitig	0,068092
p-2-seitig	0,136184
Signifikanz 1-seitig	
Signifikanz 2-seitig	
Effektstärke (geschätzter Determinationskoeffizient der PunktBiserialen Korrelation)	0,0117977

T-Test GStat Zwischenergebnisse zum F-Test

Zwischenergebnisse	
Pooled Varianz	69,4445
F-Test $p > 0,2$	nein
F-Test p	1,08804e-05

Da der F-Test für den T-Test fehlschlägt (Varianzhomogenität liegt nicht vor), wird automatisch die Korrektur berechnet.

Daher ergeben sich bei den Freiheitsgraden andere Werte als $n_1 + n_2 - 2$ (vgl. vorherige Folie).

Excel

T-Test

T-Test im Excel

TTEST			
Matrix1	B1:B22		= {0;0;0;0;0;0;0;0;0;0}
Matrix2	C1:C15		= {0;0;0;0;0;0;0;0;0;0}
Seiten	2		= 2
Typ	2		= 2
		= TTEST(B1:B22;C1:C15;2;	
Liefert die Teststatistik eines Student'schen t-Tests.			
Seiten bestimmt die Anzahl der Endflächen (Schwänze).			
	Formelergbnis = TTEST(B1:B22;C1:C15;2;:	OK	Abbrechen

Mit = eine Formel beginnen und als Funktion den T-Test wählen.
 Matrix1 enthält die Daten der einen und Matrix2 der anderen Stichprobe.
 Seiten gibt an, ob 1- oder 2-seitig getestet werden soll.
 Typ (1: gepaarte Daten; 2: Varianzhomogenität liegt vor; 3: Varianzhomogenität liegt nicht vor)
 Varianzhomogenität muss extra geprüft werden.
 Die Formel liefert nur die Wahrscheinlichkeit, also nur eine einzige Zahl.

χ^2 -Test

χ^2 -Test

Vergleich erwarteter und beobachteter Häufigkeiten:

$$\text{Testgröße: } \chi^2 = \sum_{j=1}^k \frac{(f_{b(j)} - f_{e(j)})^2}{f_{e(j)}}$$

χ^2 -Methoden basieren auf einen Vergleich von beobachteten f_b und erwarteten Häufigkeiten f_e .

Die Freiheitsgrade FG eines χ^2 -Wertes entsprechen der Anzahl der Summanden k abzüglich der Bestimmungsstücke für die Berechnung der erwarteten Häufigkeiten, die aus den beobachteten Daten abgeleitet werden (in der Regel $FG=k-1$).

Der Test darf nur angewendet werden, wenn nicht mehr als 20% der erwarteten Häufigkeiten Werte kleiner 5 aufweisen und keine der erwarteten Häufigkeiten Werte kleiner 1 aufweisen.

Varianzanalyse (ANOVA)

Einfaktorielle Varianzanalyse (ANOVA, Analysis Of Variance) 1

- Die einfaktorielle Varianzanalyse untersucht den Zusammenhang einer mehrkategorialen Gruppierungsvariablen (unabhängigen Variablen) mit einer metrischen abhängigen Variablen.
- Die Alternativ-Hypothese behauptet, dass signifikante Mittelwertsunterschiede zwischen den p Populationsgruppen bestehen. Dagegen sind nach der Nullhypothese die Mittelwerte der abhängigen Variablen in den Populationsgruppen gleich.
- Prinzipiell könnten hier auch mehrere T-Tests gerechnet werden. Dabei ist aber zu beachten, dass die Wahrscheinlichkeit für ein signifikantes Ergebnis dadurch ansteigt und entsprechend korrigiert werden muss. (**Alpha-Fehler-Korrektur**)

Einfaktorielle Varianzanalyse (ANOVA, Analysis Of Variance) 2

- **Die Gesamtvarianz wird aufgrund der Varianz der Stichprobe(n) geschätzt:**

$$\hat{\sigma}_{tot}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

X_i meint hier alle gemessenen Werte unabhängig vom Treatment (Gruppe, Kategorie)

- **Bestimmung der Treatmentvarianz:**

Hat das Treatment einen Einfluss, so sollten sich die Mittelwerte der einzelnen Faktorstufen vom Gesamtmittelwert unterscheiden.

Man ersetzt daher bei allen Faktorstufen die Messwerte x_i durch den jeweiligen Mittelwert der Faktorstufe um die Abweichungen der Faktorstufen vom Gesamtmittelwert zu bestimmen:

$$\hat{\sigma}_{treat}^2 = \frac{\sum_i^p n_i \cdot (\bar{x}_i - \bar{x})^2}{p-1}$$

j steht hier für die Faktorstufen.

Einfaktorielle Varianzanalyse (ANOVA, Analysis Of Variance) 3

- **Bestimmung der Fehlervarianz:**

Wenn die abhängige Variable ausschließlich von der unabhängigen Variablen und von keinen anderen Faktoren (Störgrößen) beeinflusst würde, müssten alle Messwerte innerhalb einer Faktorstufe gleich sein, also dem Mittelwert der entsprechenden Stichprobengruppe entsprechen. Der Einfluss der Störvariablen wird also gerade durch das Ausmaß charakterisiert, in dem Messwerte vom Mittelwert der jeweiligen Faktorstufe abweichen. Hieraus folgt für die geschätzte Fehlervarianz aller Faktorstufen:

$$\hat{\sigma}_{Fehler}^2 = \frac{\sum_i^p \sum_m^{n_i} (x_m - \bar{x}_i)^2}{n-p}$$

j steht hier für die Faktorstufen.
 m steht für die Messwerte in den Faktorstufen.

Einfaktorielle Varianzanalyse (ANOVA, Analysis Of Variance) 4

	Gesamt	Treatment	Fehler
Varianz	$\hat{\sigma}_{tot}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$	$\hat{\sigma}_{treat}^2 = \frac{\sum_i n_i \cdot (\bar{x}_i - \bar{x})^2}{p-1}$	$\hat{\sigma}_{Fehler}^2 = \frac{\sum_i \sum_m^{n_i} (x_m - \bar{x}_i)^2}{n-p}$
Quadratsumme	$QS_{tot} = \sum_{i=1}^n (x_i - \bar{x})^2$	$QS_{treat} = \sum_i n_i \cdot (\bar{x}_i - \bar{x})^2$	$QS_{Fehler} = \sum_i \sum_m^{n_i} (x_m - \bar{x}_i)^2$
Freiheitsgrade	$n - 1$	$p - 1$	$n - p$

Der prozentuale Anteil an der Aufklärung der Gesamtvarianz, die auf die Varianz der unabhängigen Variable zurückzuführen ist (Varianzaufklärung), beträgt:

$$\frac{QS_{treat}}{QS_{tot}} \cdot 100\%$$

Einfaktorielle Varianzanalyse (ANOVA, Analysis Of Variance) 5

- **Der eigentliche Signifikanztest:**

Im Signifikanztest wird das Verhältnis von Treatmentvarianz zur Fehlervarianz getestet. Als Prüfgröße wird ein F-Wert ermittelt.

Eine relativ große Varianzaufklärung (und damit eine große Treatmentvarianz) führt zu einem hohen F-Wert.

Wenn es dagegen keine Unterschiede zwischen den Mittelwerten der einzelnen Faktorstufen und damit keinen Zusammenhang zwischen der unabhängigen und abhängigen Variablen gibt, so ist die Treatmentvarianz im Vergleich zur Fehlervarianz relativ klein und wir erhalten einen relativ kleinen F-Wert:

$$F = \frac{\hat{\sigma}_{treat}^2}{\hat{\sigma}_{Fehler}^2}$$

$$FG_{Zähler} = p - 1, FG_{Nenner} = n - p$$

Vergleich mehrerer Mittelwerte/Gruppen/Treatments 2

- Für mehrfaktorielle Designs gibt es mehrfaktorielle Varianzanalysen.
- Spezielle Varianzanalysen sind für Messwiederholungsdesigns entwickelt worden.
- Der Unterschied zwischen *fixed* und *random factors* spielt bei der mehrfaktoriellen Varianzanalyse eine wichtige Rolle.
- Bei der mehrfaktoriellen Varianzanalyse können Haupteffekte und Interaktionseffekte einzeln geprüft werden.
- Voraussetzungen sind Intervallskalenniveau, Normalverteilung (bei n pro Stufe > 25 immer gegeben und kein Problem), Varianzhomogenität zwischen den Gruppen. Diese kann getestet werden mit einem F-Test, bei dem $F = s^2_{\max} / s^2_{\min}$; FG sind jeweils n der Gruppe minus Eins.
- Bei schlechterer Datenqualität oder anderen Verletzungen der Voraussetzungen kann der **Kruskal-Wallis-Test** gerechnet werden.
- Eine Varianzanalyse sagt nur, dass Unterschiede zwischen den Gruppen bestehen, nicht jedoch wo. Dafür gibt es so genannte **Kontraste**.

Inferenzstatistik – Zusammenhangshypothesen

Korrelationsrechnung

- Für die Überprüfung des Zusammenhanges zwischen zwei intervallskalierten Datensätzen wird die Produkt-Moment-Korrelation (auch Pearson- oder Bravais-Pearson-Korrelation genannt) berechnet.

$$r = \frac{\text{COV}(x,y)}{s_x s_y} \quad \text{COV}(x,y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

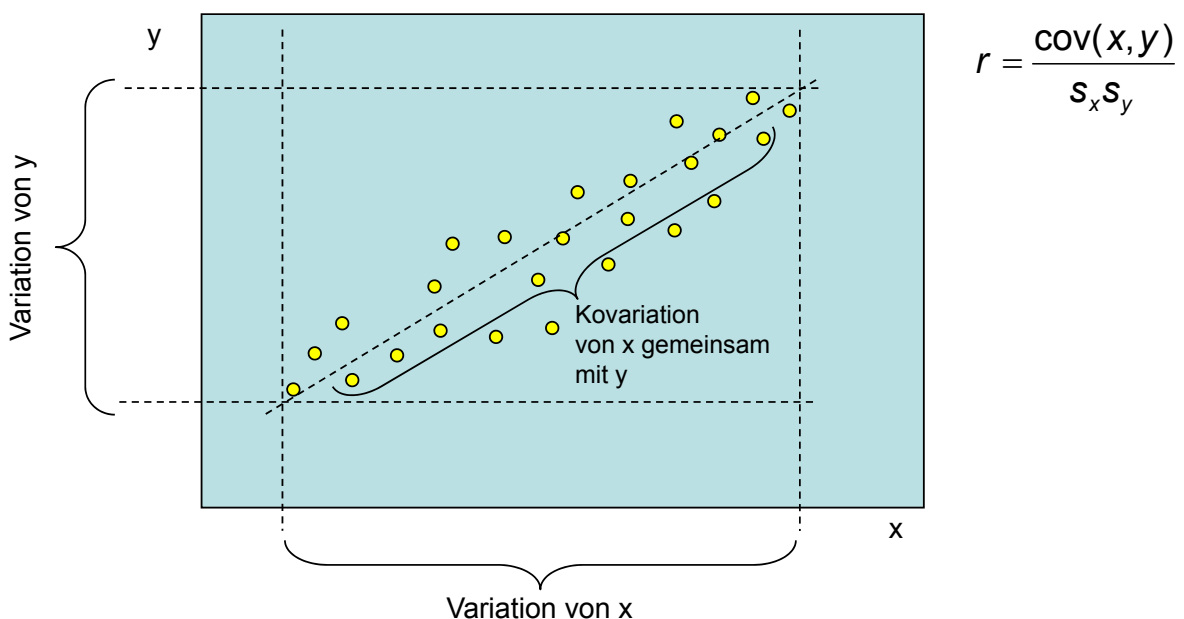
r steht für die Korrelation,

s sind die Standardabweichungen der beiden Variablen x und y

Cov(x,y) ist die Kovarianz, also die gemeinsame Varianz der beiden Variablen.

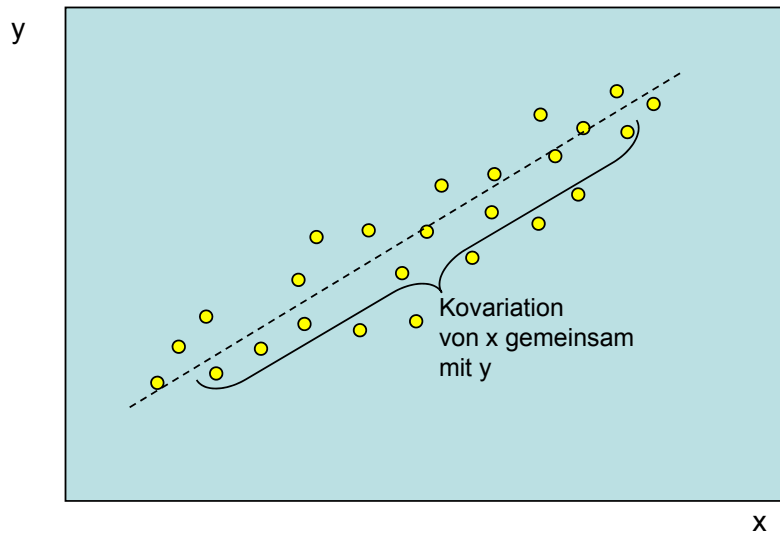
- Ist $r < 0$ liegt eine negative Korrelation vor. Ist $r > 0$ eine positive. Ist $r = 0$ liegt keine Korrelation vor.
- r kann maximal -1 bzw. 1 betragen.

Korrelationsrechnung 2



Die eingezeichnete Variationsbreite ist nicht identisch mit der Streuung oder der Varianz, weil zu groß. Dadurch wird die Grafik jedoch klarer.

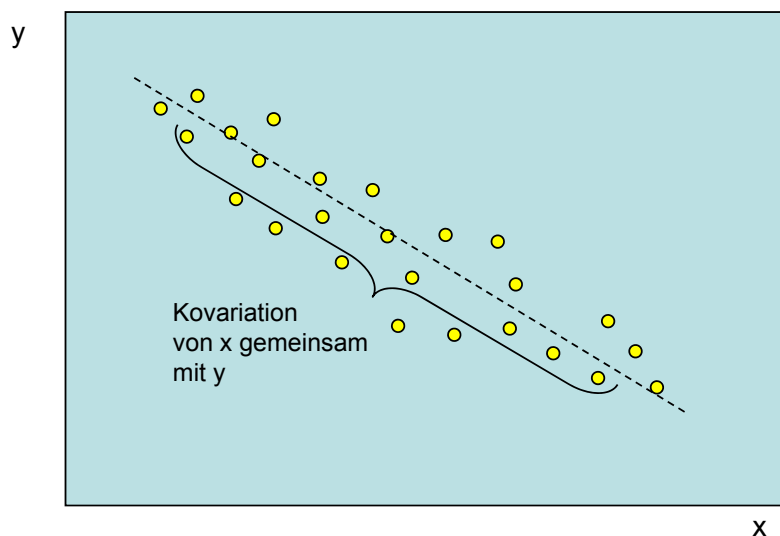
Korrelationsrechnung 3 Positive Korrelation



$$r = \frac{\text{cov}(x,y)}{s_x s_y}$$

positive Korrelation,
 $r > 0$

Korrelationsrechnung 4 Negative Korrelation

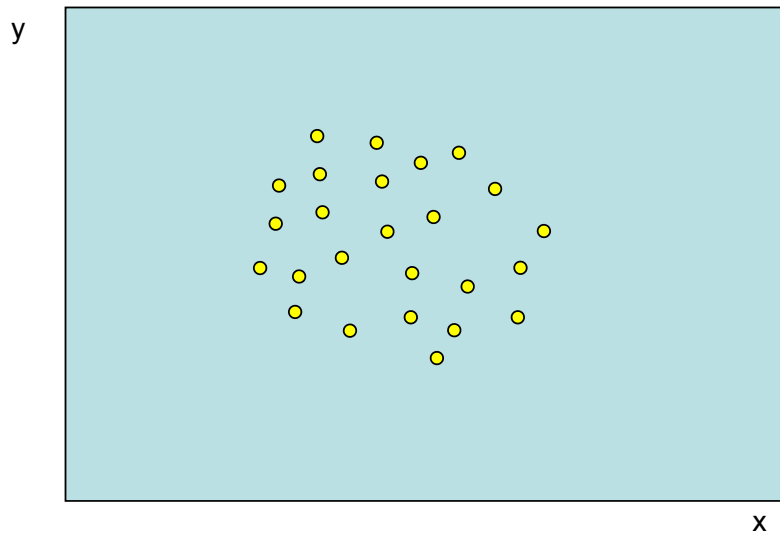


$$r = \frac{\text{cov}(x,y)}{s_x s_y}$$

negative Korrelation,
 $r < 0$

Korrelationsrechnung 5

Null-Korrelation



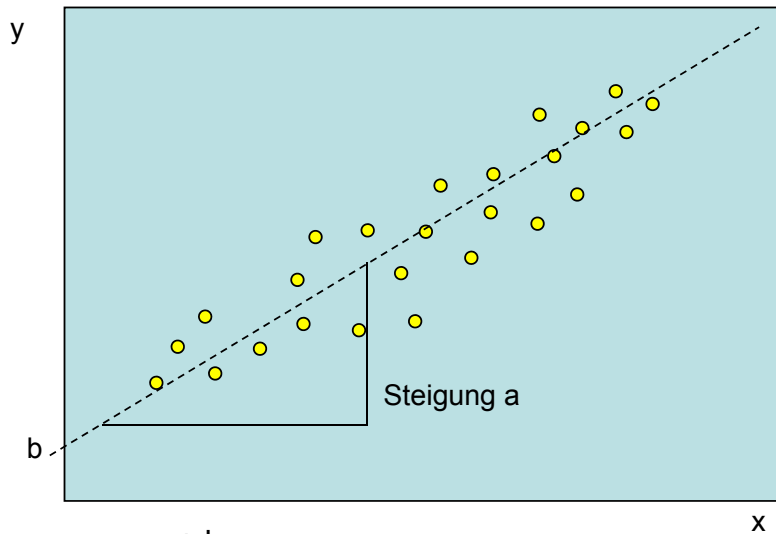
$$r = \frac{\text{cov}(x,y)}{s_x s_y}$$

keine Korrelation,
 $r = 0$

Beurteilung von Korrelationen

- Korrelationen müssen auf Signifikanz geprüft werden.
- Der Determinationskoeffizient $D = r^2$ gibt zudem die Varianzaufklärung an.
- In der Psychologie liegen viele Korrelationen nicht höher als 0,3.
- In den Naturwissenschaften sind 0,9 keine Seltenheit.
- Auch Reliabilität und Validität (siehe oben) sind Korrelationen. Eine Reliabilität ist z.B. erst ab 0,8 ausreichend.

Regressionsrechnung



$$y = ax + b$$

$$a = \frac{\text{cov}(x,y)}{s_x^2} = \frac{rs_x s_y}{s_x^2}$$

$$b = \bar{y} - a\bar{x}$$

Korrelationsverfahren

y \ x	Intervall	Dichotom	Ordinal
Intervall	Produkt-Moment-Korrelation (Pearson)	Punktbiseriale Korrelation (Alternativ: t-Test)	Rangkorrelation (Spearman)
Dichotom		Phi-Koeffizient (über Chi-Quadrat). Bei 1/0 Kodierung ist die Produkt-Moment-Korrelation identisch mit Phi.	Biseriale Rangkorrelation
Ordinal			Rangkorrelation

Die Produkt-Moment-Korrelation kann auch im Excel bestimmt werden.

Korrelationsverfahren 2

- **Produkt-Moment-Korrelation**

$$r = \frac{\text{COV}(x,y)}{s_x s_y} \quad \text{COV}(x,y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- **Punktbiseriale Korrelation**

$$r_{pb} = \frac{\bar{y}_1 - \bar{y}_0}{s_y} \cdot \sqrt{\frac{n_0 \cdot n_1}{n^2}}$$

Die Indizes 0 und 1 bezeichnen die beiden Ausprägungen der x-Variable.
 $n = n_0 + n_1$

- **Spearman's Rangkorrelation**

$$r_s = 1 - \frac{6 \cdot \sum_{i=1}^n d_i^2}{n \cdot (n^2 - 1)}$$

d_i bezeichnet die Differenz zwischen x_i und y_i .
 x und y sind dabei Ränge.
 Liefert gleiche Ergebnisse wie die Produkt-Moment-Korrelation, wenn die Ränge mit 1, 2, 3, 4... bezeichnet wurden.
 Probleme bei Rangbindungen (gleichen Rängen sog. **ties**).
 n = Zahl der Differenzen.

Kontingenztafeln

	männlich	weiblich	
mit Brille	25 a	10 b	35
ohne Brille	c 25	d 40	65
	50	50	100

$$\chi^2 = \frac{n \cdot (ad - bc)^2}{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)} \quad \text{FG} = 1$$

$$\text{Phi} = \Phi = \frac{ad - bc}{\sqrt{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)}}$$

Phi ist in etwa mit einem r vergleichbar, liegt aber nicht immer zwischen -1 und 1.

Multiple Korrelation / Regression

- Soll eine Kriteriumsvariable y durch viele verschiedene Prädiktorvariablen x_1, x_2, \dots, x_p beschrieben werden, so ergibt sich eine Gleichung der folgenden Form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_p x_p$$

- Die β -Gewichte werden mit dem Verfahren der multiplen Regression berechnet.
- Für jedes β -Gewicht muss die Signifikanz bestimmt werden.
- Es ergibt sich zudem eine multiple Korrelation, da jede Prädiktorvariable einen Teil zur Varianzaufklärung beiträgt ist die Gesamtkorrelation R größer als die einzelnen Korrelationen zwischen Prädiktor und Kriterium.

Software

Excel

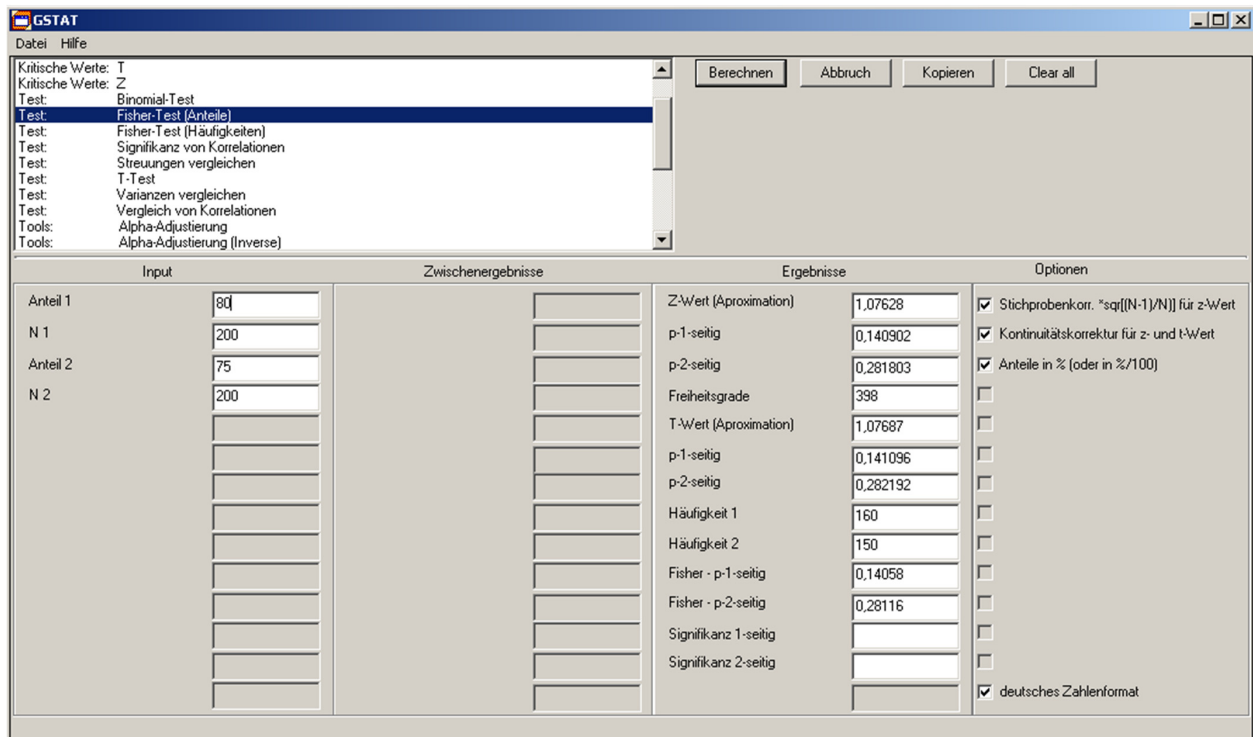
Was geht im Excel?

- Im Excel können Mittelwert, Median, Standardabweichung und Varianz (jeweils für Stichprobe oder Population) berechnet werden.
- Als Testverfahren sind der t-Test und der F-Test vorhanden.
 - Beim t-Test muss angegeben werden, ob von Varianzhomogenität ausgegangen werden kann.
 - Dazu muss zunächst der F-Test für die Varianzen gerechnet werden. Ist das Ergebnis (Wahrscheinlichkeit p) größer 0,2, so kann von Varianzhomogenität ausgegangen werden. Ist $p \leq 0,2$, dann sollte der t-Test für ungleiche Varianzen gerechnet werden.
- Auch die Chi-Verteilung ist im Excel vorhanden. Damit kann die Wahrscheinlichkeit (1-Alpha) für einen berechneten χ^2 -Wert bestimmt werden.
- Korrelationen: Eine einfache Produktmoment-Korrelation kann bestimmt werden.

GStat

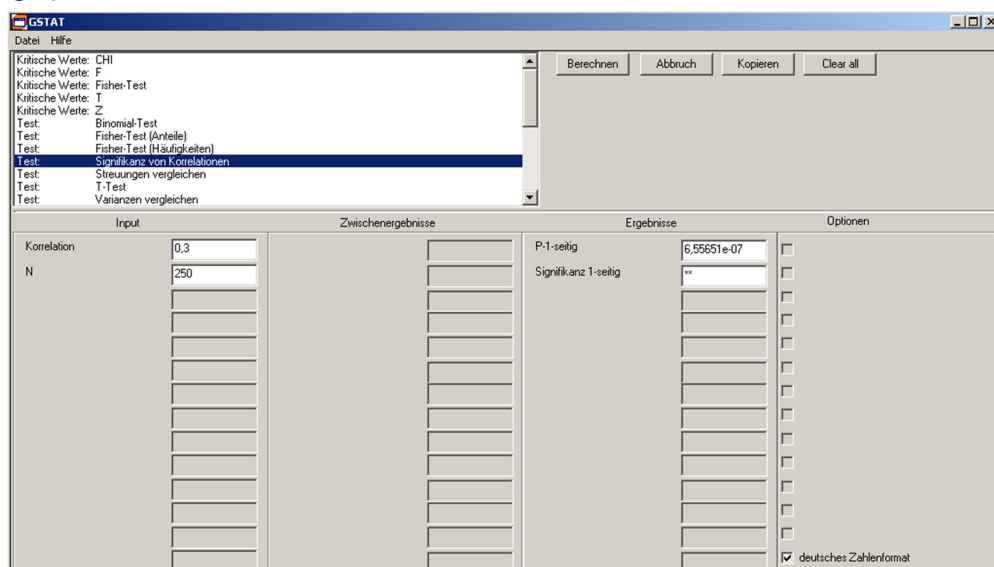
GStat

- Unter www.complexity-research.com, Menü: Downloads, Software kann Gstat heruntergeladen werden.
 - Sind Mittelwerte oder andere Kennwerte bekannt, kann mit Gstat die Signifikanz bestimmt werden.
 - Beispiel: Fisher-Test für den Vergleich von zwei Prozentwerten.
 - Wenn in zwei Gruppen jeweils 200 Personen befragt wurden und in der einen Gruppe 80% die Küche des Spitals für miserabel halten und in der anderen Stichprobe 75% stellt sich die Frage, ob der Unterschied hier signifikant ist.



Korrelationsrechnung Signifikanztestung in GStat

- Die Null-Hypothese geht davon aus, dass eine Null-Korrelation vorliegt.
 $H_0: r = 0$.
- Es kann 1-seitig (Vorhersage der Korrelationsrichtung) oder 2-seitig geprüft werden..



Statistisches Arbeiten mit SPSS

Priv.-Doz. Dr. Dr. Guido Strunk
guido.strunk@complexity-research.com

Themenübersicht

- Struktur und Aufbau des Programmpaketes SPSS.
- Grundsätzliches zum Arbeiten mit SPSS.
- Erstellung von Datenfiles und Kodierung von Fragebögen.
- Grundsätzliches zur Datenauswertung mit SPSS.
- Deskriptive Statistiken in SPSS.
- T-Test, 1-faktorielle Varianzanalyse.
- Korrelationen.
- Nichtparametrische Verfahren.
- Hilfsfunktionen, Anleitungen und Literatur.

Struktur und Aufbau von SPSS

Daten-Editor

Jede Zeile ist ein Fall (Case) Jede Spalte ist eine Variable Fehlende Angaben bleiben leer

1:auf

	auf	dname	pre_4	pre_5	pre_6	pre_7	pre_8	pre_9	pre_10	pre_11	pre_12	pre_13
1	,00	10000000011003447800	,00335	,00278	,00231	,00608	,00642	,00542	,00483	,00262	,00256	.
2	1,00	10000000021000008000	,00967	,01632	,00909	,02914	,07805	,05772	,01527	,01984	,03887	.
3	,00	10000000021003129000	,01798	,01008	,00908	,02464	,04574	,02504	,02892	,07680	,02400	.
4	,00	10000000021003129002	,01369	,01284	,01036	,02582	,04574	,01488	,10846	,02987	,01369	.
5	1,00	10000000031000020800	,04884	,03309	,03323	,00631	,01493	,05772	,00718	,03120	,01414	,03628
6	,00	10000000031000321801	,00335	,00278	,00231	,00608	,00642	,00542	,00483	,00262	,00256	.
7	,00	10000000041003404059	,09380	,03256	,04314	,09198	,04574	,05893	,05543	,03130	,08271	,04516
8	1,00	10000000051000039000	,02555	,05478	,09004	,03710	,02658	,12425	,02192	,02113	,01414	,02165
9	,00	10000000051000902000	,07742	,02047	,03939	,07417	,01121	,01488	,01671	,01712	,01415	,04915
10	1,00	10000000061000040000	,04577	,15021	,07516	,09317	,02453	,02451	,11855	,02625	,04668	,12286
11	,00	10000000061002160000	,00734	,01714	,02063	,03583	,04488	,05772	,00903	,02743	,01777	.
12	,00	10000000071003252000	,02409	,02434	,04864	,01254	,00591	,02451	,01569	,01166	,01415	.
13	,00	10000000071003252059	,06266	,01113	,00980	,01793	,01493	,02451	,04191	,02113	,01414	.
14	,00	10000000071003252801
15	,00	10000000081001659800	,00335	,00278	,00231	,00608	,00642	,00542	,00483	,00262	,00256	.
16	1,00	10000000091000046000	,05876	,11082	,03937	,06420	,04334	,02451	,02953	,04591	,04668	,06569
17	,00	10000000091004390000	,05342	,03758	,02341	,01177	,01493	,01488	,01718	,03799	,04332	,08574
18	,00	10000000091004390002
19	,00	10000000091004390004	,00462	,00790	,00764	,00920	,01493	,01372	,00953	,01639	,01369	.
20	1,00	10000000101000050000	,06171	,03247	,07249	,04495	,01493	,02451	,22049	,06719	,03478	.
21	,00	10000000101004350800	,00335	,00278	,00231	,00608	,00642	,00542	,00483	,00262	,00256	.

Eine Zelle enthält Werte (Values) In der Regel wird mit Zahlen gearbeitet

Daten-Editor 2

	id	gender	bdate	edu	jobcat
1	1	m	02/03/52	15	3
2	2	m	05/23/58	16	1
3	3	f	07/26/29	12	1
4	4	f	04/15/47	8	1
5	5	m	02/09/55	15	1
6	6	m	08/22/58	15	1
7	7	m	04/26/56	15	1
8	8	f	05/06/88	12	1
9	9	f	01/23/46	15	1
10					

Datenansicht

Variablenansicht

Variablenansicht

	Name	Typ	Spaltenformat	Dezimalstellen	Variablenlabel
1	id	Numerisch	4	0	Employee Cod
2	gender	String	1	0	Gender
3	bdate	Datum	8	0	Date of Birth
4	edu	Numerisch	2	0	Educational Lev
5	jobcat	Numerisch	1	0	Employment Ca
6	salary	Dollar	8	0	Current Salary
7	salbegin	Dollar	8	0	Beginning Salar
8	jobtime	Numerisch	2	0	Months since Hi
9	prevexp	Numerisch	6	0	Previous Experi
10	minority	Numerisch	1	0	Minority Classifi

- Variablenname
- Datentyp
- Anzahl Ziffern oder Zeichen
- Anzahl Dezimalstellen
- Beschreibende Variablen- und Wertelabels
- Benutzerdefinierte fehlende Werte
- Spaltenbreite
- Messniveau

Programm-Struktur

The screenshot displays the SPSS interface with three main windows:

- Dateneditor:** Shows a data grid with columns 'auf', 'dname', 'pre_4', 'pre_5', and 'pre_6'. The 'auf' column contains values like .00, 1.00, .00, .00, 1.00, .00, .00, 1.00.
- Output-Viewer:** Displays a tree view for 'Deskriptive Statistik' and a table of descriptive statistics.

Deskriptive Statistik					
	N	Minimum	Maximum	Mittelwert	Standardabweichung
PRE_5 Predicted Value	7366	,00225	,96641	2,38E-02	3,671929E-02
PRE_6 Predicted Value	7140	,00076	,99971	2,13E-02	3,579119E-02
PRE_7 Predicted Value	7644	,00404	,43493	2,73E-02	3,479088E-02
PRE_8 Predicted Value	7366	,00188	,99970	2,67E-02	2,882713E-02
	7140				
- Syntax-Editor:** Shows the following command:


```
DESCRIPTIVES
VARIABLES=pre_5 pre_6 pre_7 pre_8
/STATISTICS=MEAN STDDEV MIN MAX.
```

Datei-Endungen

The screenshot is identical to the one above, but with callout boxes indicating file extensions:

- Dateneditor:** .sav
- Output-Viewer:** .spo
- Syntax-Editor:** .sps

Grundsätzliches zum Arbeiten mit SPSS

Statistisches Wissen

- **SPSS ist ein Statistik-Programm für Profis.**
- **Es umfasst eine Fülle von Funktionen, statistischen Methoden, Algorithmen und Varianten von Methoden.**
- **Sie müssen wissen was Sie tun. Das Programm entscheidet nicht über die Sinnhaftigkeit einer Berechnung: Es rechnet auch den Mittelwert von Hausnummern aus.**
- **Die Ergebnisse der Berechnungen müssen von Ihnen interpretiert werden.**
- **Nicht immer benutzt SPSS die gleichen Methoden und Methodenbezeichnungen, wie gängige Statistik-Bücher.**

Unterschiede zu Excel

- **SPSS ist bewusst unflexibel bei der Dateneingabe.**
- **SPSS ist für die Analyse großer Datensätze geeignet. Auch einfache Fragebögen führen schnell zu über 255 Variablen. Internationale Datensätze können schnell auch mal über 100.000 Cases enthalten.**
- **Berechnungen liefern im SPSS eine Fülle von Detail- und Zwischenergebnissen, die automatisch in Tabellen und Grafiken dargestellt werden. Im Excel wird nur eine Zahl berechnet.**
- **SPSS verfügt über keine automatische Aktualisierung von Ergebnissen oder von Zahlen im Dateneditor.**
- **(in älteren Versionen ist SPSS in den Variablennamen beschränkt auf 8 Zeichen – keine Umlaute – und kann nur jeweils ein Tabellenblatt geöffnet haben.)**

Sonstige Besonderheiten

- **Man kann nur rechnen, wenn die Daten in einem Datenblatt vorhanden sind. Der Vergleich zweier Datenblätter oder eines Datenblattes mit Werten aus der Literatur ist nur schwer oder gar nicht möglich.**
- **Tabellen und Grafiken sind nur selten direkt in wissenschaftlichen Arbeiten nutzbar. Sie enthalten in der Regel mehr Infos als man benötigt oder sind nur schwer zu formatieren. Ein Austausch der Ergebnisse über die Zwischenablage ist aber gut möglich.**
- **SPSS kostet zwischen 1.000 und 5.000 Euro. Studentenversionen und Demoversionen sind jedoch erhältlich. PSPP kann als Alternative schon recht viel und wird beständig weiterentwickelt.**

Erstellung von Datenfiles und Kodierung von Fragebögen

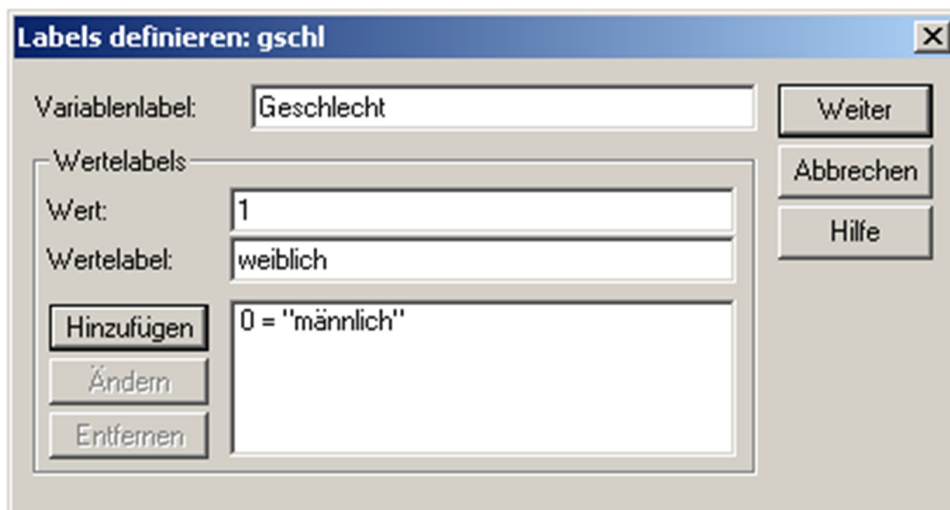
Definition der Variablen in SPSS

- **Variablenname (8 Zeichen, keine deutschen Sonderzeichen, jeden Namen nur ein mal vergeben), z.B.:**
 - lfnr
 - item1
 - item2
 - item3
 - item4_1
 - item4_2
 - item4_3
 - geschl
 - Alter
- **Datentyp (in der Regel „Numerisch“, aber auch „String“ für Texte und „Datum“ üblich)**

Definition der Variablen in SPSS

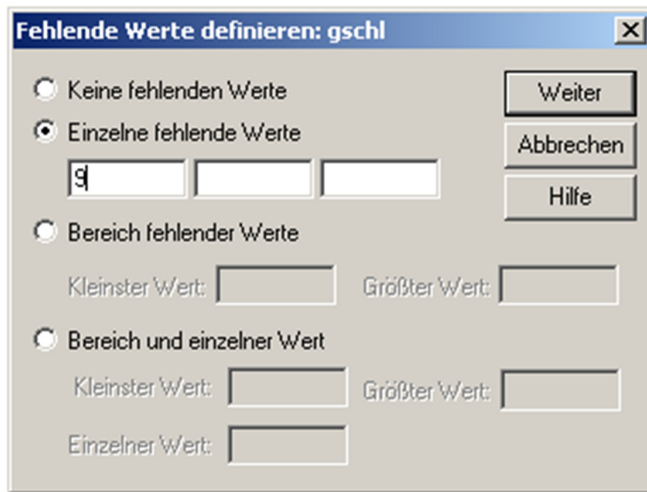
- **Anzahl Ziffern oder Zeichen (bei Text maximal 255)**
- **Anzahl Dezimalstellen (bei ganzen Zahlen sieht es schöner aus, wenn hier 0 eingegeben wird, ist aber egal)**
- **Beschreibende Variablen- und Wertelabels**
- **Benutzerdefinierte fehlende Werte**
- **Spaltenbreite (nur für die Anzeige relevant)**
- **Messniveau (Nominal, Ordinal, Metrisch). Dient als Gedächtnisstütze, hat aber sonst keine Bedeutung.**

Beschreibende Wertelabels



Zur Vergabe von Wertelabels den Wert und den Label eingeben und mit „Hinzufügen“ zur Liste hinzufügen. Am Schluss mit „Weiter“ den Dialog verlassen.

Benutzerdefinierte fehlende Werte



„weiß ich nicht“ ist eine Angabe der befragten Person, die kodiert werden sollte, aber bei Berechnungen ignoriert werden muss. Es handelt sich um einen „Definierten fehlenden Wert“ (Missing).

Fehlen Angaben im Fragebogen, kann bei der Dateneingabe einfach die Zelle im Dateneditor übersprungen werden. Man spricht hier von einem „System Missing“.

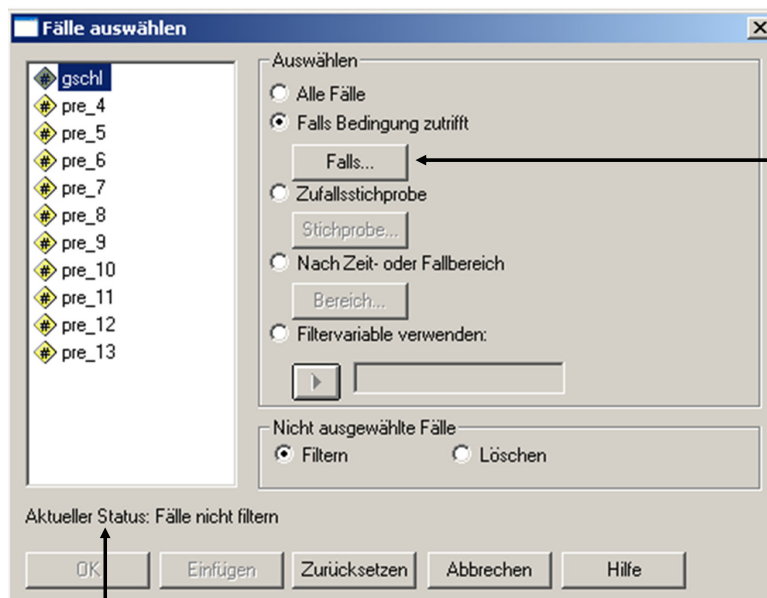
Eine der vier Optionen wählen und geforderte Angaben ausfüllen.

Am Schluss mit „Weiter“ den Dialog verlassen.

Vorbereitende Schritte

- Soll nur mit einem Teil der Daten gerechnet werden?
(Falls ja => Menü: Daten / Fälle auswählen ...)
- Müssen aus den Daten zunächst neue Variablen erzeugt werden?
(Falls ja => Menü: Transformieren / Berechnen ...)

Fälle auswählen ...



Zum
Festlegen der Bedingung
auf „Falls...“ klicken

Status

Fälle auswählen ... Falls...

Formel für die Bedingung

~ bedeutet „nicht“

Variablenliste

Kopiert eine ausgewählte Variable in den Formel-Editor

Sobald die Formel steht, Kann der Dialog mit „Weiter“ geschlossen werden.

Fälle auswählen ... Falls ... beendet

Bedingung wird angezeigt

Löschen?

„OK“ und „Einfügen“ wird verfügbar

Filter an

The screenshot shows the SPSS Daten-Editor window with a data table. The table has columns: gschl, dname, pre_4, pre_5, pre_6, and pre_7. The rows are numbered 1 to 8. The 'gschl' column contains values 0, 1, 0, 0, 1, 0, 0, 1. The status bar at the bottom right indicates 'Filter an'.

Annotations:

- Auswahl der Fälle**: Points to the first three rows of the data table.
- Filterstatus**: Points to the 'Filter an' status in the bottom right corner of the window.

Transformieren / Berechnen ...

The screenshot shows the 'Variable berechnen' dialog box. The 'Zielvariable:' field contains 'gesamt'. The 'Numerischer Ausdruck:' field contains the formula 'pre_4 + pre_5 + pre_6 + pre_7'. A list of variables is shown on the left, including 'gschl', 'dname', 'pre_4', 'pre_5', 'pre_6', 'pre_7', 'pre_8', 'pre_9', 'pre_10', 'pre_11', and 'pre_12'. A list of functions is shown on the right, including 'ABS', 'ANY', 'ARSIN', 'ARTAN', 'CDFNORM', and 'CDF.BERNOULLI'.

Annotations:

- Formel**: Points to the 'Numerischer Ausdruck:' field containing the formula.
- Variablenliste**: Points to the list of variables on the left.
- neue Zielvariable definieren**: Points to the 'Zielvariable:' field.
- Kopiert eine ausgewählte Variable in den Formel-Editor**: Points to the mouse cursor clicking on the 'pre_4' variable in the list.

Transformieren / Berechnen ... Falls...

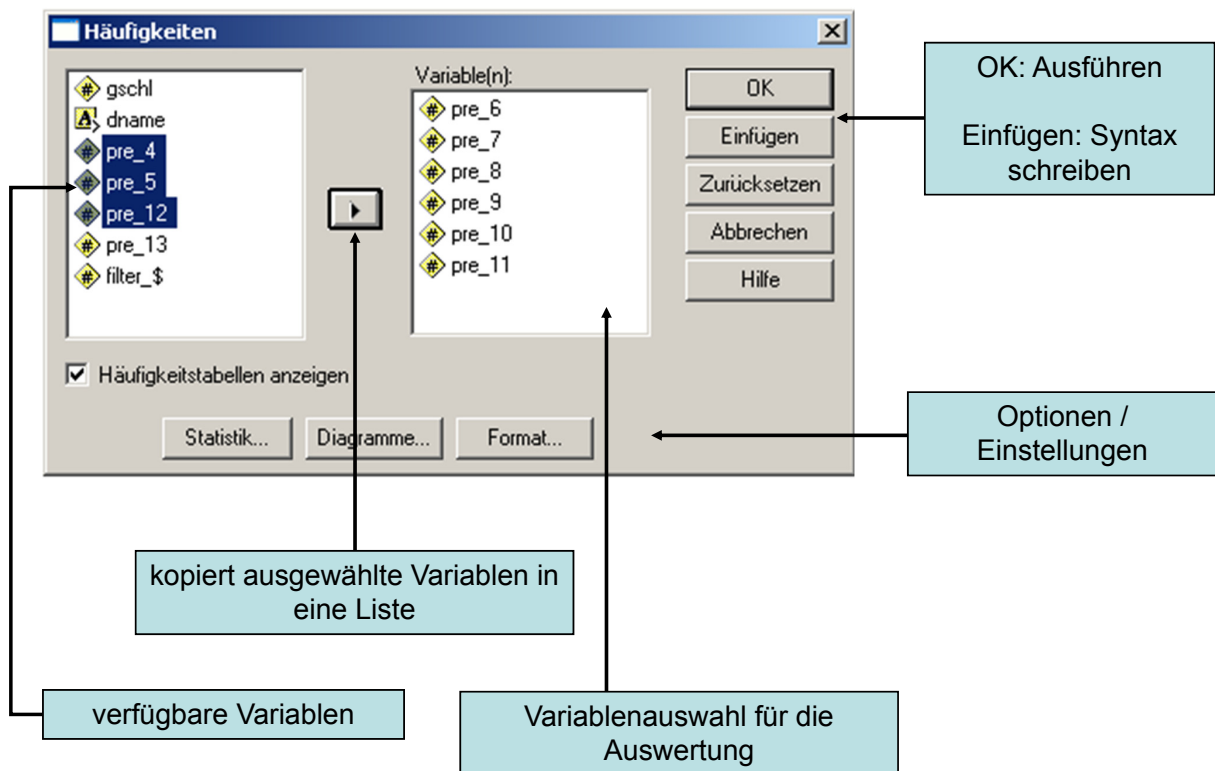
neue Zielvariable definieren
Gruppe=1
Die neue Variable „gruppe“ bekommt den Wert „1“ falls...

Transformieren / Berechnen ... Falls...

1. Fall einschließen, wenn...
2. Bedingung
3. „Weiter“

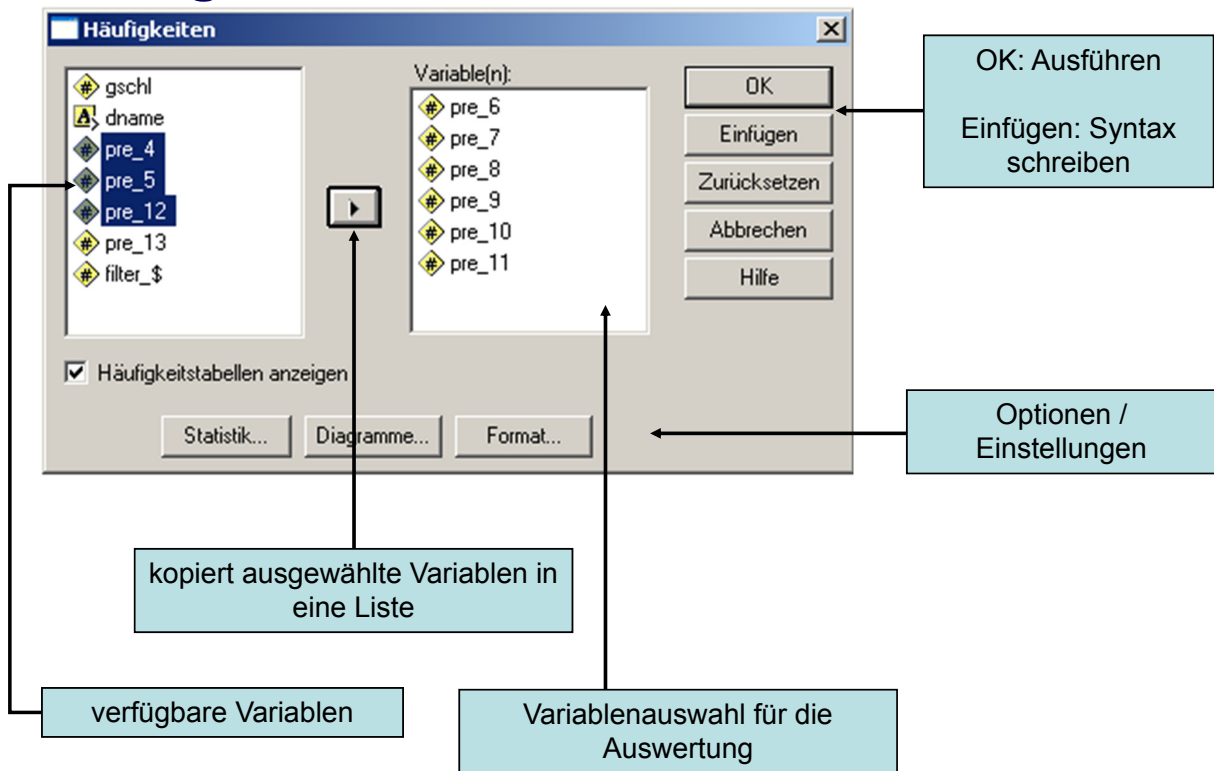
Und-Verknüpfung: AND
Oder-Verknüpfung: OR

Typischer SPSS-Dialog

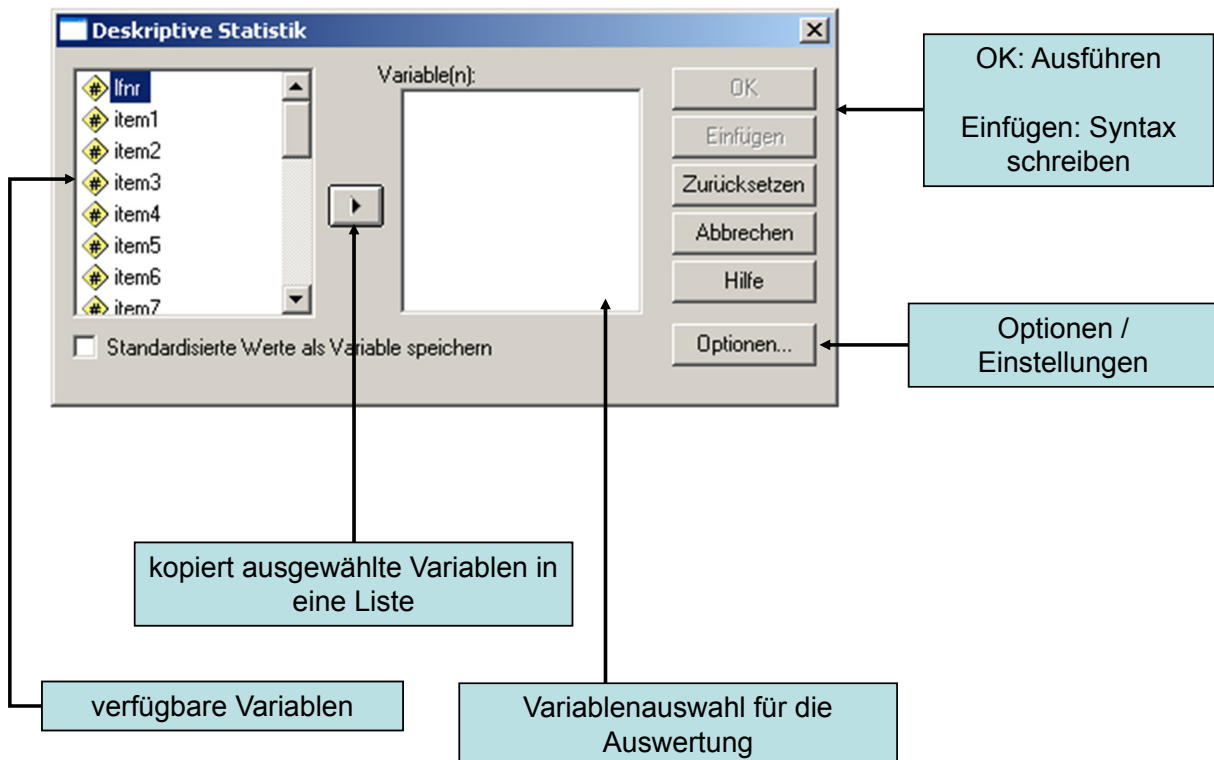


Deskriptive Statistiken in SPSS

Statistik \ Zusammenfassen \ Häufigkeiten



Statistik \ Zusammenfassen \ Deskriptive Statistiken



Hilfefunktionen, Anleitungen und Literatur

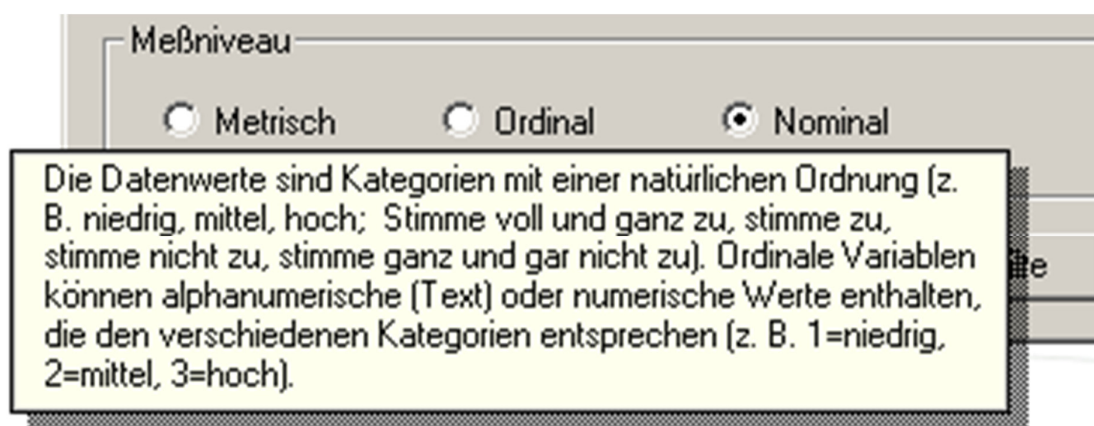
Hilfen...

- Die Hilfe von SPSS ist durchaus komfortabel. Es gibt einen Statistik-Assistenten und für alle Teile des Programmpaketes gibt es Online-PDF-Handbücher.
- Die Pop-Up-Hilfe bietet eine gute Unterstützung während der Arbeit.
- Zahlreiche Handbücher zum Arbeiten mit SPSS sind im Buchhandel zu haben. Auch zahlreiche SPSS-Handbücher zu aufwändigen Verfahren sind erschienen. Da ständig neue SPSS-Versionen erscheinen wird hier keine Empfehlung gegeben.
- Einige Statistikkbücher geben auch eine kurze Einführung zu Rechnen in SPSS, so z.B. in: Bortz, J. (1989). Statistik für Sozialwissenschaften. Berlin, Heidelberg, New York, Springer.

Hilfen...

- Viel Hilfe findet sich im Internet, so z.B. auf:
<http://www.uni-koeln.de/themen/statistik/software/spss/index.html>
- Besonders hilfreich sind die SPSS-PDF-Dokumente des Universitäts-Rechenzentrum Trier. Suchworte „SPSS Universität Trier“ in google bringt gute Treffer.

Popup-Hilfe



Ein Klick mit der rechten Maustaste auf ein Dialogelement aktiviert die Popup-Hilfe für dieses Dialogelement.

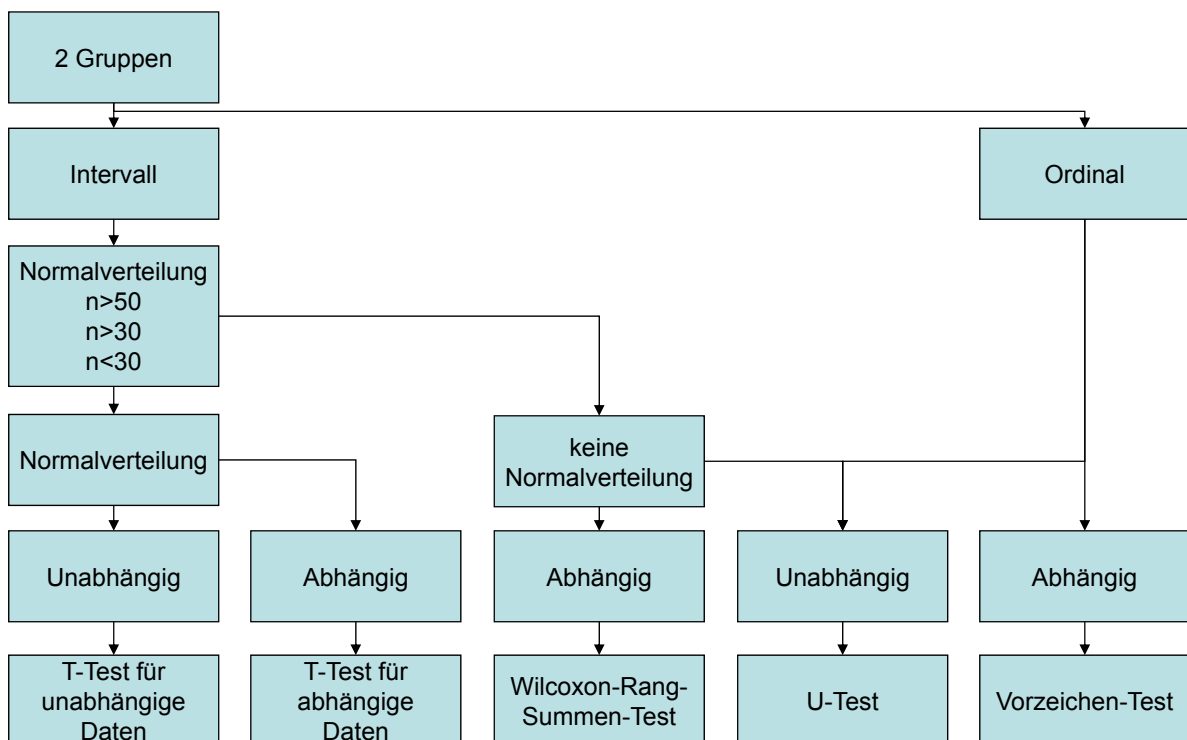
Test-Finder

<http://www.complexity-research.com/TestFinder/>

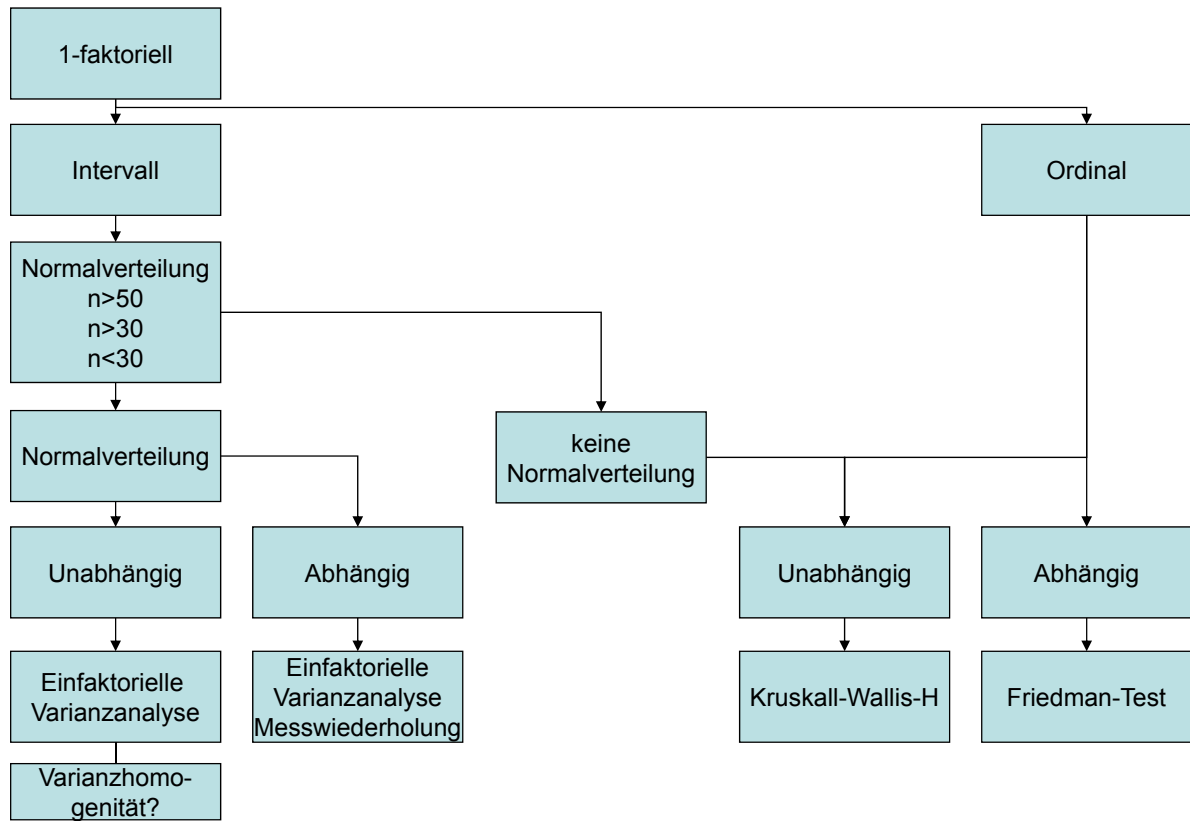
Orientierung im Test-Chaos

Priv.-Doz. Dr. Dr. Guido Strunk

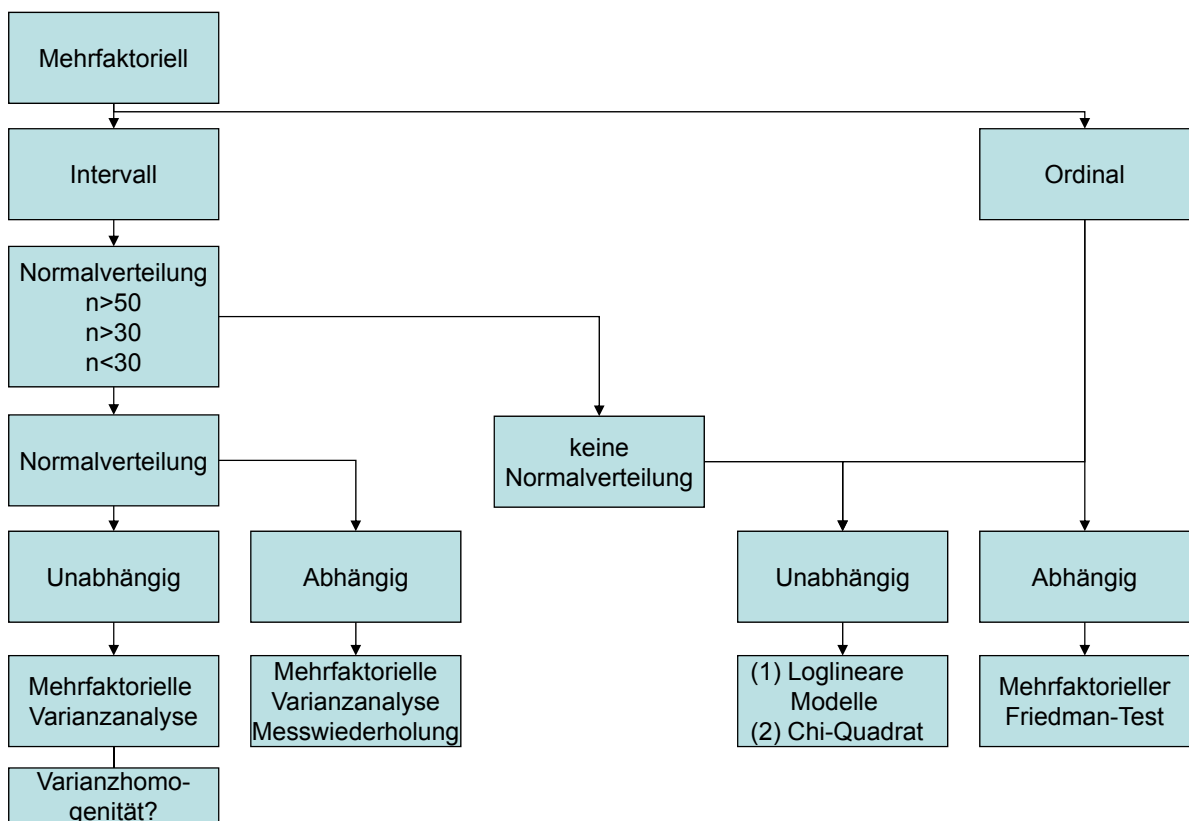
1 Faktor / 2 Gruppen / zentrale Tendenz



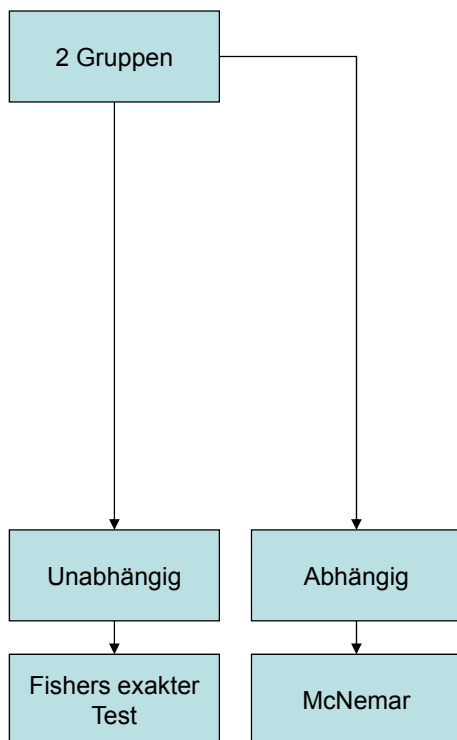
1 Faktor / > 2 Gruppen / zentrale Tendenz



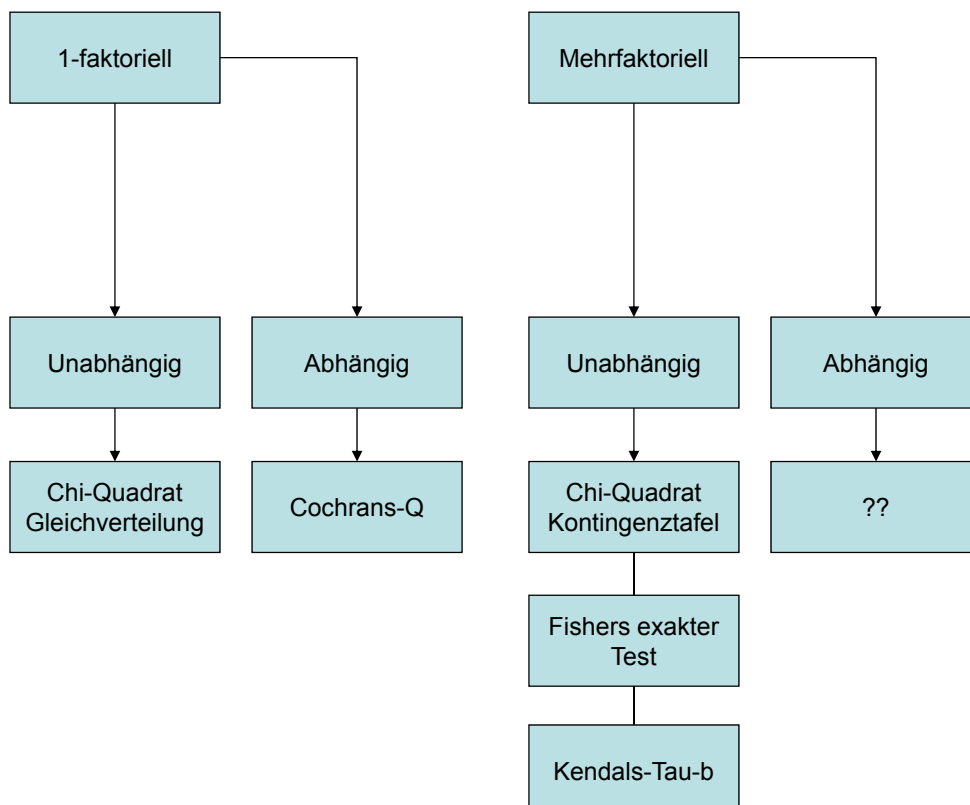
>1 Faktor / > 2 Gruppen / zentrale Tendenz



2 Gruppen / Häufigkeiten



>2 Gruppen / Häufigkeiten



Literatur

- Bortz, J. (1999) *Statistik für Sozialwissenschaften (5. Auflage)*. Berlin, Heidelberg, New York: Springer
- Bortz, J. & Döring, N. (2002) *Forschungsmethoden und Evaluation*. Berlin, Heidelberg: Springer
- Bühner, M. (2004) *Einführung in die Test- und Fragebogenkonstruktion*. München u. a.: Pearson Studium
- Churchill, G. A. & Iacobucci, D. (2002) *Marketing Research: Methodological Foundations*. Orlando, FL: Harcourt College Publishers
- Friedrichs, J. (1990) *Methoden empirischer Sozialforschung*. Opladen: Westdeutscher Verlag
- Langewitz, W., Keller, A. & Denz, M. (1995) Patientenzufriedenheits-Fragebogen (PZF): Ein taugliches Mittel zur Qualitätskontrolle der Arzt-Patient-Beziehung? *Zeitschrift für Psychotherapie Psychosomatik und medizinische Psychologie*, 45, 351-357
- Lienert, G., A. & Raatz, U. (1994) *Testaufbau und Testanalyse*. Weinheim: Beltz
- Mayring, P. (2003) *Qualitative Inhaltsanalyse*. Weinheim und Basel: Beltz Verlag
- Weick, K. E. (1985) *Der Prozeß des Organisierens*. Frankfurt am Main: Suhrkamp