

Kopiervorlagen zur Vorlesung

Forschungsmethodik: Grundbegriffe der Statistik

"Die Statistik ist eine große Lüge, die aus lauter kleinen Wahrheiten besteht." Lionel Strachey

"Mit Statistik kann man alles beweisen - auch das Gegenteil davon." James Callaghan

"Leider gehören Lüge und Statistik daher für viele zusammen wie Dick und Doof oder Pat und Patachon. Seit Benjamin Disraelis notorischem Verdikt [=Urteil] 'There are three kinds of lies: lies, damned [verfluchte] lies, and statistics' hageln die einschlägigen Bonmots auf die armen Statistiker nur so herab.

Diese Verunglimpfungen sind zwar billig, aber wahr. Billig, weil jedes Werkzeug sich sowohl ge- als auch mißbrauchen läßt. Hier steht die Statistik sicher nicht allein. Wahr, weil wir alle nur zu gern die Welt so sehen, wie wir sie gerne hätten, statt wie sie wirklich ist. Wir benutzen nur zu oft, wie der amerikanische Literat Andrew Lang einmal formulierte, 'die Statistik wie ein Betrunkener einen Laternenpfahl: vor allem zur Stütze unseres Standpunktes und weniger zum Beleuchten eines Sachverhaltes'." (KRÄMER 1991, S. 8)

Zur Einstimmung - Auszüge aus:

Prof. Dr. **Bernhard RÜGER**:

Was muß ein(e) PsychotherapeutIn über statistische Methoden in der Therapieforschung wissen?
Münsterschwarzach: Vier-Türme Verlag. (Audi-Torium, 261)

Die Frage ist schwer, vor allen Dingen weil ich erst einmal Hemmschwellen überwinden will oder muß, die ich Ihnen und vielleicht auch Sie mir gegenüber haben. Und ich habe mir gedacht, dass ich die Hemmschwellen am besten dadurch überwinde, dass ich einige Zitate an den Anfang setze, die belegen, wie stark das Lager der Psychoanalytiker und Psychotherapeuten vielleicht auch überhaupt gespalten ist, gegenüber quantitativ empirischer Psychotherapieforschung.

Das erste Zitat stammt von James HILLMAN, 1995, Psychologie Heute, Januar, S. 49:

„Mathematisch statistische Effizienzerhebungen gehören zum Engineering, nicht zur Psychotherapie, nicht zum menschlichen Wesen und nicht zur Seele, und wenn Sie in diese Richtung mit der Psychotherapie in Deutschland gehen..., dann werden Sie wieder zum Faschismus kommen.“

Das zweite Zitat stammt von Wolfgang MERTENS in seiner Entgegnung auf GRAWE: „Psychoanalyse auf dem Prüfstand“, schreibt er auf Seite 82:

„Ich denke, es ist deutlich geworden, dass diese Methodologie eine gänzlich andere Vorgehensweise als die experimentell statistische Forschung darstellt. Auf der einen Seite die strikte Trennung zwischen dem Erkenntnisobjekt und dem Erkenntnisgegenstand in der Absicht, physikalistische Forschungsmethoden in der Psychologie einzusetzen und jegliche Subjektivität des Forschers zu minimieren und auf der anderen Seite die Einfühlung und Identifizierung in den zu beforschenden Menschen, was nur im humanwissenschaftlichen Bereich überhaupt möglich ist, mit dem Ziel einer größtmöglichen Subjektivität.“

Als drittes möchte ich Ulrich RÜGER und Wolfgang SENF zitieren; es handelt sich bei dem Erstgenannten um meinen Bruder, in der Zeitschrift für psychosomatische Medizin und Psychoanalyse, 1994 :

„Die Reserve vieler Psychoanalytiker gegen eine quantifizierende Erfassung von Psychotherapieergebnissen hat eine fatale Folge gehabt. Über lange Zeit haben sich Psychoanalytiker bei der Operationalisierung von Outcome-Variablen kaum beteiligt und waren dann in ihrem Legitimationsbedürfnis gezwungen, für die Erfassung psychoanalytischer Behandlungen vorgefundene, fremdbestimmte, oft außerordentliche reduktionistische Erfassungsinstrumente anzuwenden, was zwangsläufig zur Enttäuschung und Verfestigung des ursprünglichen Vorurteils führen mußte.“

...

In dieses Spannungsfeld zwischen einer doch recht deutlichen Verneinung und Ablehnung empirisch-wissenschaftlicher Methoden und in einer Begrüßung unter dem Vorbehalt, dass sie sach-adäquat angewandt werden, fehlt nun leider noch ein dritter Komplex von Problemen, die auch Ursachen sind für den komplizierten Umgang und das aversive Verhältnis, was zwischen praktischen Psychotherapeuten und empirischen Forschern mathematisch-statistischer Art besteht. Ein Mißverhältnis oder auch eine Spannung, wie sie eigentümlicherweise ja bei praktischen Ärzten nicht besteht. Die praktischen Ärzte verlassen sich ja auf klinisch erprobte Studien, auf die Ergebnisse, die dort geliefert werden.

Ich hab mir einmal als absoluter Laie überlegt, was da die Gründe sein können.

Erster Grund vielleicht: Analytiker verabreichen keine Medikamente, auf deren klinische Erprobung sie sich mehr oder weniger verlassen müssen. Sie sind vielleicht eher selbst das Medikament, oder zumindest die Therapie ist das „Medikament“, so daß sie eine viel stärkere subjektive Einfühlung in die Wirkungsweise des Medikamentes haben und sich in dieser Einfühlung selbst ein Bild immer wieder machen von der Wirkungsweise.

Der zweite Grund: das statistische Instrumentarium der Psychotherapieforschung wurde historisch und bis auf den heutigen Tag fast noch ausschließlich von Psychologen entwickelt. Die psychologische Statistik hat eine lange Tradition und befindet sich auch außerhalb und kontrovers zu der Statistik, die wir, ich, vertreten. Die Statistik der Psychologie oder die psychologische Statistik ist nicht die Statistik schlechthin, selbst wenn Sie in ihrem Studium diese Statistik mehr oder weniger haben lernen müssen, wollen, sollen – je nachdem.

Und der dritte, wirklich wohl gravierendste Grund ist der, dass die statistischen Methoden in der Psychotherapieforschung nicht nur fahrlässig, sondern grundsätzlich fragwürdig angewandt werden; und dass die Ergebnisse, die dabei erzielt werden, für praktisch arbeitende Psychotherapeuten und Psychoanalytiker undurchschaubar sind, weil es sich um Methoden handelt, zu denen sie keinen Zugang haben, weil aber die Ergebnisse selbst, d.h. die Zahlen, die da stehen, scheinbar ja eine so klare Sprache sprechen.

Was macht man dann mit einem Ergebnis?

Diese Fahrlässigkeit in der Anwendung statistischer Methoden ist sicherlich, meine ich, die Hauptursache für ein Mißverhältnis zwischen beiden Lagern.

Ganz nebenbei noch eine weitere Ursache; ich habe mir einmal angeschaut die Flut der Veröffentlichungen auf dem Gebiet der Psychotherapieforschung im engeren Sinn. Und ich muß wirklich gestehen, nicht nur Ihnen sondern auch mir wird schlecht dabei, weil ich den Eindruck hab, dass nicht nur auf dem Gebiet sondern überhaupt heute schneller geschrieben wird als gelesen wird, und schneller gelesen wird, als gedacht wird. Es ist unvorstellbar, mit welcher Vielfältigkeit und Raffinement die Veröffentlichungen sich selbst wieder reproduzieren und immer wieder neue Kinder gebären, die eigentlich geklont sind.

Ich habe einmal die deutschsprachige Datenbank aufgerufen, Psyndex heißt die, zum Thema „Erfolgskontrollen“ sind von 1977 bis 1996 2440 Veröffentlichungen erschienen. Das ist noch gar nicht einmal so erschreckend. Zum Thema „Evaluationsforschung“ aber im gleichen Zeitraum schon 9847 Veröffentlichungen. Im englischsprachigen Sprachraum, in der Datenbank Psych-Lit, sind zum Thema „Evaluation“, in den letzten 6 Jahren 13 123 Veröffentlichungen erschienen. Davon kann man höchstens noch ein Promill lesen. Höchstens! Auf die Zahl bin ich nicht gekommen, glaub ich, wenn ich so richtig überlege; 130 habe ich nicht geschafft. Ich habe ein halbes Promille gelesen. Es ist unvorstellbar, man kann einfach bei der Flut von Veröffentlichungen – nicht nur auf Ihrem Gebiet, überhaupt auch - nur noch sagen: Ich weigere mich, die zur Kenntnis zu nehmen. Ich arbeite und mach meine Arbeit, und ich weiß, ob die gut ist, manchmal weiß ich auch, daß sie vielleicht nicht so gut ist, - ich beurteile das selber - aber es ist unmöglich, weil vor allen Dingen auch widersprüchliche Ergebnisse da sind, die ja nicht einfach so zur Kenntnis genommen werden. Und weil auch - bei der Flut von Veröffentlichungen - eigentlich erforderlichenmaßen Metaanalysen, die diese Veröffentlichungen wieder zum Gegenstand der Forschung machen, damit endlich zusammenfassend ein vielleicht klares Urteil herauskommt, die gleichen Fehler und methodischen Mängel aufweisen, wie zum Teil die Veröffentlichungen selber.

Soweit meine Präambel. Ich möchte zum Schluß dieser Präambel aber noch Jürg WILLI zitieren; das aber nur indirekt, aus der Zeitschrift, aus der Schweizerischen Ärztezeitung, genauer gesagt aus einem Aufsatz von Gion Condreau, 1995. Dort schreibt er, er, aber zitiert indirekt Jürg Willi:

„ Jürg WILLI meint, auch eine gute, deskriptiv phänomenologische Untersuchung sollte gewisse methodische Grundsätze der empirischen statistischen Forschung beachten, selbst wenn es sich um eine Einzelfallstudie handelt.“

Und um diese methodischen Grundsätze geht es mir. Die methodischen Grundsätze empirisch-statistischer Forschung entscheiden schließlich über den Wert und die Glaubwürdigkeit eines statistischen Ergebnisses und nicht der Inhalt des Ergebnisses. Das ist **mein** Standpunkt. Der Inhalt ist für Sie zwar aufregend – unter Umständen, aber für mich eigentlich schon nicht mehr. Dieser Rückzug auf Methoden, den ich hier einnehme, ist für einen verantwortungsvollen Statistiker aber notwendig, weil er nicht ergebnis- sondern verfahrensorientiert arbeitet. Und wenn vielleicht dieser Standpunkt sich stärker durchsetzt, könnte auch eine Vertrauensbildung wieder stattfinden, hoffe ich, zur methodisch gut abgesicherten Statistik.

Einleitendes

Was ist Statistik?

lat. status (Stand, Zustand, Lage, Beschaffenheit)

Statistik ist eine mathematische Disziplin - Hilfswissenschaft für Psychologie und Psychotherapie, die zahlenmäßige Untersuchungen von Massenerscheinungen vornimmt.

Zusammenstellung numerischer Information

Statistik ist auf Quantifizierung und Verallgemeinerung ausgerichtet.

"Die Statistik dient - dies ist eine banale Aussage - der Analyse von Daten. Wer Daten sagt, der muß auch Empirie sagen: Die Grundlage aller Statistik ist die Beobachtung empirischer Sachverhalte und die Erzeugung von Daten aufgrund solcher Beobachtungen. Die Statistik spielt somit eine Rolle in denjenigen Wissenschaften, die zumindest in Teilbereichen als empirische konzipiert sind, d.h. in denen wissenschaftliche Beobachtung der "Realität" als wesentliche Methode zum Erkenntnisgewinn eingesetzt wird. Nun wurden in den letzten beiden Sätzen stichwortartig bereits mehrere Themen angesprochen, die in der Methodenlehre einer empirischen Wissenschaft sozusagen noch "vor" der Statistik stehen. Dazu gehört zunächst die wissenschaftstheoretische Diskussion des Konzepts der empirischen Wissenschaft und schließlich die Fragen, was unter wissenschaftlicher Beobachtung zu verstehen ist bzw. wie solche Beobachtungen angestellt werden können, wie dann aus solchen Beobachtungen Daten erzeugt werden können bzw. welche Eigenschaften diese Daten haben sollten, damit sie für die wissenschaftliche Forschung überhaupt brauchbar sind." (SCHILLING 1998, S. 8/9)

Damit Untersuchungsergebnisse statistisch bearbeitet werden können, müssen sie zwei Bedingungen erfüllen:

1. Untersuchungsergebnisse beziehen sich auf Mengen

Beispiele:

- 100 Würfe mit einem Würfel
- die Patienten der Univ.Kliniken Innsbruck im Jahre 1994
- die Bewohner des Ötztales
- die an Depressionen leidenden Bewohner des Ötztales
- die zur selbständigen Ausübung der Psychotherapie in Österreich Berechtigten

2. die Untersuchungsergebnisse müssen die Unterscheidung quantifizierbarer Eigenschaften zulassen

Beispiele:

- das Auftreten von 6 Augen
- die Unterscheidung nach männlich / weiblich;
- das Ausmaß der Einnahme von Beruhigungsmitteln
- das Ausmaß der freiberuflichen psychotherapeutischen Tätigkeit pro Woche

Ziel der Statistik (vgl Erdfelder 1994, S 49):

- Entdeckungszusammenhang
- Begründungszusammenhang

"Es können aus der Beobachtung erzeugte Daten mit dem Ziel analysiert werden, überhaupt erst zu theoretischen Aussagen zu kommen, mögliche gesetzesmäßige Zusammenhänge aufzuspüren u. dgl. (Entdeckungszusammenhang), oder um zu überprüfen, ob theoretische Aussagen überhaupt gültig sind (Begründungszusammenhang)." (SCHILLING 1998, S.11)

"Salopp formuliert bedeutet das: Man schaut nach, ob in der Realität die Dinge sich auch so verhalten, wie eine Theorie behauptet. Diese Vorgehensweise ist jedoch nicht so einfach, wie der letzte Satz klingt, wenn dieses Nachschauen von subjektiven Verzerrungen und logischen Irrtümern frei sein soll. Die Statistik ist hierbei ein unverzichtbares Instrument, mit dem eben nicht nur dargestellt werden soll, wie sich die Dinge verhalten, sondern auch überprüft, ob diese Verhältnisse durch eine Theorie erklärt werden können." (SCHILLING 1998, S.11)

Aufgaben der Statistik

Man unterscheidet zwischen *deskriptiver Statistik* und *Inferenzstatistik*.

Die **deskriptive** (beschreibende, chronistische) **Statistik** ermöglicht die übersichtliche Darstellung und Zusammenfassung der Beobachtungsergebnisse.

Erheben, Ordnen, Aufbereiten und Darstellen von Daten sowie das Bestimmen von Kenngrößen (Kennwerten) dieses Datenmaterials und deren Interpretation.

Die beschreibende Statistik "ist für uns sehr wertvoll, weil der menschliche Verstand große Mengen von Daten nur zu verarbeiten vermag, wenn diese geordnet, gruppiert, extrahiert und verdichtet werden. Die beschreibende Statistik befähigt uns, das Allgemeine, Typische der statistischen Menge zu erkennen." (Clauß/Ebner 1977, 19)

Aber Vorsicht vor scheinbarer Präzision! (Steuererklärung; Statistiken in den Medien)

Die **Inferenzstatistik** (schließende, schlussfolgernde, operative, analytische, Prüfstatistik) hat die Aufgabe, Verfahren für den Schluss von den Ergebnissen einer Stichprobe auf die Grundgesamtheit bereitzustellen.

Inferenz (lat. infero - hineinragen, eintragen, in Rechnung stellen)

Inferenz = ein auf andere Urteile gegründetes Urteil, logischer Schluss (DORSCH 121994)

Dieser Schluss von der Stichprobe auf die Grundgesamtheit ist ein Schluss vom Besonderen auf das Allgemeine. (= induktive Schlüsse; vs. deduktive Schlüsse = vom Allgemeinen auf das Besondere)

Die Menge aller gleichartigen Fälle, auf die sich eine Aussage oder Hypothese bezieht, heißt Population oder Grundgesamtheit. Die Gruppe, an der die empirische Untersuchung durchgeführt wird, stellt lediglich eine Teilmenge oder Population dar. Diese Teilmenge wird Stichprobe genannt. (HOPPE/LIEPMANN 1974, 11)

Die Anzahl der Werte in einer Stichprobe nennt man Stichprobenumfang und gebraucht dafür das Symbol:

$$N =$$

Bei der Auswertung empirischer Untersuchungen ist der statistische Schluss notwendig, weil lediglich Ergebnisse für eine Stichprobe vorliegen, jedoch Aussagen über die Population gemacht werden sollen.

"Das Stichprobenergebnis informiert über die zur Stichprobe gehörenden Elemente und gestattet mehr oder minder zuverlässige Aussagen über die Grundgesamtheit, der die Stichprobe entnommen wurde. Die Gültigkeit einer solchen Verallgemeinerung hängt entscheidend davon ab, dass bei der Auswahl der Stichprobe keine Fehler begangen wurden." (Clauß/Ebner 1977, 18)

"Die Stichprobe muss repräsentativ für die Grundgesamtheit sein, welcher sie entstammt."
(ebd.)

"Mit der Repräsentativität einer Stichprobe bezeichnet man ... den Sachverhalt, dass die Stichprobe in ihrer Zusammensetzung der Zusammensetzung der Grundgesamtheit entspricht." (Hoppe/Liepman 1974, 16)

Zur Bewältigung dieser beiden Aufgaben (beschreibende - beurteilende Statistik) hat die Statistik ein komplexes Repertoire an Methoden und Verfahren entwickelt:

- statistische Darstellungsverfahren (Kreuztabellen, Grafiken)
- statistische Kennwerte (Mittelwert, Streuung)
- statistische Prüfverfahren (Faktorenanalyse, t-Test)

Zusammenfassend: Definition von POSPESCHILL (1996, 2):

"Der Begriff Statistik umfasst alle quantitativen Analysetechniken, mit denen empirische Daten (im Sinne von "Messergebnissen")

- nach ihrer (kategorialen) Verteilung zusammenfassend beschrieben,
- in ihrer Lage und Streuung charakterisiert und
- übersichtlich z.B. in Form von Graphiken dargestellt werden können.

Dies ist Aufgabe der Deskriptiven Statistik.

Weiters bietet die Statistik Verfahren, mit deren Hilfe auf Grund empirischer Daten

- Hypothesen aufgestellt und
- Hypothesen (statistisch) geprüft werden können.

Dies ist das Aufgabengebiet der Inferenzstatistik.

Daten

Was sind Daten, was heißt "messen"?

Daten (urspr. = Gegebenes, Singular = Datum) sind ein Ausschnitt der Wirklichkeit, der als eine Grundlage empirischer wissenschaftlicher Erkenntnisse benötigt wird. Daten sind aber - trotz des ursprünglichen Wortsinnes - in der Wissenschaft nichts einfach Gegebenes. Der Wissenschaftler stellt vielmehr Daten her, er findet sie nicht vor! Zu Daten wird etwas nur durch den Zugriff des Wissenschaftlers.

Beispiele

Geschlecht des Patienten

Alter des Patienten

Familienstand (ledig, verheiratet, getrennt lebend, geschieden)

höchster Schulabschluss (Hauptschule, Mittlere Reife, Abitur/Fachhochschulreife)

höchster Berufsabschluss (ohne Abschluss, in Ausbildung, abgeschl. Lehre,

Meister/Fachschule, FH/Universität, sonst. Abschluss)

Beruf des Patienten

Haushaltsgröße

Wohnort (Großstadt, Stadt, Kleinstadt, regionale Zentren, Land)

Störung des Patienten (Diagnose: Psychose [ICD 9 295], Neurose [300],

Persönlichkeitsstörung [301], körperliche Funktionsstörung psychischen Ursprungs [306], EB-,

Schlafstörungen, Schmerzen[307], Anpassungsstörung [309])

Schwere der Störung (krankheitswertigkeit)

Dauer der Störung

psychotherapeutische Vorbehandlung (ja/nein, ambulant/stationär)

Psychopharmaka- / Medikamentenkonsum

Geschlecht des Therapeuten

Alter des Therapeuten

Erfahrung des Therapeuten (Ausbildungskandidat, Neuling, erfahrener Therapeut,

Lehrtherapeut)

Psychotherapie"schule"

konkrete Intervention (Deutung; Heißer Stuhl, Rollenspiel mit Perspektivenwechsel, kreative Medien, Awareness-Übungen)

Person des Therapeuten (Einfühlsamkeit, Wärme, Geduld, Vertrauen in Kompetenz d.

Patienten, Ermutigung, Selbstreflexionsförderung- und Selbstwertförderung)

Beziehungsqualität Patient/Therapeut

Zugang zur Therapie

Bezahlung der Therapie

Behandlungsdauer (in Sitzungen, Tagen, Wochen)

Behandlungsetting (ambulant/stationär; Einzel-, Paar- / Gruppentherapie)

Behandlungsende (regulär, vorzeitig durch Patient, einvernehmlich, sonstiges)

Behandlungserfolg (Patientenzufriedenheit, Dauer der Arbeitsunfähigkeit, Erreichen vordefinierter Ziele, Kommentare des sozialen Umfelds, Globalschätzungen, Rückfälle)

Messen

Was heißt "Messen"?

"Messen besteht im Zuordnen von Zahlen zu Objekten, so daß bestimmte Relationen zwischen den Zahlen analoge Relationen zwischen den Objekten reflektieren." (DIEHL/KOHR ¹¹1994, S.7)

Zitat KRIZ (1996, 116-118): [empirisch = aus der Erfahrung / Beobachtung entnommen]

"Die zu reduzierende Information als Ausgangspunkt dieser Forschungsphase - üblicherweise "Datenauswertung" genannt - bezieht sich auf empirische Elemente und empirische Beziehungen zwischen diesen. Man nennt eine solche Struktur ein "empirisches relationales System" oder "empirisches Relativ". Die empirischen Elemente können z.B. Menschen, Gruppen von Menschen, Institutionen, Städte, Kunstwerke, Kreuze auf Fragebögen oder auf Tonträgern gespeicherte Gefühlsäußerungen von Menschen sein. Die empirischen Relationen können z.B. beobachtbare Aussagen, wie "ist später geboren als", "sitzt zwischen", "sind sich ähnlicher als" usw. sein.

Die Notwendigkeit für die Reduktion der Information ... läßt sich leicht verdeutlichen, wenn man sich eine sehr kleine Fragebogenuntersuchung vorstellt, in der 100 Personen jeweils 5 Fragen zur Reform des Abtreibungsparagraphen vorgelegt wurden und zusätzlich nach Alter, Geschlecht, Einkommen, Familienstand und Religionszugehörigkeit gefragt wurde. Das empirische Relativ besteht dann aus 1000 Angaben auf den Fragebögen und bestimmten Beziehungen zwischen diesen Angaben, z.B. hinsichtlich des Alters, des Einkommens, einer bestimmten Fragebeantwortung etc. Stellen wir uns weiterhin vor, man würde den Untersuchungsleiter bitten, mitzuteilen, was bei dieser Untersuchung herausgekommen ist. Wenn nun der Untersuchungsleiter daraufhin die 100 Fragebogen nähme und beim ersten beginnend einfach sämtliche Antworten vorlesen würde, so würde man ihn wahrscheinlich spätestens beim 10. Fragebogen mit Worten unterbrechen wie: "Was soll ich mit dieser Detailinformation, können Sie nicht das Wesentliche zusammenfassen?" Diese Forderung würde man aber nicht nur aus Zeitmangel stellen, sondern einfach deshalb, weil man mit der detaillierten Information kaum etwas anzufangen wüsste, d.h. man könnte wahrscheinlich bereits nach 10 Fragebögen nicht einmal so banale Fragen beantworten wie, "ob die weiblichen Befragten älter gewesen sind als die männlichen", oder "ob die Katholiken ein höheres Einkommen hatten als die Protestanten" usw. Das bedeutet, bei unsystematischer Präsentation der Information eines empirischen Relativs ist unser menschliches Gehirn ohne Hilfsmittel nicht in der Lage, die Information adäquat, d.h. im Hinblick auf die Beantwortung bestimmter Fragen, zu reduzieren. Dabei muß bedacht werden, daß dieses Beispiel mit nur 100 Befragten und jeweils 10 Fragen sogar ein untypisch kleines empirisches Relativ darstellt. Das Beispiel sollte zeigen, daß Wissenschaft offensichtlich nicht darin besteht, die gesammelten Einzelfakten des empirischen Relativs möglichst vollständig und unverfälscht in den Kommunikationsprozeß einzubringen, sondern insbesondere darin, die Information im Hinblick auf bestimmte Fragen adäquat zu reduzieren."

Es geht darum, Einzelerfahrungen, einzelnen Fakten zunächst zu erfassen und dann die Fülle von solchen Fakten zu strukturieren und zu reduzieren, Gesetzmäßigkeiten aufzudecken. "Es sollte auch deutlich werden, daß unser menschliches Gehirn in den meisten Fällen ohne Hilfsmittel zu einer solchen Informationsreduktion nicht in der Lage ist, d.h. daß der direkte Weg, von einem komplexen empirischen Relativ zu kommunizierbaren Ergebnissen zu gelangen, in der Regel auf Grund der beschränkten menschlichen Intelligenz versperrt ist. [wohl besser auf Grund der beschränkten Aufnahmekapazität, C.B.]

Eine Möglichkeit, diese Intelligenzbarriere zu umgehen - und dabei ist sowohl das Wort "eine" wie auch "Möglichkeit" zu betonen - ist die Zuhilfenahme mathematisch-statistischer Modelle.

empirisches Relativ = der Erfahrung zugängliche Elemente und ihre Beziehung zueinander
homomorph = spezielle Abbildung einer Struktur in oder auf eine andere = ungefähr "grob ähnlich"

= ein System A ist einem System B homomorph, wenn von jedem Bestandteil von A und von jeder Relation zwischen Bestandteilen von A auf einen Bestandteil von B bzw. auf eine Relation von Bestandteilen von B geschlossen werden kann, jedoch nicht umgekehrt (DORSCH ¹²1994, 324)

mit anderen Worten:

- es müssen Ausprägungen da sein (unterschiedliche Werte)
- diese müssen in "Zahlen" (Werten) abgebildet werden können
- es müssen Beziehungen / Zusammenhänge / Relationen zwischen diesen Werten feststellbar sein (gleich/ungleich, größer/kleiner doppelt so groß)

"In der Tat bedienen sich diese Wissenschaften [gemeint sind Pädagogik, Psychologie u. Soziologie, C.B.] in zunehmenden Umfang quantifizierender Methoden. Das ist prinzipiell zu begrüßen. Aber man sollte die Augen nicht vor einer Gefahr verschließen, die damit verbunden ist: Es gibt gegenwärtig Tendenzen zur Überbewertung solcher Methoden und zu ihrem nicht gegenstandslosen Einsatz. Mancher Gesellschaftswissenschaftler läßt sich - beeindruckend von der Exaktheit statistischer Verfahren - dazu verführen, einen Sachverhalt empirisch anzugehen, nicht weil er praktisch oder für die Theoriebildung bedeutsam ist, sondern weil er sich leicht quantifizieren und statistisch elegant bearbeiten läßt. Das Ergebnis einer solchen Untersuchung kann formal unanfechtbar, für den Leser zunächst imponierend, dennoch aber belanglos oder gar irreführend sein. ((Entscheidend für den Wert einer empirischen Untersuchung ist stets, ob sie unser sicheres Wissen von einem gesellschaftlich wichtigen Wirkungsbereich in verlässlicher Weise erweitert, dadurch einen Beitrag liefert zur Theorie dieses Bereiches und so letzten Endes eine verbindliche Anleitung zum richtigen Handeln gibt.)) Auch statistische Methoden sind - wie empirische Methoden überhaupt - beurteilbar nur im Zusammenhang mit der theoretischen Relevanz der Problemstellungen, zu deren Klärung sie verwendet werden. Quantifizierung um jeden Preis ist von Übel. Empirismus und Praktizismus - das heißt Vernachlässigung der Theorie, verbunden mit Überschätzung formaler Techniken der Datengewinnung und ihrer Verarbeitung - sind Gefahren, vor denen sich jeder Wissenschaftler hüten muß, wenn er quantifizierende Verfahren anwendet." (Clauß/Ebner 1977, 23/24

- Faszination, die von klaren Ergebnissen ausgeht
- Gefahr, das zu untersuchen was leicht messbar ist - aber nicht das, was aufgrund theoretischer Überlegungen als bedeutsam angesehen wird
- u.U. belanglose Daten – Datenmüll, Datenfriedhöfe

Skalen

Zum Messen braucht man bestimmte Meßinstrumente und Skalen (vgl. Küchenwaage, Zollstock mit Zentimetereinteilung, Stoppuhr)

Skalen - Niveaus:

- | | |
|-------------------|------------------------------|
| - Nominalskala | (Un-) Gleichheit |
| - Ordinalskala | Größer-kleiner-Relation |
| - Intervallskala | Gleichheit von Differenzen |
| - Verhältnisskala | Gleichheit von Verhältnissen |

Nominalskala

"Die Zuordnung zu Ausprägungen einer nominalskalierten Variablen lassen sich durch beliebige Ziffern oder Zeichen vornehmen. Relevant ist nur die Bestimmung, ob Merkmale gleich oder ungleich sind (z.B. männlich/weiblich)."

Zahl als Name, nomen
Grundsatz der Identität; dasselbe Ereignis wird derselben Zahl zugeordnet
männlich / weiblich
Telefonnummer
Postleitzahlen
Staatsbürgerschaft
Hauptwohnsitz
Quellenberuf

Krankheitsklassifikationen

Ordinalskala

"Die Stufen einer ordinalskalierten Variablen lassen sich entlang einer Dimension anordnen, wobei nur die Relation größer oder kleiner - unabhängig vom Ausmaß - bzw. gleich relevant ist (z.B. Testrohwerte)."

Ordinal- oder Rangskala
größer - kleiner; nach - vor
Bspl. Windstärken
Olympiade 1., 2., 3.
Beliebtheitsrangfolge
Berufsprestige

Intervallskala

"Bei der Intervallskala lassen sich die Stufen der Variablen entlang einer Dimension mit gleichen Abständen zwischen den Stufen anordnen, allerdings ohne absoluten Nullpunkt (z.B. Standardtestscores).

gleiche Differenzen
Kalenderzeiten (Datum)
Temperaturskalen (Abstand zwischen -5 und +5 ist gleich Abstand zwischen +20 und +30;
aber nicht: +20 ist doppelt so warm wie +10 Grad)

Verhältnisskala

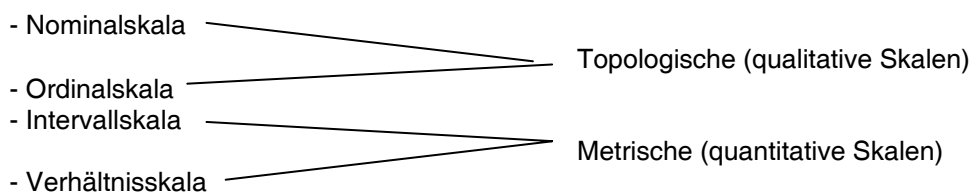
Bei der Verhältnisskala kann die Einheit frei bestimmt werden, der (natürliche) Nullpunkt ist fixiert (z.B. Körpergröße, Gewicht, elektrische Spannung).

Gleichheit der Verhältnisse
Länge, Gewicht, Zeit, Preise, Behandlungsdauer

Über verschiedene Skalen:

"Eine vergleichende Betrachtung der vier Skalen zeigt, daß mit zunehmender Ordnungsziffer der Skalen die Abbildung des empirischen Relativs in das numerische Relativ genauer wird, bzw. daß zunehmend mehr Eigenschaften des numerischen Relativs auf das empirische Relativ übertragbar sind. Dies wird deutlich, wenn wir uns vor Augen führen, daß Ordinalskalen, die größer-kleiner Relationen richtig abbilden, auch die Gleichheits-Ungleichheitsbedingung der Nominalskalen erfüllen, bzw. daß "Gleichheit der Differenzen" (Intervallskala) sowohl größer-kleiner, als auch gleich-ungleich-Relationen beinhaltet und daß "Gleichheit der Verhältnisse" (Verhältnisskala) alle drei genannten Bedingungen impliziert. So gesehen, stellt die Klassifikation der vier Skalen eine Ordinalskala dar, wobei mit zunehmender Rangnummer der Skalen mehr Informationen des empirischen Relativs im numerischen Relativ abgebildet werden."

BORTZ (1977, S. 29)



Zur Erinnerung:

Es geht im Grunde um 2 Aspekte:

1. nur das anwenden, was methodisch zulässig (mathematisch-statistischer Aspekt)

Frage: was ist erlaubt ?

s. Zitat RÜGER (1996, 136):

"Die erste [Leitlinie] betrifft den wissenschaftlichen Standpunkt eines Statistikers, demzufolge die Beachtung der methodischen Grundsätze statistischer Verfahren über den Wert und die Glaubwürdigkeit eines statistischen Ergebnisses entscheidet und nicht der Inhalt des Ergebnisses. Die Ausrichtung (der "Rückzug") auf die Methoden ist für einen verantwortungsvoll arbeitenden Statistiker absolut notwendig: Er arbeitet nicht ergebnis- sondern verfahrenorientiert. Um es verschärft auszudrücken: Wenn man den statistischen Nachweis eines Ergebnisses nicht methodengerecht durchführt, ist das Ergebnis selbst (statistisch!) wertlos, unabhängig davon, wie plausibel es ansonsten sein mag. Nur unter der Wahrung dieses Standpunktes kann empirisch-statistischen Untersuchungen Vertrauen entgegengebracht werden."

"Tendenziell gilt: Je schärfer die Aussagen sind, die das statistische Verfahren ermöglicht, desto einschneidender sind die Voraussetzungen ... (an die Stichproben und Effektgröße)

2. keine Untersuchung ohne inhaltliche Vorentscheidungen ! (wissenschaftstheoretischer Aspekt)

Frage: was ist geeignet ?

d.h. was ist erlaubt und dem Studienziel gerecht werdend ?

„Die Beurteilung der Skalenqualität hängt direkt von der Genauigkeit unserer Kenntnisse und theoretischen Formulierungen über das zu messende Merkmal und dessen Äußerung ab: Können wir z.B. sagen, daß eine Person deshalb ängstlicher sei, weil sie für mehr Situationen Angst eingesteht als eine andere Person, oder müssen wir die Angststärke nach der Gewichtigkeit des Anlasses zur Ängstlichkeit einstufen? Wissen wir hierüber nichts, so sind nur kategoriale Vergleiche möglich. Eine hinreichende Vorkenntnis mag ordinale Einstufungen zulassen. Erst eine gründliche Überprüfung eines Meßinstruments für Angst rechtfertigt es etwa, eine lineare Beziehung des Angstmaßes zu bestimmten Vermeidungsreaktionen anzunehmen. Dieses Beispiel zeigt, daß die Qualität einer Messung von theoretischen Entscheidungen abhängt. Nicht nur in die Aussage über den Untersuchungsgegenstand selbst, sondern auch in dessen Erfassung gehen Hypothesen ein. Sozialwissenschaftliche Messung ist nie ein rein technisches, sondern stets zugleich ein theoriegeleitetes Unterfangen."

(BORTZ 1977, 31)

- ich will Angst messen

- wie operationalisiere ich Angst?

- Angst vorhanden - nicht vorhanden (Nominalskala)

- mehr Angst oder weniger Angst (Ordinalskala)

- mehr Angst, wenn jemand selber für mehr Situationen Angst eingesteht [d.h. mehr bedeutet „in mehr Situationen“ = Anzahl = ich gehe nur von der Quantität aus]

- oder mehr Angst, wenn der Anlass gewichtiger

[d.h. mehr Angst bedeutet - qualitativer Unterschied; es gibt Situationen, die mehr Angst erregen als andere / Lebensangst - Schulangst - Prüfungangst - Angst vor Spinnen ...

Entscheidungen hängen davon ab, welche Theorie ich im Hinterkopf habe und welche Hypothese ich prüfen will.

Statistische Kennwerte

Häufigkeiten

- absolute Häufigkeit
- kumulierte Häufigkeit
- Prozentwertverteilung
- kumulierte Prozentwertverteilung

- absolute Häufigkeit

die Anzahl der Fälle, die für eine Variable einen bestimmten Wert (oder Bereich von Werten) besitzen
Strichliste - zusammenzählen

- kumulierte Häufigkeit

die absoluten Häufigkeiten sukzessiv zusammenzählen
um zu kontrollieren, ob alle Meßwerte berücksichtigt wurden
die letzte Kategorie muß den Wert n =Stichprobenumfang erhalten

- Prozentwertverteilung

Formel:

$$f_{1\%} = \frac{f_1}{n} \times 100$$

wobei

- $f_{1\%}$ = zu errechnender Prozentwert für die Kategorie 1
- f_1 = Häufigkeit (Frequenz) in der Kategorie 1
- n = Stichprobenumfang

um die Merkmalsverteilung überschaubarer zu machen oder um unterschiedlich große Stichproben miteinander vergleichen zu können

zu unterscheiden (vgl. POSPESCHILL 1996, 11-12):

- Prozente ohne Ausschluß von Missings
- gültige Prozente

Prozente ohne Ausschluß von Missings:

Die nicht bereinigten Prozente werden berechnet, indem die Häufigkeit eines Wertes durch die gesamte Anzahl der Fälle in der Stichprobe dividiert wird (incl. fehlender Werte). Diese weichen von den gültigen Prozenten (valid percent) ab, wenn fehlende Werte existieren, die bei der Berechnung der gültigen Prozente nicht berücksichtigt werden.

Gültige Prozente:

Der prozentuale Anteil an Fällen, die einen bestimmten Wert besitzen, wenn nur Fälle mit nicht fehlenden Werten einbezogen werden. Er wird berechnet durch die Division der Zahl an gültigen Fällen mit einem bestimmten Wert durch die gesamte Anzahl an gültigen Fällen und mit 100 multipliziert. Sie können sich von den einfachen Prozenten unterscheiden, bei denen auch die fehlenden Werte in die Prozentberechnung einbezogen werden. Sie werden auch korrigierte relative Häufigkeit genannt.

- kumulierte Prozentwertverteilung

die absoluten Prozentwertverteilungen sukzessiv zusammenzählen
wenn richtig gerechnet, muß die kumulierte Prozentwertverteilung in der letzten Kategorie 100% erhalten

Abweichung von 100% kann durch Rundungsfehler hervorgerufen werden (z.B. 99,9% oder 101%)

Bei einer Häufigkeitsverteilung, die nur in Prozentwerten angegeben wird, ist unbedingt darauf zu achten, daß der Stichprobenumfang "N = " mitgeteilt wird. Nur so ist gewährleistet, daß die Prozente nicht über Gebühr verallgemeinernd interpretiert werden

Darstellung in Häufigkeitstabellen und Grafiken

Quoten = pro ; Dichte

Bspl. Anzahl der Psychotherapeuten pro Einwohner
(Bedenbecker 1996.)

Psychotherapeuten in Österreich
Verteilung nach Bundesländer (Berufssitz/Dienstort)
Eingetragen in die Therapeutenliste; Stand: Oktober 1995; Gesamtzahl=3633

Bundesland	absolute Anzahl an Therapeuten	Einwohnerzahl	Therapeuten pro 100.000 Einwohner
Burgenland	32	273.600	11,7
Kärnten	136	559.700	24,3
Niederösterreich	270	1.511.600	17,9
Oberösterreich	332	1.383.600	24,0
Salzburg	408	504.300	80,9
Steiermark	329	1.204.000	27,3
Tirol	306	654.800	46,7
Vorarlberg	180	342500	52,6
Wien	1.632	1.595800	102,3
ohne Angabe	8		

Werte der zentralen Tendenz - Mittelwerte

"Mittelwerte sind lageangegebende Kenngrößen (Kennwerte) der deskriptiven Statistik; sie kennzeichnen die zentrale Tendenz einer Häufigkeitsverteilung auf dem Merkmals-(Beobachtungs- oder Maß-)kontinuum, der X-Achse." (BARTEL 1971, 31)

Die wichtigsten Mittelwerte:

- Modus
- Median
- arithmetisches Mittel

Modus (Modalwert, Gipfelwert, Dichtemittel, häufigster Wert)

"Der Modalwert ist der in einer Verteilung am häufigsten vorkommende Meßwert." (Claus/Ebner 1977, 81)

ähnlich "Mode"
größte Häufigkeit in der Verteilung

D=

Bestimmen des Modus:

- unmittelbar durch Ablesen in der Häufigkeitsverteilung

3,4,4,5,5,5,5,6,6,7,8 D = 5

Sind 2 oder mehrere nebeneinanderliegende Maßzahlen gleich oft und am häufigsten vorhanden, so ist ihr arithmetisches Mittel der Modalwert:

3,4,4,5,5,5,6,6,6,7 D = 5,5

Sind zwei nicht nebeneinanderliegende Maßzahlen gleich oft und am häufigsten vorhanden, so handelt es sich um eine bimodale Verteilung; es werden dann beide Maßzahlen als Modalwerte bezeichnet.

3,4,4,5,5,5,5,6,6,7,8,8,8,8,9 D = 5; 8

Voraussetzung für den Modalwert:

- Nominalskalenniveau

Aussagekraft:

- aussageschwächstes Maß der Zentraltendenz
- bringt die Verteilung nicht zum Ausdruck
- wird leicht von Zufallsfehlern beeinflusst
- eignet sich besonders zur Kennzeichnung zwei- oder mehrgipfliger Häufigkeitsverteilungen

Median (Zentralwert, 50. Zentil)

Median = die mittlere Maßzahl in der nach Größe geordneten Rangreihe der Messwerte. Sie halbiert die Reihe derart, dass ebenso viele Maßzahlen über ihr wie unter ihr liegen.

Der Median ist der Punkt auf dem Maßzahlkontinuum, unter dem 50% und über dem 50% aller Messwerte der Verteilung liegen.

Bestimmen des Median:

- durch Ablesen der mittleren Maßzahl (s. prozentuale Häufigkeitsverteilung)
- kann auch zeichnerisch bestimmt werden

- bei ungeradem N ist der Median ein in der Verteilung tatsächlich auftretender Wert
- bei geradem N liegt er zwischen den beiden zentralen Messwerten (arithmetisches Mittel zwischen beiden Messwerten)

Voraussetzung für den Median:

- Ordinalskalenniveau

Aussagekraft:

- mehr Informationen über die Verteilung als der Modalwert
- als "mittlerer Fall" weniger von Zufallsfehlern beeinflusst
- angebracht vor allem bei ordinalskalierten Werten
- angebracht bei offenen Maßzahlklassen (am Anfang oder am Ende der Verteilung)
- angebracht bei extrem geringer Anzahl von Messwerten oder bei sehr schiefen Verteilungen

Arithmetisches Mittel (Durchschnitt, Mittelwert)

Formel:

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

Bestimmen des arithmetischen Mittels:

Quotient aus der Summe aller Maßzahlen durch deren Anzahl

$\bar{x} =$

\bar{x}

\bar{x} ist der Punkt auf dem Merkmalskontinuum, für den die Summe aller Abweichungen der einzelnen Maßzahlen gleich null ist ("verschwindet").

x_i

ist Repräsentativwert, ist Schwerpunkt (im physikalischen Sinne) der Verteilung

z.B. man bedient sich des Durchschnittsgewichtes, um bei einem Aufzug auf Grund der höchstzulässigen Belastung die Maximalzahl der Aufzugsgäste auszurechnen.

Voraussetzungen

- mindestens Intervallskala

Aussagekraft:

mehr Informationen als Modus und Median (es geht nicht nur jede Maßzahl ein, sondern auch die Größe der Maßzahl = ihre Stellung in Bezug auf die anderen Maßzahlen);

Streuung

Die Streuungswerte haben die Aufgabe, die Variabilität (=Unterschiedlichkeit) des gemessenen Merkmals zu beschreiben. Dies ist zur Beurteilung empirischer Verteilungen insofern wichtig, als selbst solche Verteilungen, die bezüglich ihres Mittelwertes völlig übereinstimmen, stark voneinander abweichen können.

Variationsweite (absoluter Streubereich, Streuungsbreite, Spannweite)

$R = \text{Range}$

Differenz zwischen der kleinsten und der größten Maßzahl: $R = X_{(max)} - X_{(min)}$

d.h. ich schaue mir die Extremwerte an, und wie weit sie auseinander liegen
je stärker sie auseinander liegen, desto größer die Streuung

Vorteile:

- leicht bestimmbar
- leicht verständlich
- wird von den Extremwerten der Verteilung bestimmt

Nachteile:

- stark von Zufallseinflüssen abhängig
- hängt stark von Stichprobengröße ab (Variationsweiten aus verschiedenen großen Stichproben nicht unmittelbar vergleichbar)
- liefert unzulässige Schätzwerte für die Streuung in der Grundgesamtheit

Voraussetzung:

setzt Ordinalskalenniveau voraus;
denn bei einer Nominalskala kann man eigentlich nicht gut von Dispersion sprechen

Sollte nur bei kleinen Stichproben verwendet werden ($n \leq 12$).

Quartilabstand

Der Quartilabstand (oder Hälftespielraum) ist das Intervall auf der Merkmalsachse, das durch das untere Quartil Q_1 und das obere Quartil Q_3 begrenzt wird.

$Q_{dif} = Q_3 - Q_1 = P_{75} - P_{25}$ Voraussetzung: Ordinalskalenniveau

In diesem Bereich liegen 50% aller Werte. Diese streuen umso stärker, je breiter der Quartilabstand ist.

QA = halber Quartilabstand

$$QA = \frac{Q_3 - Q_1}{2}$$

- wenn kein Intervall- oder Verhältnisskalenniveau, d.h. wenn kein arithm. Mittel berechnet werden kann

Varianz und Standardabweichung

Die Varianz (s^2) ist die Summe der Abweichungsquadrate aller Meßwerte einer Verteilung von ihrem arithmetischen Mittel, dividiert durch die um 1 verminderte Anzahl der Messungen.

Die Standardabweichung (s) [=mittlere quadratische Abweichung] ist die Quadratwurzel aus der Varianz.

Mittelwert der Mittelwertsabweichungen
gibt den Grad der Abweichung einzelner Werte in einer Liste vom Mittelwert aller Werte in der Liste an.

je kleiner die Standardabweichung desto weniger weichen die einzelnen Werte vom Mittelwert ab und desto zuverlässiger ist der Mittelwert.

ist die STABW = 0, dann sind alle Werte gleich

Vorzüge

- wird von zufälligen Extremwerten der Stichprobe kaum beeinflusst
- hängt von allen Meßwerten der Verteilung ab
- stellt eine algebraische Funktion aller Meßwerte der Verteilung dar
- zuverlässiger Schätzwert für die Streuung in der Grundgesamtheit
- für die Prüfstatistik gut geeignet

Voraussetzung:

Intervallskalenniveau

Berechnung:

- heute: PC

Vergleich:

Zentraltendenz bedeutet die Lage der Verteilung durch die in mehrfacher Weise angebbare Mitte auf dem Merkmalskontinuum

Dispersion bezeichnet die Breite der Verteilung

Interpretation:

Zentraltendenz als Maß und Norm für Pegel, Rangstufe, Durch- und Querschnitt

Dispersion als Maß für die Abweichungen der Elemente der statistischen Menge voneinander (z.B. als Maß für die Homogenität einer Gruppe, interindividuelle Übereinstimmung)

Zusammenfassung:

statistische Kennwerte wie Häufigkeitsverteilungen, Mittelwerte, Verteilung usw.:

es geht um

- Reduktion unübersichtlicher Einzelwerte
- auf einige aussagekräftige Kennwerte;
- dabei müssen Erfahrungsnahe (Empirisches Relativ)
- und mathematische Handhabbarkeit (Numerische Relativ) unter einen Hut gebracht werden

Beispiel:

Häufigkeit oder Modalwert sind erfahrungsnah;

Varianz oder Standardabweichung eher erfahrungsfern, dafür aber mathematisch stabiler und dadurch bei Rechenoperationen wie Wahrscheinlichkeitsrechnungen eher einsetzbar.

Variablen

Variable = ein Merkmal eines Systems, das durch Veränderlichkeit charakterisiert ist - zumindest zwei verschiedene Zustände müssen möglich sein

In jedem Experiment spielen drei Arten von Variablen eine Rolle:

"(1) Unabhängige Variable: Ihr Einfluß soll untersucht werden. Dazu werden sie vom Versuchsleiter planmäßig variiert. ...

(2) Abhängige Variable: Die Variable, deren Abhängigkeit von der unabhängigen Variable Gegenstand der Untersuchung ist.

(3) Störvariable: Alle Variablen, die sonst noch (d.h. außer den planmäßig variierten unabhängigen Variablen) einen Einfluß auf die abhängige Variable haben. Sie müssen kontrolliert werden, da sie sonst die Eindeutigkeit der Interpretation gefährden."

(STELZL 1995, 108)

Zusammenhänge von Variablen

"Nach dem Prinzip der Erfahrung kann man zwar Zusammenhänge verschiedener Variablen beobachten, dabei besteht aber auch die Gefahr aus Beobachtungen falsche Schlüsse zu ziehen. Aus diesem Grund ist es wichtig, erstens die Frage zu klären, ob das gemeinsame Auftreten systematisch oder zufällig ist und zweitens zu klären, ob der Zusammenhang interpretiert werden kann." (TRIMMEL 1994, S. 64)

Arten von Zusammenhängen:

- **Kausalzusammenhang**
wörtl. = ursächlich zusammenhängend
ein Ereignis verursacht ein anderes Ereignis;
kritische Überprüfung anhand eines Experimentes;
aufgefundene Kausalzusammenhänge erlauben exakte Voraussagen
- **Korrelation**
wörtl. = einander wechselseitig erforderlich und bedingend;
zwei (oder mehr) Ereignisse treten gleichzeitig auf und hängen irgendwie miteinander zusammen;
diese korrelative Beziehung muß konsistent sein und theoretisch untermauert sein;
- **Koinzidenz**
wörtl. = zusammenfallend; Zufall
zwei (oder mehr) Ereignisse treten zufällig auf; es besteht zwischen ihnen kein theoretisch begründbarer Zusammenhang;
erlaubt keinerlei wissenschaftlich fundierte Voraussagen

Beispiel ZIMBARDO:

"Wenn Sie sich für ein Tennisspiel umziehen, ziehen sie vielleicht einmal zufällig den linken Socken und den linken Schuh zuerst an, dann erst den rechten Socken und den rechten Schuh. Sie gewinnen das schwierige Spiel und entwickeln danach ein Ritual des Umkleidens, das streng mit der linken Seite anfängt. Sie fürchten zu verlieren, wenn Sie das nicht tun. Dieses Risiko gehen Sie lieber nicht ein. Ein rein zufälliges Zusammentreffen - eine *Koinzidenz* - hat eine (abergläubische) Kontrolle über ihr Verhalten erlangt, weil Sie sie wie eine bedeutungsvolle Korrelation behandelt haben.

Während es sich bei einer Koinzidenz um eine zufällige Verbindung von Ereignissen handelt, ist eine Korrelation eine konsistente, stabile Beziehung zwischen zwei oder mehr Variablen [aber keine Kausalbeziehung, Anm. C.B.]. Psychologische Forschung besteht oft in der Suche nach Korrelationen - danach, welche Variablen zusammenhängen und wie stark die Verbindung ist." (ZIMBARDO ⁶1995, S. 20)

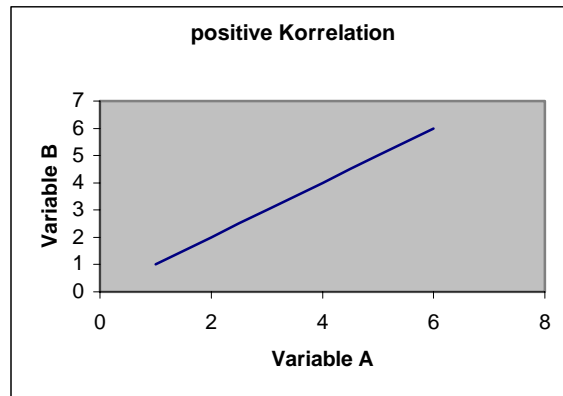
Zusammenhänge, die untersucht wurden:

- Depression und Selbstmord
- Ängstlichkeit und Stottern
- Konsum von Pornographie und Aggression gegenüber Frauen
- Stress und Arbeitsleistung

Arten von Korrelationen (vgl. ZIMBARDO ⁶1995, S. 20-21):

- positive (lineare) Korrelation

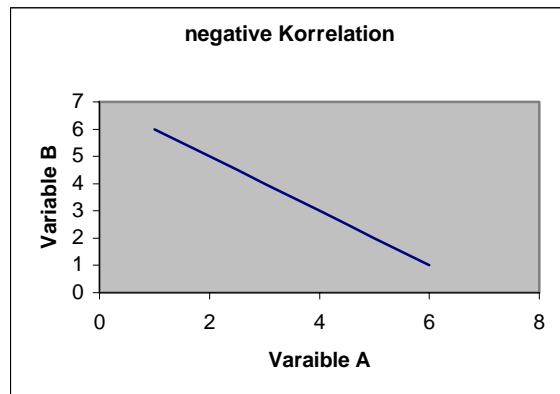
Niedrige Ausprägungen bei der einen Variablen gehen im allgemeinen mit niedrigen Ausprägungen bei der anderen einher, mittlere mit mittleren und hohe mit hohen Ausprägungsgraden.



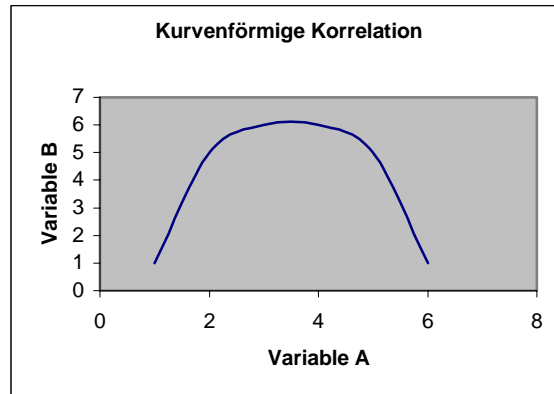
Beispiel: Rauchen und Lungenkrebs; Körpergröße und Gewicht ...

- negative (lineare) Korrelation; gegenläufige Korrelation

Niedrige Ausprägungen der einen Variablen entsprechen im allgemeinen hohe Ausprägungen der anderen Variablen und umgekehrt.



- kurvenlineare Korrelation
siehe Grafik:



Beispiel: Erregungsniveau und Leistungsgrade (auf mittlerem Erregungsniveau ist die Leistung besser als auf einem sehr hohen oder sehr niedrigen Erregungsniveau)

"Korrelationen lassen sich für *Vorhersagen* nutzen. Liegt eine hohe Korrelation, d.h. ein enger Zusammenhang vor, so kann von der einen Variablen ... auf die andere Variable ... geschlossen werden. **Vorhersagen beruhen auf Zusammenhangsbeobachtungen bei einer ganzen Gruppe von Personen. Deshalb liefern sie selten genaue Voraussagen für einzelne Personen.**" (ZIMBARDO ⁶1995, S. 20)

Vorsicht bei der Interpretation von Korrelationen !

(Beispiele aus: WIRTZ/NACHTIGALL 1998, S. 142-143)

z.B. Depressive Menschen haben weniger soziale Kontakte. (Ausmaß der Depressivität korreliert mit der Anzahl der sozialen Kontakte)

Erklärungsmöglichkeiten:

⇒ Depressivität führt zum Rückzug oder zu mangelnder Kompetenz, soziale Kontakte zu pflegen.

⇒ Der Rückgang sozialer Kontakte führt zur Depressivität.

⇒ Depressive Menschen leben in einer reizärmeren Umgebung: dies führt zu erhöhter Depressivität und resultiert in Isolation.

z.B. Die Dauer einer Therapie ist positiv korreliert mit der Befindlichkeit.

⇒ Hieraus folgt nicht sofort, dass die Therapie wirksam ist: Meist beginnen Patienten eine Therapie wenn es ihnen extrem schlecht geht. Es ist aber dann möglicherweise zu erwarten, dass auch ohne Therapie eine gewisse Besserung eintritt.

z.B. Menschen, die Tabletten nehmen fühlen sich unwohler.

⇒ Hieraus lässt sich nicht folgern, dass die Tabletten Ursachen für das Missbefinden sind. Denn, nur diejenigen, denen es schlecht geht, nehmen Tabletten.

Anhang

Beispiele:

Altersverteilung - Hörerbefragung

Messwert	Häufigkeit	kumulierte Häufigk.	Prozent	kumulierte Prozente	Messwert x Häufigkeit	Abstand Messwert v. arithm. Mittel.	quadrierter Abstand d. Messwertes v. arithm. Mittel.	quadrierter Abstand d. Messwertes v. arithm. Mittel. x Häufigkeit
Alter in Jahren	f	f_k	$f\%$	$f\%_K$	$f \times x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f \times (x_i - \bar{x})^2$
20		0	0	0				
21	1	1	3,7	3,7	21	-13,11	171,90	171,90
22	1	2	3,7	7,4	22	-12,11	146,68	146,68
23	1	3	3,7	11,1	23	-11,11	123,46	123,46
24		3		11,1				
25	3	6	11,1	22,2	75	-9,11	82,99	248,97
26		6		22,2				
27		6		22,2				
28		6		22,2				
29		6		22,2				
30	4	10	14,8	37,0	120	-4,11	16,89	67,56
31	2	12	7,4	44,4	62	-3,11	9,67	19,34
32	3	15	11,1	55,5	96	-2,11	4,45	13,35
33	1	16	3,7	59,2	33	-1,11	1,23	1,23
34	1	17	3,7	62,9	34	-0,11	0,01	0,01
35	1	18	3,7	66,6	35	0,89	0,79	0,79
36	2	20	7,4	74,0	72	1,89	3,57	7,14
37		20		74,0				
38		20		74,0				
39		20		74,0				
40		20		74,0				
41	1	21	3,7	77,7	41	6,89	47,47	47,47
42		21		77,7				
43	1	22	3,7	81,4	43	8,89	79,03	79,03
44		22		81,4				
45	1	23	3,7	85,1	45	10,89	118,59	118,59
46	1	24	3,7	88,8	46	11,89	141,37	141,37
47	1	25	3,7	92,5	47	12,89	166,15	166,15
48		25		92,5				
49	1	26	3,7	96,2	49	14,89	221,71	221,71
50		26		96,2				
				96,2				
57	1	27	3,7	99,9	57	22,89	523,95	523,95
					921			2098,70

Wert	Formel	Berechnung	Ergebnis
Modus	häufigster Wert	ablesen „Häufigkeit“	D = 30
Median	teilt in obere u. untere 50%	N = 27 14. Wert, ablesen „kumulierte Häufigkeit“	Z = 32
arithm. Mittel	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$\frac{921}{27}$	$\bar{x} = 34,11$
Variationsweite	$R = X_{(max)} - X_{(min)}$	ablesen Tabelle 57 - 21	R = 36
Quartil ₁	25% aller Werte liegen drunter	ablesen „kumulierte Prozente“	Q1 = 30
Quartil ₃	75% aller Werte liegen drunter	ablesen „kumulierte Prozente“	Q3 = 41
Quartilabstand	Differenz Q ₃ - Q ₁	41 - 30	11
Varianz	$s_x^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}$	$\frac{2098,70}{27 - 1}$	$s^2 = 80,72$
Standardabweichung	$s_x = \sqrt{s_x^2}$	$\sqrt{80,72}$	s = 8,98

Beispiel:

GRAWE/CASPAR/AMBÜHL (1990, S. 301)

geben für die Behandlungsbedingung "Interaktionelle Gruppentherapie" einen Mittelwert für die Anzahl der Sitzungen an: 34,
(Minimum 5, Maximum 94)

Fiktive Möglichkeiten bei n= 16 Patienten:

10x 2 - 4x 31 - 2x 200

15x 3 - 1x 499

Fiktive Möglichkeiten bei n= 16 Patienten: und Angabe der Minimal-/Maximalwerte (5 bzw. 94)

10x 5 - 1x 24 - 5x 94

1x 5 - 2x 7 - 3x 15 - 2x 28, 1x 32, 2x 36, 1x 63, 1x 78, 1x 85, 1x 94

Tabelle: Anzahl der Therapiesitzungen

	Beispiel 1	Beispiel 2	Beispiel 3	Beispiel 4
Patient 1	2	3	5	5
Patient 2	2	3	5	7
Patient 3	2	3	5	7
Patient 4	2	3	5	10
Patient 5	2	3	5	10
Patient 6	2	3	5	15
Patient 7	2	3	5	15
Patient 8	2	3	5	15
Patient 9	2	3	5	28
Patient 10	2	3	5	32
Patient 11	31	3	24	36
Patient 12	31	3	94	49
Patient 13	31	3	94	58
Patient 14	31	3	94	78
Patient 15	200	3	94	85
Patient 16	200	499	94	94
Modus	2	3	5	15
Median	2	3	5	21,5
Arithm.Mittel	34	34	34	34
R	198	496	89	89
QA	29	0	89	43,5
S ²	4.359,20	15.376	1.767,33	901,07
s	66,02	124	42,04	30,02

Empfohlene Literatur

Pä-c-33

Trimmel, M. 1994. Wissenschaftliches Arbeiten in der Psychologie. Wien: Universitätsverlag.

M-b-48

Clauß, G.; Finze, F.-R. & Partzsch, L. 1995. Statistik. Für Soziologen, Pädagogen, Psychologen und Mediziner. Band 1. 2. Auflage. Thun: Deutsch.

M-b-57

Wirtz, M. & Nachtigall, Ch. 1998. Deskriptive Statistik. Statistische Methoden für Psychologen. Teil 1. München: Juventa.

M-b-58

Nachtigall, Ch. & Wirtz, M. 1998. Wahrscheinlichkeitsrechnung und Inferenzstatistik. Statistische Methoden für Psychologen. Teil 2. München: Juventa

M-b-56

Schilling, O. 1998. Grundkurs: Statistik für Psychologen. München: Fink.

Leonhart, R. 2004. Lehrbuch Statistik. Einstieg und Vertiefung. Bern: Huber.

U-w-2

Krämer, W. 1991. So lügt man mit Statistik. 3. Auflage. Frankfurt: Campus.

Krämer, W. 2001. Denkste! Trugschlüsse aus der Welt der Zahlen und des Zufalls. 4. Auflage. München: Piper.

Krämer, W. 2002. Statistik verstehen. 2. Auflage. München: Piper.