

**Ablehnungsbereich:** → *Signifikanzniveau*

**abhängige Gruppen:** Gruppen von Versuchspersonen, in denen jede einzelnen Versuchsperson aus Gruppe A einer äquivalenten Versuchsperson aus Gruppe B entspricht (oder tatsächlich die gleiche Versuchsperson ist).

**abhängige Variable:** Eine der Variablen, deren Änderungen bei einem *Experiment* gemessen wird.

**Abweichungsquadrate:** Das Quadrat der Differenz zwischen einem Rohwert und dem *Mittelwert*, wobei der Mittelwert je nach Kontext derjenige der Gruppe oder der Mittelwert aller Rohwerte sein kann.

**α-Fehler:** Der α-Fehler ist die irrtümliche Annahme von der *Alternativehypothese*  $H_1$  (aufgrund der Stichprobe), obwohl die *Null-Hypothese*  $H_0$  in der Grundgesamtheit vorliegt. Die Wahrscheinlichkeit für einen α-Fehler ist gleich dem *Signifikanzniveau*.

**α-Fehler-Inflation:** Wird eine Vielzahl von statistischen Test gerechnet (z.B. mehrere *t-Test* statt einer *ANOVA*), muss der *Ablehnungsbereich* nach unten korrigiert werden, da die Wahrscheinlichkeit, rein zufällig ein signifikantes Ergebnis zu erhalten, sich erhöht.

**Alternativhypothese:** Behauptet einen Unterschied/Zusammenhang zwischen den verschiedenen Gruppen eines *Designs* und ist damit die Gegenthese zu  $H_0$ .

**ANOVA:** Analysis Of Variance, Varianzanalyse. Ein parametrisches Verfahren, um eine Prüfgröße zu ermitteln, anhand der über die Gültigkeit der  $H_0$  entschieden werden kann. Kommt zum Einsatz, sobald das *Design* über ein Zwei-Gruppen-Design hinausgeht (ansonsten könnte z.B. der *t-Test* zur Anwendung kommen). Die ANOVA wird in folgenden Schritten ausgeführt:

1. Gesamtvarianz ermitteln
2. Treatmentvarianz (Varianz, die vom Treatment abhängt) ermitteln
3. Fehlervarianz (Grad der nicht erklärbaren Streuung innerhalb einer Gruppe) ermitteln
4. F-Bruch berechnen

**Anpassungstests:** Bei einem Anpassungstest werden die erwarteten (aus der Theorie vorhergesagten) Werte mit den tatsächlich empirisch ermittelten verglichen. Beispiele: *Chi-Quadrat-Test*, Kolmogoroff-Smirnov-Test (falls sehr viele Kategorien auftreten).

**A posteriori Wahrscheinlichkeit:** Das Ergebnis des *Bayes-Theorems* wird als A posteriori W. bezeichnet, da es durch Modifizierung von *A priori Wahrscheinlichkeiten* unter Verwendung von bedingten Wahrscheinlichkeiten zustande kommt.

**A priori Wahrscheinlichkeit:** Wahrscheinlichkeiten, die im *Bayes-Theorem* auftauchen und bereits vor einer empirischen Untersuchung ermittelt werden können.

**Bayes-Statistik:** Die Bayes-Statistik beruht auf dem *Bayes-Theorem*. Der Hauptunterschied zwischen der frequentistischen und der bayesianischen Schule ist, dass in der frequentistischen Schule danach gefragt wird, wie wahrscheinlich das Auftreten der empirisch ermittelten Daten unter Annahme der *Nullhypothese* ist und somit nur indirekt *Hypothesen* testet, während der bayesianische Ansatz direkt nach der Wahrscheinlichkeit einer Hypothesen unter den ermittelten Daten fragt.

**Bayes-Theorem:** Das Bayes-Theorem ist die Grundlage der *Bayes-Statistik*. Es beantwortet die Frage, wie wahrscheinlich eine Hypothese ( $A_j$ ) unter den ermittelten Daten (B) ist.

$$\text{Formel: } p(A_j | B) = \frac{p(A_j) \cdot p(B | A_j)}{\sum_{i=1}^k p(A_i) \cdot p(B | A_i)}, \text{ mit } k \text{ als Anzahl der Hypothesen } A_i \text{ und den Daten } B$$

**Bestimmtheitsmaß:** → *Determinationskoeffizient*

**β-Fehler:** Die irrtümliche Annahme der *Null-Hypothese*  $H_0$  (aufgrund der Stichprobe), obwohl die *Alternativehypothese*  $H_1$  in der Grundgesamtheit vorliegt.

**between-subjects-Design:** Ein Untersuchungsplan (*Design*), bei dem in den verschiedenen Gruppen auch verschiedene Versuchspersonen untersucht werden (keine abhängigen Messungen).

**Blickbewegungsmessungen:** Können die Frage beantworten, wann welche Informationen aufgenommen werden. Es gibt invasive (z.B. mit Kontaktlinsen) und nicht-invasive (beruhen meist auf Lichtreflexmessung – wie Corneal Reflection oder Prukinje Image Tracking – oder sichtbaren Merkmalen des bewegten Auges, z.B. Limbus Tracker) Methoden; auch bei den Aufzeichnungsapparaten gibt es mehrere Alternative (z.B. head-mounted, remote eye tracking devices, table-mounted).

**Chi-Quadrat Test:** Ein Verfahren, um *Unterschiedshypothesen* zu überprüfen, die sich auf Nominaldaten beziehen. Beim Chi-Quadrat-Test wird überprüft, in wie weit sich die erwarteten Werte mit den tatsächlich gemessenen decken (*Anpassungstest*); es kann eingesetzt werden, um die angenommene *Normalverteilung* der Werte zu überprüfen.. Die entsprechende *Prüfgröße* errechnet sich nach folgender Formel:

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - f_{e(i)})^2}{f_{e(i)}}, \text{ wobei } k \text{ die Anzahl der Kategorien, } f_k \text{ den gemessenen Wert und } f_{e(i)}$$

den erwarteten Wert für die Kategorie  $i$  angibt.

**Cohens' d:** Maß für die *Effektstärke* zwischen zwei Gruppen (Mittelwertdifferenz durch Standardabweichung).

Formel:  $Cohens' d = \frac{\bar{x}_1 - \bar{x}_2}{s}$ , mit  $\bar{x}_1$  und  $\bar{x}_2$  Gruppenmittelwerte und  $s$  Standardabweichung aller Werte in den beiden Gruppen.

**Design:** Der Aufbau eines Experiments (auch Untersuchungsplan genannt), eine „konzeptionelle Brücke“ zwischen der Untersuchung und ihrer statistischen Auswertung. Von „Designs“ oder auch „faktoriellen Designs“ spricht man besonders bei Untersuchungen, die mittels varianzanalytischer Verfahren ausgewertet werden, ansonsten wird eher „Untersuchungsplan“ benutzt.

**Determinationskoeffizient:** Gibt die Effektstärke für korrelationsanalytische Ansätze an; berechnet sich als das Quadrat des *Korrelationskoeffizienten*:  $r^2/R^2$ ;  $r^2$  wird für eine bivariate Korrelation (zwei Variablen) herangezogen,  $R^2$  bei einer multivariaten.

**Dispersionsmaße:** → *Streuungsmaße*

**Drittvariable:** Eine statistisch signifikante *Korrelation* zwischen zwei Variablen  $x$  und  $y$  lässt sich grundsätzlich stets vierfach interpretieren:

1.  $x$  beeinflusst  $y$  kausal
2.  $y$  beeinflusst  $x$  kausal
3.  $x$  und  $y$  werden von einer dritten oder weiteren Variablen kausal beeinflusst
4.  $x$  und  $y$  beeinflussen sich wechselseitig kausal

Da der *Korrelationskoeffizient* nichts darüber aussagt, welche der vier Interpretationen richtig ist, lässt sich Kausalität nur widerlegen und nie eindeutig bestätigen. Analog kann bei varianzanalytischen Untersuchungsplänen eine *Konfundierung* auftreten.

**Effektstärke:** Bezeichnet einen Unterschied zwischen min. zwei Parametern der Population. Ein Maß für die Effektstärke ist u.a. *Cohens d* (varianzanalytisch) und der *Determinationskoeffizient* (Korrelationsanalyse).

**Experiment:** In einem Experiment werden die *unabhängigen Variablen* (UV) systematisch variiert und der resultierende Effekt auf die *abhängigen Variablen* (AV) ermittelt. Dabei sind Kontrolltechniken anzuwenden, die Verzerrungen durch Störvariablen entgegenwirken und

trolltechniken anzuwenden, die Verzerrungen durch Störvariablen entgegenwirken und der sachgerechten Interpretation der Effekte dienen. Ein Experiment fordert die randomisierte Zuweisung der Versuchspersonen durch den Untersuchungsleiter, ansonsten liegt ein *Quasi-Experiment* vor.

**Externe Validität:** Eine Untersuchung gilt als extern valide, wenn ihr Ergebnis über die besonderen Bedingungen der Untersuchungssituation und über die untersuchten Personen hinausgehend generalisierbar ist. Die externe Validität sinkt mit wachsender Unnatürlichkeit der Untersuchungsbedingungen (z.B. im Labor) bzw. mit abnehmender Repräsentativität der untersuchten Stichproben.

**F-Bruch:** Die *Prüfgröße*, die z.B. beim der *ANOVA* ermittelt wird; der Quotient aus der *Treatmentvarianz* und der *Fehlervarianz*.

**Falsifikation:** *Hypothesen* können nach dem *kritischen Rationalismus* nie bewiesen, sondern nur falsifiziert, also widerlegt werden. Sie gelten also „nur vorläufig“, solange niemand sie falsifizieren kann.

**Fehlervarianz:** Ermittelt die Streuung der Messergebnisse zwischen den Versuchspersonen innerhalb einer Gruppe (die also nicht durch die Experimentbedingungen erklärt werden kann). Ist umso größer, je kleiner der Haupteffekt des Treatments ist.

Formel: 
$$\text{Fehlervarianz} = \frac{\sum_{i=1}^p \sum_{m=1}^n (x_{mi} - \bar{x})^2}{p(n-1)}$$
, wobei  $p$  die Anzahl der Stufen des untersuchten

Faktors,  $n$  die Anzahl der Versuchspersonen (und damit  $p(n-1)=df$ , also die *Freiheitsgrade* angibt),  $x_{mi}$  den Rohwert und  $\bar{x}_i$  den Gruppenmittelwert bezeichnet.

**Fragebogen:** Untersuchungsmethode.

**Freiheitsgrad:** Die Anzahl der bei der Berechnung eines Kennwerts frei variierbaren Werte. Bsp.: Die Summe der Differenzen aller Werte von ihrem Mittelwert ergibt 0. Es können also von  $n=10$  Werten nur 9 ( $=n-1$ ) beliebig gewählt werden.

**frequentistische – bayesianische Schule:** Zwei verschiedene, miteinander nicht kompatible Varianten der statistischen Entscheidungstheorie. Der Hauptunterschied zwischen der frequentistischen und der bayesianischen Schule ist, das in der frequentistischen Schule danach gefragt wird, wie wahrscheinlich das Auftreten der empirisch ermittelten Daten unter Annahme der *Nullhypothese* ist und somit nur indirekt *Hypothesen* testet, während der bayesianische Ansatz direkt nach der Wahrscheinlichkeit einer Hypothesen unter den ermittelten Daten fragt.

**gemischtes Design:** Ein faktorielles *Design*, bei dem unterschiedliche Gruppen jeweils mehrfach untersucht werden, z.B. indem man vier Gruppen vor und nach Eintritt eines Ereignisses testet.

**geschichtete Stichprobe:** Eine Stichprobe, in der die Versuchspersonen so ausgewählt sind, dass sie vermeintlich relevante Verhältnisse der *Grundgesamtheit* widerspiegeln, um die Stichprobe repräsentativ zu machen. Beispiele: Berufsgruppen, Altersverteilung, etc.

**Grundgesamtheit:** Alle potentiell untersuchbaren Einheiten, die ein bestimmtes Merkmal aufweisen (auch als *Population* bezeichnet).

**Gütekriterien:** 1. *Reliabilität:* Wie verlässlich ist das Ergebnis? re-test reliability: Liefert das Experiment bei Wiederholung das gleiche Ergebnis? Inter-rater reliability: Liefert das Experiment auch bei einem anderen Experimentator das gleiche Ergebnis?

2. *Objektivität*: Sind die Ergebnisse unabhängig von der Versuchsdurchführung?
3. *Validität*: Untersucht das Verfahren wirklich das, was es zu untersuchen behauptet? Kriteriumsvalidität: Korrelation der Ergebnisse mit „objektiven“ Außenkriterien; Konstruktvalidität: Zusammenhänge zu verwandten theoretischen Konstrukten.

**Haupteffekt:** Ein statistisch signifikanter Effekt über alle Stufen eines Faktors. („Alle Gruppen unterscheiden sich.“)

**Hypothese:** Eine empirisch überprüfbare Behauptung, die (meistens) aus einer Theorie abgeleitet wurde. Es wird zwischen *Zusammenhangs-* und *Unterschiedshypothesen* unterschieden. Eine spezielle Abart sind die beiden statistischen Hypothesen: die *Null-Hypothese* und die *Alternativhypothese*, die zur Auswertung der ermittelten Daten aufgestellt werden.

**Internetforschung:** Online-*Experimente* weisen im Gegensatz zu Laborexperimenten eine höhere *externe Validität* auf, allerdings stehen den Vorteilen wie einfacher Beschaffung von Versuchspersonen, Wiederverwendbarkeit der Steuerungsprogramme und dem Zugang zu speziellen Gruppen (z.B. Häftlingen) auch gravierende Nachteile gegenüber, da einerseits die *Grundgesamtheit* nicht bekannt ist und andererseits keine *Repräsentativität* gegeben ist.

**interne Validität:** Eine Untersuchung ist intern valide, wenn ihr Ergebnis eindeutig interpretierbar ist. Die interne Validität sinkt mit wachsender Anzahl plausibler Alternativverklärungen für das Ergebnis aufgrund nicht kontrollierbarer Störvariablen.

**Intervallschätzung:** Bei der Intervallschätzung wird ein *Konfidenzintervall* (auch Vertrauensintervall genannt) angegeben, das den Bereich eines Merkmals kennzeichnet, in dem sich 95% (oder 99%) aller möglichen Parameter der *Grundgesamtheit* befinden, die den empirisch ermittelten Kennwert erzeugen können.

**Intervallskala:** *Skalenniveau*. Eine Intervallskala kann sinnvoll Differenzen zwischen Werten angeben.

**Klumpenstichprobe:** Eine Klumpenstichprobe ist nicht repräsentativ, da hier eine bereits Gruppe nach Erreichbarkeit ausgesucht wird, z.B. die Schulklasse, in der ein befreundeter Lehrer unterrichtet.

**Konfidenzintervalle:** Ein Intervall, in dem sich 95% bzw. 99% aller möglichen Parameter der *Grundgesamtheit* befinden, die den empirischen Kennwert (z.B. *Mittelwert* etc.), dem das Konfidenzintervall zugrunde liegt, erzeugt haben können. Durch Vergrößern des Stichprobenumfangs lassen sich die Konfidenzintervalle verkleinern, die Schätzung also präzisieren.

**Konfundierung:** Ein im Aufbau des *Untersuchungsplans* liegender Fehler, der für Unklarheit sorgt, ob die ermittelten Effekte durch die Variation der *unabhängigen Variablen* zustande gekommen sind oder durch die Untersuchungsmethode. Von einer Konfundierung spricht man im wesentlichen im Zusammenhang mit varianzanalytischen Designs; das Pendant bei der Korrelations- bzw. Regressionsrechnung ist das Problem der *Drittvariable*.

**Kontraste:** Kontraste sind Einzelvergleiche zwischen zwei Stufen eines Faktors. **A priori Kontraste(=lineare Kontraste)** sind von vornherein geplante Kontraste, **post hoc Kontraste** sind nachträglich vorgenommene Vergleiche, bei denen das *Signifikanzniveau* abgesenkt werden sollte.

**Korrelation:** Die Korrelation ist ein Maß für die Stärke des Zusammenhangs zwischen zwei Variablen und gehört zu den parametrischen Verfahren.

**Kovarianz:** Ein Maß für das Miteinander-Variieren zweier Messwertreihen  $x$  und  $y$ . Errechnet sich durch:

$$\text{cov}(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{n}, \text{ mit } n \text{ verschiedenen Messwerten in den Messwertreihen } x \text{ und } y \text{ und deren Mittelwerten } \bar{x} \text{ und } \bar{y}.$$

**Kriteriumsvariable:** Die *abhängige Variable* bei Korrelations- und Regressionsrechnungen; eine Variable, die mittels einer *Prädiktorvariablen* und einer Regressionsgleichung vorhergesagt werden kann.

**Kritischer Rationalismus:** Hauptvertreter war Sir Karl Popper; der kritische Rationalismus behauptet, dass es eine Außenwelt gibt; allerdings ist die einzig naturwissenschaftliche Methode die *Falsifikation*.

**Likelihood Wahrscheinlichkeit:** Eine bedingte Wahrscheinlichkeit, die im *Bayes-Theorem* auftritt und die Wahrscheinlichkeit der Daten ( $B$ ) unter der zu überprüfenden *Hypothese*( $A$ ) angibt:  $P(B|A)$ .

**Messwiederholungsdesign:** *Untersuchungsplan*, bei dem die Messung bei derselben Gruppe nach einer bestimmten Zeitspanne/Ereignis wiederholt wird. → *within-subject-design*

**Modalwert:** Der häufigste Wert einer Messreihe, auch als Modus bezeichnet. Der Modalwert gehört zu den Maßen der zentralen Tendenz in der deskriptiven Statistik und kann auf allen Skalen ermittelt werden.

**Multivariat:** In multivariaten Verfahren werden mehr als eine *unabhängige Variable* untersucht

**Median:** Der Wert, der die Messreihe in zwei Hälften teilt. Der Modalwert gehört zu den Maßen der zentralen Tendenz in der deskriptiven Statistik und kann auf allen Skalen ermittelt werden

**Nicht-Parametrische Verfahren:** Verfahren der Inferenzstatistik, die nicht so viele Voraussetzungen (und meist eine geringere *Teststärke*) haben wie die *parametrischen Verfahren*; werden auch als *verteilungsfreie Verfahren* bezeichnet. Beispiele: Mann-Whitney U-Test, Wilcoxon Signed Rank, Kruskal-Wallis, Chi-Quadrat

**Nominalskala:** *Skalenniveau*. Skala, auf der nur Gleichheit/Ungleichheit ermittelt werden kann.

**Normalverteilung:** Die wichtigste Verteilung der Statistik mit einem glockenförmigen Verlauf (asymptotische Annäherung an die x-Achse), bei der *Modalwert*, *Median* und *Mittelwert* zusammenfallen. Die **Standartnormalverteilung** hat einen Mittelwert von Null und eine Standardabweichung von Eins. Jede Normalverteilung kann durch die *z-Transformation* in die Standartnormalverteilung überführt werden.

**Objektivität:** *Gütekriterium*. Sind die Ergebnisse unabhängig von der Versuchsdurchführung?

**Operationalisierung:** Die Verfügbarmachung eines theoretischen Konstrukts für die empirische Überprüfung.

**Ordinalskala:** → *Rangskala*

**Panel (z.B. Online-Panel):** Geschichtete Pools von Versuchspersonen, die mehrfach benutzt werden.

**Parametrische Verfahren:** Verfahren der Inferenzstatistik, die über eine relative große *Teststärke* verfügen, für deren Anwendung die Daten allerdings eine Reihe von Voraussetzungen erfüllen müssen: (1) Varianzhomogenität: Die Varianzen in den verschiedenen Gruppen dürfen sich nicht signifikant unterscheiden, (2) die Fehlerkomponenten (Abweichung eines Messwertes vom jeweiligen Stichprobenmittel) müssen normalverteilt und unabhängig sein, d.h. alle

Gruppen gleichermaßen betreffen, und (3) die Variablen müssen *Intervallskalenniveau* haben. Beispiele: *t-Test*, *ANOVA*, *Korrelations- u. Regressionsrechnung*.

**Partialkorrelation:** Das rechnerische Eliminieren von bekannten Drittvariablen aus den Variablen einer Korrelations-/Regressionsrechnung. Das funktioniert allerdings nur, wenn ein Zusammenhang zwischen den Variablen und der Drittvariable, die „rauspartialisiert“ werden soll, bekannt ist. Eine alternative Methode (falls kein solcher Zusammenhang bekannt ist) besteht darin, die Drittvariable beim Erheben der Daten konstant zu halten. Bei einer **Semipartialkorrelation** wird die Drittvariable nur aus einer der Variablen herausgerechnet.

**Population:** → *Grundgesamtheit*.

**Power:** Die Teststärke gibt die Wahrscheinlichkeit an, einen *Effekt* bestimmter Größe nachzuweisen, falls dieser existiert. Die Power steigt mit zunehmendem Stichprobenumfang, sinkt allerdings bei wachsender *Streuung*. *Parametrische* Tests haben normalerweise eine höhere Power als *nicht-parametrische*, ebenso einseitige Test gegenüber zweiseitigen und *within-subject-Designs* im Vergleich zu *between-subject-Designs*.

**Prädiktorvariable:** Bezeichnung für die *unabhängigen Variablen* bei Regressions- und Korrelationsrechnungen.

**Produkt-Moment Korrelation:** ist ein Maß für die Stärke eines Zusammenhangs zwischen zwei Variablen  $x$  und  $y$ , das gegenüber Maßstabsveränderungen der untersuchten Merkmale invariant ist (im Prinzip eine *z-transformierte Kovarianz*). Die Abkürzung für den sog. Korrelationskoeffizienten ist  $r$ , errechnet wird er durch folgende Formel:

$$r = \frac{\text{cov}(x, y)}{s_x \cdot s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{n \cdot s_x \cdot s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}, \text{ mit } s_x \text{ und } s_y \text{ Standardab-}$$

weichung für  $x$  bzw.  $y$  und  $n$  der Anzahl der Messwerte.

**Prüfgröße:** Ein Wert, in dem die Informationen der Messung verdichtet werden, um dann anhand der durch Tabellen zugeordneten Wahrscheinlichkeit für den Wert der Prüfgröße über die Annahme oder Ablehnung der  $H_0$  zu entscheiden, je nachdem ob die tabellierte Wahrscheinlichkeit der Prüfgröße in den *Ablehnungsbereich* fällt oder nicht. Der vom statistischen Verfahren ermittelte Wert wird als **empirische Prüfgröße** bezeichnet, die dazugehörige Wahrscheinlichkeit als **theoretische/kritische Prüfgröße**.

**Punktschätzung:** Die Schätzung von Parametern der *Grundgesamtheit* (wie *Mittelwert*, *Streuung* etc.) anhand eines einzigen Wertes, der aus den ermittelten Daten, also aus der Stichprobe, errechnet wurde. Eine weitere Möglichkeit ist die *Intervallschätzung*.

**Quasi-Experiment:** Ein *Experiment* ohne randomisierte Verteilung der Versuchspersonen auf die Gruppen und daher von niedriger *interner Validität*; *Online-Experimente* sind immer Quasi-Experimente.

**Range:** Streubreite der Messwerte, ergibt sich aus der Differenz zwischen dem größten und dem kleinsten Messwert; gehört zu den *Dispersionsmaßen*.

**Rangskala:** *Skalenniveau*. Eine Skala, auf der Relationen zwischen Werten (größer/kleiner) ermittelt werden können. Auch als *Ordinalskala* bezeichnet.

**Regression:** Ein Vorhersagemodell für den Zusammenhang zwischen den *Prädiktor-* und den *Kriteriumsvariablen*. Bei der **einfachen Regression** gibt es nur eine Prädiktor- und eine Kriteriumsvariable, bei der **multiplen Regression** mehrere Prädiktorvariablen und eine Kriteriumsvariable und bei der **schrittweisen multiplen Regression** werden nach und nach die

nicht-signifikanten Prädiktorvariablen eliminiert. Die **logistische Regression** bezieht sich auf nominalskalierte Kriteriumsvariablen.

**Reliabilität:** *Gütekriterium:* Wie verlässlich ist das Ergebnis? re-test reliability: Liefert das Experiment bei Wiederholung das gleiche Ergebnis?. inter-rater reliability: Liefert das Experiment auch bei einem anderen Experimentator das gleiche Ergebnis?

**Repräsentativität:** Eine Stichprobe gilt als repräsentativ, wenn sie einige (spezifische R.) oder alle (globale R.) Merkmale der Grundgesamtheit widerspiegelt. Zwei Verfahren, um eine repräsentative Stichprobe zu bekommen, sind zum einen die zufällige *Stichprobenziehung*, bei der eine große Anzahl von Versuchspersonen benötigt wird, oder das Arbeiten mit *geschichteten Stichproben*. Die *Klumpenstichprobe* ist nicht repräsentativ.

**Signifikanz:** Eine empirischen Untersuchung gilt als signifikant, wenn die Prüfgröße in den Ablehnungsbereich fällt, wenn also die Wahrscheinlichkeit für das Ergebnis unter Annahme der  $H_0$  kleiner als 5% (signifikant) oder 1% (sehr signifikant) ist.

**Signifikanzniveau:** Das Signifikanzniveau (oder Ablehnungsbereich) legt fest, bei welchem Ergebnis der *Prüfgröße* die  $H_0$  verworfen wird und legt damit gleichzeitig auch die  $\alpha$ -Fehler-Schranke fest (Standartwerte in den Sozialwissenschaften sind 5% bzw. 1%).

**Skalenniveau:** Gibt an, welche Operationen mit den gemessenen Werten der Variablen durchgeführt werden können, ohne das dabei die Aussagebarkeit verloren geht. Es gibt die *Nominalskala*, die *Rangskala*, die *Intervallskala* und die *Rationalskala*

**Standardabweichung:** Wurzel aus der *Varianz*,  $s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ , wobei  $x_i$  die Rohwerte angibt,  $\bar{x}$  den Mittelwert und  $n$  die Anzahl dieser Rohwerte.

**Stichprobenziehung:** Entscheidend für die *Repräsentativität* einer Untersuchung ist das Verfahren, mit dem die Stichprobe ermittelt wird, von der auf die Grundgesamtheit geschlossen werden soll. Eine große Anzahl von Versuchspersonen alleine ist dafür nicht ausreichend.

**Streuungsmaße:** Begriff aus der deskriptiven Statistik; Streuungsmaße oder **Dispersionsmaße** sind Angaben über die Verteilung der Messwerte.

**Summe der Abweichungsquadrate:** Die Summe der quadrierten Differenzen zwischen den Messwerten und dem *Mittelwert*. Je nach Kontext kann dabei der Gruppenmittelwert oder das Gesamtmittel zum Einsatz kommen.

**Test-Stärke:**  $\rightarrow$  *Power*

**Theorie:** Kohärentes System über einen Sachverhalt, das zwar nie bewiesen werden kann, aus dem aber empirisch falsifizierbare *Hypothesen* abgeleitet werden können.

**t-test:** Ein *parametrisches Verfahren* der Inferenzstatistik, um einen Zwei-Gruppen-Vergleich anzustellen. Sollen mehr als zwei Gruppen verglichen werden, empfiehlt sich eine Varianzanalyse (*ANOVA*).

**unabhängige Variable:** Eine Variable in einem *Experiment*, die kontrolliert verändert wird, um den Effekt auf davon *abhängige Variablen* zu bestimmen.

**Univariat:** In univariaten Verfahren wird nur eine *unabhängige Variable* untersucht.

**Unterschiedshypothese:** Eine *Hypothese*, die einen Unterschied zwischen den verschiedenen Gruppen in einem *Experiment* behauptet; die Formulierung einer Unterschiedshypothese führt normalerweise zu einem *faktoriellen Design* und einer Auswertung mittels *Varianzanalyse (ANOVA)*.

**Untersuchungsplan:**  $\rightarrow$  *Design*

**Validität: Gütekriterium:** Untersucht das Verfahren wirklich das, was es zu untersuchen behauptet?  
 Kriteriumsvalidität: Korrelation der Ergebnisse mit „objektiven“ Außenkriterien; Konstruktvalidität: Zusammenhänge zu verwandten theoretischen Konstrukten.

**Varianz:** Die durchschnittliche Abweichung der Messwerte vom Mittelwert.

Formel:  $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ , mit  $n$  als Anzahl der Messwerte,  $x_i$  den einzelnen Messwerten und  $\bar{x}$  ihrem (arithmetischem) Mittel.

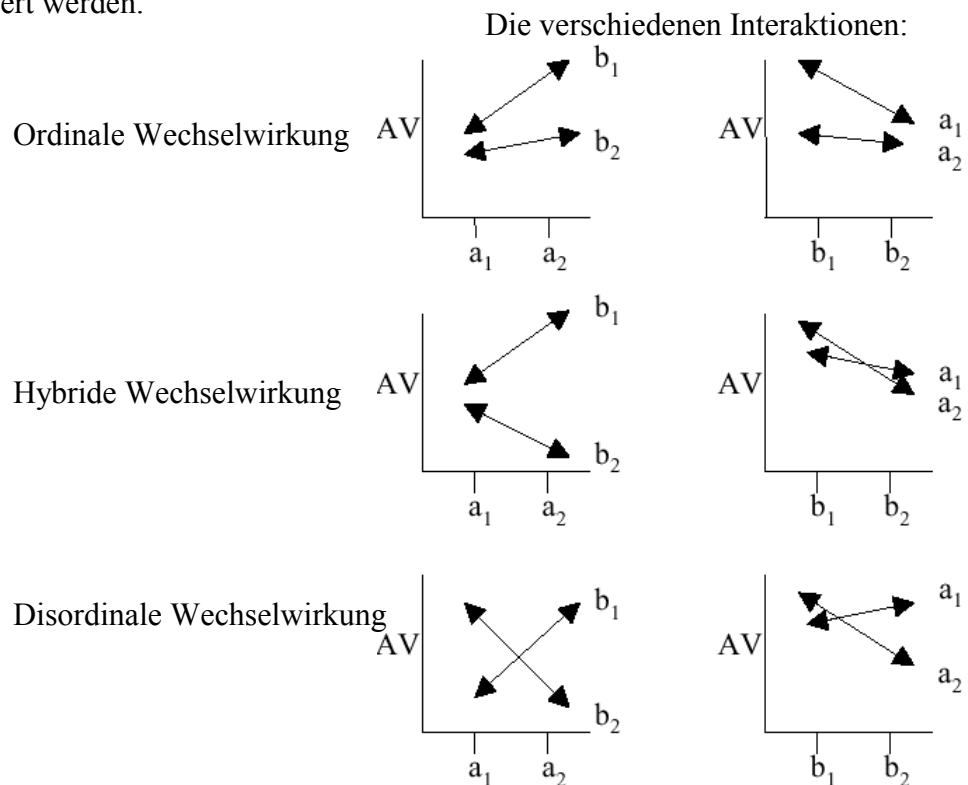
**Varianzanalyse:** → ANOVA

**Verifikation:** Begriff aus dem Neopositivismus (Wiener Kreis); nach dem Verifikationsprinzip kann die Geltung von *Hypothesen* durch *Experimente* bewiesen werden.

**Verteilungsfreie Verfahren:** Verfahren, die nicht als Voraussetzung eine bestimmte Verteilung (meist die *Normalverteilung*) der Daten erwarten, also alle *nicht-parametrischen Verfahren*.

**Vertrauensintervalle:** Konfidenzintervall → *Intervallschätzung*

**Wechselwirkung:** Bei einer Wechselwirkung (oder Interaktion) ist der Unterschied zwischen dem *Effekt* einzelner Faktorstufen abhängig von den Faktorstufen eines anderen *Faktors*, d.h. der Unterschied zwischen A und B fällt für X anders aus als für Y, wobei A und B Stufen des ersten und X und Y Stufen des zweiten Faktors sind. (Bsp: Der Unterschied zwischen dem Einfluss eines Medikaments auf Männer bzw. Frauen (Factor Geschlecht) ist bei Gewichtsklasse A anders als bei Gewichtsklasse B (Factor Gewicht)). Man unterscheidet drei Typen von Wechselwirkungen, die sich am einfachsten anhand von Mittelwertsverlaufkurven klassifizieren lassen: Ordinale, hybride und disordinale Interaktionen. Unter Umständen können Wechselwirkungen Haupteffekt bedeutungslos werden lassen und sollten daher vor der Auswertung kontrolliert werden.



**within-subjects-Design:** Ein Versuchsdesign, bei dem *abhängige Gruppen* untersucht werden, indem z.B. zwei Messungen zu verschiedenen Zeitpunkten an derselben Gruppe vorgenommen werden..

**Z-Transformation:** Von jedem Wert einer Verteilung wird der Mittelwert subtrahiert und diese Differenz durch die *Standardabweichung* geteilt. Die resultierende Verteilung ist die Standardnormalverteilung mit einem Mittelwert von Null und einer Standardabweichung von eins.

$$\text{Formel: } \forall x : z = \frac{x - \hat{x}}{s} = \frac{x - \hat{x}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x})^2}}$$

**Zufallsstichprobe:** Eine Art der *Stichprobenziehung*, bei der zufällig ausgewählte Versuchspersonen untersucht werden; um damit eine repräsentative Stichprobe zu erhalten, braucht man eine große Anzahl von Versuchspersonen

**Zusammenhangshypothese:** Ein *Hypothese*, die einen Zusammenhang zwischen zwei Variablen behauptet.