

Ein kleines Glossar von Grundbegriffen der Statistik

Arithmetisches Mittel (das, was man normalerweise unter «Durchschnitt» versteht)

Summe aller Werte geteilt durch die Anzahl der Werte.

Beispiel: Summe aller Noten einer Klasse geteilt durch die Zahl der Schüler/innen. (Es gibt auch ein *geometrisches Mittel*, auf das ich hier nicht eingehe – bei Wikipedia ist es recht gut erklärt.)

Erwartungswert

Der Wert, der sich bei einem Zufallsprozess «erwarten lässt», präziser: der sich als Durchschnitt herausstellt, wenn man den Prozess oft wiederholt. Wenn ich eine Münze werfe, habe ich keine Ahnung, ob sie auf Kopf oder Zahl fällt, aber auf lange Sicht werde ich in 50 Prozent der Würfe Kopf und in 50 Prozent der Würfe Zahl werfen.

Die physikalische Größe der Halbwertszeit des radioaktiven Zerfalls beruht beispielsweise auf Erwartungswerten: Darüber, wann ein einzelnes Atom radioaktiv zerfällt, kann man keine Aussage machen. Man kann aber angeben, mit welcher Wahrscheinlichkeit ein Atom in einem bestimmten Zeitraum zerfällt – oder anders ausgedrückt: Wie lange es dauert, bis es mit 50-prozentiger Wahrscheinlichkeit zerfallen ist. Weil Atome so klein sind und man es eigentlich immer mit extrem vielen Atomen zu tun hat, ist die Halbwertszeit ein sehr präziser Wert.

Externe Validität

Verallgemeinerbarkeit. Resultate einer Untersuchung einer Stichprobe sind dann extern valide, wenn sie sich auf die Grundgesamtheit übertragen lassen.

Korrelation

Maß der Übereinstimmung zweier Indikatoren.

Die Korrelation wird mit dem Korrelationskoeffizienten r ausgedrückt. r kann Werte zwischen -1 und 1 annehmen. $r = 1$ bedeutet, dass die Werte vollständig korrelieren: Steigt Wert 1, steigt Wert 2 proportional dazu. $r = -1$ bedeutet, dass die Werte umgekehrt korrelieren: Steigt Wert 1, sinkt Wert 2 um den gleichen Faktor. $r = 0$ bedeutet, dass es keine Korrelation gibt. Zwei Reihen von Zufallswerten weisen (bei ausreichender Stichprobengröße) ein r von 0 auf.

Eine hohe Korrelation kann ein Hinweis auf eine Kausalität sein – muss aber nicht!

Median

Der Median teilt eine Stichprobe in zwei gleich große Hälften: Die Hälfte aller Werte einer Stichprobe sind größer, die andere Hälfte kleiner als der Median. Der Median ist das 50er-Perzentil.

Modus

Häufigster Wert einer Stichprobe.

Normalverteilung

Ein häufiger Typus von Wahrscheinlichkeitsverteilungen, wie er beispielsweise in der Verteilung von Intelligenzquotienten in einer Bevölkerungsgruppe auftritt. Als Graph aufgezeichnet, ergibt sich eine «[gauß'sche] Glockenkurve». Weil die Glockenkurve symmetrisch ist und der Zenit der Kurve in der Mitte liegt, sind bei der Normalverteilung *arithmetisches Mittel*, *Median* und *Modus* identisch.

Operationalisierung (Messbarmachung)

Legt fest, wie ein Wert gemessen wird. Wenn zwei Statistiken zum selben Thema eine unterschiedliche Aussage treffen («Die Ungleichheit nimmt zu» vs. «Die Ungleichheit nimmt ab»), liegt das oft daran, dass unterschiedlich operationalisiert wurde.

Prozente und Prozentpunkte

Mit Prozenten zu rechnen, ist tricky: 100 Prozent minus 50 Prozent gibt 50 Prozent plus 50 Prozent gibt 75 Prozent – denn «plus 50 Prozent» bedeutet «mal 1,5». Will man mit den Prozentzahlen wie mit «normalen» Maßeinheiten rechnen, muss man von *Prozentpunkten* sprechen: 50 Prozent plus 50 Prozentpunkte gibt 100 Prozent.

Will man nach einer Wahl angeben, wie viel eine Partei gegenüber der letzten Wahl gewonnen oder verloren hat, spricht man in der Regel von den Prozentpunkten (auch wenn viele Journalist/innen diese Prozentpunkte fälschlicherweise «Prozente» nennen).

Perzentil

Ein Perzentil teilt eine Stichprobe in 100 gleich große Segmente. Das erste Perzentil ist der Wert, der von einem Prozent der Stichprobe erreicht und von 99 Prozent überschritten wird, usw. Das 50er-Perzentil ist der *Median*. Man spricht auch von Dezilen (10er-Perzentil), Quintilen (20er-Perzentil), Quartilen (25er-Perzentil) usw.

Wer Kinder hat, kennt den Begriff vom Kinderarzt: Kinderärzte sprechen von Perzentilen, um die Körpergröße oder das Gewicht eines Kindes zu den Kindern seines Alters in Bezug zu setzen.

Repräsentativität

Eine Stichprobe gilt als repräsentativ, wenn sie bezüglich wichtiger Parameter (bei Umfragen beispielsweise: Geschlecht, Alter, Bildungsniveau, Wohnort ...) die Proportionen der Gesamtheit abbildet. Von einer repräsentativen Stichprobe lässt sich aufs Ganze schließen – wobei die Fehlermarge (das *Vertrauensintervall*) umso kleiner ist, je größer die Stichprobe ist.

Signifikanz

Ein Resultat eines statistischen Tests gilt als signifikant, wenn die Wahrscheinlichkeit, dass das Resultat rein zufällig hätte zustande kommen können, klein ist. *Wie* klein die Wahrscheinlichkeit sein soll, wird mit dem *p*-Wert ausgedrückt. In medizinischen Studien sind *p*-Werte von 5 Prozent üblich. Das bedeutet: Wenn 20 Resultate nach Definition signifikant sind, ist zu erwarten, dass eines davon ein Zufallsbefund ist.

Surrogatmarker

Wenn sich eine Größe nicht oder nur schlecht messen lässt, kann man stattdessen eine Ersatzgröße – den Surrogatmarker – messen, sofern zwischen dieser und jener ein statistisch signifikanter Zusammenhang besteht: Man will zum Beispiel Angst messen und misst stattdessen die Pulsfrequenz, oder man misst anstelle eines Krankheitserregers die Abwehrstoffe, die der Körper gegen diesen bildet.

Vertrauensintervall (Konfidenzintervall)

Bei Werten, die nicht präzise gemessen, sondern aufgrund von Hochrechnungen oder Modellierungen ermittelt werden, gibt das Vertrauensintervall an, in welchem Bereich ein Wert mit hoher (beispielsweise 95-prozentiger) Wahrscheinlichkeit liegt. So heißt es beispielsweise bei Abstimmungen nach Hochrechnungen, der Ja-Anteil liege bei 53 Prozent, plus-/minus 3 *Prozentpunkte*. Das bedeutet: Mit hoher Wahrscheinlichkeit wird der Ja-Anteil am Ende zwischen 50 und 56 Prozent liegen. Je größer die Datenbasis, desto kleiner ist das Vertrauensintervall.