



Statistik-Tutorial ^[1]

[Explorable.com](#) ^[2] 11K reads

Dieses Statistik-Tutorial ist eine Anleitung, die Ihnen helfen soll, die zentralen Begriffe der Statistik sowie deren Funktion in wissenschaftlicher Methodik und Forschung zu verstehen.

Wissenschaftler verwenden häufig Statistiken zur Analyse ^[3] ihrer Ergebnisse. ^[3] Warum verwenden Forscher Statistiken? ^[4] Die Statistik kann dabei helfen, ein Phänomen durch die Bestätigung oder Widerlegung einer Hypothese zu verstehen. Sie spielt für die meisten wissenschaftlichen Theorien eine wesentliche Rolle bei der Wissenserschließung.

Sie brauchen allerdings kein Wissenschaftler sein; für jeden, der mehr darüber lernen will, wie Forscher Statistiken für ihre Arbeit nutzen, kann dieses Statistik-Tutorial interessant sein.

Was ist Statistik? ^[5]

Forschungsdaten

In diesem Abschnitt des Statistik-Tutorials wird erklärt, wie Daten erhoben und verwendet werden.

Die Ergebnisse einer wissenschaftlichen Untersuchung enthalten oft viel mehr Daten oder Informationen als der Forscher benötigt. Dieses Datenmaterial wird Primärdaten oder Rohdaten genannt.

Um die Daten sinnvoll analysieren zu können, werden die Primärdaten verarbeitet ^[6] zu Sekundärdaten ^[7]. Es gibt viele Methoden, um Daten zu verarbeiten, aber im Grunde ordnet und fasst der Wissenschaftler die Primärdaten zu einem sinnvollen Datenblock zusammen. Jede Art geordneter Informationen kann als Datenmenge ^[8] bezeichnet werden.

Anschließend wenden Forscher unterschiedliche statistische Verfahren an, um die Daten besser (und genauer) verstehen zu können. Je nach Forschungsprojekt lassen sich die statistischen Daten zu deskriptiven ^[9] Zwecken oder für eine explorative Forschung verwenden.

Primärdaten haben den Vorteil, dass man zur Grundlage zurückgehen und die Dinge nochmals überprüfen kann, wenn der Verdacht besteht, dass etwas anders ist, als man dachte. Das passiert in der Regel, nachdem die Bedeutung der Ergebnisse analysiert wurde.

Die Primärdaten können Ideen für neue Hypothesen liefern, da man einen besseren Blick auf

den Sachverhalt hat. Sie können auch die Variablen kontrollieren, die Ergebnis [10] beeinflussen könnten (z. B. Drittvariablen [11]). In der Statistik ist ein Parameter [12] jede numerische Größe, die eine bestimmte Grundgesamtheit oder einen bestimmten Aspekt derselben charakterisiert.

Zentrale Tendenz und Normalverteilung

Dieser Teil des Statistik-Tutorials soll Ihnen helfen, die Begriffe Verteilung, zentrale Tendenz und deren Zusammenhang mit Datenmengen [8] zu verstehen.

Viele Daten aus der realen Welt sind normal verteilt [13], das heißt, sie folgen einer Häufigkeitskurve oder einer Häufigkeitsverteilung [14], bei welcher die häufigsten Punkte in der Nähe der Mitte liegen. Viele Experimente beruhen auf Normalverteilungsannahmen [15]. Dies ist ein Grund, warum in der statistischen Forschung sehr oft die zentrale Tendenz gemessen [16] wird, etwa anhand des Mittels [17] (arithmetisches Mittel [18] oder geometrisches Mittel [19]), Medians [20] (Zentralwert) oder Modus [21] (Modalwert).

Die zentrale Tendenz kann einen recht guten Überblick über die Art der Daten bieten (Mittel, Median und Modus zeigen den "mittleren Wert"), insbesondere dann, wenn sie mit Messungen der Datenverteilung kombiniert wird. Wissenschaftler berechnen normalerweise die Standardabweichung [22], um zu messen, wie die Daten verteilt sind.

Es gibt allerdings verschiedene Methoden, um die Datenverteilung zu messen: Varianz [23], Standardabweichung [24], Standardfehler des Mittelwertes [25], Standardfehler der Schätzung oder "Spannweite [26]" (d. h. die Extremwerte der Daten).

Um die Normalverteilung grafisch darzustellen, verwendet man normalerweise das arithmetische Mittel [18] einer Probe, die groß genug ist [27], und man berechnet die Standardabweichung.

Allerdings liegt bei der Verteilung der Stichproben [28] in der Regel keine Normalverteilung vor, wenn die Verteilung schief (assymmetrisch) ist oder Ausreißer [29] hat (meist seltene Ergebnisse oder Messfehler), die die Daten durcheinander bringen. Ein Beispiel einer Verteilung, die nicht normal verteilt ist, ist die F-Verteilung [30], die nach rechts verzerrt ist.

Forscher prüfen daher mit Hilfe von Spannweite, Median und Modus doppelt, ob ihre Ergebnisse normal verteilt sind. Liegt keine Normalverteilung vor, beeinflusst dies die Wahl des statistischen Tests bzw. der Methode für die Analyse.

Andere Instrumente

- Quartil [31]
- Trimean [32]

Hypothesenprüfung - Statistik-Tutorial

Wie können wir herausfinden, ob eine Hypothese richtig ist oder falsch ist?

Warum verwendet man statistische Methoden zur Hypothesenprüfung? [4]

Um Statistiken korrekt für Forschungszwecke anzuwenden, ist es nicht damit getan, statistische Formeln zu verwenden oder sich in Statistikprogramme einzuarbeiten.

Folgende Punkte bilden die Grundlage für ein erfolgreiches Arbeiten mit Statistiken:

1. Erlernen statistischer Grundlagen [33]
2. Verständnis der Beziehung zwischen Wahrscheinlichkeit und Statistik [34]
3. Verständnis der beiden großen Zweige der Statistik [35]: deskriptive Statistik [9] und induktive Statistik [36].
4. Verständnis, wie Statistiken sich zur wissenschaftlichen Methodik verhalten. [4]

Bei der Statistik zu wissenschaftlichen Zwecken geht es nicht nur um Formeln und Berechnungen. (Das fehlende Verständnis der zentralen Begriffe der Statistik hat schon zu vielen falschen Schlüssen geführt)

Statistische Inferenzen helfen uns, aus Stichproben einer Grundgesamtheit Schlüsse zu ziehen [10].

Ein wichtiger Schritt bei der Durchführung von Experimenten [37] ist das Gegenprüfen von Hypothesen [38]. Daher bildet es ein wichtiges Element dieses Statistik-Tutorials.

Das Prüfen von Hypothesen findet anhand der Formulierung einer Alternativhypothese [39] statt, die gegen die Nullhypothese [40], also die herrschende Meinung, geprüft wird. Hypothesen werden statistisch gegeneinander geprüft. [41]

Der Forscher kann ein Konfidenzintervall [42] herausarbeiten, wodurch die Grenzen definiert werden, ab wann ein Ergebnis die Nullhypothese oder die Alternativhypothese [39] stützt.

Das bedeutet, dass nicht alle Abweichungen zwischen der Versuchsgruppe und der Kontrollgruppe [43] als Stützung der Alternativhypothese akzeptiert werden können - das Ergebnis muss für den Forscher statistisch signifikant [44] abweichen, um die Alternativhypothese akzeptieren zu können. Dies geschieht mit Hilfe eines Signifikanztests [45] (weiterer Artikel [46]).

Allerdings ist Vorsicht geboten, denn durch "data dredging [47]", also das Sammeln von Daten ohne spätere Prüfung der Hypothese, kann man in einem kontrollierten Experiment leicht auf eine Ursache-Wirkungs-Beziehung [48] schließen, [48] obwohl kein Wahrheitsbezug [49] besteht.

Je nach Hypothese sind einseitige oder zweiseitige Tests nötig.

Manchmal wird die Kontrollgruppe durch experimentelle Wahrscheinlichkeit [50] ersetzt - wenn das Forschungsthema aus ethischer Sicht problematisch [51], aus wirtschaftlicher Sicht zu kostspielig oder zu zeitaufwendig ist, wird oft statt eines echten Experiments [52] ein quasiexperimenteller Ansatz [53] verfolgt.

Oft kommt es zu einem Publikationsbias, wenn der Forscher die Alternativhypothese für richtig hält, anstatt ein "Null-Ergebnis" zu erhalten und die Schlussfolgerung zu ziehen, dass die Nullhypothese die beste Erklärung liefert.

Bei richtiger Anwendung ermöglichen Statistiken es, Ursache-Wirkungs-Beziehungen zwischen Variablen [54] zu verstehen.

Sie ermöglichen auch das Erkennen von Drittvariablen, wenngleich Statistiken auch dafür benutzt werden können, um Drittvariablen [11] zu manipulieren und zu vertuschen, wenn derjenige, der die Zahlen vorlegt, mit seinen Ergebnissen keine ehrlichen Absichten verfolgt (oder nicht über ausreichende Kenntnisse verfügt).

Der Missbrauch von Statistiken ist ein weit verbreitetes Phänomen und wird wahrscheinlich so lange fortauern, wie Menschen die Absicht haben, andere Menschen zu beeinflussen. Die ordnungsgemäße statistische Auswertung [55] von experimentellen Daten kann dabei helfen, die Verwendung ethisch nicht vertretbarer statistischer Verfahren zu vermeiden. Die Philosophie der Statistik [56] beschäftigt sich mit der Rechtfertigung der ordnungsgemäßen Verwendung statistischer Verfahren, der Sicherung der statistischen Validität [57] sowie der Gewährleistung der ethischen Korrektheit von Statistiken [58].

Hier ein weiteres gutes Statistik-Tutorial mit Erklärungen zu Statistiken und wissenschaftlicher Methodik.

Reliabilität und experimentelle Fehler

Für statistische Tests werden aus Strichproben erhobene Daten verwendet. Diese Ergebnisse werden dann auf die Grundgesamtheit generalisiert. [59] Wie können wir sicher sein, dass sie sich korrekt in der Ergebnis [10] widerspiegeln?

Im Gegensatz zur weitläufigen Ansicht sind Fehler in der Forschung [60] ein wesentlicher Bestandteil des Signifikanztests [46]. Ironischerweise ist gerade die Möglichkeit, dass Forschungsfehler auftreten, charakteristisch für die wissenschaftliche Forschung. Wenn eine Hypothese nicht falsifiziert [61] werden kann (z. B. wenn die Hypothese eine zirkuläre Logik aufweist), ist sie nicht überprüfbar [62] und damit per Definition nicht wissenschaftlich.

Wenn eine Hypothese überprüfbar ist, ist es möglich, dass sie widerlegt wird. In der Statistik eröffnet dies die Möglichkeit, dass ein Ergebnis aufgrund von zufälligen Fehlern oder anderen Problemen in der Untersuchung experimentelle Fehler [63] aufweist. Experimentelle Fehler lassen sich kategorisieren in Fehler 1. Art und Fehler 2. Art [60]. ROC-Kurven [64] werden verwendet, um die Sensitivität zwischen echten Positiven und falschen Positiven zu berechnen.

Eine Teststärke-Analyse [65] eines statistischen Tests kann bestimmen, wie viele Stichproben ein Test benötigt, um einen akzeptablen p-Wert [66] zu haben, damit eine falsche Nullhypothese [40] abgelehnt werden kann.

Die Fehlerspanne [67] hängt mit dem Konfidenzintervall [42] und der Beziehung zwischen statistischer Signifikanz, Stichprobenumfang und den erwarteten Ergebnisse [68] zusammen. Die Effektstärke [69] schätzt die Stärke der Beziehung zwischen zwei Variablen innerhalb einer Grundgesamtheit ein. Sie ist hilfreich bei der Bestimmung des erforderlichen Stichprobenumfangs, [27] um die Ergebnisse auf die Grundgesamtheit generalisieren [59] zu können.

Wenn die Untersuchungsergebnisse generalisierbar waren oder auf ein zufälliges "Ausreißer-Experiment" zurückgehen, ist es außerdem wichtig, die Replikation [70] einer durchgeführten Untersuchung zu verstehen. Die Replikation kann bei der Identifizierung sowohl von zufälligen Fehlern [71] als auch von systematischen Fehlern [72] (Testvalidität [73]) helfen.

Cronbachs Alpha [74] wird verwendet, um die interne Konsistenz oder Reliabilität [75] eines Testergebnisses zu messen.

Die Replikation des Experiments bzw. der Untersuchung gewährleistet die statistische Zuverlässigkeit der Ergebnisse [76].

Wenn die Ergebnisse Ausreißer haben, tritt oft eine Regression zum Mittelwert [77] auf, was die Ergebnisse zwischen Versuchs- und Kontrollgruppe nicht statistisch voneinander abweichen lässt.

Statistische Tests

Hier stellen wir Ihnen einige häufig verwendete statistische Tests und Verfahren vor, die von Forschern genutzt werden.

Beziehung zwischen Variablen

Die Beziehung zwischen Variablen [78] ist von großer Bedeutung für Wissenschaftler. Sie hilft ihnen, das Wesen ihres Untersuchungsgegenstandes zu verstehen. Eine lineare Beziehung [79] liegt vor, wenn zwei Variablen proportional zueinander variieren. Das heißt, wenn eine Variable steigt, steigt auch die andere Variable im gleichen Verhältnis. Eine nichtlineare Beziehung [80] liegt vor, wenn Variablen nicht proportional zueinander variieren. Korrelation [81] ist eine Möglichkeit, eine Beziehung zwischen zwei Datenmengen oder zwei Variablen auszudrücken.

Messskalen [82] werden verwendet, um Variablen zu klassifizieren, zu kategorisieren und (falls zutreffend) zu quantifizieren.

Der Pearson-Korrelationskoeffizient [83] (oder Pearson-Produkt-Moment-Korrelation) drückt nur die lineare Beziehung zwischen zwei Variablen aus. Spearman's Rangkorrelationskoeffizient [84] wird vor allem für lineare Beziehungen verwendet, wenn mit ordinalen Variablen gearbeitet wird. Kendalls Tau kann für die Messung nichtlinearer Beziehungen verwendet werden.

Partielle Korrelation [85] (und Multiple Korrelation) kann bei der Steuerung einer Drittvariable [11] verwendet werden.

Prognosen

Prognosen dienen dazu, Ursachen zu verstehen. Korrelation bedeutet nicht zwangsläufig Kausation [86]. Anhand der linearen Regression wird oft die manipulierte Variable [87] gemessen.

Was ist der Unterschied zwischen Korrelation und linearer Regression [88]? Grundsätzlich wird bei einer Korrelationsstudie [89] die Stärke zwischen den Variablen [90] untersucht, während es bei der linearen Regression um die am besten passende Linie in einem Diagramm geht.

Regressionsanalyse und andere Modellierungsinstrumente

- Lineare Regression [91]
- Multiple Regression [92]
- Eine Pfadanalyse ist eine Erweiterung des Regressionsmodells
- Eine Faktorenanalyse [93] versucht, zugrunde liegende Faktoren aufzudecken.
- In der Metaanalyse [94] wird häufig die Effektstärke [69] verwendet

Die Bayessche Wahrscheinlichkeit [95] ist eine Möglichkeit, die Wahrscheinlichkeit künftiger Ereignisse auf interaktive Weise vorherzusagen, anstatt mit Messungen zu beginnen und anschließend Ergebnisse bzw. Prognosen zu erhalten.

Statistische Hypothesenprüfung

Der t-Test [96] ist ein Test, der aufzeigen kann, ob die Nullhypothese [40] richtig oder falsch ist. In der Forschung wird er häufig verwendet, um Unterschiede zwischen zwei Gruppen (z. B. zwischen einer Kontrollgruppe [43] und einer Versuchsgruppe) zu testen.

Der t-Test nimmt an, dass die Daten mehr oder weniger normal verteilt sind und daß die Varianz gleich ist (dies lässt sich durch den F-Test [97] testen).

t-Test [98]:

- Unabhängiger Einstichproben-t-Test [99]
- Unabhängiger Zweistichproben-t-Test [100]
- Abhängiger t-Test für zwei verbundene Stichproben [101]

Der Wilcoxon-Vorzeichen-Rang-Test [102] kann für nichtparametrische Daten verwendet werden.

Ein z-Test [103] ist ähnlich wie ein t-Test, wird aber gewöhnlich nicht bei Stichprobenumfängen unter 30 verwendet.

Ein Chi-Quadrat-Test [104] kann verwendet werden, wenn die Daten eher qualitativer als quantitativer Natur sind.

Vergleich zwischen mehr als zwei Gruppen

Eine Varianzanalyse [105] oder ANOVA wird durchgeführt, wenn es wünschenswert ist, zu testen, ob eher eine unterschiedliche Variabilität zwischen Gruppen vorliegt als andere Mittelwerte. Die Varianzanalyse kann auch bei mehr als zwei Gruppen angewandt werden. Die F-Verteilung [30] kann für die Berechnung der p-Werte [66] für die ANOVA benutzt werden.

Varianzanalyse

- Einfaktorielle ANOVA [106]
- Zweifaktorielle ANOVA [107]

- Mehrfaktorielle ANOVA mit mehr als zwei Faktoren [108]
- Messwiederholungen und ANOVA [109]

Nichtparametrische Statistik

Einige gängige Methode, bei der nichtparametrische Statistik [110] angewandt wird:

- Cohens Kappa [111]
- Mann-Whitney-U-Test [112]
- Spearman's Rangkorrelationskoeffizient [84]

Weitere wichtige Begriffe in der Statistik

- Diskrete Variablen [113]

Quell-URL: <https://staging.explorable.com/de/statistik-tutorial>

Links

- [1] <https://staging.explorable.com/de/statistik-tutorial>
- [2] <https://staging.explorable.com/en>
- [3] <https://staging.explorable.com/statistical-analysis>
- [4] <https://staging.explorable.com/why-statistics-matter>
- [5] <https://staging.explorable.com/what-is-statistics>
- [6] <https://staging.explorable.com/raw-data-processing>
- [7] <https://staging.explorable.com/data-output>
- [8] <https://staging.explorable.com/statistical-data-sets>
- [9] <https://staging.explorable.com/descriptive-statistics>
- [10] <https://staging.explorable.com/drawing-conclusions>
- [11] <https://staging.explorable.com/confounding-variables>
- [12] <https://staging.explorable.com/parameters-and-statistics>
- [13] <https://staging.explorable.com/normal-probability-distribution>
- [14] <https://staging.explorable.com/frequency-distribution>
- [15] <https://staging.explorable.com/normal-distribution-assumptions>
- [16] <https://staging.explorable.com/measures-of-central-tendency>
- [17] <https://staging.explorable.com/statistical-mean>
- [18] <https://staging.explorable.com/arithmetical-mean>
- [19] <https://staging.explorable.com/geometric-mean>
- [20] <https://staging.explorable.com/calculate-median>
- [21] <https://staging.explorable.com/statistical-mode>
- [22] <https://staging.explorable.com/calculate-standard-deviation>
- [23] <https://staging.explorable.com/statistical-variance>
- [24] <https://staging.explorable.com/measurement-of-uncertainty-standard-deviation>
- [25] <https://staging.explorable.com/standard-error-of-the-mean>
- [26] <https://staging.explorable.com/range-in-statistics>
- [27] <https://staging.explorable.com/sample-size>
- [28] <https://staging.explorable.com/sampling-distribution>
- [29] <https://staging.explorable.com/statistical-outliers>
- [30] <https://staging.explorable.com/f-distribution>
- [31] <https://staging.explorable.com/quartile>
- [32] <https://staging.explorable.com/trimean>
- [33] <https://staging.explorable.com/learn-statistics>
- [34] <https://staging.explorable.com/probability-and-statistics>
- [35] <https://staging.explorable.com/branches-of-statistics>
- [36] <https://staging.explorable.com/inferential-statistics>
- [37] <https://staging.explorable.com/conducting-an-experiment>

[38] <https://staging.explorable.com/hypothesis-testing>
[39] <https://staging.explorable.com/research-hypothesis>
[40] <https://staging.explorable.com/null-hypothesis>
[41] <https://staging.explorable.com/statistical-hypothesis-testing>
[42] <https://staging.explorable.com/statistics-confidence-interval>
[43] <https://staging.explorable.com/scientific-control-group>
[44] <https://staging.explorable.com/statistically-significant-results>
[45] <https://staging.explorable.com/significance-test-2>
[46] <https://staging.explorable.com/significance-test>
[47] <https://staging.explorable.com/data-dredging>
[48] <https://staging.explorable.com/cause-and-effect>
[49] <https://staging.explorable.com/truth-and-theory>
[50] <https://staging.explorable.com/experimental-probability>
[51] <https://staging.explorable.com/de/ethik-in-der-forschung>
[52] <https://staging.explorable.com/true-experimental-design>
[53] <https://staging.explorable.com/quasi-experimental-design>
[54] <https://staging.explorable.com/research-variables>
[55] <https://staging.explorable.com/statistical-treatment-of-data>
[56] <https://staging.explorable.com/philosophy-of-statistics>
[57] <https://staging.explorable.com/statistical-validity>
[58] <https://staging.explorable.com/ethics-in-statistics>
[59] <https://staging.explorable.com/what-is-generalization>
[60] <https://staging.explorable.com/type-i-error>
[61] <https://staging.explorable.com/falsifiability>
[62] <https://staging.explorable.com/testability>
[63] <https://staging.explorable.com/experimental-error>
[64] <https://staging.explorable.com/roc-curve-analysis>
[65] <https://staging.explorable.com/statistical-power-analysis>
[66] <https://staging.explorable.com/p-value>
[67] <https://staging.explorable.com/statistics-margin-of-error>
[68] <https://staging.explorable.com/statistical-significance-sample-size>
[69] <https://staging.explorable.com/effect-size>
[70] <https://staging.explorable.com/replication-study>
[71] <https://staging.explorable.com/random-error>
[72] <https://staging.explorable.com/systematic-error>
[73] <https://staging.explorable.com/test-validity>
[74] <https://staging.explorable.com/cronbachs-alpha>
[75] <https://staging.explorable.com/statistics-reliability>
[76] <https://staging.explorable.com/statistical-reliability>
[77] <https://staging.explorable.com/regression-to-the-mean>
[78] <https://staging.explorable.com/relationship-between-variables>
[79] <https://staging.explorable.com/linear-relationship>
[80] <https://staging.explorable.com/non-linear-relationship>
[81] <https://staging.explorable.com/statistical-correlation>
[82] <https://staging.explorable.com/measurement-scales>
[83] <https://staging.explorable.com/pearson-product-moment-correlation>
[84] <https://staging.explorable.com/spearman-rank-correlation-coefficient>
[85] <https://staging.explorable.com/partial-correlation-analysis>
[86] <https://staging.explorable.com/correlation-and-causation>
[87] <https://staging.explorable.com/independent-variable>
[88] <https://staging.explorable.com/correlation-and-regression>
[89] <https://staging.explorable.com/correlational-study>
[90] <https://staging.explorable.com/variables-and-statistics>
[91] <https://staging.explorable.com/linear-regression-analysis>
[92] <https://staging.explorable.com/multiple-regression-analysis>
[93] <https://staging.explorable.com/factor-analysis>

- [94] <https://staging.explorable.com/meta-analysis>
- [95] <https://staging.explorable.com/bayesian-probability>
- [96] <https://staging.explorable.com/students-t-test>
- [97] <https://staging.explorable.com/f-test>
- [98] <https://staging.explorable.com/students-t-test-2>
- [99] <https://staging.explorable.com/independent-one-sample-t-test>
- [100] <https://staging.explorable.com/independent-two-sample-t-test>
- [101] <https://staging.explorable.com/dependent-t-test-for-paired-samples>
- [102] <https://staging.explorable.com/wilcoxon-signed-rank-test>
- [103] <https://staging.explorable.com/z-test>
- [104] <https://staging.explorable.com/chi-square-test>
- [105] <https://staging.explorable.com/anova>
- [106] <https://staging.explorable.com/one-way-anova>
- [107] <https://staging.explorable.com/two-way-anova>
- [108] <https://staging.explorable.com/factorial-anova>
- [109] <https://staging.explorable.com/repeated-measures-anova>
- [110] <https://staging.explorable.com/nonparametric-statistics>
- [111] <https://staging.explorable.com/cohens-kappa>
- [112] <https://staging.explorable.com/mann-whitney-u-test>
- [113] <https://staging.explorable.com/discrete-variables>