

Kapitel 3: STATISTISCHE METHODEN

Inhalt:

EINLEITUNG	35
BESCHREIBENDE STATISTIK	35
BEWERTENDE STATISTIK.....	46
LITERATUR	50
LINKS	51

Einleitung

"Ich glaube nur die Statistik, die ich selbst gefälscht habe."

Diese Aussage, die fälschlicherweise Winston Churchill zugeschrieben wurde, wird oft als Argument gegen die Statistik zitiert. Sie ließe aber nur den Schluss zu, dass Churchill bewusst falsche Daten und/oder falsche Methoden zuließ. Unredlichkeit und Voreingenommenheit beim Umgang mit Daten können nur die betreffenden Anwender, nicht aber die naturwissenschaftliche Methode abqualifizieren.

Beschreibende Statistik

Statistik liefert Informationen in Form empirischer Zahlen (Statistiken - Umfrageergebnisse). Sie ist eine Zusammenfassung von Methoden, welche uns erlauben, vernünftige Entscheidungen im Falle von Ungewissheit zu treffen. Es sind Verfahren, nach denen empirische Zahlen gewonnen, dargestellt, verarbeitet, analysiert und für Schlussfolgerungen, Prognosen und Entscheidungen verwendet werden (Wahrscheinlichkeitsrechnung). Mit Statistik lassen sich keine Beweise führen, nur Hypothesen bekräftigen. Neben der Unzulänglichkeit des Beobachters sowie der Instrumente kommen in der Biologie noch die natürlichen Schwankungen der lebenden Systeme hinzu, die eine eindeutige Aussage erschweren. Eine mathematisch-statistische Auswertung der Beobachtungen lässt sich daher kaum vermeiden.

Mit der mathematischen Statistik analysiert man also Massenerscheinungen. Dabei zeigt sich oft, dass die Massenerscheinungen gewisse Gesetzmäßigkeiten aufweisen, die sich für Einzelercheinungen nicht formulieren lassen, da sie dort nur als zufällige Unregelmäßigkeit auftreten. Werden z.B. 100 Bohnen einer bestimmten Sorte einzeln gewogen, so streuen die einzelnen Werte zufällig und sind somit nicht vorhersagbar. Das mittlere Gewicht und die Streuung der Werte jedoch sind auch nach dem Auszählen einer zweiten Stichprobe nahezu

identisch. Diese charakteristischen Werte erlauben somit eine Aussage über die Grundgesamtheit aller Bohnen dieser Sorte.

In der Wahrscheinlichkeitsrechnung wird der Mittelwert als Erwartungswert interpretiert. In der Bundesrepublik sterben jährlich 100 Personen aufgrund eines technischen Defekts an einem elektrischen Haushaltsgerät. Glaubhaft, da ein eng definiertes, singuläres Schadensereignis. Auf der Welt sterben jährlich über zwei Millionen Menschen an den Folgen ihres Nikotinkonsums. Dies wird nicht geglaubt, da es sich um ein Multikomponenten Ereignis handelt, von dem jeder eine Ausnahme kennt.

Die biologische Statistik besteht in erster Linie in einer kritischen Bewertung von Stichprobenergebnissen. Messungen biologischer Parameter schwanken nicht um einen wahren Wert, wie etwa eine physikalische Größe im unbelebten System, sondern sie haben eine beträchtliche Streuung, die durch die biologische Variabilität bedingt ist. Das Ziel ist, allgemeine Aussagen über spezielle Merkmale gleicher Individuen zu machen. Die Gesamtheit aller Individuen des zu untersuchenden Materials stellt die Grundgesamtheit dar. Da es meist nicht möglich ist, alle Individuen zu untersuchen, werden Stichproben durchgeführt. Die Auswahl dieser Stichproben muss zufällig erfolgen.

Tabellarische und graphische Darstellung

Üblicherweise werden die n Beobachtungsereignisse (Massen einzelner Bohnen) der Reihe nach aufgelistet (Urliste). Aus dieser Stichprobe vom Umfang n kann man Schlüsse auf die zugehörige Grundgesamtheit ziehen. Wären gleichzeitig mehrere Merkmale gemessen worden, z.B. Masse und Größe, so hätte man eine Stichprobe erhalten, die aus n Zahlenpaaren besteht.

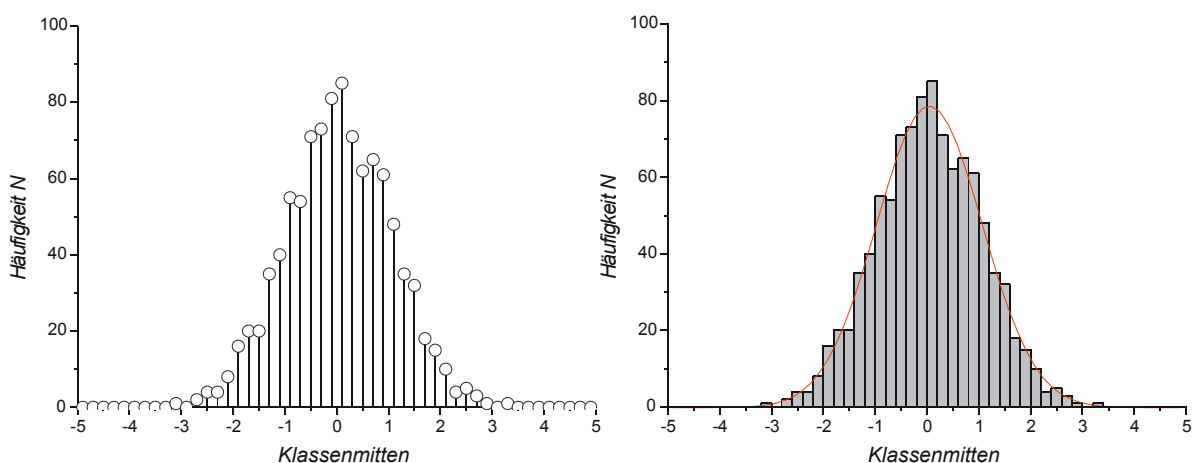


Abb. 19: Häufigkeitsverteilung einer Gaußschen Normalverteilung. links: Stabdiagramm. rechts: Balkendiagramm mit Kurvenzug der theoretischen Funktionsgleichung.

Bei kleinen Stichproben hilft es schon, wenn man die Werte der Größe nach ordnet, um einen Überblick zu bekommen. Besser ist es, zahlenmäßig gleiche Werte zusammenzufassen und sie

graphisch darzustellen. Dabei wird die Anzahl N über dem Messwert x_i aufgetragen. Die Anzahl N heißt die absolute Häufigkeit des betreffenden Wertes in der Stichprobe. Dividiert man die absolute Häufigkeit durch den Stichprobenumfang n , so erhält man die relative Häufigkeit. Die relative Häufigkeit ist somit immer eine Zahl zwischen 0 und 1. Die Auftragung kann als Punkt-, Stab- oder Balkendiagramm (Abb. 19) erfolgen. Eine direkte Verbindung der Punkte untereinander ergibt ein Häufigkeitspolygon (Abb. 19 *links*). Diese Graphiken stellen Häufigkeitsverteilungen oder Histogramme der Stichprobe dar.

Klassenbildung

Kommen in einer Stichprobe sehr viele zahlenmäßig verschiedene Werte vor, so ist die Tabelle oder die Zeichnung der Häufigkeitsverteilung meist noch recht unübersichtlich. Man kann in diesem Fall die Stichprobe weiter vereinfachen, und zwar durch die sog. Gruppierung oder Klassenbildung im Gegensatz zu den oben genannten nicht gruppierten Werten. Dabei geht man von dem Intervall aus, in dem alle Stichprobenwerte liegen. Dieses unterteilt man in Teilintervalle (Abb. 19 *rechts*), die als Klassenintervalle bezeichnet werden. Die Mitten dieser Intervalle heißen Klassenmitten (Abb. 19 *links*). Alle Stichprobenwerte in einem solchen Intervall bilden zusammen jeweils eine Klasse von Werten. Die ursprünglichen Stichprobenwerte treten nicht mehr einzeln in Erscheinung. Man nimmt an, dass alle Werte einer Klasse in der zugehörigen Klassenmitte liegen (Abb. 19 *links*). Je weniger Klassen man bildet, desto mehr Information, die in den ursprünglichen Stichprobenwerten steckt, geht aber verloren. Man sollte so klassifizieren, dass nur unwesentliche Einzelheiten ausgeschieden werden. In der Praxis wählt man meist 10 - 20 Klassen und mehr als 20 höchstens bei sehr umfangreichen Stichproben. Unnötige Komplikationen bei späteren Rechnungen lassen sich vermeiden, wenn man die folgenden Regeln beachtet:

- Die Klassenintervalle wählt man gleich lang.
- Die Klassenmitten sollen möglichst einfachen Zahlen, d.h. Zahlen mit möglichst wenigen Ziffern, entsprechen.
- Ein Wert, der auf einen Intervallendpunkt fällt, wird je zur Hälfte in jedem der beiden angrenzenden Klassenintervalle mitgezählt.

Oftmals kann man die genannten Endpunkte ohne Mühe so wählen, dass sie nicht mit Stichprobenwerten zusammenfallen. Durch die beschränkten Messgenauigkeiten von Messgeräten ergeben sich die besten Klassengrenzen oft von selbst, da das Messgerät selbst die Klasseneinteilung vorgibt.

Summenhäufigkeitsfunktion einer Stichprobe

Die Häufigkeitsverteilung einer Stichprobe gibt die Häufigkeiten an, mit der die einzelnen Zahlenwerte in der Stichprobe vorkommen (Abb. 19 *rechts*). Wenn es 30 Bohnen mit dem Gewicht 3,2 g gibt, man jedoch wissen möchte, wie viel Bohnen 3,2 g oder weniger wiegen, so

erhält man die Antwort durch aufsummieren der einzelnen Häufigkeiten bis $x = 3,2$ g. Man erhält auf diese Weise die Summenhäufigkeitsfunktion oder Verteilungsfunktion einer Stichprobe. Die Summenhäufigkeitsfunktion stellt das Integral der Häufigkeitsfunktion dar (Abb. 20). Aus unstetigen Häufigkeitsfunktionen erhält man Treppenfunktionen, aus stetigen Häufigkeitsverteilungen Sigmoide. Jede der beiden genannten Funktionen bestimmt die Stichprobe in allen Einzelheiten. Die Summenhäufigkeitsfunktion (Abb. 20) ist weniger anschaulich als die Häufigkeitsfunktion (Abb. 19 rechts).

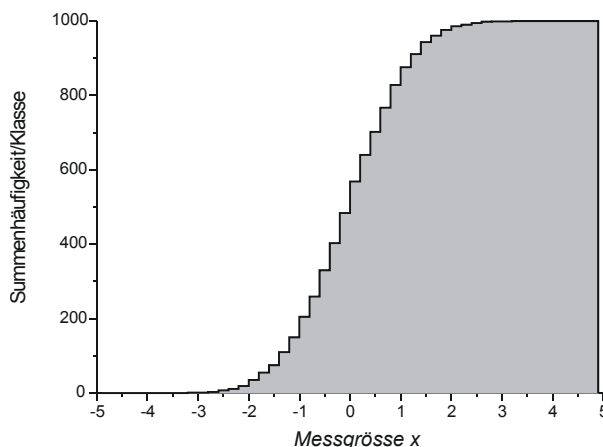


Abb. 20: Summenhäufigkeit (Treppenfunktion und Sigmoide) der Häufigkeitsverteilung aus Abb. 19 (rechts)

Mittelwert, Varianz, Standardabweichung, Standardfehler

Neben der Häufigkeits- bzw. Summenhäufigkeitsfunktion kann man eine Grundgesamtheit oder eine Stichprobe auch durch Maßzahlen charakterisieren. Die beiden in der Praxis wichtigsten Maßzahlen sind der Mittelwert, der die durchschnittliche Größe der Grundgesamtheit N oder der Stichprobe n kennzeichnet, und eine Angabe über die Streuung der Werte. Im weiteren wird die Annahme gemacht, dass die Messwerte eine Normalverteilung nach Gauß ergeben (s.u.). Eine genügend große Stichprobe wird vorausgesetzt.

Der arithmetische Mittelwert ist definiert als:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

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

μ = Mittelwert der Grundgesamtheit, \bar{x} = Mittelwert der Stichprobe

Dieser allein reicht jedoch nicht aus, um z.B. eine Stichprobe zu beschreiben, wie folgendes Beispiel zeigt:

Stichprobe 1:	1; 2; 4; 5	$\bar{x} = 3$
Stichprobe 2:	2,7; 3,0; 3,1; 3,2	$\bar{x} = 3$

Beide Stichproben haben den Mittelwert $\bar{x} = 3$. Sie unterscheiden sich aber trotzdem wesentlich voneinander, denn die Werte der ersten Stichprobe liegen viel weiter auseinander (und auch weiter vom Mittelwert entfernt) als die Werte der zweiten Stichprobe. Um diesen Unterschied zu erfassen, braucht man noch eine weitere Maßzahl. Geeignet ist hierzu offenbar eine Zahl, die die Abweichung der Stichprobenwerte vom Mittelwert misst. Man könnte die Spannweite der Stichprobe, d.h. die Differenz zwischen dem größten (Maximum) und kleinsten (Minimum) Stichprobenwert ermitteln (Abb. 19 links: Minimum = 0, Maximum = 100). Es wird jedoch gefordert, dass ähnlich wie beim Mittelwert jeder Einzelwert in gewisser Weise mitberücksichtigt wird. Die wohl am nächsten liegende Möglichkeit, die Summe der Einzelabweichungen $x_i - \bar{x}$ scheidet allerdings aus, da die Summe aus negativen und positiven Gliedern besteht und diese somit immer Null ist. Dies könnte vermieden werden, wenn man die Absolutbeträge der Einzelabweichungen bilden würde. Aus mathematischen Ableitungen hat sich jedoch die Bildung der Quadrate der Einzelabweichungen als günstiger erwiesen. Diese werden auch als die kleinsten Gaußschen Fehlerquadrate (engl. least square) bezeichnet. Die Maßzahl, die man auf diesem Weg erhält heißt Varianz oder Streuung (engl. variance). Sie berechnet sich für die Grundgesamtheit nach

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

und für die Stichprobe nach

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

Aus der Wahrscheinlichkeitstheorie lässt sich die unterschiedliche Berechnung der Varianz für Grundgesamtheit und Stichprobe ableiten. Man muss im Allgemeinen bei der Berechnung nur wissen, ob es sich um eine Grundgesamtheit oder eine Stichprobe handelt. $(n - 1)$ bezeichnet man als die Anzahl der Freiheitsgrade, sie ergeben sich aus der Anzahl unabhängiger Einzelwerte. Die nichtnegative Quadratwurzel der Varianz heißt Standardabweichung (engl. standard deviation, S.D.)

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$$

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

Bei Taschenrechnern mit statistischen Programmen muss der Unterschied bei der Standardabweichung zwischen Grundgesamtheit und Stichprobe durch Auswahl der entsprechenden Funktionstaste beachtet werden. Die Größen Varianz und Standardabweichung sind mit demselben Formelbuchstaben belegt, da beide in der Praxis gleichwertig verwendet werden. Die Varianz hat den Vorteil, dass man sich nicht mit Quadratwurzeln herumärgern muss. Die Standardabweichung hat den Vorteil, dass sie dieselbe Dimension der Größeneinheit (z.B. cm oder kg) wie der Mittelwert besitzt. Für die obigen Beispiele ergeben sich somit:

$$\text{Stichprobe 1: } \quad \bar{x} = 3 \quad s^2 = 3,3 \quad s = 1,8$$

$$\text{Stichprobe 2: } \quad \bar{x} = 3 \quad s^2 = 0,05 \quad s = 0,22$$

Die Streuung der zweiten Stichprobe ist also wesentlich kleiner. Durch Angabe von Mittelwert und Varianz bzw. Standardabweichung sind Stichproben meist ausreichend beschrieben. Die Berechnung der Standardabweichung (bzw. Varianz) nach den Definitionsformeln ist ungünstig. Durch die Differenzbildung $(x_i - \bar{x})$ von den relativ großen Zahlen entstehen sehr kleine Differenzen, die dann auch noch quadriert werden müssen. Durch Rundungsfehler entstehen Genauigkeitsverluste, die beim elektronischen Rechnen nicht einmal bemerkt werden. Es gibt deshalb Berechnungsformeln für die Praxis. Bei ihnen werden die Differenzbildungen vermieden. Für die Standardabweichung einer Stichprobe ergibt sich somit

$$s = \sqrt{\frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right]}$$

Eine ebenfalls verwendete Formel ist:

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

Bei der Bestimmung von Stichproben möchte man gerne wissen, mit welcher Wahrscheinlichkeit sich die bei einer Stichprobe gefundenen Größen auf die Grundgesamtheit ausweiten lassen. Im Beispiel der Bohnen möchte man also eine Aussage über alle Bohnen einer Sorte machen. Die für eine Stichprobe ermittelten Werte (Mittelwert, Varianz, Standardabweichung) sind also nur Schätzwerte für die Grundgesamtheit. Man möchte z.B. wissen, wie weit der Stichprobenmittelwert \bar{x} vom Mittelwert der Grundgesamtheit μ abweicht. Diese Abweichung bezeichnet man als Standardfehler (= Fehler des Mittelwertes = Standardabweichung des Mittelwertes; engl. standard error of the mean, S.E.M.). Wenn keine extremen Abweichungen der Stichprobenwerte x_i von der Normalverteilung um den Stichprobenmittelwert \bar{x} vorliegen, darf man annehmen, dass sich auch die Mittelwerte annähernd gleich großer Stichproben gleichmäßig um den Grundgesamtheitsmittelwert μ verteilen. Die Abweichung kann durch den Standardfehler abgeschätzt werden. Er berechnet sich aus der Standardabweichung s .

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

Das zusätzliche n in der Formel für den Standardfehler s (im Gegensatz zu Varianz und Standardabweichung) liefert eine Angabe über die Größe der Stichprobe. Je größer eine Stichprobe ist, desto genauer wird die Schätzung für die Grundgesamtheit. Der Standardfehler verkleinert sich dabei (n steht im Nenner), geht somit gegen μ . (Die Genauigkeit ist dem Geduldsfaden des Experimentators direkt proportional). Der Standardfehler wird oft zusammen mit dem Stichprobenmittelwert zur Charakterisierung einer Stichprobe bezüglich der Grundgesamtheit angegeben:

$$\bar{x} \pm s_{\bar{x}} \quad (\text{z.B. } 5\text{g} \pm 0,6\text{g})$$

Gewichteter Mittelwert, Zentralwert, Häufigster Wert

Teilt man die Messwerte in k Klassen ein (Gruppierung, Abb. 19 links), so lässt sich der arithmetische Mittelwert auch als gewichteter Mittelwert $\bar{x}_{\text{gew.}}$ (gewogenes Mittel) berechnen.

Dazu wird jede mittlere Klassengröße x_i mit ihrer Klassenhäufigkeit a_i multipliziert.

$$\bar{x}_{\text{gew.}} = \frac{a_1 x_1 + a_2 x_2 + \dots + a_k x_k}{n} \quad \left| \quad n = \sum_{i=1}^k a_i \right.$$

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

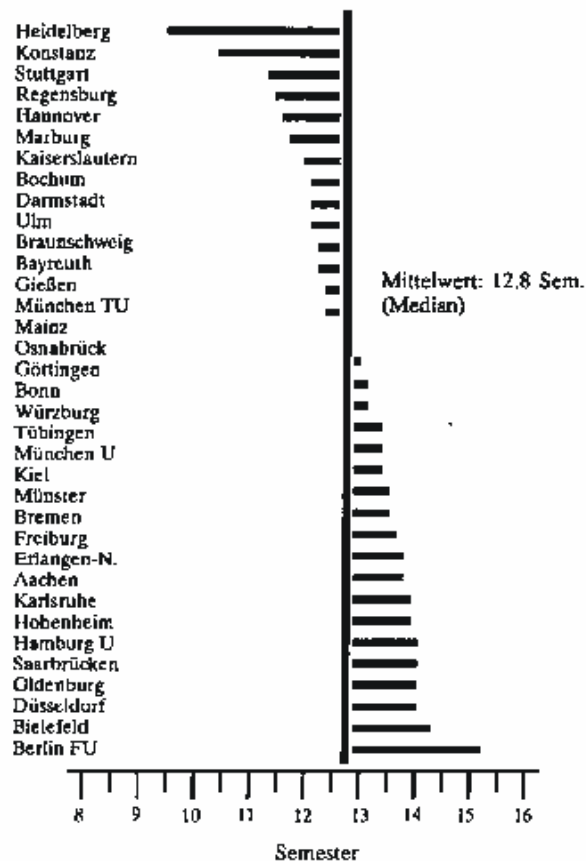


Abb. 21: Beispiel für einen Medianwert: Durchschnittliche Studienzeiten im Fach Biologie, Diplom 1988 (Quelle: Wissenschaftsrat). Die Heidelberger hatten "vergessen" die Zeit für die Anfertigung einer Diplomarbeit mit einzurechnen (siehe [Churchill](#)).

Der Zentralwert oder Medianwert stellt ebenfalls einen charakteristischen Wert einer Häufigkeitsverteilung dar. Er wird für bestimmte statistische Verfahren benötigt. Er teilt die Häufigkeitsverteilung flächengleich auf, so dass sich links und rechts vom Zentralwert genau gleich viele Ereignisse befinden. Studienzeiten eines Jahrgangs sind meist nicht normalverteilt ("Langzeitstudenten"), daher wird der Wert angegeben, bis zu dem die erste Hälfte ihren Abschluss geschafft hat.

Der häufigste Wert oder Modalwert stellt, wie sein Name schon sagt, den Wert mit der größten Häufigkeit dar. Er ist also der Gipfel ("Peak") in einer Häufigkeitsverteilung. In einer Normalverteilung (Gaußverteilung, Abb. 22) sind infolge der Symmetrie der Verteilung arithmetischer Mittelwert, Medianwert und Modalwert identisch.

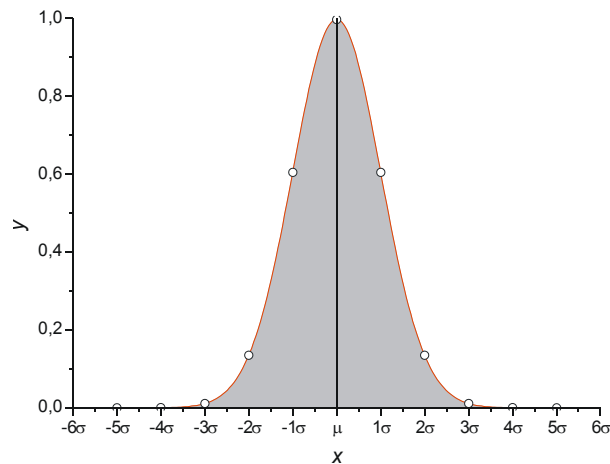


Abb. 22: Normalverteilung (vergl. **Abb. 19**) mit Angabe des Mittelwertes und der Standardabweichung s

Gaußverteilung, Poissonverteilung

Als Beispiele für theoretische Häufigkeitsverteilungen sollen die Normalverteilung (Gaußsche Glockenkurve) und die Poissonverteilung als Beispiel für eine schiefe Verteilung besprochen werden. Viele Messwerte aus Experimenten sind nach diesen beiden theoretischen Mustern verteilt. Die Normalverteilung (Abb. 22) wurde von Gauß im Zusammenhang mit der Theorie der Messfehler eingeführt. Aus verschiedenen Gründen ist sie die wichtigste stetige Verteilung:

1. Viele Zufallsvariablen, die bei Experimenten und Beobachtungen auftreten, sind normalverteilt.
2. Andere Zufallsvariablen sind annähernd normalverteilt. In vielen Fällen führt dann die Annahme einer Normalverteilung zu sinnvollen und praktisch brauchbaren Ergebnissen.
3. Gewisse, nicht normalverteilte Variablen lassen sich auf einfache Weise so transformieren, dass die sich daraus ergebende Variable normalverteilt ist.



Abb. 23: Der Mathematiker Carl Friedrich Gauss und seine Normalverteilung auf dem ehemaligen 10 Markschein.

Die Funktionsgleichung der Gaußverteilung lautet:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2}$$

μ = Mittelwert σ = Standardabweichung

Die Standardabweichungen $\pm \sigma$ sind definitionsgemäß die Wendepunkte der Glockenkurve, projiziert auf die x-Achse (Abb. 22). Im Bereich zwischen $\pm \sigma$ liegen 68 % aller beobachteten Werte. Im Bereich $\pm 2 \sigma$ liegen 95,5 % und im Bereich $\pm 3 \sigma$ so gut wie alle Werte, nämlich 99,7%.

Die Standardabweichung σ ist ein Maß für die Streuung der Werte um den Mittelwert μ . Je größer die Standardabweichung, desto weiter streuen also die Werte um den Mittelwert μ (Abb. 24). Die Summenhäufigkeitsfunktion oder Verteilungsfunktion (Integral der Glockenkurve) ergibt eine Sigmoidale (Abb. 20).

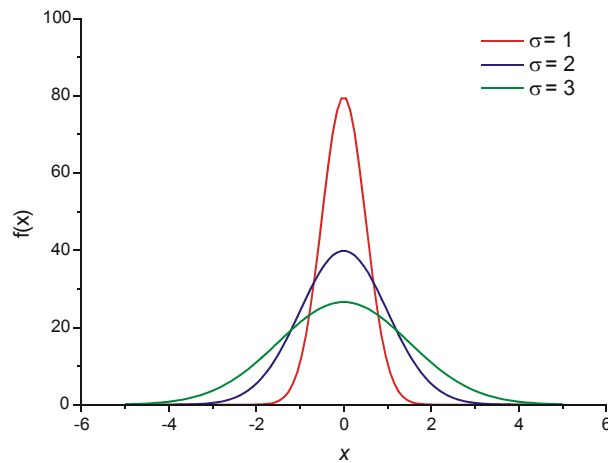


Abb. 24: Normalverteilungen mit gleich großen Mittelwerten und verschieden großen Standardabweichungen.

Ein Beispiel für eine schiefe diskrete Verteilung ist die Poissonverteilung (Abb. 25) mit der Funktionsgleichung:

$$f(x) = \frac{\mu^x}{x!} e^{-\mu}$$

Für Mittelwerte nahe Null nähert sich die Poissonverteilung einer abnehmenden geometrischen Verteilung an, für größere Mittelwerte einer Binomialverteilung (Abb. 25).

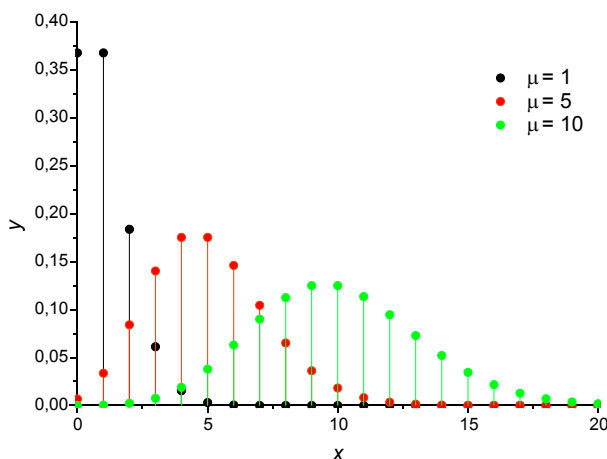


Abb. 25: Poissonverteilungen mit unterschiedlich großen Mittelwerten. Für $\mu = 1$: Annäherung an eine geometrische Verteilung; für $\mu = 10$: Annäherung an Binomialverteilung.

Bewertende Statistik

Bei wissenschaftlichen Untersuchungen muss man meist Vergleiche anstellen. Man möchte z.B. wissen, ob sich zwei Stichproben (Mittelwerte und Standardabweichungen) tatsächlich, d.h. signifikant, oder nur rein zufällig unterscheiden (Abb. 26). Für solche Fragestellungen stehen eine Reihe von statistischen Tests zur Verfügung.

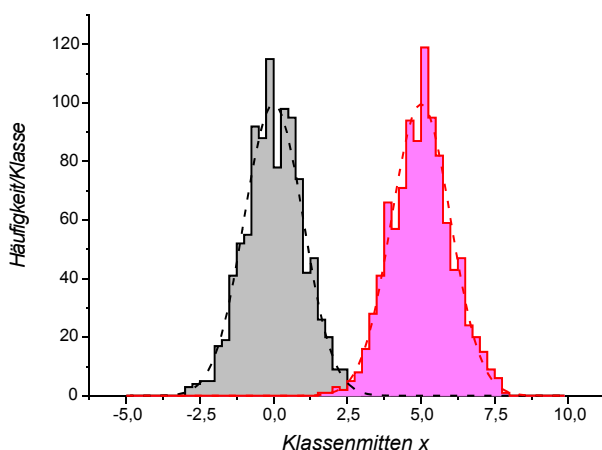


Abb. 26: Häufigkeitsverteilungen von zwei verschiedenen Stichproben mit unterschiedlich großen Mittelwerten und gleich großen Standardabweichungen.

Man kann sie grob in verteilungsabhängige und verteilungsunabhängige Tests einteilen. Bei verteilungsabhängigen Tests muss die Art der Verteilung (z.B. Normalverteilung) bekannt sein oder sie wird vorausgesetzt. Als Beispiel soll wieder das Gewicht der Bohnen dienen. Zwei

Stichproben derselben Bohnensorte werden sich nur rein zufällig, d.h. nicht signifikant voneinander unterscheiden, während sich zwei unterschiedliche Bohnensorten tatsächlich, d.h. signifikant unterscheiden könnten.

Der t-Test nach Student

Der t-Test dient zur Überprüfung, ob sich die Mittelwerte zweier Stichproben mit ansonsten gleichen Verteilungen unterscheiden. Um das zu prüfen, muss eine Bewertung für die Differenz der beiden arithmetischer Mittelwerte ($D = | \bar{x}_1 - \bar{x}_2 |$) der beiden Stichproben über dasselbe Merkmal gefunden werden. Dabei wird auch für diese Differenz eine Standardabweichung s_d gebildet, die sich aus den beiden Standardabweichungen der beiden Mittelwerte berechnen lässt:

$$s_d^2 = s_{\bar{x}_1}^2 + s_{\bar{x}_2}^2$$

Die Genauigkeit, die man für die Unterscheidung zweier Stichproben vorgibt, wird durch Vielfache der Standardabweichung angegeben:

$D > 2,576 s_d$: Abweichung ist signifikant (Vertrauensintervall für $s = 99\%$)

$D < 2,576 s_d$: keine sichere Aussage möglich

$D > 1,96 s_d$: keine sichere Aussage möglich

$D < 1,96 s_d$: Abweichung ist zufällig (Vertrauensintervall für $s = 95\%$)

Der so genannte t-Test nach Student beruht auf der t-Verteilung, die von W.S. Gosset unter dem Pseudonym Student veröffentlicht wurde. Mit dem t-Test wird geprüft, ob die Mittelwerte \bar{x}_1 und \bar{x}_2 mit ihren Standardabweichungen s_1 und s_2 zweier normalverteilter Stichproben (Voraussetzung beim t-Test) gleich oder verschieden sind (Abb. 26).

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s_d}$$

Das Ergebnis dieses Tests ist eine Fehlerwahrscheinlichkeit α in %, die angibt, ob der Unterschied der beiden Stichproben signifikant ist. Oft ist es nicht möglich, in beiden Messreihen den gleichen Probenumfang herzustellen, dann liegt also $n_1 \neq n_2$ vor. Bei kleineren Messreihen (unter 50 Varianten) oder bei Unterschieden zwischen n_1 und n_2 von mehr als 5 - 10 % muss dies berücksichtigt werden. Man verwendet dann die Berechnungsformel:

$$t = \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}}$$

Auch hier sind die wichtigsten Zahlenwerte schon aus den Berechnungen der Mittelwerte und Standardabweichungen der einzelnen Proben bekannt.

Die weitere Auswertung erfolgt mit einer [t-Tafel](#) (siehe Tabellenwerke der Statistik) oder mit einem t-Wert-Diagramm, in dem die t-Werte, die α -Werte und die Anzahl der Freiheitsgrade tabellarisch oder graphisch dargestellt sind. Mit dem berechneten Wert für t und der Anzahl der Freiheitsgrade ($m = n_1 + n_2 - 2$) kann man die Fehlerwahrscheinlichkeit ablesen. Zur Beurteilung des Versuchsergebnisses gilt:

- $\alpha < 0,01$: Es besteht ein signifikanter Unterschied zwischen den Proben
- $\alpha < 0,05$: Mit großer Wahrscheinlichkeit besteht ein Unterschied
- $\alpha > 0,05$: Ein Unterschied ist nicht anzunehmen
- $\alpha > 0,5$: Sehr wahrscheinlich besteht kein Unterschied

Paare von Messungen, Regression, Korrelation

Bisher sind Zufallsexperimente behandelt worden, in denen nur eine einzelne Variable x vorkam. Bei Problemen mit zwei Variablen (x, y) prüft man meist, ob eine Beziehung zwischen den Variablen besteht und welcher Gesetzmäßigkeit sie folgt. So kann man beispielsweise nach der Abhängigkeit zwischen Durchmesser x und Gewicht y bei Bohnen fragen. In der Analysis heißt y dann eine Funktion von x . In der Statistik spricht man stattdessen von der Regression y bezüglich x . Diese nichts sagende Beziehung (Regress = Rückschritt) hat sich leider allgemein eingebürgert und erhalten. Interessiert man sich nur für die Beziehung zwischen x und y , ohne nach der Abhängigkeit zu fragen, so spricht man von der Korrelation zwischen x und y .

Regressionsgerade, Prinzip der kleinsten Quadrate

Liegt eine Stichprobe von Beobachtungen $(x_1, y_1), (x_2, y_2), \dots (x_n, y_n)$ aus einer zweidimensionalen Grundgesamtheit vor, so trägt man sie am besten in ein kartesisches Koordinatensystem ein (Abb. 27). Man erhält dann entweder eine Punktwolke oder einen mehr oder weniger guten Zusammenhang der Punkte, die vielleicht schon optisch eine Gerade ergeben könnten. Man könnte dann subjektiv eine Ausgleichsgerade oder Regressionsgerade durch die Punkte zeichnen und zu einem beliebigen x -Wert den zugehörigen y -Wert ablesen (Abb. 27).

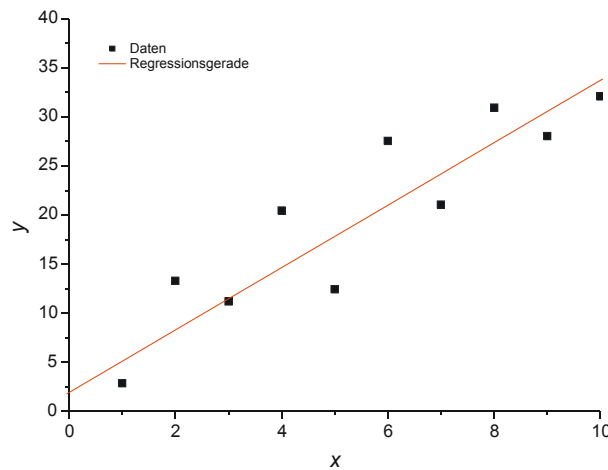


Abb. 27: Regressionsgerade durch verstreut liegende Punkte.

Liegen die Punkte jedoch nicht mehr so ideal, so werden verschiedene Personen im Allgemeinen verschiedener Meinung darüber sein, wie die Ausgleichsgerade zu legen ist. Um subjektive Einflüsse auszuschalten, muss wieder eine objektive Methode herangezogen werden. Eine solche ist das Gaußsche Prinzip der kleinsten Quadrate (least square). Es besagt bezüglich der Regressionsgerade (lineare Regression) folgendes:

Die Gerade $y = a x + b$ ist so zu legen, dass die Summe der Quadrate aller Abstände der Punkte von der Geraden möglichst klein (Minimum) wird (Abb. 27).

Unter dem Abstand eines Punktes von einer Geraden versteht man üblicherweise die Länge des Lotes von dem Punkt auf die Gerade, also den senkrechten Abstand. Man benutzt aber aus rechentechnischen und theoretischen Gründen den Abstand parallel zur y-Richtung. Beide sind trigonometrisch einfach ineinander umrechenbar. Die Summe der Abstände wird minimal, wenn man die Steigung a der Regressionsgeraden wie folgt berechnet:

$$a = \frac{\frac{1}{n} \sum_i x_i y_i - \bar{x} \bar{y}}{\frac{1}{n} \sum_i x_i^2 - \bar{x}^2}$$

Der Achsenabschnitt b der Regressionsgeraden berechnet sich mit dieser Steigung a :

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

Man kann sich die Suche nach dem Minimum mit einem Modell aus der Mechanik anschaulich erklären. Die Regressionsgerade stellt einen Stab dar, auf den von den einzelnen Punkten aus Kräfte, vermittelt durch Gummiringe (bzw. Federn), einwirken. Die Gummiringe müssen dabei parallel zur y-Richtung gehängt werden. Der Stab bewegt sich automatisch in eine Lage, in der das Gesamtdrehmoment auf ihn gleich Null ist, d.h. der Stab kommt zur

Ruhe. Diese Lage entspricht genau dem gesuchten Minimum für die Summe der Abstände und somit der theoretisch berechneten Regressionsgeraden.

Als Maß dafür, wie gut die einzelnen Punkte eine Gerade ergeben, wie gut sie also korrelieren, dient der Korrelationskoeffizient r . Er berechnet sich nach:

$$r = \frac{\sum_i (x_i - \bar{x}) \sum_i (y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

Der Wert für r bewegt sich zwischen -1 und +1.

$r = -1$: Regressionsgerade mit negativer Steigung und beste Korrelation der Messwerte.

$r = 0$: Die Messwerte zeigen überhaupt keine Korrelation.

$r = +1$: Regressionsgerade mit positiver Steigung und beste Korrelation der Messwerte.

Viele Taschenrechner besitzen auch eine direkte Möglichkeit zur statistischen Berechnung von Regressionsgeraden. Sie geben als Ergebnis die Steigung a , den Achsenabschnitt b und den Korrelationskoeffizienten r aus.

Der Vollständigkeit halber soll noch erwähnt werden, dass es auch nicht lineare Regressionen für beliebige Funktionsgleichungen nach unterschiedlichen Verfahren gibt. Da sie rechnerisch aufwendig sind, werden sie meist mit Computerprogrammen (Statistik Software) realisiert. Außerdem gibt es noch die Möglichkeit, verschiedene Funktionen in Geraden zu transformieren und dann eine lineare Regression durchzuführen. Hierbei muss man jedoch beachten, dass die Abstände von den Punkten zur Kurve entsprechend mittransformiert werden. Man spricht bei der Regression auch von Approximation, Anpassung, Annäherung, Ausgleichsrechnung oder Kurvenfitting (kurz Fitting). Alle Begriffe bringen zum Ausdruck, dass theoretische Kurven an vorhandene Messwerte angepasst werden sollen.

Literatur

1. Elementare Einführung in die Wahrscheinlichkeitsrechnung
Bosch, K.
Vieweg Studium Basiswissen, Braunschweig, 1986

2. Elementare Einführung in die angewandte Statistik
Bosch, K.
Vieweg Studium Basiswissen, Braunschweig, 1987
3. Mathematische Funktionen zur Beschreibung von Vorgängen in Natur und Technik
Hopp, V., Berninger, G.
GIT Fachz. Lab. 8, 682-691, 1987
4. Statistische Methoden und ihre Anwendungen
Kreyszig E.
Vandenhoeck u. Ruprecht, Göttingen, 1975
5. Basic-Programme für die angewandte Statistik
Müller, G.W., Kick, T.:
R.Oldenburger Verlag, München, 1985
6. Funktionstabellen und statistische Tabellen
Vogel, A.
Verlag Konrad Wittwer, Stuttgart, 1979
7. Methoden der Statistik
Wallis W.A., Roberts H.V.
rororo, 1969

Links

- [Elektronisches Statistiklehrbuch](#)
- [Statistikressourcen](#)