

Bernd Lederer: Quantitative Auswertungsmethoden

Es folgt ein erster einführender Überblick auf die **gängigsten Verfahren** bei der Auswertung quantitativer Daten, wie sie etwa bei einer Beobachtung, einer Befragung oder einem Experiment anfallen. Hierbei lernen Sie die allerwichtigsten **Kenngrößen** und **Vorgehensweisen** sowohl der **beschreibenden** als auch der **vergleichenden Statistik** kennen.

Sie haben eine Befragung, eine Beobachtung, eine Inhaltsanalyse oder gar ein Experiment durchgeführt und dabei jede Menge hierfür interessanter Daten erhoben. Jetzt geht es darum, diese Daten auszuwerten und sozusagen „in Form zu bringen“. Dies kann auf qualitative (siehe Abschnitt V.1) oder eben quantitative Art und Weise geschehen (wobei sich diese beiden Ansätze aber keinesfalls gegenseitig ausschließen!).

Zur Auswertung quantitativer Daten gehören unvermeidlich zumindest *grundlegende* Kenntnisse der Statistik. **Statistik bezeichnet die Befassung mit größeren Datenmengen**, wie sie für die quantitative empirische Sozialforschung üblich sind. Sie sucht nach möglichen **Regelmäßigkeiten** und **Gesetzmäßigkeiten**.

Erfahrungsgemäß schreckt allein schon das Wort „Statistik“ viele Studierende (womöglich auch Sie?) regelrecht ab: Haben sich manche von Ihnen nicht sogar deshalb für ein Studium der Sozial- und/oder Geisteswissenschaften entschieden, um eben nicht mehr mit mathematischen Prozeduren belästigt zu werden? Keine Angst: Im Folgenden geht es hier auch gar nicht darum, Sie mit einem „Crashkurs“ in Statistik zu bedienen (was auf den wenigen Seiten ohnehin nicht einmal denkbar ist!). Vielmehr soll Ihnen hier ein allererster Überblick auf die allerwichtigsten Kenngrößen, Begrifflichkeiten und Prozeduren einer Datenauswertung im Rahmen quantitativer empirischer Forschung geboten werden. Es wird Ihnen hier in aller Kürze sozusagen „die Spitze des Eisbergs“ präsentiert, wobei sich diese Spitze aber durchaus als **Einstiegshilfe** in die Thematik eignet. Damit soll Ihnen auch – so vorhanden – der Argwohn gegen alles, was mit Zahlen und Formeln zu tun hat genommen und so die weitere Annäherung an das Thema erleichtert werden. Schließlich werden Sie sehen, dass die Grundlagen einer **beschreibenden** und einer **vergleichenden Statistik** keine höhere Mathematik erfordern und im Grunde genommen viel einfachere Sachverhalte umfassen, als viele Studierende oft befürchten.

A) Deskriptive Statistik

Sie haben im Rahmen eines Forschungsprojekts also empirische Daten gesammelt, etwas durch eine Befragung oder eine Beobachtung, die es nun auszuwerten und letztlich auch zu präsentieren gilt. **Daten sind beobachtete Merkmalsausprägungen von empirischen Beobachtungseinheiten**, die letztlich auf **Messskalen** abgebildet werden (z.B. „Metermaß“, „Temperatur in Grad Celsius“, „Richterskala“ etc., auch Diagramme gehören hierzu, etwa solche mit „X- und Y-Achse“ usw.).

Die „**deskriptive Statistik**“ („beschreibende Statistik“) beschäftigt sich mit der **Zusammenfassung von Daten** und mit der **Präsentation aussagekräftiger Kennzeichen** einer Datenmenge. Dabei bieten sich Ihnen verschiedene Möglichkeiten an, Daten zusammenfassend und aussagekräftig darzustellen.

1. Tabellarische Darstellungen von Daten

Eine der gängigsten Möglichkeiten, eine Auszählung von Daten **optisch** ansprechend darzustellen, liefert etwa folgendes Beispiel. Es wurde das Bruttoeinkommen von insgesamt 200 Arbeitnehmer/innen erhoben. Die **Stichprobe** umfasst also 200 Personen, man sagt: **n (Stichprobenumfang) = 200** .

(Die **Grundgesamtheit**, aus der heraus die Stichprobe gezogen wurde, wird mit **N** bezeichnet. In unserem Beispiel also etwa alle Arbeitnehmer/innen eines bestimmten Unternehmens, also bspw. $N = 2500$).

Der besseren Übersicht halber wurden die Ergebnisse in unserem Beispiel in 5 **Kategorien** von Einkommensuntergruppen unterteilt. Die **absolute Häufigkeit** von „Treffern“ in jeder einzelnen Kategorie wird als f_i bezeichnet, wobei „ f “ für „*frequency*“ (Frequenz) steht und „ i “ für *Index* steht (also hier die Rangzahl der jeweiligen Kategorie).

Häufigkeitentabelle 1

Kategorie	Brutto- verdienst von ... bis	<i>absolute</i> Häufigkeit f_i	<i>relative</i> Häufigkeit (%)	<i>kumulierte</i> (d.h. aufsummierte) absolute Häufigkeit	<i>kumulierte</i> relative Häufigkeit
1	1200–1599	8	4	8	4
2	1600–1999	46	23	54	27
3	2000–2399	68	34	122	61
4	2400–2799	64	32	186	93
5	2800–3200	14	7	200	100
		Summe=200	Summe=100		

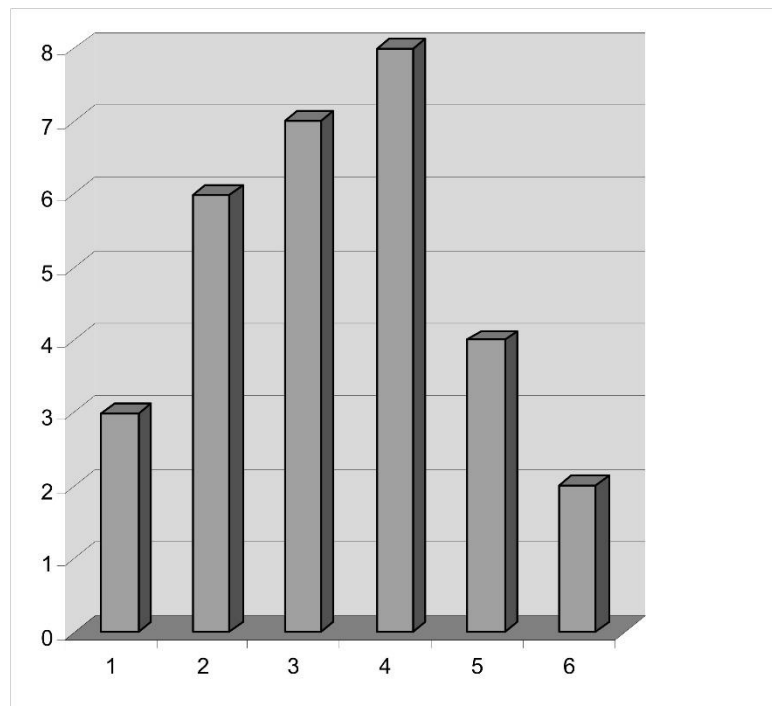
Ein anderes Beispiel für die mögliche tabellarische Darstellung von Daten liefert die folgende Tabelle. In ihr sind die Schulnoten (in sechs möglichen Kategorien in Form der Noten von 1 bis 6) und die Häufigkeiten, mit denen diese in einer Schulklasse mit 30 Schülern ($n = 30$) vergeben wurden, gegenüber gestellt.

Häufigkeitentabelle 2

Noten x_i (x steht für die Kategorie)	Häufigkeit f_i
1	3
2	6
3	7
4	8
5	4
6	2

Eine solche ganz und gar typische Tabelle kann auch in Form eines **Balkendiagramms** (in der Fachsprache ein so g. „**Histogramm**“) dargestellt werden. Dies ist optisch weitaus ansprechender und vor allem aussagekräftiger, wie Sie sicher bestätigen werden:

Histogramm 1



Verbindet man nun noch die **Mittelpunkte der Balkenspitzen** miteinander, erhält man ein sogenanntes „**Polygon**“, das gleich auf den ersten Blick nochmal etwas aussagekräftiger erscheint, wie folgendes Beispiel zeigt:

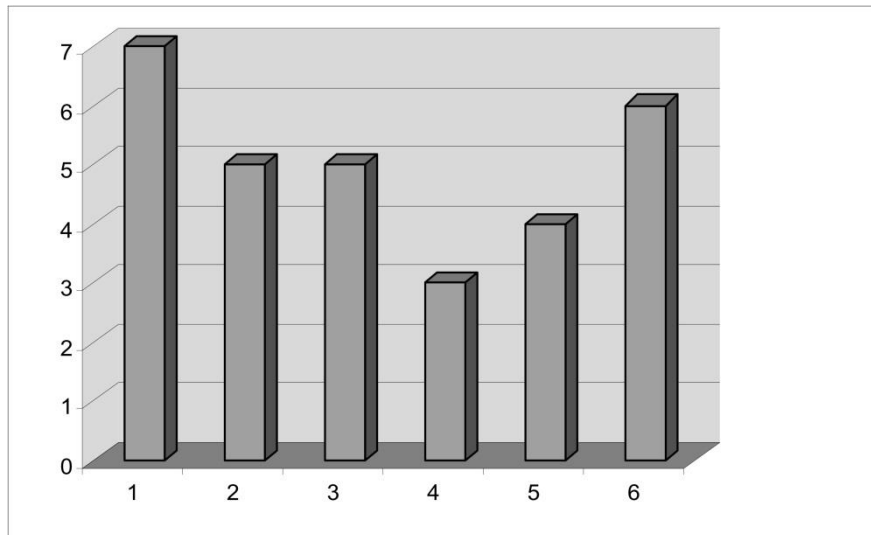
Zunächst die zugehörige Häufigkeitstabelle (Der Stichprobenumfang bleibt der Gleiche, d.h. **n = 30**), die Häufigkeiten sind diesmal jedoch andere:

Häufigkeitentabelle 3

Noten x_i	Häufigkeit f_i
1	7
2	5
3	5
4	3
5	4
6	6

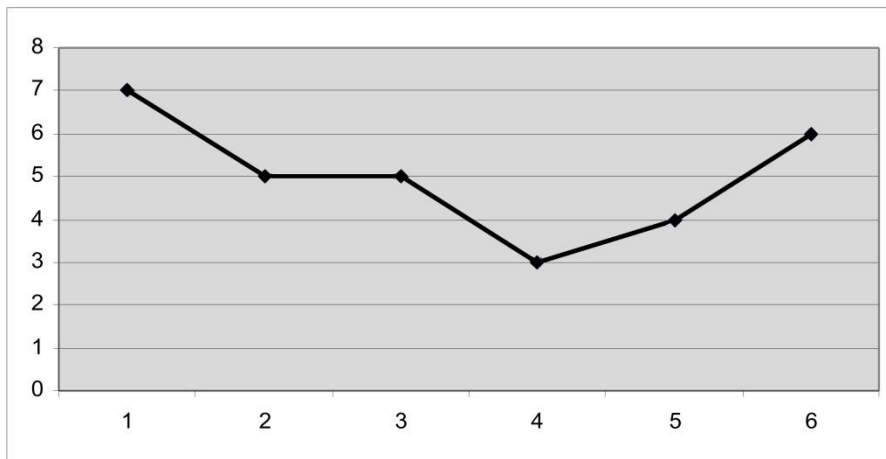
Die Daten lassen sich zunächst wieder als **Histogramm** darstellen:

Histogramm 2



Oder in Form eines **Polygons**:

Polygon



Bei den vorangegangenen Beispielen handelt es sich nur um einige der gängigsten Möglichkeiten, Datenmengen grafisch ansprechend und aussagekräftig aufzubereiten. Geläufig sind Ihnen z.B. sicher auch die sog. „**Tortendiagramme**“ (die etwa oft bei der Visualisierung von Wahlergebnissen genutzt werden).

Die hier verwendeten Tabellen wurden mit dem gerade für Einsteiger sehr einfach zu bedienenden Programm „Excel“ erstellt. Jedoch sind auch viele andere Anwendungen möglich, etwa mit dem Statistikprogramm „SPSS“ – näheres hierzu unter XYZ).

Im folgenden erhalten Sie einen Überblick auf die wichtigsten **statistischen Kenngrößen** der **beschreibenden („deskriptiven“) Statistik**, die aus solchen Grafiken wie den gezeigten nicht direkt

ablesbar sind. Es handelt sich dabei um Angaben und Werte, die als **zentrale Kennziffern wichtige Eigenschaften einer Datenmenge** ausdrücken und diese dadurch beschreiben helfen.

Welches also sind nun die wichtigsten und meist angeführten Größen der deskriptiven Statistik?

2. Der MODUS bzw. MODALWERT

Der grundlegendste Kennwert, der zur Beschreibung größerer Datenmengen herangezogen werden kann, ist der **Modalwert**, auch **Modus** genannt (die Abkürzung dafür ist **X_{mod}**). Der Modalwert ist einfach derjenige einzelne Messwert, der am **häufigsten** Auftritt, also die Kategorie mit der größten Häufigkeit.

In den Beispielen 2 und 3 ist der Modus 4; in den Beispielen 4, 5 und 6 hingegen gibt es keinen Modus, weil es keinen häufigsten Wert gibt, da zwei Kategorien gleichauf liegen. Gibt es mehrere Merkmalsausprägungen mit der gleichen maximalen Häufigkeit, so existiert also kein Modalwert!

Versuchen Sie doch einmal, den Modalwert in folgender Reihung zu bestimmen:

3 / 4 / 4 / 5 / 5 / 5 / 5 / 6 / 6 / 6 / 7 / 8

Richtig! Es ist die 5

In folgender Reihung hingegen

3 / 4 / 4 / 4 / 4 / 5 / 6 / 6 / 6 / 7

ist der Modalwert die 4

3. Der MEDIAN

Die zweite wichtige statistische Kenngröße ist der **Median**, auch **Zentralwert** genannt (Kurzzeichen: **X_{med}**). Er ist derjenige Wert, also diejenige Merkmalsausprägung, die **in der Mitte** steht, sofern alle Daten nach ihrer Größe sortiert und aufgereiht sind (bei Daten, die nicht in eine Reihenfolge gebracht werden können, gibt es entsprechend auch keinen Median!). Der Median **halbiert** aufgrund seiner Mittelstellung die Messwerte in **zwei gleich große Teile**.

Im folgenden Beispiel (gesammelte Punkte in einem Test mit $n = 9$):

3 / 5 / 6 / 7 / 9 / 10 / 11 / 13 / 14

ist der Median wegen der ungeraden Zahl der Stichprobe „aufgerundet“ die 5. Stelle in der Reihung, also die 9.

Bei einer geraden Zahl von Werten ($n = 8$), etwa

3 / 5 / 6 / 7 / 9 / 10 / 11 / 13

ist der Median genau in der Mitte, also zwischen dem 4. und 5. Rangplatz; in diesem Beispiel liegt er also genau zwischen 7 und 9, der Median ist somit 8.

Ein weiteres Beispiel soll Ihnen nun noch den großen Nutzen bzw. die große Aussagekraft des Medians verdeutlichen (und zudem vom nachfolgend vorgestellten Mittelwert abgrenzen helfen).

In einer Belegschaft wurden folgende Bruttoverdienste (in Euro) erhoben ($n = 9$):

970 / 980 / 1050 / 1090 / 1160 / 1180 / 1200 / 1800 / 6600

Der Median liegt am 5. Rangplatz, ist also der Wert 1160,- Euro. Wie Sie sehen können, gibt es einen „Ausreißer“, der den errechneten Durchschnitt der Einkommen (also die Summe der Gesamteinkommen geteilt durch 9) auf 1781, 11 Euro anhebt. Der Durchschnitt täuscht aber gewaltig, weil 7 von 9 Angestellten höchstens 1200 Euro verdienen!

Der Median als derjenige Wert, der in der Mitte einer Rangreihenfolge steht, ist hier sicher aussagekräftiger, weil atypische Ausreißer nach oben oder unten keine Rolle spielen!

4. Das arithmetische Mittel

An dieser Stelle wird es ein (klein) wenig komplexer, was die verwendeten Zeichen und Formeln angeht, nicht aber, was die Inhalte betrifft.

Das „arithmetische Mittel“ ist umgangssprachlich „der Durchschnitt“ einer Datenmenge. Das Zeichen dafür ist \bar{x} . Die mathematische Formel für das arithmetische Mittel ist bei einfachen Datenreihen

$$x_1 + x_2 + x_3 \dots \text{usw. bis } + x_n : n$$

Mathematisch formuliert lautet die Formel:

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

bzw. (anders geschrieben)

$$\bar{x}_{\text{arithm}} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Es besteht übrigens kein Grund zur Beunruhigung: Bereits Taschenrechner, ganz zu schweigen von Statistikprogrammen wie Excel oder SPSS, ersparen Ihnen komplexe Rechenoperationen und liefern die Lösung per Tastendruck bzw. Mausklick!

Hier ein einfaches Beispiel zur Berechnung des arithmetischen Mittels: 4 Personen ($n = 4$) haben folgende Punktwerte in einem Test erzielt:

7 / 3 / 5 / 5

$$\bar{x} \quad \text{ist dann } 7 + 3 + 5 + 5 : 4 ; \quad 20 : 4 = 5$$

Wenn mehrere Kategorien mit unterschiedlichen zugehörigen Daten vorhanden sind (wie im Beispiel 2 bzw. 3), gilt die Formel

$$\bar{x} = \sum x_i \cdot f_i : n$$

(x_i ist dabei die „i-te“ Beobachtungskategorie, f_i die Anzahl der zugehörigen Daten)

Im Beispiel der Häufigkeitentabelle 2 bzw. des Histogramms 1:

$$\begin{aligned} \bar{x} &= (1 \cdot 3 + 2 \cdot 6 + 3 \cdot 7 + 4 \cdot 8 + 5 \cdot 4 + 6 \cdot 2) : 30 \\ &= 3 + 12 + 21 + 32 + 20 + 12 : 30 \end{aligned}$$

$$= 100 : 30 = 3,33$$

5. Varianz und Standardabweichung

Die **Varianz** (S^2 , manchmal auch einfach mit **V** bezeichnet) und die **Standardabweichung** (s = die Wurzel aus S^2 bzw. V) sind ein Maß dafür, wie stark die einzelnen Messwerte einer Untersuchung durchschnittlich **um den Mittelwert verteilt** sind.

Schauen Sie sich zur Verdeutlichung noch einmal die beiden *Histogramme* in den Beispielen 3 und 5 an: Ob Sie es auf den ersten Blick glauben oder nicht: das *arithmetische Mittel* beträgt in beiden Fällen das Gleiche! (Rechnen Sie es ruhig nach: Sie kommen auf den Wert 3,33) Dabei sind doch die Messwerte/Daten ganz offensichtlich völlig unterschiedlich um diesen Mittelwert herum verteilt! Hier nun kommen die **Varianz** bzw. die **Standardabweichung** ins Spiel: Sie gibt Auskunft darüber, ob die erhobenen Daten weit um den jeweiligen Mittelwert herum gestreut sind oder ob sie eher eng um den Mittelwert herum zu liegen kommen.

Die obigen Beispiele (3 und 5) verdeutlichen Ihnen den Nutzen und die Aussagekraft dieser statistischen Kenngrößen: Es macht eben einen großen Unterschied, ob das Ergebnis einer Klausur darin besteht, dass die meisten Noten um „die Mitte“ herum zu liegen kommen oder ob es statt dessen einige sehr gute Zensuren auf der einen und einige sehr schlechte auf der anderen Seite des Spektrums gibt, hingegen kaum Durchschnittsleistungen – auch wenn der Mittelwert jeweils der Gleiche ist.

Die zugegeben etwas abschreckende Formel für die Varianz lautet ausgeschrieben:

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

Einen Spezialfall bildet die Formel bei sehr großen Grundgesamtheiten:

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

Auch in diesen beiden Fällen gilt wiederum, dass entsprechende Programme Ihnen jedwede Rechenarbeit abnehmen!

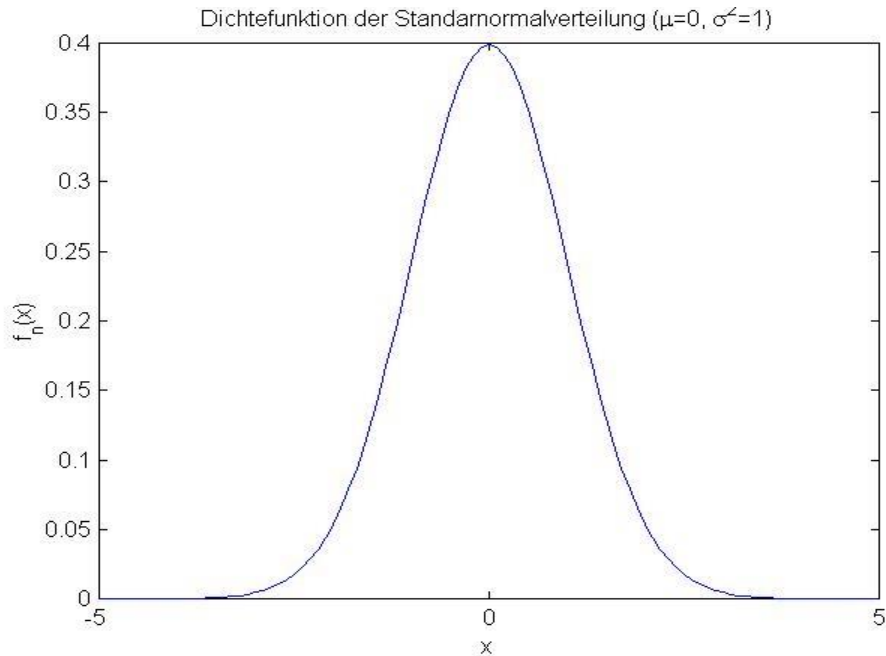
Da die **Varianz** (S^2) als quadrierte Größe vorliegt, wird der besseren Anschauung zuliebe die **Wurzel** gezogen und man erhält die **Standardabweichung** s (auch **Streuung** genannt), die aber die gleiche grundsätzliche Bedeutung für die Interpretation zahlenmäßiger Daten hat.

Noch ein Beispiel, um Ihnen die Aussagekraft dieser Kenngrößen zu verdeutlichen: In einer Gruppe von in etwa gleichgroßen männlichen Basketballspielern dürfte die durchschnittliche Körpergröße relativ eng um das arithmetische Mittel herum **streuen**. D.h., dass die einzelnen Abweichungen nach unten und nach oben nicht allzu stark ausfallen sollten. Ganz anders schaut das aber auf die Gesamtbevölkerung bezogen aus, wo es weitaus mehr (sehr) kleine und (sehr) große Menschen gibt und die Messwerte deshalb insgesamt *breiter* (um den Mittelwert herum) *streuen* als in der Stichprobe der Basketballspieler.

Am besten lässt sich die Streuung anhand einer „**Normalverteilungskurve**“ erklären: Sehr viele Datenerhebungen sind nämlich „**normalverteilt**“, wie es in der Statistik heißt: Ob Körpergröße, Intelligenzquotient oder durchschnittliche Studiendauer: in allen Fällen gibt es einen Mittelwert, um den herum die meisten Einzelfälle zu liegen kommen, wohingegen Ausreißer nach oben und unten eher die Ausnahme bilden: Die allermeisten Menschen sind in Europa zwischen 1,60 m und 1,90 m groß, nur relativ wenige sind sehr klein oder sehr groß und kommen darunter oder darüber zu liegen. Als **Diagramm** (mit „X- und Y-Achse“) aufgetragen entsteht so eine „**Normalverteilung**“. (In

Anlehnung an den Mathematiker Carl-Friedrich Gauß wird dabei auch von einer „**Gauß'schen Glockenkurve**“ gesprochen.) Sie sieht der Form nach folgendermaßen aus:

Standardnormalverteilung



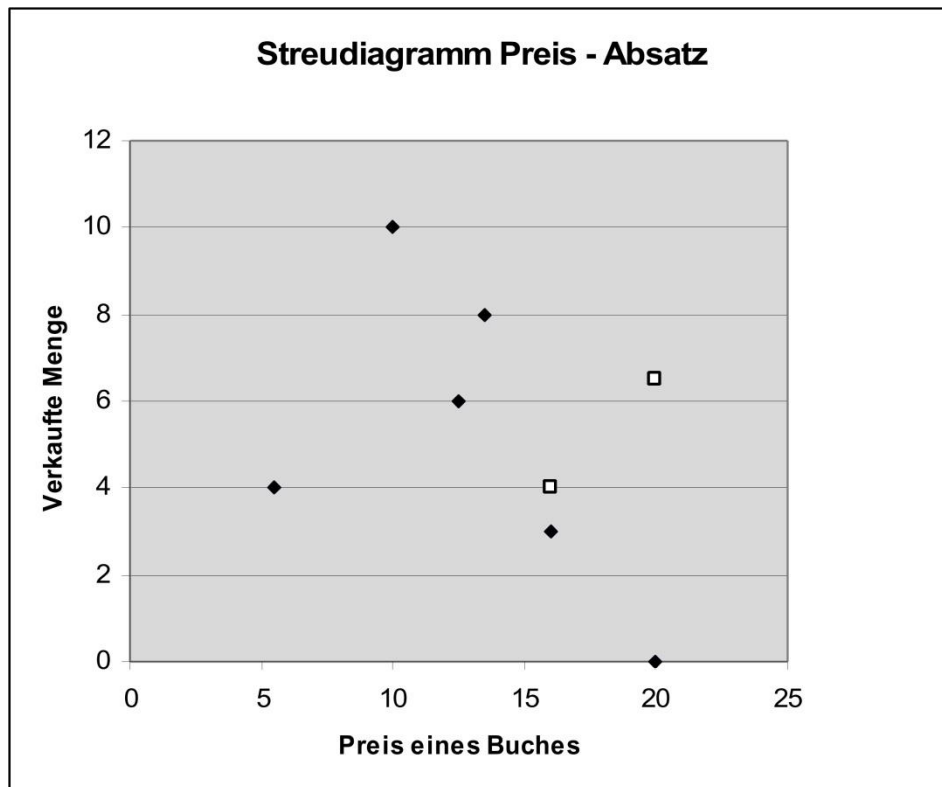
Definitionsgemäß werden bei einer Normalverteilung der Mittelwert mit μ und die Varianz mit σ^2 , die Standardabweichung mit σ symbolisiert. Betragen diese Werte 0 (Mittelwert) und 1 (Varianz), handelt es sich um eine **Standardnormalverteilung**. Unterscheiden sich die Werte der Glockenkurve von 0 und 1, handelt es sich um keine Standardnormalverteilung, sondern eine einfache Normalverteilung.

Je *enger* bzw. *steiler* und *schmäler* nun diese Glockenkurve ist, desto geringer ist die Varianz bzw. Standardabweichung, desto enger sind also die Daten um den Mittelwert versammelt. Je *breiter* und *weniger steil* die Verteilung hingegen, desto größer ist die Standardabweichung.

6. Der Korrelationskoeffizient

Da gerade von zweidimensionalen Diagrammen die Rede war (also solchen mit einer „x- und einer y-Achse“): Ein weiteres wichtiges Kriterium der beschreibenden Statistik ist der sog. „**Korrelationskoeffizient**“. Immer wenn zwei einander zugehörige Merkmale gegeneinander aufgetragen werden (z.B. die Anzahl von Personen auf der x-Achse und deren jeweiliges Einkommen auf der y-Achse), haben Sie es mit Diagrammen zu tun, die etwa die nachfolgende Form annehmen können.

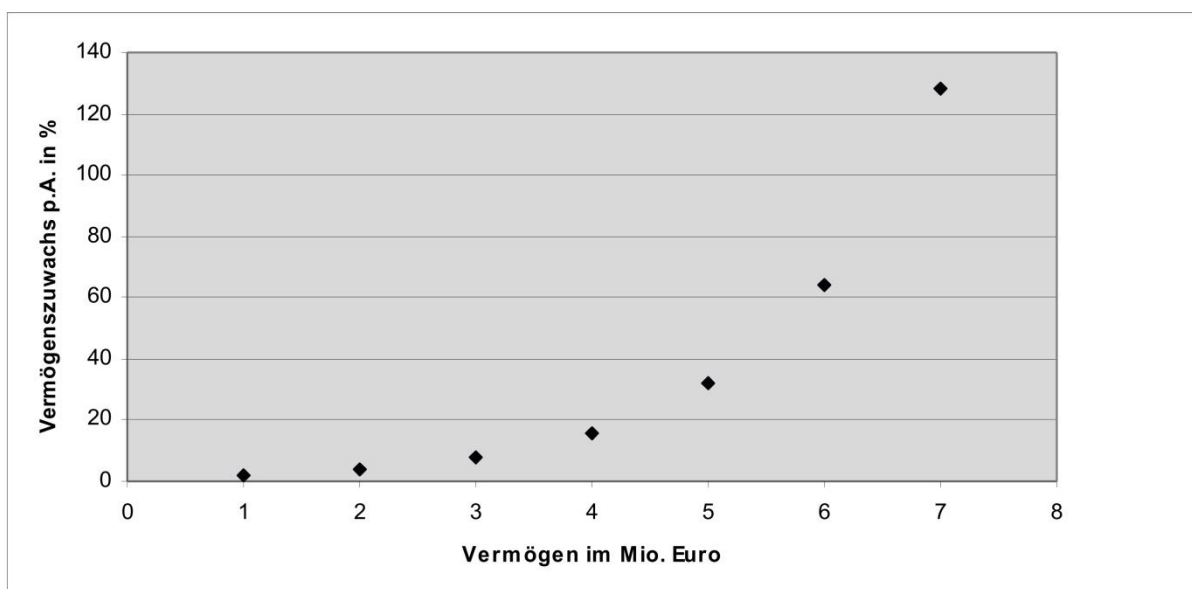
Streudiagramm



In diesem Beispiel weisen die beiden **Variablen** (einerseits der Preis einer Ware, andererseits deren verkaufte Stückzahl) offenbar *keinen* engeren Zusammenhang auf. Das Diagramm zeigt einen sog. „**Schwarm**“ von Daten, Fachleute sprechen in diesem Zusammenhang von einem „**Streudiagramm**“.

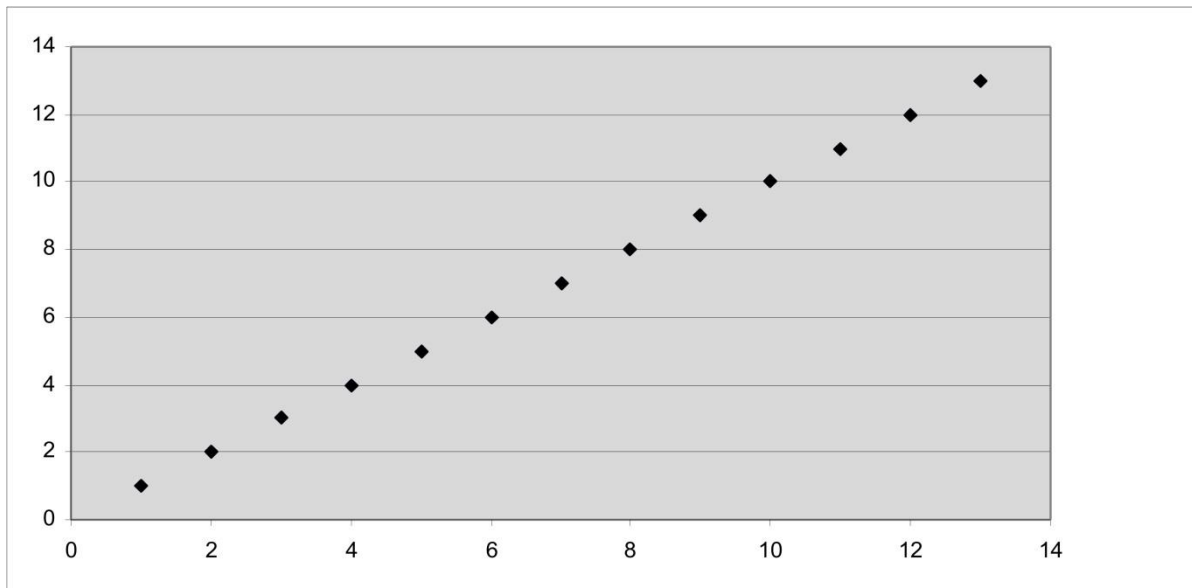
Anders stellt sich Ihnen das im nachfolgenden Beispiel-Diagramm dar, wo die Kurve („**Graph**“ ist der Fachausdruck hierfür) ein steiles, „**exponentielles**“ **Wachstum** an den Tag legt.

Exponentielles Wachstum



Im nächsten Beispiel sind die beiden Variablen so miteinander verbunden („**korreliert**“), dass der Graph eine gleichmäßig steigende Gerade darstellt. Man sagt, der Zusammenhang ist **linear**.

Lineares Wachstum



In den beiden Beispielen springt jeweils der Zusammenhang der beiden Größen geradezu ins Auge. Wie aber lässt sich unabhängig von solchen Graphen der jeweilige Zusammenhang zweier Merkmale *berechnen*? Schließlich kann dieser Zusammenhang offensichtlich ganz unterschiedlicher Art sein und wie in Beispiel 8 oft auch überhaupt keine Gerade ergeben. Anders gefragt: Wie kann berechnet werden, wie die jeweiligen **Variablen miteinander korreliert** sind?

Auskunft darüber gibt der sog. „**Korrelationskoeffizient**“ (sein Zeichen ist **r**). Die mathematische Formel ist zugegeben etwas kompliziert (sie weist Klammern, einen Bruch und eine Wurzel auf), erfordert aber dennoch keinerlei höhere Mathematikkenntnisse und hat uns an dieser Stelle auch nicht weiter zu interessieren (zumal verschiedene leicht zu bedienenden Statistikprogramme für den PC dergleichen „mir nichts, dir nichts“ auszurechnen in der Lage sind.)

Ist **r** berechnet, kann das Ergebnis jeden Wert zwischen -1 und $+1$ annehmen (d.h.: $-1 \leq r \leq +1$). Ist $r = 0$, heißt das, dass es keinen Zusammenhang zwischen den Werten gibt (wie im Beispiel 7). Ist **r** gleich $+1$, sind die Daten **positiv korreliert** (d.h., große **x**-Werte entsprechen großen **y**-Werten, wie in den Beispielen 8 und 9), ist **r** gleich -1 , sind die Daten umgekehrt **negativ korreliert**: Kleine **y**-Werte entsprechen dann großen **x**-Werten. Grundsätzlich gilt: **Je kleiner r, desto schwächer ist der Zusammenhang zweier Merkmale.**

B) Inferenzstatistik

Bisher ging es darum, Ihre empirisch gewonnen Daten anhand markanter Merkmalsgrößen zu *beschreiben*. Hierfür zeichnet die vorangehend skizzierte *deskriptive Statistik* verantwortlich. Die **Inferenzstatistik** („schließende Statistik“) hingegen beschäftigt sich mit der Frage, ob wissenschaftliche Ergebnisse dem **Zufall** geschuldet sind oder ob ihnen eine wissenschaftliche **Gesetzmäßigkeit** zugrunde liegt - in letzterem Fall spricht man von einem **signifikanten Ergebnis**.

Die Frage nach „Zufall und Notwendigkeit“ in der Wissenschaft hat auch viel zu tun mit dem **Zusammenhang von Stichproben und den Grundgesamtheiten**, aus denen diese Stichproben gezogen werden. Ist ein bestimmtes Untersuchungsergebnis zufällig nur für eine bestimmte Stichprobe gültig oder auch für andere, vergleichbare Stichproben einer interessierenden Grundgesamtheit von Untersuchungsobjekten (Menschen, Ereignisse usw.)?

Es geht also auch um die Frage nach *möglichen Rückschlüssen von der Stichprobe auf die Gesamtheit* eines zu untersuchenden Phänomens (also z.B. danach, inwieweit eine bestimmte Merkmalsausprägung der Schüler einer bestimmten Schule in einer bestimmten Stadt typisch ist für alle Schüler eines bestimmten Bundeslandes). Von großer Bedeutung ist innerhalb der schließenden Statistik somit der Aspekt der **Repräsentativität**. Was aber bedeutet es aus Sicht der Wissenschaft eigentlich, wenn es beispielsweise heißt: „Das Ergebnis ist **repräsentativ**?“

1. Repräsentativität

Prinzipiell gibt es verschiedene Arten von Stichproben: z.B. **Zufallsstichproben**, bei denen jede Untersuchungseinheit einer Grundgesamtheit die gleiche Chance hat, in die Stichprobe zu gelangen. Das beste Beispiel einer solchen rein zufälligen Auswahl ist das *Losverfahren*. Zufallsstichproben weisen dabei die höchste **Repräsentativität** auf. **Systematische Verfahren** hingegen, etwa *Quotenverfahren* oder eine *systematische Auswahl*, „ziehen“ eine Stichprobe nicht zufällig, sondern berücksichtigen (je nach Forschungsinteresse) bestimmte Gruppen bevorzugt oder auch ausschließlich. (Beispielsweise kann eine Gesamtheit von Personen nach Alter, Geschlecht, Bildungsabschluss etc. aufgeteilt werden. Die Stichproben werden dann gemäß der prozentualen Anteile dieser Gruppen an der Gesamtbevölkerung zusammengestellt).

Repräsentativität heißt, dass die Stichprobe(n) so ausgewählt wurde(n), dass sich die Merkmale, um die es in der Untersuchung geht, hinsichtlich ihrer statistischen Kenngrößen (wie *Median*, *Modalwert*, *arithmetisches Mittel*, *Streuung* usw.) in Stichprobe *und* Gesamtpopulation nicht zu stark unterscheiden, so dass sich **von der Stichprobe auf diese Grundgesamtheit rückschließen** lässt. Wie groß muss nun aber eine Stichprobe sein, um sie als repräsentativ bezeichnen zu können? Grundsätzlich gilt: Eine ganz bestimmte Zahl (etwa: immer 10% einer Gesamtheit) gibt es nicht, vielmehr hängt diese immer auch vom Untersuchungsgegenstand ab. (Bei manchen Merkmalen wie dem Wahlverhalten reichen schon relativ kleine Fallzahlen [gemessen an der Gesamtbevölkerung], um aussagekräftige Rückschlüsse ziehen zu können, denn die Zahl potentieller Wahlentscheidungen ist durch das Parteienspektrum relativ eng begrenzt. Hinsichtlich einer Befragung des „Lieblingsessens der Deutschen“ bedürfte es hingegen umfangreicherer Stichproben.)

Oft sind es auch wissenschaftlich-methodische **Erfahrungswerte**, mit der die notwendige Anzahl befragter (oder beobachteter) Menschen festgelegt wird, die mindestens erforderlich ist, um die gleichen (oder sehr ähnliche) Ergebnisse zu erzielen, wie sie größere (und damit repräsentivere) Stichproben erzielen. Darüber hinaus gibt es in der Statistik aber auch (relativ komplizierte) **mathematische Verfahren**, um die optimale Stichprobengröße zu bestimmen.

Ob Hypothesen, die anhand einer Stichprobe überprüft werden, nur *zufällig* bestätigt oder widerlegt werden, oder tatsächlich Rückschlüsse auf die Grundgesamtheit zulassen, darüber geben bestimmte Testverfahren Auskunft.

2. Hypothesentest

Hypothesentests sagen also etwas darüber aus, ob von einer gezogenen Stichprobe auf die untersuchte Grundgesamtheit geschlossen werden kann. Bei diesem Verfahren geht es nun aber schon ein klein wenig stärker „ans Eingemachte“ der Statistik. Entsprechend sollen Ihnen hier im Rahmen einer ersten Einführung in die Thematik auch nur die Grundzüge skizziert werden.

Wie Sie bereits wissen, bezeichnet Wissenschaft im Grunde nichts anderes, als das **Überprüfen von Hypothesen**, die wiederum Behauptungen über die Beschaffenheit der Welt sind, an der empirischen, d.h. wahrnehmbaren Realität. Ein Hypothesentest hilft festzustellen, wie groß die **Irrtumswahrscheinlichkeit** dafür ist, eine bestimmte Hypothese fälschlicherweise anzunehmen oder abzulehnen. Voraussetzung für einen solchen Hypothesentest ist aber, dass die **Messdaten normal verteilt** sind, in einem Diagramm also eine **Normalverteilung** („Glockenkurve“) bilden, was eigentlich bei den meisten statistisch interessanten Verteilungen der Fall ist.

Das Praktische an einer solchen Normalverteilung ist nun: Die **Gesamtfläche unterhalb der Glockenkurve beträgt immer gleich 1**. Deshalb lässt sich anhand einer vorgegebenen Tabelle die

Wahrscheinlichkeit (in Prozent) ablesen, mit der ein bestimmter Messwert unterhalb einer bestimmten Grenze zu liegen kommt oder innerhalb eines bestimmten Intervalls liegt. Bspw. können Sie ohne Probleme eine Aussage darüber treffen, mit welcher Wahrscheinlichkeit ein Mensch zwischen 1,70m und 1,80m groß ist, indem Sie die entsprechende Fläche zwischen den Werten 1,70 und 1,80 *unterhalb* des Graphen der besagten Tabelle ablesen (und in Prozentwerte übersetzen).

Ohne hier im Rahmen dieser überblicksartigen Einführung auf die *Détails* des Verfahrens einzugehen, müssen Sie zunächst aber noch die Normalverteilung eines vorliegenden Datensatzes in eine sog. „**Standardnormalverteilung**“ überführen (deren **Mittelwert** ist, wie Sie schon wissen, **stets 0** und die **Standardabweichung** ist **stets 1**), da in der Regel die erwähnten Tabellen nur für Standardnormalverteilungen in den einschlägigen Lehrbüchern (oder auch im Internet) anzufinden sind. Diese Umwandlung geschieht mittels eines bestimmten mathematischen Verfahrens, welches aber keinerlei höhere Mathematikkenntnisse voraussetzt.

Doch wie läuft so ein **Hypothesentest** nun im einzelnen ab? Die Ausgangsfrage dabei lautet zunächst ganz grundsätzlich: Ist eine bestimmte Hypothese gültig oder muss sie durch eine andere ersetzt werden? Die bisherige bzw. bisher als gültig erachtete Hypothese (sozusagen der „Titelverteidiger“) wird üblicherweise als **Nullhypothese** bezeichnet (**H₀**), die **Alternativhypothese** ist sozusagen die „Herausforderin“ (**H₁**): Wer hat recht, welche Hypothese ist die Zutreffende? Im Kern funktioniert der Hypothesentest nun so, dass der **Mittelwert der H₀-Verteilung mit dem Mittelwert der H₁-Verteilung verglichen** wird. Behauptet H₁ beispielsweise, dass dieser größer ist, wird H₁ *verworfen* und H₀ beibehalten, wenn dem tatsächlich so ist.

Aber um wie viel bzw. wie stark müssen sich diese Mittelwerte voneinander unterscheiden, so dass Ihre Entscheidung, eine Hypothese anzunehmen oder aber abzulehnen, nicht nur dem Zufall geschuldet ist? Man spricht hierbei von der „**Irrtumswahrscheinlichkeit**“ oder auch vom „**Signifikanzniveau**“. Wann ist mit anderen Worten ein Ergebnis **signifikant** (also **nicht auf Zufall gründend**)? Wiederum vereinfacht gesagt: Wenn der Mittelwert der Alternativhypothese in den *äußeren Grenzbereichen* der Glockenkurve der H₀-Verteilung zu liegen kommt, wird dieser abgelehnt. Genauer gesagt: Ist die Irrtumswahrscheinlichkeit für eine fälschliche Ablehnung der Nullhypothese kleiner gleich 5% (liegt also in den äußersten 5% der Fläche unterhalb der Verteilungskurve) wird ein Testergebnis als signifikant bezeichnet. Ist die Irrtumswahrscheinlichkeit sogar kleiner gleich 1%, spricht man von einem **hoch signifikanten** Ergebnis.

Die Hypothesenprüfung liefert also wie gesagt letztlich Aussagen darüber, wie groß die sog. „**Irrtumswahrscheinlichkeit**“ ist, d.h.: wie groß die Wahrscheinlichkeit ist, dass die **Alternativhypothese fälschlicherweise angenommen oder aber abgelehnt** wird.

Falls Ihnen das etwas zu komplex anmutet, hier ein **Beispiel** zum besseren Nachvollziehen: Die Befragung aller Teilnehmer/innen einer der meist gebuchten Fortbildungsveranstaltungen („Grundlagen quantitativer Datenauswertung“) in einer von Ihnen geleiteten Volkshochschule ergab, dass nur 10% der Teilnehmer/innen mit der Veranstaltung *nicht* zufrieden waren. Sie haben nunmehr die Entscheidung zu treffen, ob es einer (nicht ganz billigen) didaktischen und medientechnischen Überarbeitung der Fortbildung bedarf. Damit Sie eine aussagekräftige Entscheidungsgrundlage gewinnen, beschließen Sie, 100 Teilnehmer/innen per standardisiertem Fragebogen zu befragen. Sie legen vorab folgendes fest:

Sind mindestens 10 Teilnehmer/innen mit der Qualität der Veranstaltung unzufrieden, soll diese überarbeitet und verbessert werden (*Hypothese H₀*), sind es weniger, bedarf es keiner größeren Änderungen (*Hypothese H₁*). Sagen wir, in Ihrer ersten Befragung erklärten 12 Teilnehmer/innen ihre Unzufriedenheit.

Da Sie sich der Zufälligkeit von Stichproben und darauf gründenden Ergebnissen im klaren sind, machen Sie die 2. Umfrage (n=100). Sie entschließen sich, mit einer 95%igen Wahrscheinlichkeit zufrieden zu sein (d.h. mit *einfacher*, nicht hoher *Signifikanz*), dass Ihr Befragungsergebnis *nicht* dem Zufall geschuldet ist. In dieser Befragung erklären 8 Fortbildungsteilnehmer/innen ihre Unzufriedenheit.

Was heißt das nun mit Blick auf Ihre beiden Hypothesen?

Bei Ihrer ersten Befragung erklärten wie gesagt 12 Teilnehmer/innen, mit der Qualität der Veranstaltung unzufrieden zu sein. Ihre H₀- Hypothese wäre damit angenommen. Jedoch könnte dies aufgrund des zufälligen Charakters Stichprobe auch täuschen: Dann nämlich, wenn der tatsächliche Anteil der nicht-zufriedenen in der **Grundgesamtheit** (alle Fortbildungsteilnehmer/innen) im

Gegensatz zur Stichprobe doch kleiner als 10% ist. Sie würden also bei der Annahme der Hypothese H_0 mit einer gewissen **Irrtumswahrscheinlichkeit** einen Fehler begehen.

Dieser Fehler berechnet sich aus der sog. **Ablehnungswahrscheinlichkeit**, die anhand der Verteilungsfunktion Ihrer Befragungsergebnisse ablesbar ist (nachdem diese, wie gesagt, auf mathematischem Wege in eine „Standardnormalverteilung“ überführt wurde). Anhand einer schon erwähnten Tabelle („die **Standardnormalverteilungstabelle**“), die sich in jedem Statistik-Lehrbuch bzw. im Internet findet, kann nun die Wahrscheinlichkeit abgelesen werden, dass es einem Stichprobenzufall geschuldet ist, dass weniger oder gleich 9 Teilnehmer/innen unzufrieden sind. Sie beträgt in unserem Fall 0,451. Was aber bedeutet dieser Wert? – Er bedeutet, dass es unter der Annahme, dass tatsächlich 10% aller Teilnehmer/innen unzufrieden sind, bei der Befragung mit 45,1% Wahrscheinlichkeit zu einer Fehlentscheidung hinsichtlich der Annahme von H_0 kommt. Da dieses Ergebnis natürlich nicht akzeptabel ist, können Sie H_0 verwerfen.

Natürlich gibt es über die vorangehend kurz vorgestellten statistischen Ansätze und die ihnen zugrundeliegenden Fragestellungen hinaus noch viele weitergehende Auswertungsmethoden quantitativer Sozialforschung: So ist es z.B. durchaus üblich, **mehrere Variablen auf einmal** miteinander in Bezug zu setzen („zu *korrelieren*“): Bspw. könnten Sie bei einer Befragung zum Thema „Politische Einstellungen von jungen Frauen“ deren verschiedene Antworten noch mit weitergehenden Faktoren wie Ausbildungsstand oder Wohnort (Stadt/Land) in Zusammenhang setzen. Hierfür gibt es sogenannte „**Multivarianz- und Diskriminanzanalysen**“, die jedoch den Rahmen eines ersten Überblicks an diese Stelle überschreiten.

Die vorangehenden Ausführungen zur beschreibenden und schließenden Statistik sollten Ihnen, wie eingangs schon erwähnt, lediglich einen aller ersten Eindruck davon vermitteln, wie Sie bei der Auswertung quantitativer Daten *im wesentlichen* vorzugehen haben. Dabei wurde Ihnen eine Auswahl der gängigsten Begriffe und Verfahrensweisen unterbreitet, die keinesfalls das gesamte Repertoire statistischer Verfahren abdeckt.

Vielmehr ist hier zumindest eines hoffentlich gelungen: Sie davon zu überzeugen, dass auch bei der Auswertung quantitativer Forschungsergebnisse im Grunde nur „mit Wasser gekocht wird“. Einer intensiveren Beschäftigung und Einarbeitung in das Thema (falls Ihr Forschungsprojekt größere und komplexere Datenmengen „abwirft“), sollte nun nichts mehr im Wege stehen.

Zusammenfassung

Die **beschreibende Statistik (deskriptive Statistik)** fasst empirische Daten anhand **aussagekräftiger Kennzeichen** zusammen. Gängige Verfahren sind dabei die **tabellarische Darstellung** von Daten, entweder in Form einer **Häufigkeitstabelle** oder mit Hilfe eines **Histogramms** bzw. **Polygons**. Die allerwichtigsten Kenngrößen der beschreibenden Statistik sind der

- **Modalwert/Modus** (der Messwert, der am häufigsten auftritt),
- der **Median** (er halbiert die Datenmenge in zwei gleich große Teile),
- das **arithmetische Mittel** (der „Durchschnitt“),
- die **Varianz** und die **Standardabweichung** (sie geben darüber Auskunft, wie weit die Daten um den Mittelwert „streuen“) und nicht zuletzt
- der **Korrelationskoeffizient** (der den Zusammenhang zwischen zwei Variablen charakterisiert).

Die **schließende Statistik (Inferenzstatistik)** stellt die Frage, ob wissenschaftliche Ergebnisse dem Zufall einer Stichprobe geschuldet sind oder ob es sich dabei vielmehr um wissenschaftliche Gesetzmäßigkeiten handelt. Entsprechend geht es auch darum, inwieweit eine Stichprobe Rückschlüsse auf die jeweilige **Grundgesamtheit**, aus der die jeweilige **Stichprobe** „gezogen“ wurde, zulässt; d.h.: inwieweit sie **repräsentativ** ist. Das wichtigste Verfahren hierbei ist der **Hypothesentest**, der die **Mittelwerte von Normalverteilungen** vergleicht und Aussagen darüber zulässt, ob die „alte“ Hypothese (**H_0**) oder die „Herausforderin“ (**H_1**) **signifikant** (oder gar **hoch signifikant**) anzunehmen oder abzulehnen ist.

- Bortz, Jürgen & Döring, Nicola (2006): Forschungsmethoden und Evaluation für Sozialwissenschaftler. 4. überarbeitete Auflage. Berlin: Springer.
- Bosbach, Gerd & Korff, Jens Jürgen (2011): Lügen mit Zahlen. Wie wir mit Statistiken manipuliert werden. 3. Auflage. München: Heyne Verlag
- Friedrichs, Jürgen (1990): Methoden empirischer Sozialforschung. 14. Auflage. Opladen: Westdeutscher Verlag.
- Kromrey, Helmut (2006): Empirische Sozialforschung. Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung. 11. Auflage. Stuttgart: Lucius & Lucius.
- Schnell, Rainer; Hill, Paul B. & Esser, Elke (2008): Methoden der Empirischen Sozialforschung. 8. Auflage. München: Oldenbourg.
- Urban, Klaus (1996): Statistik. Einführung in die statistische Methodenlehre. 4. Auflage. München/Wien: Oldenbourg.