

# MUFFINS: Statistik mit komplexen Datensätzen – Freizeitgestaltung und Mediennutzung von Jugendlichen

ROLF BIEHLER & KLAUS KOMBRINK, UNIVERSITÄT KASSEL STEFAN SCHWEYNOCH, FREIHERR-  
VOM-STEIN-GYMNASIUM BÜNDE

---

## 1. Einleitung

Mit diesem Aufsatz wollen wir einen komplexen Datensatz und dazugehörige Materialien vorstellen, die wir für computergestützten Stochastikunterricht entwickelt haben, der sich stark an den Ideen der Explorativen Datenanalyse orientiert (Biehler & Weber 1995, Biehler 1999), welche in jüngster Zeit auch unter dem Gesichtspunkt der Entwicklung statistischer Kompetenz und von Daten orientierten Zugängen zur Mathematik aufgegriffen wird (Borovcnik, Engel, Wickmann 2001). Beispielhafte Auswertungsmöglichkeiten sowie didaktische Ideen für neue Akzente im Stochastikunterricht werden vorgestellt. Ausführlichere Informationen und die Daten in verschiedenen Formaten findet man über unser Internetportal

[www.mathematik.uni-kassel.de/didaktik/muffins](http://www.mathematik.uni-kassel.de/didaktik/muffins).

Wir haben einen Fragebogen entwickelt, mit dem wir Schülerinnen und Schüler nach ihrer Mediennutzung und nach ihrem Freizeitverhalten befragen. Die bereits von 540 Schülerinnen und Schülern vorliegenden Daten können als authentische Beispieldaten im Unterricht verwendet werden. Didaktisch günstig ist es, wenn an der jeweiligen Schule mit unserem Fragebogen weitere Daten erhoben werden, so dass die Schülerinnen und Schüler auf persönliche Daten zurückgreifen können und diese mit anderen schon vorliegenden Daten vergleichen können. Über unser Portal steht unser Fragebogen online zur Verfügung. Interessierte Schulen können dies nutzen, um eine Online-Erhebung durchzuführen werden. Die Daten können dann im passenden Format zur Verfügung gestellt werden.

Erfahrungsgemäß ist die Konzipierung einer komplexen eigenen Fragebogenuntersuchung und die anschließende Datenerhebung und -eingabe so aufwendig, dass im Allgemeinen zu wenig Zeit für eine gründliche Analyse bleibt. Mit unserem Vorschlag, auf vorhandenen Daten und Fragebögen aufzubauen, aber weitere eigene Daten dazu zu erheben, glauben wir, dass ein vernünftiger didaktischer Kompromiss möglich ist.

Wir haben das Thema Medien und Freizeitgestaltung schon einmal in einer 8. Klasse zum Gegenstand des Unterrichts in Beschreibender Statistik gemacht und Ergebnisse dazu publiziert (Kombrink

1997, Biehler & Kombrink 1999, 1999a). Die Daten und der Fragebogen sind auch im Internet verfügbar<sup>7</sup>. Den Fragebogen haben wir gründlich überarbeitet und Schülerinnen und Schülern aus der 11. Klasse vorgelegt, die vorwiegend aus NRW stammen, um hiermit Materialien für den Unterricht in der 11. Klasse zu entwickeln, in der seit der letzten Richtlinienänderung die Beschreibende Statistik einen neuen Schwerpunkt bildet.

Unser Aufsatz ist aber auch für all diejenigen interessant, die in anderen Klassenstufen und Schulformen nach interessanten Datensätzen suchen, welche aus der Lebenswelt von Schülerinnen und Schülern stammen. Unser Projekt haben wir „MUFFINS“: *Medien- und Freizeitgestaltung für interessanten Stochastikunterricht* getauft.

Der Datensatz kann im Unterricht zu verschiedenen Zwecken eingesetzt werden: als Beispielmateriale für die Entwicklung von Begriffen, als Material für „Minianwendungen“ oder um komplexere „Entdeckungsreisen“ in die Daten im Sinne der Explorativen Datenanalyse zu unternehmen. Der neue Lehrplan in NRW unterstützt diesen didaktischen Ansatz besonders, und wir wollen zunächst genauer darauf eingehen.

1999 wurden in Nordrhein-Westfalen neue Richtlinien für die Oberstufe (Ministerium für Schule und Weiterbildung, Wissenschaft und Forschung, 1999) veröffentlicht. Obligatorisch enthalten diese für die Jahrgangsstufe 11 das Thema „Beschreibende Statistik“ und für die Jahrgangsstufen 12/13 die „Beurteilende Statistik“. Der Computereinsatz wird an vielen Stellen als sinnvoll herausgestellt, ist aber nicht verpflichtend festgeschrieben. Die Schulbuchverlage können nicht davon ausgehen, dass an allen Schulen der PC-Einsatz in Mathematik gesichert ist und müssen gewährleisten, dass ihre Bücher auch ohne Computer eingesetzt werden können. Untersucht man die drei am häufigsten in der gymnasialen Oberstufe Nordrhein-Westfalens eingesetzten Schulbücher für die Jahrgangsstufe 11, so stellt man fest, dass der Computereinsatz in der Beschreibenden Statistik gar nicht oder nur wenig un-

---

<sup>7</sup> <http://www.mathematik.uni-kassel.de/didaktik/mathelehren/mathelehren.htm>

terstützt wird (siehe Jahnke, T. u.a. 2000, Griesel, Postel 1999, Baum, M. u.a. 2000).

Um das Anforderungsniveau in der Beschreibenden Statistik auf ein für die Oberstufe angemessenes Niveau zu bringen, wird gefordert, mit großen und unübersichtlichen Datenmengen zu arbeiten. Datensätze mit nur zwei oder drei Variablen können diesem Anspruch kaum genügen. Untersucht man in unserem Datensatz den Zusammenhang zwischen zwei Variablen, so stößt man in der Regel sehr schnell auf weitere Variablen, die für den Zusammenhang relevant sein können. Das „Weiterfragen“ wird wesentlich erleichtert, wenn schon viele Variablen zur Verfügung stehen und nicht erst neu erhoben werden müssen.

Der von uns erhobene Datensatz bietet neben den vielfältigen Einsatzmöglichkeiten im Mathematikunterricht auch sinnvolle Inhalte zum fächerübergreifenden Arbeiten. Dies wird schon lange von den Richtlinien gefordert. Der thematische Rahmen unseres Datensatzes legt ein Fächerübergreifen von Sozialwissenschaften und Mathematik sowohl bei der Beschreibenden (Jahrgangsstufe 11) als auch bei der Beurteilenden Statistik (Jahrgangsstufe 12/13) nahe

## 2. Die „MUFFINS“-Umfrage zur Freizeit- und Mediennutzung

Analysiert man die in den Kapiteln zur Beschreibenden Statistik verwendeten Daten in den einschlägigen Unterrichtswerken für die Jahrgangsstufe 11, stellt man fest, dass weitgehend mit Daten aus realen Kontexten gearbeitet wird. Die verwendeten Daten haben aber kaum zusammenhängende inhaltliche Bezüge, was die Realisation von Ansprüchen und Intentionen der eingangs vorgestellten Richtlinien deutlich erschwert. Der vorliegende komplexe Datensatz überbrückt dieses Defizit: er ermöglicht neben Einzelanalysen auch eine längere abwechslungsreiche Auseinandersetzung mit einem zusammenhängenden Thema, das für die Jugendlichen interessant und relevant ist. Auch unter motivationalen Gesichtspunkten, so unsere bisherigen Erfahrungen mit dem Datensatz, erscheint das Potenzial, einzelne Ergebnisse in Richtung differenzierter und facettenreicher Gesamtergebnisse zu kumulieren, als wichtiger Aspekt.

### 2.1 Die Erhebung der Daten

Inhaltlicher Schwerpunkt des Datensatzes ist die Freizeit- und Mediennutzung von Jugendlichen. Wir haben unsere Daten zunächst in Klasse 11 erhoben, also von Schülern und Schülerinnen, die

i.d.R. zwischen 16 und 18 Jahre alt sind. Bei der Entwicklung des Fragebogens haben wir versucht, möglichst viele in diesem Kontext für die Mädchen und Jungen relevanten Aspekte anzusprechen. Alle Zeit- und Häufigkeitsangaben von Aktivitäten sind Angaben und Schätzungen der Jugendlichen selber und nicht objektiv gemessene Größen. Die drei in Bezug auf die Erhebung der Daten wesentlichen Grundsätze Objektivität, Reliabilität, Validität müssen und können thematisiert werden. Bei der Interpretation von Ergebnissen muss man sich bewusst sein, dass alle Angaben von den Schülern selber geschätzt wurden. Der Fragebogen kann deshalb auch immer dazu verwendet werden, diese drei Aspekte im unterrichtlichen Gespräch zu thematisieren und die Umfrage kritisch zu beleuchten. Das für uns primäre Ziel, nämlich die Produktion eines attraktiven und unterrichtlich verwendbaren Datensatzes, sollte dabei aber nicht aus den Augen verloren werden.

Die folgende Übersicht umreißt das berücksichtigte Spektrum des Themas:

#### *Beliebte Freizeitaktivitäten*

Wie häufig gehen Jugendliche verschiedensten Aktivitäten nach, z.B. Disco-, Kino-, Kneipenbesuch, aber auch Teilnahme an ehrenamtlichen Aktivitäten etc.

#### *Fernsehen*

Wie lange und wofür nutzen Jugendliche ein TV-Gerät? Welche Sendungen favorisieren sie etc.?

#### *Computer und Internet*

Wann, wofür und wie lange verwenden die Jugendlichen einen Computer? Haben sie Internetzugang? Welche Internetdienste sind beliebt (E-mail, Chat-ten etc.)?

#### *Telefonieren*

Fragen zur Dauer und Häufigkeit von Telefongesprächen. Besitzen die Mädchen und Jungen eigene Anschlüsse bzw. Handys. Beteiligen sie sich an Kosten etc.?

#### *Sport*

Welche Sportarten werden ausgeübt, und wieviel Zeit wird dafür verwandt? Welche Jugendliche gehören einem Sportverein an?

#### *Musik*

Wie lange hören Jugendliche Musik? Musizieren sie selbst? Welche Instrumente spielen sie?

#### *Jobben*

Haben die Befragten Nebenjobs? Wie lange arbeiten sie in diesen Jobs? Wieviel Geld verdienen sie etc.?

#### *Lesen*

Fragen nach Lesegewohnheiten, z.B. Lesedauer,

welche Formate sind beliebt? Wie häufig und wofür werden Tageszeitungen genutzt?

Wie bereits vorne erwähnt, sind in das Design des Fragebogens wesentliche Erfahrungen aus dem Unterrichtsprjekt zur Freizeit- und Mediennutzung in einer 8. Klasse (Kombrink 1997, Biehler & Kombrink 1999) eingearbeitet. Dieses betrifft insbesondere das differenzierte Zeitraster, mit dessen Hilfe die Jugendlichen zunächst die für verschiedene Aktivitäten überhaupt zur Verfügung stehende Freizeit für jeden Tag der Woche ermitteln müssen. Die alte Umfrage hatte gezeigt, dass die Schülerinnen und Schüler nicht in der Lage sind, diese Zeit sinnvoll abzuschätzen. Durch die durchzuführende Rechnung erhalten die Schülerinnen und Schüler einen Orientierungswert über die ihnen zur Verfügung stehende Freizeit, die sie dann auf verschiedene Aktivitäten verteilen können. Nach unserem Eindruck hat dieses Verfahren zu deutlich realistischeren Werten geführt.

Ferner haben wir versucht, die Inhalte des Fragebogens bezogen auf die Gewohnheiten der Jugendlichen thematisch zu aktualisieren bzw. zu differenzieren sowie altersspezifischer zu gestalten.

Auf Grundlage dieses Fragebogens konnten wir dann eine Umfrage unter Schülerinnen und Schülern des Jahrgangs 11 im Frühjahr 2000 an neun überwiegend nordrhein-westfälischen, aber auch niedersächsischen und Hamburger Schulen durchführen. Es handelt sich also nicht um eine Zufallsstichprobe aus einer klar definierten Grundgesamtheit. Auch das muss man bei der inhaltlichen Interpretation von Analyseergebnissen berücksichtigen. Die Ergebnisse sind im Datensatz MUFFINS zusammengefasst

## 2.2 Der Datensatz „MUFFINS“

Der entstandene Datensatz besteht aus mehr als 140 Variablen. Durch das Design des Fragebogens war sichergestellt, dass sowohl qualitative als auch quantitative Merkmale in ausreichender Zahl enthalten sind, so dass alle relevanten Inhalte der Beschreibenden Statistik mit den vorhandenen Merkmalen thematisiert werden können. Darüber hinaus bieten sich die Variablen des Datensatzes auch für anwendungsorientierte und in großem Maße auf Schüleraktivität zielende Zugänge zur Beurteilenden Statistik an, z.B. durch *Randomisierungstests* (das ist ein Thema, das weiter unten genauer dargestellt wird).

Die Übersicht kennzeichnet kurz die verschiedenen in der Statistik auftauchenden Variablentypen und exemplifiziert diese mit konkreten Fragestellungen aus der Umfrage.

### Qualitative Merkmale

Bei qualitativen Merkmalen ist es nur möglich, die Variablen in verschiedene Kategorien einzuordnen, wir sprechen deshalb auch von kategorialen Variablen. Innerhalb qualitativer Variablen unterscheidet man nominale und ordinale Skalenniveaus.

*Nominale* Daten haben keine natürlich vorgegebene Ordnung der Antwortkategorien. Alle Kategorien stehen wertfrei nebeneinander, die Reihenfolge der Aufzählung ist beliebig. Das klassische Beispiel für dieses Skalenniveau ist die Variable „Geschlecht“, ein weiteres Beispiel liefert der folgende Ausschnitt aus dem Fragebogen:

Wann musst du normalerweise zu Hause sein?

während der Woche:

- egal, es ist mir während der Woche keine Zeit vorgeschrieben.
- ich muss während der Woche um \_\_\_\_\_ Uhr zuhause sein.
- ich spreche es jeden Abend individuell mit meinen Eltern ab

Es gibt in diesem Beispiel drei Antwortkategorien, die wir mit „nein“, „ja“ und „Absprache“ bezeichnen, um gewissermaßen die Antwort auf die Frage „Hast du eine Ausgehbeschränkung?“ zu kodieren. Die Kategorien haben keine natürliche Reihenfolge. Bei der Antwort „ja“ kann eine Uhrzeit angegeben werden, die wir in einer eigenen Variablen erfassen.

Bei *ordinal* skalierten Daten ist dagegen eine bestimmte Reihenfolge vorgegeben und inhaltlich wesentlich. Wenn z.B. das Interesse an Fernsehsendungen erfragt wird, so kann man die Antwortkategorien nach wachsender oder fallender Beliebtheit sortieren. Wichtig ist die Reihenfolge der Kategorien auch bei der Frage, wie häufig die Jugendlichen bestimmte Freizeitaktivitäten ausüben. Im Beispiel dazu sind die Kategorien danach sortiert, mit welcher Häufigkeit bestimmte Beschäftigungen ausgeübt werden:

|  | fast jeden Tag | 2-3 mal pro Woche | etwa einmal pro Woche | 1-2 mal pro Monat | seltener | nie |
|--|----------------|-------------------|-----------------------|-------------------|----------|-----|
| mit Freunden Essen oder Trinken gehen (in Kneipen, Cafés o.a.)                 | 1              | 2                 | 3                     | 4                 | 5        | 6   |
| bei Freunden übernachten   | 1              | 2                 | 3                     | 4                 | 5        | 6   |
| Rumhängen mit Freunden oder Bekannten  | 1              | 2                 | 3                     | 4                 | 5        | 6   |
| Karten-, Gesellschafts- oder Fantasyspiele mit Freunden (keine Computerspiele) | 1              | 2                 | 3                     | 4                 | 5        | 6   |
| Beteiligung in Parteien oder Bürgerinitiativen                                 | 1              | 2                 | 3                     | 4                 | 5        | 6   |

### Quantitative Merkmale

Quantitative Merkmale werden auf einer kontinuierlichen metrischen Skala gemessen. Sie werden deshalb auch *metrische* oder *kontinuierliche* Daten genannt. Der Name *Zahlvariable*, abgeleitet vom englischen *number*, wird ebenso benutzt. Auch hier werden wieder zwei Arten unterschieden, die beide im Datensatz enthalten sind.

*Intervallskalierte* Daten besitzen gleiche Differenzen zwischen zwei Werten auf der Skala, allerdings existiert kein wirklicher Nullpunkt. Die von den Jugendlichen in der Umfrage genannten Aufstehzeiten stehen hierfür: Man kann nur durch Differenzenbildung Unterschiede zwischen zwei Werten sinnvoll beschreiben, z.B. bei den Aufstehzeiten.

*Wie viel Zeit bleibt dir wöchentlich für verschiedenste Freizeitaktivitäten? Eine Hilfestellung zur Ermittlung dieser Zeit bietet der folgende Fahrplan.*

1. Notiere zunächst, wann du üblicherweise aufstehst, aus der Schule kommst und zu Bett gehst. Falls du nach der Schule nicht nach Hause gehst, sondern einer anderen Aktivität nachgehst, gib den Zeitpunkt an, an dem du den entsprechenden Ort erreichst.  
Gehst du nach Mitternacht zu Bett, trage die Zeit trotzdem beim vorangegangenen Tag ein.

**Runde alle Zeiten auf Viertelstunden!**

|                                  | Mo | Di | Mi | Do | Fr | Sa  | So  |
|----------------------------------|----|----|----|----|----|-----|-----|
| Aufstehen                        |    |    |    |    |    |     |     |
| aus der Schule nach Hause kommen |    |    |    |    |    | --- | --- |
| zu Bett gehen                    |    |    |    |    |    |     |     |

Diese Tabelle dient zur Ermittlung der „*Nettozeit*“, welche jeweils für Freizeit und andere Aktivitäten zur Verfügung steht.

Bei *quotientenskalierten* Daten drücken gleiche Quotienten von zwei Werten auch gleich große Unterschiede aus, z.B. ein Jugendlicher mit 10 Std. Sport pro Woche treibt doppelt so lange Sport wie derjenige, der nur 5 Std. angibt:

Wir möchten nun wissen, wie viele Stunden deiner wöchentlichen Nettozeit von dir für die folgenden Aktivitäten verwendet wird.

(Hinweis: „gezielt“ soll heißen, nicht zur Untermalung oder nebenbei!)

|   | Stunden pro Woche |   | Stunden pro Woche |                            | Stunden pro Woche |
|---|-------------------|---|-------------------|----------------------------|-------------------|
| <i>gezielt ein TV Gerät nutzen</i>                            |                   | <i>Sport ausüben (ohne Sportunterricht)</i> |                   | <i>Helfen zu Hause</i>     |                   |
| <i>gezielt Musik hören (ohne Disco)</i>                       |                   | <i>Musizieren/ Singen</i>                   |                   | <i>Jobben</i>              |                   |
| <i>gezielt einen Computer nutzen (nicht für Hausaufgaben)</i> |                   | <i>Lesen (ohne Schulbücher)</i>             |                   | <i>Hausaufgaben machen</i> |                   |

Für die Analyse der quantitativen Variablen des Datensatzes, die Stundenangaben für Freizeitaktivitäten enthalten, sollte man im Unterricht in jedem Fall berücksichtigen, dass alle Zeitangaben bei der Befragung auf Viertelstunden gerundet werden sollten. Dieses verfälscht in verschiedenen Kontexten die Ergebnisse. Eine Möglichkeit hiermit konkret umzugehen, stellen wir im Kontext unseres Beispiels zur Beurteilenden Statistik, Stichwort „scrambling“, vor.

### 3. Exemplarische Fragestellungen und Auswertungen

#### 3.1 Fragestellungen und Überblick

Der Datensatz bietet eine Fülle von Themen; einige komplexe und offene Fragestellungen und Entdeckungsreisen, die man am Ende einer Unterrichtseinheit unternehmen könnte, sind beispielsweise:

- (1) Welche Hauptunterschiede gibt es im Freizeitverhalten zwischen Jungen und Mädchen?
- (2) In welchem Umfang und zu welchen Zwecken nutzen die Schülerinnen und Schüler ein TV-Gerät?
- (3) Thema Jobben: Wie viele jobben? In welchen Branchen? Was wird tendenziell verdient? Gibt es einen Zusammenhang zwischen dem Verdienst und dem Geschlecht oder der Branche? Hängt der Verdienst vom Geschlecht, von der Branche ab?
- (4) Computer und Internet: Wie lange? Wie verbreitet überhaupt? Wozu im Einzelnen? Von welchen Faktoren abhängig? (Geschlecht, Verfügbarkeit).

In Abschnitt 4 stellen wir beispielhaft eine komplexe Entdeckungsreise zum Thema (3) vor. Sowohl bei den Entdeckungsreisen als auch bei den Minianwendungen können die Schülerinnen und Schüler eigene Fragestellungen einbringen und untersuchen. Die Reichhaltigkeit des Datensatzes erlaubt das „Weiterfragen“ und zumindest teilweise das Beantworten, ohne dass weitere Daten erhoben werden. In dem Unterrichtsexperiment der 8. Klasse etwa (Kombrink 1997, Biehler & Kombrink 1999) stellten die Schüler keinen Unterschied in der Zeitdauer des Lesens fest zwischen Fernsehbesitzern und Nicht-Fernsehbesitzern. Die Hypothese der Schülerinnen und Schüler, dass Nicht-Fernsehbesitzer tendenziell mehr lesen als Fernsehbesitzer, konnte dann in modifizierter Weise „gerettet“ werden. Sie konnten nämlich einen deutlichen Unterschied zwischen diesen Gruppen hinsichtlich des Lesens von *Büchern* nachweisen. Diese Möglichkeit des Weiterfragens ist bei den üblicherweise für den Schulgebrauch vorgeschlagenen Datensätzen, die nur aus wenigen Variablen bestehen, nicht gegeben.

In diesem Abschnitt geht es allerdings zunächst um begrenzte Anwendungen. Beispielhafte Themen für „Minianwendungen“ könnten sein:

1. Wie viel Freizeit (Nettozeit) steht den Schülerinnen und Schülern in der untersuchten Gruppe zur Verfügung und wie kann sich ein einzelner Schüler darin einordnen? (statistische Verteilung)
2. Wie unterscheiden sich Jungen und Mädchen hinsichtlich des Shopping-Verhaltens? (Vergleich von Verteilungen)
3. Wie unterscheidet sich der Umfang der Computernutzung zwischen Computerbesitzern und denen, die keinen eigenen Rechner haben? (Vergleich von Verteilungen)

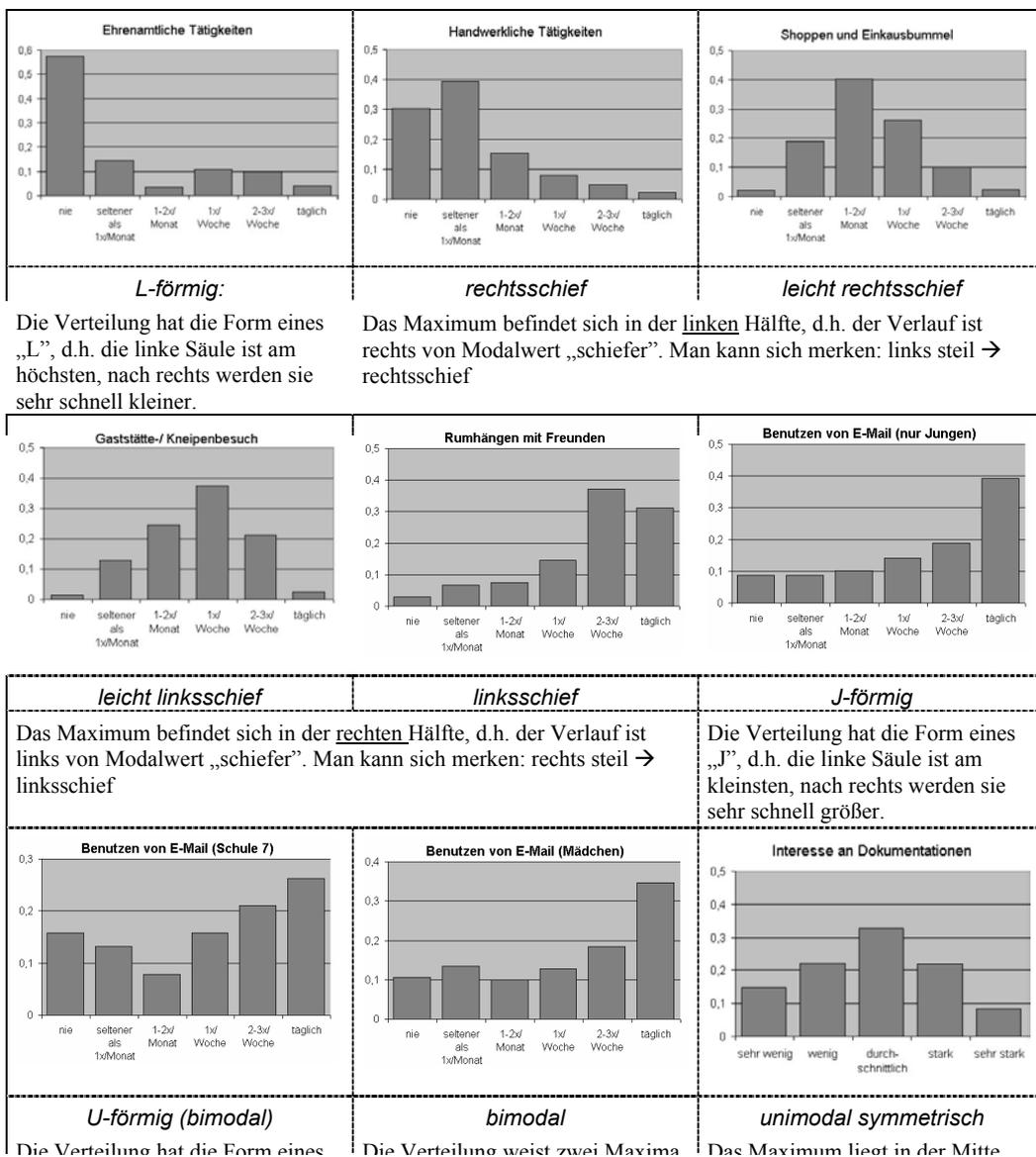
4. Stehen diejenigen, die freitags spät ins Bett gehen, tendenziell auch am Samstag später auf? (Korrelation und Regression).

Die Fragestellungen haben wir den großen thematischen Blöcken der Beschreibenden Statistik zugeordnet. Zur Beantwortung werden idealerweise alle Methoden, Graphiken, Begriffe herangezogen, die gelernt wurden, z. B. bei Verteilungen Mittelwerte, Streuungsmaße, andere Kennzahlen, Histogramme, Boxplots, evtl. kumulative Häufigkeitsdiagramme.

werden, sondern in dem eine sinnvolle Zusammenfassung und Interpretation gegeben wird.

Den Aspekt des Vergleichens von Verteilungen halten wir für zentral, da hierüber zahlreiche interessante Fragen aufgeworfen werden und sich das Repertoire der Beschreibenden Statistik erst richtig funktional bewährt. Vergleiche beziehen sich dabei zunächst nur auf die untersuchte Schülergruppe, eine Verallgemeinerung muss zunächst offen bleiben. Wir stellen aber unten eine Möglichkeit vor, einen Schritt in Richtung Beurteilende Statistik zu gehen.

## Klassifizierung von Verteilungsformen



Eine Antwort erfordert einen „statistischen Report“, in dem nicht einfach Computerausdrücke kopiert

Auch in Einführungsphasen kann man die Daten als Beispiele benutzen. Aus Platzmangel können wir auf diese Möglichkeiten nur exemplarisch eingehen.

### 3.2 Software zur Datenanalyse

Um solche Fragestellungen flexibel untersuchen zu können, bedarf es einer Software, in die sich Schüler leicht einarbeiten können und die eine interaktive graphische Datenanalyse unterstützt. Die Verteilungsgraphiken müssen auf Knopfdruck verfügbar, ein Gruppenvergleich soll leicht möglich sein, indem man eine kategoriale Variable als gruppierende Variable nutzt, die Analyse muss flexibel auf Teilmengen beschränkbar sein (Filterfunktionen), einzelne Ausreißer müssen sich leicht identifizieren lassen. Der Output soll gezielt gewählt werden können, multiple Tabellen und Graphiken sollen den Vergleich zwischen verschiedenen Darstellungen ermöglichen. Eine wesentliche Anforderung ist es auch, dass eine Software angemessen mit „missing values“ umgehen kann, die in allen realen Datensätzen vorkommen. Das weit verbreitete EXCEL hat hier verschiedene Schwächen. Auch aus diesem Grunde haben wir die Software MEDASS LIGHT entwickelt, welche viel Flexibilität für elementare Anwendungen liefert und auch von Schülerinnen und Schülern der Sekundarstufe I leicht erlernbar ist. Die Software FATHOM bietet wesentlich mehr Möglichkeiten, auch zum Experimentieren mit Graphiken und Methoden und für die Simulation und Beurteilende Statistik. Wir streben an, eine deutsche Adaptation zu entwickeln. Die folgenden Graphiken und Analysen sind im Wesentlichen mit MEDASS LIGHT oder FATHOM gemacht worden.

### 3.3 Beispiele für Verteilungsformen

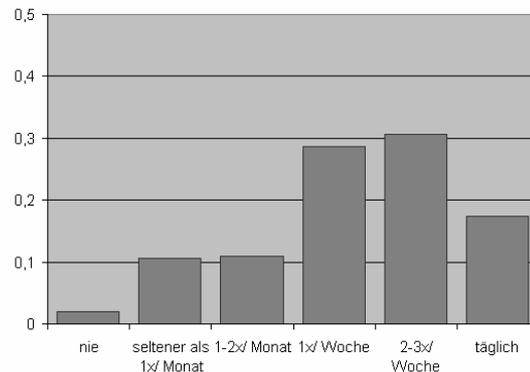
Ein Thema des Unterrichts sollten die vielfältigen Formen der Verteilungen sein, die in der Welt vorkommen. Unsere Ordinalvariablen zur Häufigkeit von Freizeitaktivitäten liefern hier einen interessanten Mikrokosmos (siehe vorhergehende Seite). Schülerinnen und Schüler können daran lernen, dass auch die Form der Verteilung interessante Unterschiede zwischen den verschiedenen Aktivitäten angibt. Ein Vergleich alleine hinsichtlich des Modalwertes wäre überhaupt nicht instruktiv.

### 3.4 Vergleich Mädchen und Jungen – Beispiele „Sport“ und „Disco“

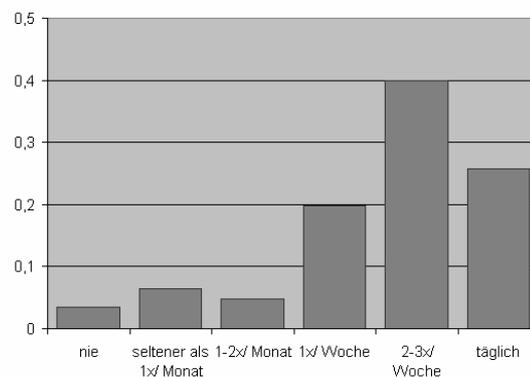
*Wie oft gehen Mädchen verglichen mit Jungen sportlichen Betätigungen nach?*

Zur Analyse dieser Frage ziehen wir je ein Häufigkeitsdiagramm (relative Häufigkeiten) für die Mädchen und Jungen heran.

Wie oft treibst du Sport? - Mädchen



Wie oft treibst du Sport? - Jungen



Wenn man die Diagramme Säule für Säule vergleicht, erkennt man, dass bei den Mädchen die erste Säule niedriger, die 3 anschließenden Säulen höher sind als bei den Jungen und die beiden rechten niedriger. Wichtig ist es, die Verteilungsformen miteinander zu vergleichen: Bei den Jungen hat sich die Verteilung mehr zu den Kategorien 2-3x/Woche und täglich verschoben. Dies ist der wesentliche Unterschied. Ein paar Prozente sind aber auch gleichsam gegenläufig in die Kategorie nie gewandert.

Somit könnte man festhalten, dass die Mädchen aus dieser Stichprobe tendenziell seltener Sport treiben als die Jungen, als zusammenfassendes Kriterium könnte man den jeweiligen Anteil nehmen, der mehr als 2-3x die Woche Sport treibt. Wenn man es für wichtig hält, kann man noch auf die Ausnahme verweisen und hinzufügen, dass es aber etwa doppelt so viele Jungen als Mädchen gibt, die nie Sport treiben. Die Formulierung „tendenziell“ schlagen wir generell für Gruppenvergleiche bei quantitativen oder ordinalen Variablen (wie hier) vor, um statistische Unterschiede zunächst qualitativ auszudrücken, um sie ggf. im nächsten Schritt geeignet zu präzisieren.

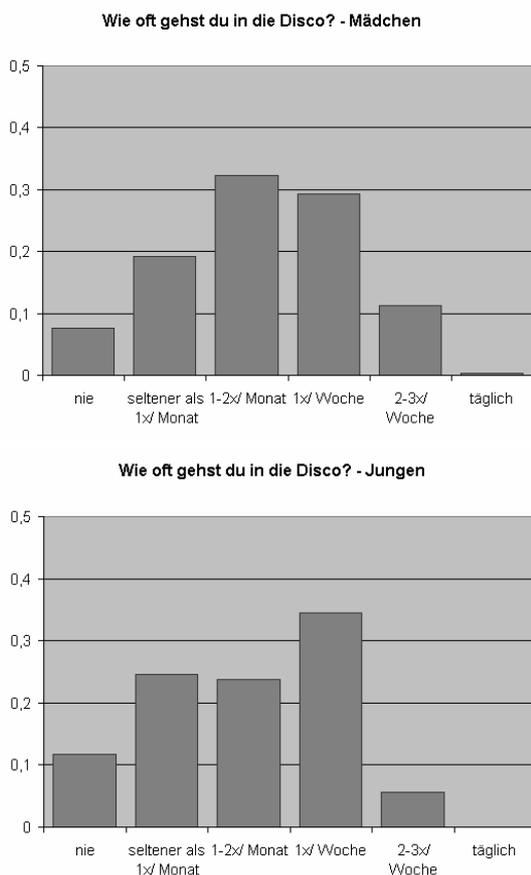
Der Datensatz bietet an dieser Stelle zahlreiche Möglichkeiten für vertiefende Analysen, z.B.:

Ist der (in Wochenstunden auch quantitativ angegebene) Zeitaufwand für Sport bei den Jungen tendenziell auch größer als bei den Mädchen?

- Lassen sich entsprechende Unterschiede auch für das Besuchen von Sportveranstaltungen, für das Interesse an Sportsendungen im Fernsehen oder Sportartikeln in Zeitungen feststellen?
- Gibt es bestimmte Schulen, die dieser allgemeinen Tendenz widersprechen?
- Gibt es einen Zusammenhang zwischen Häufigkeit des Sporttreibens und Mitgliedschaft im Verein?

Wie oft besuchen Mädchen verglichen mit Jungen eine Diskothek?

Erstellt man zu dieser Frage die gleichen Graphiken wie oben, so ergibt sich kein eindeutiges Bild:

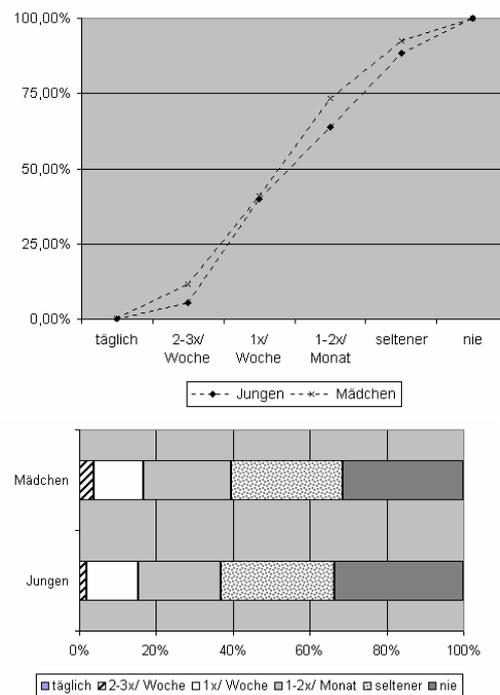


Man sieht, dass die 2. und 4. Säule bei den Jungen höher ist, die 1., 3. und 5. bei den Mädchen. Vergleicht man jedoch nicht die einzelnen Kategorien miteinander, sondern die kumulativen Häufigkeiten, so ergibt sich ein viel deutlicheres Bild (s. u.). Der erste Punkt der Graphik gibt an, wie viele Schülerinnen bzw. Schüler täglich die Disco besuchen, der zweite, wie viele täglich oder 2-3x pro

Woche, der dritte täglich, 2-3x pro Woche oder 1x pro Woche die Disco besuchen. Die „Kurve“ der Mädchen liegt immer oberhalb der der Jungen. Somit kann man festhalten, dass die befragten Mädchen tendenziell etwas häufiger in die Disco gehen als die Jungen. Der Vergleich der kumulativen Häufigkeiten, mit der man entscheiden kann, ob eine Variable in einer Gruppe „statistisch größer“ ist, ist hier die angemessene Präzisierung (vgl. auch Biehler 2001).<sup>8</sup>

Leider kann man in keinem der von uns genannten Programme eine solche Graphik ohne Umwege erzeugen. Man muss immer erst die kumulativen Häufigkeiten errechnen und kann sie dann darstellen.

Das zusätzlich abgebildete Blockdiagramm ermöglicht auch diesen Vergleich. Man erkennt, dass bei den Jungen die aneinander gereihten Blöcke jeweils links von denen der Mädchen enden. Also liegen die Mädchen in der Summe stets unterhalb der Summe der Jungen. Zusätzlich kann man in dieser Darstellungsform aber auch noch die einzelnen Kategorien paarweise miteinander vergleichen. Ein großer Nachteil dieser Darstellungsform ist es, dass die Verteilungsformen nicht mehr zu erkennen sind.



Auch bei dieser Auswertung schließen sich wieder zahlreiche weitere Fragestellungen an:

- Gehen häufige Discobesucher tendenziell später ins Bett und stehen sie später auf?

<sup>8</sup> Bei der zuvor betrachteten Variable zum Sporttreiben würden sich die kumulativen Kurven am Anfang schneiden, wegen der Gegenläufigkeit bei der Kategorie *nie*.

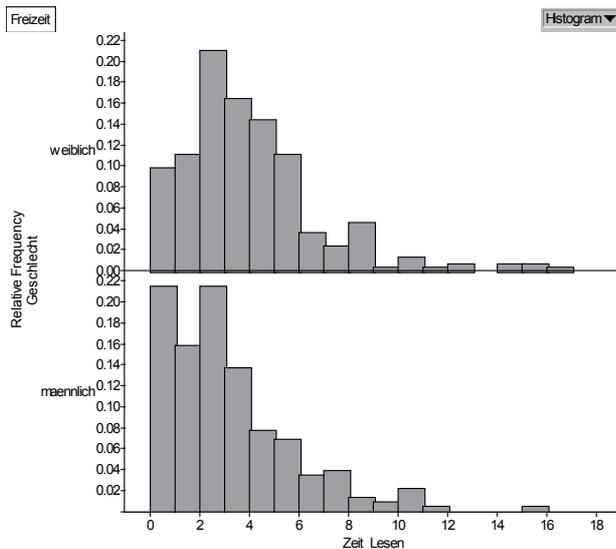
- Besuchen die Schülerinnen und Schüler, die von ihren Eltern eine Zeit vorgeschrieben bekommen, wann sie zu Hause sein müssen, tendenziell seltener die Disco?
- Besteht bei häufigen Discogängern tendenziell ein größeres Interesse an Musiksendungen?
- Machen häufige Discobesucher tendenziell weniger Hausaufgaben?

Zusammenfassend kann man festhalten, dass viele interessante Fragestellungen schon durch Einsatz elementarster Darstellungsformen beantwortet werden können. Für einige nicht sehr offensichtliche Unterschiede muss der Werkzeugkasten mit Analysetools erweitert werden. Auf jeden Fall ergeben sich grundsätzlich weitere Fragestellungen, die eine vertiefende Analyse anregen.

Zur Wahl der Software ist anzumerken, dass die Auszählung der Häufigkeiten mit EXCEL nur sehr umständlich zu leisten ist, hier bieten Statistiktools wie MEDASS LIGHT und FATHOM einen deutlich größeren Komfort.

### 3.5 Vergleich von Verteilungen bei kontinuierlichen Variablen

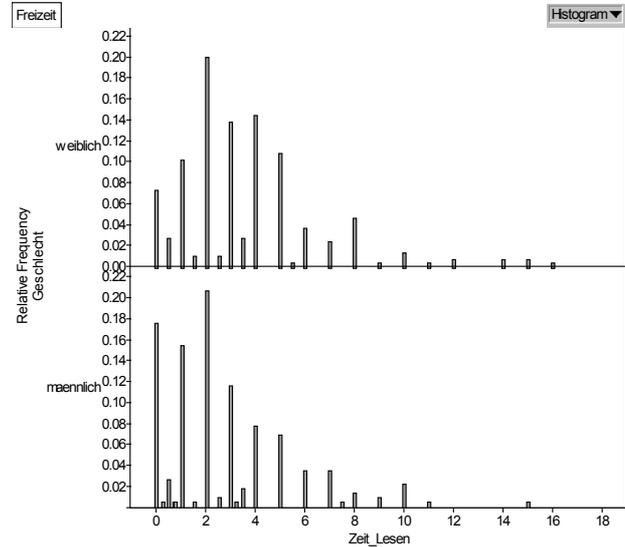
Wir untersuchen beispielhaft die Frage, wie sich Jungen und Mädchen hinsichtlich der wöchentlichen Lesezeit unterscheiden. Eine solche „Analyse nach Gruppen“ soll in einer Software einfach zu erreichen sein (sehr gut in FATHOM und MEDASS LIGHT, sehr umständlich in EXCEL). Wir sehen uns zunächst Histogramme mit relativen Häufigkeiten an. Gemäß der Devise „Look at the data“ vermeiden wir zu frühe Zusammenfassungen, z.B. eine Reduktion auf Mittelwerte, ohne dass man sich Klarheit über die Verteilung als Ganzes verschafft hätte.



Wir sehen, dass die Verteilung der Mädchen verglichen mit der der Jungen ihren Schwerpunkt weiter

rechts besitzt: Tendenziell lesen Mädchen mehr. Wir nehmen einen einzelnen getrennt liegenden Wert bei 15 Std. bei den Jungen wahr, ferner mehrere hohe Werte bei den Mädchen. Auffällig ist die erste hohe Säule bei 0 bis 1 Std. (Jungen)

Wir wählen die feinste Klasseneinteilung



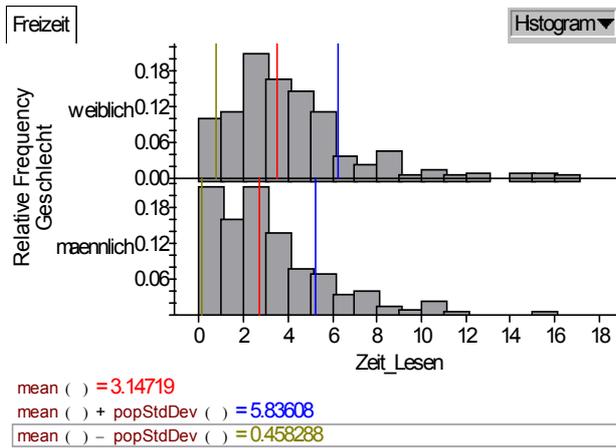
und sehen, dass fast 18% der Jungen gar nicht lesen, bei den Mädchen sind es nur gut 7%. Das ist bereits eine interessante Detailinformation. Weiter fallen sog. populäre Werte auf: Es werden im Wesentlichen ganze Stunden als Antwort angegeben.

Unterschiede in den Verteilungen kann man mit Kennzahlen präziser beschreiben. Wir wählen zunächst die traditionellen Maße arithmetisches Mittel und Standardabweichung.

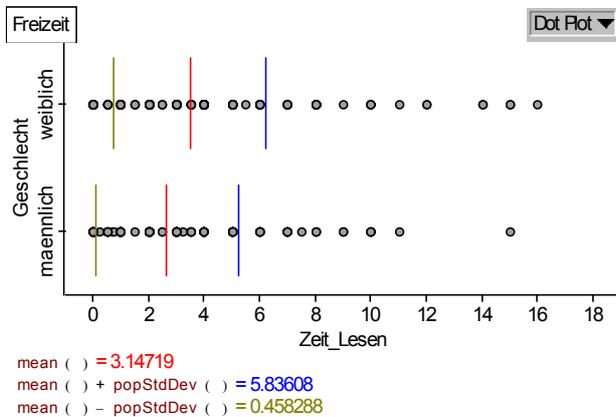
|          | arith. Mittel | Standardabweichung | Anzahl |
|----------|---------------|--------------------|--------|
| männlich | 2,69          | 2,54               | 232    |
| weiblich | 3,5           | 2,74               | 301    |
| Gesamt   | 3,15          | 2,69               | 533    |

Wir sehen, dass die Mädchen „im Schnitt“ etwa 0,8 Stunden pro Woche mehr lesen. Die Streuung in der Mädchengruppe ist etwas größer. Um die graphische Bedeutung dieser Kennzahlen zu erkennen, kann man die Histogramme damit anreichern<sup>9</sup>.

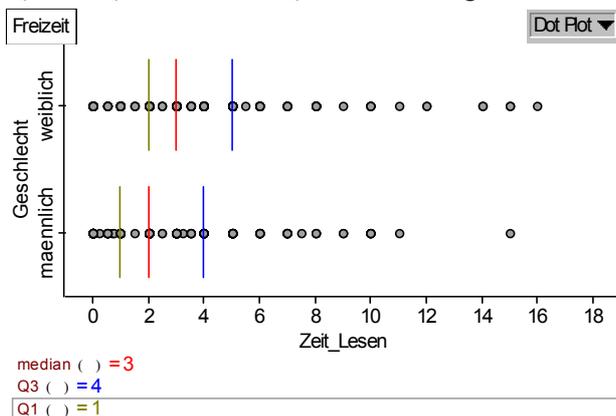
<sup>9</sup> Die Screenshots in diesem Abschnitt sind mit FATHOM gemacht. Die Software erlaubt das Einzeichnen von Linien, deren Lage mit einem „Formeleditor“ definiert werden kann.



Für den übersichtlichen Vergleich, vor allem wenn man noch mehr Gruppen hat, kann es sinnvoll sein, ein eindimensionales Streudiagramm (Dotplot) mit angereicherten Kennzahlen anzugeben.

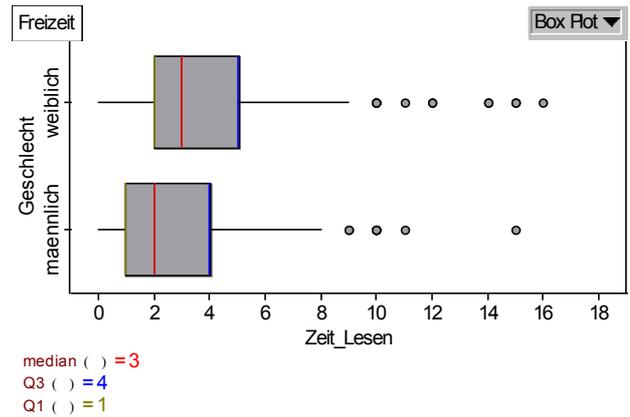


Verwendet man die gegenüber Ausreißern robusten Kennzahlen Median, unteres Quartil Q1 und oberes Quartil Q3 - also die 3 Quartile -, so ergeben sich:



Die Werte der Quartile sind bei Mädchen um etwa 1 Std. zu höheren Werten verschoben: Die Verteilung hat sich insgesamt um 1 Std. verschoben. An der Quartilgraphik erkennt man deutlich die Schiefe der Verteilung: rechts des Medians liegen die Daten weniger dicht als links vom Median. Die Verwendung von „plus/minus eine Standardabweichung“ stülpt den Daten eine Symmetrie in der Abweichung von der Mitte über, die gar nicht vorhanden ist. Eine zusammenfassende Darstellung dafür ist

der Boxplot, den wir hier in der Variante verwenden, dass Ausreißer getrennt dargestellt werden, wenn sie außerhalb der Zäune liegen.<sup>10</sup>



Wir haben hier ausführlich auch „Vorformen“ oder Alternativen zum Boxplot thematisiert, die man bei der Entwicklung der neuen Darstellungsform mitbenutzen kann und bei Bedarf auch in der realen Analyse. Als wesentliche Verteilungsunterschiede können wir mit der Boxplotdarstellung festhalten, dass die durch den Quartilabstand gemessene Streuung (Länge der Box) gleich ist, im Wesentlichen ist die Verteilung der Mädchen um eine Stunde zu höheren Werten verschoben. In diesem Sinne könnten wir festhalten, dass in der Gruppe der Mädchen tendenziell eine Stunde mehr in der Woche gelesen wird. Beide Gruppen enthalten einige Ausreißer (Vielleser). Aus den Histogrammen halten wir die zusätzliche Information fest, dass fast 18% der Jungen gar nicht lesen, bei den Mädchen sind es nur gut 7%. Als weiterführende Frage könnte man untersuchen, welche zusätzlichen Besonderheiten die nicht lesenden Jungen noch haben.

### 3.6 Beurteilende Statistik mit Randomisierungstests

An den Verteilungsvergleich schließt sich die natürliche Frage an: Wie zuverlässig sind denn die festgestellten Unterschiede zwischen Jungen und Mädchen? Wir haben ja nur eine Stichprobe vorliegen; können wir mit einer gewissen Sicherheit schließen, dass es auch einen vergleichbaren Unterschied in der Gesamtgruppe aller Elftklässler in NRW gibt? Um einen statistisch abgesicherten Vergleich nach höchsten Ansprüchen durchführen zu können, hätten wir eine Zufallsstichprobe aus

<sup>10</sup> Die Zäune sind wie folgt definiert

$$f_u = Q_1 - 1.5 \cdot (Q_3 - Q_1), f_o = Q_3 + 1.5 \cdot (Q_3 - Q_1).$$

Dies ist aber nicht mehr als eine erste Faustregel für die Darstellung von weit weg liegenden Werten. Ob inhaltliche Gründe dafür vorliegen, dass ein Wert relativ vom Zentrum entfernt liegt, muss erst geklärt werden.

allen Schülern ziehen müssen. Dies haben wir nicht getan. Man könnte jetzt die Arbeit einstellen. In der Praxis ist es aber auch in solchen Fällen durchaus üblich, einmal hypothetisch von der Annahme auszugehen, es würde eine Zufallsstichprobe vorliegen. Erweisen sich unter dieser Annahme Unterschiede als nicht signifikant, so wird von vorneherein von einer Verallgemeinerung abgesehen. Falls unter dieser Annahme Signifikanz ermittelt wird, ist dies ein positiver Hinweis, man darf aber trotzdem nicht von einem definitiven statistischen Nachweis ausgehen. Wenn man diesen differenzierten Hinweis zu kompliziert findet, kann man sich in der Schule durchaus legitimiert fühlen, den Fall unter der Annahme, es wäre eine Zufallsstichprobe, einmal durchzuspielen. Natürlich sollte man die Verfahren auch noch an echten Stichproben anwenden.

Die von der Statistik bereit gestellten Verfahren für den 2-Stichproben-Vergleich sind aber in der Schule normalerweise kein Thema und auch zu kompliziert, um sie vollständig zu entwickeln. Eine sehr schöne Idee, die u.a. Arthur Engel (1987) bereits vor Jahren für den Unterricht vorgeschlagen hat, ist es, Randomisierungstests zu verwenden. Bei unseren Daten können wir uns das folgendermaßen vorstellen.

Wir machen ein Gedankenexperiment. In unserem Datensatz haben 232 Jungen und 301 Mädchen Angaben zur Lesezeit gemacht. Wir nehmen eine Urne und füllen sie mit 232 grauen und 301 roten Kugeln. Diese Kugeln teilen wir zufällig an die 533 Personen aus und vergleichen nun die Lesezeiten der roten Gruppe mit der der grauen Gruppe. Unterschiede im Leseverhalten, die wir auf diese Weise feststellen, sind durch Zufall entstanden, durch zufällige Aufteilung der gesamten Gruppe in graue und rote Personen. Natürlich kann es bei manchen zufälligen Aufteilungen sein, wenn auch selten, dass einmal die Vielleser vorwiegend rot werden, während die Wenigleser vorwiegend grau werden. Wenn wir das zufällige Aufteilen sehr oft wiederholen, dann lässt sich z.B. der Anteil der Fälle ermitteln, bei dem Differenz „Median im Lesen der roten Gruppe“ minus „Median im Lesen der grauen Gruppe“ eine Stunde oder größer ist. Diesen Unterschied von einer Stunde hatten wir bei unseren Personen zwischen Schülerinnen und Schülern beobachtet. Damit schätzt man die Wahrscheinlichkeit, dass rein zufällig ein solcher Unterschied oder ein noch größerer auftritt.

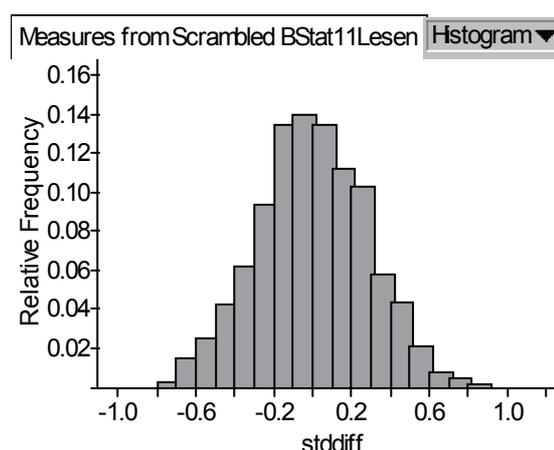
Mit der Software FATHOM lässt sich dieses zufällige Wiederverteilen von Merkmalsausprägungen ganz leicht durch ein Kommando „Scrambling attributes“ realisieren.

Wir interessieren uns nun in unserem Fall für die Unterschiede im arithmetischen Mittel (*meandiff*), im Median (*mediandiff*) und in der Standardabweichung (*stddiff*).

Wie im vorangehenden Abschnitt festgestellt, haben wir die folgenden Differenzen in unseren Daten beobachtet:

|                    | <i>Measures:</i><br>Kennzahl für Mädchen<br>minus Kennzahl für Jungen<br>(in Std. pro Woche) | Name des Measures |
|--------------------|--|-------------------|
| arith. Mittel      | 0,818193   | meandiff          |
| Median             | 1  | mediandiff        |
| Standardabweichung | 0,20177  | stddiff           |

In FATHOM kann man jetzt wie folgt vorgehen. Die obigen Maßzahlen kann man als sog. „measures“ für unseren Datensatz definieren. „Measures“ sind Eigenschaften einer ganzen Kollektion und nicht einzelner Schüler. Man muss nun in der Kollektion das Etikett „Geschlecht“ zufällig neu verteilen (Kommando: *Scramble Attributes* in Fathom). In der randomisierten Kollektion haben die obigen measures immer andere Werte. Wenn man das Randomisieren wiederholt, kann man diese 3 measures als Zufallsgrößen auffassen. Mit einem einfachen Befehl *Collect measures* erlaubt Fathom, deren Werte wieder in einem neuen Datensatz zu sammeln. Wir haben dies  $n = 1900$  mal durchgeführt. Die nächste Graphik zeigt die Verteilung der Variable *stddiff*.



Den Anteil, der größer als die beobachteten Werte ist, ermitteln wir zu <sup>11</sup>

| Measure    | Value       | Formula                               |
|------------|-------------|---------------------------------------|
| meansign   | 0.000526... | proportion (meandiff $\geq$ 0.818193) |
| mediansign | 0.291579    | proportion (mediandiff $\geq$ 1)      |
| stddevsign | 0.237368    | proportion (stddiff $\geq$ 0.20177)   |
| <new>      |             |                                       |

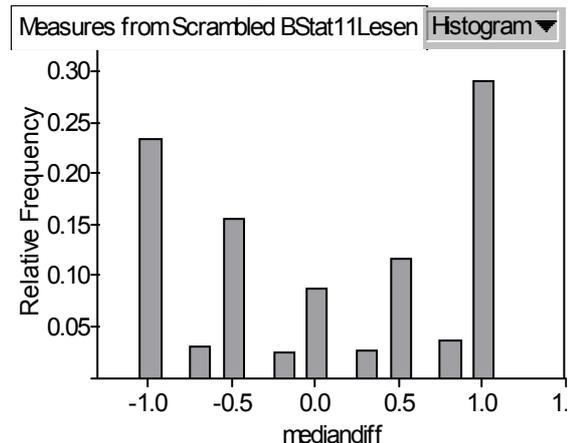
Was haben wir jetzt gewonnen? Wir beobachten als Ergebnis des Randomisierungstests eine relative Häufigkeit (eine Wahrscheinlichkeit) von 0,000526... dafür, dass der Unterschied des arithmetischen Mittels Jungen vs. Mädchen größer als die in der Stichprobe beobachtete Differenz von 0,818... ist. Wir haben also in unserer Stichprobe von 540 Schülern etwas vergleichsweise sehr seltenes beobachtet, wenn „in Wirklichkeit“ kein Geschlechtsunterschied vorhanden wäre. Wir können somit die Annahme, dass kein „Geschlechtseffekt“ vorhanden ist, begründet ablehnen oder anzweifeln. Die obige Wahrscheinlichkeit von 0,000526... nennt man auch den „beobachteten P-Wert“. Ist der P-Wert kleiner gleich 5% (p%), sagt man der Unterschied ist signifikant auf dem 5% (p%) –Niveau. Damit findet man Anschluss an die übliche Terminologie des Hypothesentestens. <sup>12</sup>

Lediglich die Abweichungen der arithmetischen Mittelwerte wären signifikant auf dem 5%-Niveau, sogar auf dem 1%-Niveau. Für die Mediandifferenz und die Differenz der Standardabweichungen gilt das nicht: Die beobachtete Differenz von 0,20177 oder größer bei der Standardabweichung kommt etwa mit der Wahrscheinlichkeit von 24% vor - unter der Annahme vor, dass das „Geschlecht“ zufällig zugeordnet wird. Man würde nach diesem Ergebnis erstmal nicht davon ausgehen, dass ein statistisch signifikanter Unterschied in den Standardabweichungen zwischen der Jungen- und der Mädchengruppe besteht.

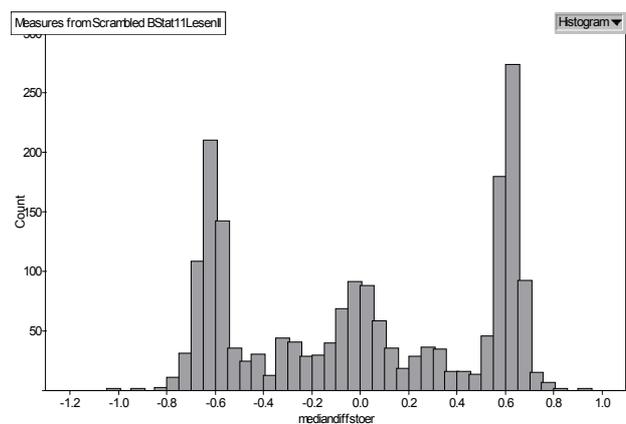
<sup>11</sup> Der Fathom Screenshot zeigt exemplarisch, wie man diesen Anteil über die Eingabe von Formeln ermitteln kann.

<sup>12</sup> Engel (1987) plädiert etwa in Anlehnung an Teile der statistischen Praxis für die Benutzung von P-Werten statt von vorgegebenen Signifikanzniveaus. Hierin steckt eine komplexe Thematik, die aber für die statistische Praxis relevant ist. Sie hängt auch mit unterschiedlichen Test-Konzepten von R. A. Fisher einerseits und Neyman & Pearson andererseits zusammen, auf die wir in einem gesonderten Aufsatz eingehen werden, vgl. hierzu auch den Leserbrief von Dieppen in Stochastik in der Schule 22 (2002), Heft 3, S. 34ff.

Überraschend ist vielleicht die Tatsache, dass der Medianunterschied nicht signifikant ist. Wir schauen uns die Verteilung an:



Es kommen bei den Differenzen immer nur wenige feste Zahlen, nämlich Vielfache von 0,25 vor, weil die Schüler und Schülerinnen im Wesentlichen ganze und halbe Stunden angeben, also praktisch runden, und Medianwerte immer Datenwerte oder solche zwischen 2 Werten sind. Dies erschwert hier die statistische Bewertung. <sup>13</sup> Man kann die Daten künstlich wieder realistischer machen, indem man zur Variable *Zeit\_Lesen* jeweils eine Zufallskomponente, z.B. eine zufällig aus dem Intervall (-0.25,+0.25) gezogene Zufallszahl, hinzu addiert, um das vorgenommene Runden auf halbe Stunden quasi rückgängig zu machen. Wir haben das einmal durchgeführt und so die Variable *Zeit\_LesenStoer* erhalten. Als Mediandifferenz der Jungen und Mädchen in der untersuchten Schülerkollektion erhalten wir dann 0,917185 Stunden. Nach dem wiederholten Randomisieren (Scrambling) ergibt sich als Verteilung der Mediandifferenzen folgendes Bild ( $n = 1900$ ).



<sup>13</sup> Genauere Überlegungen sollen einem eigenen Aufsatz vorbehalten sein. Zum Scrambling und zum Randomisieren finden sich weitere Informationen auf unserer Webpage.

Wir erhalten dann als Anteil, der rein zufällig größer/gleich dem beobachteten Medianunterschied von 0,917 ist, den Wert 1/1900, nämlich nur einen einzigen Fall. Also haben wir jetzt, ähnlich wie beim arithmetischen Mittel, einen signifikanten Unterschied der Mediane mindestens auf dem 1%-Niveau ermittelt. Der Streuungsunterschied bei der Standardabweichung erwies sich als nicht signifikant, das kann man durchaus als kompatibel damit ansehen, dass der Quartilabstand (als robustes Streuungsmaß) ja ohnehin schon gleich war.

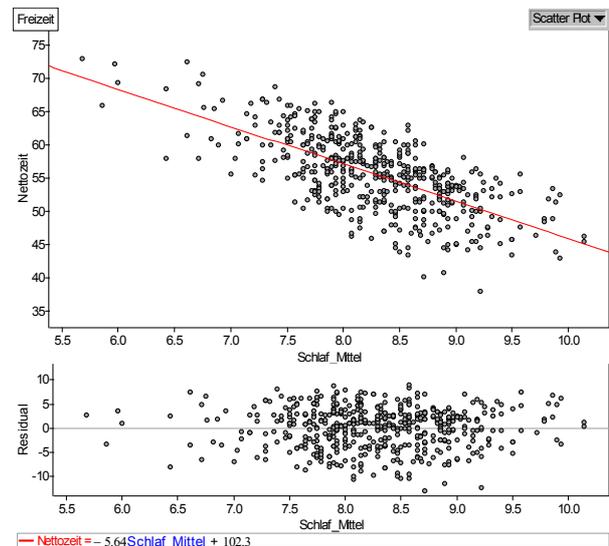
Wir wollen festhalten, dass das Verfahren des wiederholten Rerandomisierens (Scrambling) auf alle Unterschiedsmaße anwendbar und insofern universell ist. Dies Verfahren wird auch in der statistischen Praxis dort angewendet, wo noch keine maßgeschneiderten Verfahren entwickelt worden sind und man nicht auf Signifikanzaussagen verzichten möchte.

### 3.7 Korrelation und Regression

Ein wesentlicher Inhalt der Beschreibenden Statistik wurde in der bisherigen Vorstellung des Datensatzes noch nicht behandelt: Korrelation und Regression. Im Folgenden wird dargestellt, wie auch statistische Zusammenhänge differenziert mit den Freizeitdaten analysiert werden können. Wir konzentrieren uns hierbei bewusst auf grafische Analysemethoden, um auch Alternativen zu üblichen Zugängen aufzuzeigen.

#### Wie gewinnt man Zeit für Freizeitaktivitäten?

Eine Analyse der oben definierten „Nettozeiten“, die für Freizeit- und andere Aktivitäten zur Verfügung stehen, bringt das überraschende Ergebnis, dass Jungen wöchentlich im Mittel (sowohl Median als auch arithm. Mittel) etwa 2 Std. mehr Zeit für Freizeitaktivitäten haben als Mädchen. Die Frage, die wir uns stellen wollen, lautet: Welche Faktoren beeinflussen die den Jugendlichen zur Verfügung stehende Nettozeit besonders stark bzw. vielleicht überhaupt nicht? Überprüft man, wie die Variable Nettozeit entsteht, reduziert sich die Zahl der Faktoren auf zwei: die Schlafzeiten und die Zeit für Schulweg/Schule. Wir wollen an dieser Stelle den Zusammenhang zwischen der mittleren täglichen Schlafzeit<sup>14</sup> und der Nettozeit näher analysieren.



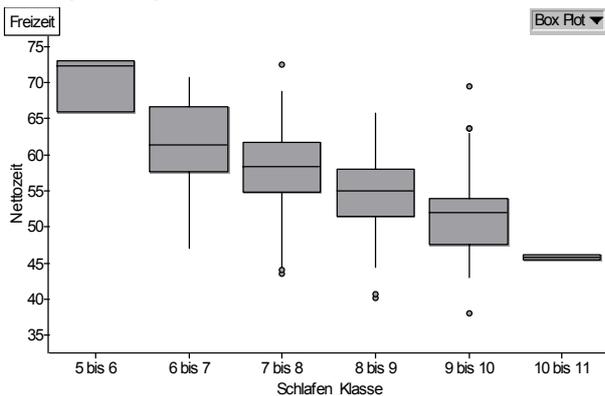
In der Punktwolke kann man einen Zusammenhang zwischen den Variablen erkennen: je höher die mittleren Schlafzeiten, desto weniger Nettozeit bleibt den Jugendlichen tendenziell für ihre Freizeitaktivitäten. Im Residuenplot darunter werden die Abweichungen von der kQ-Geraden (Optimale Gerade nach der Methode der kleinsten Quadrate) dargestellt, die im oberen Teil der Graphik eingezeichnet wurde. Es wird deutlich, dass diese Gerade insbesondere für Werte oberhalb von 8,75 Std. mittlerer Schlafzeit den Zusammenhang aber nur unbefriedigend beschreibt, denn hier ist noch ein systematischer Trend in den Residuen zu erkennen. Eine flachere Gerade würde in diesem Bereich die Beschreibung des Zusammenhangs verbessern. Grundsätzlich kann man durch eine differenzierte Analyse von Residuen die Auseinandersetzung mit dem Thema sinnvoll bereichern und insbesondere durch moderne Softwaretools offene Zugänge zum Thema Korrelation und Regression anbieten. Anregungen hierzu findet man in Schmidt & Noll (1994) und Biehler & Schweynoch (1999, 1999a).

An dieser Stelle kann man sich auch einen Exkurs „Schlafgewohnheiten“ vorstellen und die mittleren Schlafzeiten weiter thematisieren. Beispielsweise können die Zusammenhänge zwischen Aufstehzeiten und Schlafzeiten oder zwischen Zubettgehzeiten und Schlafzeiten betrachtet werden. Man stellt bei Letzterem eine relativ hohe Korrelation fest.

Eine weitere elementare Möglichkeit, den Zusammenhang zwischen Schlaf- und Nettozeit zu untersuchen, ohne dass man die lineare Regression benötigt, ist die Zerlegung der Verteilung der mittleren Schlafzeiten in Gruppen. Für diese Gruppen kann man die Verteilung der Nettozeiten z.B. durch Boxplots darstellen und diese vergleichen. Man hat die Analyse der Beziehung zweier numerischer Variabler somit auf ein „Gruppen-Vergleichs-

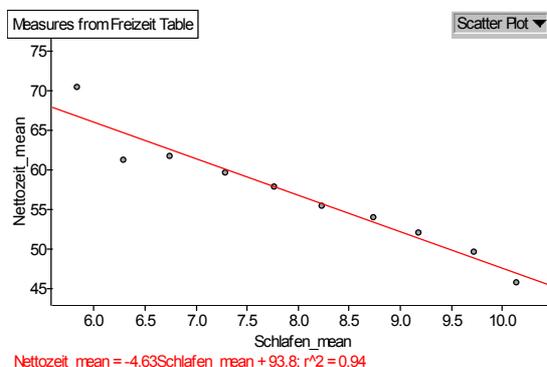
<sup>14</sup> Aus der oben aufgeführten Tabelle kann man die Gesamtschlafzeit ermitteln. Wir haben diesen Wert dann durch 7 (Tage) geteilt.

Problem“ zurückgeführt. Im folgenden Boxplot haben wir den Wertebereich in 6 gleichlange Intervalle aufgeteilt (jeweils 1 Std.).



Die Boxplots zeigen: Je mehr geschlafen wird, desto geringer ist tendenziell oder „im Durchschnitt“ die verfügbare Nettozeit. Unsere Vermutung wird bestätigt, dass Jugendliche ihre Nettozeiten insbesondere dadurch steigern, dass sie weniger schlafen. Die Boxplots der mittleren vier Gruppen – die äußeren sollte man wegen der geringen Gruppengröße vernachlässigen – verschieben sich systematisch sukzessive nach unten.

Die Zerlegung einer kontinuierlichen Variable in Gruppen ist insbesondere bei großen Datensätzen eine sinnvolle elementare Technik, um mögliche funktionale Abhängigkeiten herauszuarbeiten, ohne bestimmte Funktionstypen (etwa Linearität) vorauszusetzen. Engel (1998, 1999) hat komplexere Methoden dargestellt, bei denen nicht eine feste, sondern eine „wandernde“ Klasseneinteilung vorgenommen wird, damit man sich noch besser an Verläufe anpassen kann. Wir haben abschließend die Klassenbreite für die Aufteilung der mittleren Schlafzeiten verkleinert (0,5 Std.) und für die einzelnen Gruppen arithmetische Mittel und Mediane von mittlerer Schlafzeit und Nettozeit berechnet. Im Streudiagramm sind die ermittelten arithmetischen Mittel gegeneinander geplottet und probenhalber eine kQ-Gerade eingezeichnet.



Bisherige Ergebnisse werden bestätigt: Es zeigt sich, dass die Annahme eines linearen Zusammen-

hangs mit gewissen Unsicherheiten in den Randbereichen vernünftig erscheint. Bei dem entsprechenden „Medianplot“ ist dies übrigens genauso.

## 4. Eine Entdeckungsreise: Jugendliche und Nebenjobs

Wir wollen exemplarisch eine Entdeckungsreise durchführen. Welche Fragen bieten sich zu dem Thema „Jugendliche und Nebenjobs“ an? Zunächst kann man untersuchen, wie lange Jugendliche wöchentlich in Nebenjobs arbeiten. Wie viele der Jugendlichen haben überhaupt einen Job? Welchen Stundenlohn bekommen sie? Interessant ist sicherlich auch, in welchen Bereichen sie tätig sind: vielleicht privat als Babysitter, in der Nachhilfe oder im Einzelhandel. Bei allen diesen Fragen sind Gruppenvergleiche möglich: z.B. zwischen Jungen und Mädchen, zwischen Älteren und Jüngeren, zwischen verschiedenen Schulen. Selbstverständlich kann man diese Kategorien auch noch miteinander kombinieren.

### 4.1 Wie lange arbeiten Jugendliche in Nebenjobs?

Eine erste Orientierung verschafft uns die Betrachtung der beiden Mittelwerte arithmetisches Mittel und Median für die Variable *Zeit\_Jobben*. Das arithmetische Mittel der Gesamtgruppe beträgt 3,5 Std. wöchentliche Arbeitszeit, der Median liegt bei nur einer Stunde. Unser besonderes Interesse gilt aber dem Vergleich von Mädchen und Jungen:

*Durchschnittliche Arbeitszeit in Stunden*

|               | Jungen | Mädchen | Gesamt |
|---------------|--------|---------|--------|
| arith. Mittel | 3,09   | 3,82    | 3,5    |
| Median        | 0      | 2       | 1      |

Solche deutlichen Unterschiede zwischen Median und arithm. Mittel weisen auf ungewöhnliche Verteilungsformen hin, i.d.R. auf schiefe Verteilungen. Aus Platzgründen verzichten wir hier auf eine Verteilungsgraphik, die man sich bei jeder realen Analyse ansehen sollte. Man erkennt beim arithmetischen Mittel eine deutliche Differenz: die Jungen arbeiten durchschnittlich ca. eine dreiviertel Stunde (0,73 Std.) weniger pro Woche als die Mädchen. Beim Vergleich der Mediane fällt der Unterschied mit zwei Stunden noch stärker aus. Die ermittelten Medianwerte verdienen jedoch etwas mehr Aufmerksamkeit: der Wert von 0 Std. für die Jungengruppe bedeutet auch, dass mindestens die Hälfte der Jungen überhaupt nicht in Nebenjobs arbeitet. Bei den Mädchen ist es umgekehrt ( $\tilde{x} = 2$ ): mindestens die Hälfte dieser Gruppe arbeitet in Nebenjobs. Die folgende Übersicht liefert uns die genauen

Anteile (eine neue Variable *Nebenjob* ordnet jedem Wert *Zeit\_Jobben* > 0 den Wert „ja“ zu, sonst „nein“).

| Nebenjob |        | Jungen | Mädchen | Gesamt |
|----------|--------|--------|---------|--------|
| ja       | Anzahl | 103    | 181     | 284    |
|          | Anteil | 45,6%  | 60,7%   | 54,2%  |
| nein     | Anzahl | 123    | 117     | 240    |
|          | Anteil | 54,4%  | 39,3%   | 45,8%  |

In der Gesamtgruppe arbeitet in einem Nebenjob etwas mehr als die Hälfte der Befragten (ca. 54%). Für die Mädchen scheint es aber typischer zu sein als für die Jungen: 181 von 298 Mädchen, das entspricht 61%, jobben regelmäßig. Bei ihren männlichen Altersgenossen tut dies mit fast 46% ein viel kleinerer Anteil (103 von 226 Jungen).

Berücksichtigt man jetzt in beiden Gruppen nur noch diejenigen, die auch tatsächlich Jobs haben, führt dieses zu modifizierten Mittelwerten:

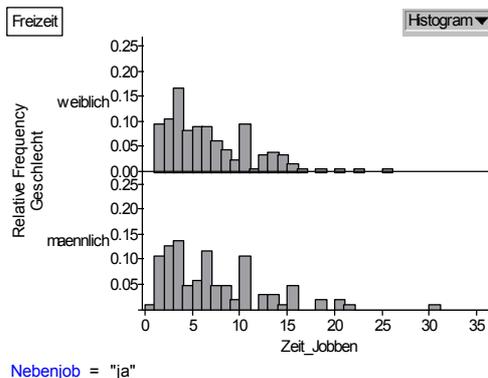
*Durchschnittliche Arbeitszeit in Stunden der tatsächlich Jobbenden*

|               | Jungen | Mädchen |
|---------------|--------|---------|
| arith. Mittel | 6,78   | 6,30    |
| Median        | 6      | 5       |

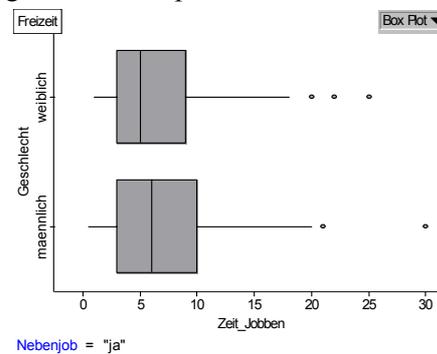
Es ergibt sich ein anderes Bild, denn sowohl das arithmetische Mittel wie auch der Median liegen nun bei der Jungengruppe höher. Auf die genaueren Werte können wir an dieser Stelle verzichten. Festhalten muss man jedoch ein wichtiges Zwischenergebnis als Grundlage für weitere Analysen: „Mädchen arbeiten häufiger in einem Nebenjob als Jungen – wenn Jungen aber in Nebenjobs arbeiten, dann tun sie es tendenziell länger als Mädchen!“

## 4.2 Wie unterscheiden sich Mädchen und Jungen?

An dieser Stelle bietet sich als nächster Schritt an, die Verteilungen zu visualisieren, um weitere Unterschiede, Gemeinsamkeiten und Besonderheiten herauszufinden, z.B. durch Histogramme (Klassenbreite 1 Std.):



Wir nutzen die Histogramme, um innerhalb der beiden Verteilungen kleinere Gruppen zu identifizieren und diese miteinander zu vergleichen: die erste besteht aus denjenigen Jugendlichen, die weniger als 4 Std. pro Woche jobben. Bis zur Klasse [3,4) steigen in beiden Histogrammen die Häufigkeiten an. Kumuliert man die Häufigkeiten bis hierhin, so erreichen sie jeweils einen Anteil von ungefähr 36%. Sowohl bei Jungen wie auch bei Mädchen arbeiten also etwas mehr als ein Drittel weniger als vier Stunden pro Woche. Bei den Jungen sieht man allerdings etwas häufiger Werte in den unteren Intervallen als bei den Mädchen. Je nach Interesse können weitere Klassen zusammengefasst und verglichen werden. Der detaillierte Vergleich bleibt dem Leser überlassen, ebenso der Vergleich der dargestellten Boxplots für die beiden Gruppen.



Zusammenfassend kann man aber bis hierher festhalten, dass die von den Jungen genannten Werte stärker streuen als die der Mädchen – der Vergleich verschiedener Streuungsmaße unterstreicht das –, daneben nennen die Jungen gegenüber den Mädchen häufiger Werte, die besonders niedrig oder besonders hoch sind. Ferner arbeiten ungefähr 10 % der Jungen mindestens 15 Std. pro Woche, also nahezu zwei Arbeitstage eines Berufstätigen; bei den Mädchen ist dieser Anteil nur halb so groß.

## 4.3 Arbeiten ältere Jugendliche häufiger in Nebenjobs?

Ergänzen kann man die bisherigen Analysen dadurch, dass man innerhalb der Mädchen- und Jungengruppen zusätzlich nach dem Alter zu unterscheidet. Auch hier macht man weitere wichtige Entdeckungen.

In der folgenden Kreuztabelle sind jeweils die absolute Häufigkeit von Jugendlichen sowie die absolute und die relative Häufigkeit von Jungen und Mädchen mit Nebenjob für jede Altersgruppe bzw. für die Gesamtgruppe dargestellt. Wir betrachten dabei nur die Jugendlichen zwischen 16 und 18 Jahren, die sechs Jugendlichen im Alter von 19 und 20 Jahren, die mit einer Ausnahme alle einen Nebenjob haben, vernachlässigen wir.

| Alter | Jungen  | Mädchen | Gesamt |        |
|-------|---------|---------|--------|--------|
| 16    | Anzahl  | 54      | 90     | 144    |
|       | mit Job | 24      | 51     | 75     |
|       | Anteil  | 44,44%  | 56,67% | 52,08% |
| 17    | Anzahl  | 136     | 180    | 316    |
|       | mit Job | 62      | 115    | 177    |
|       | Anteil  | 45,59%  | 63,89% | 56,01% |
| 18    | Anzahl  | 27      | 15     | 42     |
|       | mit Job | 12      | 11     | 23     |
|       | Anteil  | 44,44%  | 73,33% | 54,76% |

Bevor wir Mädchen und Jungen getrennt betrachten, richten wir zunächst noch einmal den Blick auf die Gesamtgruppe (rechte Spalte). Unterscheiden sich die Jugendlichen in den verschiedenen Altersstufen? Überraschenderweise tun sie es kaum, denn der Anteil von Jugendlichen mit Nebenjobs bleibt bei zunehmendem Alter nahezu gleich: die relativen Häufigkeiten der Jugendlichen mit Nebenjobs differieren bei den 16-, 17- und 18-Jährigen nicht sehr stark. Er liegt zwischen 52% und 56% und entspricht ungefähr dem Anteil der Gesamtgruppe, der fast 55% beträgt.

Gruppieren wir demgegenüber wieder nach Mädchen und Jungen, so gibt es eine interessante Entdeckung: bei den Jungen (linke Spalte) bleibt in den Alterstufen zwischen 16 und 18 Jahren der Anteil fast unverändert bei ca. 45%. Bei den Mädchen (mittlere Spalte) ist das anders, denn hier nehmen die Anteile mit höherem Alter zu: sie betragen 57% bei den 16-Jährigen, 64% bei den 17-Jährigen und 73% bei den 18-Jährigen. Dieses bedeutet, je älter die Mädchen sind, desto relativ mehr von ihnen haben einen Nebenjob. Bei den 17-Jährigen sind es knapp zwei Drittel, bei den 18-Jährigen sogar fast drei Viertel. Allerdings ist die Gruppe der befragten 18jährigen sehr klein für eine allgemeinere Aussage. Daneben unterstreichen die Angaben der Kreuztabelle auch die Behauptung, dass Mädchen tendenziell häufiger in einem Nebenjob arbeiten als Jungen, denn in jeder Alterstufe – insbesondere bei den 18-Jährigen – sind die entsprechenden Anteile in der Mädchengruppe höher.

Ein nächster Schritt könnte hier wieder sein, die festgestellten Unterschiede mit Randomisierungstests auf Signifikanz hin zu analysieren. Bei diesen einfachen Anteilen kann man aber auch Konfidenzintervalle mit schulüblichen Mitteln errechnen und die Überlappung prüfen.

Die dargestellten Explorations können nur beispielhaft darstellen, wie man das Thema „Nebenjobs“ untersuchen kann. Wir haben an dieser Stelle überwiegend numerische Methoden bzw. Darstellungen gewählt und interessante Ergebnisse sam-

eln können. Das Thema bietet aber genug Potenzial, auch durch eine stärkere Akzentuierung graphischer Methoden vielfältige Entdeckungen zu machen.

## 5. Schlussbemerkungen

Wir haben in unserem Aufsatz das inhaltliche und didaktische Potential unseres Datensatzes darzustellen versucht. Im Hinblick auf Schülerleistungen muss unterstrichen werden, dass es auf reflektierte Datenanalysen mit schriftlich formulierten Interpretationen ankommen soll. Das kann schon an kleineren Aufgaben, z. B. bei Gruppenvergleichen, geübt werden. Im Idealfall ist am Ende einer Unterrichtseinheit von allen Schülerinnen und Schülern eine Projektarbeit zu erbringen, eine Art Gesellenstück zur Datenanalyse. Diese Prinzipien haben wir in den Unterrichtsversuchen, die wir bisher mit den Daten durchgeführt haben, realisiert. Stefan Schweynoch berichtet in dieser Ausgabe von einem Unterrichtsversuch im Rahmen des Schüleraustausches, in dem unser Fragebogen verwendet wurde.

Unseren Datensatz haben wir in jüngster Zeit weiterhin in der Lehrveranstaltung „Elementare Stochastik“, einer Fachvorlesung für Lehramtsstudierende (Grund-, Haupt-, und Realschulen) an der Universität Kassel erfolgreich erprobt. Ferner wurden die Daten zusammen mit der Software FATHOM im Zeitraum Februar – April 2003 in einer 13. Klasse eines Kasseler Oberstufengymnasiums erprobt. Über die Evaluation dieser Unterrichtsexperimente werden wir an anderer Stelle berichten.

## 6. Die Muffins-Webpage

Auf der Webpage zum Muffinsprojekt unter <http://www.mathematik.uni-kassel.de/didaktik/muffins> finden Sie zum jetzigen Zeitpunkt die hier vorgestellten Daten in unterschiedlichen Formaten sowie den Fragebogen zur Umfrage in verschiedenen Sprachen und in einer Online-Version. Das Angebot unseres Portal soll in den kommenden Monaten ausgeweitet werden. Die geplanten Inhalte sind z.B. konkrete unterrichtliche Erfahrungsberichte, Unterrichtsbeispiele und auch weitere interessante exemplarische Datenanalysen. Interessierte Lehrerinnen, Lehrer und Schulen, die den On-line-Fragebogen nutzen wollen, werden gebeten, sich mit uns in Verbindung zu setzen.

## Literatur

Baum, M. u.a. (2000): Lambacher Schweizer 11, Stuttgart, Ernst Klett Verlag

Biehler, R. (2001): Statistische Kompetenz von Schülerinnen und Schülern – Konzepte und Ergebnisse empirischer Studien am Beispiel des Vergleichens empirischer Verteilungen. In Borovcnik et al. (2001), 97- 114

Biehler, R. & Weber, W. (Hrsg.) (1995): Explorative Datenanalyse. Computer+Unterricht 17 (März 1995)

Biehler, R. (Hrsg.) (1999): Daten und Modelle. Mathematik lehren 97 (Dezember 1999)

Biehler, R. & Kombrink, K. (1999): Mediennutzung von Schülerinnen und Schülern - Auswertung von Umfragen, Mathematik lehren 97, 6-11

Biehler, R. & Schweynoch, S. (1999): Trends und Abweichungen von Trends. Mathematik lehren 97, 17-22, Friedrich-Verlag

Biehler, R. & Schweynoch, S. (1999a): Holen Frauen auf? Prognosen zu Weltrekorden im Sport. Mathematik lehren 97 (Dezember 1999) Internet-Beilage ([http://www.friedrich-verlagsgruppe.de/ftp.pub/500/ml\\_97.pdf](http://www.friedrich-verlagsgruppe.de/ftp.pub/500/ml_97.pdf))

Borovcnik, M., Engel, J. & Wickmann, D. (Hrsg.) (2001): Anregungen zum Stochastikunterricht. Hildesheim Franzbecker

Engel, A. (1987): Stochastik. Stuttgart: Klett

Engel, J. (1998): Zur stochastischen Modellierung funktionaler Abhängigkeiten: Konzepte, Postulate, Fundamentale Ideen. Mathematische Semesterberichte 45(1), 95-112

Engel, J. (1999): Kurvenanpassung im Streudiagramm mit gleitenden gewichteten Mittelwerten. Praxis der Mathematik 41(2), 69-73

Jahnke, T. u.a. (2000): Mathematik 11. Schuljahr, Berlin, Cornelsen Verlag

Griesel, H., Postel, H. (1999): Elemente der Mathematik 11, Hannover, Schroedel Verlag

Kombrink, K. (1997): Kompetenzen von Schülerinnen und Schülern im Umgang mit Darstellungsmitteln und Begriffen der beschreibenden Statistik, Universität Bielefeld, Fakultät für Mathematik

Ministerium für Schule und Weiterbildung, Wissenschaft und Forschung des Landes Nordrhein-Westfalen (1999): Richtlinien und Lehrpläne für die Sekundarstufe II – Gymnasium und Gesamtschule – Mathematik, Frechen, Ritterbach Verlag

Schmidt, G. & Noll, G. (1994): Trends und statistische Zusammenhänge. Materialien zur Explorativen Datenanalyse und Statistik in der Schule. Hrsg. vom Landesinstitut für Schule und Weiterbildung, Soest, Druck Verlag Kettler.

## Software

FATHOM: Version 1.16. Key Curriculum Press. Testversion unter <http://www.keypress.com/fathom/index.html>

MEDASS LIGHT: Autoren R. Biehler, W. Rach, S. Bauer. Kostenloser Download und Registrierung für Lehrende unter <http://www.mathematik.uni-kassel.de/didaktik/biehler/meddown.html>

## Anschriften der Verfasser

Prof. Dr. Rolf Biehler  
Klaus Kombrink  
Universität Kassel, FB Mathematik/Informatik  
34109 Kassel

[biehler@mathematik.uni-kassel.de](mailto:biehler@mathematik.uni-kassel.de)  
[kombrink@mathematik.uni-kassel.de](mailto:kombrink@mathematik.uni-kassel.de)

Stefan Schweynoch  
Schongauer Str. 33; 33615 Bielefeld

[stefan.schweynoch@gmx.de](mailto:stefan.schweynoch@gmx.de)