

Kapitel 3

Lügen mit
Statistik

Die graphische Darstellung von Häufigkeiten bietet vielfältige Möglichkeiten, den Betrachter zu manipulieren. Eine immer wieder verwendete optische Manipulation beruht darauf, dem Betrachter einen Flächen- oder Raumvergleich anstelle eines Längenvergleichs zu präsentieren. Außer den graphischen Darstellungen werden noch weitere Methoden präsentiert, mit denen bei Statistik gelogen werden kann.

3.1 Manipulation graphischer Darstellungen

In der Statistik »Deutscher Wein Statistik 2014/2015« findet man Angaben über geprüfte Weinmengen nach Qualitätsstufen. Für das Anbaugebiet Pfalz sind es 1708 000 hl, Württemberg kommt auf 862 000 hl. In der Darstellung durch ein Säulendiagramm ergibt sich damit das folgende Bild:

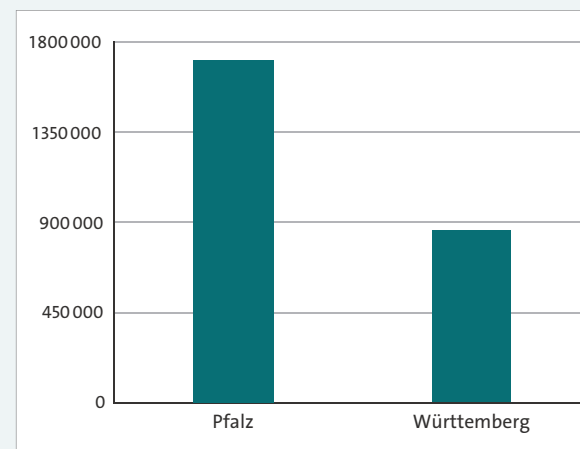


Abbildung 3.1 Qualitätswein in hl 2014/2015

Diese Darstellung ist natürlich langweilig. Der durch Medien verwöhnte Betrachter möchte lieber Weinflaschen statt Säulen sehen:

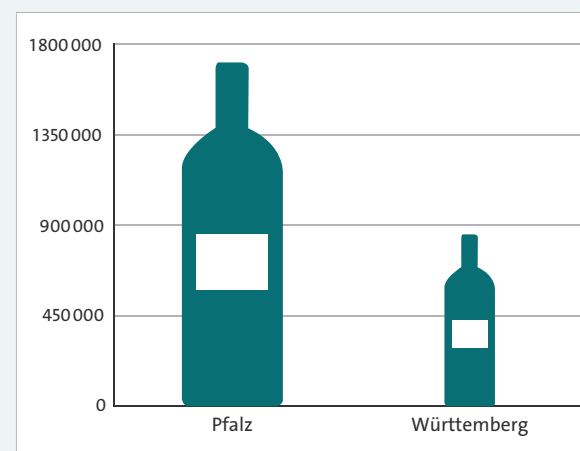


Abbildung 3.2 Piktogramm

Also ersetzt man die Säulen durch Weinflaschen gleicher Höhe. Für den Betrachter ergibt sich dadurch ein ganz anderer Blickwinkel. Die Unterschiede zwischen Pfalz und Württemberg erscheinen jetzt gravierend. Das rührt daher, dass dem Betrachter ein Flächenvergleich statt des Vergleichs der Höhen aufgedrängt wird. Während das Verhältnis der Längen der Weinflaschen von Pfalz und Württemberg etwa 2:1 beträgt, so ist das wahrgenommene Verhältnis der Flächen 4:1 – zur Erläuterung: Wenn die Länge und die Breite verdoppelt werden, dann vervierfacht sich der Flächeninhalt.

Wer gezielt manipulieren möchte, der wählt eine solche Darstellung – oder besser noch eine räumliche Darstellung, welche die Unterschiede noch größer erscheinen lässt. Im Jahr 2014 betrug der Hühnerfleischkonsum in Deutschland 12,4 kg pro Person. Das ist etwa doppelt so viel wie der Konsum im Jahr 1970 betragen hatte [BOE]. Stellt man dieses Verhältnis durch ein dreidimensional wirkendes Huhn graphisch dar, dann wird optisch der Eindruck eines Verhältnisses von 8:1 erzeugt.

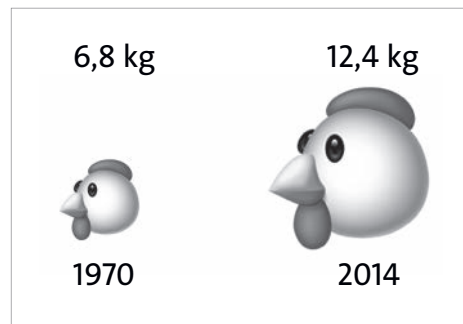


Abbildung 3.3 Hühnerfleischkonsum 1970 und 2014

Eine beliebte Methode zur Manipulation eines Diagramms ist die verkürzte y-Achse. Die folgende Graphik zeigt die erwartete Entwicklung der Schülerzahlen bis zum Schuljahr 2010/2011, wie sie in dem vom rheinland-pfälzischen Kultusministerium herausgegebenen »Elternjournal« im Januar 1993 prognostiziert wurde:

Das Diagramm ist um einen Sockel von 500 000 Schülern verkürzt, d. h. die y-Achse beginnt nicht bei 0, sondern bei 500 000. Dadurch wird der Anstieg der Schülerzahlen besonders drastisch dargestellt. Die nächste Graphik zeigt das ungekürzte Diagramm, in dem die wahre Relation des Anstiegs zu sehen ist.

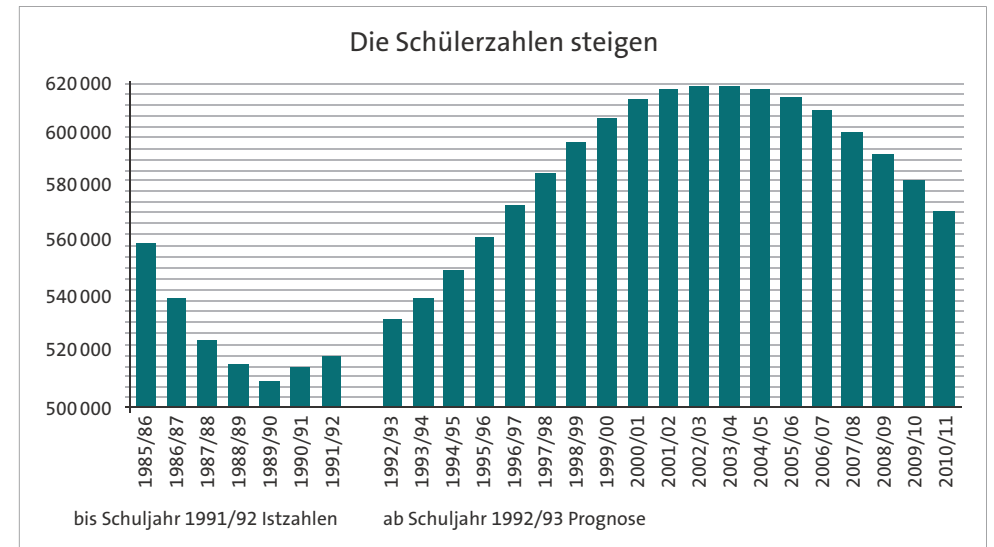


Abbildung 3.4 Die Schülerzahlen steigen – drastische Darstellung

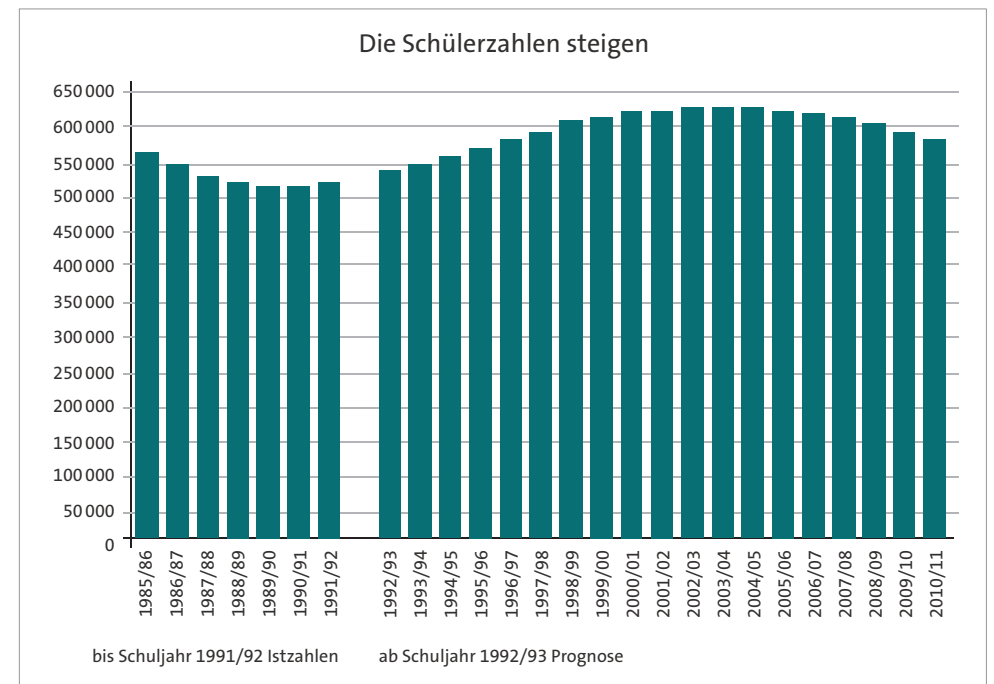


Abbildung 3.5 Die Schülerzahlen steigen – ungekürzte Darstellung

Die Verkürzung eines Diagramms ist erlaubt, wenn durch eine Unterbrechung an der y-Achse der Hinweis gegeben wird, dass abgeschnitten worden ist. Fehlt wie im betrachteten Beispiel dieser Hinweis, so handelt es sich um eine Manipulation.

3.2 Losbuden und Krankenhäuser: Das Simpson-Paradoxon

Wer Statistik verstehen will, der sollte das *Simpson-Paradoxon* kennen. Das Wissen um dieses Paradoxon ist für den Umgang mit Daten unerlässlich.

Peter und Paul verkaufen auf dem Jahrmarkt Lose. Jeder hat zwei Eimer, die Gewinne und Nieten enthalten. Wer ein Los bei Peter kaufen will, sucht sich erst einen der Eimer A oder B von Peter aus und zieht dann ein Los. Analoges gilt für einen Kauf bei Paul. (Die Anregung zu diesem Beispiel stammt aus [KRA].)

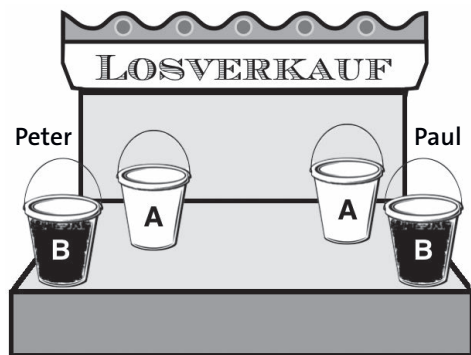


Abbildung 3.6 Losbude

Könnte man in die Eimer hineinsehen und auch erkennen, wie oft Gewinne und Nieten in den Eimern vorkommen, dann würde sich das folgende Bild ergeben:

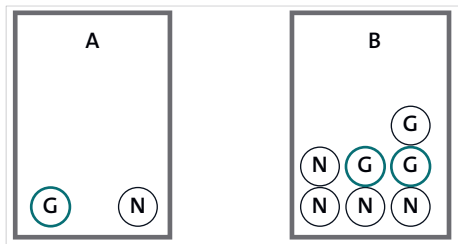


Abbildung 3.7 Die Eimer von Peter

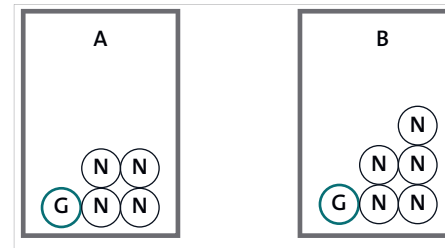


Abbildung 3.8 Die Eimer von Paul

Wüsste man als Loskäufer über die Verteilung von Gewinnen (G) und Nieten (N) Bescheid, dann würde man sowohl bei Peter als auch bei Paul aus Eimer A ziehen. Die folgende Tabelle zeigt, warum dies der Fall ist. Dazu werden für jeden Eimer die Gewinnchancen ermittelt. Diese sind für die Eimer mit der Aufschrift A größer als für die B-Eimer:

		Gewinne	Nieten	Gesamtzahl der Lose	Gewinnchance in Prozent
Peter	Eimer A	1	1	2	$1/2 = 50\%$
	Eimer B	3	4	7	$3/7 = 42,86\%$
Paul	Eimer A	1	4	5	$1/5 = 20\%$
	Eimer B	1	5	6	$1/6 = 16,67\%$

Tabelle 3.1 Gewinnchancen

Für Peter und Paul ist das Ende ihrer Schicht an der Losbude erreicht. Sie werden von Mary abgelöst. Damit Mary nicht mit 4 Eimern hantieren muss, werden kurzerhand die Eimer A zusammenschüttet. Dasselbe erfolgt auch für die Eimer B:

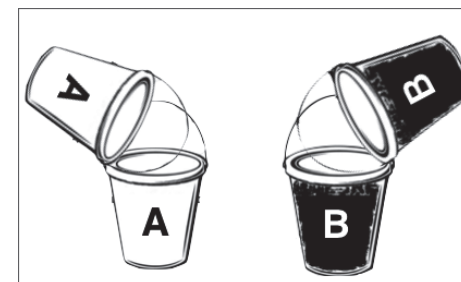


Abbildung 3.9 Zusammenschütten der Eimer

Stellen Sie sich nun vor, Sie wollten bei Mary ein Los kaufen. Aus welchem Eimer würden Sie dann ziehen, um eine möglichst große Gewinnchance zu haben?

Vor dem Zusammenschütten hatte man bei den Eimern A eine größere Gewinnchance als bei den Eimern B. Also scheint es klar zu sein, dass auch nach dem Zusammenschütten der Eimer A eine bessere Gewinnchance bietet. Aber Vorsicht! Wie es nach dem Zusammenschütten in den Eimern A und B von Mary aussieht, zeigt die nächste Abbildung:

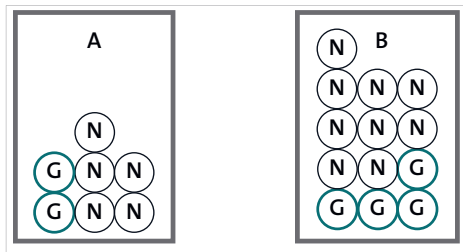


Abbildung 3.10 Eimer A und B von Mary

Eimer A bietet eine Gewinnchance von $2/7$, für Eimer B ist die Gewinnchance gleich $4/13$.

Nun ist $2/7 \approx 0,286$ und $4/13 \approx 0,308$. Folglich ist es jetzt günstiger, aus Eimer B zu ziehen! Die nächste Abbildung zeigt eine Gesamtübersicht:

		Gewinne	Nieten	Gesamtzahl der Lose	Gewinnchance in Prozent
Peter	Eimer A	1	1	2	$1/2 = 50\%$
	Eimer B	3	4	7	$3/7 = 42,86\%$
Paul	Eimer A	1	4	5	$1/5 = 20\%$
	Eimer B	1	5	6	$1/6 = 16,67\%$
Gesamt	Eimer A	2	5	7	$2/7 = 28,57\%$
	Eimer B	4	9	13	$4/13 = 30,77\%$

Tabelle 3.2 Gewinnchancen nach dem Zusammenschütten

Wie kann man das verstehen? Rein mathematisch betrachtet ist das nicht schwer:

Aus $\frac{3}{7} < \frac{1}{2}$ und $\frac{1}{6} < \frac{1}{5}$ folgt eben *nicht* die Gültigkeit von $\frac{3+1}{7+6} < \frac{1+1}{2+5}$ bzw. $\frac{4}{13} < \frac{2}{7}$.

Vielmehr ist $\frac{4}{13} > \frac{2}{7}$ richtig.

Für ganze Zahlen a, b, c und d, für die $\frac{a}{b} < \frac{c}{d}$ und $\frac{p}{q} < \frac{r}{s}$ gilt, kann aber $\frac{a+p}{b+q} < \frac{c+r}{d+s}$

durchaus richtig sein. Man kann also weder folgern, dass diese Ungleichung gilt, noch, dass sie nicht gilt. Dementsprechend gibt es Zahlen, für die sich der Überraschungseffekt beim Zusammenschütten einstellt, aber auch Zahlen, für die er ausbleibt.

Aufgabe 1: Summenvergleich

Am nächsten Tag füllen Peter und Paul ihre Eimer wie folgt:

		Gewinne	Nieten	Gesamtzahl der Lose	Gewinnchance
Peter	Eimer A	3	7	10	$3/10$
	Eimer B	2	3	5	$2/5$
Paul	Eimer A	1	6	7	$1/7$
	Eimer B	1	5	6	$1/6$

Untersuchen Sie, indem Sie Brüche vergleichen:

- In welchen Eimer sollte ein Kunde bei Peter greifen?
- In welchen Eimer sollte ein Kunde bei Paul greifen?
- Wie sieht es aus, wenn die beiden Losverkäufer wieder ihre beiden Eimer A und ihre beiden Eimer B zusammenschütten?

Auch wenn sich hier etwas mathematisch begründen lässt, so richtig verstanden hat man doch noch nicht, was hier vorgeht. Es lässt sich festhalten, dass die Bewertung der beiden Teilgruppen, die dann zu einer Gesamtgruppe zusammengeführt wurden, anders ausfällt als die Bewertung der Gesamtgruppe. Dieses Phänomen, dass nämlich das Zusammenlegen von Teilgruppen zu einer Umkehrung der Bewertung führen kann, wurde von Edward Hugh Simpson 1951 gründlich untersucht. Er war nicht der Erste, der sich mit dieser Problematik beschäftigt hat, aber schließlich wurde dieses Phänomen nach ihm benannt. Mit dem folgenden Beispiel kommt man den Ursachen für das *Simpson-Paradoxon* auf die Spur.

Beispiel 1: Heilungsraten in zwei Krankenhäusern

Stellen Sie sich vor, Sie wollten einen erkrankten Angehörigen zur Behandlung in ein Krankenhaus bringen. Nach einiger Recherche haben Sie zwei Krankenhäuser ausfindig gemacht, die sich auf die Behandlung der vorliegenden Krankheit spezialisiert haben. Sie bringen in Erfahrung, dass von 1000 in Krankenhaus A behandelten Patienten 900 geheilt entlassen werden. Die restlichen 100 Patienten versterben. In Krankenhaus B überleben 800 von 1000 Patienten, 200 versterben. Damit scheint die Angelegenheit für Sie klar zu sein. Den 90 % der Überlebenden von Krankenhaus A stehen nur 80 % Überlebende in Krankenhaus B gegenüber. Das Krankenhaus A bietet damit die besseren Überlebenschancen.

Da Sie besonders gründliche Nachforschungen anstellen, erkennen Sie, dass in Krankenhaus A überwiegend leicht erkrankte Patienten behandelt werden. In Krankenhaus B überwiegt dagegen die Behandlung der besonders schwer erkrankten Patienten.

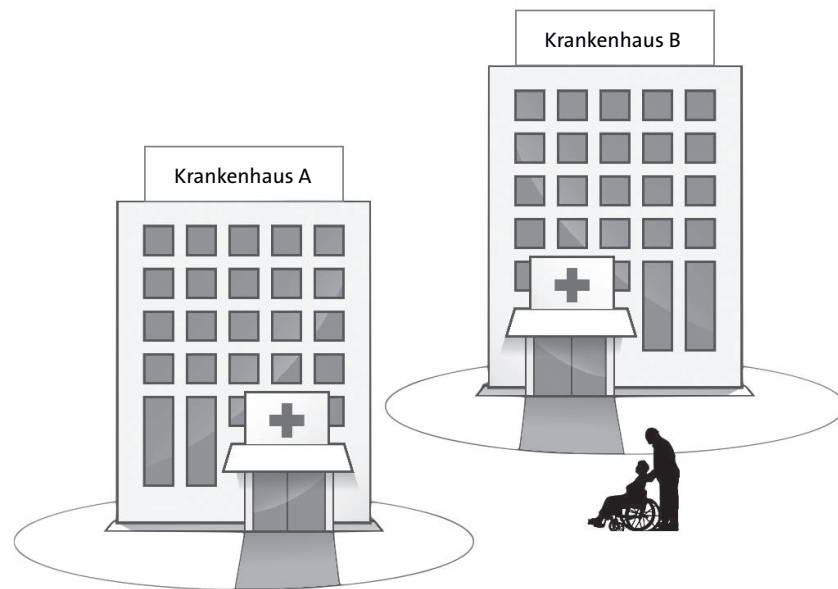


Abbildung 3.11 Welches Krankenhaus ist besser?

In Krankenhaus A überlebten von 900 leicht erkrankten Patienten 870. In Krankenhaus B wurden 600 leicht erkrankte Patienten behandelt, von denen 590 überlebten.

Schwer erkrankte Patienten wurden in Haus A nur in 100 Fällen behandelt, von denen 30 überlebten. In Haus B wurden dagegen 400 schwer Erkrankte behandelt, von denen 210 überlebten.

Aus den genannten Zahlen kann man die relativen Häufigkeiten ermitteln, mit denen leicht erkrankte und schwer erkrankte Patienten in den beiden Krankenhäusern überleben. Die prozentualen Anteile der überlebenden Patienten sind in Tabelle 3.3 aufgelistet.

Überraschend ist, dass sowohl bei den leicht Erkrankten als auch bei den schwer Erkrankten das Krankenhaus B die besseren Heilungschancen bietet.

In der anfänglichen Gesamtbetrachtung beider Krankenhäuser wird nicht zwischen den leicht und den schwer Erkrankten unterschieden. Es werden also leicht Erkrankte, die eine gute Heilungsaussicht haben, zusammen mit schwer Erkrankten in einen Topf geworfen. Das Krankenhaus A hat zudem viele leicht Erkrankte und nur wenige schwer Erkrankte. Dies erklärt das gute Abschneiden von Krankenhaus A in der Gesamtbetrachtung. Das bedeutet schließlich, dass ein Patient bei Einlieferung in Krankenhaus B die besseren Überlebenschancen hat.

		Überlebt	Tot	Patienten	Überlebende
Leichte Erkrankung	Haus A	870	30	900	$870/900 = 96,67\%$
	Haus B	590	10	600	$590/600 = 98,33\%$
Schwere Erkrankung	Haus A	30	70	100	$30/100 = 30\%$
	Haus B	210	190	400	$210/400 = 52,50\%$
Gesamt	Haus A	900	100	1000	$900/1000 = 90\%$
	Haus B	800	200	1000	$800/1000 = 80\%$

Tabelle 3.3 Relative Häufigkeiten für Überlebende getrennt nach leichter und schwerer Erkrankung

Das letzte Beispiel hat gezeigt, dass das Simpson-Paradoxon widersprüchliche Ergebnisse aus ein und demselben Datensatz produziert, je nachdem ob man die gesamte Population oder Teilgruppen davon betrachtet. Dabei handelt es sich eigentlich gar nicht um ein Paradoxon, denn schließlich kann man alles widerspruchsfrei auflösen, wenn man den zugrunde liegenden verborgenen Faktor findet, dessen Nichtbeachtung für die Verwirrung sorgt. Im letzten Beispiel war der verborgene Faktor die Schwere der Erkrankung.

Das im nächsten Beispiel behandelte Problem ist ein klassischer Fall des Simpson-Paradoxons. Im Jahr 1972 wurde in Großbritannien eine Studie zu Gewohnheiten und Lebenserwartung der Menschen gestartet. In diesem Rahmen wurden 1314 Frauen über

einen Zeitraum von 20 Jahren beobachtet, unter ihnen Raucherinnen und Nichtraucherinnen. Nach diesen 20 Jahren wurde ermittelt, wie viele Überlebende in beiden Gruppen vorkamen [APP].

Beispiel 2: Nichtraucherinnen und Raucherinnen verschiedener Lebensalter

1972 wurden 1314 Frauen in Raucherinnen und Nichtraucherinnen eingeteilt [APP]. 20 Jahre später ergab sich, was Überlebende und Gestorbene in beiden Gruppen anging, das folgende Bild:

	Überlebt	Tot	Anzahl Frauen	Überlebende
Raucherinnen	443	139	582	443/582 = 76,12 %
Nichtraucherinnen	502	230	732	502/732 = 68,58 %

Tabelle 3.4 Überlebende Raucherinnen und Nichtraucherinnen nach 25 Jahren

Betrachtet man diese Tabelle, so liegen mit 76 % der Überlebenden die Raucherinnen klar vor den Nichtraucherinnen, denn von diesen überlebten nur 69 %. Demnach ist das Rauchen scheinbar gesund! Können Sie sich vorstellen, wie dieses offenbar unsinnige Ergebnis zustande kommt?

Der verborgene Faktor bei dieser Studie ist das Alter der Versuchspersonen. Geht man vereinfachend von nur zwei Altersgruppen (»zu Beginn der Studie zwischen 18 und 64« und »älter als 64 zu Beginn der Studie«) aus, dann erhält man die unten abgebildete Situation. In beiden Altersgruppen sind es, ganz im Gegensatz zu den zuvor betrachteten Gesamtgruppen, offenbar die Nichtraucherinnen, die die größeren Chancen haben, die nächsten 20 Jahre zu erleben.

		Überlebt	Tot	Anzahl Frauen	Überlebende
18–64 J.	Raucherinnen	436	97	533	436/533 = 81,80 %
	Nicht-raucherinnen	474	65	539	474/539 = 87,94 %
> 64 J.	Raucherinnen	7	42	49	7/49 = 14,29 %
	Nicht-raucherinnen	28	165	193	28/193 = 14,51 %

Tabelle 3.5 Unterteilung nach Altersgruppen

		Überlebt	Tot	Anzahl Frauen	Überlebende
Gesamt	Raucherinnen	443	139	582	443/582 = 76,12 %
	Nicht-raucherinnen	502	230	732	502/732 = 68,58 %

Tabelle 3.5 Unterteilung nach Altersgruppen (Forts.)

Unterteilt man in die beiden Altersgruppen, dann sind in der jüngeren Gruppe 533 von 1072 Frauen Raucherinnen, d. h. ungefähr 50 %. In der Gruppe der älteren Frauen gibt es unter 242 Frauen nur 35 Raucherinnen, das sind etwa 14 %. Salopp ausgedrückt besteht diese Gruppe hauptsächlich aus alten Nichtraucherinnen. Beim Zusammenlegen der beiden Gruppen werden demnach junge Raucherinnen mit alten Nichtraucherinnen verglichen. Es ist einleuchtend, dass die jungen Raucherinnen trotz ihres Lasters noch länger als die alten Nichtraucherinnen zu leben haben.

Aufgabe 2: Das Simpson-Paradoxon ausnutzen

Stellen Sie sich vor, dass Sie der Hersteller eines neuen Medikaments X sind, das am Markt mit dem Medikament Y konkurriert. Um die Wirksamkeit der beiden Medikamente zu prüfen, wurde bereits eine Studie an 120 Personen durchgeführt. Leider schnitt dabei Ihr Medikament nicht gut ab. Es wurde 20 Personen verabreicht und war bei nur 2 Personen, d. h. bei 10 %, wirksam. Das Medikament Y wurde 100 Personen verabreicht. Bei 20 Personen hat sich Y als wirksam erwiesen, d. h. bei 20 %. Als Kenner des Simpson-Paradoxons denken Sie sich nun einen gemeinen Plan aus. Sie wissen, dass beide Medikamente bei jungen Patienten wirksamer als in dieser Studie sind. Bei jungen Patienten wirkt Medikament X in 40 % der Fälle und Medikament Y sogar in 50 % der Fälle. Also führen Sie eine zweite Studie durch, wieder mit 120 Patienten, dieses Mal aber nur mit jungen Patienten. Sie wollen erreichen, dass nach dem Zusammenlegen der beiden Datensätze Ihr Medikament X wirksamer als das Medikament Y erscheint.

Überlegen Sie sich, wie vielen der 120 Personen Ihr Medikament verabreicht werden soll, damit Ihr Plan gelingt.

Anmerkung: Es gibt eine ganze Reihe verschiedener Lösungen.

Zum Ausprobieren können Sie auch das Tabellenblatt »3.2 Medikament« verwenden und Ihre vermutete Personenanzahl in die dick umrahmte Zelle eintragen. Die restlichen Zellen werden dann automatisch ausgefüllt, sodass Sie sofort sehen können, ob Sie erfolgreich waren:

		Wirksam	Unwirksam	Anzahl	Wirksam
Erste Studie	X	2	18	20	10 %
	Y	20	80	100	20 %
Junge Patienten	X				40 %
	Y				50 %
Gesamt	X				
	Y				

Tabelle 3.6 Tabelle zum Ausfüllen

3.3 Der wohlgewählte Mittelwert

Sie sind der Chef einer Firma mit 9 Angestellten. Jeder Angestellte verdient im Monat 3 000 €. Wegen Ihrer besonderen Verantwortung gönnen Sie sich natürlich einen besseren Verdienst. 23 000 € ist das für Sie angemessene Salär.

Sie sind daran interessiert, die Arbeitsbedingungen in Ihrer Firma möglichst gut darzustellen. Wenn es dabei um den Verdienst geht, dann könnten Sie natürlich angeben, dass jeder Ihrer Angestellten 3 000 € verdient. Sie wollen aber, dass Ihre Firma besser dasteht. Aus diesem Grund berechnen Sie das mittlere Einkommen aller in der Firma beschäftigten Personen, wozu natürlich auch Sie gehören, auf folgende Weise:

$$(9 \cdot 3\,000 \text{ €} + 23\,000 \text{ €}) / 10 = 50\,000 \text{ €} / 10 = 5\,000 \text{ €}$$

Es klingt natürlich gut, wenn Sie sagen können, dass der durchschnittliche Verdienst in ihrer Firma 5 000 € beträgt. Das ist aber weit von der Wirklichkeit entfernt. Sie haben für die Berechnung das aus der Schule bekannte arithmetische Mittel verwendet. Dieses hat die Eigenschaft, dass es sehr empfindlich auf *Ausreißer* reagiert, und Ihr Gehalt ist ein solcher Ausreißer.

Das Beispiel zeigt: Das arithmetische Mittel eignet sich nicht, um die Verdienstsituation in Ihrer Firma zu beschreiben, es sei denn, Sie wollen bewusst manipulieren. Im nächsten Kapitel werden verschiedene Mittelwerte betrachtet, die für diese Situation besser geeignet sind.

3.4 Lösungen zu den Aufgaben

Aufgabe 1: Summenvergleich

a) $\frac{2}{5} = \frac{4}{10} > \frac{3}{10}$

Also bietet bei Peter der Eimer B die besseren Gewinnchancen.

b) Es ist $\frac{1}{7} < \frac{1}{6}$, denn der Nenner des ersten Bruchs ist größer als der des zweiten Bruchs.

Damit bietet auch bei Paul der Eimer B die besseren Gewinnchancen.

c) Die Gewinnchance für Eimer A beträgt $\frac{3+1}{10+7} = \frac{4}{17}$, die für Eimer B ist $\frac{2+1}{5+6} = \frac{3}{11}$.

Erweitert man den ersten Bruch mit 3 und den zweiten Bruch mit 4, dann ergibt sich

$$\frac{4}{17} = \frac{12}{51} \text{ und } \frac{3}{11} = \frac{12}{44}$$

Da bei gleichem Zähler der Nenner 51 größer als der Nenner 44

ist, so gilt: $\frac{12}{51} < \frac{12}{44}$. Auch nach dem Zusammenschütten ist es also günstiger, in Eimer

B zu greifen.

Aufgabe 2: Das Simpson-Paradoxon ausnutzen

Die einfachste Möglichkeit, das Ziel zu erreichen (Idee nach [DUB]), wäre, dass alle 120 jungen Patienten das Medikament X verabreicht bekommen. Das hätte dann zur Folge, dass in der Gesamtstichprobe das Medikament X in ca. 36 % der Fälle erfolgreich ist. Medikament Y hätte dann nur noch 20 % Wirksamkeit. Dieses Vorgehen wäre jedoch zu auffällig. Um doch noch die Gesamtstatistik zu retten, reicht es allerdings aus, wenn weniger als 120 Personen Medikament X bekommen. Die folgende Abbildung zeigt, dass auch für den Fall, dass 100 Personen das Medikament X und 20 Personen das Medikament Y bekommen, die Gesamtstatistik auf der Seite von X ist:

		Wirksam	Unwirksam	Anzahl	Wirksam
Erstes Krankenhaus	X	2	18	20	10 %
	Y	20	80	100	20 %
Junge Patienten	X	40	60	100	40 %
	Y	10	10	20	50 %
Gesamt	X	42	78	120	35 %
	Y	30	90	120	25 %

Tabelle 3.7 Simpson-Paradoxon