

# Das Grosse im Kleinen : Anmerkungen zum Begriff der Repräsentativität

Autor(en): **Riedo, Christof**

Objekttyp: **Article**

Zeitschrift: **Schweizerische Zeitschrift für Kriminologie = Revue suisse de criminologie = Rivista svizzera di criminologia = Swiss Journal of Criminology**

Band (Jahr): **12 (2013)**

Heft 1

PDF erstellt am: **09.08.2024**

Persistenter Link: <https://doi.org/10.5169/seals-1050760>

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

## **Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Christof Riedo

## Das Grosse im Kleinen

### Anmerkungen zum Begriff der Repräsentativität\*

#### Zusammenfassung

Um ein bestimmtes Phänomen zahlenmässig zu erfassen, können in der Regel nicht sämtliche Personen der zu untersuchenden Gesamtheit befragt werden. Meistens wird daher eine Stichprobe erhoben, die in Bezug auf die relevanten Merkmale dem Ganzen entspricht und damit «repräsentativ» ist. Die Ergebnisse einer Untersuchung der Stichprobe lassen sich nach dieser Annahme auf die Gesamtheit übertragen. Der nachfolgende Artikel setzt sich damit auseinander, wann diese Repräsentativität vorliegt und wie sie erreicht werden kann.

*Schlüsselwörter:* Repräsentativität, Statistik, Stichprobe, empirische Untersuchung.

#### Résumé

Afin de saisir un certain phénomène en termes numériques, il est normalement impossible de questionner toutes les personnes de l'entité en question. Pour cela, les sondages se basent régulièrement sur des échantillons qui, par rapport aux caractéristiques pertinents, correspondent le plus possible à l'entité entière et sont ainsi représentatifs. Dans cette hypothèse, les résultats d'une enquête de l'échantillon peuvent être reportés à l'entité entière. L'article suivant analyse quand il y a représentativité et comment celle-ci peut être atteinte.

*Mots-clés:* Représentativité, statistique, échantillon, enquête empirique.

#### Summary

To acquire data on a specific phenomenon, it is usually impossible to question everybody who is part of the entity under investigation. Therefore, surveys are usually based on samples that conform as much as possible to the entity regarding the relevant attributes. According to this hypothesis, the results of an inquiry of the sample can be transferred on the entire entity. The following article analyzes when representativeness exists and how it can be attained.

*Keywords:* Representativeness, statistics, sample, empirical research.

## I. Der König und sein Schlosspark

Es war einmal ein König, der sich für nichts so sehr interessierte wie für seinen geliebten Schlosspark. Und so kam es, dass er eines Tages seinen Hofgärtner zu sich rief und ihm auftrag herauszufinden, mit wie vielen Grashalmen die königliche Gartenanlage bestockt sei. Der Gärtner bedankte sich mit einem tiefen Bückling für die Ehre und zog sich zurück. Erst draussen wagte er es, sein Schicksal zu verfluchen und seinen geistig offenbar verwirrten König ins Pfefferland zu wünschen – allein, es half nichts: Er musste seine Pflicht erfüllen.

Allerdings, so dachte er sich, konnte es ja kaum die Meinung des Königs sein, dass er und seine Helfer monatelang auf Knien herumrutschten. Deshalb fragte er den Hofgelehrten um Rat, und der fand rasch eine einfache Lösung: Der Schlosspark – das wusste man aus früheren Vermessungen – hatte eine Fläche von rund 40 000 m<sup>2</sup>. Auf einem Teil der Parkfläche befanden sich Wege, kleine Bäche und verschiedene Nebengebäude, so dass sich die Rasenfläche auf geschätzte 35 000 m<sup>2</sup> belief. Er empfahl dem Gärtner, einen ganz beliebigen Quadratmeter Fläche auszuwählen und dort die Halme zu zählen. Das war eine mühsame und langweilige Tätigkeit, aber nach drei Tagen vermeldete der Gärtner, er habe auf der ausgewählten Rasenfläche genau 87 397<sup>1</sup> Halme gezählt. Nun war die Rechnung schnell gemacht: 35 000 Quadratmeter zu 87 397 Halmen, das ergab eine Gesamtzahl von 3 058 895 000. Als der Hofgärtner dem König dieses Ergebnis hinterbrachte, war dieser nicht wenig erstaunt:

\* Dr. Michael Beer und Dr. Manrico Glauser haben den vorliegenden Text mehrfach kritisch gelesen und mich auf zahlreiche Unzulänglichkeiten hingewiesen. Ich bedanke mich sehr herzlich. Was sich weiterhin an Blödsinn findet, habe ich natürlich selbst zu verantworten.

1 Die Angabe ist nicht ganz aus der Luft gegriffen: Wer das (riesige) Sortiment an Rasensamen durchforstet, stösst regelmässig auf Angaben zwischen 80 000 bis 100 000 Triebe pro m<sup>2</sup>. Da sich diese Zeitschrift nicht durch product-placement finanziert, wird auf entsprechende Hinweise verzichtet.



«Und wie hast Du das so schnell herausgefunden?» Als der Hofgärtner die angewandte Methode<sup>2</sup> erläutert hatte, reagierte der König ungehalten: «Und wie willst Du wissen, dass der ausgewählte Quadratmeter gleich viele Grashalme besitzt wie die anderen?» «Nun», meinte der Gärtner, «das weiss ich nicht. Aber warum sollte es anders sein?» «Keine Ahnung», erwiderte der König resigniert, «aber genügt das?»

## II. Ein Grundproblem empirischer Kriminologie

Kriminologie ist bekanntlich die Wissenschaft des kriminellen, abweichenden, devianten Verhaltens, und als solche interessiert sie sich *auch* für die Anzahl der tatsächlich begangenen Delikte. Insofern ist die Kriminologie eine empirische Wissenschaft.<sup>3</sup>

Nun lassen sich Straftaten nicht einfach zählen wie die Äpfel in der Früchteschale – die Aufgabe ist ungleich komplizierter. Die amtlichen (Anzeige-, Urteils-, Vollzugs- und Opferhilfe-) Statistiken erfassen jedenfalls von vorneherein nur das Hellfeld und damit einen mehr oder weniger kleinen Teil aller Straftaten. Der Filterungsprozess führt zu zahlreichen Verzerrungen; das Dunkelfeld ist nicht erfasst.<sup>4</sup>

2 Diese Methode wird in der Agrarwissenschaft nach wie vor angewandt, vgl. dazu etwa Spickermann/Plümer, *Gesunde Pflanzen* 2003, 222–227.

3 Zum Ganzen statt vieler: Göppinger, *Kriminologie*, § 1; Kaiser, *Kriminologie*, § 1 und 2; Killias/Kuhn/Aebi, *Kriminologie*, N 101 ff.; Kunz, *Kriminologie*, § 1; Schwind, *Kriminologie*, § 1 N 14 ff.

4 Hierzu wiederum statt vieler: Göppinger, *Kriminologie*, § 23 und § 30 N 25 ff.; Kaiser, *Kriminologie*, § 37; Killias/Kuhn/Aebi, *Kriminologie*, N 229 ff.; Kunz, *Kriminologie*, § 21; Kunz, *Zugänglichkeit*, 1 ff.; Schwind, *Kriminologie*, § 2, v. a. N 83 f.

5 Die Durchführung solcher Befragungen ist praktisch und methodologisch in hohem Masse anspruchsvoll. Auf die damit verbundenen Probleme soll hier nicht näher eingegangen werden. Vgl. zum Ganzen statt vieler: Diekmann, *Sozialforschung*, 434 ff.; Killias/Kuhn/Aebi, *Kriminologie*, N 229 ff.; Schirmer, *Sozialforschung*, 180 ff.; Schnell/Hill/Esser, *Sozialforschung*, 299 ff.; ferner Kunz, *Zugänglichkeit*, 18 ff., zum «doppelten Dunkelfeld», das sich daraus ergibt, dass in Zuge einer Befragung bestimmte Delikte allenfalls verschwiegen werden.

6 Das ist schon aus ökonomischen Gründen regelmässig ausgeschlossen, vgl. dazu nur Schnell/Hill/Esser, *Sozialforschung*, 254; Tiede/Voß, *Statistik*, 83.

7 Letztlich handelt es sich um eine Variante des Induktionsprinzips, also des Grundsatzes, dass es zulässig ist, vom Einzelnen auf das Allgemeine zu schliessen. Logisch gültig ist eine solche Schlussfolgerung nur dann, wenn die Generalisierbarkeit (eben: das Induktionsprinzip) als Prämisse akzeptiert wird – klassisch: Hume, *Enquiry*, § 4; Mill, *System*, Book III; Popper, *Logik*, 3 ff.; Russell, *Problems*, Chapter VI.; ferner etwa Chalmers, *Wissenschaft*, 7 ff.; Føllesdal/Walløe/Elster, *Argumentation*, 76 ff.; Kutschera, 189 ff.; Losee, 140 ff.

8 Vgl. zu diesem «statistischen Schluss» vom Teil auf das Ganze etwa Hartung, *Statistik*, 123 ff.; Schnell/Hill/Esser, *Sozialforschung*, 286; Bortz/Schuster, *Statistik*, 79.

9 <http://www.aerztekasse.ch/de/p12003424.html>.

Das tatsächliche Ausmass der Kriminalität lässt sich also (wenn überhaupt) nur durch Dunkelfeldforschung ermitteln. Deshalb werden heute (Täter-, Opfer- oder Informanten-) Befragungen durchgeführt.<sup>5</sup>

## III. Sample-Bildung

Weil in aller Regel nicht sämtliche Personen einer an sich interessierenden Grundgesamtheit befragt werden können<sup>6</sup>, wird zunächst ein sog. Sample gebildet. Es wird also eine bestimmte Anzahl von Personen als Teil der Grundgesamtheit ausgewählt, es wird eine Stichprobe erhoben.

Die Crux der Sample-Bildung besteht darin, dass nicht jede beliebige Untergruppe gleichsam eine verkleinerte Variante der Grundgesamtheit, eine Abbildung des Grossen im Kleinen bildet: Wer herausfinden will, wie viele Personen im Kanton Zürich eine Tätowierung aufweisen, wird *prima vista* keine zuverlässigen Daten generieren, wenn er seine Erhebung in einer Primarschule oder in einem Altersheim macht.

Exakte Angaben lassen sich demnach nur dann erzielen, wenn die Stichprobe mit Bezug auf die zu erforschenden Merkmale dem Ganzen entspricht. Deshalb basiert jeder Schluss von einer Stichprobe auf die Grundgesamtheit auf der unausgesprochenen Prämisse, die gewonnenen Ergebnisse liessen sich von der Stichprobe auf die Grundgesamtheit übertragen. Man könnte diese unausgesprochene Zusatzprämisse als Übertragbarkeitsprinzip<sup>7</sup> bezeichnen.<sup>8</sup>

## IV. Die Verwendung des Begriffs der Repräsentativität: Eine nicht repräsentative Medien- und Politschau

Die (behauptete) Übertragbarkeit erscheint terminologisch in unterschiedlichen Gewändern, häufig ist aber davon die Rede, eine Umfrage oder eine Stichprobe sei «repräsentativ»:

«Fast drei Viertel der Schweizerinnen und Schweizer sind zumindest teilweise der Meinung, dass nicht mehr alle Patienten die bestmögliche medizinische Behandlung erhalten. Dies zeigt eine repräsentative Studie...»<sup>9</sup>. «Sie haben Konzentrationsprobleme, schlafen nicht



genug und leiden an Entzugserscheinungen, sobald das Mobiltelefon nicht in Griffnähe ist. Die Symptome gleichen einer klassischen Verhaltenssucht, wie etwa beim Glücksspiel. Zu diesem Schluss kommt eine repräsentative Studie...».<sup>10</sup> «Wie eine repräsentative Umfrage im Auftrag des Apothekenmagazins «Diabetes Ratgeber» herausfand, litt im vergangenen Jahr 2011 jeder zweite Deutsche an Verdauungs- und Stoffwechselkrankheiten (48,9%).»<sup>11</sup> «Eine vom Institut GFS Zürich Anfang 2011 im Auftrag der Fondation Franz Weber durchgeführte repräsentative Umfrage hat ergeben, dass 89 Prozent der Befragten dem Verkauf von oder dem Handel mit Robbenprodukten kritisch bis ablehnend gegenüberstehen.»<sup>12</sup> «Angesichts des Umstandes, dass mehrere repräsentative Umfragen zu Fragen der Suizidhilfe (so zuletzt in der Zeitschrift «Reformiert» im August 2008 als auch in der Zeitschrift «L'Hebdo» im April 2009) ergeben haben, dass sich zwischen 72 und 75 Prozent der befragten Personen für die Möglichkeit des durch eine Organisation begleiteten sicheren und schmerzlosen Suizids in unserem Lande aussprechen...».<sup>13</sup> «Die Resultate aus ähnlichen Untersuchungen zeigen jedoch, dass die Beteiligung mit rund 40 Prozent äusserst schwach ist, sodass die Resultate nicht als repräsentativ betrachtet werden können.»<sup>14</sup>

Offenkundig handelt es sich bei der jeweils behaupteten «Repräsentativität» um eine Art Qualitätssiegel.<sup>15</sup> Eine Erklärung dessen, was das eigentlich heissen soll, sucht der interessierte Leser regelmässig vergebens. Grund genug, die Sache genauer zu betrachten...

## V. Ein Blick ins Wörterbuch

Der Ausdruck «repräsentieren» stammt ab vom lateinischen «repraesentare», vergegenwärtigen. «Repräsentativ» meint zunächst «würdig vertretend», teilweise auch «eindrucksvoll» («eine repräsentative Villa»). Eine «repräsentative Demokratie» ist eine Staatsform, bei der gewählte Repräsentanten den Willen des Volkes zum Ausdruck bringen. Oft wird damit aber auch die Eigenschaft umschrieben, ein Einzelner sei mit Bezug auf ein bestimmtes Merkmal ein typisches Mitglied einer Gruppe («ein repräsentatives Beispiel für die italienische Architektur des 16. Jahrhunderts»).<sup>16</sup>

## VI. Zur Repräsentativität von Stichproben

Eine empirische Untersuchung gilt dann als «repräsentativ», wenn sich die Ergebnisse einer Teilerhebung auf eine sog. Grundgesamtheit (noch in einem vagen Sinne) «übertragen» lassen.<sup>17</sup>

Das ist der Fall, wenn die Stichprobe der Grundgesamtheit hinsichtlich bestimmter (vordefinierter) Merkmale (mehr oder weniger genau) entspricht.<sup>18</sup> Offen bleibt freilich, wie weit diese Entsprechung tatsächlich bestehen muss. Repräsentativität in einem strikten Sinne würde eine exakte Entsprechung voraussetzen. Am einfachsten lässt sich das Gesagte anhand eines Beispiels erläutern, mit dem üblicherweise Gymnasiasten gequält werden:

In einem (undurchsichtigen) Behälter befinden sich 100 Kugeln, 60 weisse und 40 schwarze. Nun werden (im Sinne einer Stichprobe) fünf Kugeln gezogen. Die gezogenen Kugeln sind für die Grundgesamtheit dann und nur dann exakt repräsentativ, wenn 3 der 5 Kugeln weiss und 2 schwarz sind, denn dann und nur dann wäre die relative Häufigkeit<sup>19</sup> der

10 <http://www.20min.ch/wissen/news/story/31269176>.

11 [http://www.focus.de/gesundheit/diverses/ernaehrung-schwierige-verdauung-umfrage-durchfall-verstopfung-oder-blhungen-viele-deutsche-berichten-von-beschwerden\\_aid\\_807150.html](http://www.focus.de/gesundheit/diverses/ernaehrung-schwierige-verdauung-umfrage-durchfall-verstopfung-oder-blhungen-viele-deutsche-berichten-von-beschwerden_aid_807150.html).

12 Motion Freysinger (11.3635 – Importverbot für Robbenprodukte), abrufbar unter: [http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch\\_id=20113635](http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch_id=20113635).

13 Anfrage Gross (09.1154 – Bericht des EJPD über vertiefte Abklärungen zur organisierten Suizidhilfe), abrufbar unter: [http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch\\_id=20091154](http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch_id=20091154).

14 Antwort des Bundesrates vom 30.5.2008 auf die Interpellation Marra (08.3122 – Teilnahme der Schweiz an der nächsten internationalen Studie zum Illetrismus [sic!]), abrufbar unter: [http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch\\_id=20083122](http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch_id=20083122).

15 Entsprechend selten wird mit *nicht* repräsentativen Studien für ein Anliegen geworben, vgl. aber etwa die Begründung der Interpellation der FDP (98.3289 – Regionale Arbeitsvermittlungszentren. Verbesserung der Aktivitäten), abrufbar unter: [http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch\\_id=19983289](http://www.parlament.ch/d/suche/seiten/geschaefte.aspx?gesch_id=19983289): «Eine *nicht* repräsentative Umfrage der Freisinnig-demokratischen Partei der Schweiz brachte ein etwas differenzierteres Bild zutage: Rund 10 Prozent der befragten Arbeitgeber, die im damaligen Zeitraum (Februar bis April 1998) offene Stellen nicht besetzen konnten, stellten eine erfolglose Tätigkeit der RAV fest» (Hervorhebung durch den Autor).

16 Die fraglichen Informationen finden sich in jedem beliebigen Wörterbuch. Konsultiert wurden der «Brockhaus», der «Wahrig» und der «Duden».

17 Vgl. auch hierzu die bereits in Fussnote 16 erwähnten Nachschlagewerke. In den Details bestehen beträchtliche begriffliche Unterschiede, vgl. Kruskal/Mosteller, International Statistical Review 1980, 169 ff.; für eine historische Einführung Desrosières, Zahlen, 235 ff.

18 Angeschrieben ist also von vornherein lediglich eine spezifische Repräsentativität (mit Bezug auf bestimmte Merkmale) und nicht eine globale Repräsentativität (Bortz/Schuster, Statistik, 80; Quatember, Repräsentativität, 236 f.).

19 Analoges würde gelten, wenn man sich etwa für Durchschnittswerte interessieren würde: Exakte Repräsentativität läge vor, wenn die Durchschnittswerte von Stichprobe und Grundgesamtheit identisch wären.



interessierenden Merkmale bei Grundgesamtheit und Stichprobe identisch.<sup>20</sup>

Unter pragmatischen Gesichtspunkten wäre man geneigt, sich mit weniger zufrieden zu geben, aber zu diesem abgeschwächten Begriff der Repräsentativität dann später ...

## VII. Zufallsstichproben

Üblicherweise soll «Repräsentativität» erzielt werden, indem aus der an sich interessierenden Grundgesamtheit durch eine *Zufallsauswahl* eine Teilgesamtheit ausgewählt wird; es wird eine sog. Zufallsstichprobe gezogen (ein «Randomsample» gebildet).

Eine (einfache<sup>21</sup>) Zufallsstichprobe liegt genau dann vor, wenn jedes Element der Grundgesamtheit die gleiche Chance hat, «gezogen» zu werden, also zur Stichprobe zu gehören<sup>22</sup>, wenn es mithin vom Zufall abhängt, ob ein Element der Grundgesamtheit zur Stichprobe gehört.

Nun ist aber das Bilden einer echten Zufallsstichprobe faktisch häufig ausgeschlossen: Wenn mit Hilfe eines Zufallsgenerators 1000 Personen auf «tel.search.ch» ausgewählt werden, können sich von vorneherein nur Personen äussern, die über einen registrierten Telefonanschluss verfügen. Wenn dann versucht wird, die ausgewählten Personen zu erreichen, wird dies bei all jenen scheitern, die in der fraglichen Zeit ortsabwesend sind usw. Zusätzlich vorausgesetzt ist schliesslich die Bereitschaft, sich an einer Umfrage überhaupt zu beteiligen. All dies sind unausgesprochene Auswahlkriterien; mit der zufälligen Auswahl erfolgt eine nicht zufällige Vorauswahl.<sup>23</sup>

Das gilt auch für weite Bereiche der empirisch orientierten kriminologischen Forschung: Mit Täter- und Opferbefragungen beispielsweise werden unter Umständen Personen nicht erreicht, die an sich besonders interessieren würden, bei Täterbefragungen etwa Strafgefangene und Kriminaltouristen, bei Opferbefragungen Tote oder Schwerverletzte.<sup>24</sup> Es fehlen dann unter Umständen Teile der Bevölkerung, die wichtige Angaben machen könnten.

Die tatsächlich *potentiell* zum Sample gehörigen Elemente decken also die Grundgesamtheit regelmässig gar nicht vollständig ab. Zufällig ist die Auswahl dann nur mit Bezug auf eine andere Bezugsgrösse, die sog. Stichprobenbasis – und die ist für die Grundgesamtheit nicht (oder nicht zwingend) repräsentativ.<sup>25</sup>

Selbst wenn man aber davon ausgeht, dass die Stichprobenbasis der Grundgesamtheit entspricht<sup>26</sup> (oder für diese repräsentativ ist), führt die anschliessende Stichprobenerhebung nicht ohne weiteres zu verlässlichen Ergebnissen:

In einem (undurchsichtigen) Behälter befinden sich 100 Kugeln. Jede dieser Kugeln ist entweder vollständig weiss oder vollständig schwarz (und keine hat eine andere Farbe). Nun werden Kugeln gezogen. Wie gross ist die Wahrscheinlichkeit, dass das Sample der gezogenen Kugeln ein genau zutreffendes Bild von der tatsächlichen Farbverteilung aller Kugeln vermittelt?

Offenkundig hängt das Ergebnis zunächst davon ab, wie sich die Grundgesamtheit zusammensetzt. Sind sämtliche Kugeln schwarz, werden sich auch in der Stichprobe ausschliesslich schwarze Kugeln finden – die

20 Quatember, Repräsentativität, 237.

21 Die wichtigste Alternative zur einfachen Zufallsstichprobe ist die systematische Stichprobe (dazu unten, VIII.). Teilweise werden auch sog. Klumpenstichproben gezogen, vgl. zum Ganzen etwa Bortz/Schuster, Statistik, 81 f.; Diekmann, Sozialforschung, 380 ff.; Hartung, Statistik, 273 ff.; Quatember, Repräsentativität, 237 ff.; Schirmer, Sozialforschung, 174 ff.; Schnell/Hill/Esser, Sozialforschung, 264 ff.; ferner Freedman/Pisani/Purves, Statistics, 333 ff.

22 Diese Definition scheint sich durchgesetzt zu haben, vgl. etwa Bortz/Schuster, Statistik, 80; Diekmann, Sozialforschung, 380 ff.; Hartung, Statistik, 273 f.; Quatember, Repräsentativität, 237; Schnell/Hill/Esser, Sozialforschung, 259; Tiede/Voß, Statistik, 87; Schwind, Kriminologie, § 9 N 34; Stegmüller, Probleme und Resultate IV/2, 135 f.

23 Zu dieser Problematik Bortz/Schuster, Statistik, 81; Diekmann, Sozialforschung, 376 ff. und 410 ff.; Hartung, Statistik, 271 f.; Schnell/Hill/Esser, Sozialforschung, 257 ff.; lesenswert auch Krämer, Statistik, 97 ff.; ferner Görnitz/Moser, der markt 2000, 156 ff.

24 Dazu Kunz, Zugänglichkeit, 20 ff., mit einer schönen Parabel. Ferner Vignando/Haas, Crimiscopes 2001, 4: «die Repräsentativität unserer Stichprobe könnte in dieser Hinsicht etwas eingeschränkt sein, denn Männer mit nicht mehr kompensierbaren psychischen Defiziten und solche, welche zu einer langjährigen Freiheitsstrafe verurteilt wurden, befinden sich nicht in der Rekrutenschule [und konnten deshalb nicht befragt werden].»

25 Klassisches Beispiel ist das «Literary Digest Disaster»: 1936 versuchte das Magazin «Literary Digest» den Ausgang der US-Präsidentenwahl vorherzusagen. Zu diesem Zweck wurde eine Befragung durchgeführt. Obwohl sich am Ende nicht weniger als 2.3 Mio. Stimmberechtigte an der Umfrage beteiligten, erwiesen sich die Ergebnisse als völlig falsch: Statt des vorausgesagten Erdrutschsieges des Republikaners Landon wurde Roosevelt in seinem Amt bestätigt. Die enorme Abweichung zwischen Stichprobe und Grundgesamtheit dürfte in erster Linie darauf zurückzuführen sein, dass an der telefonischen Befragung nur Personen teilnehmen konnten, die ein Telefon besaßen – und das waren nur Gutbetuchte (die traditionell für die Republikaner stimmen). Vgl. dazu v. a. Freedman/Pisani/Purves, Statistics, 334 ff.; ferner Diekmann, Sozialforschung, 373 ff.; Schwind, Kriminologie, § 9 N 33; Desrosières, Zahlen, 229.

26 Technischer ausgedrückt: Wenn jedes Element  $x$  der Grundgesamtheit  $G$  auch Element der Stichprobenbasis  $S_b$  ist und umgekehrt:  $(\forall x) (x \in G \iff x \in S_b)$ .



Wahrscheinlichkeit  $p$  eines genau zutreffenden Ergebnisses beträgt 1. Umgekehrt lassen sich Verteilungen, die Primzahlen enthalten, durch eine Zufallsstichprobe überhaupt nicht exakt ermitteln: Die Verteilung: «37 weisse und 63 schwarze Kugeln» lässt sich mit einer Stichprobe von  $k < 100$  nicht exakt abbilden.

Nehmen wir nun an, in Tat und Wahrheit seien 60 Kugeln der Grundgesamtheit weiss, 40 schwarz. Bei dieser Ausgangslage hängt die Wahrscheinlichkeit eines vollständig richtigen Ergebnisses von der Grösse der Stichprobe ab: Das Ziehen von weniger als 5 Kugeln führt zwingend zu einem falschen Ergebnis (denn eine Verteilung von 3:2 lässt sich offenkundig mit weniger als fünf Kugeln gar nicht erzielen). Werden  $k=5$  von  $N=100$  Kugeln gezogen, beträgt die Wahrscheinlichkeit  $p$  eines exakten Ergebnisses (bei  $k_1=3$  von  $n_1=60$  weissen und  $k_2=2$  von  $n_2=40$  schwarzen Kugeln)<sup>27</sup>:

$$p = \frac{\binom{n_1}{k_1} * \binom{n_2}{k_2}}{\binom{N}{k}} = \frac{\binom{60}{3} * \binom{40}{2}}{\binom{100}{5}} = 0.35$$

Mit anderen Worten: In rund zwei Dritteln aller Fälle führt die Stichprobe zu einem unzutreffenden Ergebnis.

Wer nun meinen würde, mit dem Ziehen jeder weiteren Kugel nehme die Wahrscheinlichkeit eines exakten Ergebnisses stetig zu, unterliegt einem Irrtum: Auch das Ziehen von sechs, sieben, acht oder neun Kugeln muss nämlich ein falsches Bild vermitteln, denn die bestehende Verteilung lässt sich nur erzielen, wenn sich die Anzahl der gezogenen Kugeln ohne Rest durch fünf teilen lässt.

Selbst das wiederholte Ziehen fünf weiterer Kugeln führt indessen nicht zu höheren Trefferwahrscheinlichkeiten<sup>28</sup>:

Anzahl Ziehungen	Wahrscheinlichkeit
5	0.35
10	0.26
15	0.22
20	0.20
25	0.19
30	0.18
35	0.17
40	0.16
45	0.16
50	0.16
55	0.16
60	0.16

Anzahl Ziehungen	Wahrscheinlichkeit
65	0.17
70	0.18
75	0.19
80	0.20
85	0.22
90	0.26
95	0.35
100	1

Die Wahrscheinlichkeit einer vollständigen Übereinstimmung nimmt also mit fortschreitender Vergrößerung des Samples zunächst ab, und steigt erst später wieder an, bleibt aber für Stichproben von  $k < 100$  unter 0.4.

Noch weit dramatischer wird die Situation, wenn die Zusammensetzung der Grundgesamtheit stärker fragmentiert ist, wenn sie «inhomogen aufgebaut» ist. Nehmen wir an, von den 100 Kugeln seien deren 17 schwarz, 8 grün, 11 blau, 34 rot, 9 weiss und 21 gelb. Bei solchen Verteilungen beträgt die Wahrscheinlichkeit einer vollständigen Übereinstimmung unabhängig von der Grösse des Samples<sup>29</sup> 0.

Die Zufallsstichprobe wird der Grundgesamtheit nur ganz ausnahmsweise vollständig entsprechen – exakte Repräsentativität ist durch Zufall kaum zu erzielen.

## VIII. Systematische Stichproben

Ein möglicher Ausweg scheint nun prima vista darin zu bestehen, die Reichweite des Zufalls bei der Sample-Bildung nach sinnvollen Kriterien einzugrenzen, also die Varianz zu minimieren. Eine angepasste Version unseres Beispiels mag dies verdeutlichen:

In einem (undurchsichtigen) Behälter befinden sich 50 Kugeln und 50 Würfel mit folgender Farbverteilung:

	Kugeln	Würfel	Total
Weiss	15	45	60
Schwarz	35	5	40
Total	50	50	100

<sup>27</sup> Es handelt sich um ein «ungeordnetes Ziehen ohne Zurücklegen». Die entsprechende Formel findet sich in jedem beliebigen Grundlagenwerk der Kombinatorik bzw. Wahrscheinlichkeitsrechnung – auf entsprechende Hinweise wird man deshalb verzichten können.

<sup>28</sup> Die Berechnung richtet sich nach der bereits erwähnten Formel.

<sup>29</sup> Eine Ausnahme gilt lediglich, wenn die Grösse des Samples jener der Grundgesamtheit entspricht (wenn also sämtliche Kugeln gezogen werden). Aber dann handelt es sich nicht mehr um eine Stichprobe, sondern um eine Vollerhebung.



Die Aufgabe besteht wiederum darin, anhand einer Stichprobe die Farbverteilung der Grundgesamtheit zu ermitteln. Nun hängt das Ergebnis offenkundig stark davon ab, welche Formen die gezogenen Gegenstände aufweisen: Ziehen wir 10 Kugeln, werden diese wahrscheinlich in der Mehrzahl schwarz sein – ziehen wir 10 Würfel, werden wir häufiger überwiegend weisse Gegenstände in den Händen halten.

Das Ziehen einer gleichen Anzahl von Kugeln und Würfeln verspricht deshalb bessere Ergebnisse. Das lässt sich rechnerisch ohne weiteres aufzeigen, wenn man die Wahrscheinlichkeiten *exakter* Repräsentativität miteinander vergleicht:

Anzahl Ziehungen $n$	Wahrscheinlichkeit $p_1$ bei Zufallsstichprobe <sup>30</sup>	Wahrscheinlichkeit $p_2$ bei systematischer Stichprobe <sup>31</sup>
10	0.26	0.34
20	0.20	0.25
30	0.18	0.22
40	0.16	0.21
50	0.16	0.20
60	0.16	0.21
70	0.18	0.22
80	0.20	0.25
90	0.26	0.34
100	1	1

Man wird also bei der Sample-Bildung sinnvollerweise darauf achten, dass die entsprechende Verteilung insoweit der Grundgesamt-

heit entspricht.<sup>32</sup> Man erhebt dann eine «systematische» Stichprobe<sup>33</sup>.

Eine solche Vorgehensweise setzt das Wissen voraus:

- welche bekannten Merkmale mit den zu untersuchenden Merkmalen signifikant zusammenhängen (in unserem Beispiel: das Wissen darum, dass Form und Farbe signifikant korrelieren), und
- wie diese bekannten Merkmale in der Grundgesamtheit verteilt sind (in unserem Beispiel: wie viele Gegenstände in unserem Behälter kugel- bzw. würfelförmig sind).

Das scheint teilweise unproblematisch: So glauben wir etwa mit ausreichender Sicherheit zu wissen, dass die Wahrscheinlichkeit, dass eine Person eine Straftat begangen hat, massgeblich von deren Geschlecht und deren Alter abhängt.<sup>34</sup> Man wird deshalb auch hier bei der Sample-Bildung für eine proportionale Verteilung hinsichtlich Geschlecht und Alter sorgen – und das ist mindestens theoretisch ohne weiteres möglich, denn die Zusammensetzung der Bevölkerung hinsichtlich Geschlecht und Alter ist bekannt<sup>35</sup>.

In anderen Bereichen tappen wir völlig im Dunkeln. So wäre es etwa durchaus denkbar, dass die Anzahl durch eine Person begangener Delikte gegen Leib und Leben signifikant mit deren Körpergrösse zusammenhängt. Allein: Wir haben keine Ahnung, ob dieser Zusammenhang tatsächlich besteht, wir verfügen noch nicht einmal über verlässliche Zahlen hinsichtlich der Körpergrösse innerhalb der Bevölkerung. Eine entsprechende Quotenbildung scheidet aus. Mit anderen Worten: Eine Gewichtung anhand massgeblicher Merkmale vermag die Verlässlichkeit der stichprobeweise erhobenen Zahlen nur teilweise zu erhöhen, weil ein entsprechendes Vorgehen Wissen voraussetzt, über das wir allenfalls gar nicht verfügen.

## IX. Eine erste Zwischenbilanz

Wie man es auch dreht und wendet: Es ist höchst unwahrscheinlich, dass eine Teilerhebung zu (strikte) repräsentativen Ergebnissen führt.<sup>36</sup>

«Repräsentativität» im Sinne einer exakten Entsprechung zwischen Stichprobe und Grundgesamtheit mit Bezug auf bestimmte vordefinierte Merkmale scheint kaum erreich-

30 Die Formel ist die bereits verwendete, entsprechend sind auch die Werte insoweit mit jenen gemäss Tabelle in Abschnitt VII. identisch.

31 Die Grundformel ist wiederum dieselbe, wobei hier mit bedingten Wahrscheinlichkeiten operiert werden muss: Bei einer Stichprobe  $k$  ziehen wir zuerst  $k/2$  Kugeln. Damit am Ende die exakt repräsentative Farbverteilung resultiert, ist beim Ziehen der Würfel eine besondere Farbverteilung vorausgesetzt – je nachdem, welche Farben die bereits gezogenen Kugeln haben. Die jeweiligen Wahrscheinlichkeiten sind alsdann zu multiplizieren und die Resultate in der Folge zu addieren.

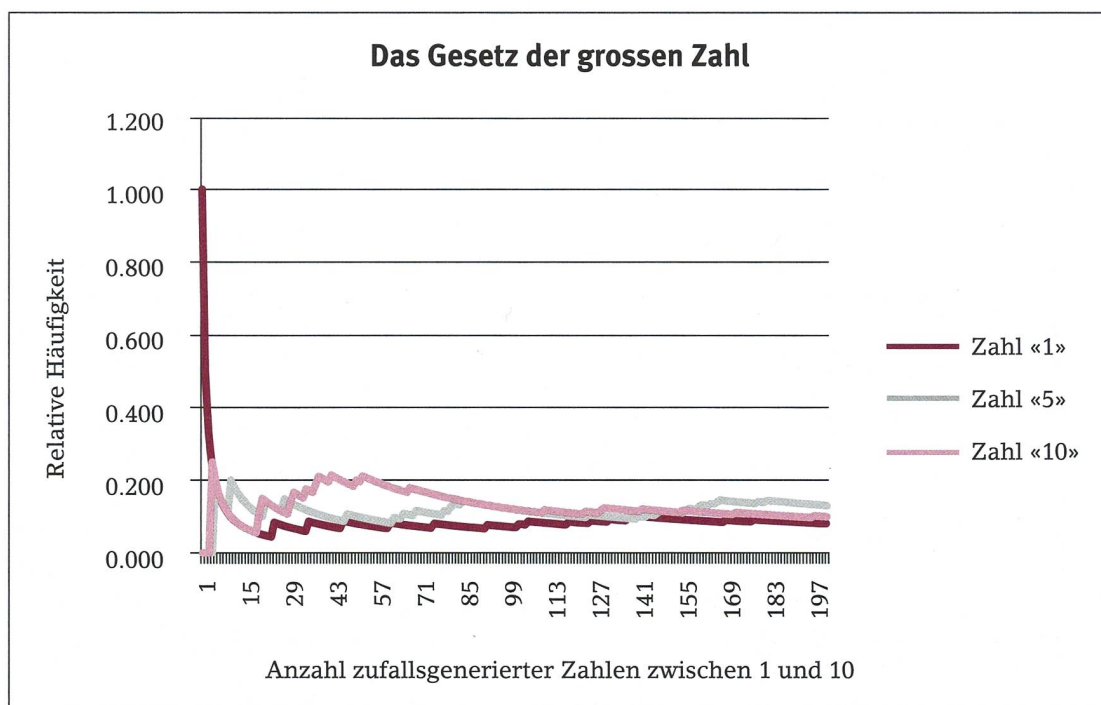
32 Gemeint ist hier eine proportionale Entsprechung. In unserem Beispiel wäre also eine 50:50-Quote zu beachten. Für eine mathematisch korrekte Definition der proportionalen Entsprechung vgl. etwa Hartung, Statistik, 285.

33 Die Terminologie ist nicht einheitlich, teilweise ist auch von einer «geschichteten Zufallsstichprobe», einer «Quotenstichprobe» oder einem «Quota-Sample» die Rede.

34 Durchaus korrekt deshalb die Antwort des Bundesrates vom 4. 3. 2011 auf die Interpellation der SVP (10.3962 – Problematische Einwanderung aus dem afrikanischen Kontinent): «Die Studie erfasste lediglich die Altersgruppe von 18 bis 34 Jahren und stützte sich somit auf eine äusserst kleine und demnach wenig repräsentative Stichprobe. Generelle Aussagen zur Kriminalität von Staatsangehörigen Nigerias oder Angolas in der Schweiz lassen sich daher daraus nicht ziehen.»

35 Das Bundesamt für Statistik BFS verfügt über entsprechendes Zahlenmaterial.

36 Vgl. auch Ludwig-Mayerhofer, 1.



bar. Das vermag freilich weniger die Qualität von Forschungsergebnissen in Frage zu stellen – es weckt vielmehr Zweifel am so gefassten Begriff der Repräsentativität...

## X. Pragmatismus oder Das Gesetz der grossen Zahl

Repräsentativität im Sinne exakter Entsprechung scheint nämlich nicht nur weitgehend illusorisch, sondern auch allzu ambitiös. Statistische Erhebungen verfolgen ein spezifisches Erkenntnisinteresse, und je nach Zweck des Forschungszieles mögen weniger genaue Zahlen durchaus genügen. So wird es etwa kriminalpolitisch kaum eine Rolle spielen, ob in einem bestimmten Gebiet während eines bestimmten Zeitraumes 1387 oder 1398 Raubüberfälle verübt wurden.

Bei nüchterner Betrachtung ist deshalb die Wahrscheinlichkeit exakter Übereinstimmung von Stichprobe und Grundgesamtheit weniger entscheidend als die Hoffnung, dass mit einer Vergrösserung des Samples die Wahrscheinlichkeit einer erheblichen Abweichung von der Grundgesamtheit zusehends sinkt. Wir kennen es aus anderen Bereichen:

Eine geworfene Münze zeigt mit einer Wahrscheinlichkeit von 0.5 Kopf bzw. Zahl. Je öfter die Münze geworfen wird, desto mehr

werden sich die tatsächlich erzielten Ergebnisse dem errechneten Erwartungswert annähern.<sup>37</sup>

Für die oben stehende Abbildung wurden mit Hilfe eines Zufallsgenerators<sup>38</sup> (ganze) Zahlen zwischen 1 und 10 gebildet. Die Kurve zeigt das relative Auftreten der Zahl 1, 5 und 10 im Verhältnis zur Anzahl der gebildeten Zufallszahlen.

Dabei lässt sich unschwer erkennen, dass sich der Wert immer mehr dem erwarteten Wert von 0.1 annähert.

Nicht anders verhält es sich mit Stichproben: Je grösser die Stichprobe (oder je mehr Stichproben), desto geringer wird die Abweichung von der Grundgesamtheit. Wir scheinen deshalb davon ausgehen zu dürfen, dass sich mit Hilfe von Stichproben zwar kaum je absolut präzise, aber unter pragmatischen Gesichtspunkten allenfalls brauchbare Ergebnisse erzielen lassen.<sup>39</sup>

<sup>37</sup> Vorausgesetzt ist freilich, dass das Ergebnis tatsächlich vom Zufall abhängt. Es muss sich also um eine «Laplace-Münze» handeln.

<sup>38</sup> Der Ausdruck «Zufallsgenerator» ist an sich missverständlich: Die Zahlen werden durch einen (per definitionem deterministischen) Algorithmus erzeugt. Das Thema soll hier indes nicht weiter vertieft werden – unter heuristischen Gesichtspunkten scheint das Beispiel genügend.

<sup>39</sup> Vgl. etwa auch AB N 2002 319 (Votum Janiak): «Sie kennen die Ergebnisse dieser Untersuchung, welche zweifellos als repräsentativ eingestuft werden kann, ist doch von einer statistischen Auswertungsquote von gut zwei Dritteln aller Ratsmitglieder – Stände- und Nationalräte – auszugehen».



## XI. Eine zweite Zwischenbilanz

Wenn also Repräsentativität im strikten Sinne nicht erreicht werden kann, empirische Sozialforschung aber unter gewissen Voraussetzungen in der Lage ist, mit Hilfe von Stichproben *mehr oder weniger* genaue Aussagen über eine Grundgesamtheit zu machen, liegt es nahe, den Begriff der Repräsentativität weiter zu fassen.<sup>40</sup>

Als repräsentativ muss also eine Stichprobe auch dann gelten, wenn sie der Grundgesamtheit mit Bezug auf die interessierenden Merkmale *in gewissem Masse* entspricht. Zu lösen bleibt dann freilich eine Spielart des Sorites-Paradoxes<sup>41</sup>: Wenn eine Stichprobe auch dann repräsentativ sein soll, wenn sie mit Bezug auf die interessierenden Merkmale nicht exakt der Grundgesamtheit entspricht, verliert der Begriff jede Schärfe, es sei denn, es werde ein Mass tolerierbarer Abweichung festgelegt, dessen Einhaltung sich verlässlich überprüfen lässt.

Dieses Vorgehen setzt dreierlei voraus:

- 1) ein Mass der Abweichung zwischen Stichprobe und Grundgesamtheit – ein solches Mass lässt sich ohne weiteres finden, denn die entsprechenden Ergebnisse werden regelmässig quantifiziert;
- 2) die Festsetzung einer tolerierbaren Abweichung – idealerweise würde diesbezüglich Konsens bestehen, denkbar wäre aber auch die mehr oder weniger willkürliche Festsetzung eines zulässigen Abweichungsniveaus und dessen Offenlegung<sup>42</sup>;
- 3) eine Methode, die es erlauben würde, verlässlich festzustellen, ob ein bestimmtes Stichprobenergebnis das festgesetzte Ausmass der zulässigen Abweichung von der Grundgesamtheit einhält oder nicht – eine solche Methode gibt es nicht und kann es auch nicht geben, denn das würde eine Vollerhebung voraussetzen, und dann wäre die Stichprobenerhebung sinnlos.

Auch ein weiterer Begriff der Repräsentativität ist deshalb untauglich. Das Konzept erscheint als funktionales Äquivalent eines obskuren Biolabels auf einer Packung *Marshmallows*.<sup>43</sup>

40 Ähnlich Quatember, Repräsentativität, 237.

41 Es handelt sich um eine der sieben Fangfragen, die dem Megariker Eubulides von Milet zugeschrieben werden:

A: «Bildet ein Korn einen Haufen?»

B: «Nein.»

A: «Bilden zwei Körner einen Haufen?»

B: «Nein.»

A: «Bildet das Hinzufügen eines Kornes einen Haufen?»

B: «Nein.»

A: «Also bildet keine Anzahl von Körnern einen Haufen?»

Vgl. zum Ganzen Bult/Schmidt, Sorites.

42 Beispiele: Der Durchschnittswert der Stichprobe darf vom Durchschnittswert der Grundgesamtheit um nicht mehr als X% abweichen; die einzelnen Anteilswerte der Stichprobe dürfen von den entsprechenden Anteilswerten der Grundgesamtheit um nicht mehr als X% abweichen o. ä.

43 Schnell/Hill/Esser, Sozialforschung, 287: Es handle sich «um kein exakt definiertes Gütekriterium»; ähnlich Schirmer, Sozialforschung, 74; ferner Tiede/Voß, Statistik, 83 ff.; Diekman, Sozialforschung, 430 ff.; Quatember, Repräsentativität, 236. Der Begriff wird denn auch in erster Linie im Zuge der medialen Aufbereitung von Forschungsergebnissen verwendet (vgl. die Zitate in Abschnitt IV.). Er findet sich aber durchaus auch in wissenschaftlichen Publikationen. Aus dem Bereich der Kriminologie etwa den – übrigens höchst interessanten – Beitrag von Kuhn/Villettaz/Jayet/Willi, Crimscope 2002, 1 ff., wo mehrfach von «Repräsentativität» die Rede ist, ohne dass auch nur im Ansatz erklärt würde, was das heissen soll. Nicht anders verhält es sich in anderen Wissenschaften. Spickermann/Plümer, Gesunde Pflanzen 2003, 222–227, untersuchen, ob die angewandten Stichprobenverfahren von landwirtschaftlichen Expertensystemen repräsentativ seien – erklären aber nicht, was sie mit «repräsentativ» meinen. Wer schliesslich den Suchbegriff «repräsentativ» auf der Webseite der «Zeitschrift für Psychologie» eingibt, erhält 114 Treffer, und auch hier fehlt eine eigentliche Definition mit schöner Regelmässigkeit.

44 Vgl. zum Nachfolgenden statt vieler etwa Ludwig-Mayerhofer, 3 ff.; Diekmann, Sozialforschung, 401 ff.; Hartung, Statistik, 129 ff.; Bortz/Schuster, Statistik, 93 f.; Killias/Kuhn/Aebi, Kriminologie, N 253; Schnell/Hill/Esser, Sozialforschung, 259 ff.; Polasek, Statistik, 211 ff.; Tiede/Voß, Statistik, 131 ff.; Freedman/Pisani/Purves, Statistics, 381 ff.; Stegmüller, Probleme und Resultate IV/2, 189 f.

## XII. Raffinesse und Verwirrung

Gearbeitet wird deshalb heute insbesondere mit sog. Vertrauensintervallen, die es erlauben sollen, die Verlässlichkeit empirischer Erhebungen bis zu einem gewissen Grade einzuschätzen.<sup>44</sup>

In einem (undurchsichtigen) Behälter befinden sich 100 Kugeln. Nun ziehen wir 10 Kugeln und stellen fest: 5 davon sind weiss, 5 haben eine andere Farbe. Mit welcher Sicherheit und mit welcher Genauigkeit dürfen wir nun davon ausgehen, dass die Hälfte der Kugeln weiss ist?

Die Grundidee des Lösungsansatzes ist simpel: Würden wir das Experiment mehrfach wiederholen (also die 10 gezogenen Kugeln jeweils wieder zurücklegen, alles gut mischen und erneut 10 Kugeln ziehen), würden wir nach den Gesetzen der Wahrscheinlichkeit häufiger Stichproben erheben, die hinsichtlich der Farbverteilung mit der Grundgesamtheit übereinstimmen oder dieser mindestens nahekommen. Auch dies ergibt sich aus dem Gesetz der grossen Zahl. Wie nahe die Stichprobenergebnisse im Durchschnitt der Grundgesamt-



heit entsprechen, hängt ab von der «Standardabweichung» (vom «Standardfehler») und damit von der möglichen Streuung der Stichprobenergebnisse. Diese wiederum ist abhängig von der Stichprobengrösse und der Streuung der interessierenden Merkmale in der Grundgesamtheit.<sup>45</sup>

Mit Bezug auf *Anteilswerte*<sup>46</sup> beträgt der Standardfehler  $\sigma$  bei einem Stichprobenumfang  $k=10^{47}$ , einem Stichprobenanteil weisser Kugeln  $Pw=5/10=0.5$ :

$$\sigma = \frac{\sqrt{Pw*(1-Pw)}}{\sqrt{k}} = \frac{\sqrt{0.5*(1-0.5)}}{\sqrt{10}} = 0.1581$$

Das sog. Vertrauens- oder Konfidenzintervall wird nun in der Regel für eine «Sicherheit» von 95% berechnet. Diese Verlässlichkeit ist erreicht, wenn man zum Stichprobenwert den 1.96-fachen Standardfehler hinzu- bzw. abzählt, also für unser Beispiel:

$$Pw \pm 1.96*\sigma = 0.5 \pm 1.96*0.1581 = 0.5 \pm 0.3099 = 50\% \pm 30.99\%$$

Basierend auf der erhobenen Stichprobe lässt sich also feststellen:

Mit einer Wahrscheinlichkeit von 95% wird der Anteil weisser Kugeln innerhalb der Grundgesamtheit zwischen 19.0% und 81.0% betragen. Das scheint wenig beeindruckend. Allerdings verbessern sich die Werte massgeblich, wenn die Stichprobengrösse ausgeweitet wird:

Stichprobenumfang k	Standardfehler $\sigma$	$\sigma*1.96$
10	0.158	0.309
20	0.112	0.219
30	0.091	0.178
40	0.079	0.154
50	0.071	0.138
60	0.065	0.126
70	0.059	0.117
80	0.056	0.109
90	0.053	0.103
100	0.05	0.098 <sup>48</sup>

Raum für Zweifel bleibt indessen allemal, denn es ist eine Frage der Willkür, welche «Sicherheit» vorausgesetzt werden soll.<sup>49</sup> Der zum Standard gewordene Wert von 95% ist jedenfalls nicht a priori angemessen. Je nach Bedeutung der Fragestellung bzw. je nach Bedeutung der an die Zahlen geknüpften Konsequenzen mag man sich damit zufrieden geben – oder eben nicht.

### XIII. Schlussbemerkungen

Grund zum Zweifel besteht immer. Das gilt nicht nur für die Ergebnisse kriminologischer Forschung, sondern auch für jene der Geschichtswissenschaft, der Physik oder der Chemie.<sup>50</sup> Der hartgesottene Skeptiker mag sogar behaupten, noch niemand habe ihm die Existenz der Aussenwelt schlüssig beweisen können. Solches Philosophieren mag ab und an ganz lustig sein, es eignet sich vielleicht sogar für eine kleine Causerie<sup>51</sup> – gewonnen ist damit nur wenig.

Letztlich ist die Frage stets dieselbe: Was wollen wir als ausreichenden Beweis für eine bestimmte Behauptung betrachten? Hier lassen sich mit gutem Gewissen unterschiedliche Positionen einnehmen.<sup>52</sup> Die öffentliche Debatte über die Ergebnisse kriminologischer Erhebungen (und über «Statistiken») ganz allgemein wird dadurch freilich nicht einfacher, denn der Umgang mit Wahrscheinlichkeiten (und also mit Unsicherheiten) ist des Men-

45 Schnell/Hill/Esser, Sozialforschung, 260 ff.

46 Eine analoge Vorgehensweise wird auch für Durchschnittswerte empfohlen, vgl. dazu die bereits in Fussnote 44 zitierten Werke.

47 Diese Stichprobe ist unter methodischen Gesichtspunkten zu klein.  $k$  wird hier absichtlich zu niedrig angesetzt, um den Einfluss der Stichprobengrösse auf die Verlässlichkeit der Ergebnisse zu zeigen (vgl. die nachfolgende Tabelle).

48 Prima vista würde man hier einen Wert von 0 erwarten. Die Abweichung rührt daher, dass die verwendete Formel kein Wissen um die Grösse der Grundgesamtheit voraussetzt. Die Formel «weiss» also nicht, dass die 100 gezogenen Einheiten bereits die ganze Grundgesamtheit ausmachen. Grundsätzlich liesse sich die Länge des Konfidenzintervalls auch exakt berechnen (und hier entsprechend 0 erhalten). Diese Vorgehensweise wäre aber immer dann ausgeschlossen, wenn die Grösse der Grundgesamtheit nicht bekannt ist.

49 Ludwig-Mayerhofer, 6.

50 Anders ist die Ausgangslage höchstens bei der (reinen) Mathematik: Interpretiert man mathematische Gleichungen als Tautologien, sind sie per definitionem logisch wahr.

51 Wittgenstein, Über Gewissheit, Nr. 467: «Ich sitze mit einem Philosophen im Garten; er sagt zu wiederholten Malen «Ich weiss, daß das ein Baum ist», wobei er auf einen Baum in unsrer Nähe zeigt. Ein Dritter kommt daher und hört das, und ich sage ihm: «Dieser Mensch ist nicht verrückt: Wir philosophieren nur.»»

52 Inakzeptabel scheint indes das willkürliche Anlegen von Massstäben je nach Untersuchungsergebnis. Namentlich in der Politik kommt es durchaus vor, dass bestimmte Statistiken als Beweis angerufen werden, wenn sie die eigene Position unterstützen, dieselbe Statistik im Folgejahr aber als belanglos abgetan wird, weil die Zahlen nicht mehr dem entsprechen, was man gerne hören möchte. Man wird auch hier an Wittgenstein erinnern müssen (Über Gewissheit, Nr. 185): «Es käme mir lächerlich vor, die Existenz Napoleons bezweifeln zu wollen; aber wenn Einer die Existenz der Erde vor 150 Jahren bezweifelte, wäre ich vielleicht eher bereit aufzuhorchen, denn nun bezweifelt er unser ganzes System der Evidenz. Es kommt mir nicht vor, als sei dies System sicherer als eine Sicherheit in ihm.»



schen Sache nicht:<sup>53</sup> Ungenaue oder mit Unsicherheiten behaftete Informationen sind ihm allemal lieber als gar keine.

## Epilog: Des Königs Herold

Es war einmal ein König, der trug seinem zuverlässigsten Herold auf, dem Bischof von Trier einen Brief zu überbringen. Schon nach wenigen Wochen traf der Herold am Bischofssitz ein, allein man beschied ihm, der Bischof habe sich nach Koblenz begeben, um daselbst im Konzilgebäude einen Handelsstreit zu schlichten. Unverzüglich machte sich der Herold auf den Weg, und bald schon war er in Koblenz angekommen. Dort allerdings irrte er – unentwegt auf seine Karte starrend – durch die Gassen, ohne das Konzilgebäude zu finden. Zufällig traf er auf einen hilfsbereiten Franziskaner: «Kann ich Ihnen helfen, werter Herr? Sie scheinen sich verirrt zu haben?» «In der Tat», meinte der Herold, «ich suche das Konzilgebäude». «Nun, ich werde es Ihnen gerne auf der Karte zeigen.» Der Mönch betrachtete die Karte und stellte leicht belustigt fest: «Aber Herr, Sie befinden sich in Koblenz, und das ist eine Stadtkarte von Trier». «Nun, das weiss ich», erwiderte der Herold, «aber eine Karte von Koblenz besitze ich nicht».

## Literatur

- Bortz J., Schuster C., *Statistik für Human- und Sozialwissenschaftler*, 7. Aufl., Berlin, 2010.
- Bult B., Schmidt E. G., Sorites, in: Ritter J., Gründer K. (Hrsg.), *Historisches Wörterbuch der Philosophie*, Bd. 9, Basel, 1996, Spalten 1090–1099.
- Chalmers A. F., *Wege der Wissenschaft*, 2. Aufl., Berlin, 1989.
- Desrosières A., *Die Politik der grossen Zahlen, Eine Geschichte der statistischen Denkweise*, Berlin, 2005.
- Diekmann A., *Empirische Sozialforschung*, 20. Aufl., Reinbek bei Hamburg, 2009.
- Føllesdal D., Walløe L., Elster J., *Rationale Argumentation*, Berlin, 1988.
- Freedman D., Pisani R., Purves R., *Statistics*, 4<sup>th</sup> ed., New York, 2007.
- Hume D., *An Enquiry concerning Human Understanding*, London, 1748, abrufbar unter: <http://www.gutenberg.org/files/9662/9662-h/9662-h.htm>.
- Göppinger H., *Kriminologie*, 6. Aufl., München, 2008.
- Göritz A. S., Moser K., Repräsentativität im Online-Panell, *der markt*, 2000, 156–162, abrufbar unter: [http://www.goeritz.net/Der\\_Markt.pdf](http://www.goeritz.net/Der_Markt.pdf).
- Hartung J., *Statistik*, 15. Aufl., München, 2009.
- Kahnemann D., *Schnelles Denken, langsames Denken*, München, 2012.
- Kaiser G., *Kriminologie*, Ein Lehrbuch, 3. Aufl., Heidelberg, 1996.
- Killias M., Auch auf die Schwere der Strafen kommt es an – ein Zwischenruf, in: Hilgendorf E., Rengier R. (Hrsg.), *Festschrift für Wolfgang Heinz zum 70. Geburtstag*, Baden-Baden, 2012, 652–662.
- Killias M., Kuhn A., Aebi M. F., *Grundriss der Kriminologie*, 2. Aufl., Bern, 2011.
- Klein S., *Alles Zufall*, 2. Aufl., Reinbek bei Hamburg, 2004.
- Krämer W., *Denkste!*, 5. Aufl., München, 2003.
- Krämer W., *So lügt man mit Statistik*, 12. Aufl., München, 2009.
- Kruskal W., Mosteller F., Representative Sampling, IV: the History of the Concept in Statistics, 1895–1939, *International Statistical Review*, 1980, 169–195.
- Kuhn A., Villettaz P., Jayet A., Willi F., Öffentliche Meinung und strenge Richter, *Crimiscope*, 2002, 1–6, abrufbar unter: [http://www.unil.ch/webdav/site/esc/shared/Crimiscope/Crimiscope019\\_2002\\_D.pdf](http://www.unil.ch/webdav/site/esc/shared/Crimiscope/Crimiscope019_2002_D.pdf)
- Kunz K.-L., *Kriminologie*, 5. Aufl., Bern, 2008.
- Kunz K.-L., *Die wissenschaftliche Zugänglichkeit von Kriminalität. Ein Beitrag zur Erkenntnistheorie der Sozialwissenschaften*, Wiesbaden, 2008.
- Kutschera F. von, *Wissenschaftstheorie I*, München, 1972.
- Losee J., *Wissenschaftstheorie*, München 1977.
- Ludwig-Mayerhofer W., *Konfidenzintervalle so einfach wie möglich erklärt*, abrufbar unter: [http://www.uni-siegen.de/phil/sozialwissenschaften/soziologie/mitarbeiter/ludwig-mayerhofer/statistik/statistik\\_downloads/konfidenzintervalle.pdf](http://www.uni-siegen.de/phil/sozialwissenschaften/soziologie/mitarbeiter/ludwig-mayerhofer/statistik/statistik_downloads/konfidenzintervalle.pdf).
- Makridakis S., Hogarth R., Gaba A., *Tanz mit dem Glück*, Berlin, 2010.
- Mill J. S., *A System of Logic, Ratiocinative and Inductive*, London, 1843, abrufbar unter: [http://books.google.ch/books?id=y4MEAAAAQAAJ&printsec=frontcover&hl=de&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](http://books.google.ch/books?id=y4MEAAAAQAAJ&printsec=frontcover&hl=de&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false).
- Mlodinow L., *Wenn Gott würfelt*, Reinbek bei Hamburg, 2009.
- Polasek W., *Schließende Statistik*, Berlin, 1997.
- Popper K. R., *Logik der Forschung*, 10. Aufl., Tübingen, 1994.
- Quatember A., Das Problem mit dem Begriff der Repräsentativität, *Allgemeines Statistisches Archiv*, 1996, 236–241.

<sup>53</sup> Entsprechend zahl- und erfolgreich sind die Werke, die für Abhilfe sorgen wollen. Dazu statt vieler: Taleb, Schwan; Taleb, Narren; Klein, Zufall; Makridakis/Hgarth/Gaba, Tanz; Mlodinow, Gott; in Teilen auch Kahnemann, Denken; für den Juristen besonders interessant Schweizer, Täuschungen, v. a. § 11.

- Russell B., *The Problems of Philosophy*, Oxford 1912. Abrufbar unter: <http://www.ditext.com/russell/russell.html>.
- Schirmer D., *Empirische Methoden der Sozialforschung*, Paderborn, 2009.
- Schnell R., Hill P. B., Esser E., *Methoden der empirischen Sozialforschung*, 5. Aufl., München, 1995.
- Schweizer M., *Kognitive Täuschungen vor Gericht*, Diss. Zürich, 2005.
- Schwind H.-D., *Kriminologie, Eine praxisorientierte Einführung mit Beispielen*, 21. Auflage, Heidelberg, 2011.
- Spickermann G., Plümer L., Wie repräsentativ sind die Stichprobenverfahren von landwirtschaftlichen Expertensystemen?, *Gesunde Pflanzen*, 203, 222–227, abrufbar unter: <http://link.springer.com/article/10.1007/s10343-003-0010-y?null>.
- Stegmüller W., *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*, Band IV: Personelle und Statistische Wahrscheinlichkeit, Zweiter Halbband: Statistisches Schliessen – Statistische Begründung – Statistische Analyse, Berlin, 1973.
- Taleb N. N., *Der schwarze Schwan*, München, 2008.
- Taleb N. N., *Narren des Zufalls*, 2. Aufl., Weinheim, 2005.
- Tiede M., Voß W., *Schließen mit Statistik – Verstehen*, München, 2000.
- Vignando R., Haas H., Die Skinhead Bewegung: Eine empirische Studie, *Crimiscope*, 2001, 1–9, abrufbar unter: [http://www.unil.ch/webdav/site/esc/shared/Crimiscope/Crimiscope015\\_2001\\_D.pdf](http://www.unil.ch/webdav/site/esc/shared/Crimiscope/Crimiscope015_2001_D.pdf).

**Christof Riedo**

Prof. Dr. iur., Fürsprecher

Lehrstuhl für Strafrecht

und Strafprozessrecht

Universität Freiburg

Av. Beauregard 11

1700 Fribourg

[christof.riedo@unifr.ch](mailto:christof.riedo@unifr.ch)