

Jeske, Roland

Working Paper

Stichprobenziehung nach dem Prinzip des "Schiffeversenkens" - Über eigentümliche Hochrechnungspraktiken des Bundesamtes für Finanzen

CoFE Discussion Paper, No. 00/25

Provided in Cooperation with:

University of Konstanz, Center of Finance and Econometrics (CoFE)

Suggested Citation: Jeske, Roland (2000) : Stichprobenziehung nach dem Prinzip des "Schiffeversenkens" - Über eigentümliche Hochrechnungspraktiken des Bundesamtes für Finanzen, CoFE Discussion Paper, No. 00/25, University of Konstanz, Center of Finance and Econometrics (CoFE), Konstanz, <https://nbn-resolving.de/urn:nbn:de:bsz:352-opus-5769>

This Version is available at:

<https://hdl.handle.net/10419/85201>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Stichprobenziehung nach dem Prinzip des "Schiffeversenkens" – Über eigentümliche Hochrechnungspraktiken des Bundesamtes für Finanzen

Roland Jeske
Universität Konstanz

Abstract:

Der Gesetzgeber läßt im Handelsgesetzbuch (HGB) zahlreiche mathematisch-statistische Stichprobenverfahren zur Inventur zu. Die dort vorgeschriebenen Richtlinien finden auch in anderen Sparten der Wirtschaftsprüfung bei statistischen Hochrechnungen Verwendung. Eigentümliche Vorstellungen existieren in der Wirtschaftsprüfung darüber, wie eine einfache Zufallsstichprobe gezogen werden sollte. Anhand eines konkreten Beispiels wird die Verzerrung des Ergebnisses bei falscher Stichprobenziehung dargelegt.

Keywords:

Einfache Zufallsstichprobe, geschichtete Zufallsstichprobe, Stichprobenplan, Verzerrung, nonresponse, Wirtschaftsprüfung

1 Einleitung

Wenige Anwendungen der Statistik unterliegen einer so reglementierten Aufsicht wie Stichprobenverfahren in der Wirtschaftsprüfung. Seit 1977 regelt §39, Abs. 2a HGB etwa die Stichprobeninventuren mit Hilfe mathematisch-statistischer Methoden. Aus §241 Abs. 1 (HGB) geht hervor, daß der Gesetzgeber anerkannte mathematisch-statistische Verfahren auf Grund von Stichproben zuläßt. Mit dieser eher unpräzisen Formulierung läßt der Gesetzgeber bewußt einen Spielraum, um auch neuere Entwicklungen im Bereich der Stichprobentheorie zuzulassen (vgl. Nies, 1976). Die

Vorgaben des HGB motivieren auch außerhalb der eigentlichen Lagerinventur Stichprobenverfahren in der Wirtschaftsprüfung (Frank/Schneeweis, 1984).

Das einfachste der nach dem Handelgesetzbuch zulässigen Verfahren stellt die einfache Zufallsstichprobe dar. Dabei wird unterschieden, ob die Elemente der Stichprobe mit oder ohne Zurücklegen gezogen werden. In beiden Fällen definiert sich eine einfache Zufallsstichprobe darüber, daß alle Stichproben vom Umfang n aus einer Grundgesamtheit vom Umfang N gleiche Ziehungswahrscheinlichkeit besitzen (siehe etwa Hansen/Hurwitz/Madow, 1966). Beim Ziehen mit Zurücklegen ist diese Forderung erfüllt, wenn sämtliche Untersuchungseinheiten gleiche Auswahlwahrscheinlichkeit besitzen.

In der Praxis erfolgt das Ziehen in der Regel über Zufallszahlen. Dabei wird auf eine vorhandene Liste mit sämtlichen Untersuchungseinheiten der Grundgesamtheit zurückgegriffen oder eine solche eigens für die Stichprobenziehung erstellt. Darin sind die N Untersuchungseinheiten eindeutig über ihre Position festgelegt. Mit Hilfe eines Zufallszahlengenerators, der gleichverteilte Zufallszahlen über $\{1, \dots, N\}$ generiert, werden nacheinander n Untersuchungseinheiten gezogen. Sollen Doppeltziehungen ausgeschlossen werden, so unterstellt man in der Praxis häufig, daß bei Eliminierung mehrfach gezogener Untersuchungseinheiten eine Zufallsstichprobe mit Zurücklegen verwirklicht wird (siehe etwa Schaich et. al. 1975, S.11).

2 Ziehen der Stichprobe mittels mehrdimensionaler Zufallsvektoren

In der Wirtschaftsprüfung gibt es Vorschläge, die diese Vorgehensweise der Stichprobenziehung vereinfachen möchten. Es steht ganz außer Frage, daß eine Durchnummerierung der einzelnen Elemente der Grundgesamtheit oder gar die Neuerstellung einer Liste sämtlicher Untersuchungseinheiten der Grundgesamtheit einen nicht unerheblichen Arbeitsaufwand erzeugt. Die Praxis sieht vielfach so aus, daß mehrere Elemente der Grundgesamtheit auf Belegen wie Rechnungen, Lieferscheinen etc. zusammengefaßt sind, so daß etwa Roberts (1978) vorschlägt, statt einer einstelligen Zufallszahl mehrdimensionale Zufallszahlen zu erzeugen, wobei eine Zufallszahl die Belegnummer, eine weitere die Position (in der Regel die Zeile) des zu erhebenden Elementes auf dem Beleg angibt. Ein solches Vorgehen ist auch bei Newman (1976) beschrieben. Roberts (1978) fügt ergänzend hinzu, daß bei dieser Vorgehensweise dafür Sorge zu tragen ist, daß jeder Untersuchungseinheit die gleiche Ziehungswahr-

scheinlichkeit zukommt.¹

Sieht man einmal davon ab, daß diese Voraussetzung nur für die Zufallsauswahl mit Zurücklegen gültig ist, liegt dieser Annahme trotz allem ein Trugschluß zugrunde: Man betrachte den Fall einer einfachen Zufallsstichprobe mit Zurücklegen. Dann ist die Grundgesamtheit gegeben durch $\{X_1, X_2, \dots, X_N\}$. Aus dieser Grundgesamtheit wird eine Stichprobe vom Umfang n gezogen, die mit $\{x_1, x_2, \dots, x_n\}$ bezeichnet sei. Beim Ziehen mit Zurücklegen wird nach obiger Definition einer einfachen Zufallsstichprobe angenommen, daß die Stichprobenelemente x_1, x_2, \dots, x_n unabhängig identisch verteilt sind.

Geht man nun davon aus, daß die Untersuchungseinheiten der Grundgesamtheit auf K Belege verteilt sind, wobei der i -te Beleg über N_i Untersuchungseinheiten verfügen möge ($i = 1, \dots, K$) mit $\sum_{i=1}^K N_i = N$, so würde das Vorgehen Roberts (1978) darin bestehen, zweidimensionale Zufallszahlen zu generieren, wobei die erste Komponente über $\{1, \dots, K\}$ gleichverteilt wäre und unabhängig davon eine zweite Komponente einer Gleichverteilung über $\{1, \dots, \max_{i \in \{1, \dots, K\}} \{N_i\}\}$ entstammt.

Auf den ersten Blick erzeugt dieses Vorgehen nur das Problem, daß Paare von Zufallszahlen generiert würden, die zu zahlreichem "Nonresponse" führen, da auf dem Beleg i möglicherweise eine Positionsnummer größer als N_i gezogen wird und damit die Stichprobengröße n zufällig wird.

In der Praxis taucht jedoch noch ein anderes, viel gravierenderes Problem auf: Es steht außer Frage, daß bei unterschiedlichen Beleggrößen mehr Untersuchungseinheiten vom Beginn der Belege in die Stichprobe gelangen als Untersuchungseinheiten, die eher am Ende der Belege stehen (siehe Abbildung 1 unter Abschnitt 3). Damit dieses Verfahren zulässig ist, bedarf es daher der Annahme, daß die einzelnen Untersuchungseinheiten eines jeden Beleges unabhängig über den Beleg verteilt sind. Diese Unabhängigkeitsannahme dürfte in der Praxis allerdings so gut wie nie erfüllt sein: Rechnungen haben fast immer eine Struktur, indem beispielsweise die einzelnen Positionen einer Rechnung bezüglich der Artikelbezeichnung oder der Artikelnummer lexikographisch geordnet sind. Mit letzterem geht in der Regel eine Anordnung verschiedener Warengruppen einher. Im schlimmsten Fall könnte man sich Rechnungen vorstellen, in denen die Einzelpositionen der Größe nach geordnet vorliegen. Bei derartigen Sortierreihenfolgen ist es somit wahrscheinlicher, daß eine Abhängigkeit zur erhebenden Größe besteht als daß kein Einfluß besteht. In den oben genannten Fällen trifft man mit dem Verfahren von Roberts (1978) systematisch auf zuviele Untersuchungseinheiten, die mit Buchstaben aus dem Anfang des Alphabets begin-

¹"In deciding on the correspondence scheme, two important requirements are to [...] make sure the scheme gives each sampling unit an equal chance of being selected [...] (S. 20)

nen, niedrige Artikelnummern haben bzw. allgemein den Positionen am Beleganfang entspringen. Im Fall, daß die Untersuchungseinheiten der Größe nach sortiert sind, trifft man systematisch auf zu kleine oder zu große Merkmalsausprägungen, je nach Sortierreihenfolge.

Um ein weiteres Beispiel zu nennen, sei auf Rechnungen hingewiesen, die sowohl Warenwerte wie auch Dienstleistungen umfassen. In der Kfz-Branche ist es beispielsweise üblich, bei Reparaturen zunächst die ausgetauschten Ersatzteile und anschließend die geleisteten Arbeitsstunden zu fakturieren. Würde man den Vorschlag von Roberts (1978) auf diesen Fall anwenden, so würden die Warenwerte überschätzt und die Dienstleistungen unterschätzt.

Um diesen Effekt zu vermeiden, müßten die Untersuchungseinheiten der einzelnen Belege vor der Erhebung zufällig permutiert werden, was einen Arbeitsaufwand erfordert, der der üblichen Vorgehensweise, eine Liste mit sämtlichen Untersuchungseinheiten zu erstellen, in keinsten Weise nachstehen dürfte und zusätzlich das Problem eines zufälligen Stichprobenumfangs enthält, so daß dieses Verfahren kaum eine ernsthafte Alternative zur üblichen Ziehungstechnik einer einfachen Zufallsstichprobe darstellen dürfte.

3 Ein Beispiel aus der Praxis

Eine Steigerung dieses Verfahrens findet sich in einem Gutachten des Bundesamtes für Finanzen (BfF) aus dem Jahr 1995. Dort wurden im Zuge der Ermittlung von Verkaufserlösen einer Apotheke in Süddeutschland im Wirtschaftsjahr 1991/92 Belege in Form von Rechnungen und Gutschriften sowie Inventurlisten zu Beginn und Ende des untersuchten Wirtschaftsjahres herangezogen. Es wurde vereinfachend angenommen, daß die Ware im Inventuranfangsbestand im Laufe des Wirtschaftsjahres verkauft wurde, ebenso die über Rechnungen gelieferte Ware. Für den Wareneinsatz wurden Gutschriften und der Inventurendbestand zum Ende des Wirtschaftsjahres in Abzug gebracht, so daß Untersuchungseinheiten der letzten beiden Belegarten mit negativem Vorzeichen in die Untersuchung eingingen.

Rechnungen wie auch Gutschriften umfassen stark variierende Anzahlen an einzelnen Buchungspositionen. Abhängig davon, ob Bestellungen bei einem Großhändler oder direkt beim Hersteller erfolgten, setzten sich Rechnungsbelege aus mehr oder weniger Einzelpositionen zusammen.

Die Gutschriften wiederum wurden teilweise einzeln erstattet, zum Teil in Form einer Monatsabrechnung gesammelt und dann erstattet, so daß auch hier eine stark variierende Anzahl von Einzelpositionen auf den Belegen zu finden ist. Insgesamt

beinhalten die Gutschriften durchschnittlich jedoch deutlich weniger Elemente pro Beleg als etwa die Rechnungen (vgl. Tabelle 1).

Die Inventurbücher bestehen aus Vordrucken, die sämtliche über Apotheken zu vertreibende Präparate gelistet haben, wobei darauf hinzuweisen ist, daß der weitaus größte Teil aller dort gelisteten Präparate nicht zum ständigen Sortiment gehört (andernfalls wäre der pharmazeutische Großhandel überflüssig), viele Artikel wurden durch die betreffende Apotheke gar nicht vertrieben, da im Einzugskreis kein Bedarf für diese Arzneimittel bestand.

Tabelle 1: Maßzahlen der Grundgesamtheit

Belegart	Anzahl der Belege	Anzahl Einzelpos.	durchschn. Anzahl an Einzelpos./Beleg.
Inventuranfangsbestand	442	5.607	12,69
Rechnungen	6.665	70.541	10,58
Inventurendbestand	454	5.930	13,06
Gutschriften	807	1.690	2,09

Das Bundesamt für Finanzen hat nun die jeweiligen Belege paginiert und für jede Belegart die maximale Anzahl von Einzelpositionen bestimmt, um anschließend mit Hilfe des Verfahrens von Roberts (1978) durch Vektoren von gleichverteilten Zufallszahlen eine Stichprobenziehung durchzuführen.

Bei den Inventurlisten wurden die maximal vorkommenden Präparate pro Seite (132 Stück) als maximale Position verwendet, wenngleich wie oben erwähnt die betreffende Apotheke nur eine geringe Anzahl aller über eine Apotheke vertreibbaren Artikel im ständigen Sortiment führte.

Diese Vorgehensweise führte dazu, daß das BfF auf eine Vielzahl von nicht anzutreffenden Positionen in den Belegen stieß. Jeder Nonresponse wurde gemäß dem Spiel "Schiffeversenken" als sog. "Wassertreffer" bezeichnet und konnte nicht in die Stichprobe aufgenommen werden. Um eine Stichprobe vom Umfang $n = 782$ zu erhalten, waren insgesamt 4000 Zufallszahlen notwendig. Somit führten nicht einmal 20% der Zufallszahlen zu einem Element der Grundgesamtheit. Die zahlreichen Befürworter von Quotenverfahren aus der Meinungsforschung (siehe etwa Noelle-Neumann/Petersen, 1996) könnten einen derartigen Nonresponse-Anteil zum Anlaß nehmen, die Auseinandersetzung um das Für und Wider von Quotenverfahren und Zufallsstichproben in der Meinungsforschung auch auf Warenstichproben auszudehnen.

Da das Verfahren von Roberts (1978) separat für die unterschiedlichen Belegarten

verwendet wurde, entstanden für die einzelnen Schichten unterschiedliche Trefferwahrscheinlichkeiten (vgl. Tabelle 2), die dazu führten, daß die verschiedenen Belegarten nicht mit den Gewichten der Grundgesamtheit in die Stichprobe eingingen (vgl. Tabelle 3).

Tabelle 2: Trefferwahrscheinlichkeiten für die einzelnen Belegarten

Belegart	durchschn. Anzahl Unters.-Einh./Beleg	mögliche Zufalls- ausgänge/Beleg	Trefferwahr- scheinlichkeit
Inventuranfangsbestand	12,69	132	0,096
Rechnungen	10,58	44	0,240
Inventurendbestand	13,06	132	0,099
Gutschriften	2,09	34	0,061

Tabelle 3: Anteile der einzelnen Belegarten in Grundgesamtheit und Stichprobe

Belegart	Anteil der Untersuchungseinh. (in %)	
	der Grundgesamtheit	der Stichprobe
Inventuranfangsbestand	6,69	2,56
Rechnungen	84,21	91,94
Wareneingang gesamt	90,90	94,50
Inventurendbestand	7,08	1,79
Gutschriften	2,02	3,71
Warenausgang gesamt	9,10	5,50

Ungeachtet dieser Tatsache hat das BfF die gezogene Stichprobe aber als ungeschichtete Stichprobe analysiert mit der Folge, daß der durchschnittliche Einkaufspreis in der Stichprobe den wahren durchschnittlichen Einkaufspreis von 26,74 um 10,96% überschätzt (vgl. Tabelle 4).

Prinzipiell wäre es zwar möglich, eine einfache Zufallsstichprobe nachträglich zu schichten (siehe etwa Stenger, 1986). Schwierig dürfte es allerdings sein, die teilweise extremen Verzerrungen der Mittelwerte, die einerseits auf die geringen Stichprobenumfänge in den Schichten, andererseits auf die systematische Verzerrung durch das Stichprobenverfahren zurückzuführen sind, zu eliminieren, so daß eine erneute Stichprobenziehung unumgänglich erscheint.

Der "Roberts"-Effekt, daß die einzelnen Untersuchungseinheiten auf den Belegen einer Sortierreihenfolge unterliegen, kommt hier auf verschiedene Art zum Tragen: Abbildung 1 zeigt die Zeilenpositionen der gezogenen Stichprobenelemente bei

den Rechnungen. Diese sind zumindest bezogen auf den weitaus größten Teil der Großhändler lexikographisch nach den Namen der pharmazeutischen Produkte geordnet, ebenso die Inventurlisten. Die Gutschriften sind nur teilweise lexikographisch geordnet, zum Teil aber auch willkürlich angeordnet. Bei Sammelgutschriften liegt die Sortierung zum Teil nach Eingangsdaten, teilweise nach Hersteller getrennt vor.

Tabelle 4: Vergleich der mittleren Einkaufspreise unterteilt nach Belegarten

Belegart	mittlerer Einkaufspreis (in DM)	
	der Grundgesamtheit	der Stichprobe
Inventuranfangsbestand	29,20	19,44
Rechnungen	32,50	33,10
Wareneingang gesamt	32,26	32,73
Inventurendbestand	-27,72	-21,04
Gutschriften	-30,87	-26,43
Warenausgang gesamt	-28,42	-24,68
Gesamt	26,74	29,67

Abbildung 1: Verteilung der Zeilenpositionen in der Stichprobe (nur Rechnungen)

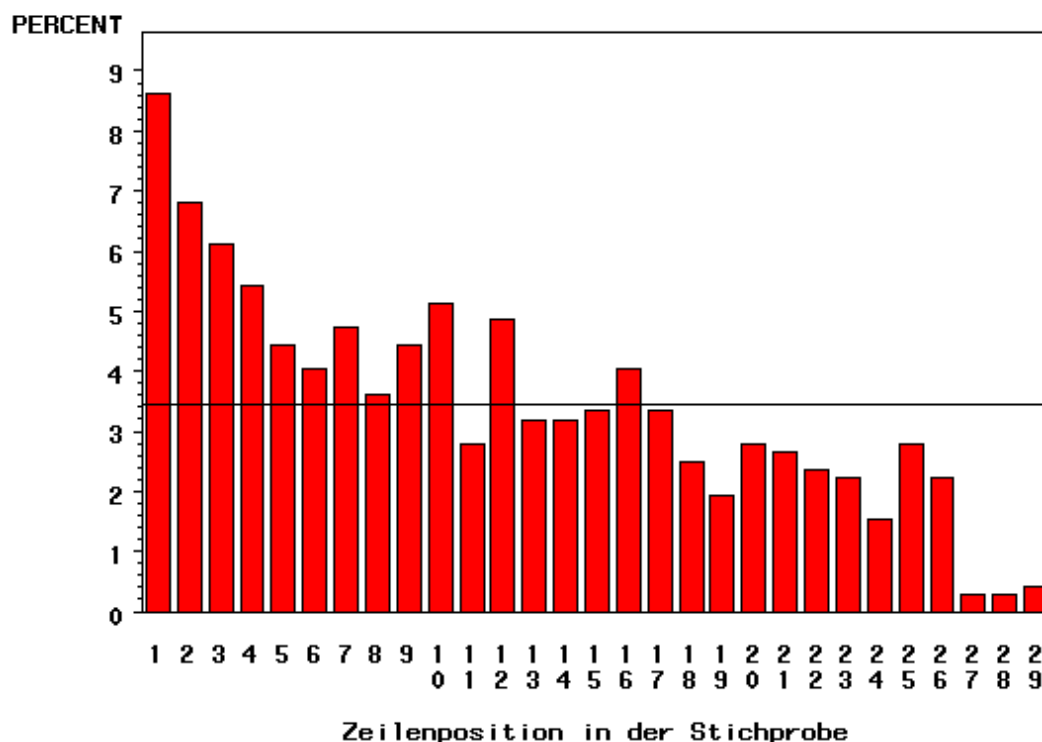
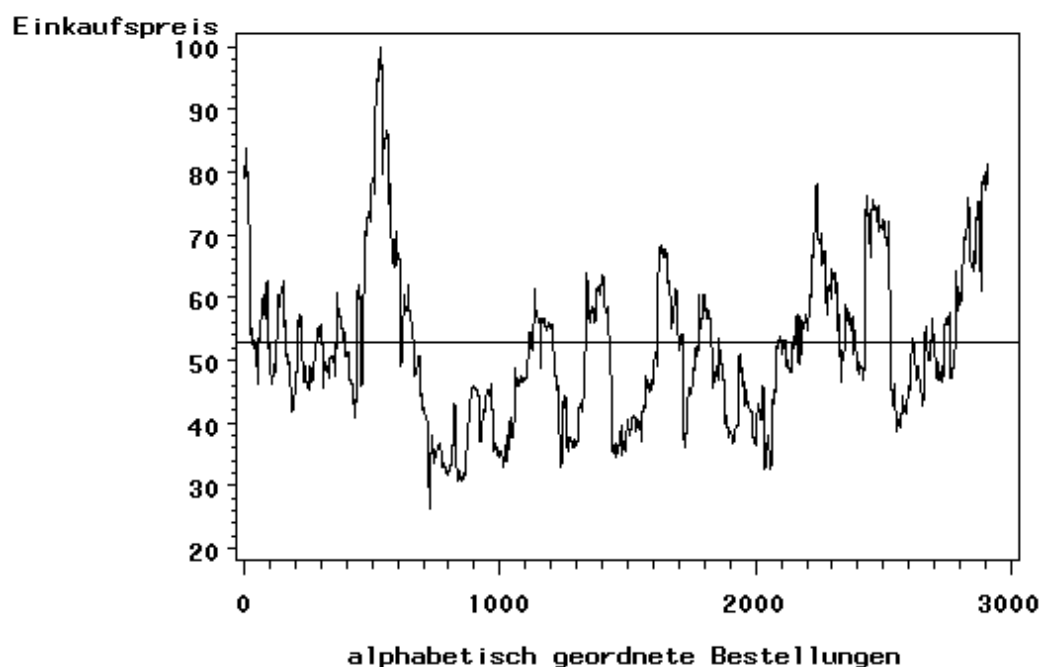


Abbildung 2 stellt einen einfachen gleitenden 99er-Durchschnitt über die nach Präparaten alphabetisch sortierten Einkaufspreise aller im Monat Juni 2000 beim meist frequentierten Großhändler bezogenen Präparate der betrachteten Apotheke dar. Wenngleich dieser gleitende Durchschnitt mit großer Bandbreite nach wie vor starke Schwankungen offenbart, so sind deutliche Buchstabeneffekte erkennbar. Insbesondere die mit größerer Wahrscheinlichkeit gezogenen Positionen am Anfang liegen eher über der Referenzlinie des Gesamtmittelwertes.

Abbildung 2: Schwankungen des Einkaufspreises in Abhängigkeit des Präparatenamens (einfacher gleitender 99er-Durchschnitt)



4 Zusammenfassung:

Eine einfache Zufallsstichprobe sollte stets mit Hilfe eindimensionaler Zufallszahlen aus einer Liste gezogen werden. Das von Roberts (1978) vorgeschlagene Verfahren einer Stichprobenziehung mit mehrdimensionalen Zufallszahlen führt zu erheblichen Verzerrungen, wenn auf den Belegseiten Abhängigkeiten zum erhebenden Merkmal bestehen.

In Verbindung mit einer nicht als solchen erkannten Schichtenbildung wird der

Wareneinsatz in einer Stichprobe des Bundesamtes für Finanzen um etwa 11% überschätzt.

Für die Wirtschaftsprüfung allgemein wären verbindliche Normen zur Erstellung von Stichprobenplänen wünschenswert. Darin sollte dem Anwender auch die für einen Statistiker trivial anmutende Frage, wie eine einfache Zufallstichprobe zu erheben ist, detailliert erläutert werden.

Literatur:

Bundesamt für Finanzen (1995): Behördengutachten über ein mathematisch-statistisches Verfahren zur Ermittlung der Verkaufserlöse im Wirtschaftsjahr 1991/1992 einer süddeutschen Apotheke,²

Frank, G und L. Schneeweis (1984): Stichproben im Bereich von Revision und Wirtschaftsprüfung, Der Betrieb, S. 2629-2631

Hansen, M.H., W.N. Hurwitz, W.G. Madow (1966): Sample Survey Methods and Theory Volume 1, John Wiley & Sons, Inc., New York

HGB (1999): Gebundene Ausgabe, C.H.Beck

Newman, M.S. (1976): Financial Accounting Estimates through Statistical Sampling by Computer, Wiley, New York

Nies (1976): Mathematisch-statistische Methoden als Hilfsmittel zur Aufstellung des Inventars, Die steuerliche Betriebsprüfung 11/1976, S. 245-248.

Noelle-Neuman, E. und Th. Petersen (1996): Alle, nicht jeder, München

Roberts, D.M. (1978): Statistical Auditing, New York

Schaich, E., D. Köhle, W. Schweitzer, F. Wagner (1975): Statistik II für Volkswirte, Betriebswirte und Soziologen, Vahlen, München.

Stenger, H. (1986): Stichproben, Physica-Verlag Heidelberg, Wien

²aus Datenschutzgründen wurden Name der Apotheke und Inhabernamen durch die Bezeichnung "einer süddeutschen Apotheke" ersetzt