

Schriften des Europäischen
Instituts für Sozioökonomie e.V. Band 8

Cheater-Detection bei der Randomized Response-Technik

Herleitung, Analyse und Anwendung

Sascha Feth

Monika Frenger

Werner Pitsch

Patrick Schmelzeisen



universaar

Universitätsverlag des Saarlandes
Saarland University Press
Presses Universitaires de la Sarre

Schriften des Europäischen
Instituts für Sozioökonomie e.V.
Band 8

Sascha Feth, Monika Frenger, Werner Pitsch und
Patrick Schmelzeisen

Cheater-Detection bei der Randomized Response-Technik

Herleitung, Analyse und Anwendung



universaar

Universitätsverlag des Saarlandes
Saarland University Press
Presses Universitaires de la Sarre

© 2014 *universaar*
Universitätsverlag des Saarlandes
Saarland University Press
Presses Universitaires de la Sarre



Postfach 151150, 66041 Saarbrücken

ISBN gedruckt: 978-3-86223-165-2

ISBN online: 978-3-86223-166-9

URN mit Prüfziffer: urn:nbn:de:bsz:291-universaar-1343

Projektbetreuung *universaar*: Susanne Alt, Matthias Müller

Satz: Patrick Schmelzeisen

Umschlaggestaltung: Julian Wichert

Gedruckt auf säurefreiem Papier von Mosenstein & Vannerdat

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Inhaltsverzeichnis

1	Einleitung	11
1.1	Theoretischer Hintergrund und Entwicklung der Randomized Response-Technik	13
1.2	Das Unrelated Question-Modell	14
1.3	Das Forced Answer-Modell	15
1.4	Das Cheater Detection-Modell	17
2	Modellierung und Notation	21
2.1	Generelle Notation	21
2.2	Population	21
2.3	Randomized Response-Technik	22
2.4	Übersicht der Verfahren	23
3	Herleitung der Verfahren	25
3.1	Herleitung der Teilschätzer	25
3.2	Herleitung der TCD	30
3.3	Herleitung der NCD	35
3.4	Herleitung der YCD	37
3.5	Praktische Anwendung	39
4	Analyse der Schätzer	41
4.1	Berechnung der Varianzen der Schätzer	41
4.2	Erwartungswerte der Schätzer	49
4.3	Optimierung der Forced Answer-Wahrscheinlichkeiten	53
4.4	Darstellung der simulierten Daten	55
4.5	Analyse der TCD	56
4.6	Analyse der NCD	62
4.7	Analyse der YCD	68

5	Vergleich der Verfahren	73
5.1	Notation	73
5.2	Zusammenhang TCD mit DCD	74
5.3	Vergleich zum Verfahren von Clark und Desharnais	75
6	Praktische Anwendungen	77
6.1	Auswirkungen unterschiedlicher Umfänge in den RRT-Gruppen	77
6.2	Dynamische Steuerung einer Umfrage	80
6.3	Unterschiedliches Antwortverhalten in den einzelnen RRT-Gruppen	84
7	Ausblick	87
A	Herleitung der Dreiecksdarstellung	89
B	Abbildungen und Tabellen	93
C	Literaturverzeichnis	109

Zum Geleit

Quantitative Methoden zur Analyse des Antwortverhaltens im Falle sensibler Fragen in Umfragestudien sind für die sozialwissenschaftliche Forschung von großer Bedeutung. Sie erlauben Einblicke in Forschungsfelder, welche auf anderen Wegen nur schwer zugänglich sein dürften. Mit der Randomized Response-Technik (RRT) steht für derartige Analysen ein langjährig erprobtes und in verschiedenen Varianten bewährtes Instrument zur Verfügung. Angewendet wurde das Instrument in Umfragestudien z. B. zum Suchtverhalten, zu Täuschungsversuchen im Studium, zum Umfang der Schattenwirtschaft, zum Einsatz des Dopings im Sport oder zum Wahlverhalten der Bevölkerung. Allen diesen Forschungsfeldern gemeinsam ist, dass eine „wahre“ Antwort für die Antwortenden peinlich, unangenehm und bisweilen sogar strafrechtlich sanktionsbeladen ist. Die Antworten müssen daher als „sensitiv“ eingestuft und daher die einer Einzelperson zuordenbaren Daten robust „verschlüsselt“ werden. Interessant ist aus sozialwissenschaftlicher Sicht ja gerade nicht die Verfolgung der Handlungen einer Einzelperson, sondern vielmehr die Quantifizierung des Anteils der Süchtigen, Täuscher und Betrüger, Schwarzarbeiter und Doper an einer Gesamtpopulation. Genau dies leistet die RRT. Der vorliegende Band zeichnet die Entwicklung der unterschiedlichen RRT-Verfahren nach, gibt einen Überblick über ihre statistischen Eigenschaften, widmet sich insbesondere Varianten der so genannten „Cheater Detection“ (also jener Gruppe von Antwortenden, die sich trotz zugesicherter Verschlüsselung nicht an die Instruktionen der RRT halten) und illustriert mittels numerischer Simulationen das Verhalten der RRT in praktischen Anwendungen. Damit wird dem interessierten Leser ein kompakter und doch zugleich fundierter Überblick über die bei Anwendungen der RRT zu beachtenden Aspekte und Probleme geboten. Dies sollte zu einem besseren Design von Umfragestudien und damit zur Ableitung belastbarer empirischer Befunde führen. Im Rahmen von Forschungsprojekten des Europäischen Instituts für Sozioökonomie (EIS) wurde die RRT in der Vergangenheit wiederholt eingesetzt. Der vorgelegte Band fasst die Erfahrungen von EIS-Forschern im Umgang mit der RRT zusammen und macht sie einem

breiten Leserkreis zugänglich. Die Schriftenreihe des EIS ist als Open Access Publikation verfügbar, so dass interessierte Leser erst gar nicht in Versuchung gebracht werden, eine Raubkopie anzulegen oder sich auf anderen dunklen Kanälen eine Kopie des Bandes zu verschaffen. Insofern können Leser des Bandes gelassen etwaigen zukünftigen mittels der RRT durchgeführten Umfragen des EIS unter den Lesern unserer Publikationsreihe entgegensehen.

Christian Pierdzioch
Eike Emrich

Vorwort der Autoren

Das vorliegende Buch bewegt sich an der Grenze zwischen sozialwissenschaftlicher Methodologie und mathematisch-statistischer Methodenentwicklung. Aus der Untersuchung abweichenden Verhaltens im Sport entstand die Notwendigkeit, vorhandene statistische Modelle zur Analyse abweichenden Verhaltens zu erweitern. Dieser Doppelcharakter ist dem Buch auch deutlich anzumerken. Beginnend mit praktischen Fragen der Soziologie der Devianz wendet es sich den mathematischen Grundlagen der neu entwickelten Methode zu, um dann über die Analyse und numerische Simulation der Grenzen des Einsatzes der Methode zu eher praktischen Fragen ihrer Anwendung zurückzukehren. Wir hoffen damit, sowohl den an der einzelnen Disziplin interessierten Leser zufrieden stellen zu können, als auch ein Beispiel für eine gelungene Zusammenarbeit über Grenzen der Disziplinen hinweg vorzulegen.

Sascha Feth
Monika Frenger
Werner Pitsch
Patrick Schmelzeisen

1 Einleitung

Will man herausfinden, wie häufig eine peinliche Eigenschaft in einer Population auftritt, gehen damit etliche Probleme einher, denen man mit bestimmten Befragungsmodellen entgegenwirken oder sie in der Rechnung kompensieren kann. Die Idee, welcher der Entwicklung indirekter Befragungsmethoden zugrunde liegt, ist die in der Forschung gut bekannte Tatsache, dass eine signifikante Verzerrung durch sozial erwünschtes Antworten im Falle, dass nach peinlichen Themen gefragt wird, besteht (vgl. hierzu Lee, 1993; Holbrook & Krosnik, 2010; 2012). Um dieser Verzerrung durch eine hohe Rate von Leugnern (Rasinski, Willis, Baldwin, Yeh & Lee, 1999) entgegenzuwirken und sich dem tatsächlichen Wert anzunähern, werden unterschiedliche Techniken angewandt wie bspw. unaufdringliche Forschung, die frei ist von Rückwirkungen auf den Befragten (s. Lee, 2000), soziale Erwünschtheitskalen (z. B. Stoeber, 2001; Thompson & Phua, 2005) oder Techniken, die darauf ausgelegt sind, das Vertrauen der Antwortenden soweit zu steigern, dass diese selbst sozial unerwünschtes Verhalten zugeben (z. B. mit der Unmatched Count-Technik oder der Item Count-Technik, siehe hierzu Ahart & Sackett, 2004; Coutts & Jann, 2011; Tourangeau & Yan, 2007; Randomized Count-Technik, siehe Frenger, Pitsch & Emrich, 2013 oder Single Sample Count-Technik, siehe James, Nepusz, Naughton & Petróczi, 2013). Unter den zuletzt genannten Techniken war die Randomized Response-Technik (RRT) die erste. Auch derzeit ist sie nach wie vor die am weitesten verbreitete, am meisten evaluierte und am meisten erforschte Technik im Forschungsfeld sensitiver Themen.

Dass diese Methoden die Bereitschaft, eine peinliche Eigenschaft zuzugeben, wesentlich gegenüber direkten Befragungen steigern, wurde bereits vielfach nachgewiesen. In der Logik dieser Untersuchungen werden die Ergebnisse direkter Fragen mit den Ergebnissen von Randomized Response-Fragen verglichen unter der Annahme, dass die peinliche Eigenschaft aufgrund eines „social desirability bias“ in direkten Fragen häufig gelehnet wird. Damit weist also ein

höherer Schätzer der Rate von Eigenschaftsträgern in Randomized Response-Fragen darauf hin, dass es gelungen ist, diese Verzerrung zumindest partiell zu kompensieren. Solche Untersuchungen wurden zum Beispiel

- zu illegalen Abtreibungen (Abernathy, Greenberg & Horvitz, 1970; Greenberg, Kuebler, Abernathy & Horvitz, 1971; Krotki & Fox, 1974; Rider, Harper, Chow & I-Cheng, 1976; I-Cheng, Chow & Rider, 1972),
- zum Konsum legaler und illegaler Drogen (Barth & Sandler, 1976; Brewer, 1981; Akers, Massey, Clarke & Lauer, 1983; Danermark & Swenson, 1987; Goodstadt, Cook & Gruson, 1987),
- zum Betrug bei der Inanspruchnahme von Sozialversicherungsleistungen (Böckenholt & van der Heijden, 2007),
- zum Einhalten von Regeln zum Umwelt- und Artenschutz (Chaloupka, 1985),
- zu Plagiaten (Coutts, Jann, Krumpal & Näher, 2011; Jerke & Krumpal, 2013) oder anderen Formen studentischen Fehlverhaltens (Kerkvliet, 1994; Shotland & Yankowski, 1982; Scheers & Dayton, 1987),
- zum Fehlverhalten im Beruf (Reckers, Wheeler & Wong-On-Wing, 1997; Soeken & Macready, 1986), sowie
- zum Sexualverhalten (Fidler & Kleinknecht, 1977; Williams & Suen Hoi, 1994)

durchgeführt. Verschiedentlich wurden auch mehrere peinliche Eigenschaften innerhalb eines Instruments erfragt (Coutts & Jann, 2011; Edgell, Himmelfarb & Duchan, 1982; Fisher, Kupferman & Lesser, 1992). Mit wenigen Ausnahmen (Duffy & Waterton, 1988; Krotki & Fox, 1974) führte die RRT zu höheren Schätzungen der Prävalenz der peinlichen Eigenschaft(en) als die Verwendung direkter Fragen. Für eine Vielzahl sensitiver Fragen wurde damit gezeigt, dass die RRT die Antwortverzerrungen bei sozial unerwünschten Verhaltensweisen reduzieren kann. Wie Wolter (2012, 96) feststellte, ist allerdings im Bereich der RRT ein deutliches Missverhältnis zwischen einer Vielzahl methodologischer Arbeiten (s.o.) im Vergleich zu sehr wenigen Anwendungen zur Bearbeitung wissenschaftlicher Fragestellungen (zum Doping s. Pitsch & Emrich, 2012; Simon, Striegel, Aust, Dietz & Ulrich, 2006 und zu Wettbewerbsverzerrungen im Sport Pitsch, Emrich & Pierdzioch, 2013) zu verzeichnen.

Im Folgenden soll die Idee und die Entwicklung der RRT dargestellt werden, um deutlich zu machen, wie diese Methode verlässliche Antworten bei

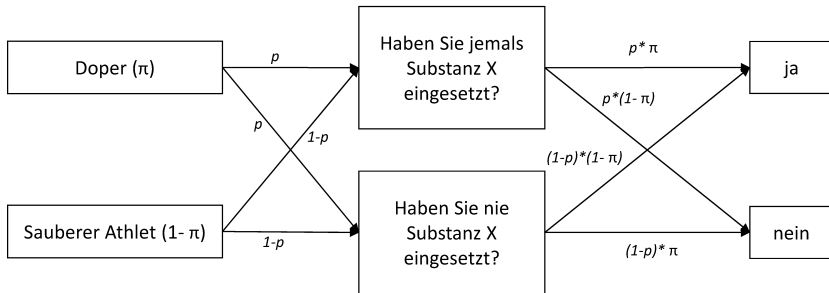


Abbildung 1.1: Warners Modell (ohne Cheater Detection) am Beispiel des Dopings bzw. der Einnahme von Dopingsubstanzen (vgl. Warner, 1965).

gleichzeitiger Diskretion erzeugen kann. Wir werden hier lediglich einen Abriss über die Entwicklung der Methode darstellen ohne bei jeder Variation auf die mathematischen Hintergründe einzugehen. Dem Leser wird an dieser Stelle empfohlen zu den jeweiligen Entwicklungen und der zugehörigen mathematischen Basis die Originalliteratur hinzuzuziehen (vgl. Lensvelt-Mulders, Hox, van der Heijden & Maas, 2005; Lee, 1993; Coutts & Yann, 2011; Antonak & Livneh, 1995; Wolter, 2012).

1.1 Theoretischer Hintergrund und Entwicklung der Randomized Response-Technik

Im ersten RRT-Modell von Warner (1965) werden zwei unterschiedliche Fragen (hinsichtlich der Zustimmung zur Aussage) gestellt, um herauszufinden, ob der Befragte zur Gruppe X gehört oder nicht dazu gehört (vgl. Abbildung 1.1 am Beispiel des Dopings).

Basierend auf den Ergebnissen eines Zufallsprozesses werden die Antwortenden mit einer bestimmten Wahrscheinlichkeit dazu aufgefordert eine Antwort auf Frage 1 oder Frage 2 zu geben. Da das Ergebnis des Zufallsprozesses für den Forscher unbekannt bleibt, kann eine „Ja“-Antwort nicht gleich gesetzt werden mit der Tatsache, dass die befragte Person tatsächlich die peinliche Eigenschaft X besitzt. Erst die Differenz in der Rate der „Ja“-Antworten kann dazu genutzt werden, um die Prävalenz π_X der Eigenschaft X in der untersuchten Population zu schätzen. Warner (1965) konnte zeigen, dass der Schätzer π_X ein unverzerrter und normalverteilter Schätzer für den wahren Anteil von X in der Population ist.

Seit Warner (1965) die Methode eingeführt hat, gab es zahlreiche Verbesserungen und es tauchten neue Entwicklungen auf. Drei Modelle, die sich später in vielen Bereichen sozial abweichenden Verhaltens als effektiv erwiesen haben, waren

- das Unrelated Question-Modell
- das Forced Answer-Modell und
- das Cheater Detection-Modell.

Diese drei Modelle der RRT sollen im Folgenden kurz dargestellt werden.

1.2 Das Unrelated Question-Modell

Die Tatsache, dass in Warners Modell alle Befragten die peinliche Frage beantworten sollen (positiv oder negativ formuliert), führt zu einem begründeten Verdacht, dass Antwortende eher bereit sind ehrlich zu antworten, wenn es zwei unterschiedliche Fragen gibt. Hierbei sollte eine Frage auf die peinliche Eigenschaft ausgerichtet sein und die zweite völlig unabhängig davon eine nicht peinliche Eigenschaft betreffen (Horvitz, Shah & Simmons, 1967). Hierdurch kann in Verbindung mit dem Randomisierungseffekt weder eine „Ja“- noch eine „Nein“-Antwort eindeutig damit in Verbindung stehen, dass der Antwortende die peinliche Eigenschaft hat oder dass er/sie diese nicht hat. Der theoretische und mathematische Hintergrund zu diesem Modell geht auf Greenberg, Abul-Ela, Simmons & Horvitz (1969) zurück. Wenn dieses Modell zum Einsatz kommt, werden den Befragten zwei unterschiedliche Fragen gestellt (vgl. Abbildung 1.2). Er/Sie wird dann aufgefordert, einen Zufallsgenerator (z. B. eine Münze oder einen Würfel) einzusetzen, ohne hierbei dem Versuchsleiter das Ergebnis des Zufallsprozesses mitzuteilen. In Abhängigkeit vom Ergebnis wird der Befragte mit einer Wahrscheinlichkeit von p aufgefordert auf die erste Frage zu antworten oder mit einer Wahrscheinlichkeit von $(1 - p)$ auf die zweite Fragen zu antworten.

Hierbei stellt X die peinliche Eigenschaft dar und Y die unverbundene, nicht peinliche Eigenschaft (alternativ könnte diese Frage auch lauten: „Waren Sie schon einmal in New York?“ oder „Sind Sie in North Carolina geboren?“). Greenberg et al. (1969) haben gezeigt, dass das Unrelated Question-Modell zu besseren Ergebnissen führt als Warners ursprüngliches Modell. Dies gilt, sofern die (unbekannte) Prävalenz π_X von X deutlich von 0,5 verschieden ist, einige Beschränkungen für p beachtet werden und die unverbundene Eigenschaft Y so gewählt wird, dass π_X und π_Y beide entweder größer oder kleiner 0,5

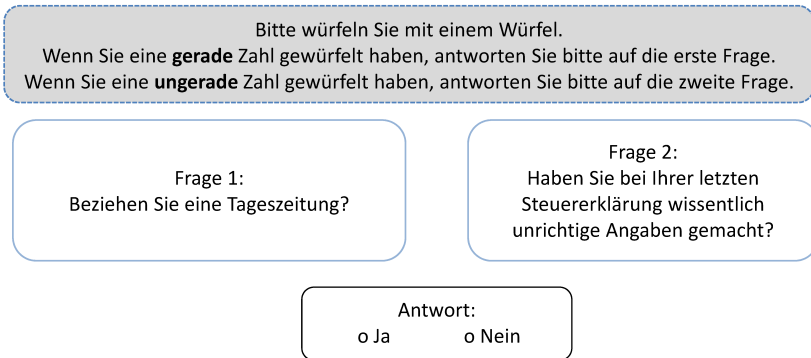


Abbildung 1.2: Beispielhafter Fall der Unrelated Question (ohne Cheater Detection).

sind. Die Forscher konnten die relative Effizienz für das Unrelated Question-Modell auch im Fall, dass das Auftreten der peinlichen Eigenschaft nicht völlig wahrheitsgemäß berichtet wird, simulieren. In späteren Anwendungen wurden Unrelated Question-Modelle ebenso für bekannte Häufigkeiten der nicht peinlichen Eigenschaft Y (z. B. „Sind Sie im Monat April geboren?“ Scheers, 1992) als auch für Eigenschaften mit unbekannter Prävalenz (z. B. „Waren Sie jemals in New York?“ ; Tracy & Fox, 1981) eingesetzt. Eine schematische Darstellung dieses Modells ist in Abbildung 1.3 zu finden.

1.3 Das Forced Answer-Modell

Das Forced Answer-Modell (verschiedentlich auch als Forced Response-Modell bezeichnet) wurde erstmals von Greenberg et al. im Jahr 1969 explizit benannt, jedoch ohne es formal zu entwickeln. Fidler und Kleinknecht (1977) berichteten über die erste Anwendung, wobei sowohl die „Ja“-Antwort als auch die „Nein“-Antwort als Forced Answers vorgegeben wurden. Abweichend davon und ohne Bezug auf diese Quellen wurde später von Dawes und Moore (1980) eine Variante vorgeschlagen, in der lediglich die peinliche Antwort, meist also die Antwort „Ja“ als Forced Answer vorgegeben wird mit dem Ziel, die Varianz des Schätzers π_X zu reduzieren. Eine Anwendung dieses Modells könnte beispielsweise wie in Abbildung 1.4 aussehen.

Mit diesem Modell kann durch eine „Ja“-Antwort also nicht eindeutig auf das Vorhandensein der peinlichen Eigenschaft X geschlossen werden, während

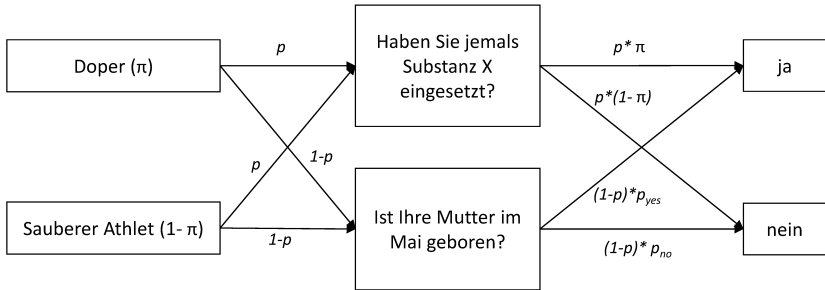


Abbildung 1.3: Unrelated Question-Modell (ohne Cheater Detection) nach Greenberg et al., 1969.

Frage:
Haben Sie in den vergangenen 12 Monaten Marihuana konsumiert?

Bitte würfeln Sie mit einem Würfel.
Wenn Sie eine **eins** oder **zwei** gewürfelt haben, sagen Sie bitte "nein".
Wenn Sie eine **andere** Zahl gewürfelt haben, antworten Sie bitte ehrlich auf die Frage.

Antwort:
o Ja o Nein

Abbildung 1.4: Forced Answer ohne Cheater Detection bei peinlicher „Nein“-Antwort.

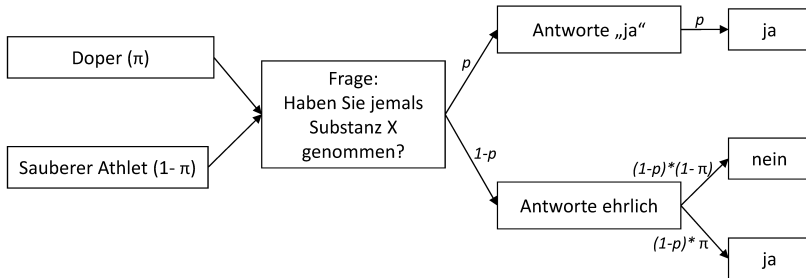


Abbildung 1.5: Forced Answer-Modell ohne Cheater Detection.

die Anonymität der Befragten völlig zugesichert werden kann. Alternativ könnte auch durch die Anweisung „... , dann sagen Sie bitte ja“, die „Nein“-Antwort geschützt werden, sofern diese die peinliche Antwort wäre (bspw. bei der Frage „Haben Sie einen Schulabschluss?“ oder bei Befragungen in delinquenten Subkulturen nach Cohen & Short, 1958). Obwohl die peinliche Eigenschaft X durch diese Methode perfekt geschützt ist, liegt die wahrgenommene Sicherheit trotzdem nicht so hoch wie die Sicherheit von „symmetrischen“ Forced Answer-Modellen (siehe Bourke, 1984). In der ähnlichen Darstellungsformen wie bereits beim Warner-Modell 1.1 und dem Unrelated Question-Modell 1.3 ist in Abbildung 1.5 die schematische Darstellung des Forced Answer-Modells nach Dawes und More (1980) zu finden.

1.4 Das Cheater Detection-Modell

Die Tatsache, dass auch bei RRT-Befragungen noch immer eine gewisse Menge von gefährdeten (sich bedroht fühlenden) Befragten vorliegt, wurde bereits bei Greenberg, Kuebler, Abernathy & Horvitz (1977) diskutiert und von Tracy und Fox (1981) sowie von van der Heijden, van Gils, Bouts, & Hox (2000) empirisch nachgewiesen. So kann es vorkommen, dass einzelne Personen der Anonymität des Verfahrens nicht vertrauen und daher trotz der speziellen Technik nicht wahrheitsgemäß antworten. Außerdem könnten Befragte bewusst versuchen, den Ausgang der Befragung zu manipulieren (zu Auswirkungen dieser Verzerrungen s. Becker, 2010 sowie Böckenholt, Barlas & van der Heijden, 2009; Edgell et al., 1982 sowie im Überblick Wolter, 2012, 90ff.). Neben vielen anderen Ideen zur RRT-Entwicklung, die zu einem verstärkten Vertrauen der Antwortenden führen sollten, haben daher Clark und Desharnais (1998) das

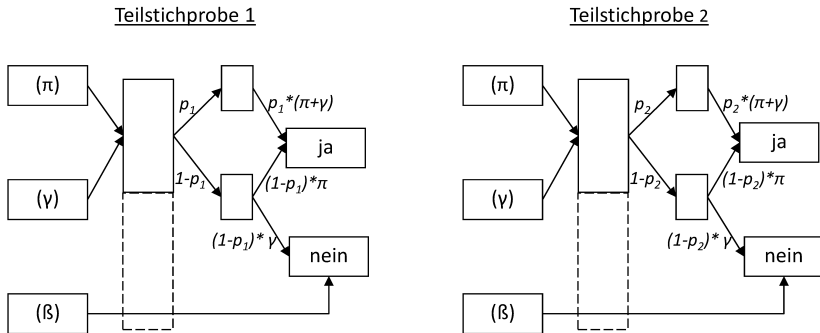


Abbildung 1.6: Schematische Darstellung des Forced Answer-Modells mit Cheater Detection. Doppelten Forced Answer mit $p_1 \neq p_2$ und $\beta + \gamma + \pi = 1$.

Cheater Detection-Modell entwickelt (vgl. schematische Darstellung in Abbildung 1.6). „Cheating“ bedeutet hier, die RRT-Instruktion nicht korrekt anzuwenden, unabhängig davon, ob die peinliche Eigenschaft vorliegt, oder nicht. Um den Anteil von Cheatern zu schätzen, wird die Stichprobe zufällig in zwei Teilstichproben unterteilt. Jede Teilstichprobe erhält eine Forced Answer-Frage mit unterschiedlichen Wahrscheinlichkeiten für diese Frage. Beim Einsatz dieser Technik erlaubt die Differenzierung in zwei Teilstichproben nicht nur den Anteil von zwei Gruppen (π und $1 - \pi$) in der Population zu schätzen, sondern den Anteil der folgenden drei Gruppen zu schätzen:

- π – die Rate der ehrlichen „Ja“-Sager, welche die niedrigste mögliche Prävalenz der peinlichen Eigenschaft X angibt,
- β – die Rate der ehrlichen „Nein“-Sager, welche die geringst mögliche Rate der Populationsmitglieder darstellt, welche die Eigenschaft X nicht haben und
- γ – gibt den Anteil der Cheater an, d.h. jene, die nicht gemäß der Anweisung in der RRT-Frage antworten.

In diesem Zusammenhang ist es wichtig, dass Cheating dann auftaucht, wenn Antwortende die die Eigenschaft X haben „nein“ antworten in den Fällen, wenn sie entweder „ja“ ehrlich oder als eine Forced Answer antworten sollten. Cheating kann aber auch auftauchen, wenn Antwortende die die Eigenschaft nicht haben „nein“ antworten, obwohl der Randomisierungsprozess zu

dem Ergebnis geführt hat, dass eine „ja“ Antwort hätte gegeben werden sollen. Das Nicht-Befolgen der RRT-Anweisungen, kann auch aufgrund von Missverstehen der Anweisungen, Missinterpretieren des Randomisierungsergebnisses oder dem Misstrauen gegenüber der Anonymität der Antwort herrühren. Die Methode liefert daher keine Informationen über die Cheater, was auch immer der Grund für das Cheating sein mag. In der ersten Veröffentlichung zur Methode bieten Clark und Desharnais (1998) hierfür einen χ^2 -Test mit der Hypothese $\gamma > 0$ und schlagen vor, auf die Ergebnisse zu verzichten, wenn dieser Test einen signifikanten Anteil an Cheatern anzeigt. In späteren Anwendungen wurde der geschätzte Anteil an ehrlichen „Ja“-Sagern als untere Grenze eines Prävalenzintervalls berichtet. Hierzu wurde der Cheateranteil addiert, um die obere Grenze des Intervalls, in dem der wahre Wert der Prävalenz von X in der Population liegt, zu erhalten (z. B. Musch, Bröder & Klauer, 2001; Pitsch & Emrich, 2012).

In der Logik der klassischen RRT ist dabei immer klar, welche Eigenschaft peinlich ist und zu Verzerrungen in Folge sozialer Erwünschtheit führt. In bestimmten Subkulturen kann gerade das Ausbleiben der Eigenschaft X jedoch peinlich sein. Ein Beispiel hierfür wäre etwa der Konsum illegaler Drogen unter Jugendlichen. Allgemein ist immer bei der Untersuchung in devianten Subkulturen die Möglichkeit der Überschätzung der in der Hauptkultur peinlichen Eigenschaft aufgrund subkultureller, von der Hauptkultur devianter, Normen gegeben (Cohen & Short, 1958; Cohen, 1957; Mays, 1957). In diesem Fall würden Befragte sich die Eigenschaft X fälschlicherweise zuschreiben und entgegen anderslautender Instruktionen immer mit „ja“ antworten. Darüber hinaus können diese Verzerrungen auch bei strategischem Antwortverhalten auftreten, wenn Befragte sich aus dieser Verzerrung Vorteile erhoffen oder, aus welchen Gründen auch immer, eine Befragung sabotieren wollen. Diese Überlegungen führten zur Verallgemeinerung der Idee von Clark und Desharnais (1998) zur Schätzung des Anteils von „Ja“-Cheatern, also jenen, die sich fälschlich die peinliche Eigenschaft zuschreiben. Ausgehend von diesen Überlegungen wird dieses Problem an dieser Stelle mathematisch bearbeitet. Ziel ist es, eine „Total Cheater Detection“ zu entwickeln und zu validieren, sodass beide Formen des Cheatings abgebildet werden können. Zunächst wird die Notation für die kommenden Kapitel und die grundlegende Modellierung geklärt, um darauf aufbauend das neue, erweiterte Verfahren herzuleiten. Anschließend erfolgt eine Analyse der Schätzer, ein Vergleich der unterschiedlichen Verfahren und schließlich Erläuterungen zu praktischen Anwendungen. Im abschließenden Ausblick wird auf mögliche Weiterentwicklungen des Verfahrens, über den hier beschriebenen Stand hinaus kurz angesprochen.

2 Modellierung und Notation

2.1 Generelle Notation

- Für $v \in \mathbb{R}^n$ sei v_i die i -te Komponente von v .

- Für $n \in \mathbb{N}$ sei $0_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n$ und $1_n := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$.

- Für $\alpha \in \mathbb{R}, n \in \mathbb{N}$ schreiben wir an einigen Stellen statt „ $\alpha \cdot 1_n$ “ lediglich „ α “, sofern n im Kontext offensichtlich ist.

- Es sei $M \in \mathbb{R}^{n \times m}$ eine Matrix mit vollem Spaltenrang. Die Spalten von M sind linear unabhängig und zu M existiert die Pseudoinverse $M^+ = (M^T M)^{-1} M^T$. Ist M quadratisch, so existiert die inverse M^{-1} und wegen $(M^T M)^{-1} M^T = M^{-1} (M^T)^{-1} M^T = M^{-1}$ fällt die Pseudoinverse M^+ mit der inversen Matrix M^{-1} zusammen. Die Pseudoinverse dient dazu, die Lösung überbestimmter Gleichungssysteme zu approximieren:

Zu einer gegebenen Koeffizientenmatrix $M \in \mathbb{R}^{n \times m}$ mit $n \geq m$ und den Variablen $x \in \mathbb{R}^m$ und $y \in \mathbb{R}^n$, wird durch $\hat{x} = M^+ y$ die Lösung des Gleichungssystems $Mx = y$ approximiert in dem Sinne, dass \hat{x} die Summe der Fehlerquadrate $\sum_{i=1}^n ((M\hat{x})_i - y_i)^2$ minimiert.

2.2 Population

Wir teilen die Population in vier Gruppen auf:

- α - Personen mit der Eigenschaft E, welche den Instruktionen stets folgen (ehrlche „Ja“-Sager).

- β - Personen mit der Eigenschaft E, welche stets „Nein“ antworten („Nein“-Cheater).
- γ - Personen ohne die Eigenschaft E, welche den Instruktionen stets folgen (ehrlische „Nein“-Sager).
- δ - Personen ohne die Eigenschaft E, welche stets „Ja“ antworten („Ja“-Cheater).

Gleichzeitig bezeichne der Gruppenname auch die relative Häufigkeit dieser Gruppe in der Gesamtpopulation. Es ergeben sich also

$$\alpha + \beta + \gamma + \delta = 1 \quad \text{und} \quad 0 \leq \alpha, \beta, \gamma, \delta \leq 1$$

als natürliche Randbedingungen.

2.3 Randomized Response-Technik

Als grundlegende Befragungstechnik bedienen wir uns der RRT mit Forced Answer. Hierbei sollen sowohl die „Ja“- als auch die „Nein“-Antworten geschützt werden. Wie schon bei dem Verfahren von Clark und Desharnais (1998), müssen die Befragten in mehrere Gruppen mit unterschiedlichen Forced Answer-Wahrscheinlichkeiten unterteilt werden. Um n Populationsparameter schätzen zu können, benötigt man mindestens $n - 1$ Gruppen: Naiv ausgedrückt erhalten wir aus jeder RRT-Gruppe eine Anforderung an die Populationsparameter. Zusammen mit der Randbedingung $\alpha + \beta + \gamma + \delta = 1$ haben wir dann n Informationen, aus denen sich die Populationsparameter berechnen lassen.

Für eine RRT mit R Gruppen seien

- $P^Y \in \mathbb{R}^R$ die Forced Yes-Wahrscheinlichkeiten,
- $P^N \in \mathbb{R}^R$ die Forced No-Wahrscheinlichkeiten und
- $P^S \in \mathbb{R}^R$ die Wahrscheinlichkeiten für die sensitive Frage.

Wir erhalten

$$0 \leq P_i^Y, P_i^N, P_i^S \leq 1, \quad 1 \leq i \leq R \quad \text{und} \quad P^Y + P^N + P^S = 1_R$$

als natürliche Bedingungen an diese Wahrscheinlichkeiten.

Ein weiteres Modell der RRT basiert auf der Methode der Unrelated Question. Hier entscheidet der Zufallsgenerator, ob der Befragte die sensitive Frage gestellt bekommt, oder ob ihm eine harmlose Frage, deren Antwortverhältnis

bekannt ist, gestellt wird wie z.B.: „Sind sie in einem ungeraden Monat geboren?“. Mathematisch lässt sich das Modell der Unrelated Question sehr einfach in das Forced Answer-Modell transformieren: Ist P die Wahrscheinlichkeit, die harmlose Ausweichfrage zu stellen und ϕ die Wahrscheinlichkeit, diese mit „Ja“ zu beantworten, erhält man $P^Y = P\phi$, $P^N = P(1 - \phi)$ und $P^S = 1 - P$.

2.4 Übersicht der Verfahren

Im Folgenden werden drei Verfahren hergeleitet:

1. Total Cheater Detection (TCD): Sie kann alle der obigen Populationsparameter schätzen.
2. No-Cheater Detection (NCD): Als Grundlage wird $\delta = 0$ angenommen. Sie kann die übrigen drei Parameter schätzen.
3. Yes-Cheater Detection (YCD): Analog zur NCD wird hier $\beta = 0$ vorausgesetzt und sie schätzt die restlichen drei Parameter.

Die letzten beiden Verfahren fassen wir unter dem Oberbegriff Directed Cheater Detection (DCD) zusammen. Diese beiden Verfahren können zwar weniger Parameter schätzen als die TCD und werden ungenau, wenn der als 0 angenommene Parameter in Wirklichkeit nicht 0 ist, liefern aber bei weniger Teilnehmern bereits zuverlässigere Schätzungen als die TCD.

3 Herleitung der Verfahren

3.1 Herleitung der Teilschätzer

Die Herleitung der Schätzer funktioniert für alle Verfahren nach dem selben Schema. Wir gehen also zunächst ganz allgemein von einer RRT-Befragung mit R unterschiedlichen Gruppen, und einem Verfahren, das s Populationsparameter schätzt, aus und verwenden die allgemeine Lösung danach für die Herleitung jedes Verfahrens. Wie bereits in Abschnitt 2.3 angemerkt, muss $R \geq s - 1$ sein, um die Schätzung durchführen zu können. Weiter müssen wir bei der Herleitung der Verfahren verschiedene Teilschätzer betrachten:

- Jedes Verfahren hat genau einen Schätzer, der $s - 1$ Populationsparameter berechnet. Der verbleibende Parameter wird über $\alpha + \beta + \gamma + \delta = 1$ ermittelt. Wir nennen diesen Schätzer den **nativen Schätzer** bzw. die durch ihn geschätzten Parameter die **native Lösung**. Diese Parameter fassen wir als $\Theta = (\alpha, \beta, \gamma, \delta)^T$ zusammen.
- Weiter müssen wir bei jedem Verfahren Randfälle betrachten, bei denen eine Teilmenge der Populationsparameter den Wert Null annimmt. Die Anzahl der als Null angenommen Parameter sei mit t bezeichnet.

Anders formuliert müssen wir für jede t -elementige Teilmenge der Populationsparameter mit $0 \leq t < s$ einen Teilschätzer betrachten. Für jedes Verfahren müssen also

$$\sum_{t=0}^{s-1} \binom{s}{t} = \sum_{t=0}^s \binom{s}{t} - 1 = 2^s - 1$$

Teilschätzer hergeleitet werden.

Für jeden Teilschätzer lässt sich die Wahrscheinlichkeit einer „Ja“-Antwort in jeder RRT-Gruppe aus den Populationsparametern und den Forced Answer-Wahrscheinlichkeiten berechnen. Es ergibt sich ein Gleichungssystem der Form $\lambda = M\Theta$, wobei $M \in \mathbb{R}^{R \times (s-t-1)}$ eine Koeffizientenmatrix abhängig von den

Forced Answer-Wahrscheinlichkeiten, $\Theta \in \mathbb{R}^{(s-t-1)}$ der Vektor der zu schätzenden Populationsparameter und $\lambda \in \mathbb{R}^R$ die errechneten „Ja“-Wahrscheinlichkeiten sind.

Beispiel

Wir betrachten den Randfall $\alpha = 0$ der NCD mit $R = 2$ RRT-Gruppen. Die NCD schätzt $s = 3$ Parameter (α, β, γ). Gemäß dem Randfall ist $t = 1$. Der einzige Beitrag zu den „Ja“-Antworten wird von ehrlichen „Nein“-Sagern erbracht, welche der Instruktion „Antworte Ja“ folgen und wir erhalten somit $\lambda = \gamma P^Y \in \mathbb{R}^R = \mathbb{R}^2$. Der zu schätzende Parameter ist $\Theta = \gamma \in \mathbb{R}^1$, da $\beta = 1 - \alpha - \gamma$ und $\alpha = 0$ gemäß dem Randfall. Aus $\lambda = M\Theta$ können wir nun

$$M = \begin{pmatrix} P_1^Y \\ P_2^Y \end{pmatrix} \in \mathbb{R}^{R \times (s-t-1)} = \mathbb{R}^{2 \times 1}$$

herleiten.

3.1.1 ML-Gleichung der Wahrscheinlichkeiten einer „Ja“-Antwort

Sind die Anzahlen der „Ja“-Antworten $k \in \mathbb{R}^R$ in den Gruppen und die Gruppenumfänge $N \in \mathbb{R}^R$ gegeben, können wir die Wahrscheinlichkeit einer „Ja“-Antwort in jeder Gruppe über die Maximum-Likelihood-Methode berechnen. Innerhalb jeder Gruppe ergibt sich eine Binomialverteilung für die Anzahl der „Ja“-Antworten. Weiter geschieht die Befragung der Gruppen unabhängig voneinander. Wird mit X_i die Anzahl der „Ja“-Antworten in Gruppe i bezeichnet, lässt sich die Likelihood somit als

$$\begin{aligned} L_\lambda(\lambda) &= \mathbb{P}_\lambda(X_1 = k_1, \dots, X_R = k_R) \\ &= \prod_{i=1}^R \mathbb{P}_\lambda(X_i = k_i) && \text{(Unabhängigkeit)} \\ &= \prod_{i=1}^R \binom{N_i}{k_i} \lambda_i^{k_i} (1 - \lambda_i)^{N_i - k_i} && \text{(Binomialverteilung)} \end{aligned}$$

darstellen. Für die weitere Rechnung ist die Verwendung der logarithmierten Likelihood

$$l_\lambda(\lambda) := \ln(L_\lambda(\lambda)) = \sum_{i=1}^R \ln \binom{N_i}{k_i} + k_i \ln \lambda_i + (N_i - k_i) \ln(1 - \lambda_i)$$

sinnvoll.

3.1.2 Lösung der ML-Gleichung

Lemma 1. Die logarithmierte Likelihood l_λ nimmt ihr globales Maximum bei $\hat{\lambda}_i = \frac{k_i}{N_i}$ an.

Beweis. Zunächst bestimmen wir die partiellen Ableitungen

$$\frac{\partial}{\partial \lambda_i} l_\lambda(\lambda) = \frac{k_i}{\lambda_i} - \frac{N_i - k_i}{1 - \lambda_i}.$$

Es ergeben sich $\lambda_i = \frac{k_i}{N_i}$ als Nullstellen der partiellen Ableitungen. Um nachzuweisen, dass dies wirklich ein Maximum ist, genügt es zu zeigen, dass die Hesse-Matrix $D^2 l_\lambda$ an dieser Stelle negativ definit ist. Die zweiten partiellen Ableitungen sind

$$\frac{\partial^2}{\partial \lambda_i \partial \lambda_j} l_\lambda(\lambda) = \begin{cases} -\frac{k_i}{\lambda_i^2} - \frac{N_i - k_i}{(1 - \lambda_i)^2} & i = j \\ 0 & i \neq j \end{cases}.$$

Weiter gilt

$$\operatorname{sgn} \left(-\frac{k_i}{\lambda_i^2} - \frac{N_i - k_i}{(1 - \lambda_i)^2} \right) = -1$$

wegen $k_i, N_i - k_i \geq 0$ und $N_i > 0 \Rightarrow k_i \neq 0 \vee N_i - k_i \neq 0$. Die Hesse-Matrix hat also die Form

$$D^2 l_\lambda(\lambda) = \begin{pmatrix} -\frac{k_1}{\lambda_1^2} - \frac{N_1 - k_1}{(1 - \lambda_1)^2} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\frac{k_R}{\lambda_R^2} - \frac{N_R - k_R}{(1 - \lambda_R)^2} \end{pmatrix}$$

mit ausschließlich negativen Elementen auf der Hauptdiagonalen und sonst überall Nullen. Deshalb haben die führenden Hauptminoren alternierende Vorzeichen und die Hesse-Matrix ist somit sogar überall negativ definit. Die logarithmierte Likelihood l_λ nimmt ihr Maximum also bei $\hat{\lambda}_i = \frac{k_i}{N_i}$ an.

Um nun daraus einen ML-Schätzer für Θ zu bestimmen, müssen wir zusätzlich fordern, dass die oben errechnete Koeffizientenmatrix M des Verfahrens vollen Spaltenrang $s - t - 1$ hat. Für die weitere Herleitung nehmen wir diese Forderung als erfüllt an und untersuchen in Kapitel 4.3, wie die Forced Answer-Wahrscheinlichkeiten zu wählen sind, damit diese Annahme auch wirklich erfüllt ist. Wir müssen nun zwei Fälle unterscheiden:

Fall 1: M quadratisch

Für quadratisches $M \neq 0$ ist die Abbildung $\Theta \mapsto M\Theta$ bijektiv (da M vollen Spaltenrang hat) und wir erhalten $l(\Theta) := l_\lambda(M\Theta)$ als interessierende Likelihood-Funktion. Da die Abbildung bijektiv ist, erhalten wir dann aus $\hat{\Theta} = M^{-1}\hat{\lambda}$ den ML-Schätzer für die Populationsparameter.

Fall 2: M nicht quadratisch

Ist M nicht quadratisch, erhalten wir mit $\hat{\Theta} = M^+\hat{\lambda}$ keinen reinen ML-Schätzer. Es wurde zwar $\hat{\lambda}$ nach der ML-Methode berechnet, die Verwendung der Pseudoinversen entspricht hier aber der Kleinste-Quadrate-Lösung des Gleichungssystems $M\hat{\Theta} = \hat{\lambda}$. Die Kleinste-Quadrate-Methode ist hierbei selbst nur dann ein ML-Schätzer, wenn die rechte Seite des Gleichungssystems normalverteilt ist. In unserem Fall folgen die $\hat{\lambda}_i$ aber einer Binomialverteilung.

Wir erhalten also nur bei einer regulären Koeffizienten-Matrix einen echten ML-Schätzer. Dies ist nur im Fall $R = s - t - 1$ möglich. Wegen $R \geq s - 1$ folgt weiter, dass dies nur beim nativen Schätzer möglich ist. Damit zumindest die native Lösung ein echter ML-Schätzer ist, muss also dem Verfahren entsprechend die Anzahl der RRT-Gruppen $R = s - 1$ gewählt werden.

3.1.3 Berechnung der Populationsparameter

Als **Verfahrensmatrix** wollen wir die Koeffizientenmatrix des nativen Schätzers bezeichnen. Die Vorgehensweise zur Berechnung der Populationsparameter hängt davon ab, ob die Verfahrensmatrix M quadratisch ist oder nicht.

Fall 1: M quadratisch

Obige Lösung $\hat{\Theta} = M^{-1}\hat{\lambda}$ maximiert zwar die Likelihood-Funktion, beachtet aber lediglich die Randbedingung $\alpha + \beta + \gamma + \delta = 1$. Da es sich bei diesen Parametern um Anteile der Population handelt, sind nur $\alpha, \beta, \gamma, \delta \in [0, 1]$ zulässig. Ist diese Anforderung nicht erfüllt, müssen alle Randlösungen und deren Likelihood berechnet werden. Die beste Schätzung für die Populationsparameter liefert dann genau die Randlösung mit der höchsten Likelihood, welche den Anforderungen an α, β, γ und δ genügt.

Fall 2: M nicht quadratisch

In diesem Fall muss auch die Likelihood der nativen Lösung $\hat{\Theta} = M^+ \hat{\lambda}$ mit allen Randlösungen verglichen werden, unabhängig davon, ob $\hat{\Theta}$ die Randbedingungen erfüllt, da der nativen Lösung in diesem Fall kein ML-Schätzer zugrunde liegt.

3.1.4 ML-Gleichung der Wahrscheinlichkeiten einer „Nein“-Antwort

Bei der Herleitung der Verfahren werden wir gelegentlich anstelle des Gleichungssystems $\lambda = M\Theta$ für die Wahrscheinlichkeiten einer „Ja“-Antwort ein Gleichungssystem $1_R - \lambda = M'\Theta$ erhalten, welches die Wahrscheinlichkeiten einer „Nein“-Antwort wiedergibt.

Lemma 2. *Durch $\hat{\mu} := 1_R - \hat{\lambda} = 1 - \frac{k_i}{N_i}$ ist ein ML-Schätzer für die Wahrscheinlichkeiten einer „Nein“-Antwort gegeben.*

Beweis. Mit k und N wie oben berechnet sich die Likelihood durch

$$\begin{aligned}
 L_{\mu}(\mu) &= \mathbb{P}_{\lambda}(X_1 = N_1 - k_1, \dots, X_R = N_R - k_R) \\
 &= \prod_{i=1}^R \mathbb{P}_{\lambda}(X_i = N_i - k_i) \\
 &= \prod_{i=1}^R \binom{N_i}{N_i - k_i} \mu_i^{N_i - k_i} (1 - \mu_i)^{k_i} \\
 &= \prod_{i=1}^R \binom{N_i}{k_i} (1 - \lambda_i)^{N_i - k_i} \lambda_i^{k_i} \\
 &= L_{\lambda}(\lambda)
 \end{aligned}$$

und wir erhalten gemäß Lemma 1 bei $\lambda_i = \frac{k_i}{N_i}$ das globale Maximum.

Im Folgenden werden wir $1 - \hat{\lambda}$ anstelle von $\hat{\mu}$ schreiben, weil die Resultate der Rechnungen diese Schreibweise suggerieren und um die Notation möglichst gering zu halten.

3.2 Herleitung der TCD

3.2.1 Berechnung der „Ja“-Wahrscheinlichkeiten

Wir setzen für dieses Verfahren einen Umfang von drei RRT-Gruppen an. Mehr Gruppen sind auch möglich, aber wie wir oben gesehen haben, können wir so zumindest sicher stellen, dass die native Lösung auf einem ML-Schätzer basiert.

Zur Möglichkeit einer „Ja“-Antwort haben wir die folgenden Anteile:

- Eine Person aus Gruppe α antwortet auf die sensitive Frage immer mit „Ja“.
- Eine Person aus Gruppe α befolgt die Instruktion „Antworte Ja“.
- Eine Person aus Gruppe γ befolgt die Instruktion „Antworte Ja“.
- Eine Person aus Gruppe δ antwortet immer „Ja“.

Korollar 1. *Bei der TCD sind die Wahrscheinlichkeiten der „Ja“-Antworten durch*

$$\begin{aligned}\lambda &= \alpha (P^S + P^Y) + \beta \cdot 0_3 + \gamma P^Y + \delta \cdot 1_3 \\ &= \alpha (1_3 - P^N) + \gamma P^Y + \delta\end{aligned}$$

gegeben.

3.2.2 Berechnung der Teilschätzer

Zunächst wollen wir den zulässigen Wertebereich der Schätzer geometrisch interpretieren. Wir sehen, dass β aufgrund der „Ja“-Wahrscheinlichkeiten nur implizit über $\beta = 1 - \alpha - \gamma - \delta$ geschätzt wird. Der zulässige Wertebereich der Schätzer kann also über die restlichen drei Parameter im \mathbb{R}^3 veranschaulicht werden. Für α, γ und δ gelten stets die Bedingungen $\alpha, \gamma, \delta \geq 0$ und $\alpha + \gamma + \delta \leq 1$. Jede dieser vier Bedingungen definiert einen abgeschlossenen Halbraum im \mathbb{R}^3 . Der gültige Bereich für α, γ, δ ergibt sich als Überschneidung all dieser Halbräume. Diese erzeugt ein Tetraeder (siehe Abbildung 3.1).

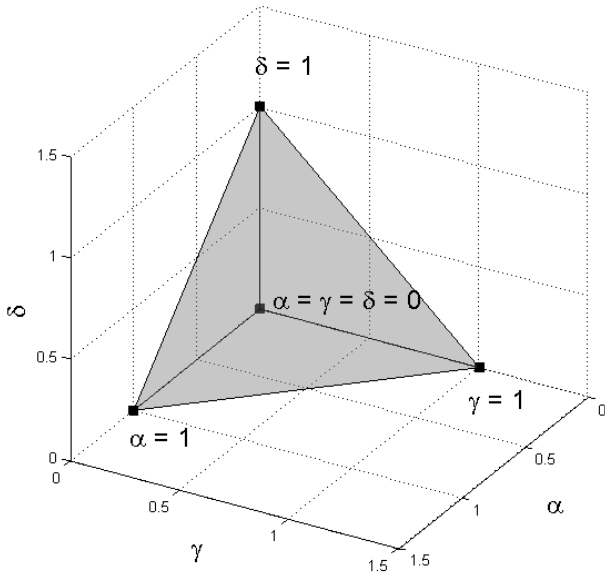


Abbildung 3.1: Darstellung des Definitionsbereichs von α , β und γ . Alle zulässigen Punkte (α, β, γ) liegen im Inneren und auf den Begrenzungsflächen des Tetraeders.

Native Lösung

Zulässige Werte der nativen Lösung liegen genau im Inneren des Tetraeders. Mit $M = (1_3 - P^N, P^Y, 1_3)$ erhalten wir die Parameter

$$\hat{\Theta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \\ \hat{\delta} \end{pmatrix} = M^{-1} \hat{\lambda}, \quad \hat{\beta} = 1 - \hat{\alpha} - \hat{\gamma} - \hat{\delta}.$$

Randlösungen 1. Ordnung

Bei diesen Randlösungen ist jeweils genau einer der vier Populationsparameter gleich Null. Die restlichen drei summieren sich zu Eins auf, deswegen kann man die zulässigen Lösungen genau auf den Begrenzungsflächen des Tetraeders finden.

1. Randfall $\alpha = 0$

Der Ansatz vereinfacht sich zu

$$\lambda = \gamma P^Y + \delta.$$

Mit $M = (P^Y, 1_3)$ erhalten wir die Parameter

$$\hat{\Theta} = \begin{pmatrix} \hat{\gamma} \\ \hat{\delta} \end{pmatrix} = M^+ \hat{\lambda}, \quad \hat{\alpha} = 0, \quad \hat{\beta} = 1 - \hat{\gamma} - \hat{\delta}.$$

2. Randfall $\beta = 0$

Wegen $\delta = 1 - \alpha - \gamma$ lässt sich der Ansatz umschreiben zu

$$\lambda = \alpha(1_3 - P^N) + \gamma P^Y + 1_3 - \alpha - \gamma \Leftrightarrow 1_3 - \lambda = \alpha P^N + \gamma (1_3 - P^Y).$$

Mit $M = (P^N, 1_3 - P^Y)$ erhalten wir die Parameter

$$\hat{\Theta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \end{pmatrix} = M^+ (1 - \hat{\lambda}), \quad \hat{\beta} = 0, \quad \hat{\delta} = 1 - \hat{\alpha} - \hat{\gamma}.$$

3. Randfall $\gamma = 0$

Wegen $\delta = 1 - \alpha - \beta$ lässt sich der Ansatz umschreiben zu

$$\lambda = \alpha(1_3 - P^N) + 1_3 - \alpha - \beta \Leftrightarrow 1_3 - \lambda = \alpha P^N + \beta.$$

Mit $M = (P^N, 1_3)$ erhalten wir die Parameter

$$\hat{\Theta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = M^+ (1 - \hat{\lambda}), \quad \hat{\gamma} = 0, \quad \hat{\delta} = 1 - \hat{\alpha} - \hat{\beta}.$$

4. Randfall $\delta = 0$

Der Ansatz vereinfacht sich zu

$$\lambda = \alpha(1_3 - P^N) + \gamma P^Y.$$

Mit $M = (1_3 - P^N, P^Y)$ erhalten wir die Parameter

$$\hat{\Theta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \end{pmatrix} = M^+ \hat{\lambda}, \quad \hat{\delta} = 0, \quad \hat{\beta} = 1 - \hat{\alpha} - \hat{\gamma}.$$

Randlösungen 2. Ordnung

Hier sind jeweils genau zwei Populationsparameter gleich Null. Sie liegen in der Schnittmenge von jeweils zwei Begrenzungsflächen, sie liegen also genau auf den Kanten des Tetraeders.

5. Randfall $\alpha = \beta = 0$

Wegen $\delta = 1 - \gamma$ lässt sich der Ansatz umschreiben zu

$$\lambda = \gamma P^Y + 1_3 - \gamma \Leftrightarrow 1_3 - \lambda = \gamma (1_3 - P^Y).$$

Mit $M = (1_3 - P^Y)$ erhalten wir die Parameter

$$\hat{\gamma} = M^+ (1 - \hat{\lambda}), \quad \hat{\alpha} = \hat{\beta} = 0, \quad \hat{\delta} = 1 - \hat{\gamma}.$$

6. Randfall $\alpha = \gamma = 0$

Der Ansatz vereinfacht sich zu

$$\lambda = \delta.$$

Mit $M = 1_3$ erhalten wir die Parameter

$$\hat{\delta} = M^+ \hat{\lambda}, \quad \hat{\alpha} = \hat{\gamma} = 0, \quad \hat{\beta} = 1 - \hat{\delta}.$$

7. Randfall $\alpha = \delta = 0$

Der Ansatz vereinfacht sich zu

$$\lambda = \gamma P^Y.$$

Mit $M = P^Y$ erhalten wir die Parameter

$$\hat{\gamma} = M^+ \hat{\lambda}, \quad \hat{\alpha} = \hat{\delta} = 0, \quad \hat{\beta} = 1 - \hat{\gamma}.$$

8. Randfall $\beta = \gamma = 0$

Wegen $\delta = 1_3 - \alpha$ lässt sich der Ansatz umschreiben zu

$$\lambda = \alpha(1_3 - P^N) + 1_3 - \alpha \Leftrightarrow 1_3 - \lambda = \alpha P^N.$$

Mit $M = P^N$ erhalten wir die Parameter

$$\hat{\alpha} = M^+ (1 - \hat{\lambda}), \quad \hat{\beta} = \hat{\gamma} = 0, \quad \hat{\delta} = 1 - \hat{\alpha}.$$

9. Randfall $\beta = \delta = 0$

Wegen $\gamma = 1_3 - \alpha$ lässt sich der Ansatz umschreiben zu

$$\lambda = \alpha(1_3 - P^N) + (1_3 - \alpha) P^Y \Leftrightarrow \lambda = \alpha(1_3 - P^Y - P^N) + P^Y.$$

Mit $M = 1_3 - P^Y - P^N$ erhalten wir als Zwischenschritt

$$\hat{\lambda} = M\alpha + P^Y \Leftrightarrow \hat{\lambda} - P^Y = M\alpha.$$

Somit ergeben sich die Parameter

$$\hat{\alpha} = M^+ \left(\hat{\lambda} - P^Y \right), \quad \hat{\beta} = \hat{\delta} = 0, \quad \hat{\gamma} = 1 - \hat{\alpha}.$$

10. Randfall $\gamma = \delta = 0$

Der Ansatz vereinfacht sich zu

$$\lambda = \alpha(1_3 - P^N).$$

Mit $M = 1_3 - P^N$ erhalten wir die Parameter

$$\hat{\alpha} = M^+ \hat{\lambda}, \quad \hat{\gamma} = \hat{\delta} = 0, \quad \hat{\beta} = 1 - \hat{\alpha}.$$

Randlösungen 3. Ordnung

Diese Lösungen liegen genau in den Ecken des Tetraeders.

11. Randfall $\alpha = \beta = \gamma = 0$

Die Parameter sind

$$\hat{\alpha} = 0, \quad \hat{\beta} = 0, \quad \hat{\gamma} = 0, \quad \hat{\delta} = 1.$$

12. Randfall $\alpha = \beta = \delta = 0$

Die Parameter sind

$$\hat{\alpha} = 0, \quad \hat{\beta} = 0, \quad \hat{\gamma} = 1, \quad \hat{\delta} = 0.$$

13. Randfall $\alpha = \gamma = \delta = 0$

Die Parameter sind

$$\hat{\alpha} = 0, \quad \hat{\beta} = 1, \quad \hat{\gamma} = 0, \quad \hat{\delta} = 0.$$

14. Randfall $\beta = \gamma = \delta = 0$

Die Parameter sind

$$\hat{\alpha} = 1, \quad \hat{\beta} = 0, \quad \hat{\gamma} = 0, \quad \hat{\delta} = 0.$$

3.3 Herleitung der NCD**3.3.1 Berechnung der „Ja“-Wahrscheinlichkeiten**

Bei diesem Verfahren sollen nur die „Nein“-Cheater erkannt werden. Wir nehmen an, dass in den Populationsparametern $\delta = 0$ gilt. Ist diese Annahme nicht erfüllt, erhalten wir einen verzerrten Schätzer. Siehe hierzu auch Kapitel 4.6.1. Für dieses Verfahren benötigen wir zwei RRT-Gruppen.

Für eine „Ja“-Antwort bestehen folgende Möglichkeiten:

- Eine Person aus Gruppe α antwortet auf die sensitive Frage immer mit „Ja“.
- Eine Person aus Gruppe α befolgt die Instruktion „Antworte Ja“.
- Eine Person aus Gruppe γ befolgt die Instruktion „Antworte Ja“.

Korollar 2. *Bei der NCD sind die Wahrscheinlichkeiten der „Ja“-Antworten durch*

$$\begin{aligned} \lambda &= \alpha (P^S + P^Y) + \beta \cdot 0_2 + \gamma P^Y \\ &= \alpha (1_2 - P^N) + \gamma P^Y \end{aligned}$$

gegeben.

3.3.2 Berechnung der Lösungen

Analog zur TCD lässt sich der Bereich der zulässigen Lösungen geometrisch interpretieren. Er wird hier durch ein Dreieck im \mathbb{R}^3 gebildet. Die native Lösung liegt im Inneren des Dreiecks und die zulässigen Randlösungen entsprechend auf den Seiten bzw. in den Ecken.

Native Lösung

Mit $M = (1_2 - P^N, P^Y)$ ergeben sich die Parameter

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \end{pmatrix} = M^{-1} \hat{\lambda}, \quad \hat{\beta} = 1 - \hat{\alpha} - \hat{\gamma}.$$

Randlösungen 1. Ordnung**1. Randfall $\alpha = 0$**

Der Ansatz vereinfacht sich zu

$$\lambda = \gamma P^Y.$$

Mit $M = P^Y$ erhalten wir die Parameter

$$\hat{\gamma} = M^+ \hat{\lambda}, \quad \hat{\alpha} = 0, \quad \hat{\beta} = 1 - \hat{\gamma}.$$

2. Randfall $\beta = 0$

Wegen $\gamma = 1 - \alpha$ lässt sich der Ansatz umschreiben zu

$$\lambda = \alpha(1_2 - P^N) + \gamma P^Y \Leftrightarrow \lambda - P^Y = \alpha(1_2 - P^N - P^Y).$$

Mit $M = 1_2 - P^Y - P^N$ erhalten wir die Parameter

$$\hat{\alpha} = M^+ (\hat{\lambda} - P^Y), \quad \hat{\beta} = 0, \quad \hat{\gamma} = 1 - \hat{\alpha}.$$

3. Randfall $\gamma = 0$

Der Ansatz vereinfacht sich zu

$$\lambda = \alpha (1_2 - P^N).$$

Mit $M = 1_2 - P^N$ erhalten wir die Parameter

$$\hat{\alpha} = M^+ \hat{\lambda}, \quad \hat{\gamma} = 0, \quad \hat{\beta} = 1 - \hat{\alpha}.$$

Randlösungen 2. Ordnung**4. Randfall $\alpha = \beta = 0$**

Die Parameter sind

$$\hat{\alpha} = 0, \quad \hat{\beta} = 0, \quad \hat{\gamma} = 1.$$

5. Randfall $\alpha = \gamma = 0$

Die Parameter sind

$$\hat{\alpha} = 0, \quad \hat{\beta} = 1, \quad \hat{\gamma} = 0.$$

6. Randfall $\beta = \gamma = 0$

Die Parameter sind

$$\hat{\alpha} = 1, \quad \hat{\beta} = 0, \quad \hat{\gamma} = 0.$$

3.4 Herleitung der YCD**3.4.1 Berechnung der „Nein“-Wahrscheinlichkeiten**

Nun sollen die „Ja“-Cheater erkannt werden. Analog zur NCD nehmen wir an, dass in den Populationsparametern $\beta = 0$ gilt. Wir erhalten ebenfalls einen verzerrten Schätzer, falls diese Annahme nicht erfüllt ist (Kapitel 4.7.1). Auch hier benötigen wir zwei RRT-Gruppen.

Für eine „Ja“-Antwort bestehen folgende Möglichkeiten:

- Eine Person aus Gruppe α antwortet auf die sensitive Frage immer mit „Ja“.
- Eine Person aus Gruppe α befolgt die Instruktion „Antworte Ja“.
- Eine Person aus Gruppe γ befolgt die Instruktion „Antworte Ja“.
- Eine Person aus Gruppe δ antwortet immer mit „Ja“.

Korollar 3. *Bei der YCD sind die Wahrscheinlichkeiten der „Ja“-Antworten durch*

$$\lambda = \alpha (P^S + P^Y) + \gamma P^Y + \delta \cdot 1_2$$

gegeben. Mit $\delta = 1 - \alpha - \beta$ erhalten wir

$$1_2 - \lambda = \alpha P^N + \gamma (1_2 - P^Y)$$

als Wahrscheinlichkeiten für die „Nein“-Antworten.

3.4.2 Berechnung der Lösungen

Da die YCD und die NCD die selbe Anzahl Parameter schätzen, erhalten wir hier genau die gleiche geometrische Interpretation wie bei der NCD.

Native Lösung

Mit $M = (P^N, 1_2 - P^Y)$ ergeben sich die Parameter

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\gamma} \end{pmatrix} = M^{-1} \begin{pmatrix} 1 - \hat{\lambda} \end{pmatrix}, \quad \hat{\delta} = 1 - \hat{\alpha} - \hat{\gamma}.$$

Randlösungen 1. Ordnung**1. Randfall $\alpha = 0$**

Der Ansatz vereinfacht sich zu

$$1_2 - \lambda = \gamma (1_2 - P^Y).$$

Mit $M = 1_2 - P^Y$ erhalten wir die Parameter

$$\hat{\gamma} = M^+ \begin{pmatrix} 1 - \hat{\lambda} \end{pmatrix}, \quad \hat{\alpha} = 0, \quad \hat{\delta} = 1 - \hat{\gamma}.$$

2. Randfall $\gamma = 0$

Der Ansatz vereinfacht sich zu

$$1_2 - \lambda = \alpha P^N.$$

Mit $M = P^N$ erhalten wir die Parameter

$$\hat{\alpha} = M^+ \begin{pmatrix} 1 - \hat{\lambda} \end{pmatrix}, \quad \hat{\gamma} = 0, \quad \hat{\delta} = 1 - \hat{\alpha}.$$

3. Randfall $\delta = 0$

Wegen $\gamma = 1 - \alpha$ lässt sich der Ansatz umschreiben zu

$$\lambda = \alpha(1_2 - P^N) + \gamma P^Y \Leftrightarrow \lambda - P^Y = \alpha(1_2 - P^N - P^Y).$$

Mit $M = 1_2 - P^Y - P^N$ erhalten wir die Parameter

$$\hat{\alpha} = M^+ \begin{pmatrix} \hat{\lambda} - P^Y \end{pmatrix}, \quad \hat{\delta} = 0, \quad \hat{\gamma} = 1 - \hat{\alpha}.$$

Randlösungen 2. Ordnung**4. Randfall $\alpha = \gamma = 0$**

Die Parameter sind

$$\hat{\alpha} = 0, \quad \hat{\gamma} = 0, \quad \hat{\delta} = 1.$$

5. Randfall $\alpha = \delta = 0$

Die Parameter sind

$$\hat{\alpha} = 0, \quad \hat{\gamma} = 1, \quad \hat{\delta} = 0.$$

6. Randfall $\gamma = \delta = 0$

Die Parameter sind

$$\hat{\alpha} = 1, \quad \hat{\gamma} = 0, \quad \hat{\delta} = 0.$$

3.5 Praktische Anwendung

Der Einsatz dieser Verfahren wird nun am Beispiel der TCD (siehe 3.1.3) geschildert. Für die NCD und YCD gilt dieses Vorgehen analog.

- Berechne die native Lösung. Sind alle Parameter im zulässigen Bereich, ist dies die Schätzung der Populationsparameter.
- Ist die native Lösung nicht zulässig, so berechne alle Randlösungen und deren Likelihood (oder Log-Likelihood). Bestimme aus allen zulässigen Randlösungen diejenige mit der höchsten Likelihood. Sie ist die Schätzung der Populationsparameter. Solch eine Lösung existiert, da die Randlösungen 3. Ordnung immer zulässig sind.

Um die Likelihood einer Lösung zu ermitteln berechnet man im ersten Schritt $\lambda = M\Theta$, wobei M die Verfahrensmatrix (Koeffizienten-Matrix des nativen Schätzers) ist und $\Theta = (\alpha, \gamma, \delta)^T$ die von der Lösung geschätzten Populationsparameter. Dieses λ setzt man dann zusammen mit N und k in die oben hergeleitete Formel der Likelihood oder Log-Likelihood ein. Die Berechnung der Binomialkoeffizienten ist dabei nicht nötig, da N und k bei allen Teilschätzern gleich sind und somit die Likelihood um einen festen Faktor strecken bzw. die Log-Likelihood um einen festen Summanden vergrößern.

Wie die Forced Answer-Wahrscheinlichkeiten (und somit die Verfahrensmatrix) zu wählen sind, wird im nächsten Kapitel, speziell in 4.3 untersucht.

4 Analyse der Schätzer

In diesem Kapitel untersuchen wir die Schätzer zunächst mathematisch, indem wir die Varianzen und Erwartungswerte der Lösungen analytisch herleiten. Die Berechnung der Varianzen beschränken wir hierbei auf die nativen Lösungen. Bei jedem der Verfahren zeigt sich, dass die einzelnen Teilschätzer der nativen Lösung und der Randlösungen erwartungstreu sind. Die zusammengesetzten Schätzer, also die Verfahren selbst, sind jedoch in Nähe der Randfälle verzerrt, da etwa in Nähe des Randfalls $\alpha = 0$ der native Schätzer $\hat{\alpha}$ den Parameter überschätzen kann, bei einer Unterschätzung mit $\hat{\alpha} < 0$ greift dann aber der Teilschätzer für den Randfall, welcher $\hat{\alpha} = 0$ liefert, so dass im Mittel eine Überschätzung vorliegt. Diese Verzerrung in Nähe der Randfälle werden wir aufgrund ihrer Komplexität mittels Monte-Carlo-Simulation untersuchen.

4.1 Berechnung der Varianzen der Schätzer

Um die Varianz der Schätzer darzustellen, benutzen wir die Beziehung

$$M^{-1} = \frac{1}{\det M} \text{adj}(M)$$

da so der Zusammenhang mit der Determinante der Verfahrensmatrix offensichtlich wird. Die allgemeine Form der Adjunkten für (3×3) -Matrizen ist durch

$$\text{adj} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{pmatrix}$$

gegeben.

Weiter benötigen wir auch die Varianzen $\text{Var}(\hat{\lambda}_i)$ der Schätzer und die Kovarianzen $\text{Cov}(\hat{\lambda}_i, \hat{\lambda}_j)$. Für $i \neq j$ ist $\text{Cov}(\hat{\lambda}_i, \hat{\lambda}_j) = 0$, da die Datenerhebungen in den RRT-Gruppen unabhängig voneinander sind. Bei der Herleitung der Schätzer erhielten wir $\hat{\lambda}_i = \frac{k_i}{N_i}$, wobei k_i die absoluten Häufigkeiten der „Ja“-Antworten in den RRT-Gruppen und N_i die Gruppenumfänge sind. Dabei sind

die k_i Zufallsvariablen, welche einer Binomialverteilung mit N_i Wiederholungen und den Erfolgswahrscheinlichkeiten λ_i folgen.

Lemma 3. *Somit erhalten wir*

$$\text{Var}(\hat{\lambda}_i) = \text{Var}\left(\frac{k_i}{N_i}\right) = \frac{\text{Var}(k_i)}{N_i^2} = \frac{N_i \lambda_i (1 - \lambda_i)}{N_i^2} = \frac{\lambda_i (1 - \lambda_i)}{N_i}$$

wobei die Erfolgswahrscheinlichkeiten λ_i von dem gewählten Verfahren, den Forced Answer-Wahrscheinlichkeiten und den Populationsparametern abhängen (siehe Korollar 1, 2 und 3).

4.1.1 Varianzen der TCD

Die Verfahrensmatrix der TCD ist

$$M = \begin{pmatrix} 1 - P_1^N & P_1^Y & 1 \\ 1 - P_2^N & P_2^Y & 1 \\ 1 - P_3^N & P_3^Y & 1 \end{pmatrix}.$$

Somit erhalten wir

$$\begin{aligned} M^{-1} &= \frac{1}{\det M} \text{adj}(M) = \frac{1}{\det M} (m_{ij})_{1 \leq i, j \leq 3} \\ &= \frac{1}{\det M} \begin{pmatrix} P_2^Y - P_3^Y & P_3^Y - P_1^Y & P_1^Y - P_2^Y \\ P_2^N - P_3^N & P_3^N - P_1^N & P_1^N - P_2^N \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \end{aligned}$$

mit

$$\begin{aligned} m_{31} &= (1 - P_2^N) P_3^Y - P_2^Y (1 - P_3^N) \\ m_{32} &= P_1^Y (1 - P_3^N) - (1 - P_1^N) P_3^Y \\ m_{33} &= (1 - P_1^N) P_2^Y - P_1^Y (1 - P_2^N). \end{aligned}$$

Die Wahrscheinlichkeiten für „Ja“-Antworten berechnen sich durch

$$\lambda_i = \alpha (1 - P_i^N) + \gamma P_i^Y + \delta.$$

Varianz von $\hat{\alpha}$

Mit obiger Darstellung der Inversen der Verfahrensmatrix lässt sich der Schätzer $\hat{\alpha}$ wie folgt darstellen:

$$\hat{\alpha} = \frac{1}{\det M} \left((P_2^Y - P_3^Y) \hat{\lambda}_1 + (P_3^Y - P_1^Y) \hat{\lambda}_2 + (P_1^Y - P_2^Y) \hat{\lambda}_3 \right)$$

Für die Varianz des Schätzers erhalten wir

$$\begin{aligned}\text{Var}(\hat{\alpha}) &= \text{Var} \left(\frac{1}{\det M} \left((P_2^Y - P_3^Y) \hat{\lambda}_1 + (P_3^Y - P_1^Y) \hat{\lambda}_2 \right. \right. \\ &\quad \left. \left. + (P_1^Y - P_2^Y) \hat{\lambda}_3 \right) \right) \\ &= \frac{1}{(\det M)^2} \text{Var} \left((P_2^Y - P_3^Y) \hat{\lambda}_1 + (P_3^Y - P_1^Y) \hat{\lambda}_2 + (P_1^Y - P_2^Y) \hat{\lambda}_3 \right).\end{aligned}$$

Da $\text{Cov}(\hat{\lambda}_i, \hat{\lambda}_j) = 0$ für $i \neq j$ gilt für die Gleichung

$$\begin{aligned}\text{Var}(\hat{\alpha}) &= \frac{1}{(\det M)^2} \left(\text{Var} \left((P_2^Y - P_3^Y) \hat{\lambda}_1 \right) + \text{Var} \left((P_3^Y - P_1^Y) \hat{\lambda}_2 \right) \right. \\ &\quad \left. + \text{Var} \left((P_1^Y - P_2^Y) \hat{\lambda}_3 \right) \right) \\ &= \frac{1}{(\det M)^2} \left((P_2^Y - P_3^Y)^2 \text{Var} \left(\hat{\lambda}_1 \right) + (P_3^Y - P_1^Y)^2 \text{Var} \left(\hat{\lambda}_2 \right) \right. \\ &\quad \left. + (P_1^Y - P_2^Y)^2 \text{Var} \left(\hat{\lambda}_3 \right) \right) \\ &= \frac{1}{(\det M)^2} \left((P_2^Y - P_3^Y)^2 \frac{\lambda_1 (1 - \lambda_1)}{N_1} + (P_3^Y - P_1^Y)^2 \frac{\lambda_2 (1 - \lambda_2)}{N_2} \right. \\ &\quad \left. + (P_1^Y - P_2^Y)^2 \frac{\lambda_3 (1 - \lambda_3)}{N_3} \right) \\ &\stackrel{(*)}{\approx} \frac{3}{N (\det M)^2} \left((P_2^Y - P_3^Y)^2 \lambda_1 (1 - \lambda_1) + (P_3^Y - P_1^Y)^2 \lambda_2 (1 - \lambda_2) \right. \\ &\quad \left. + (P_1^Y - P_2^Y)^2 \lambda_3 (1 - \lambda_3) \right),\end{aligned}$$

wobei die letzte Abschätzung (*) nur gültig ist, wenn alle drei RRT-Gruppen etwa gleich groß sind, also $N_1, N_2, N_3 \approx \frac{1}{3}N$. Diese Abschätzung erlaubt uns in den folgenden Abschnitten das Aufstellen einer Faustformel zur Entwicklung der Varianz bezüglich der Stichprobengröße.

Varianz von $\hat{\gamma}$

Wir erhalten

$$\hat{\gamma} = \frac{1}{\det M} \left((P_2^N - P_3^N) \hat{\lambda}_1 + (P_3^N - P_1^N) \hat{\lambda}_2 + (P_1^N - P_2^N) \hat{\lambda}_3 \right)$$

als neue Darstellung des Schätzers. Analog zu $\text{Var}(\hat{\alpha})$ lässt sich

$$\begin{aligned}\text{Var}(\hat{\gamma}) &= \frac{1}{(\det M)^2} \left(m_{21}^2 \frac{\lambda_1 (1 - \lambda_1)}{N_1} + m_{22}^2 \frac{\lambda_2 (1 - \lambda_2)}{N_2} \right. \\ &\quad \left. + m_{23}^2 \frac{\lambda_3 (1 - \lambda_3)}{N_3} \right) \\ &\approx \frac{3}{N (\det M)^2} (m_{21}^2 \lambda_1 (1 - \lambda_1) + m_{22}^2 \lambda_2 (1 - \lambda_2) \\ &\quad + m_{23}^2 \lambda_3 (1 - \lambda_3))\end{aligned}$$

berechnen.

Varianz von $\hat{\delta}$

$$\hat{\delta} = \frac{1}{\det M} (m_{31} \hat{\lambda}_1 + m_{32} \hat{\lambda}_2 + m_{33} \hat{\lambda}_3)$$

Auch hier verläuft die Rechnung vollkommen analog.

$$\begin{aligned}\text{Var}(\hat{\delta}) &= \frac{1}{(\det M)^2} \left(m_{31}^2 \frac{\lambda_1 (1 - \lambda_1)}{N_1} + m_{32}^2 \frac{\lambda_2 (1 - \lambda_2)}{N_2} \right. \\ &\quad \left. + m_{33}^2 \frac{\lambda_3 (1 - \lambda_3)}{N_3} \right) \\ &\approx \frac{3}{N (\det M)^2} (m_{31}^2 \lambda_1 (1 - \lambda_1) + m_{32}^2 \lambda_2 (1 - \lambda_2) \\ &\quad + m_{33}^2 \lambda_3 (1 - \lambda_3))\end{aligned}$$

Varianz von $\hat{\beta}$

Die Varianz wird durch

$$\begin{aligned}\text{Var}(\hat{\beta}) &= \text{Var}(1 - \hat{\alpha} - \hat{\gamma} - \hat{\delta}) \\ &= \text{Var}(\hat{\alpha}) + \text{Var}(\hat{\gamma}) + \text{Var}(\hat{\delta}) \\ &\quad + 2 \left(\text{Cov}(\hat{\alpha}, \hat{\gamma}) + \text{Cov}(\hat{\alpha}, \hat{\delta}) + \text{Cov}(\hat{\gamma}, \hat{\delta}) \right)\end{aligned}$$

berechnet. Die Berechnung der Kovarianzen ist jedoch recht aufwendig und der resultierende Ausdruck lässt sich nicht kompakt zusammenfassen. Deswegen wollen wir hier nur das grobe Vorgehen skizzieren. Am Ende dieses Kapitels

werden wir ein empirisches Hilfsmittel einführen, mit dem wir von der komplizierten Form der Varianzen abstrahieren können.

Seien $\chi, \psi \in \{\alpha, \gamma, \delta\}$. Es gilt $E(\hat{\chi}) = \chi$ (siehe Kapitel 4.2). Wir definieren

$$i(\chi) := \begin{cases} 1 & \text{für } \chi = \alpha \\ 2 & \text{für } \chi = \gamma \\ 3 & \text{für } \chi = \delta \end{cases}$$

Die Kovarianzen lassen sich mittels

$$\text{Cov}(\hat{\chi}, \hat{\psi}) = E(\hat{\chi} \cdot \hat{\psi}) - E(\hat{\chi})E(\hat{\psi}) = E(\hat{\chi} \cdot \hat{\psi}) - \chi\psi$$

berechnen.

$$\begin{aligned} E(\hat{\chi} \cdot \hat{\psi}) &= \frac{1}{(\det M)^2} E\left(\left(m_{i(\chi),1}\hat{\lambda}_1 + m_{i(\chi),2}\hat{\lambda}_2 + m_{i(\chi),3}\hat{\lambda}_3\right) \cdot \left(m_{i(\psi),1}\hat{\lambda}_1 + m_{i(\psi),2}\hat{\lambda}_2 + m_{i(\psi),3}\hat{\lambda}_3\right)\right) \\ &= \frac{1}{(\det M)^2} \sum_{k=1}^3 \sum_{l=1}^3 m_{i(\chi),k} m_{i(\psi),l} E(\hat{\lambda}_k \cdot \hat{\lambda}_l) \end{aligned}$$

Hierbei ist

$$E(\hat{\lambda}_k \hat{\lambda}_l) = \begin{cases} E(\hat{\lambda}_k) \cdot E(\hat{\lambda}_l) = \lambda_k \lambda_l & \text{für } k \neq l \text{ (siehe Kapitel 4.2)} \\ \text{Var}(\hat{\lambda}_k) + \left(E(\hat{\lambda}_k)\right)^2 & \text{für } k = l, \text{ wegen} \\ \text{Var}(X) = E(X^2) - (E(X))^2 & \end{cases} .$$

Unter der Annahme $N_1, N_2, N_3 \approx \frac{1}{3}N$ ergibt sich für die Varianz von $\hat{\beta}$ ein Ausdruck der Form

$$\text{Var}(\hat{\beta}) = \frac{A(\alpha, \gamma, \delta, P^N, P^Y, N)}{N \cdot (\det M)^2},$$

wobei A nur von den Populationsparametern, den Forced Answer-Wahrscheinlichkeiten und dem Befragungsumfang abhängt.

4.1.2 Varianzen der NCD

Die Verfahrensmatrix ist

$$M = \begin{pmatrix} 1 - P_1^N & P_1^Y \\ 1 - P_2^N & P_2^Y \end{pmatrix} .$$

Die allgemeine Form der Adjunkten für (2×2) -Matrizen ist durch

$$\text{adj} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

gegeben und wir erhalten

$$M^{-1} = \frac{1}{\det M} \text{adj}(M) = \frac{1}{\det M} \begin{pmatrix} P_2^Y & -P_1^Y \\ -(1 - P_2^N) & 1 - P_1^N \end{pmatrix}$$

als Inverse der Verfahrensmatrix.

$$\lambda_i = \alpha (1 - P_i^N) + \gamma P_i^Y$$

Varianz von $\hat{\alpha}$

Nun können wir $\hat{\alpha}$ anders darstellen:

$$\hat{\alpha} = \frac{1}{\det M} \left(P_2^Y \hat{\lambda}_1 - P_1^Y \hat{\lambda}_2 \right)$$

Die bei der TCD durchgeführte Rechnung führt man analog durch:

$$\begin{aligned} \text{Var}(\hat{\alpha}) &= \frac{1}{(\det M)^2} \left((P_2^Y)^2 \frac{\lambda_1 (1 - \lambda_1)}{N_1} + (P_1^Y)^2 \frac{\lambda_2 (1 - \lambda_2)}{N_2} \right) \\ &\approx \frac{2}{N (\det M)^2} \left((P_2^Y)^2 \lambda_1 (1 - \lambda_1) + (P_1^Y)^2 \lambda_2 (1 - \lambda_2) \right) \end{aligned}$$

Auch hier wird für die letzte Abschätzung angenommen, dass die RRT-Gruppen in etwa gleich groß sind, also $N_1, N_2 \approx \frac{1}{2}N$.

Varianz von $\hat{\gamma}$

Es gilt

$$\hat{\gamma} = \frac{1}{\det M} \left(-(1 - P_2^N) \hat{\lambda}_1 + (1 - P_1^N) \hat{\lambda}_2 \right)$$

und wir erhalten

$$\begin{aligned} \text{Var}(\hat{\gamma}) &= \frac{1}{(\det M)^2} \left((1 - P_2^N)^2 \frac{\lambda_1 (1 - \lambda_1)}{N_1} + (1 - P_1^N)^2 \frac{\lambda_2 (1 - \lambda_2)}{N_2} \right) \\ &\approx \frac{2}{N (\det M)^2} \left((1 - P_2^N)^2 \lambda_1 (1 - \lambda_1) + (1 - P_1^N)^2 \lambda_2 (1 - \lambda_2) \right). \end{aligned}$$

Varianz von $\hat{\beta}$

Wie bei der TCD erhalten wir einen Ausdruck der Form

$$\text{Var}(\hat{\beta}) = \frac{1}{N \cdot (\det M)^2} A(\alpha, \gamma, P^N, P^Y, N).$$

4.1.3 Varianzen der YCD

Die Verfahrensmatrix ist durch

$$M = \begin{pmatrix} P_1^N & 1 - P_1^Y \\ P_2^N & 1 - P_2^Y \end{pmatrix}$$

gegeben. Unter Verwendung obiger Form der Adjunkten für (2×2) -Matrizen erhalten wir

$$M^{-1} = \frac{1}{\det M} \text{adj}(M) = \frac{1}{\det M} \begin{pmatrix} 1 - P_2^Y & -(1 - P_1^Y) \\ -P_2^N & P_1^N \end{pmatrix}$$

als Inverse der Verfahrensmatrix. Weiter haben wir

$$1 - \lambda_i = \alpha P_i^N + \gamma (1 - P_i^Y)$$

als Wahrscheinlichkeit einer „Nein“-Antwort. Für diesen Schätzer gelten die selben Bemerkungen wie für den Schätzer der „Ja“-Antworten und seine Varianz berechnet sich ebenso.

Varianz von $\hat{\alpha}$

$$\hat{\alpha} = \frac{1}{\det M} \left((1 - P_2^Y) (1 - \hat{\lambda}_1) - (1 - P_1^Y) (1 - \hat{\lambda}_2) \right)$$

Wie oben rechnet man:

$$\begin{aligned} \text{Var}(\hat{\alpha}) &= \frac{1}{(\det M)^2} \left((1 - P_2^Y)^2 \frac{\lambda_1 (1 - \lambda_1)}{N_1} + (1 - P_1^Y)^2 \frac{\lambda_2 (1 - \lambda_2)}{N_2} \right) \\ &\approx \frac{2}{N (\det M)^2} \left((1 - P_2^Y)^2 \lambda_1 (1 - \lambda_1) + (1 - P_1^Y)^2 \lambda_2 (1 - \lambda_2) \right) \end{aligned}$$

Varianz von $\hat{\gamma}$

Analog erhalten wir:

$$\hat{\gamma} = \frac{1}{\det M} \left(-P_2^N \left(1 - \hat{\lambda}_1 \right) + P_1^N \left(1 - \hat{\lambda}_2 \right) \right)$$

$$\text{Var}(\hat{\gamma}) = \frac{1}{(\det M)^2} \left((P_2^N)^2 \frac{\lambda_1 (1 - \lambda_1)}{N_1} + (P_1^N)^2 \frac{\lambda_2 (1 - \lambda_2)}{N_2} \right)$$

$$\approx \frac{2}{N (\det M)^2} \left((P_2^N)^2 \lambda_1 (1 - \lambda_1) + (P_1^N)^2 \lambda_2 (1 - \lambda_2) \right)$$

Varianz von $\hat{\delta}$

Hier erhalten wir ebenfalls die Form

$$\text{Var}(\hat{\delta}) = \frac{1}{N \cdot (\det M)^2} A(\alpha, \gamma, P^N, P^Y, N).$$

4.1.4 Überlegungen zur Optimierung des Verfahrens

Für einen Populationsparameter $\chi \in \{\alpha, \beta, \gamma, \delta\}$ bestimmen wir den mittleren quadratischen Fehler $\text{MSE}(\hat{\chi})$ von $\hat{\chi}$ als $\text{MSE}(\hat{\chi}) := \text{Var}(\hat{\chi}) + (\text{Bias}(\hat{\chi}))^2$. Um diesen zu minimieren, genügt es, $\text{Var}(\hat{\chi})$ zu minimieren, da die Schätzer (wie wir im nächsten Abschnitt sehen werden) erwartungstreu sind. Will man dazu die oben hergeleiteten Formen verwenden, ergeben sich zwei Probleme. Zum einen ist der Ausdruck im Zähler sehr komplex. Zum Anderen gehen die unbekanntenen Populationsparameter ein, auf die der Experimentator keinen Einfluss hat. Die Faustformeln

$$\text{std}(\text{Bias}(\hat{\chi}))_{\text{TCD}} \approx \frac{1}{\sqrt{|\det M| N}}$$

$$\text{std}(\text{Bias}(\hat{\chi}))_{\text{DCD}} \approx \frac{1}{|\det M| \sqrt{N}}$$

für die Standardabweichungen der Schätzfehler $\text{Bias}(\hat{\chi})$, $\chi \in \{\alpha, \beta, \gamma, \delta\}$ verschaffen hierbei Abhilfe. Diese sind zwar sehr grobe Abschätzungen für die Proportionalität, die je nach Forced Answer-Wahrscheinlichkeiten und Populationsparametern Abweichungen von $\pm 30\%$ in der Proportionalitäts-, Konstante“ haben können, sie bieten zur Optimierung der Schätzer jedoch eine für

den praktischen Einsatz ausreichende Entkopplung der Forced Answer-Wahrscheinlichkeiten von den Populationsparametern.

In Tabelle 4.1, 4.2 und B.2 wird der Zusammenhang zwischen Standardabweichung der Schätzfehler und der Determinante der Verfahrensmatrix dargestellt. Dazu wurden 1.000 zufällige Populationen erzeugt, zu jeder wurden 1.000 Monte-Carlo-Simulationen mit Stichprobenumfang $N = 1.500$ durchgeführt. Dann wurden zu jeder dieser Populationen die Standardabweichungen der Schätzfehler und anschließend die 10%-, 50%- und 90%-Quantile ermittelt. Der Zusammenhang mit $\frac{1}{\sqrt{N}}$ ist in Abbildung 4.4, 4.9 und 4.12 dargestellt.

Tabelle 4.1: Zusammenhang der Standardabweichungen der Schätzfehler bei der TCD mit der Determinante der Verfahrensmatrix. Die Determinante ist exakt, die Quantile sind um den Faktor 1.000 skaliert und gerundet.

χ	det M	std (Bias ($\hat{\chi}$))			std (Bias ($\hat{\chi}$)) $\cdot \sqrt{ \det M }$		
		10%	50%	90%	10%	50%	90%
α	0,36	38,97	49,91	62,22	23,38	29,95	37,33
	0,20	51,53	69,37	88,71	23,04	31,02	39,67
	0,11	65,94	92,67	120,70	21,87	30,74	40,03
	0,06	87,39	123,48	164,03	21,41	30,25	40,18
β	0,36	25,86	32,19	34,94	15,52	19,31	20,96
	0,20	35,48	42,86	46,52	15,87	19,17	20,81
	0,11	45,88	64,10	71,97	15,22	21,26	23,87
	0,06	49,27	66,91	75,77	12,07	16,39	18,56
γ	0,36	39,79	49,13	52,69	23,88	29,48	31,61
	0,20	46,68	59,30	63,47	20,88	26,52	28,38
	0,11	78,97	112,99	127,25	26,19	37,47	42,20
	0,06	89,25	125,85	147,26	21,86	30,83	36,07
δ	0,36	25,20	32,53	42,69	15,12	19,52	25,62
	0,20	34,67	45,14	60,29	15,51	20,19	26,96
	0,11	34,15	47,08	65,02	11,33	15,62	21,56
	0,06	47,72	66,76	94,58	11,69	16,35	23,17

4.2 Erwartungswerte der Schätzer

Mit obigen Überlegungen zu Lemma 3 können wir $E(\hat{\lambda}_i)$ berechnen:

Korollar 4.

$$E(\hat{\lambda}_i) = E\left(\frac{k_i}{N_i}\right) = \frac{E(k_i)}{N_i} = \frac{N_i \lambda_i}{N_i} = \lambda_i$$

Tabelle 4.2: Zusammenhang der Standardabweichungen der Schätzfehler bei der NCD mit der Determinante der Verfahrensmatrix. Die Determinante ist exakt, die Quantile sind um den Faktor 1.000 skaliert und gerundet.

χ	det M	std (Bias ($\hat{\chi}$))			std (Bias ($\hat{\chi}$)) · det M		
		10%	50%	90%	10%	50%	90%
α	0,54	16,33	21,96	23,97	8,82	11,86	12,94
	0,40	18,35	24,90	27,56	7,34	9,96	11,02
	0,22	30,87	40,29	43,77	6,79	8,86	9,63
	0,16	43,89	60,35	65,78	7,02	9,66	10,52
β	0,54	22,80	26,70	28,22	12,31	14,42	15,24
	0,40	27,00	31,92	33,75	10,80	12,77	13,50
	0,22	41,68	55,61	59,65	9,17	12,23	13,12
	0,16	44,67	61,29	66,25	7,15	9,81	10,60
γ	0,54	32,78	40,04	42,74	17,70	21,62	23,08
	0,40	38,23	47,88	51,44	15,29	19,15	20,58
	0,22	67,05	91,42	98,52	14,75	20,11	21,67
	0,16	84,01	117,47	128,14	13,44	18,79	20,50

Um die Erwartungstreue aller Teilschätzer zu zeigen könnten wir genau so vorgehen, wie bei der Berechnung der Varianzen. Stattdessen beweisen wir die ersten drei Spezialfälle ($1 \leq n_0 \leq 3$) folgender Vermutung, da sie uns die Erwartungstreue sämtlicher Teilschätzer (also auch der Schätzer der Randlösungen) für alle drei Verfahren ebenfalls mitliefern.

Vermutung 1. Seien $n, n_0 \in \mathbb{N}$ mit $n \geq n_0 > 0$ und seien $\chi_1, \dots, \chi_{n_0+1} \in \mathbb{R}$. Weiter seien $v_1, \dots, v_{n_0} \in \mathbb{R}^n$ so gegeben, dass $M := (v_1, \dots, v_{n_0})$ vollen

Spaltenrang n_0 hat und seien $\begin{pmatrix} \hat{\lambda}_1 \\ \vdots \\ \hat{\lambda}_{n_0} \end{pmatrix} := \hat{\lambda}$ Schätzer mit $E(\hat{\lambda}) = \sum_{i=1}^{n_0} v_i \chi_i$.

Dann sind die Schätzer

$$\begin{pmatrix} \hat{\chi}_1 \\ \vdots \\ \hat{\chi}_{n_0} \end{pmatrix} := M^+ \hat{\lambda}$$

erwartungstreu, d.h. es gilt $E(\hat{\chi}_i) = \chi_i$ für $1 \leq i \leq n_0$.

Gilt zudem $\sum_{i=1}^{n_0+1} \chi_i = 1$, dann ist offensichtlich auch $\hat{\chi}_{n_0+1} := 1 - \sum_{i=1}^{n_0} \hat{\chi}_i$ erwartungstreu.

Wir werden nun die Vermutung für die Fälle $n_0 = 1$, $n_0 = 2$ und $n_0 = 3$ zeigen. Wie üblich sei das reelle Standardskalarprodukt mit $\langle \cdot, \cdot \rangle$ bezeichnet.

Lemma 4. *Die Vermutung gilt im Fall $n_0 = 1$.*

Beweis. Es ist $M^+ = (M^T M)^{-1} M^T = (v_1^T v_1)^{-1} v_1^T = \frac{1}{\langle v_1, v_1 \rangle} v_1^T$. Wir erhalten $E(\hat{\chi}_1) = M^+ E(\hat{\lambda}) = \frac{1}{\langle v_1, v_1 \rangle} \cdot \chi_1 \cdot v_1^T v_1 = \chi_1 \frac{\langle v_1, v_1 \rangle}{\langle v_1, v_1 \rangle} = \chi_1$.

Lemma 5. *Die Vermutung gilt für $n_0 = 2$.*

Beweis. Es ist

$$\begin{aligned} M^+ &= (M^T M)^{-1} M^T = \left(\begin{pmatrix} v_1^T \\ v_2^T \end{pmatrix} (v_1 \quad v_2) \right)^{-1} \begin{pmatrix} v_1^T \\ v_2^T \end{pmatrix} \\ &= \begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_1, v_2 \rangle & \langle v_2, v_2 \rangle \end{pmatrix}^{-1} \begin{pmatrix} v_1^T \\ v_2^T \end{pmatrix} \\ &= \frac{1}{D} \begin{pmatrix} \langle v_2, v_2 \rangle & -\langle v_1, v_2 \rangle \\ -\langle v_1, v_2 \rangle & \langle v_1, v_1 \rangle \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \end{pmatrix} \\ &= \frac{1}{D} \begin{pmatrix} \langle v_2, v_2 \rangle v_1^T - \langle v_1, v_2 \rangle v_2^T \\ -\langle v_1, v_2 \rangle v_1^T + \langle v_1, v_1 \rangle v_2^T \end{pmatrix} \end{aligned}$$

mit

$$D = \langle v_1, v_1 \rangle \langle v_2, v_2 \rangle - \langle v_1, v_2 \rangle^2.$$

Wir erhalten

$$\begin{aligned} E \begin{pmatrix} \hat{\chi}_1 \\ \hat{\chi}_2 \end{pmatrix} &= M^+ E(\hat{\lambda}) = M^+ (\chi_1 v_1 + \chi_2 v_2) \\ &= \frac{1}{D} \begin{pmatrix} \chi_1 (\langle v_2, v_2 \rangle \langle v_1, v_1 \rangle - \langle v_1, v_2 \rangle \langle v_2, v_1 \rangle) \\ \quad + \chi_2 (\langle v_2, v_2 \rangle \langle v_1, v_2 \rangle - \langle v_1, v_2 \rangle \langle v_2, v_2 \rangle) \\ \chi_1 (-\langle v_1, v_2 \rangle \langle v_1, v_1 \rangle + \langle v_1, v_1 \rangle \langle v_2, v_1 \rangle) \\ \quad + \chi_2 (-\langle v_1, v_2 \rangle \langle v_1, v_2 \rangle + \langle v_1, v_1 \rangle \langle v_2, v_2 \rangle) \end{pmatrix} \\ &= \frac{1}{D} \begin{pmatrix} \chi_1 \cdot D + \chi_2 \cdot 0 \\ \chi_1 \cdot 0 + \chi_2 \cdot D \end{pmatrix} = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \end{aligned}$$

Lemma 6. Die Vermutung gilt im Fall $n_0 = 3$.

Beweis. Es ist

$$\begin{aligned}
 (M^T M)^{-1} &= \left(\begin{pmatrix} v_1^T \\ v_2^T \\ v_3^T \end{pmatrix} (v_1 \ v_2 \ v_3) \right)^{-1} \\
 &= \begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \langle v_1, v_3 \rangle \\ \langle v_1, v_2 \rangle & \langle v_2, v_2 \rangle & \langle v_2, v_3 \rangle \\ \langle v_1, v_3 \rangle & \langle v_2, v_3 \rangle & \langle v_3, v_3 \rangle \end{pmatrix}^{-1} \\
 &= \frac{1}{D} \begin{pmatrix} \langle v_2, v_2 \rangle \langle v_3, v_3 \rangle & \langle v_1, v_3 \rangle \langle v_2, v_3 \rangle & \langle v_1, v_2 \rangle \langle v_2, v_3 \rangle \\ -\langle v_2, v_3 \rangle^2 & -\langle v_1, v_2 \rangle \langle v_3, v_3 \rangle & -\langle v_1, v_3 \rangle \langle v_2, v_2 \rangle \\ \langle v_2, v_3 \rangle \langle v_1, v_3 \rangle & \langle v_1, v_1 \rangle \langle v_3, v_3 \rangle & \langle v_1, v_2 \rangle \langle v_1, v_3 \rangle \\ -\langle v_1, v_2 \rangle \langle v_3, v_3 \rangle & -\langle v_1, v_3 \rangle^2 & -\langle v_1, v_1 \rangle \langle v_2, v_3 \rangle \\ \langle v_1, v_2 \rangle \langle v_2, v_3 \rangle & \langle v_1, v_2 \rangle \langle v_1, v_3 \rangle & \langle v_1, v_1 \rangle \langle v_2, v_2 \rangle \\ -\langle v_1, v_3 \rangle \langle v_2, v_2 \rangle & -\langle v_1, v_1 \rangle \langle v_2, v_3 \rangle & -\langle v_1, v_2 \rangle^2 \end{pmatrix}
 \end{aligned}$$

mit

$$\begin{aligned}
 D &= \langle v_1, v_1 \rangle \langle v_2, v_2 \rangle \langle v_3, v_3 \rangle + 2 \langle v_1, v_2 \rangle \langle v_1, v_3 \rangle \langle v_2, v_3 \rangle \\
 &\quad - \langle v_1, v_3 \rangle^2 \langle v_2, v_2 \rangle - \langle v_2, v_3 \rangle^2 \langle v_1, v_1 \rangle - \langle v_1, v_2 \rangle^2 \langle v_3, v_3 \rangle.
 \end{aligned}$$

Somit erhalten wir

$$\begin{aligned}
 M^+ &= (M^T M)^{-1} \begin{pmatrix} v_1^T \\ v_2^T \\ v_3^T \end{pmatrix} \\
 &= \frac{1}{D} \begin{pmatrix} (\langle v_2, v_2 \rangle \langle v_3, v_3 \rangle - \langle v_2, v_3 \rangle^2) v_1^T \\ + (\langle v_1, v_3 \rangle \langle v_2, v_3 \rangle - \langle v_1, v_2 \rangle \langle v_3, v_3 \rangle) v_2^T \\ + (\langle v_1, v_2 \rangle \langle v_2, v_3 \rangle - \langle v_1, v_3 \rangle \langle v_2, v_2 \rangle) v_3^T \\ (\langle v_2, v_3 \rangle \langle v_1, v_3 \rangle - \langle v_1, v_2 \rangle \langle v_3, v_3 \rangle) v_1^T \\ + (\langle v_1, v_1 \rangle \langle v_3, v_3 \rangle - \langle v_1, v_3 \rangle^2) v_2^T \\ + (\langle v_1, v_2 \rangle \langle v_1, v_3 \rangle - \langle v_1, v_1 \rangle \langle v_2, v_3 \rangle) v_3^T \\ (\langle v_1, v_2 \rangle \langle v_2, v_3 \rangle - \langle v_1, v_3 \rangle \langle v_2, v_2 \rangle) v_1^T \\ + (\langle v_1, v_2 \rangle \langle v_1, v_3 \rangle - \langle v_1, v_1 \rangle \langle v_2, v_3 \rangle) v_2^T \\ + (\langle v_1, v_1 \rangle \langle v_2, v_2 \rangle - \langle v_1, v_2 \rangle^2) v_3^T \end{pmatrix}
 \end{aligned}$$

und wegen

$$\mathbb{E} \begin{pmatrix} \hat{\chi}_1 \\ \hat{\chi}_2 \\ \hat{\chi}_3 \end{pmatrix} = M^+ \mathbb{E}(\hat{\lambda}) = M^+ (\chi_1 v_1 + \chi_2 v_2 + \chi_3 v_3)$$

ergibt sich

$$\begin{aligned} \mathbb{E}(\hat{\chi}_1) &= \frac{1}{D} (\chi_1 (\langle v_2, v_2 \rangle \langle v_3, v_3 \rangle \langle v_1, v_1 \rangle - \langle v_2, v_3 \rangle^2 \langle v_1, v_1 \rangle \\ &\quad + \langle v_1, v_3 \rangle \langle v_2, v_3 \rangle \langle v_2, v_1 \rangle - \langle v_1, v_2 \rangle \langle v_3, v_3 \rangle \langle v_2, v_1 \rangle \\ &\quad + \langle v_1, v_2 \rangle \langle v_2, v_3 \rangle \langle v_3, v_1 \rangle - \langle v_1, v_3 \rangle \langle v_2, v_2 \rangle \langle v_3, v_1 \rangle) \\ &\quad + \chi_2 (\langle v_2, v_2 \rangle \langle v_3, v_3 \rangle \langle v_1, v_2 \rangle - \langle v_2, v_3 \rangle^2 \langle v_1, v_2 \rangle \\ &\quad + \langle v_1, v_3 \rangle \langle v_2, v_3 \rangle \langle v_2, v_2 \rangle - \langle v_1, v_2 \rangle \langle v_3, v_3 \rangle \langle v_2, v_2 \rangle \\ &\quad + \langle v_1, v_2 \rangle \langle v_2, v_3 \rangle \langle v_3, v_2 \rangle - \langle v_1, v_3 \rangle \langle v_2, v_2 \rangle \langle v_3, v_2 \rangle) \\ &\quad + \chi_3 (\langle v_2, v_2 \rangle \langle v_3, v_3 \rangle \langle v_1, v_3 \rangle - \langle v_2, v_3 \rangle^2 \langle v_1, v_3 \rangle \\ &\quad + \langle v_1, v_3 \rangle \langle v_2, v_3 \rangle \langle v_2, v_3 \rangle - \langle v_1, v_2 \rangle \langle v_3, v_3 \rangle \langle v_2, v_3 \rangle \\ &\quad + \langle v_1, v_2 \rangle \langle v_2, v_3 \rangle \langle v_3, v_3 \rangle - \langle v_1, v_3 \rangle \langle v_2, v_2 \rangle \langle v_3, v_3 \rangle) \\ &= \frac{1}{D} (\chi_1 \cdot D + \chi_2 \cdot 0 + \chi_3 \cdot 0) = \chi_1. \end{aligned}$$

Ebenso rechnet man $\mathbb{E}(\hat{\chi}_2) = \chi_2$ und $\mathbb{E}(\hat{\chi}_3) = \chi_3$ nach.

Aus dem Korollar (bzw. den Lemmata) folgt die Erwartungstreue aller Teilschätzer für jedes der Verfahren:

- **TCD**

Lemma 6 liefert die Erwartungstreue der nativen Schätzer, Lemma 5 die Erwartungstreue für die Schätzer der Randlösungen erster Ordnung und Lemma 4 die Erwartungstreue für die Schätzer der Randlösungen zweiter Ordnung. In den trivialen Fällen, bei denen ein Populationsparameter gleich Eins ist, ist die Erwartungstreue offensichtlich.

- **NCD/YCD**

Die Erwartungstreue der nativen Schätzer folgt aus Lemma 5, die Erwartungstreue der Schätzer für die Randlösungen erster Ordnung folgt aus Lemma 4.

4.3 Optimierung der Forced Answer-Wahrscheinlichkeiten

Beim Lösen der ML-Gleichung wurde gefordert, dass die Verfahrensmatrix vollen Zeilenrang hat, bzw. im Fall einer quadratischen Matrix invertierbar

ist. Ist ihr Zeilenrang nicht maximal, kommen, bedingt durch die Wahl der Forced Answer-Wahrscheinlichkeiten, redundante Informationen in die Schätzung, welche dann natürlich fehlschlägt. Zusätzlich sollte, wie oben angemerkt, die Determinante der Verfahrensmatrix einen möglichst großen Betrag haben, um die mittleren Schätzfehler zu minimieren. Diese rein mathematische Optimierung der Schätzer steht aber im Konflikt zu der grundsätzlichen Idee der RRT.

Da dies an den Matrizen der vorgestellten Verfahren nicht ohne Weiteres erkennbar ist, wollen wir diesen Sachverhalt am Beispiel der klassischen RRT mit einer Gruppe zeigen.

Es gelte $\beta = \delta = 0$. Mit $\gamma = 1 - \alpha$ ist die Wahrscheinlichkeit einer „Ja“-Antwort durch

$$\lambda = \alpha (P^S + P^Y) + \gamma P^Y \Leftrightarrow \lambda - P^Y = \alpha (1 - P^Y - P^N)$$

gegeben. Die Verfahrens-„Matrix“ ist also $M = (1 - P^Y - P^N)$. Dieses Ergebnis sieht man oben auch in allgemeiner Form bei den Randfällen $\beta = \delta = 0$ der TCD, NCD und YCD. Die Maximierung der Determinante entspricht offensichtlich dem Auslassen der Forced Answer-Instruktionen, was jedoch genau einer direkten Befragung entspricht.

Dieses Prinzip lässt sich auf die hier vorgestellten Verfahren verallgemeinern: Die mathematische Optimierung ergibt einen Randfall, der genau dem Auslassen der Forced Answer-Instruktion in einem Teil der Gruppen und dem Erzwingen der Forced Answer in den restlichen Gruppen gleich kommt. Dieses Optimum entspricht damit einer praktisch entarteten Lösung. Die Maximierung der Determinanten der Verfahrensmatrizen wird durch Minimierung der ε_i ($0 \leq \varepsilon_i \leq 1$) in den folgenden Forced Answer-Wahrscheinlichkeiten erreicht:

- Für die TCD:

$$P_{\text{TCD}}^Y = \begin{pmatrix} 1 - \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}, \quad P_{\text{TCD}}^N = \begin{pmatrix} \varepsilon_4 \\ 1 - \varepsilon_5 \\ \varepsilon_6 \end{pmatrix}$$

- Für die NCD:

$$P_{\text{NCD}}^Y = \begin{pmatrix} 1 - \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}, \quad P_{\text{NCD}}^N = \begin{pmatrix} \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}$$

- Für die YCD:

$$P_{\text{YCD}}^Y = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}, \quad P_{\text{YCD}}^N = \begin{pmatrix} 1 - \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}$$

4.4 Darstellung der simulierten Daten

Da die Verfahren drei bzw. vier Parameter schätzen, können wir die Schätzfehler bzw. deren Standardabweichungen nicht unmittelbar darstellen. Eine Koordinatentransformation erlaubt es uns, den kompletten Wertebereich von drei Populationsparametern auf ein Dreieck im \mathbb{R}^2 zu legen, so dass wir in einem Schaubild entlang der dritten Achse die Schätzfehler und ihre Standardabweichungen abtragen können. Dieses Dreieck repräsentiert dann die Gesamtheit aller Populationen. Jeder Punkt darin repräsentiert eindeutig eine Population. In diesem Abschnitt soll hauptsächlich die Interpretation dieser Darstellung erläutert werden, eine detaillierte Herleitung dieser Transformation ist in Anhang A zu finden. Um eine Analyse der Schätzfehler über alle möglichen Populationsparameter darzustellen, wurden zu jedem Punkt in diesem Dreieck (also zu jeder möglichen Population) eine Serie von Monte-Carlo-Simulationen durchgeführt und die daraus gewonnenen Ergebnisse auf der dritten Achse im Koordinatensystem aufgetragen.

Bei der DCD müssen drei Parameter variiert werden. Durch oben genannte Transformation des Wertebereichs lässt sich eine zu untersuchende Größe (etwa der Schätzfehler Bias ($\hat{\alpha}$)) komplett in einem einzigen Schaubild darstellen.

Die Interpretation soll exemplarisch an der NCD gezeigt werden: Die interessierenden Populationsparameter sind α , β und γ . Zur Darstellung benutzen wir ein gleichseitiges Dreieck mit Eckpunkten A , B und C . Jeder Punkt P in diesem Dreieck lässt sich eindeutig als $P = \alpha A + \beta B + \gamma C$ darstellen. Diese Darstellung ist nur eindeutig, da $\alpha, \beta, \gamma \in [0, 1]$ gilt. Man spricht hier auch von einer Konvexkombination von A , B und C mit Koeffizienten α , β und γ . Je näher P also an einem der Eckpunkte liegt, desto größer ist der zur Ecke zugehörige Koeffizient. Zum Beispiel bedeutet $P = A$ den Randfall, dass P die Population mit $\alpha = 1$ repräsentiert. Liegt P etwa auf der Seite \overline{AB} , so repräsentiert P eine Population mit $\alpha + \beta \equiv 1 \Leftrightarrow \gamma \equiv 0$. Punkte auf einer zu \overline{AB} parallelen Geraden stehen für Populationen, in denen $\alpha + \beta \equiv 1 - \gamma \equiv \text{const} \Leftrightarrow \gamma \equiv \text{const}$ gilt. Im Schwerpunkt des Dreiecks ist die Population mit $\alpha = \beta = \gamma = \frac{1}{3}$ zu finden.

Die TCD lässt sich auf diese Weise nicht komplett in einem Schaubild darstellen, da bei ihr vier Parameter geschätzt werden. Zur Darstellung werden nach obigen Verfahren Graphen zu einigen ausgewählten Werten des Parameters δ betrachtet, d.h. der Populationsparameter δ wird fixiert. Die oben dargestellte Interpretation zur Lage eines Punktes im Dreieck lässt sich hierauf mit der zusätzlichen Einschränkung $\alpha + \beta + \gamma = 1 - \delta$ übertragen. In solch einem Dreieck repräsentiert dann etwa der Eckpunkt A genau die Population, in der $\alpha = 1 - \delta, \beta = \gamma = 0$ gilt, wobei δ fixiert ist.

4.5 Analyse der TCD

Wie bereits am Anfang dieses Kapitels angedeutet, wollen wir nun die kompletten Verfahren, also die Zusammensetzung aller Teilschätzer, mit Hilfe von Monte-Carlo-Simulationen untersuchen.

Die hier vorgestellten Daten basieren auf der Auswertung von 1.000 Monte-Carlo-Simulationen zu jeder möglichen Population. Als Forced Answer-Wahrscheinlichkeiten haben wir

$$P^Y = \begin{pmatrix} 0,7 \\ 0,1 \\ 0,1 \end{pmatrix}, \quad P^N = \begin{pmatrix} 0,1 \\ 0,7 \\ 0,1 \end{pmatrix}$$

gewählt. Zunächst sollen an einem Beispiel die grundlegenden Eigenschaften der TCD erläutert werden, bevor wir die Parameter systematisch variieren.

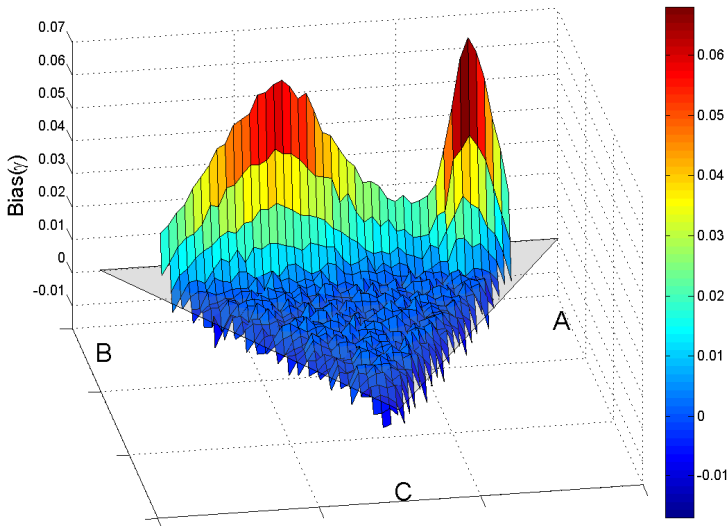


Abbildung 4.1: Schätzfehler für γ der TCD ($N = 1.500$, $\delta = 0,20$, 1.000 Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

Abbildung 4.1 zeigt den mittleren Schätzfehler $\text{Bias}(\hat{\gamma})$ für $\delta = 0,20$. Es ist deutlich zu erkennen, dass die TCD nahe den Randfällen nicht erwartungstreu

ist. Dies ist in den Positivitätsbedingungen an die Parameter begründet: Man erkennt hier sehr deutlich, dass nahe den Randfällen, in denen eine Population mit $\gamma = 0$ vorliegt, also an der Seite AB, der Schätzer $\hat{\gamma}$ stark ins Positive verzerrt ist. Der verhältnismäßig hohe Wert des Schätzfehlers nahe Punkt A auf Seite \overline{AB} lässt sich wie folgt erklären: Eine „Nein“-Antwort kann sowohl dem Populationsanteil γ (auf den sich diese Abbildung bezieht) als auch dem Populationsanteil β zugeordnet werden. Ein Punkt in der Nähe von A repräsentiert eine Population mit sehr kleinen Anteilen β und γ , wodurch der Schätzer $\hat{\gamma}$ bei der jeweiligen Simulationsserie häufig fälschlicherweise eine Schätzung aus den Randfällen $\beta = 0$ oder $\beta = \gamma = 0$ liefert. Der Randfall $\beta = 0$ bedeutet dann immer eine Überschätzung von $\hat{\gamma}$. Liefert die native ML-Lösung einen Wert $\gamma < 0$, hat (in fast allen Fällen) der Randfall $\gamma = 0$ die höchste Likelihood unter allen Randfällen. Deswegen kann eine native (bzw. eine „echte“) ML-Lösung mit $\gamma > 0$ im Mittel nicht kompensiert werden, wie es etwa in der Mitte des Dreiecks, wo alle Parameter relativ weit von einer Randlösung entfernt sind, möglich ist.

Dies spiegelt sich auch in der Standardabweichung des Schätzfehlers wieder, wie man deutlich in Abbildung 4.2 erkennen kann. Nahe den Randfällen wird die Standardabweichung im Allgemeinen geringer. Zur Mitte des Dreiecks hin wird die Standardabweichung größer, weil die Schätzung nicht durch Randfälle künstlich verzerrt wird und die TCD Spielraum für große Über- und Unterschätzungen hat. Analog zum Schätzfehler (Abb. 4.1) ist auch die Standardabweichung nahe Punkt A sehr hoch. An dieser Stelle führen die bereits diskutierten falschen Randlösungen $\beta = 0$ und die korrekte Randlösung $\gamma = 0$ zu dieser hohen Standardabweichung.

Für den Fall $\delta = 0$ kann man die Verzerrung des Schätzers nahe den Randfällen in den Abbildungen B.1 und B.2 gut erkennen. Die augenscheinliche Symmetrie der Fehler verdeutlicht, dass „Ja“-Antworten sowohl fälschlich auf den Populationsanteil α als auch δ zurückgeführt werden können.

Außerhalb der Randfälle kann man die TCD durchaus als erwartungstreu ansehen. Bei Abbildung 4.3 wird dies besonders deutlich. Sie zeigt die gleichen Schätzfehler, die auch in Abbildung 4.1 zu sehen sind, allerdings wurden hier nur Werte für Populationen mit $\alpha, \beta, \gamma > 0,05$ dargestellt. Auch hier sieht man noch deutlich die Auswirkungen der Randfälle $\gamma = 0$, wobei der größte Fehler bei $\gamma \approx 0,05$ mit Bias ($\hat{\gamma}$) $\approx 0,01$ für die Schätzung eines Populationsanteils noch in einem sehr akzeptablen Rahmen liegt.

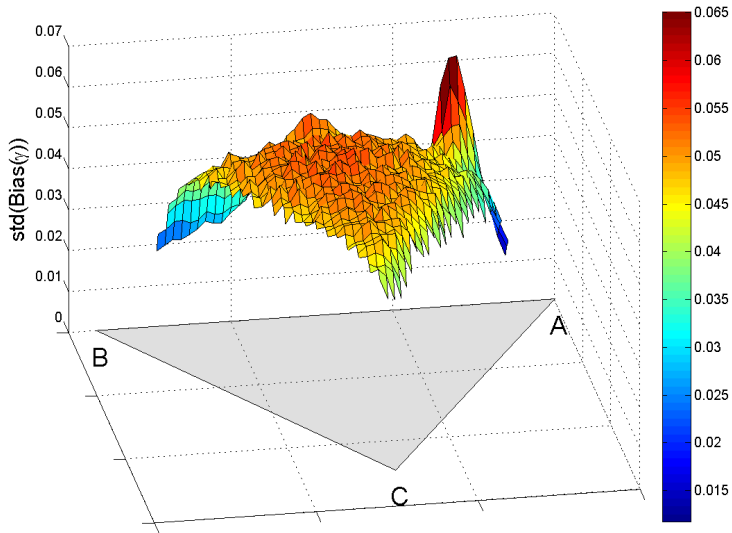


Abbildung 4.2: Standardabweichung des Schätzfehlers für γ der TCD ($N = 1.500$, $\delta = 0,20, 1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

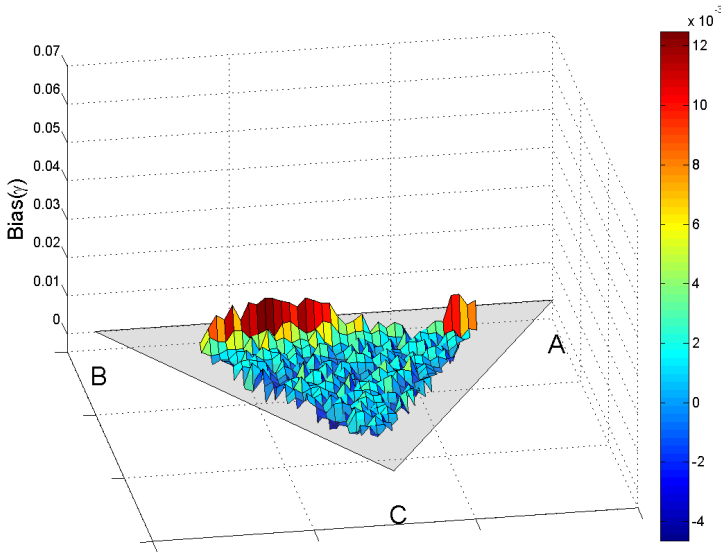


Abbildung 4.3: Schätzfehler für γ der TCD ($N = 1.500$, $\delta = 0,20$, 1.000 Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

4.5.1 Mittlere Fehler

Zu den simulierten Daten sollen nun weiter die mittleren Schätzfehler und somit die Erwartungstreu des Verfahrens dargestellt werden. Wir haben zwar schon analytisch gezeigt, dass die Teilschätzer jeweils erwartungstreu sind, wir haben jedoch auch gesehen, dass die Zusammensetzung der Teilschätzer zur TCD nahe den Randfällen verzerrt ist. Dazu wurde in der Simulationsreihe für jeden Parameter der Mittelwert über alle Schätzfehler gebildet, einmal mit Einbeziehung aller Populationen, und einmal unter Vernachlässigung der Populationen nahe den Randfällen ($\alpha, \beta, \gamma, \delta < 0,05$). So finden sich z.B. die Gesamtmittelwerte der in Abbildung 4.1 und 4.3 dargestellten Fehler in Tabelle 4.3 in Zeile 5, Spalte 3 bzw. Spalte 4 wieder. Die entsprechenden Werte für $\text{Bias}(\hat{\beta})$ und $\text{Bias}(\hat{\delta})$ sind in Tabelle B.1 zu finden.

Tabelle 4.3: Mittlere Schätzfehler der TCD zu den simulierten Daten, skaliert um den Faktor 10^6 .

δ	χ	Bias ($\hat{\chi}$)	
		mit Randfällen	ohne Randfälle
0,00	α	-112.393	-
	γ	20.690	-
0,05	α	14.190	1.839
	γ	3.182	2.733
0,10	α	23.228	7.248
	γ	12.044	7.056
0,15	α	18.439	4.265
	γ	25.918	3.959
0,20	α	301	-770
	γ	42.919	10.060
0,35	α	-57.091	-14.791
	γ	61.993	14.216
0,50	α	-40.546	-33.231
	γ	41.423	24.200
0,65	α	-55.749	-19.506
	γ	69.690	15.936
0,80	α	-27.722	-16.765
	γ	35.223	8.020
0,95	α	14.861	-
	γ	-3.707	-

Man kann erkennen, dass die Schätzfehler deutlich kleiner werden, wenn man einen Großteil der Randfälle ausschließt. Wie oben analytisch gezeigt, sehen wir also, dass die TCD erwartungstreu ist, wenn nur die innere Lösung angenommen wird. Die Regelmäßigkeiten in den Vorzeichen der Schätzfehler zeigen, dass noch ein systematischer Fehler vorliegt. Dies liegt darin begründet, dass trotz der Einschränkung $\alpha, \beta, \gamma, \delta > 0,05$ bei Weitem nicht alle Randfälle ausgeschlossen werden. Abbildung 4.3 und Tabelle 4.3, zeigen jedoch, dass die Restfehler so gering sind, dass man die TCD für ihren Einsatzzweck innerhalb dieser Einschränkungen als erwartungstreu ansehen kann.

4.5.2 Standardabweichung der Schätzfehler

Zur Darstellung, wie sich die Standardabweichung der Schätzfehler mit steigendem Stichprobenumfang entwickelt, wurden 1.000 Populationen erzeugt.

Zu jeder Population wurde an jeder Stützstelle des Stichprobenumfangs N die Standardabweichung des Schätzfehlers in 1.000 Monte-Carlo-Simulationen ermittelt und als Boxplot gegen die Stützstellen von N aufgetragen. Durch die Mediane wurde eine Regressionskurve der Form $\sqrt{\frac{a}{N}} + b$ gelegt, welche die Proportionalität der Standardabweichung zu $\frac{1}{\sqrt{N}}$ illustriert. Die Resultate bezüglich den Populationsparametern α und γ sind in den Abbildungen 4.4 und 4.5 zu finden, bzgl. β und δ in den Abbildungen B.3 und B.4 im Anhang.



Abbildung 4.4: Entwicklung der Standardabweichung des Schätzfehlers von α bei der TCD gemäß 4.5.2. Die Box kennzeichnet die Grenzen des oberen und des unteren Quartils. Der rote Strich innerhalb der Box markiert den Median der Datenpunkte. Die Endpunkte der Whiskers geben die Lage des größten bzw. kleinsten Datenpunkts an, der maximal um den 1,5-fachen Interquartilsabstand oberhalb bzw. unterhalb der Grenze des oberen bzw. unteren Quartils liegt.

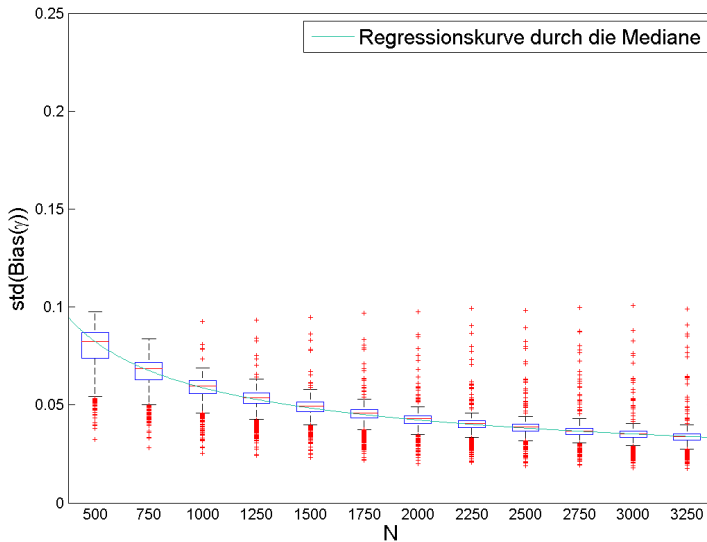


Abbildung 4.5: Entwicklung der Standardabweichung des Schätzfehlers von γ bei der TCD gemäß 4.5.2. Erläuterungen zur Darstellung siehe Abbildung 4.4.

4.6 Analyse der NCD

Prinzipiell lassen sich alle bei der TCD gewonnenen Ergebnisse übertragen. Die einzige Ausnahme hiervon bildet der Fall, dass der Parameter δ fälschlicherweise als Null angenommen wurde. Als Forced Answer-Wahrscheinlichkeiten wurden

$$P^Y = \begin{pmatrix} 0,7 \\ 0,1 \end{pmatrix}, \quad P^N = \begin{pmatrix} 0,1 \\ 0,1 \end{pmatrix}$$

benutzt.

4.6.1 Auswirkung $\delta \neq 0$

Ist der Populationsparameter $\delta \neq 0$, so ist die NCD verzerrt. Die Berechnung der Erwartungswerte funktioniert genau so wie oben durchgeführt, mit dem

einzigsten Unterschied, dass sich die Wahrscheinlichkeiten von „Ja“-Antworten durch

$$\lambda_i = \alpha (1 - P_i^N) + \gamma P_i^Y + \delta$$

berechnen. Für die native Lösung erhalten wir:

Erwartungswert von $\hat{\alpha}$

$$\begin{aligned}\hat{\alpha} &= \frac{1}{\det M} \left(P_2^Y \hat{\lambda}_1 - P_1^Y \hat{\lambda}_2 \right) \\ E(\hat{\alpha}) &= \frac{1}{\det M} \left(P_2^Y (\alpha (1 - P_1^N) + \gamma P_1^Y + \delta) \right. \\ &\quad \left. - P_1^Y (\alpha (1 - P_2^N) + \gamma P_2^Y + \delta) \right) \\ &= \frac{1}{\det M} (\alpha \det M + \gamma \cdot 0 + \delta (P_2^Y - P_1^Y)) \\ &= \alpha + \delta \frac{P_2^Y - P_1^Y}{\det M}\end{aligned}$$

Erwartungswert von $\hat{\gamma}$

$$\begin{aligned}\hat{\gamma} &= \frac{1}{\det M} \left(- (1 - P_2^N) \hat{\lambda}_1 + (1 - P_1^N) \hat{\lambda}_2 \right) \\ E(\hat{\gamma}) &= \frac{1}{\det M} (\alpha \cdot 0 + \gamma \det M + \delta (-1 + P_2^N + 1 - P_1^N)) \\ &= \gamma + \delta \frac{P_2^N - P_1^N}{\det M}\end{aligned}$$

Erwartungswert von $\hat{\beta}$

$$\begin{aligned}\hat{\beta} &= 1 - \hat{\alpha} - \hat{\gamma} \\ E(\hat{\beta}) &= E(1 - \hat{\alpha} - \hat{\gamma}) = 1 - \alpha - \delta \frac{P_2^Y - P_1^Y}{\det M} - \gamma - \delta \frac{P_2^N - P_1^N}{\det M} \\ &= \beta + \delta \left(1 - \frac{P_2^Y - P_1^Y + P_2^N - P_1^N}{\det M} \right)\end{aligned}$$

In den optimalen Forced Answer-Wahrscheinlichkeiten sind P_1^N und P_2^N beide relativ klein bzw. etwa gleich groß. Setzt man also $P_1^N = P_2^N$ bei der Umfrage an, so erhält man einen Schätzer $\hat{\gamma}$, der unabhängig von δ erwartungstreu ist. Simulationen dazu sind in den Abbildungen 4.6 und B.5 zu sehen. Mit den hier benutzten Forced Answer-Wahrscheinlichkeiten gilt $\det M = -\frac{54}{100}$, wir erhalten also $\text{Bias}(\hat{\alpha}) = \frac{10}{9} \delta$ und $\text{Bias}(\hat{\beta}) = -\frac{1}{9} \delta$. Nach Abschneiden der Randfälle (wie bei der TCD) in Abbildung 4.7 erhalten wir $\text{Bias}(\hat{\gamma}) = 0,000070$

als mittleren Schätzfehler. Wie berechnet ist die native Lösung von $\hat{\gamma}$ also erwartungstreu.

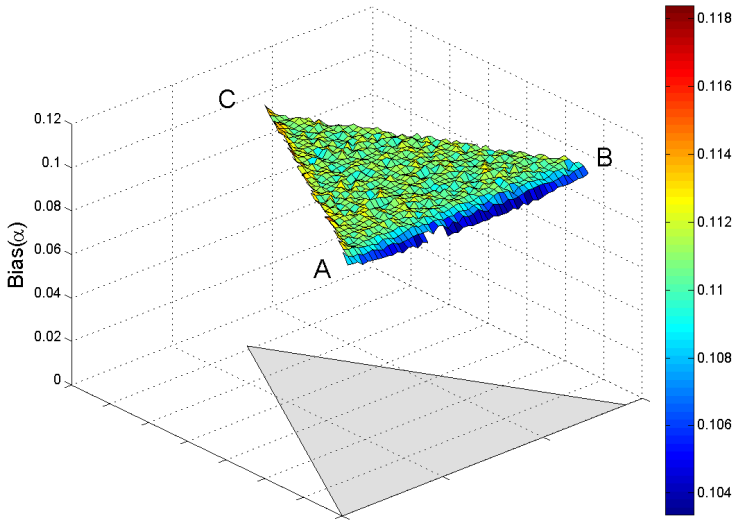


Abbildung 4.6: Schätzfehler für α der NCD ($N = 1.000, \delta = 0,10, 1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4. Die fehlenden Datenpunkte auf der Seite \overline{AB} sind durch Rundungsfehler bei der Koordinatentransformation bedingt.

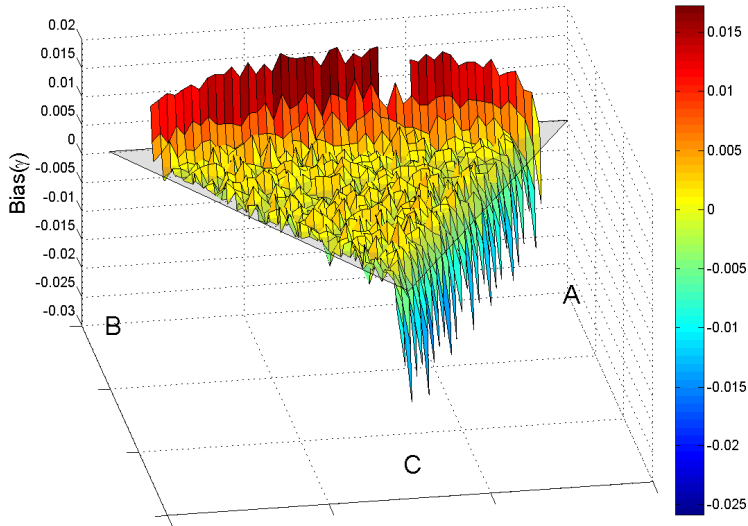


Abbildung 4.7: Schätzfehler für γ der NCD ($N = 1.000, \delta = 0,10, 1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4. Die fehlenden Datenpunkte auf der Seite \overline{AB} sind durch Rundungsfehler bei der Koordinatentransformation bedingt.

4.6.2 Mittlere Fehler

Auch hier ist zu erkennen, dass die NCD außerhalb der Randfälle als erwartungstreu eingestuft werden kann. Die Schätzfehler $\text{Bias}(\hat{\gamma})$, $\text{Bias}(\hat{\alpha})$ und $\text{Bias}(\hat{\beta})$ sind jeweils in Abbildung 4.8, B.6 und B.7 dargestellt. Die mittleren Schätzfehler über die simulierten Daten sind in Tabelle 4.4 zu finden.

Tabelle 4.4: Mittlere Schätzfehler der NCD zu den simulierten Daten, skaliert um den Faktor 10^6 . Man kann wie bei der TCD die Verzerrung durch die Randfälle deutlich erkennen.

	mit Randfällen	ohne Randfälle
$\text{Bias}(\hat{\alpha})$	-6.336	-392
$\text{Bias}(\hat{\beta})$	-3.412	-307
$\text{Bias}(\hat{\gamma})$	9.748	700

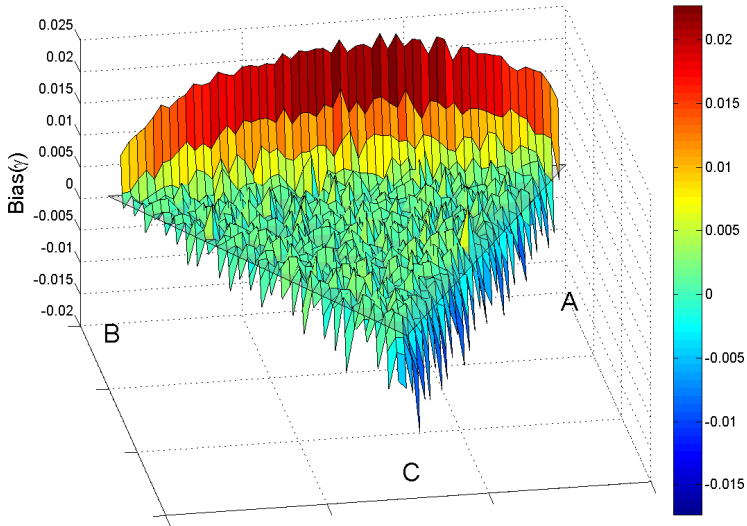


Abbildung 4.8: Schätzfehler für γ der NCD ($N = 1.000$, $\delta = 0, 1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

4.6.3 Standardabweichung der Schätzfehler

Hier wurde die selbe Vorgehensweise wie bei der TCD gewählt. Die Ergebnisse sind in Abbildungen 4.9 für α und 4.10 für γ sowie für den Cheater-Anteil in B.8 zu finden. Wieder ist die Proportionalität zu $\frac{1}{\sqrt{N}}$ deutlich zu sehen. Im Vergleich zur TCD braucht die NCD einen geringeren Stichprobenumfang, um qualitativ gleichwertige Ergebnisse zu liefern, da die Standardabweichungen durchweg geringer sind.

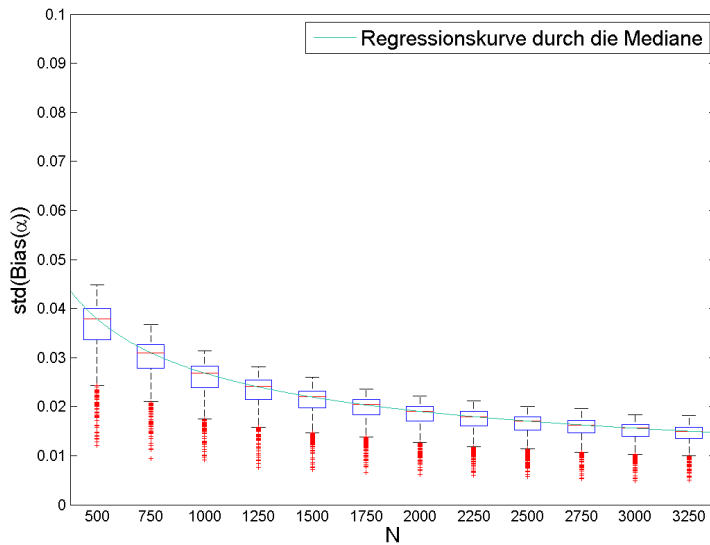


Abbildung 4.9: Entwicklung der Standardabweichung des Schätzfehlers von α bei der NCD gemäß 4.6.3. Erläuterungen zur Darstellung siehe Abbildung 4.4.

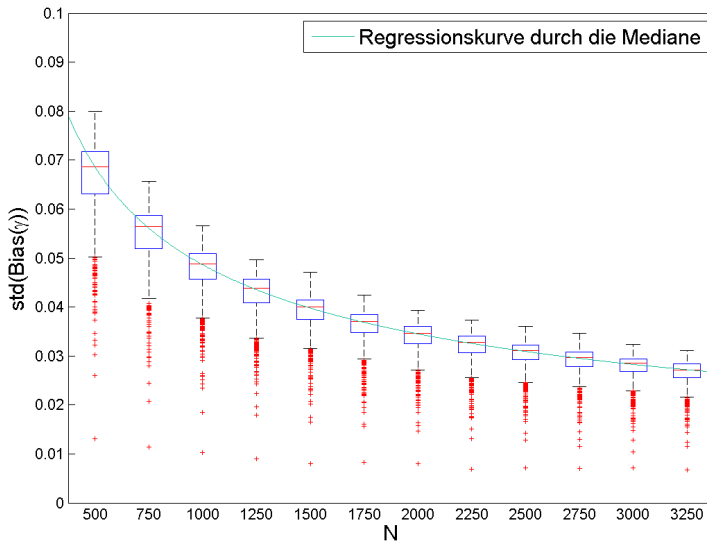


Abbildung 4.10: Entwicklung der Standardabweichung des Schätzfehlers von γ bei der NCD gemäß 4.6.3. Erläuterungen zur Darstellung siehe Abbildung 4.4.

4.7 Analyse der YCD

Die hier gewonnenen Ergebnisse entsprechen (im Rahmen der Simulationsgenauigkeit) exakt den Resultaten der NCD, wenn man α mit γ und β mit δ vertauscht. Es wurden

$$P^Y = \begin{pmatrix} 0,1 \\ 0,1 \end{pmatrix}, \quad P^N = \begin{pmatrix} 0,7 \\ 0,1 \end{pmatrix}$$

als Forced Answer-Wahrscheinlichkeiten verwendet.

4.7.1 Auswirkung $\beta \neq 0$

Ist $\beta \neq 0$, so erhalten wir hier auch eine Verzerrung der Schätzung. Zur Berechnung der Erwartungswerte benutzen wir die modifizierten Schätzer der „Nein“-Wahrscheinlichkeiten

$$1 - \hat{\lambda}_i = \alpha (1 - P_i^N) - \beta + \gamma P_i^Y.$$

Erwartungswert von $\hat{\alpha}$

$$\begin{aligned}\hat{\alpha} &= \frac{1}{\det M} \left((1 - P_2^Y) (1 - \hat{\lambda}_1) - (1 - P_1^Y) (1 - \hat{\lambda}_2) \right) \\ E(\hat{\alpha}) &= \frac{1}{\det M} (\alpha \det M + \beta (1 - P_2^Y - P_1^Y) + \gamma \cdot 0) \\ &= \alpha + \delta \frac{P_1^Y - P_2^Y}{\det M}\end{aligned}$$

Erwartungswert von $\hat{\gamma}$

$$\begin{aligned}\hat{\gamma} &= \frac{1}{\det M} \left(-P_2^N (1 - \hat{\lambda}_1) + P_1^N (1 - \hat{\lambda}_2) \right) \\ E(\hat{\gamma}) &= \frac{1}{\det M} (\alpha \cdot 0 + \beta (P_1^N - P_2^N) + \gamma \det M) \\ &= \gamma + \frac{P_2^N - P_1^N}{\det M} \delta\end{aligned}$$

Erwartungswert von $\hat{\beta}$

$$\begin{aligned}\hat{\delta} &= 1 - \hat{\alpha} - \hat{\gamma} \\ E(\hat{\delta}) &= E(1 - \hat{\alpha} - \hat{\gamma}) = 1 - \alpha - \beta \frac{P_1^Y - P_2^Y}{\det M} - \gamma - \beta \frac{P_1^N - P_2^N}{\det M} \\ &= \delta + \beta \left(1 - \frac{P_1^Y - P_2^Y + P_1^N - P_2^N}{\det M} \right)\end{aligned}$$

Auf den ersten Blick sieht es so aus, als wären die Vorzeichen der Ergebnisse genau umgekehrt zu denen der NCD. Da man aber bei symmetrischen Forced Answer-Wahrscheinlichkeiten die Verfahrensmatrix der YCD durch Vertauschen der Spalten der Verfahrensmatrix der NCD erhält, kehrt sich das Vorzeichen der Determinante um. Die Verzerrungen entsprechen also denen der NCD. Auch hier kann man mit $P_1^Y = P_2^Y$ einen erwartungstreuen Schätzer $\hat{\alpha}$ unabhängig von $\beta = 0$ erhalten.

4.7.2 Mittlere Fehler

Die Schätzfehler $\text{Bias}(\hat{\alpha})$, $\text{Bias}(\hat{\gamma})$ und $\text{Bias}(\hat{\delta})$ sind in Abbildung 4.11, B.9 und B.10 zu finden. Die YCD kann mit genügend Abstand von den Randfällen ebenfalls als erwartungstreu angesehen werden.

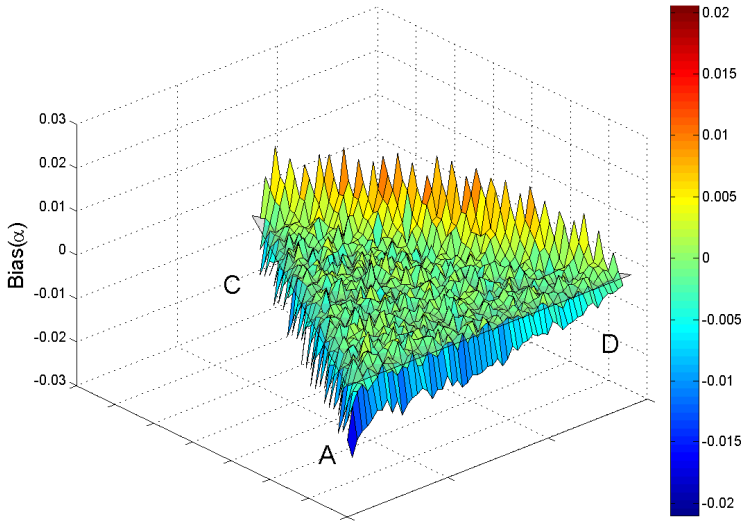


Abbildung 4.11: Schätzfehler für α der YCD ($N = 1.000$, $\beta = 0,1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

Tabelle 4.5: Mittlere Schätzfehler der YCD zu den simulierten Daten, skaliert um den Faktor 10^6 .

	mit Randfällen	ohne Randfälle
$\text{Bias}(\hat{\alpha})$	-7.865	1.705
$\text{Bias}(\hat{\gamma})$	2.619	-708
$\text{Bias}(\hat{\delta})$	5.245	-997

4.7.3 Standardabweichung der Schätzfehler

Die Entwicklung mit wachsendem Stichprobenumfang ist in Abbildung 4.12, B.11 und B.12 dargestellt. Das Vorgehen ist hierbei wieder analog zu dem bei TCD und NCD. Ein Vergleich mit Abbildungen bezüglich 4.6.3 hebt auch wieder die Symmetrie der beiden DCD-Verfahren hervor und die zur NCD formulierten Ergebnisse gelten auch hier.

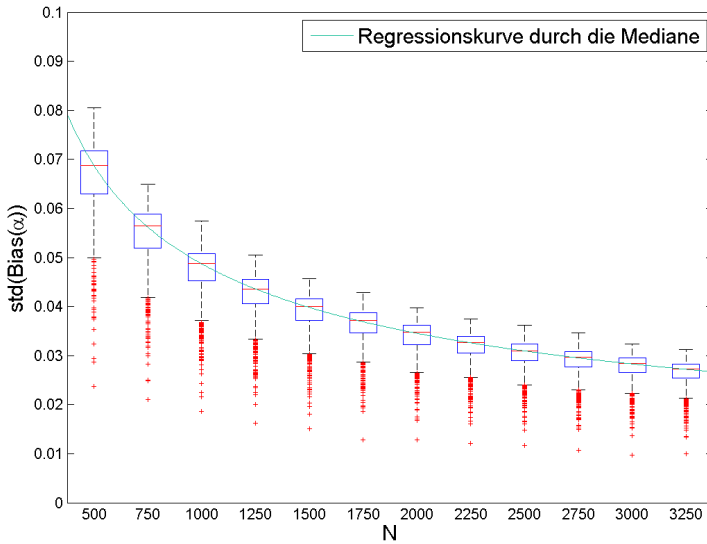


Abbildung 4.12: Entwicklung der Standardabweichung des Schätzfehlers von α bei der YCD gemäß 4.7.3. Erläuterungen zur Darstellung siehe Abbildung 4.4.

5 Vergleich der Verfahren

5.1 Notation

Bei den Analysen der Schätzer hatten wir

- für die TCD

$$P_{\text{TCD}}^Y = \begin{pmatrix} 0,7 \\ 0,1 \\ 0,1 \end{pmatrix}, \quad P_{\text{TCD}}^N = \begin{pmatrix} 0,1 \\ 0,7 \\ 0,1 \end{pmatrix}$$

- für die NCD

$$P_{\text{NCD}}^Y = \begin{pmatrix} 0,7 \\ 0,1 \end{pmatrix}, \quad P_{\text{NCD}}^N = \begin{pmatrix} 0,1 \\ 0,1 \end{pmatrix}$$

- für die YCD

$$P_{\text{YCD}}^Y = \begin{pmatrix} 0,1 \\ 0,1 \end{pmatrix}, \quad P_{\text{YCD}}^N = \begin{pmatrix} 0,7 \\ 0,1 \end{pmatrix}$$

als Forced Answer-Wahrscheinlichkeiten gewählt. Für die weiteren Analysen in diesem und im nächsten Kapitel wollen wir diese beibehalten und den zugehörigen RRT-Gruppen Namen geben:

- G^T bezeichne die Gruppe mit $P^Y = 0,1$ und $P^N = 0,1$.
- G^N bezeichne die Gruppe mit $P^Y = 0,7$ und $P^N = 0,1$.
- G^Y bezeichne die Gruppe mit $P^Y = 0,1$ und $P^N = 0,7$.

5.2 Zusammenhang TCD mit DCD

Unter bestimmten Kriterien stimmen die Schätzungen der TCD mit denen der DCD überein. Ein Blick auf die oben vorgeschlagenen Forced Answer-Wahrscheinlichkeiten suggeriert eine Untersuchung, wie sich zum Beispiel die NCD verhält, wenn man sie auf den beiden Gruppen G^T und G^N einer durchgeführten TCD-Befragung anwendet. Es hat sich herausgestellt, dass die Schätzung der TCD im Umfang von $1,5N$ Teilnehmern (wobei alle drei Gruppen gleichen Umfang $0,5N$ haben), mit den Schätzungen der DCD im Umfang von N Teilnehmern im Bezug auf α und γ unter bestimmten Kriterien übereinstimmen, sogar wenn der Populationsparameter, den die entsprechende Methode der DCD nicht schätzt, ungleich Null ist.

Ein notwendiges Kriterium für die Übereinstimmung scheint zu sein, dass die Forced Answer-Wahrscheinlichkeiten der TCD die Form

$$P^Y = \begin{pmatrix} 1 - \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_2 \end{pmatrix}, \quad P^N = \begin{pmatrix} \varepsilon_3 \\ 1 - \varepsilon_4 \\ \varepsilon_3 \end{pmatrix}$$

haben. Dies konnten wir allerdings nur empirisch beobachten. Ist dies erfüllt, kann man folgende Regelmäßigkeiten (Ausnahmen: Die Randfälle $\beta = 1$ und $\delta = 1$) beobachten:

- Nimmt die TCD einen Randfall bezüglich β an, so nimmt auch die NCD diesen Randfall an und beide Schätzer unterscheiden sich in γ minimal. Das heißt, es wurden nur absolute Abweichungen von $< 1\%$ beobachtet.
- Für δ gilt die analoge Aussage bezüglich der Schätzung von α in der YCD.
- Nimmt die TCD den Randfall $\alpha = 0$ an, so nimmt die YCD diesen auch an (und die Schätzungen stimmen offensichtlich überein) und der Schätzwert, den die NCD für γ liefert, unterscheidet sich aber von dem der TCD.
- Im Randfall $\gamma = 0$ bei der TCD gilt die analoge Aussage.

Sind die Forced Answer-Wahrscheinlichkeiten nicht in obiger Form, so besteht ein funktionaler Zusammenhang zwischen den Abweichungen der Verfahren und (mindestens) den Forced Answer-Wahrscheinlichkeiten. Dieser scheint jedoch nicht auf einfache Weise darstellbar zu sein, er ist an dieser Stelle aber auch nicht von zentralem Interesse.

5.3 Vergleich zum Verfahren von Clark und Desharnais

Die Herleitung lässt erwarten, dass das von Clark und Desharnais (1998) entwickelte Verfahren ein Spezialfall der hier vorgestellten NCD mit $P^N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ist. Simulationen bestätigen diese Vermutung. Vom Standpunkt der Versuchsplanung sind die Verfahren jedoch unterschiedlich, da wir davon ausgehen, dass das Ausbleiben der sensitiven Eigenschaft auch peinlich sein kann. Deshalb werden wir hier keine weiteren Vergleiche mit diesem Verfahren anstellen.

6 Praktische Anwendungen

Wir wollen nun noch einige Möglichkeiten und Probleme, die sich in der praktischen Anwendung der Verfahren ergeben, besprechen. Dabei beschränken wir uns auf besondere Eigenschaften der hier entwickelten Verfahren und wollen nicht weiter auf generelle Fragen zur Anwendung der RRT eingehen, da diese bereits ausführlich untersucht sind (u.a. in Übersichtsbeiträgen von Lensvelt-Mulders et al., 2005 oder Wolter, 2012).

6.1 Auswirkungen unterschiedlicher Umfänge in den RRT-Gruppen

Bei der Versuchsplanung sollte man sich die Frage stellen, wie man am Besten mit Antwortausfällen umgeht, vor Allem, wenn sie in einer der RRT-Gruppen gehäuft auftreten. Sollte man dann die Gruppen weiter zufällig zuweisen, oder sollte man Befragte in die Gruppe mit den wenigsten Antworten zuteilen?

Auch aus mathematischer Sicht ist die Frage, wie sich unterschiedliche Gruppenanteile in der Gesamtbefragung auswirken, sinnvoll. Betrachtet man etwa bei den (mathematisch) optimalen Forced Answer-Wahrscheinlichkeiten die Gruppe G^N in der NCD, so lässt sich ohne Informationen aus der Gruppe G^T zu benutzen, direkt der Parameter β mittels $\lambda_{G^N} = \alpha + \gamma = 1 - \beta$ schätzen. Im Randfall mit den (mathematisch) optimalen Forced Answer-Wahrscheinlichkeiten kann die Schätzung von β also komplett von der Schätzung der anderen Parameter entkoppelt werden. Wählt man die Forced Answer-Wahrscheinlichkeiten nun sinnvoll, liegt die Vermutung nahe, dass die Gruppe G^N einen größeren Teil zur Abschätzung von β beiträgt, als die Gruppe G^T . Demzufolge würde ein höherer Anteil des Gruppenfangs der Gruppe G^N die Qualität der Schätzung von β positiv beeinflussen.

Bei der Darstellung der Ergebnisse steht man wieder vor dem Problem einer übersichtlichen Darstellung. Da die Entwicklung der Standardabweichung der Schätzfehler mit dem Befragungsumfang N bekannt ist, wurde bei diesen

Simulationen N auf 1.500 festgesetzt. Wir wendeten die oben bereits dargestellte Monte-Carlo-Methode an, wobei ausschließlich die Umfänge der RRT-Gruppen variiert wurden.

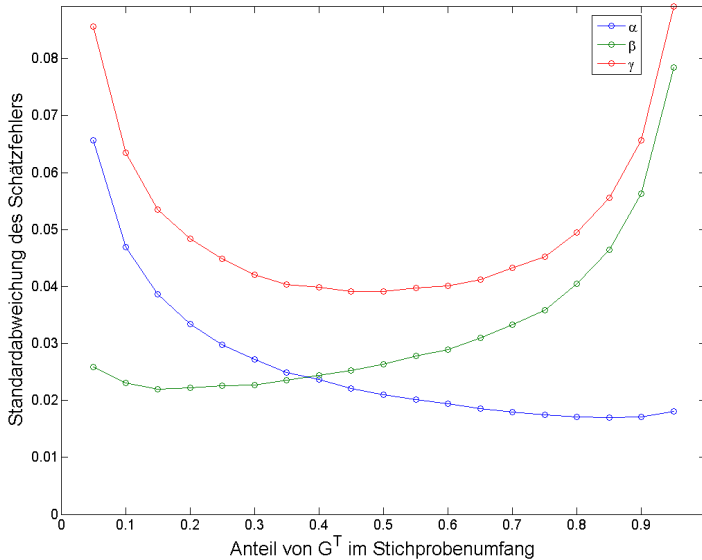


Abbildung 6.1: Auswirkung unterschiedlicher Gruppenumfänge bei der NCD ($N = 1.500$).

Bei der NCD gilt $|G^N| + |G^T| = 1$, wobei mit $|\cdot|$ die Anteile der jeweiligen Gruppen am gesamten Befragungsumfang bezeichnet sind. Wie erwartet führt eine Erhöhung von $|G^N|$ (durch eine Verringerung von $|G^T|$) zu einer Verbesserung der Schätzung von β . Der bei der NCD zentrale Parameter γ wird am Besten geschätzt, wenn beide Gruppen etwa gleich groß sind. Siehe hierzu Abbildung 6.1. Weiter kann man hier sehen, dass zum Abschätzen des Vorliegens einer extrem peinlichen Eigenschaft (so peinlich, dass es keine „Ja“-Cheater gibt, also $\delta = 0$, womit die Erwartungstreue der NCD garantiert ist), die NCD mit großem $|G^T|$ sehr gut geeignet ist, sofern man sich hauptsächlich für den Parameter α interessiert.

In Abbildung B.13 sind die Simulationsergebnisse zur YCD zu sehen, welche erneut symmetrisch zu den Ergebnissen der NCD sind.

Bei der TCD (Abbildung 6.2) wurde $|G^N| = |G^Y|$ zur weiteren Vereinfachung angenommen. Es gilt also $|G^N| = |G^Y| = \frac{1-|G^T|}{2}$. Die optimale Schätzung von α und γ liegt hier auch etwa in dem Bereich vor, wo alle Gruppen etwa gleich groß sind ($|G^T| \approx \frac{1}{3}$). Eine Verringerung von $|G^T|$ führt zu einer besseren Schätzung von β und δ , welche in den meisten Fällen aber die uninteressanten Parameter sind, weil sie keine Information über das Vorliegen der zu untersuchenden Eigenschaft bieten.

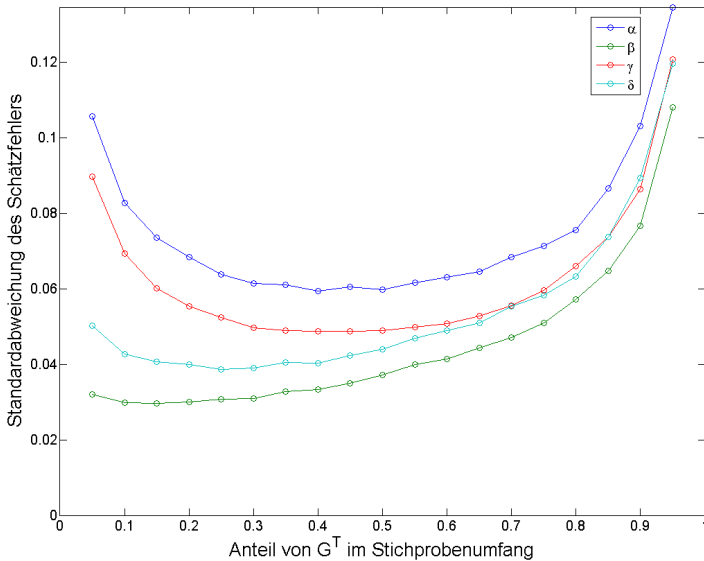


Abbildung 6.2: Auswirkung unterschiedlicher Gruppenumfänge bei der TCD ($N = 1.500$).

Allerdings kann bei der TCD nicht wie bei den beiden DCD-Verfahren durch Erhöhen von $|G^T|$ die Qualität des Schätzers α oder γ selektiv gesteigert werden.

Die Kenntnis, wie die Gruppenumfänge die Schätzung beeinflussen, kann man sich bei der Versuchsplanung zunutze machen um z.B. eine laufende NCD-Umfrage in eine TCD-Umfrage zu erweitern (siehe 6.2).

6.2 Dynamische Steuerung einer Umfrage

Grundsätzlich bietet sich die Möglichkeit, eine bereits laufende Umfrage schon auszuwerten, bevor ein Großteil des Rücklaufs vorliegt. Man hat so die Möglichkeit, die Befragung bezüglich ihrem Ziel zu optimieren bzw. den noch erwarteten restlichen Rücklauf effizienter nutzen zu können. Stellt etwa nach Beginn einer NCD-Befragung fest, dass fälschlicherweise der Populationsparameter $\delta = 0$ angenommen wurde, so sollte die Befragung zu einer TCD umgewandelt werden, da der Schätzer sonst bezüglich α und β verzerrt ist. Hier soll vordergründig die Frage behandelt werden, wie eine DCD zur Laufzeit möglichst effizient in eine TCD umgewandelt werden kann.

Wir haben in 6.1 bei der Untersuchung der TCD $|G^N| = |G^Y|$ eingehalten. Dies wollen wir hier auch tun (basierend auf den bisher vorhandenen Ergebnissen wird eine Veränderung der Gruppenumfänge vermutlich die Schätzung eines Parameters verbessern, während die Schätzung eines anderen Parameters verschlechtert wird). Unter dieser Bedingung ist die Erweiterung einer DCD auf eine TCD mit wenig zusätzlichen Antworten möglich, wenn $|G^T|$ relativ groß ist. Oben haben wir gesehen, dass die TCD jedoch α und γ schlechter schätzt, wenn die Gruppenumfänge nicht gleich groß sind. Bei dieser Erweiterung kommt aber noch ein entscheidender Punkt hinzu: Beim Übergang zur TCD wird der Befragungsumfang N vergrößert. Es muss also noch die Effizienz bezüglich N mit einbezogen werden, vor allem wenn der Rücklauf geringer ausfällt, als er ursprünglich erwartet wurde. Wir wollen dies an einem Beispiel erläutern: Es liegen die Antworten einer NCD mit 1.500 Befragten vor. Im Fall $|G^N| = |G^T| = \frac{1}{2}$ werden also noch zusätzliche 750 Antworten benötigt um eine TCD mit optimalen Gruppenumfängen zu erhalten. Nachdem man bei 2.250 Teilnehmern bei einer TCD mit optimalen Gruppenanteilen angekommen ist, kann man alle weiteren Umfrageteilnehmer gleichmäßig auf die drei Gruppen verteilen. Liegen andererseits die 1.500 Antworten mit $|G^T| = \frac{4}{5}$ vor, so werden lediglich noch $\frac{1}{5} \cdot 1500 = 300$ Antworten benötigt, bis man eine TCD mit $|G^N| = |G^Y|$ erreicht hat. Danach kann man alle weiteren Befragten den Gruppen G^N und G^Y zuordnen, bis alle drei Gruppen gleich groß sind.

Bei einer eingeschränkten Anzahl an Antworten steht man also im Allgemeinen vor dem Problem, dass man zwischen der Qualität der DCD-Schätzung und dem Aufwand zur Umwandlung in eine TCD abwägen muss. Da der Umfang der entstehenden TCD von der Größe $|G^T|$ in der DCD beeinflusst wird, ist die Standardabweichung der Schätzfehler alleine kein geeigneter Indikator für die Umwandlungseffizienz. Die Standardabweichung ist proportional zu $\frac{1}{\sqrt{N}}$.

Die Qualität der resultierenden TCD kann in Relation zur Stichprobengröße also über

$$\text{std}_{\text{TCD}}(\chi_e, N_{\text{TCD}}) \cdot \sqrt{N_{\text{TCD}}}$$

mit $N_{\text{TCD}} = N_{\text{DCD}}(1 + |G_{\text{DCD}}^{\text{N}}|)$, $\chi \in \{\alpha, \beta, \gamma, \delta\}$ abgeschätzt werden, wobei ein geringerer Wert besser ist. Abbildung 6.3 zeigt diese Abschätzung. Die durchgezogenen Linien bilden die Standardabweichungen (skaliert um den Faktor 100) der NCD ab, wie sie auch in Abbildung 6.1 zu sehen sind. Unten entlang der horizontalen Achse sind die Anteile $|G_{\text{NCD}}^{\text{T}}|$ verzeichnet, während oben die Stichprobengröße der resultierenden TCD zu finden ist. Die gestrichelten Linien sind die Schätzer für die Effizienz wie oben definiert.

Im Bereich $0,4 \leq |G_{\text{NCD}}^{\text{T}}| \leq 0,8$ bleibt die Übergangseffizienz etwa gleich gut, während auch die Schätzung der NCD bezüglich γ nicht viel an Qualität einbüßt. Die Schätzung für α der NCD wird sogar besser mit wachsendem $|G_{\text{NCD}}^{\text{T}}|$. Für die YCD gelten die analogen Ergebnisse.

Sicher gibt es eine Vielzahl an organisatorischen Problemen bei der Durchführung einer Befragung auf diese Art. Je nach Natur der Befragung kann es aber noch zu anderen, ernsthaften Problemen kommen. Sollten beispielsweise die Verzögerung, mit der eine Person sich entscheidet, an einer Befragung teilzunehmen etwa mit seinem Antwortverhalten (Cheater oder ehrlicher Befragungsteilnehmer) oder dem Vorliegen der Eigenschaft korrelieren, dann besteht die Möglichkeit, dass die Populationsgruppen in den RRT-Gruppen anteilmäßig ungleich vertreten sind. Dies führt, wie wir in 6.3 sehen werdenn, zu einem verzerrten Schätzer.

In Abbildung 6.4 ist der Verlauf einer simulierten dynamisch gesteuerten Befragung zu sehen. Die uns unbekannt Parameter sind $\alpha = 0,3$, $\beta = 0,1$, $\gamma = 0,55$, $\delta = 0,05$. In der Planungsphase ist die Entscheidung gefallen, eine NCD durchzuführen, da uns gute Gründe zur Annahme $\delta = 0$ vorliegen. Wir rechnen mit einem Umfang von ca. 2.000 Befragten, also theoretisch genug Leute um eine TCD problemlos durchführen zu können. Wir wollen uns die Option offen halten, möglichst schnell auf eine TCD umschwenken zu können und setzen deswegen $|G_{\text{NCD}}^{\text{T}}| = 0,8$ an.

Nach den ersten 1.000 Antworten können wir abschätzen, dass der Rücklauf viel höher ist als erwartet, so dass wir insgesamt mit etwa 3.000 Antworten, statt der geplanten 2.000 haushalten können. Dies veranlasst uns zu der Entscheidung, die nächsten 200 Antworten für eine Erweiterung auf die TCD zu investieren. Nun gibt es drei sinnvolle Möglichkeiten, was geschehen kann.

1. Die Vermutung $\delta = 0$ bestätigt sich, aber wir bleiben bei der TCD, da die unerwartet hohe Anzahl an Antworten im Groben die ursprünglich erwartete Qualität liefert.

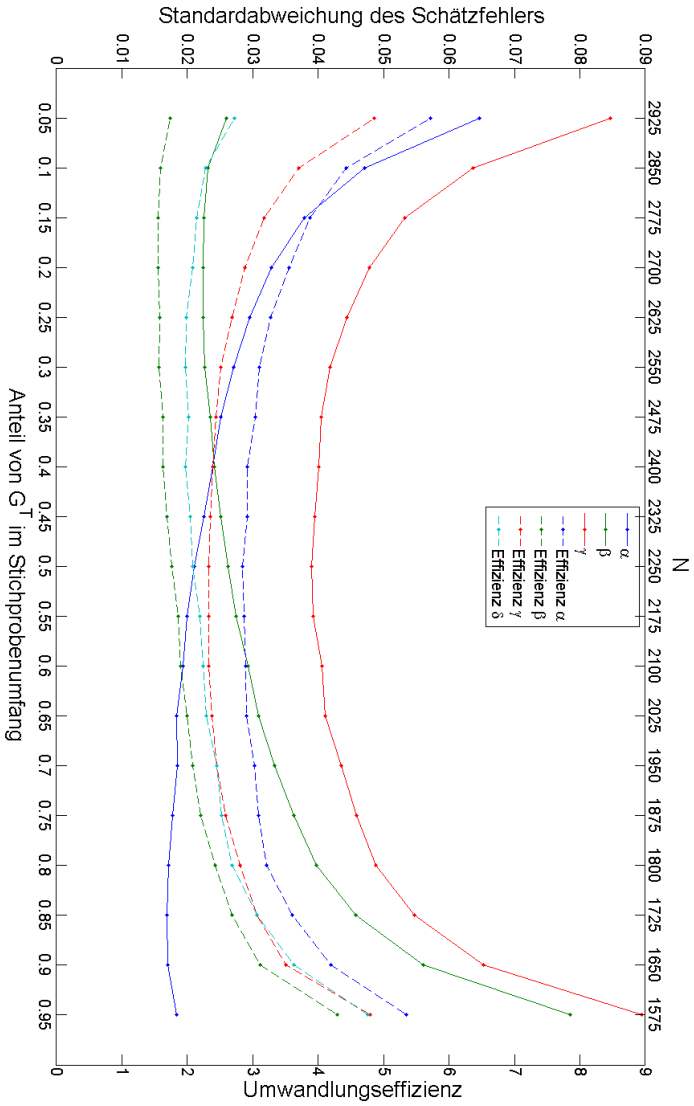


Abbildung 6.3: Effizienz bei dynamischer Steuerung.

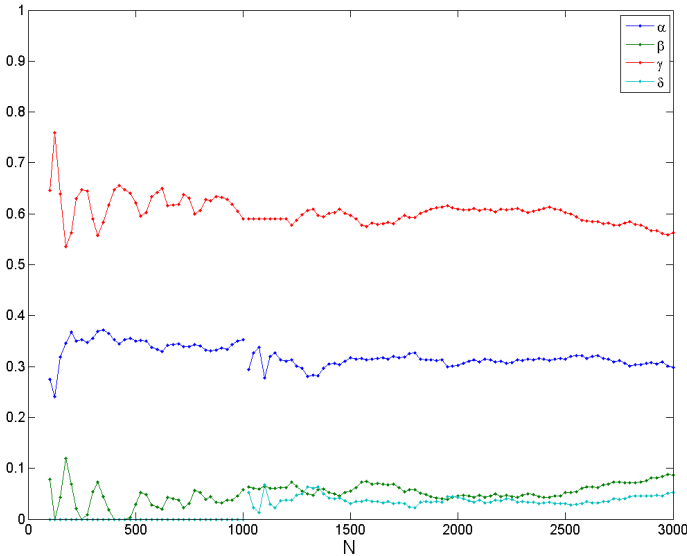


Abbildung 6.4: Beispielschätzung einer dynamisch gesteuerten Befragung.

2. Die Vermutung $\delta = 0$ bestätigt sich, wir gehen wieder zur NCD zurück und arbeiten auf $|G_{NCD}^T| = |G_{NCD}^N|$ hin. In diesem Fall ist es zwar schade um die 200 verlorenen Antworten, die wir nicht sinnvoll in der NCD mit einrechnen können, weil die Forced Answer-Wahrscheinlichkeiten absolut unpassend sind, aber die NCD mit 2.800 Antworten liefert ein zuverlässigeres Ergebnis als die TCD mit 3.000 Antworten.

Rechnen wir die NCD auf drei Gruppen, so liegt kein echter ML-Schätzer vor und der relativ kleine Umfang dieser dritten Gruppe bietet ein hohes Potential für statistische Ausreißer. Arbeiten wir die 200 Antworten in die beiden Gruppen G^T oder G^N ein, müssen wir in der Verfahrensmatrix die gewichteten Mittelwerte der Forced Answer-Wahrscheinlichkeiten benutzen. Asymptotisch betrachtet verschlechtert sich die Determinante der Verfahrensmatrix linear mit den zusätzlichen Antworten, während der steigende Stichprobenumfang nur mit Exponent $\frac{1}{2}$ zur Verbesserung der Schätzung beiträgt.

3. Unsere Annahme $\delta = 0$ wird nicht bestätigt. Wir bleiben also bei der TCD und steuern zur Verbesserung der Qualität auf gleiche Gruppenumfänge zu.

Obwohl der (uns bis dahin unbekannte) Parameter δ nur sehr gering ist, können wir schnell sehen, dass unsere Annahme $\delta = 0$ falsch war. Und wir verfahren gemäß Möglichkeit 3.

6.3 Unterschiedliches Antwortverhalten in den einzelnen RRT-Gruppen

Eine Annahme, die bei der Herleitung der Verfahren gemacht wurde, ist dass das Antwortverhalten einer Person unabhängig von der Zuteilung in eine bestimmte RRT-Gruppe ist. Für Cheating gibt es unterschiedliche Gründe. So kann es sein, dass eine Person mutwillig versucht die Umfrage zu verfälschen, oder einfach die Instruktion nicht versteht. Diese Arten von Cheating wirken sich aber nicht auf die gemachte Annahme aus, sofern sich diese Personen im Mittel gleich auf die Gruppen verteilen. Darüber hinaus kann der Mangel an Schutzgefühl zu Falschantworten führen. Dieser kann natürlich in jeder RRT-Gruppe unterschiedlich sein, da die Forced Answer-Wahrscheinlichkeiten unterschiedlich groß sind. Allgemein hängt der erlebte Schutz der Anonymität von der Wahrscheinlichkeit der Ausweichinstruktion für die als peinlich erlebte Eigenschaft ab. Ist die Antwort „Ja“ peinlich, so wird der Schutz als um so größer eingeschätzt, je mehr Befragte aufgrund der Instruktion „Ja“ antworten sollen. Betrachtet man eine TCD mit

$$P^Y = \begin{pmatrix} 0,1 \\ 0,1 \\ 0,7 \end{pmatrix}, \quad P^N = \begin{pmatrix} 0,1 \\ 0,7 \\ 0,1 \end{pmatrix},$$

so ist die „Ja“-Antwort zwar in den ersten beiden RRT-Gruppen gleichermaßen schwach geschützt, dies betrifft aber einen deutlich unterschiedlichen Anteil der Befragten. In der ersten Gruppe werden mit einem Anteil von $0,8 \cdot \alpha$ daher deutlich mehr Personen mit der peinlichen Eigenschaft in die Versuchung geraten, die Instruktionen „Antworte ja“ oder „Antworte ehrlich“ zu missachten und „nein“ zu antworten als in der zweiten und dritten RRT-Gruppe, in denen der Anteil nur $0,4 \cdot \alpha$ ist.

Unterschiedliches Antwortverhalten in den RRT-Gruppen führt dazu, dass die Schätzer (im Allgemeinen) verzerrt werden. Wir wollen dies an einem Beispiel mit der NCD zeigen: Zu einer NCD mit Gesamtumfang $N = 1.000$

werden aus zwei Populationen jeweils $N_1 = N_2 = 500$ Befragungsteilnehmer simuliert. Die Populationsparameter sind hierbei

$$\alpha_1 = 0,3, \beta_1 = 0,1, \gamma_1 = 0,6 \text{ und } \alpha_2 = 0,2, \beta_2 = 0,2, \gamma_2 = 0,6.$$

Die erste Hälfte wird der Gruppe G^N und die zweite Hälfte der Gruppe G^T zugeteilt. In der Gruppe G^T ist die „Ja“-Antwort deutlich schlechter geschützt als in der Gruppe G^N . Die unterschiedlichen Parameter spiegeln also die erhöhte Bereitschaft zum „Nein“-Cheating mangels Schutzgefühl wider. Bei einem unverzerrten Schätzer würde man in dieser Situation genau den Mittelwert der Parameter der beiden Teilpopulationen erwarten, also

$$\hat{\alpha} = 0,25, \hat{\beta} = 0,15, \hat{\gamma} = 0,6.$$

Unter genau diesen Bedingungen wurde die Simulation 10.000 mal durchgeführt und als mittlere Schätzwerte haben sich

$$\hat{\alpha} = 0,1832, \hat{\beta} = 0,0670, \hat{\gamma} = 0,7498$$

ergeben. Die Schätzung weicht also deutlich von den wahren Parametern ab.

Gelingt es, empirisch einen funktionalen Zusammenhang zwischen Forced Answer-Wahrscheinlichkeiten und Schutzgefühl bzw. Bereitschaft zum Cheating aufzustellen, kann man dies bei der Herleitung der Verfahren weiter berücksichtigen und entsprechend in der Rechnung kompensieren.

Um dieser Art des Cheatings entgegen zu wirken, sollte man beim Aufstellen der Forced Answer-Instruktionen nach Möglichkeit verschleiern, dass die Wahrscheinlichkeit für die sensitive Frage relativ hoch, bzw. das Zugeben oder Leugnen der Eigenschaft schlecht geschützt ist.

7 Ausblick

Die Herleitung der Verfahren gilt für beliebig viele RRT-Gruppen. Es ist also möglich, zu den hier vorgestellten Verfahren noch weitere Verallgemeinerungen zu entwickeln, welche mehr Parameter schätzen können. Beispielsweise könnte man die eingeführten vier Populationsgruppen um zwei erweitern:

- ε - Leute mit der peinlichen Eigenschaft, welche sich bei der Forced Answer-Instruktion korrekt verhalten und bei der sensitiven Frage mit „Nein“ antworten.
- φ - Leute ohne die peinliche Eigenschaft, welche sich bei der Forced Answer-Instruktion korrekt verhalten und bei der sensitiven Frage mit „Ja“ antworten.

Entsprechend dem Vorgehen bei der DCD kann man so auch Verfahren herleiten, welche nur einen Teil der Populationsparameter schätzen, sofern die begründete Annahme besteht, dass die restlichen Parameter gleich Null sind.

In Kapitel 5.2 haben wir gesehen, dass zum Schätzen von mehr Parametern auch ein größerer Befragungsumfang von Nöten ist, wenn die Schätzung nicht an Qualität verlieren soll. Überträgt man diese Zusammenhänge auf ein Verfahren mit fünf RRT-Gruppen, das sechs Populationsparameter schätzt, benötigt man 2,5-mal so viele Umfrageteilnehmer, um die Qualität eines Verfahrens zu erreichen, das auf nur drei Parameter ausgelegt ist.

Es stellen sich noch weitere praktische Schwierigkeiten bei der Verallgemeinerung. Vor allem müssen die Forced Yes- und Forced No-Wahrscheinlichkeiten in den Gruppen so gewählt werden, dass zum einen die Befragten sich geschützt fühlen, zum anderen aber die Determinante der Verfahrensmatrix weder singulär noch der Betrag der Determinante sehr klein wird. Je mehr RRT-Gruppen im Verfahren sind, um so komplexer wird diese Variation der Forced Answer-Wahrscheinlichkeiten, da die Verfahrensmatrix für jeden weiteren Parameter eine zusätzliche Zeile bekommt, welche sich von den anderen Zeilen lediglich in den Forced Answer-Wahrscheinlichkeiten unterscheiden

muss. Sind aber zwei Zeilen gleich, ist die Matrix bereits singulär. Unterscheiden sich zwei Zeilen nur gering, so ist der Betrag der Determinante sehr klein. Wenn die Instruktionen einfach gehalten werden sollen, sind aber „fast“ gleiche Zeilen nicht vermeidbar. Zum Anderen wird die Herleitung und Umsetzung der Verfahren exponentiell komplexer: Setzt man R RRT-Gruppen zur Berechnung von $R + 1$ Populationsparametern an, so müssen dazu $2^R - 1$ Randfälle betrachtet werden.

A Herleitung der Dreiecksdarstellung

Es sei $\delta \in [0, 1)$ gegeben (für $\delta = 1$ lässt sich die Transformation nicht durchführen, aber sie ist in diesem Fall offensichtlich nicht notwendig). Es sei

$$T := \left\{ \Theta = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \in \mathbb{R}^3 \mid \alpha + \beta + \gamma = 1 - \delta \wedge \alpha, \beta, \gamma \in [0, 1] \right\}$$

die Menge aller gültigen Populationsparameter zu dem gegebenen δ . Wir wählen

$$B := \left\{ B_a = \begin{pmatrix} 1 - \delta \\ 0 \\ 0 \end{pmatrix}, B_b = \begin{pmatrix} 0 \\ 1 - \delta \\ 0 \end{pmatrix}, B_c = \begin{pmatrix} 0 \\ 0 \\ 1 - \delta \end{pmatrix} \right\}$$

als Basis des \mathbb{R}^3 . Schreiben wir nun $\Theta \in T$ als Linearkombination der gewählten Basis, $\Theta = \lambda_a B_a + \lambda_b B_b + \lambda_c B_c$, erhalten wir

$$\lambda_a (1 - \delta) + \lambda_b (1 - \delta) + \lambda_c (1 - \delta) = 1 - \delta \Leftrightarrow \lambda_a + \lambda_b + \lambda_c = 1$$

wegen $\alpha + \beta + \gamma = 1 - \delta$. Mit $\alpha, \beta, \gamma \in [0, 1]$ folgt weiter, dass $\lambda_a, \lambda_b, \lambda_c \in [0, 1]$. Jeder Punkt $\Theta \in T$ lässt sich also als Konvexkombination der Basisvektoren darstellen und die Darstellung ist eindeutig, weil es eine spezielle Linearkombination der Basisvektoren ist. Das heißt, dass alle Punkte aus T auf der konvexen Hülle der Basisvektoren liegen, welche gemäß der Wahl von B aber gerade ein gleichseitiges Dreieck mit Kantenlänge $(1 - \delta) \sqrt{2}$ ist (vgl. Abbildung A.1). Die Koeffizienten der Konvexkombination lassen sich nun auch als Schwerpunktkoordinaten in diesem Dreieck interpretieren. Wir können jetzt ein beliebiges anderes gleichseitiges Dreieck $T' \subset \mathbb{R}^3$ gleicher Kantenlänge wählen, dort zu einem gegebenem Punkt $\Theta' \in T'$ die Schwerpunktkoordinaten $\lambda'_a, \lambda'_b, \lambda'_c$ berechnen, so dass die Punkte Θ und Θ' innerhalb

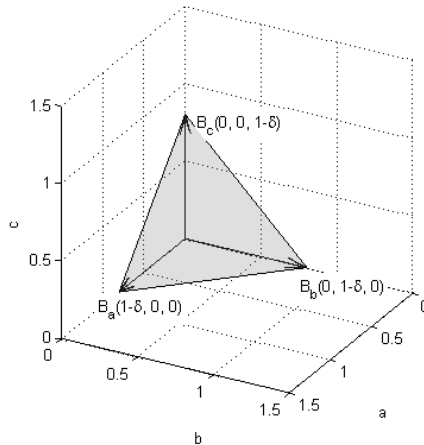


Abbildung A.1: Die konvexe Hülle von B_a, B_b und B_c (grau) ist ein gleichseitiges Dreieck der Kantenlänge $(1 - \delta)\sqrt{2}$.

des jeweiligen Dreiecks die gleiche Lage haben, wenn man ihre Schwerpunktkoordinaten gleichsetzt. Sei

$$B' := \left\{ B'_a = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, B'_b = \begin{pmatrix} (1-\delta)\sqrt{2} \\ 0 \\ 1 \end{pmatrix}, B'_c = \begin{pmatrix} (1-\delta)\frac{\sqrt{2}}{2} \\ (1-\delta)\sqrt{\frac{3}{2}} \\ 1 \end{pmatrix} \right\}.$$

Offensichtlich ist B' eine Basis des \mathbb{R}^3 und die Basisvektoren spannen ebenfalls ein gleichseitiges Dreieck der Kantenlänge $(1 - \delta)\sqrt{2}$ auf. Dieses von B' aufgespannte Dreieck T' lässt sich ohne Probleme in den \mathbb{R}^2 einbetten, indem man die dritte Koordinate einfach wegfällt lässt. Von der wegfallenden Koordinate brauchen wir uns nur die Information $\lambda'_a + \lambda'_b + \lambda'_c = 1$ zu merken. Haben wir nun einen Punkt $\Theta_2 = (x, y)$ in diesem Dreieck mit Eckpunkten

$$\begin{aligned} T_A &= (T_{A,x}, T_{A,y}) = (0, 0) \\ T_B &= (T_{B,x}, T_{B,y}) = \left((1-\delta)\sqrt{2}, 0 \right) \\ T_C &= (T_{C,x}, T_{C,y}) = \left((1-\delta)\frac{\sqrt{2}}{2}, (1-\delta)\sqrt{\frac{3}{2}} \right) \end{aligned}$$

so gilt die Beziehung

$$\begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = M \begin{pmatrix} \lambda'_a \\ \lambda'_b \\ \lambda'_c \end{pmatrix} \text{ mit } M = \begin{pmatrix} 1 & 1 & 1 \\ T_{A,x} & T_{B,x} & T_{C,x} \\ T_{A,y} & T_{B,y} & T_{C,y} \end{pmatrix}$$

über welche die Schwerpunktkoordinaten von $\Theta_2 \in T'$ berechnet werden können. Nach obiger Bemerkung entspricht dann der Punkt Θ_2 gerade dem Punkt, der aus der Linearkombination der ursprünglichen Basisvektoren von B mit den im zweidimensionalen errechneten Schwerpunktkoordinaten als Koeffizienten entsteht, also

$$\Theta_2 \hat{=} \lambda'_a B_a + \lambda'_b B_b + \lambda'_c B_c.$$

B Abbildungen und Tabellen

B.1 Abbildungen und Tabellen zur TCD

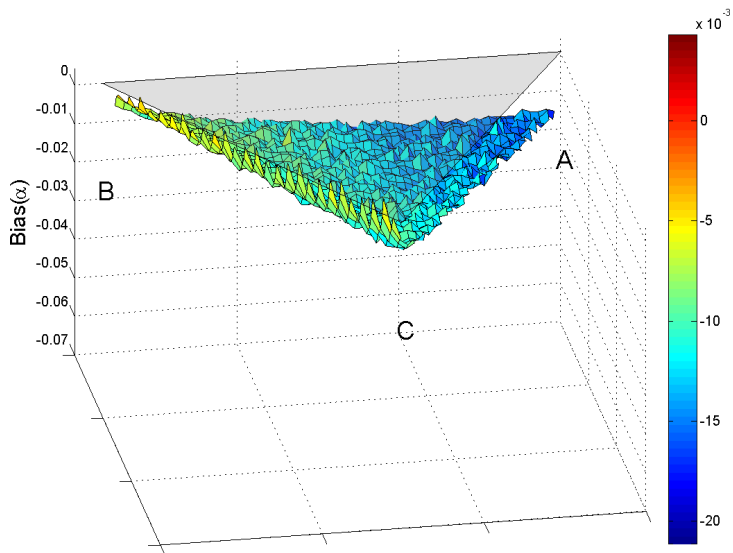


Abbildung B.1: Schätzfehler für α der TCD ($N = 1.500$, $\delta = 0,1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

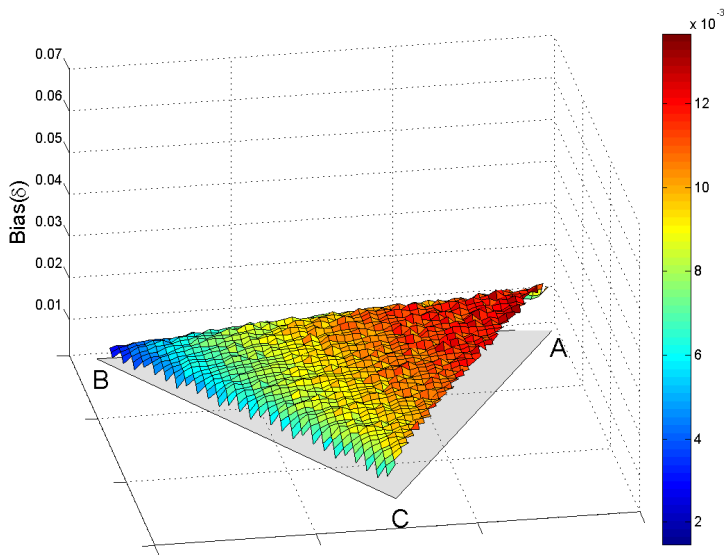


Abbildung B.2: Schätzfehler für δ der TCD ($N = 1.500$, $\delta = 0,1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

Tabelle B.1: Mittlere Schätzfehler der TCD zu den simulierten Daten, skaliert um den Faktor 10^6 .

δ	χ	Bias ($\hat{\chi}$)	
		mit Randfällen	ohne Randfälle
0,00	β	8.445	-
	δ	83.258	-
0,05	β	-3.824	-2.061
	δ	-13.548	-2.511
0,10	β	-11.509	-5.602
	δ	-23.762	-8.702
0,15	β	-19.614	-3.147
	δ	-24.743	-5.078
0,20	β	-27.075	-6.299
	δ	-16.145	-2.990
0,35	β	-28.682	-5.926
	δ	23.781	6.502
0,50	β	-19.141	-8.912
	δ	18.264	17.942
0,65	β	-35.129	-5.673
	δ	21.189	9.243
0,80	β	-17.175	-2.810
	δ	9.673	11.556
0,95	β	-388	-
	δ	-10.765	-

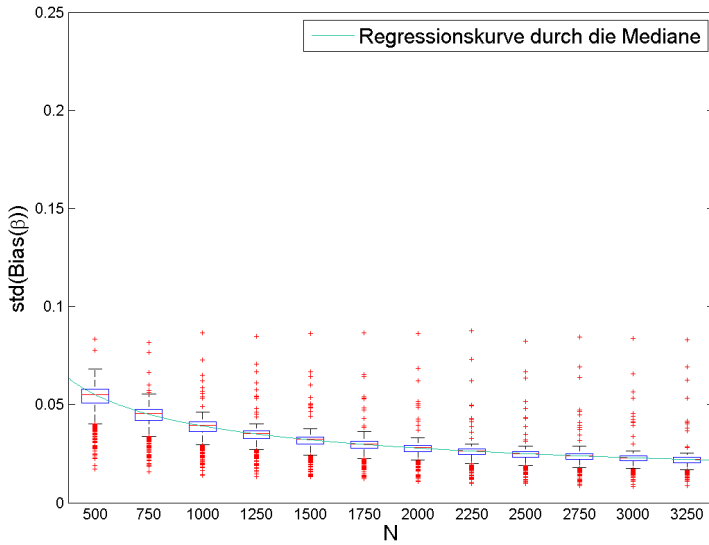


Abbildung B.3: Entwicklung der Standardabweichung des Schätzfehlers von β bei der TCD gemäß 4.5.2. Erläuterungen zur Darstellung siehe Abbildung 4.4.

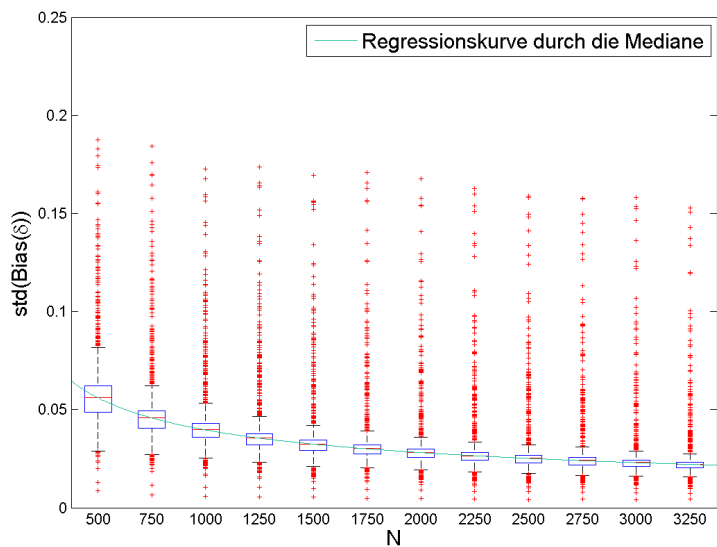


Abbildung B.4: Entwicklung der Standardabweichung des Schätzfehlers von δ bei der TCD gemäß 4.5.2. Erläuterungen zur Darstellung siehe Abbildung 4.4.

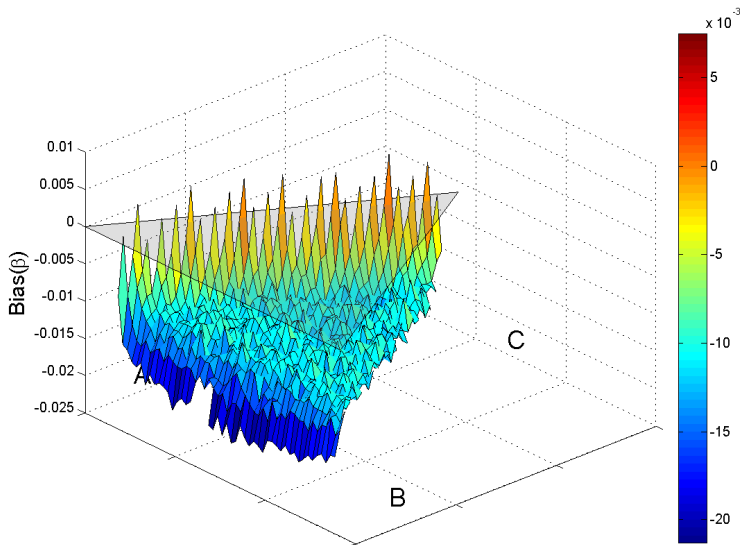
B.2 Abbildungen zur NCD

Abbildung B.5: Schätzfehler für β der NCD ($N = 1.000$, $\delta = 0, 10, 1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

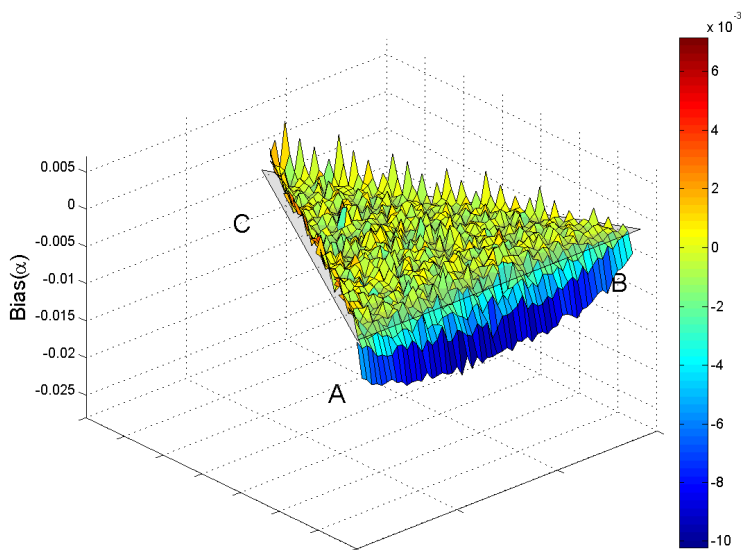


Abbildung B.6: Schätzfehler für α der NCD ($N = 1.000$, $\delta = 0,1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

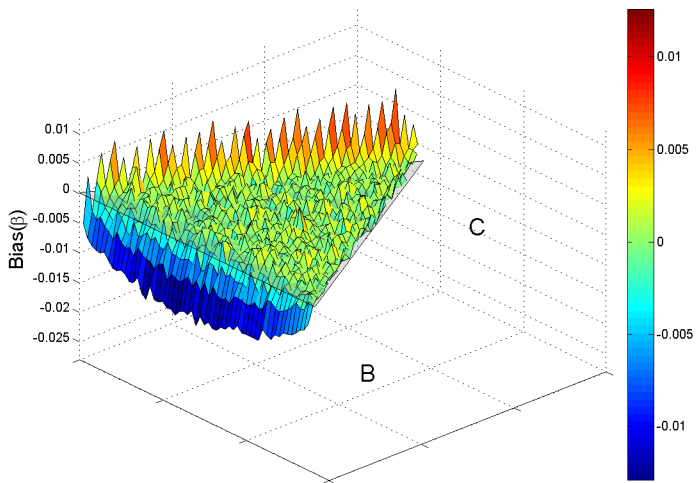


Abbildung B.7: Schätzfehler für β der NCD ($N = 1.000$, $\delta = 0,1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

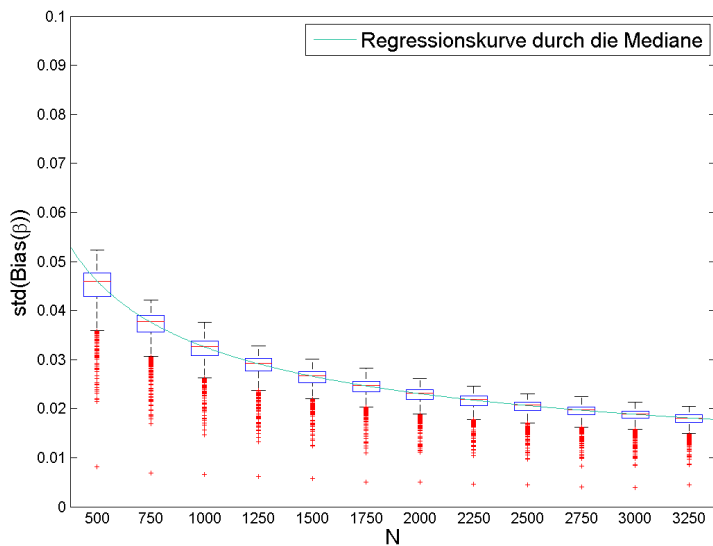


Abbildung B.8: Entwicklung der Standardabweichung des Schätzfehlers von β bei der NCD gemäß 4.6.3. Erläuterungen zur Darstellung siehe Abbildung 4.4.

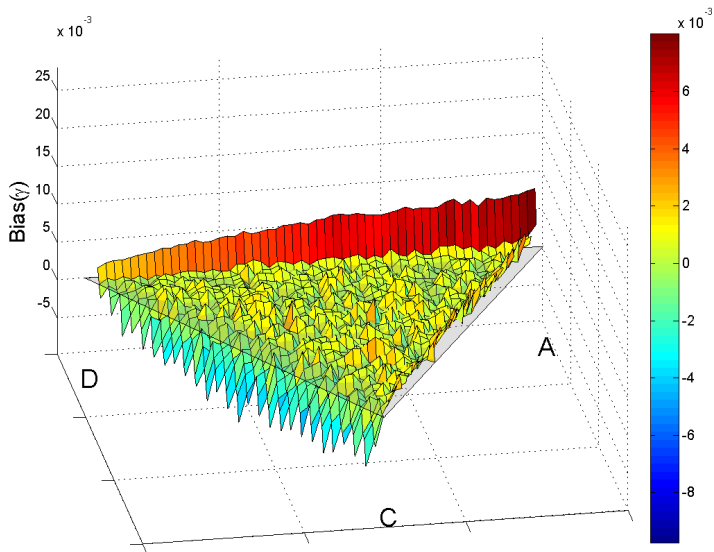
B.3 Abbildungen und Tabellen zur YCD

Abbildung B.9: Schätzfehler für γ der YCD ($N = 1.000$, $\beta = 0,1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

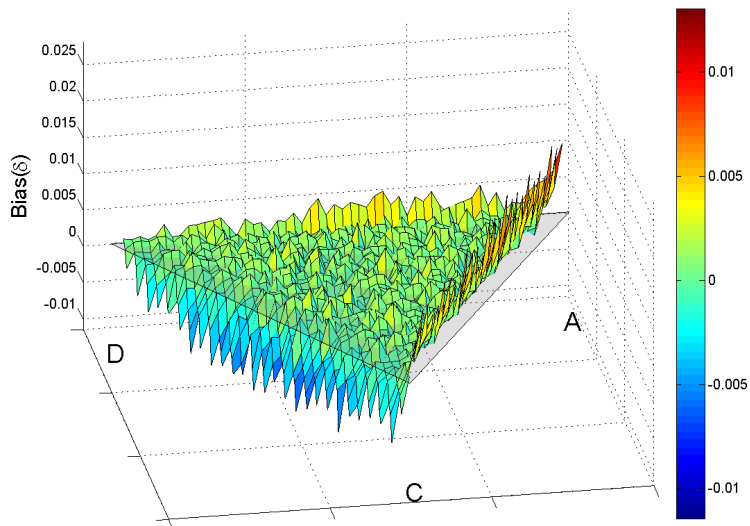


Abbildung B.10: Schätzfehler für δ der YCD ($N = 1.000$, $\beta = 0,1.000$ Wiederholungen). Darstellung in Dreieckskoordinaten gemäß 4.4.

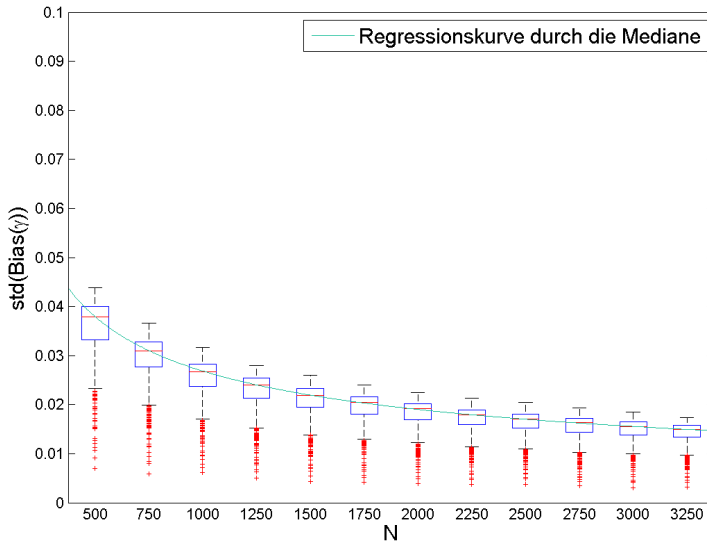


Abbildung B.11: Entwicklung der Standardabweichung des Schätzfehlers von γ bei der YCD gemäß 4.7.3. Erläuterungen zur Darstellung siehe Abbildung 4.4.

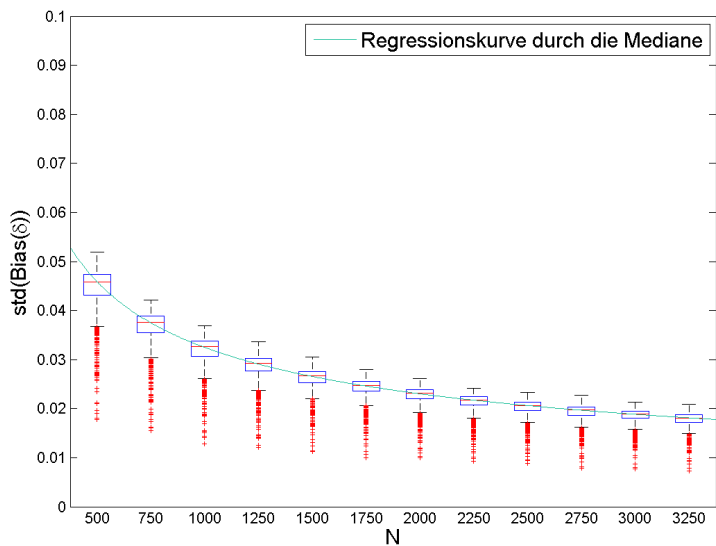


Abbildung B.12: Entwicklung der Standardabweichung des Schätzfehlers von δ bei der YCD gemäß 4.7.3. Erläuterungen zur Darstellung siehe Abbildung 4.4.

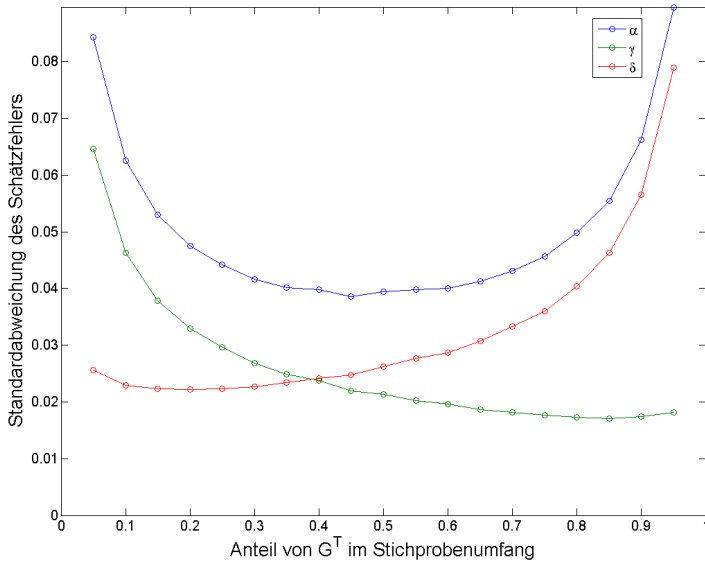


Abbildung B.13: Auswirkung unterschiedlicher Gruppenumfänge bei der YCD ($N = 1.500$).

Tabelle B.2: Zusammenhang der Standardabweichungen der Schätzfehler bei der YCD mit der Determinante der Verfahrensmatrix. Die Determinante ist exakt, die Quantile sind um den Faktor 1.000 skaliert und gerundet.

χ	$ \det M $	std (Bias ($\hat{\chi}$))			std (Bias ($\hat{\chi}$)) · $ \det M $		
		10%	50%	90%	10%	50%	90%
α	0,54	32,72	40,07	42,96	17,67	21,64	23,20
	0,40	43,59	54,69	58,30	16,56	20,78	22,15
	0,22	52,47	67,90	72,83	15,74	20,37	21,85
	0,16	68,07	91,90	100,30	12,93	17,46	19,06
γ	0,54	15,23	22,13	24,06	8,22	11,95	12,99
	0,40	21,72	28,42	30,55	8,25	10,80	11,61
	0,22	26,38	35,20	38,22	7,91	10,56	11,47
	0,16	33,77	45,63	50,46	6,42	8,67	9,59
δ	0,54	22,96	26,64	28,32	12,40	14,38	15,30
	0,40	28,04	33,60	35,51	10,66	12,77	13,50
	0,22	31,32	38,48	40,97	9,40	11,54	12,29
	0,16	39,87	51,66	55,76	7,58	9,82	10,59

C Literaturverzeichnis

- Abernathy, J. R., Greenberg, B. G. & Horvitz, D. G. (1970). Estimates of induced abortion in urban North Carolina. *Demography*, 7, 19-29.
- Ahart, A. M. & Sackett, P. R. (2004). A New Method of Examining Relationships Between Individual Difference Measures and Sensitive Behavior Criteria. Evaluating the Unmatched Count Technique. *Organizational Research Methods*, 7, 101-114.
- Akers, R. L., Massey, J., Clarke, W. & Lauer, R. M. (1983). Are Self-Reports of Adolescent Deviance Valid? Biochemical Measures, Randomized Response, and the Bogus Pipeline in Smoking Behavior. *Social Forces*, 62, 234-251.
- Antonak, R. F. & Livneh, H. (1995). Randomized response technique: A review and proposed extension to disability attitude research. *Genetic, Social & General Psychology Monographs*, 121 (1), 97.
- Barth, J. T. & Sandler, H. M. (1976). Evaluation of the randomized response technique in a drinking-survey. *Journal of Studies on Alcohol*, 37, 690-693.
- Becker, C. (2010). Exploring sensitive topics: Sensitivity, jeopardy and cheating. In H. Locarek-Junge & C. Weihs (Hrsg.), *Classification as a Tool for Research: Proceedings of the 11th IFCS Biennial Conference and 33rd Annual Conference of the Gesellschaft für Klassifikation* (S. 299-305). Springer: Berlin/Heidelberg.
- Böckenholt, U. & van der Heijden, P. G. M. (2007). Item Randomized-Response Models for Measuring Noncompliance. *Risk-Return Perceptions, Social Influences, and Self-Protective Responses. Psychometrika*, 72, 245-262.
- Böckenholt, U., Barlas, S. & van der Heijden, P. G. M. (2009). Do randomized-response designs eliminate response biases? An empirical study of non-compliance behavior. *Journal of Applied Econometrics*, 24 (3), 377-392.

- Bourke, P. D. (1984). Estimation of proportions using symmetric randomized response designs. *Psychological Bulletin*, 96(1), 166-172.
- Brewer, K. R. W. (1981). Estimating Marihuana Usage Using Randomized Response: Some Paradoxical Findings. *Australian Journal of Statistics*, 23 (1), 139-148.
- Chaloupka, M. Y. (1985). Application of the randomized response technique to marine park management: an assessment of permit compliance. *Environmental Management*, 9 (5), 393-398.
- Clarke, S. J. & Desharnais, R. A. (1998). Honest Answers to Embarrassing Questions: Detecting Cheating in the Randomized Response Model. *Psychological Methods*, 3, 160-168.
- Cohen, A. K. (1957). Kriminelle Subkulturen. In P. Heintz & R. König (Hrsg.), *Soziologie der Jugendkriminalität* (S. 103-117). Köln und Opladen. Sonderheft 2/1957 der Kölner Zeitschrift für Soziologie und Sozialpsychologie.
- Cohen, A. K. & Short, J. F. (1958). Research in Delinquent Subcultures. *Journal of Social Issues* 14 (3), 20-37.
- Coutts, E. & Jann, B. (2011). Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT). *Sociological Methods & Research*, 40 (1), 169-193.
- Coutts, E., Jann, B., Krumpal, I. & Näher, A.-F. (2011). Plagiarism in Student Papers: Prevalence Estimates Using Special Techniques for Sensitive Questions. *Jahrbücher für Nationalökonomie und Statistik*, 5-6, 749-760.
- Danermark, B. & Swensson, B. (1987). Measuring Drug Use among Swedish Adolescents. *Journal of Official Statistics*, 3, 439-448.
- Dawes, R. M., & Moore, M. (1980). Die Guttman-Skalierung orthodoxer und randomisierter Reaktionen. In F. Petermann (Hrsg.), *Einstellungsmessung, Einstellungsforschung* (S. 117-133). Göttingen, Toronto, Zürich: Verlag für Psychologie Hogrefe.
- Duffy, J. C. & Waterton, J. J. (1988). Randomize response versus direct questioning - Estimating the prevalence of alcohol-related in a field-survey. *Australian Journal of Statistic*, 30, 1-14.
- Edgell, S. E., Himmelfarb, S. & Duchan, K. L. (1982). Validity of forced responses in a randomized-response model. *Sociological Methods & Research*, 11, 89-100.
- Fidler, D. S. & Kleinknecht, R. E. (1977). Randomized response versus direct questioning - Two data-collection methods for sensitive information. *Psychological Bulletin*, 84 (5), 1045-1049.

- Fisher, M., Kupferman L. B. & Lesser M. (1992). Substance Use in a School-Based Clinic Population: Use of the Randomized Response Technique to Estimate Prevalence. *Journal of Adolescent Health, 13*, 281-285.
- Frenger, M., Pitsch, W. & Emrich, E. (2013). Ehrliche Antworten auf sensitive Fragen?! Entwicklung einer Befragungsmethode und erste Anwendungsplanung. In H. Kempf, S. Nagel & H. Dietl (Hrsg.), *Im Schatten der Sportwirtschaft* (Sportökonomie, 15, S. 241-252). Schorndorf: Hofmann.
- Goodstadt, M. S., Cook, G. & Gruson, V. (1987). The Validity of Reported Drug Use: The Randomized Response Technique. *International Journal of the Addictions, 13*, 359-367.
- Greenberg, B. G., Abul-Ela, A.-L. A., Simmons, W. R. & Horvitz, D. G. (1969). The Unrelated Question Randomized Response Model: Theoretical Framework. *Journal of the American Statistical Association, 64*, 520-539.
- Greenberg, B. G., Kuebler, R. R. jr., Abernathy, J. R. & Horvitz, D. G. (1971). Application of the Randomized Response Technique in Obtaining Quantitative Data. *Journal of the American Statistical Association, 66* (334), 243-250.
- Greenberg, B. G., Kuebler, R. R. jr., Abernathy, J. R. & Horvitz, D. G. (1977). Respondent hazards in the unrelated question randomized response model. *Journal of Statistical Planning and Inference, 1* (1), 53-60.
- Holbrook, A. L. & Krosnik, J. A. (2010). Measuring Voter turnout by Using the Randomized Response Technique. Evidence Calling into Question the Method's Validity. *Public Opinion Quarterly, 74*, 328-343.
- Holbrook, A. L. & Krosnik, J. A. (2012). Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique. *Public Opinion Quarterly, 74*, 37-67.
- Horvitz, D. G., Shah, B. V. & Simmons, W. R. (1967). The unrelated question randomized response model. *Proceedings of the American Statistical Association, Section on Survey Research Methodes*, 65-72.
- I-Cheng, C., Chow, L. P. & Rider, R. V. (1972). The randomized response techniques as used in the taiwan outcome and pregnancy study. *Studies in Family planning, 3*, 265-269.
- James, R. A., Nepusz, T., Naughton, D. P. & Petróczi, A. (2013). A potential inflating effect in estimation models: Cautionary evidence from comparing performance enhancing drug and herbal hormonal supplement use estimates. *Psychology of Sport and Exercise, 14* (1), 84-96.
- Jerke, J. & Krumpal, I. (2013). Plagiate in studentischen Arbeiten. Eine empirische Untersuchung unter Anwendung des Triangular Modells. *methoden, daten, analysen, 7* (3), 347-368.

- Kerkvliet, J. (1994). Cheating by economics students: A comparison of survey results. *Journal of Economic Education*, 25 (2), 121.
- Krotki, K. J. & Fox, B. (1974). The randomized response technique, the interview and the self-administered questionnaire - An empirical comparison of fertility report. *Proceedings of the American statistical Association, Social statistics section*, 367-371.
- Lee, R. M. (1993). *Doing Research on Sensitive Topics*. London u. a.: Sage.
- Lee, R. M. (2000). *Unobtrusive methods in social research*. Philadelphia: Open University.
- Lensvelt-Mulders, G. J. L., Hox, J. J., van der Heijden, P. G. M. & Maas, C. J. M. (2005). Meta-Analysis of Randomized Response Research. Thirty-Five Years of Validation. *Sociological Methods and Research*, 33 (3), 315-348.
- Musch, J., Bröder, A. & Klauer, K. C. (2001). Improving Survey Research on the World-Wide Web Using Randomized Response Technique. In U.-D. Reips und M. Bosnjak (Hrsg.), *Dimensions of Internet Science* (S. 179-192). Lengerich u.a.: Pabst.
- Pitsch, W. & Emrich, E. (2012). The Frequency of Doping in Elite Sport - Results of a Replication Study. *International Review for the Sociology of Sport*, 47, 559-580.
- Pitsch, W., Emrich, E. & Pierdzioch, C. (2013). Match Fixing im deutschen Fußball: Eine Validierung der Randomized-Response-Technik mit Total-Cheater-Detection mittels multinomialer Verarbeitungsbäume. In H. Kempf, S. Nagel & H. Dietl (Hrsg.), *Im Schatten der Sportwissenschaft (Sportökonomie, 15)* (S. 111-126). Schorndorf: Hofmann.
- Rasinski, K. A., Willis, G. B., Baldwin, A. K., Yeh, W. & Lee, L. (1999). Methods of data collection, perceptions of risks and losses, and motivation to give truthful answers to sensitive survey questions. *Applied Cognitive Psychology*, 13 (5), 465-484.
- Reckers, P. M. J., Wheeler, S. W. & Wong-On-Wing, B. (1997). A Comparative examination of audiotape premature sign-off using the direct and the randomized response methods. *Auditing: A Journal of practice and Theory*, 16, 69-78.
- Rider, R. V., Harper, P. A., Chow, L. P. & I-Cheng, C. (1976). A Comparison of Four Methods for Determining Prevalence of Induced Abortion. *American Journal of Epidemiology*, 103 (1), 37-50.
- Scheers, N. J. (1992). Methods, Plainly Speaking. A Review of Randomized Response Techniques. *Measurement and Evaluation in Counselling and Development*, 24, 27-41.

- Scheers, N. J. & Dayton, C. M. (1987). Improved Estimation of Academic Cheating Behavior Using the Randomized Response Technique. *Research in Higher Education*, 26, 61-69.
- Shotland, R. L. & Yankowski, L. D. (1982). The random response method - A valid and ethical indicator of the 'truth' in reactive situations. *Personality and Social Psychology Bulletin*, 8, 174-179.
- Simon, P., Striegel, H., Aust, F., Dietz, K. & Ulrich, R. (2006). Doping in fitness sport: estimated number of unreported cases and individual probability of doping. *Addiction*, 101, 1640-1644.
- Soeken, K. L. & Macready, G. B. (1986). Application of Setwise Randomized Response Procedures for Surveying Multiple Sensitive Attributes. *Psychological Bulletin*, 99 (2), 289-295.
- Stoeber, J. (2001). The social desirability scale-17 (SD-17). *European Journal of Psychological Assessment*, 17, 222-232.
- Thompson, E. R. & Phua, F. T. T. (2005). Reliability among senior managers of the Marlowe-Crowne short-form social desirability scale. *Journal of Business and Psychology*, 19, 541-554.
- Tourangeau, R. & Yan, T. (2007). Sensitive questions in Surveys. *Psychological Bulletin*, 133 (5), 859-883.
- Tracy, P. E. & Fox, J. A. (1981). The Validity of Randomized Response for Sensitive Measurements. *American Sociological Review*, 46, 187-200.
- van der Heijden, P. G. M., van Gils, G., Bouts, J., & Hox, J. J. (2000). A Comparison of Randomized Response, Computer-Assisted Self-Interview, and Face-to-Face Direct Questioning - Eliciting Sensitive Information in the Context of Welfare and Unemployment Benefit. *Sociological Methods & Research*, 28, 505-537.
- Warner, S. L. (1965). Randomized-response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60, 63-69.
- Williams, B. L. & Suen, H. (1994). A methodological comparison of survey techniques in obtaining self-reports of condom-related behaviors. *Psychological Reports*, 7, 1531-1537.
- Wolter, F. (2012). *Heikle Fragen in Interviews*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Will man herausfinden, wie häufig eine peinliche Eigenschaft in einer Population auftritt, gehen damit etliche Probleme einher, die vor allem auf verfälschte Antworten in Folge sozialer Erwünschtheit zurückgehen. Die älteste der Techniken, die entwickelt wurden um diese Verzerrungen zu kompensieren ist die Randomized Response-Technik, die auch derzeit noch die am weitesten verbreitete und am meisten evaluierte und erforschte Technik im Feld sensibler Themen darstellt. Obwohl diese Technik den Befragten vollständige Sicherheit garantiert und damit die Möglichkeit eröffnet, auch auf bedrohliche Fragen ehrlich zu antworten, tritt auch dabei „Cheating“ auf im Sinne der Nicht-Beachtung der Instruktionen. Der vorliegende Band beschreibt die mathematische Herleitung von Techniken zur Erkennung des Ausmaßes, in dem dieses „Cheating“ auftritt. Er stellt darüber hinaus Analysen zu Eigenschaften des Verfahrens dar und erweitert diese Analysen um praktische Hinweise zum flexiblen Einsatz dieser Verfahren. Dabei wird erstmals für unterschiedliche Formen des „Cheatings“ eine allgemeine Lösung entwickelt.