Schätzung von Veränderungen über die Zeit mittels koordinierter Stichprobenziehungen

Dissertation
Universität Trier -
Lehrstuhl für Wirtschafts- und Sozialstatistik

Zur Erlangung des Grades:
Dr. rer. pol

Eingereicht bei:
Herrn Professor Dr. Ralf Münnich
Herrn Professor Dr. Beat Hulliger

vorgelegt von:
Bernhard Stefan Zins
geboren am 01.06.1982
in Saarbrücken

April 2015

Vorwort

An dieser Stelle möchte ich meinem Erstgutachter Prof. Dr. Ralf Münnich danken. Er hat mich für eine Forschungskarriere im Bereich Survey Statistik begeistert und mir letztlich zu den Weg zu dieser Dissertation geebnet. Durch meine Mitarbeit in dem von Prof. Dr. Ralf Münnich geleiteten EU Forschungsprojekt AMELI am Lehrstuhl für Wirtschafts- und Sozialstatistik der Universität Trier wurde die Idee für dieser Dissertation geboren und die Kontakte zu den Partnern des Projekts hatten einen entscheidenden Anteil an der Bearbeitung der Forschungsfragestellung.

Zudem möchte mich bei meinem Zweitgutachter Prof. Dr. Beat Hulliger bedanken. Dies gilt insbesondere für seine Unterstützung und die belebenden Diskussionen aus denen wichtige Aspekte dieser Arbeit erwachsen sind.

Mein ganz besonderer Dank gilt auch Dr. Siegfried Gabler, der mich als mein Teamleiter, nach meinem Wechsel zum GESIS Leibniz-Institut für Sozialwissenschaften, tatkräftig unterstützt hat und mir half diese Arbeit zu einem erfolgreichen Abschluss zu bringen.

Weiterhin möchte ich meinen früheren Kollegen vom Lehrstuhl Wirtschafts- und Sozialstatistik, sowie meinen aktuellen Kollegen am GESIS Leibniz-Institut für Sozialwissenschaften, für die große Unterstützung und zahlreichen Diskussionen bedanken. Mein besonderer Dank gilt hier Christian Bruch, Tobias Enderle, Jan-Philipp Kolb und Matthias Sand.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Pseudocodeverzeichnis

Symbolverzeichnis

Allgemeines Symbolverzeichniss

Symbole	Beschreibung
$\vec{x}$	Vektor x
$\vec{x}^{\top}$	transponierter Vektor x
$\in$	in
$\notin$	nicht in
$\forall$	für alle
$\vee$	nicht-ausschließendes oder
$\dot{\vee}$	ausschließendes oder
$\wedge$	und
$\exists$	existiert
$\approx$	approximativ
$\binom{N}{n}$	Binomialkoeffizient N über n
$card(.)$	Mächtigkeit
$\min(.)$	Minimum
$\max(.)$	Maximum
$\lfloor . \rfloor$	nächst kleinere ganze Zahl
$\lceil . \rceil$	nächst größere ganze Zahl
$ggT(.)$	größter gemeinsamer Teiler
$n \bmod m$	Rest der Division n geteilt durch m
$\mathbb{N}$	Menge aller natürlichen Zahlen
$\mathbb{N}_0$	Menge aller natürlichen Zahlen mit Null
$\mathbb{Z}$	Menge aller ganzen Zahlen
$\mathbb{R}$	Menge aller reellen Zahlen
$\mathscr{P}(.)$	Potenzmenge
$\emptyset$	Leere Menge
$inv(.)$	Inverse
$\mathbf{I}_N$	$N \times N$ Einheitsmatrix
$diag(\vec{x})$	Diagonalmatrix mit $\vec{x}$ als Hauptdiagonale
$\circ$	Hadamard-Produkt
$o(x)$	asymptotisch gegenüber x vernachlässigbar
$O_p(.)$	asymptotische obere Schranke
$\xrightarrow{d}$	konvergiert in Verteilung
$\xrightarrow{p}$	konvergiert in Wahrscheinlichkeit
$\frac{\partial f}{\partial x}$	Partielle Ableitung der Funktion f nach x

$Pr(.)$	Wahrscheinlichkeit
$\mathbb{1}(.)$	Indikatorfunktion
$\sim$	verteilt nach
$\mathrm{E}(.)$	Erwartungswert
$\mathrm{V}(.)$	Varianz
$\widehat{\mathrm{V}}$	Varianzschätzer
$\mathrm{COV}(.,.)$	Kovarianz
$\widehat{\mathrm{COV}}$	Kovarianzschätzer
$NV(x,y)$	Normalverteilung mit Erwartungswert x und Varianz y
$\mathrm{Unif}(x,y)$	Gleichverteilung zwischen x und y
$\triangle$	Ende Beispiel

Spezielles Symbolverzeichniss

Symbole	Beschreibung	Erste Nennung
$\mathscr{U}$	Population bzw. Ziehungsrahmen über den gesamten Beobachtungszeitraum	S. 10 bzw. 12
N	Umfang der Population	S. 10 bzw. 12
N^t	Umfang der Population zum Zeitpunkt t	S. 12
T	Anzahl der Beobachtungszeitpunkte / Funktional	S. 12 / S. 76
$\mathscr{U}^t$	Population bzw. Ziehungsrahmen zum Zeitpunkt t	S. 12
$\mathscr{B}^t$	Geburten in Zeitpunkt t	S. 45
$\mathscr{D}^t$	Sterbefalle in Zeitpunkt t	S. 45
$\mathscr{C}^{\alpha\omega}$	Kohorte deren Elemente der Population zum Zeitpunkt α beitreten und sie zum Zeitpunkt ω verlassen, bzw. für $\omega = T+1$ sie nie verlassen	S. 47
$N_{\mathscr{C}^{\alpha\omega}}$	Anzahl der Elemente in Kohorte $\mathscr{C}^{\alpha\omega}$	S. 48
$\mathfrak{S}$	gemeinsame Stichprobe	S. 13
$\mathbf{S}$	Ausprägung der gemeinsamen Stichprobe	S. 13
$\mathscr{s}$	gemeinsame Stichprobe als Indexmenge	S. 10 bzw. 12
$\vec{\mathfrak{s}}^t$	Querschnittstichprobe zum Zeitpunkt t	S. 12
$\vec{s}^t$	Ausprägung der Querschnittstichprobe zum Zeitpunkt t	S. 13
$\mathscr{s}^t$	Querschnittstichprobe als Indexmenge zum Zeitpunkt t	S. 12
$\vec{\mathfrak{s}}_k$	Längsschnittstichprobe des k-ten Elements	S. 13
$\vec{s}_k$	Ausprägung der Längsschnittstichprobe des k-ten Elements	S. 13
$p(.)$	gemeinsames Stichprobendesign	S. 10 bzw. S. 13
$p^t(.)$	Querschnittdesign zum Zeitpunkt t	S. 13
$p_k(.)$	Längsschnittdesign des k-ten Elements k	S. 13
$\mathfrak{I}_k^t$	Stichprobenindikator des k-ten Elements zum Zeitpunkt t	S. 13
I_k^t	Ausprägung des Stichprobenindikators des k-ten Elements zum Zeitpunkt t	S. 13
π_k^t	Inklusionswahrscheinlichkeit des k-ten Elements zum Zeitpunkt t	14
$\pi_{k,l}^t$	gemeinsame Inklusionswahrscheinlichkeit des k-ten und l-ten Elements zum Zeitpunkt t, mit $k \neq l$	S. 14
$\pi_k^{t,u}$	gemeinsame Inklusionswahrscheinlichkeit des k-ten Elements zum Zeitpunkt t und Zeitpunkt u, mit $t \neq u$	S. 14

$\pi_{k,l}^{t,u}$	gemeinsame Inklusionswahrscheinlichkeit des k-ten Elements zum Zeitpunkt t und des l-ten Elements zum Zeitpunkt u, mit $k \neq l$ und $t \neq u$	S. 14
ϕ_k^t	nächster Ziehungszeitpunkt von Element k nach seiner Ziehung zum Zeitpunkt t	S. 17
$\mathfrak{o}_k^t$	Koordinationsvariable des k-ten Elements zum Zeitpunkt t	S. 25
$\mathfrak{p}_k^t$	Zeit außerhalb der Stichprobe des k-ten Elements nach der Ziehung zum Zeitpunkt t	S. 26
$\mathfrak{b}_k^t$	Kumulierte Erhebungslast des k-ten Elements nach der Ziehung zum Zeitpunkt t	S. 17
$\mathscr{U}_{(i)}^t$	Menge von Elementen der Population bzw. des Ziehungsrahmen zum Zeitpunkt t mit dem i-te höchsten Werte der Koordinationsvariable	S. 25
K^t	Anzahl unterschiedlicher Ausprägungen der Koordinationsvariable nach der Ziehung in Zeitpunkt t	S. 26
K_l^{t-1}	Anzahl unterschiedlicher Ausprägungen der Koordinationsvariable der Überlebenden bis Zeitpunkt t $\mathscr{U}^{t-1} \setminus \mathscr{D}^t$, vor der Ziehung in Zeitpunk t	S. 46
$O_{(k)}^t$	Ordnungsstatistik der Koordinationsvariable des k-ten Elements zum Zeitpunkt t	S. 25
τ	Totalwert	S. 60
$\hat{\tau}_\pi$	*Horvitz-Thompson* Schätzer für τ	S. 60
$IF(.)$	Einflussfunktion	S. 76
$\mathfrak{N}_{c^{\alpha\omega}(i)}^t$	Menge von Elementen in Kohorte $\mathscr{C}^{\alpha\omega}$ zum Zeitpunkt t mit dem i-te höchsten Werte der Koordinationsvariable	S. 47
$\mathfrak{n}_{c^{\alpha\omega}(i)}^t$	Anzahl der Elementen aus Kohorte $\mathscr{C}^{\alpha\omega}$ mit dem i-te höchsten Werte der Koordinationsvariable zum Zeitpunkt $t-1$, die in die Stichprobe zum Zeitpunkt t gelangen	S. 47
$\mathfrak{n}_{c^{\alpha\omega}}^t$	Anzahl der Elementen aus Kohorte $\mathscr{C}^{\alpha\omega}$, die in die Stichprobe zum Zeitpunkt t gelangen	S. 48
$\vec{\mathfrak{N}}_{c^{\alpha\omega}}^t$	Zufallsvektor mit allen Möglichen Ausprägungen von $\mathfrak{N}_{c^{\alpha\omega}(i)}^t$	S. 48
$\vec{\mathfrak{n}}_{c^{\alpha\omega}}^t$	Zufallsvektor mit allen Möglichen Ausprägungen von $\mathfrak{n}_{c^{\alpha\omega}(i)}^t$	S. 48
$g^{t,u}(x,y)$	Veränderungsmaß von x und y zwischen den Zeitpunkten t und u	S. 91

Abkürzungsverzeichnis

iid	unabhängig identisch certeilt
SRS	einfache Zuflassstichprobe
wr	mit Zurücklegen
EU-SILC	The European Union Statistics on Income and Living Conditions
ARP	Von Armut bedrohte Personen, nach Sozialleistungen
ARPR	Armutsgefährdungsquote
LWI	In Haushalten mit sehr niedriger Erwerbstätigkeit lebende Personen
DEP	Unter erheblicher materieller Deprivation leidende Personen
AROPE	Von Armut oder sozialer Ausgrenzung bedrohte Personen
AROPER	Rate von Armut oder sozialer Ausgrenzung bedrohte Personen
QSR	Quintile Share Ratio
GINI	Gini-Koeffizient
IPF	Iterative Proportional Fitting
PRN	Permanent Random Numbers
lGREG	linearer generalisierter Regressionsschätzer

Kapitel 1

Einleitung

1.1 Messung von Veränderungen über die Zeit mittels Stichprobenerhebungen

Ob in Politik oder Wirtschaft, Prozesse zur Entscheidungsfindung werden stark durch die Werte quantitativer Indikatoren, d.h. von Statistiken, gelenkt. Entsprechend werden auch Ergebnisse, die diesen Entscheidungen zugeschrieben werden, bezüglich dieser Indikatoren evaluiert. Darum ist eine der vorrangigen Aufgaben nationaler statistischer Ämter auch die Bereitstellung von Zeitreihen für einige als wichtig erachtete Indikatoren, wie z.B. die Aggregate der volkswirtschaftlichen Gesamtrechnung oder die Arbeitslosenstatistik und Erwerbsquote. Einige der meist beachteten Erhebungen, ob national oder international, werden regelmäßig wiederholt.[1] Erst durch die Beobachtung einer Population über einen Zeitraum lassen sich Veränderungen erkennen. So steigt gerade mit der Länge des Beobachtungszeitraums einer Erhebung deren Bedeutung als Datengrundlage in Forschung, Politik und Wirtschaft.

Zur Quantifizierung ihrer sog. Europa-2020-Ziele verwendet die Europäische Kommission eine Liste von Indikatoren, die Statistiken aus den Bereichen Beschäftigung, Forschung und Entwicklung, Klima/Energie, Bildung, sozialer Eingliederung und Armutsbekämpfung umfassen (Europäische Kommission, 2011). Die Indikatoren zur Messung sozialer Eingliederung und Armutsbekämpfung werden auf Grundlage der EU-Statistik über Einkommen und Lebensbedingungen (EU-SILC) gemessen. Die Erhebung zu EU-SILC ist eine jährlichen Stichprobe in den EU-Mitgliedstaaten sowie in Island, Norwegen, der Schweiz und der Türkei.[2]

---

[1] Beispiele wären:

The European Union Labour Force Survey (EU LFS)

The European Union Statistics on Income and Living Conditions (EU-SILC)

Der Mikrozensus

Die allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS)

[2] EU-SILC wird seit 2005 durchgeführt, jedoch nicht durchgehend für alle der jetzigen EU Mitgliedstaaten, bzw. der nicht EU-Länder (Eurostat, 2015).

Die drei Indikatoren, die zur Messung von Armut und sozialer Ausgrenzung herangezogen werden sind wie folgt definiert:

Definition 1.1. *Von Armut bedrohte Personen, nach Sozialleistungen* (ARP): „Personen mit einem verfügbaren Äquivalenzeinkommen unterhalb der Armutsgefährdungsschwelle, die bei 60% des nationalen verfügbaren Medianäquivalenzeinkommens (nach Sozialtransfers) liegt.“ (Eurostat, 2014b)

Definition 1.2. *In Haushalten mit sehr niedriger Erwerbstätigkeit lebende Personen* (LWI): „Personen im Alter von 0-59 Jahren, die in Haushalten leben, in denen die Erwachsenen (18-59 Jahre) im vorhergehenden Jahr insgesamt weniger als 20% gearbeitet haben.“ (Eurostat, 2014d)

Definition 1.3. *Unter erheblicher materieller Deprivation leidende Personen* (DEP): Personen die nicht in der Lage sind, für mindestens drei der folgenden neun Ausgaben aufzukommen: Miete und Versorgungsleistungen, angemessene Beheizung der Wohnung, unerwartete Ausgaben, jeden zweiten Tag eine Mahlzeit mit Fleisch, Fisch oder gleichwertiger Proteinzufuhr, einen einwöchigen Urlaub an einem anderen Ort, ein Auto, eine Waschmaschine, einen Farbfernseher oder ein Telefon (Eurostat, 2014c).

Der Leitindikator im Bereich Armutsbekämpfung ist eine Kombination aus den Indikatoren ARP, LWI und DEP und definiert als:

Definition 1.4. *Von Armut oder sozialer Ausgrenzung betroffene oder bedrohte Personen* (AROPE): Personen, die armutsgefährdet sind oder unter materieller Deprivation leiden oder in Haushalten mit sehr niedriger Erwerbstätigkeit leben (Eurostat, 2014a).

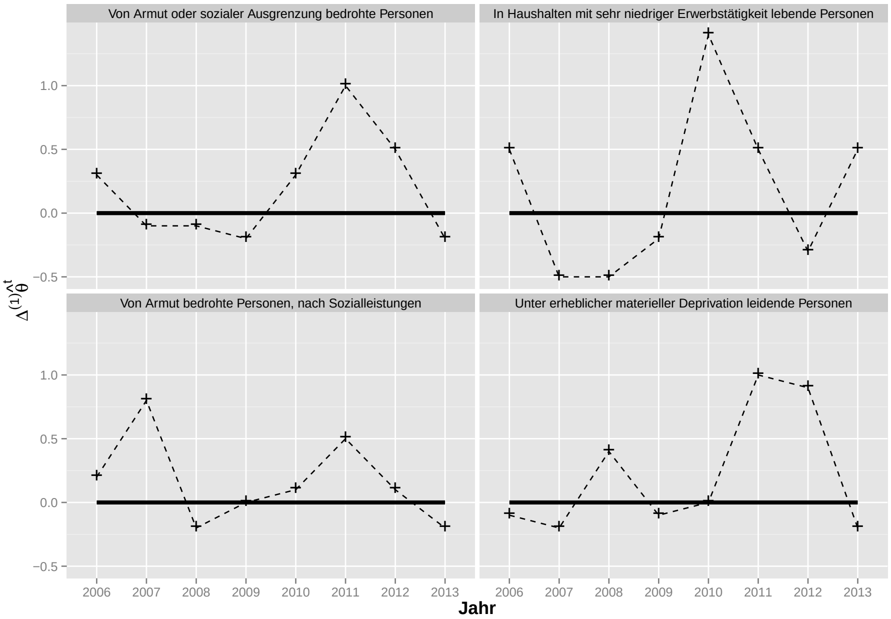
Ziel ist es, den Wert des Indikators AROPE bis zum Jahr 2020 um mindestens 20 Millionen zu senken (Europäische Kommission, 2011). Dies bedeutet, dass insbesondere die Entwicklung im Zeitablauf der obigen Indikatoren von Interesse ist. Gilt es doch zu evaluieren, ob und wenn ja in welchem Umfang sich die fraglichen Statistiken ihren Zielvorgaben nähern, um so künftige politische Entscheidungen zu planen. Da AROPE mittels eines Schätzers basierend auf den Daten der EU-SILC Erhebung gemessen wird, besteht bei jeder Bewertung berichteter Werte die Gefahr, diese zu überinterpretieren, sofern der Umstand außer Acht gelassen wird, dass es sich hierbei um Stichproben-basierte Schätzungen handelt.[3] Eine beobachtete Veränderung eines Indikator über die Zeit ist möglicherweise hauptsächlich auf die Stichprobenvarianz der Schätzungen zurückzuführen (Osier, Berger & Goedeme, 2013). Folglich sollte weniger die absolute Veränderung in den gemessenen Werten interpretiert, sondern vielmehr die Veränderung auf deren statistische Signifikanz getestet werden.

Abbildung 1.1 stellt die Entwicklung der vier Indikatoren ARP, DEP, LWI und AROPE für die Gruppe der Länder dar, die bis einschließlich 2011 den Euro als ihre offizielle Währung eingeführt hatten (EA-17 Gruppe). Dabei wird die Entwicklung als Differenz zum Vorjahr dargestellt. Ist $\hat{\theta}^t$ die Schätzung eines Indikators θ zum Zeitpunkt t, so ist $\Delta^{(1)}\left(\hat{\theta}^t\right)$ die geschätzte Veränderung zum Vorjahr, mit $\Delta^{(l)}\left(\hat{\theta}^t\right) = \hat{\theta}^t - \hat{\theta}^{t-l}$. Bei der Betrachtung der Zeitreihen fällt auf, dass die Veränderungen für die Indikatoren ARP,

---

[3] Andere Fehlerquellen sind fehlende Werte sog. *Nonresponse*, imperfekte Ziehungsrahmen oder Messfehler (Särndal, Swensson & Wretman, 1992). Diese Fehlerquellen können ebenfalls eine Rolle spielen bei Statistiken, die auf Vollerhebungen oder Auswertungen von Registern beruhen.

Abbildung 1.1: Veränderung in den Europa-2020 Indikatoren für Armut und soziale Ausgrenzung der EA-17 Gruppe



Quelle: Eurostat, 2015, http://ec.europa.eu/eurostat/web/europe-2020-indicators/europe-2020-strategy/main-tables

DEP und AROPE ab dem Jahr 2009 positiv bis 2011 sind, sowie bis 2010 für den Indikator LWI. Da LWI als ein Arbeitsmarktindikator interpretiert werden kann, ist zu erwarten, dass seine Entwicklung den der Indikatoren ARP und DEP vorgelagert ist, da diese unmittelbar vom Haushaltseinkommen abhängen. Die Hypothese wäre nun, dass sich die aufgrund der Ereignisse der Jahre 2008 und 2009 ausgelösten Wirtschaftskrise (Sinn, 2012) der Wert dieser Indikatoren zwischen den Jahren 2009 und 2011 bzw. 2010 gestiegen ist.

Für einen Indikator θ könnte demnach die folgende Nullhypothese

$$H_0 : \Delta^{(l)}(\hat{\theta}^t) \leq 0$$

gegen die Alternativhypothese

$$H_1 : \Delta^{(l)}(\hat{\theta}^t) > 0$$

getestet werden, mit $l = 2$ und $t = 2011$, oder $l = 1$ und $t = 2010$. H_0 wäre auf einem Signifikanzniveau von 5% abzulehnen, wenn

$$0 \notin \left(-\infty; \Delta^{(l)}(\hat{\theta}^t) - 1{,}645 sd_{\Delta^{(l)}(\hat{\theta}^t)}\right] . \tag{1.1}$$

Dabei ist $sd_{\Delta^{(l)}(\hat{\theta}^t)}$ die Standardabweichung des Schätzers für die Veränderung. Die Testentscheidung in (1.1) beruht auf der Annahme, dass $\Delta^{(l)}(\hat{\theta}^t)$ für die gegebenen

Stichprobenumfänge zu den Zeitpunkten t und $t-l$ approximativ normal verteilt ist. Kann H_0 abgelehnt werden, so wird die beobachtete Veränderung als signifikant positiv angesehen.

1.2 Problematik bei der statistischen Inferenz für Veränderungsmaße

Es ist davon auszugehen, dass $sd_{\Delta^{(l)}(\hat{\theta}^t)}$ unbekannt ist und zur Testentscheidung in (1.1) aus der Stichprobe geschätzt werden muss. Ein möglicher Schätzer für die Standardabweichung ist $\sqrt{\widehat{\mathrm{V}}\left(\Delta^{(l)}(\hat{\theta}^t)\right)}$, mit $\widehat{\mathrm{V}}\left(\Delta^{(l)}(\hat{\theta}^t)\right)$ als Varianzschätzer für $\Delta^{(l)}(\hat{\theta}^t)$.

Zwei Probleme ergeben sich bei der Varianzschätzung für Veränderungsmaße wie $\Delta^{(l)}(\hat{\theta}^t)$:

1. Die relevanten Statistiken, sprich die Armutsindikatoren, sind teilweise hochgradig nicht linear. Dies hat zur Folge, dass einfache Methoden zur Varianzschätzung, wie sie bei linearen Schätzen zur Anwendung kommen, nicht ohne Weiteres verwendbar sind. Auf diese Problematik wird insbesondere in Kapitel 3 eingegangen.
2. Wiederholt durchgeführte Erhebungen, wie EU-SILC, haben oft einen Längsschnittaspekt, der zu korrelierten Stichproben im Querschnitt führt. Diese Thematik der sog. Stichprobenkoordination wird in Kapitel 2 behandelt.

Zusammen bedeutet dies, dass für $\widehat{\mathrm{V}}\left(\Delta^{(l)}(\hat{\theta}^t)\right)$ die Kovarianz zwischen zwei nicht linearen, und möglicherweise nicht glatten Schätzern, zu schätzen ist. Entsprechend dieser Aufgaben sind die Kapitel der vorliegenden Arbeit strukturiert.

1.3 Literaturübersicht und Einordnung der Arbeit

Der folgende Abschnitt dient dazu, einen Übersicht über die relevante Literatur in den behandelten Themen der Arbeit zu geben, als auch die Arbeit bezüglich der bestehenden Literatur einzuordnen und abzugrenzen.

1.3.1 Kapitel 2

Um die Eigenschaften korrelierter Stichproben zu erarbeiteten, beschäftigt sich Kapitel 2 mit Koordination von Stichprobenziehungen in einer zeitlichen Abfolge. Das wiederholte Ziehen von Stichproben aus derselben[4] Population in einem zeitlichen Ablauf verlangt oftmals einen Ausgleich zwischen Zielen, die miteinander in Konflikt stehen.

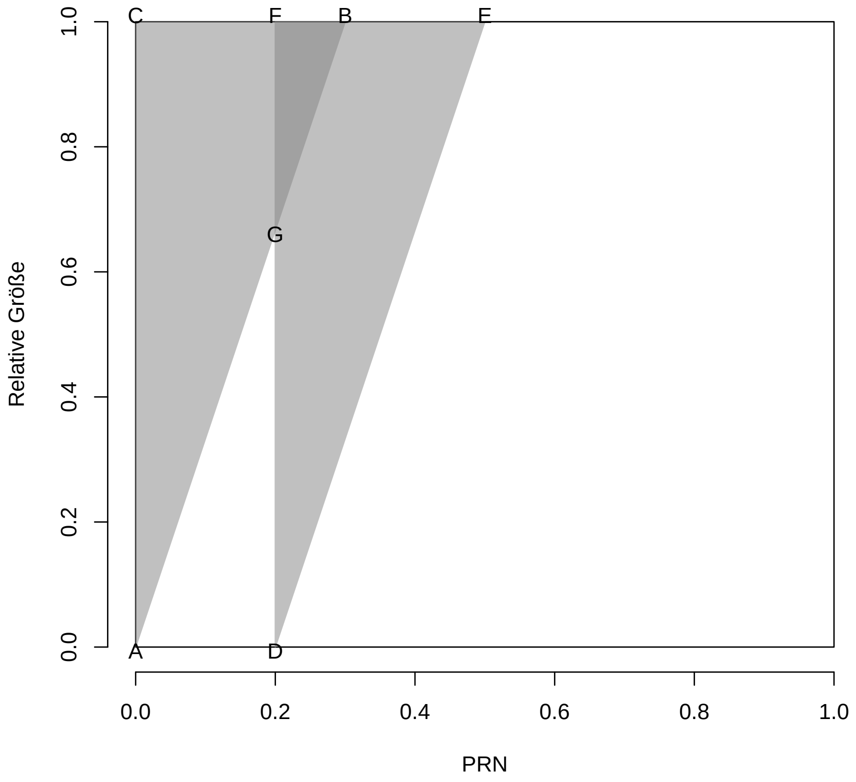
---

[4] Unter derselben Population ist zu verstehen das sich nur die zeitliche Abgrenzung der Population veränderte, sie ansonsten aber gleich definiert bleibt.

Für eine effiziente Schätzung von Veränderungen im Zeitablauf ist eine möglichst hohe Schnittmenge zwischen einzelnen Stichproben wünschenswert. In diesem Zusammenhang wird von einer positiven Koordination von Stichproben gesprochen. Demgegenüber sollen die Stichproben zu jedem Zeitpunkt aber auch erwartungstreue Schätzungen bezüglich der Parameter der aktuellen Populationszusammensetzung ermöglichen. Hierfür würde die Ziehung einer neuen Stichprobe sprechen, wobei es wünschenswert ist, dass diese eine möglichst kleine Schnittmenge mit der vorangegangenen Stichprobe hat, d.h. sie soll zu dieser negativ koordiniert sein. Ein ähnlicher Zielkonflikt besteht ebenfalls bei dem Wunsch, Elemente über eine lange Zeitspanne zu beobachteten und die Erhebungslast möglichst gleich auf die Elemente einer Population zu verteilen.

Eine Möglichkeit, einen Ausgleich zwischen diesen Zeilen zu finden, ist die Verwendung von sog. *Permanent Random Numbers* (PRN). Im Zusammenhang mit Unternehmenserhebungen schlagen Brewer et al. (1972) die Verwendung von PRN zur Koordination von Stichproben mit großen proportionalen Inklusionswahrscheinlichkeiten vor. Die von Brewer et al. (1972) vorgestellte Methode weist allen Elementen in der Population eine PRN als Ziehung aus einer stetigen Gleichverteilung auf dem Intervall $[0, 1]$ zu. Die relative Größe der Elemente (relativ in Bezug auf die Gesamtgröße) und die PRN werden dann als Punkte innerhalb eines Einheitsquadrates in einem kartesischen Koordinatensystem dargestellt. Abbildung 1.2 stellt ein Beispiel für ein PRN Koordination dar. Es werden zwei Stichproben in Abbildung 1.2 dargestellt. Die erste

Abbildung 1.2: PRN Koordination nach Brewer et al. (1972)



Stichprobe ist gegeben durch das Dreieck mit den Punkten A,B und C und die zweite Stichprobe ist gegeben durch das Dreieck mit den Punkten D, E und F. Die Schnittmenge der beiden Stichproben ist bestimmt durch das Dreieck mit den Punkten G, B und F. Durch die Wahl anderer Dreiecke, beispielsweise als weitere Parallelverschiebungen der ersten Stichprobe, lassen sich beliebige Schnittmengen zwischen unterschiedlichen Stichproben erzeugen. Wird der rechte Rand des Einheitsquadrates überschritten wird der Teil des Dreieckes der außerhalb des Einheitsquadrates liegt wieder von linkes eingefügt und umgekehrt, d.h. die Dreiecke bzw. die Stichproben sind zirkulant.

Das von Brewer et al. (1972) beschrieben Stichprobendesign ist zu jedem Zeitpunkt ist ein Poisson Design (Hájek, 1964). Somit sind die Stichproben ohne Zurücklegen gezogen. Die Flexibilität bezüglich der Koordination stellt wohl die nützlichste Eigenschaft der Koordination mit PRN dar. Denn sie erlaubt, den Erwartungswert der Schnittmenge einer Stichprobe mit einer Stichprobe zu einem anderen Zeitpunkt frei zu bestimmen. In diesem Zusammenhang wird auch von positiver bzw. negativer Koordination von Stichproben gesprochen. Eine positive Koordination liegt zumeist vor wenn die Schnittmenge zwischen zwei Stichproben maximiert wird, eine negative wenn diese minimiert wird. Durch die Flexibilität der PRN Methode kann dem Wunsch von sowohl negativer und positiver Koordination durch eine Kombination von verschiedenen mit PRN koordinierten Stichproben nachgekommen werden. So entwickelt Qualité (2009) einen sequenziellen Ziehungsalgorithmus basierend auf der Methode von Brewer et al. (1972), der negative und positive Koordination über die Zeit von verschiedenen Stichproben erlaubt, mit beliebigen Inklusionswahrscheinlichkeiten erster Ordnung (siehe hierzu auch Salamin, 2009). Nedyalkova, Qualité & Tillé (2009) geben eine weitere Möglichkeit für einen sequenziellen Ziehungsalgorithmus nach der Methode von Brewer et al. (1972) an. Ihr Algorithmus implementiert aber eine durchweg negative Koordination zwischen aufeinander folgenden Stichproben. Eine Modifikation des Verfahren schlagen auch Kröger, Särndal & Teikari (1999) vor. Ihr Ansatz besteht in einer Mischung aus Bernoulli[5] und Poisson Designs. Ein Teil der Population wird bei dieser Methode durch eine Bernoulli Ziehung erhoben. So wird erreicht, dass ein größerer Anteil von kleinen Elementen in die Stichprobe gelangt als dies bei einem reinen Poisson Design mit größenproportionaler Auswahl der Fall wäre.

Ein weiteres wichtiges Merkmal der PRN Methode besteht darin, dass Geburten und Sterbefälle sehr einfach zu handhaben sind. Die PRN von Sterbefällen werden einfach entfernt und Geburten bekommen eine neue PRN zugewiesen. Die Verwendung des Poisson Designs hat zudem den Vorteil, dass sich einfache Formen für erwartungstreue Punkt- und Varianzschätzer finden lassen. Gleichzeitig bringt dies aber auch den größten Nachteil der Methode mit sich, zufällige Stichprobenumfänge. So hat beispielsweise auch die leere Menge eine positive Wahrscheinlichkeit, als Stichprobe ausgewählt zu werden. Für entsprechend große Populationen mag dies zu vernachlässigen sein, bei kleinen Stichprobenumfängen und kleinen Populationen ist dieses Problem jedoch nicht zu ignorieren. Die Verteilung des Stichprobenumfangs ist in Tillé (2006, Abschnitt 5.5) beschrieben.

Eine weitere Variante von Stichprobendesigns mit PRN ist das sog. geordnete größenproportionale Ziehungsverfahren, entwickelt von Rosén (1997a,b). Bei dieser Variante

---

[5] Ein Bernoulli Design ist ein Poisson Design mit gleichen Inklusionswahrscheinlichkeiten für alle Elemente in der Population (Särndal et al., 1992, S. 62f)

sind die PRN auch unabhängig voneinander, aber nicht notwendigerweise gleichverteilt. Die sequenzielle Poisson Stichprobenziehung (Ohlsson, 1998) kann als spezieller Fall dieses Verfahrens dargestellt werden (siehe auch Matei & Tillé, 2007). Kröger, Särndal & Teikari (2003) stellen zudem eine Kombination aus geordneter Auswahl und ihres in Kröger et al. (1999) präsentierten Verfahrens vor.

Aufgrund ihrer Einfachheit und breiten Anwendungsmöglichkeit hat sich die PRN Technik als Methode zur Gestaltung von Stichprobendesigns wiederholter Erhebungen bei verschiedenen statistischen Ämtern etabliert. Einen Überblick liefert Hesse (1999). Es finden sich aber auch Anwendungen bei Haushaltsstichproben (siehe z.B. Qualité, 2011). Für eine Anwendung von PRN Koordination bei der Erhebung von Mietpreisen siehe Hulliger (1995).

Stichproben mit fixen Umfängen lassen sich bei gleichen Inklusionswahrscheinlichkeiten ebenfalls mit PRN ziehen. Hierzu werden die Elemente in aufsteigender oder absteigender Ordnung ihrer PRN gelistet. Wie Ohlsson (1992) zeigt, entspricht die Auswahl der ersten n Elemente der Liste einer einfachen Zufallsstichprobe, genau wie jedes andere Intervall, das genau n Elemente enthält. Wie Ohlsson (1995) beschreibt, können auch geschichtete Stichprobendesigns mittels PRN negativ und positiv koordiniert werden (siehe auch Nordberg, 2000). Eine intensive Auseinandersetzung mit der Problematik, geschichtete Stichprobendesigns über die Zeit zu koordinieren, findet sich in Nedyalkova (2009).

Die Bestimmung der Inklusionswahrscheinlichkeiten höherer Ordnung für wiederholte Erhebungen stellt sich mitunter schwieriger dar. Insbesondere gilt dies für die Wahrscheinlichkeiten, dass das gleiche Element zu zwei verschiedenen Zeitpunkten in die Stichprobe gelangt, und für die Wahrscheinlichkeiten, dass zwei verschiedene Elemente zu unterschiedlichen Zeitpunkten gezogen werden. Für koordinierte einfache Zufallsstichproben bestimmt Tam (1984) diese Wahrscheinlichkeiten. Laniel (1987) erweitert diesen Ansatz für veränderliche Populationen. In beiden Fällen werden nur Stichproben mit konstanten Schnittmengen bzw. konstanten Anteilen der Schnittmengen betrachtet (siehe hierzu auch Hidiroglou, Särndal & Binder, 1995, Abschnitt 25.4, und Forsman & Gåras, 1982, sowie Hulliger, 1995). Zumindest eines der von Tam (1984) verwendeten Stichprobendesigns kann auch als eine PRN Koordination beschrieben werden.

Ein möglicher Nachteil bei der Koordination mit PRN besteht darin, dass durch einmaliges Festlegen der PRN, bei einer unveränderlichen Population, auch alle weiteren Ziehungen festgelegt sind. Aus diesem Grund wird in der vorliegenden Arbeit ein Ziehungsalgorithmus eingeführt, der diesbezüglich mehr Flexibilität erlaubt und dessen Inklusionswahrscheinlichkeiten höherer Ordnung sich ebenfalls exakt bestimmen lassen, auch wenn sie sich analytisch komplex darstellen. Der vorgeschlagene Algorithmus sortiert die Population über eine Koordinationsvariable in Abhängigkeit von der zuletzt erfolgten Stichprobenziehung. Nach jeder Ziehung wird die Koordinationsvariable dann für alle Elemente in der Population neu bestimmt. Es werden zwei Varianten für die Konstruktion einer Koordinationsvariable vorgestellt. Für die erste Variante wird die Zeit, die seit der zuletzt erfolgten Ziehung eines Elements vergangen ist, als Koordinationsvariable verwendet. Für die zweite Variante wird die erste mit der Erhebungslast der Elemente kombiniert. Beide Varianten entsprechen einer negativen Koordination aufeinanderfolgender Stichproben.

Bei der Annahme einer gleichen Belastung für jede Ziehung und einer unveränderlichen Population hat die zweite Variante, die Koordination über die Zeit außerhalb der Stichprobe und Erhebungslast, sehr ähnliche Eigenschaften wie eine negative Koordination von einfachen Zufallsstichproben mit PRN.

Für diese Arbeit wird die Notation von Nedyalkova et al. (2009) übernommen. Nedyalkova et al. (2009) richten einen besonderen Fokus auf die Beschreibung des sog. Längsschnittdesigns, d.h. das Stichprobendesign, das die Stichproben eines einzelnen Elements über den Beobachtungszeitraum beschreibt. So werden auch hier die Längsschnittdesigns für die beiden vorgestellten Varianten der Koordination hergeleitet, die sich aufgrund der negativen Koordination zeitlich naher Stichproben als systematische Stichprobedesigns darstellen. Die Verwendung von systematischen Längsschnittdesigns bringt einige Vorteile, wie auch Nedyalkova et al. (2009) aufzeigen. So haben Stichproben, bei entsprechend geringen Auswahlsätzen, in einer gewissen Nachbarschaft keine Überlappung, d.h. keine gemeinsamen Elemente. Es wird vorgeschlagen, diesen Raum für die Implementation von Rotationspanels (Steel & McLaren, 2009) zu nutzen. Rotationspanels finden auch häufig Anwendung bei Haushalts- und Personenbefragungen. Beispielweise verwendete die EU-SILC Erhebung ein sog. 4-in Rotationspanel, bei welchem ein Element nach seiner Ziehung für jeweils vier aufeinanderfolgende Zeitpunkt in der Stichprobe bleibt (Verma, Betti & Ghellini, 2007). Es wird vorgeschlagen, durch die Kombination von mehreren aufeinanderfolgenden, nicht überlappenden Stichproben diese Art von Rotationsschemata zu erzeugen. Eine derartige Einbettung dieser *one-level* Rotation in die Konzeption einer allgemeinen Stichprobenkoordination formalisiert und vervollständigt die Theorie designbasierter Inferenz für Rotationspanels. Diese stellt somit eine Alternative zu modellbasierten Ansätzen unter Verwendung von Rotationsgruppen dar (Park, Kim & Choi, 2001).

1.3.2 Kapitel 3

Kapitel 3 beschäftigt sich mit der Problematik der Varianzschätzung im Querschnitt für nicht lineare Schätzer bei komplexen Stichprobendesigns. Es wird auf verschiedene Techniken zur Vereinfachung der Varianzschätzung beim komplexen Design eingegangen. Eine Übersicht hierzu findet sich auch in Matei & Tillé (2005a).

Zur Schätzung einer nicht linearen Statistik wird für glatte Schätzer eine Linearisierung mittels der Taylorreihe vorgestellt (Wolter, 1985, Kapitel 6). Für nicht glatte Statistiken wie Armuts- und Disparitätsmaße werden die Konzepte der Einflussfunktion (Deville, 1999) und Schätzgleichungen (Kovacevic & Binder, 1997) als Methoden der Linearisierung dargestellt. Diese Verfahren werden an einem Armutsmaß und zwei Disparitätsmaßen demonstriert. Weitere Übersichten zur Linearisierung von Armuts- und Disparitätsmaßen finden sich in Osier (2009), Münnich & Zins (2011) und Verma, Betti & Ghellini (2011), sowie in Hulliger, Alfons, Filzmoser et al. (2011) für robuste Schätzer von Disparitätsmaßen. Die in Kapitel 3 erarbeiteten Methoden dienen als Grundlage für die in Kapitel 4 dargestellte Schätzung von Veränderungen nicht linearer Statistiken.

1.3.3 Kapitel 4

In Kapitel 4 wird zuerst ein allgemeiner Ansatz zur Schätzung von Veränderungen aufgezeigt. Danach werden Varianzschätzer für die Veränderung von Querschnittschätzern basierend auf korrelierten Querschnittstichproben untersucht. Hier wird unter anderem auf die von Wood (2008) aufgestellten Bedingungen eingegangen, unter denen die Varianzschätzung von Veränderungen vereinfacht werden kann. Für den Fall veränderlicher Populationen sind diese Bedingungen jedoch für den in Kapitel 2 vorgestellten Algorithmen der Koordination mittels Koordinationsvariable nicht erfüllt.

Es wird auch die Methode von Berger (2004b), Berger & Priam (2015) zur Varianzschätzung von Veränderungen diskutiert, die den in Kapitel 3 vorgestellten Ansatz zur Vereinfachung der Varianzschätzung bei Querschnittschätzern für die Varianzschätzung bei Veränderung verallgemeinert.

Schließlich werden Ergebnisse aus Kapitel 3 aufgegriffen, um Varianzschätzer für die Veränderung von Armuts- und Disparitätsmaßen zu begründen. Als Beispiel wird die Varianz einer Differenz des AROPE Indikators dargestellt.

Kapitel 2

Stichprobenziehungen im Zeitverlauf

2.1 Notation und Konzeption von wiederholten Stichprobenziehungen

Ziel dieses Abschnitts ist die Etablierung einer geeigneten Notation für das wiederholte Ziehen von Zufallsstichproben über die Zeit, wobei maßgeblich das Modell ohne Zurücklegen betrachtet wird. Hierzu wird auf die Arbeit von Nedyalkova et al. (2009) zurückgegriffen, welche die gebräuchliche Notation für das einmalige Ziehen einer Stichprobe (vgl. etwa Tillé, 2006, Kapitel 2) auf ein wiederholtes Ziehen in einer zeitlichen Abfolge erweitern.

2.1.1 Einmaliges Ziehen einer Stichprobe

Eine endliche Population ist eine Menge von N Elementen $\{u_1, \ldots, u_k, \ldots, u_N\}$. Jedes Element der Population kann eindeutig bestimmt werden durch seinen Index aus der Indexmenge

$$\mathscr{U} = \{1, \ldots, k, \ldots, N\} ,$$

die im Folgenden auch als Ziehungsrahmen bezeichnet wird. Eine Stichprobe lässt sich als Indexmenge $\mathscr{s}$ bezeichnen mit

$$\mathscr{s} \subset \mathscr{U} .$$

Ein Stichprobendesign $p(\mathscr{s})$ beschreibt die Wahrscheinlichkeit, mit der eine bestimmte Stichprobe $\mathscr{s}$ gezogen wird. Es wird folgende Definition verwendet:

Definition 2.1. Die diskrete Wahrscheinlichkeitsverteilung $p(.)$ über $\mathscr{P}(\mathscr{U})$ heißt *Stichprobendesign* und $\mathscr{G} = \{\mathscr{s} | \mathscr{s} \in \mathscr{P}(\mathscr{U}), \, p(\mathscr{s}) > 0\}$ heißt Träger von $p(.)$ mit

$$\sum_{\mathscr{s} \in \mathscr{G}} p(\mathscr{s}) = 1$$

Somit ist $p : \mathscr{G} \mapsto (0, 1]$.

Ziehen ohne Zurücklegen

Der Träger eines Designs mit Zurücklegen ist $\mathscr{S} = \{0,1\}^N$, mit $card(\mathscr{S}) = 2^N$. Alternativ kann eine Stichprobe auch als Zufallsvektor definiert werden. Für das Ziehen mit Zurücklegen ist dieser Zufallsvektor

$$\vec{\mathfrak{s}} \in \{0,1\}^N \ ,$$

mit

$$\vec{\mathfrak{s}} = (\mathfrak{I}_1, \ldots, \mathfrak{I}_k, \ldots, \mathfrak{I}_N)^\top \ ,$$

wobei

$$\mathfrak{I}_k = \begin{cases} 1 & \text{wenn } k \in \mathscr{s} \\ 0 & \text{wenn } k \notin \mathscr{s} \end{cases} \ , \qquad \forall k \in \mathscr{U} \ ,$$

(siehe z.B. Tillé, 2006, S. 8f). Der Stichprobenumfang einer Stichprobe $\vec{\mathfrak{s}}$ ohne Zurücklegen ist gegeben durch $\sum_{k \in \mathscr{U}} \mathfrak{I}_k = \mathfrak{n}$, mit $\mathfrak{n} < N$. Eine konkrete Ausprägung von $\vec{\mathfrak{s}}$ wird mit $\vec{s}$ bezeichnet. Dabei ist

$$\vec{s} = (I_1, \ldots, I_k, \ldots, I_N)^\top \ ,$$

wobei I_k die Ausprägung von $\mathfrak{I}_k$, ist wenn $\vec{\mathfrak{s}} = \vec{s}$. Für Designs mit $\mathfrak{n} = n \ \forall \ \vec{\mathfrak{s}} \in \mathscr{S}$, d.h. mit festem Stichprobenumfang, ist der Träger gegeben mit

$$\mathscr{S}_n = \{\vec{s} \in \mathscr{S} | \sum_{k \in \mathscr{U}} I_k = n\} \ ,$$

mit $card(\mathscr{S}_n) = \binom{N}{n}$.

Ziehen mit Zurücklegen

Ziehen mit Zurücklegen besteht aus der wiederholten Entnahme jeweils eines Elements aus $\mathscr{U}$, wobei nach jedem Zug das gezogene Element wieder zurückgelegt wird. Die Beschränkung auf eine fixe Anzahl von n_{wr} Zügen führt zu dem Träger $\mathscr{R}_{n_{wr}} = \{\vec{s} \in \mathscr{R} | \sum_{k \in \mathscr{U}} I_k = n_{wr}\}$, mit $card(\mathscr{R}_{n_{wr}}) = \binom{N+n_{wr}-1}{n_{wr}}$. Eine ebenfalls übliche Mengenschreibweise für Stichproben mit Zurücklegen ist die geordnete Menge $\mathscr{s}_{wr} = \{k_1, k_2, \ldots, k_i, \ldots, k_{n_{wr}}\}$, dabei ist $k_i = l$, wenn im i-ten Zug das l-te Element gezogen wurde.

Ziehungsalgorithmus

Von zentraler Bedeutung bei der Ziehung einer Stichprobe ist Art und Weise wie ein Stichprobendesign implementiert werden kann. Die Regel zur Ziehung einer Stichprobe, im folgenden Ziehungsalgorithmus genannt, legt dabei eindeutig ein Stichprobendesign und dessen Träger fest. Eine Eineindeutigkeit liegt nicht vor, da das gleiche Stichprobendesign durch mehr als nur einen Ziehungsalgorithmus implementiert werden kann.

Inklusionswahrscheinlichkeiten

Eine wichtige Eigenschaft eines Stichprobendesigns sind die daraus resultierenden Wahrscheinlichkeiten der einzelnen Elemente, in einer Stichprobe enthalten zu sein. Die Wahrscheinlichkeit

$$\pi_k = Pr(\mathfrak{I}_k > 0)$$

ist die Inklusionswahrscheinlichkeit des k-ten Elements. Die Wahrscheinlichkeit, dass zwei Elemente k und l gemeinsam in einer Stichprobe enthalten sind, ist durch

$$\pi_{k,l} = Pr(\mathfrak{I}_k \mathfrak{I}_l > 0)$$

gegeben. $\pi_{k,l}$ wird auch Inklusionswahrscheinlichkeit zweiter Ordnung genannt.

Für das Ziehen ohne Zurücklegen ist somit $\pi_k = E(\mathfrak{I}_k) = \sum_{\vec{s}\in\mathscr{S}} p(\vec{s}) I_k$ und $\pi_{k,l} = E(\mathfrak{I}_k \mathfrak{I}_l) = \sum_{\vec{s}\in\mathscr{S}} p(\vec{s}) I_k I_l$. Der Erwartungswert des Stichprobenumfangs $\mathfrak{n}$ ist gegeben durch

$$E(\mathfrak{n}) = \mathrm{E}\left(\sum_{k\in\mathscr{U}} \mathfrak{I}_k\right) = \sum_{k\in\mathscr{U}} \mathrm{E}\left(\mathfrak{I}_k\right) ,$$

mit $E(\mathfrak{n}) = n$ für Designs mit festem Stichprobenumfängen. Speziell bei Ziehen ohne Zurücklegen mit festem Stichprobenumfang ist $\sum_{k\in\mathscr{U}} \pi_k = n$.

2.1.2 Wiederholtes Ziehen von Stichproben in zeitlicher Abfolge

Soll eine Population über einen Zeitraum $\mathscr{T} = \{1,\ldots,T\}$, also im Längsschnitt, beobachtet werden, bedarf es Stichprobenziehungen zu den Zeitpunkten $t = 1,2,\ldots,T$. Dabei wird eine Stichprobe zum Zeitpunkt t aus der Indexmenge $\mathscr{U}^t$ entnommen, wobei $\mathscr{U}^t$ den Ziehungsrahmen zum Zeitpunkt t darstellt mit $card(\mathscr{U}^t) = N^t$ als Anzahl der vorhandenen Elemente im Zeitpunkt t. Ohne Einschränkung der Aussagen in Abschnitt (2.1.1) sei, $\mathscr{U} = \cup_{t=1}^{T} \mathscr{U}^t$ und $N = card(\mathscr{U})$, mit $\mathscr{U} = \{1,\ldots,k,\ldots,N\}$. Ist die Population gleichbleibend über den Betrachtungszeitraum, ist $\mathscr{U} = \mathscr{U}^t \;\forall\, t \in \mathscr{T}$. Eine Stichprobe zum Zeitpunkt t kann dargestellt werden als $\mathscr{s}^t \subset \mathscr{U}^t$ und eine gemeinsame Stichprobe $\mathscr{s} = \cup_{t=1}^{T} \mathscr{s}^t$ über den gesamten Beobachtungszeitraum als $\mathscr{s} \subset \mathscr{U}$. Ziehen ohne Zurücklegen soll bei der Betrachtung wiederholter Stichprobenziehungen über die Zeit im Vordergrund stehen. Aus diesem Grund wird im Folgenden davon ausgegangen, dass alle Querschnittdesigns ohne Zurücklegen sind. Es wird speziell darauf hingewiesen, falls dies nicht gilt.

Ausgehend von der Notation bei einmaliger Ziehung (ohne Zurücklegen) werden nun die folgenden drei Arten von Stichproben definiert (siehe Nedyalkova et al., 2009, S. 272). Der Zufallsvektor

$$\vec{\mathfrak{s}}^t = (\mathfrak{I}_1^t,\ldots,\mathfrak{I}_k^t,\ldots,\mathfrak{I}_N^t)^\top \in \{0,1\}^N , \qquad \forall t \in \mathscr{T}$$

wird als Querschnittstichprobe bezeichnet zum Zeitpunkt t und der Zufallsvektor

$$\vec{\mathfrak{s}}_k = (\mathfrak{I}_k^1,\ldots,\mathfrak{I}_k^t,\ldots,\mathfrak{I}_k^T) \in \{0,1\}^T, \qquad \forall k \in \mathscr{U}$$

als Längsschnittstichprobe des k-ten Elements. Die Matrix

$$\mathfrak{S} = \left(\mathfrak{I}_k^t\right)_{\substack{k=1,\dots,N\\ t=1,\dots,T}} \in \{0,1\}^{N\times T}$$

ist die gemeinsame Stichprobe. Dabei ist

$$\mathfrak{I}_k^t = \begin{cases} 1 & \text{wenn } k \in \mathscr{s}^t \\ 0 & \text{wenn } k \notin \mathscr{s}^t \end{cases} , \qquad \forall\, k \in \mathscr{U} .$$

Eine konkrete Ausprägung von $\vec{\mathfrak{s}}^t$, $\vec{\mathfrak{s}}_k$ und $\mathfrak{S}^t$, wird mit $\vec{s}^t$, $\vec{s}_k$ bzw. $\mathbf{S}$ bezeichnet und es gilt

$$\begin{aligned} \vec{s}^t &= (I_1^t, \dots, I_k^t, \dots, I_N^t)^\top \in \{0,1\}^N \\ \vec{s}_k &= (I_k^1, \dots, I_k^t, \dots, I_k^t) \in \{0,1\}^T \\ \mathbf{S} &= (I_k^t)_{\substack{k=1,\dots,N\\ t=1,\dots,T}} \in \{0,1\}^{N\times T} , \end{aligned}$$

mit I_k^t als der Ausprägung von $\mathfrak{I}_k^t$ für $\mathfrak{S} = \mathbf{S}$.

Es werden folgende Stichprobendesigns definiert:

Definition 2.2. Das Stichprobendesign $Pr(\vec{\mathfrak{s}}^t = \vec{s}^t) = p^t(\vec{s}^t)$, $\forall\, t \in \mathscr{T}$ mit dem Träger $\mathscr{S}^t \subseteq \{0,1\}^N$ und mit $card(\mathscr{S}^t) = m^t$, heißt *Querschnittdesign* zum Zeitpunkt t.

Definition 2.3. $Pr(\vec{\mathfrak{s}}_k = \vec{s}_k) = p_k(\vec{s}_k)$, $\forall\, k \in \mathscr{U}$ mit dem Träger $\mathscr{S}_k \subseteq \{0,1\}^T$ und $m_k = card(\mathscr{S}_k)$, heißt *Längsschnittdesign* des k-ten Elements.

Definition 2.4. Das Stichprobendesign $Pr(\mathfrak{S} = \mathbf{S}) = p(\mathbf{S}) = p(\vec{s}^1, \dots, \vec{s}^t, \dots, \vec{s}^T)$ bzw. $Pr(\mathfrak{S} = \mathbf{S}) = p(\mathbf{S}) = p(\vec{s}_1, \dots, \vec{s}_k, \dots, \vec{s}_N)$ mit dem Träger $\mathscr{S} \subseteq \{0,1\}^{N\times T}$ heißt *gemeinsames Stichprobendesign*.

Es wird zudem folgende Definition für $\mathfrak{I}_k^t$ getroffen, falls Element k zum Zeitpunkt t nicht existiert.

Definition 2.5. $\mathfrak{I}_k^t = 0$ für alle $k \notin \mathscr{U}^t$.

Definition 2.5 wird aus praktischen Überlegungen heraus getroffen und dient dazu, die gemeinsame Stichprobe $\mathfrak{S}$ immer als Matrix darstellen zu können. Für die vorgestellten Ziehungsalgorithmen in den nachfolgenden Abschnitten ist es auch unerheblich, welche Werte $\mathfrak{I}_k^t$ für $k \notin \mathscr{U}^t$ annimmt, denn $\mathfrak{I}_k^t$ ist in diesen Fällen keine Zufallsvariable.

Der Stichprobenumfang des Querschnittdesigns $p^t(.)$ wird durch

$$\mathfrak{n}^t = \sum_{k\in\mathscr{U}} \mathfrak{I}_k^t$$

beschrieben und der des Längsschnittdesigns $p_k(.)$ mit

$$\mathfrak{n}_k = \sum_{t\in\mathscr{T}} \mathfrak{I}_k^t .$$

Der gemeinsame Stichprobenumfang der beiden Designs $p^t(.)$ und $p^u(.)$, auch Überlappung genannt, wird mit

$$\mathfrak{n}^{t,k} = \sum_{k \in \mathscr{U}} \Im_k^t \Im_k^u$$

angegeben.

Für das Querschnitt- bzw. das Längsschnittdesign ergeben sich die folgenden Inklusionswahrscheinlichkeiten erster Ordnung:

$$\pi_k^t = E(\Im_k^t) \qquad \forall \quad k \in \mathscr{U} \text{ und } t \in \mathscr{T} \ . \tag{2.1}$$

Die Inklusionswahrscheinlichkeiten zweiter Ordnung des Querschnitt- bzw. des Längsschnittdesign sind gegeben durch:

$$\pi_{k,l}^t = E(\Im_k^t \Im_l^t) \qquad \forall\, k,l \in \mathscr{U} \ k \neq l \text{ und } t \in \mathscr{T} \tag{2.2}$$

$$\pi_k^{t,u} = E(\Im_k^t \Im_k^u) \qquad \forall\, k \in \mathscr{U} \text{ und } t,u \in \mathscr{T} \ t \neq u \tag{2.3}$$

In Analogie zur einmaligen Ziehung entspricht $\pi_{k,l}^t$ der Wahrscheinlichkeit, dass die Elemente k und l gemeinsam zum gleichen Zeitpunkt t gezogen werden. Hingegen entspricht $\pi_k^{t,u}$ der Wahrscheinlichkeit, mit welcher das k-te Element sowohl zum Zeitpunkt t als auch zum Zeitpunkt u gezogen wird.

Schließlich kann für das gemeinsame Stichprobendesign die Inklusionswahrscheinlichkeit $\pi_{k,l}^{t,u}$ bestimmt werden mit

$$\pi_{k,l}^{t,u} = E(\Im_k^t \Im_l^u) \qquad \forall \quad k,l \in \mathscr{U} \text{ und } t,u \in \mathscr{T} \ , \quad \text{mit } k \neq l \text{ und } t \neq u \ . \tag{2.4}$$

$\pi_{k,l}^{t,u}$ ist die Wahrscheinlichkeit, dass das k-te Element zum Zeitpunkt t und das l-te Element zum Zeitpunkt u gezogen wird. Hierbei ist zu beachten, dass zwar $\pi_{k,l}^{t,u} = \pi_{l,k}^{u,t}$, aber im allgemeinen nicht $\pi_{k,l}^{t,u} = \pi_{l,k}^{t,u}$ gelten muss. Beispielsweise ist für den Fall einer sich veränderlichen Population, mit $k \in \mathscr{U}^t$, $l \in \mathscr{U}^t$ und $l \notin \mathscr{U}^t$ $Pr(\Im_l^t = 0) = 1$ und somit $\pi_{l,k}^{t,u} = 0$.

Des Weiteren soll $\vec{\pi}^t = (\pi_k^t)_{k=1,\dots,N}$ der Vektor der Inklusionswahrscheinlichkeiten erster Ordnung des Querschnittdesigns $p^t(.)$ sein und $\mathbf{\Pi}_{kl}^t$ die $N \times N$ Matrix seiner Inklusionswahrscheinlichkeit zweiter Ordnung, mit

$$\mathbf{\Pi}_{k,l}^t = \left[\mathrm{E}\left(\Im_k^t \Im_l^t\right)\right]_{\substack{k=1,\dots,N \\ l=1,\dots,N}} \ .$$

Analog soll $\vec{\pi}_k = (\pi_k^t)_{t=1,\dots,T}$ der Vektor der Inklusionswahrscheinlichkeiten erster Ordnung des Längsschnittdesigns $p_k(.)$ sein und $\mathbf{\Pi}_k^{tu}$ die $T \times T$ Matrix seiner Inklusionswahrscheinlichkeit zweiter Ordnung, mit

$$\mathbf{\Pi}_k^{t,u} = \left[\mathrm{E}\left(\Im_k^t \Im_k^u\right)\right]_{\substack{t=1,\dots,T \\ u=1,\dots,T}} \ .$$

Schließlich soll $\vec{\pi} = (\pi_k^t)^\top_{(t,k)=(1,1),(1,2),\dots,(1,N),(2,1),\dots,(T,N)}$ der Vektor der Inklusionswahrscheinlichkeiten erster Ordnung des gemeinsamen Designs $p(.)$ sein und $\mathbf{\Pi}$ die $TN \times TN$ Matrix seiner Inklusionswahrscheinlichkeiten zweiter Ordnung, mit

$$\mathbf{\Pi} = \left[\mathrm{E}\left(\Im_k^t \Im_l^u\right)\right]_{\substack{(t,k)=(1,1),(1,2),\dots,(1,N),(2,1),\dots,(T,N) \\ (u,l)=(1,1),(1,2),\dots,(1,N),(2,1),\dots,(T,N)}} \ .$$

Eine wichtige Eigenschaft des gemeinsamen Stichprobendesigns ist die Möglichkeit, die Stichprobe $\vec{\mathfrak{s}}^t$ zu ziehen, unabhängig davon, wie viele Stichproben $\vec{\mathfrak{s}}^{t+1}, \dots, \vec{\mathfrak{s}}^T$ noch gezogen werden. Um dies sicherzustellen wird ein sequenzieller Ziehungsalgorithmus für das Längsschnittdesign benötigt, (Nedyalkova et al., 2009, S. 274). Ein sequenzieller Ziehungsalgorithmus wird auf eine geordnete Liste von Elementen, oder wie hier von Ziehungszeitpunkten, angewendet. Zu jedem Ziehungszeitpunkt hängt die Häufigkeit, mit der das Element in die Stichprobe aufgenommen wird, von der bedingten Wahrscheinlichkeitsverteilung $Pr(\mathfrak{I}_k^t = I_k^t | \vec{\mathfrak{s}}^{t-1} = \vec{s}^{t-1}, \dots, \vec{\mathfrak{s}}^1 = \vec{s}^1)$ ab. Es lassen sich die folgenden zwei Definitionen für zeitlich sequentielle Algorithmen angeben (siehe, Nedyalkova et al., 2009 nach Tillé, 2006, S. 33f)

Definition 2.6. Der Ziehungsalgorithmus eines Elementes k ist schwach sequenziell, wenn zum Zeitpunkt $t = 1, \dots, T$ des Algorithmus die Entscheidung, wie häufig das k-te Element in die Stichprobe $\vec{\mathfrak{s}}^t$ aufgenommen wird, eindeutig feststeht.

Definition 2.7. Der Ziehungsalgorithmus eines Elements k ist strikt sequenziell, wenn er schwach sequenziell ist und darüberhinaus die Entscheidung, wie häufig das k-te Element in die Stichprobe $\vec{\mathfrak{s}}^t$ aufgenommen wird, nicht von den Inklusionswahrscheinlichkeiten des k-ten Elementes zu den Zeitpunkten $t = t + 1, \dots, T$ abhängt.

Die Verwendung eines strikt sequenziellen Algorithmus zur Umsetzung des Längsschnittdesign stellt sich in der Praxis, aufgrund einer veränderlichen Population, als notwendig dar. Ein anderer Grund können Querschnittdesigns mit Inklusionswahrscheinlichkeiten proportional zu einer Hilfsvariablen sein, die für zukünftige Ziehungszeitpunkte noch nicht beobachtbar ist, (Nedyalkova et al., 2009, S. 274).

2.2 Koordination von Zufallsstichproben

Die Anforderungen, welche an Querschnitt- und Längsschnittdesigns gestellt werden, können sehr unterschiedlich sein. Ein Querschnittdesign sollte eine unverzerrte und (nach Möglichkeit) effiziente Schätzung der interessierenden Parameter zu einem bestimmten Zeitpunkt erlauben. Ein wünschenswerte Eigenschaft von $p^t(.)$ ist zudem, dass $\pi_{k,l}^t > 0 \; \forall \, k, l \in \mathscr{U}^t$, da dies die Verwendung von wohlbekannten Varianzschätzern ermöglicht (siehe hiezu Abschnitt 3.1). Die Aufgabe eines Längsschnittdesign besteht hingegen in der Koordination mehrere Querschnittstichproben, welche in einer zeitlichen Abfolge voneinander gezogen werden. Als Folge dessen kann $\pi_k^{t,u} \geq 0 \quad \forall \quad t, u \in \mathscr{T}$ sein, wenn beispielsweise ausgeschlossen werden soll, dass ein Element sowohl im Zeitpunkt t als auch zum Zeitpunkt u in der Stichprobe enthalten ist. Für das Längsschnittdesign stellt ein variabler Stichprobenumfang oft eine Notwendigkeit dar, ist dieser doch oftmals von der Länge des betrachteten Zeitraums abhängig.

Das gemeinsame Stichprobendesign $p(\mathbf{S})$ gilt als koordiniert wenn,

$$p(\mathbf{S}) \neq p^1(\vec{s}^1) p^2(\vec{s}^2) \dots p^T(\vec{s}^T) \, .$$

Ein mögliches Maß für die Koordination stellt dabei der Umfang der Schnittmengen, d.h. die Überlappungen zwischen den Querschnittstichproben, dar. Für die Stichproben $\vec{\mathfrak{s}}^t$ und $\vec{\mathfrak{s}}^u$ ist der Erwartungswert der Überlappung $\mathfrak{n}^{t,u} = \sum_{k \in \mathscr{U}} \mathfrak{I}_k^t \mathfrak{I}_k^u$ gegeben durch

$$E(\mathfrak{n}^{t,u}) = \sum_{k \in \mathscr{U}} \pi_k^{t,u} = \sum_{s^t \in \mathscr{S}^t} \sum_{s^u \in \mathscr{S}^u} \mathfrak{n}^{t,u} p(\vec{s}^t, \vec{s}^u) \, .$$

Somit liegt $E(\mathfrak{n}^{t,u})$ in folgenden Grenzen

$$\sum_{k\in\mathcal{U}} \max(\pi_k^t + \pi_k^u - 1, 0) \leq \mathrm{E}\left(\mathfrak{n}^{t,u}\right) \leq \sum_{k\in\mathcal{U}} \min(\pi_k^t, \pi_k^u) \; .$$

Diese ergeben sich daraus, dass die Wahrscheinlichkeiten $\pi_k^{t,u}$ der logischen Verknüpfung der Ereignisse $k \in \mathscr{s}^t$ und $k \in \mathscr{s}^u$ in den Grenzen

$$\max(\pi_k^t + \pi_k^u - 1, 0) \leq \pi_k^{t,u} \leq \min(\pi_k^t, \pi_k^u) \; ,$$

liegen.[1] Positive und negative Koordination der Ziehung eines Elements k an zwei Zeitpunkten t und u werden wie folgt definiert.

Definition 2.8. Ist $\pi_k^{t,u} > \pi_k^t \pi_k^u$, so gilt die Ziehung des k-ten Elements an den Zeitpunkten t und u als positiv koordiniert.

Definition 2.9. Ist $\pi_k^{t,u} < \pi_k^t \pi_k^u$, so gilt die Ziehung des k-ten Elements an den Zeitpunkten t und u als negativ koordiniert.

Eine unkoordinierte Ziehung wird definiert durch;

Definition 2.10. Ist $\pi_k^{t,u} = \pi_k^t \pi_k^u$, so gilt die Ziehung des k-ten Elements an den Zeitpunkten t und u als unkoordiniert.

Zur Ausgestaltung des gemeinsamen Designs ist es hilfreich die folgenden Eigenschaften zu betrachten:

i) Varianz von $\mathfrak{n}^t \quad 1 \leq t \leq T$

ii) Varianz von $\mathfrak{n}^{t,u} \quad 1 \leq t < u \leq T$

iii) Verteilung der Erhebungslast

iv) Verweildauer innerhalb der Stichprobe

v) Verweildauer außerhalb der Stichprobe

Bezüglich der Eigenschaften i) und ii) ist anzumerken, dass hier zwischen Designs mit festem Stichprobenumfängen für das Querschnittdesign, bzw. festen Überlappungen, unterschieden werden kann. Für das Querschnittdesign ist es oft wünschenswert, einen festen Stichprobenumfang zu verwenden. Zum einen ist dies auf planungstechnische Gründe einer Stichprobenerhebung zurückzuführen, und zum anderen erlaubt es die Verwendung von standardisierten Verfahren bei Querschnittanalysen.[2] Die Überlappung kann als Zielgröße bei der Koordination verwendet werden, indem beispielsweise

---

[1] Diese Eigenschaft ist auch als Fréchet Ungleichung bekannt. Für einen Beweis siehe Nedyalkova (2009, S. 54)

[2] Zur Planung des Stichprobenumfangs einer Erhebung werden beispielsweise oft maximale Fehlertoleranzen bezüglich eines wichtigen Schätzers angegeben, die mit einer gegebenen Wahrscheinlichkeit nicht überschritten werden sollen, oder es wird mit minimalen effektiven Stichprobengrößen geplant. Aus Kostenerwägungen und Gründen der Planungssicherheit werden hier Designs mit fixen Stichproben bevorzugt (Valliant, Dever & Kreuter, 2013, Kapitel 1-4).

versucht wird, diese zu maximieren oder zu minimieren. Jedoch stellt sich bei derartigen Lösungen des Koordinationsproblems die Überlappung auch unmittelbar aufeinander folgender Querschnittstichproben im Allgemeinen als Zufallsvariable dar (siehe z.B., Qualité, 2009, Kap. 6). Eigenschaft iii) ist von Interesse, wenn beispielsweise eine gleiche Belastung der einzelnen Untersuchungseinheiten über den Untersuchungszeitraum hinweg erwünscht ist. Dabei ist die kumulierte Erhebungslast des k-ten Elements nach t Erhebungsperioden durch

$$\mathfrak{b}_k^t = \sum_{i=1}^{t} g_k^i \mathfrak{I}_k^i \tag{2.5}$$

gegeben, wobei g_k^t den Belastungswert des k-ten Elements bezeichnet, wenn dieses an der Erhebung zum Zeitpunkt t teilnimmt. Der Wert von g_k^t ist für alle $k \in \mathscr{U}$ und $t \in \mathscr{T}$ ein Parameter der Population und wird üblicherweise als bekannt angenommen (siehe hierzu auch Nedyalkova, 2009, Kapitel 4). Dabei kann durch ein individuelles Festlegen von g_k^t einer ungleichen Belastung, die den einzelnen Elementen aufgebürdet wird, Rechnung getragen werden. Handelt es sich beispielsweise um eine Unternehmenserhebung, entstehen bei größeren Firmen möglicherweise geringere Grenzkosten als bei kleineren, da hier bei der Zusammenstellung der Daten oft bestehende Synergien genutzt werden können. Auch kann die Befragung in einer Periode umfangreicher sein als in einer anderen.

Eigenschaft iv) lässt sich mit Hilfe der Zufallsvariable $\psi_k^t \in \{1, \ldots, T-t\}$ für $1 \leq t < T$ wie folgt beschreiben:

$$Pr(\psi_k^t = a) = \begin{cases} Pr(\mathfrak{I}_k^{t+1} = 0 | \mathfrak{I}_k^t = 1) & \text{für } a = 1 \\ Pr(\mathfrak{I}_k^{t+a} = 0, \mathfrak{I}_k^{t+a-1} = \ldots = \mathfrak{I}_k^{t+1} = 1 | \mathfrak{I}_k^t = 1) & \text{für } 1 < a < T-t \\ 1- & \\ \sum_{i=1}^{T-t-1} Pr(\mathfrak{I}_k^{t+i+1} = 0, \mathfrak{I}_k^{t+i} = \ldots = \mathfrak{I}_k^{t+1} = 1 & \\ \qquad\qquad |\mathfrak{I}_k^t = 1) & \text{für } a = T-t \end{cases}$$

ψ_k^t ist somit die Zeit, die Element k nach seiner Ziehung zum Zeitpunkt t in der Stichprobe ohne Unterbrechung verbringt, sprich seine Verweildauer in der Stichprobe.

Eigenschaft v) bestimmt die Verteilung der Zeit bis zur nächsten Ziehung eines Elements. Die Zeit, die von Zeitpunkt t bis zur nächsten Ziehung eines Elements k verstreicht, lässt sich durch die Zufallsvariable ξ_k^t beschreiben mit

$$\xi_k^t = \begin{cases} \min(T, \min(\{a | a \in \mathbb{N}, \mathfrak{I}_k^a = 1\})) & \text{für} \quad t = 0 \\ \min(T-t, \min(\{a | a \in \mathbb{N}, \mathfrak{I}_k^{t+a} = 1\})) & \text{für} \quad 1 \leq t < T \\ 0 & \text{für} \quad t = T \end{cases} \tag{2.6}$$

Entsprechend ist ξ_k^0 die Zeit bis zur erstmaligen Ziehung von k und $\xi_k^T = 0$, da nach Periode T der Beobachtungszeitraum endet.

Die Zufallsvariable $\phi_k^t = \xi_k^t | (\mathfrak{I}_k^t = 1)$ gibt die Zeit zwischen Zeitpunkt t bis zur nächsten Ziehung eines Elementes k an, gegeben dass k zuletzt zum Zeitpunkt t ausgewählt wurde. Die Wahrscheinlichkeitsverteilung von ϕ_k^t, für $t < T-2 \; \forall \; k \in \mathscr{U}$ ist gegeben durch (siehe Nedyalkova et al., 2009, S. 275):

$$Pr(\phi_k^t = a) = \begin{cases} Pr(\Im_k^{t+1} = 1 | \Im_k^t = 1) & \text{für } a = 1 \\ Pr(\Im_k^{t+a} = 1, \sum_{j=1}^{a-1} \Im_k^{t+j} = 0 | \Im_k^t = 1) & \text{für } 1 < a < T - t \\ 1 - Pr(\Im_k^{t+1} = 1 | \Im_k^t = 1) & \\ \quad + \sum_{i=2}^{T-t-1} Pr(\Im_k^{t+i} = 1, \sum_{j=1}^{i-1} \Im_k^{t+j} = 0 | \Im_k^t = 1) & \text{für } a = T - t \\ 0 & \text{sonst}. \end{cases} \tag{2.7}$$

Ist $t = T - 2$ lässt sich die Wahrscheinlichkeitsverteilung von $\phi_k^t \; \forall \, k \in \mathscr{U}$ schreiben als

$$Pr(\phi_k^t = a) = \begin{cases} Pr(\Im_k^{t+1} = 1 | \Im_k^t = 1) & \text{für } a = 1 \\ 1 - Pr(\Im_k^{t+1} = 1 | \Im_k^t = 1) & \text{für } a = T - t \\ 0 & \text{sonst}. \end{cases} \tag{2.8}$$

Schließlich gilt für $t = T - 1$, $\phi_k^t = 1 \; \forall \, k \in \mathscr{U}$ und ϕ_k^T wird definiert mit $\phi_k^T = 0 \; \forall \, k \in \mathscr{U}$.

Für $T \to \infty$ ist $E(\phi_k^t)$ die zu erwartende Zeit, die bis zur nächsten Selektion des k-ten Elements verstreicht, wenn dieses gerade zum Zeitpunkt t ausgewählt wurde. $Pr(\phi_k^t = T - t)$ ist die Wahrscheinlichkeit dafür, dass das k-te Element nicht wieder im Beobachtungszeitraum gezogen wird oder genau zum Zeitpunkt $T - t$ wieder gezogen wird. Zudem ist $\phi_k^0 = \xi_k^0$, die Verteilung von ϕ_k^0 ist gegeben durch

$$Pr(\phi_k^0 = a) = \begin{cases} Pr(\Im_k^1 = 1) & \text{für } a = 1 \\ Pr(\Im_k^a = 1, \sum\limits_{i=1}^{a-1} \Im_k^i = 0) & \text{für } 1 < a < T \\ Pr(\sum\limits_{i=1}^{T} \Im_k^i = 0) & \text{für } a = T \\ 0 & \text{sonst}. \end{cases}$$

Die Verteilung von ψ_k^t gehört zu den zentralen Eigenschaften des Längsschnittdesigns. So lässt sich beispielsweise $\pi_k^{t,u}$ hieraus ableiten. Angenommen, Element k wird zu den Zeitpunkten t und u, mit $t < u$, ausgewählt und zwischen diesen weitere zweimal, jeweils zu den Zeitpunkten $t + a_1$ und $t + a_1 + a_2$. Die Wahrscheinlichkeit für dieses Ereignis lässt sich schreiben als

$$\pi_k^t Pr(\phi_k^t = a_1) Pr(\phi_k^{t+a_1} = a_2 | \phi_k^t = a_1) Pr(\phi_k^{t+a_1+a_2} = a_3 | \phi_k^{t+a_1} = a_2) \, , \tag{2.9}$$

dabei entspricht a_3 der Zeit von der Auswahl in $t + a_1 + a_2$ bis u, d.h. $t + a_1 + a_2 + a_3 = u$. So lässt sich $\pi_k^{t,u}$ bestimmen als die Summe der Wahrscheinlichkeiten aller Kombinationen von Ziehungen von k, welche gegeben dessen Ziehung in t dazu führen, auch in $u > t$ ausgewählt zu werden. Hierzu werden alle Kombinationen von natürlichen Zahlen, deren Summe gleich $u - t$ ist, dargestellt als Mengen von m-Tupeln $\vec{a} = (a_1, \ldots, a_m)$, mit $m = 1, \ldots, (u - t)$,

$$\{\mathscr{A}^i\}_{i=1,2,\ldots,(u-t)} \quad \text{mit} \quad \mathscr{A}^i = \left\{ \vec{a} | \vec{a} \in \mathbb{N}^i, \sum_{j=1}^{i} a_j = u - t \right\}.$$

Somit gilt für $t < u < T$ und $\pi_k^{t,u} > 0$,

$$\begin{aligned} \pi_k^{t,u} &= \pi_k^t Pr(\phi_k^t = u-t) \\ &\quad + \mathbb{1}\left((u-t) > 1\right) \sum_{i=2}^{u-t} \left(\sum_{\vec{a} \in \mathscr{A}^i} \pi_k^t Pr(\phi_k^t = a_1) \right. \\ &\qquad \left. \prod_{j=2}^{i} Pr(\phi_k^{t+\sum_{\nu=0}^{j-1} a_\nu} = a_j | \phi_k^{t+\sum_{\nu=0}^{j-2} a_\nu} = a_{j-1}) \right) , \end{aligned} \tag{2.10}$$

mit $a_0 = 0$. Das heißt, für alle Stichproben mit gemeinsamen Elementen lässt sich $\pi_k^{t,u}$ durch die Verteilungen der Zeit zwischen wiederholten Ziehungen von k beschreiben. Falls $u = T$ muss noch eine entsprechende Fallunterscheidung vorgenommen werden zwischen $\mathfrak{I}_k^T = 1$ und $\mathfrak{I}_k^T = 0$, für $\phi_k^{t+\sum_{\nu=1}^{m} a_\nu} = T - t (m = 1, \ldots, (u-t))$.

Die Darstellung in (2.10) mag nicht sehr praktikabel erscheinen, wenn viele der in Betracht gezogenen Ziehungskombination von k, auch auf Grund der Stichprobenkoordination, eine Wahrscheinlichkeit von Null aufweisen. Jedoch kann es gerade bei Längsschnittdesigns mit kleinen Trägern einfacher sein, eine Regel zur sequenziellen Bestimmung aller möglichen Ziehungskombination aufzustellen, jeweils in Abhängigkeit zur den vorangegangen Ziehungen, als einen analytischen Ausdruck für $\pi_k^{t,u}$ zu finden.

In Anlehnung an die Definition von ϕ_k^t kann auch die Zeit betrachtet die seit der letzten Ziehung von Element k vergangen ist, wenn es gerade zum Zeitpunkt t gezogen würde. Hierfür soll die Zufallsvariable χ_k^t die Zeit sein die seit der letzten Ziehung von k vergangen ist, gegeben $\mathfrak{I}_k^t = 1$. Die Wahrscheinlichkeitsverteilung von χ_k^t, für $t > 3$ $\forall\, k \in \mathscr{U}$ ist gegeben durch

$$Pr(\chi_k^t = a) = \begin{cases} Pr(\mathfrak{I}_k^{t-1} = 1 | \mathfrak{I}_k^t = 1) & \text{für } a = 1 \\ Pr(\mathfrak{I}_k^{t-a} = 1, \sum_{j=1}^{a-1} \mathfrak{I}_k^{t-j} = 0 | \mathfrak{I}_k^t = 1) & \text{für } 1 < a < t-1 \\ 1 - Pr(\mathfrak{I}_k^{t-1} = 1 | \mathfrak{I}_k^t = 1) & \\ \quad + \sum_{i=2}^{t-1} Pr(\mathfrak{I}_k^{t-i} = 1, \sum_{j=1}^{i-1} \mathfrak{I}_k^{t-j} = 0 | \mathfrak{I}_k^t = 1) & \text{für } a = t-1 \\ 0 & \text{sonst}. \end{cases} \tag{2.11}$$

Für $t = 3$ gilt

$$Pr(\chi_k^t = a) = \begin{cases} Pr(\mathfrak{I}_k^{t-1} = 1 | \mathfrak{I}_k^t = 1) & \text{für } a = 1 \\ 1 - Pr(\mathfrak{I}_k^{t-1} = 1 | \mathfrak{I}_k^t = 1) & \text{für } a = 2 \\ 0 & \text{sonst}. \end{cases} \tag{2.12}$$

Schließlich ist $\chi_k^2 = 1 \;\forall\, k \in \mathscr{U}$ und χ_k^1 wird definiert mit $\chi_k^1 = 0 \;\forall\, k \in \mathscr{U}$.

2.3 Verfahren zur Koordination von Stichproben

Die meisten Verfahren zur Koordination von Zufallsstichproben lassen sich durch eine der folgenden zwei Vorgehensweisen charakterisieren:

Typ A. Es werden die Querschnittdesigns $p^t(\cdot)$ für alle Perioden in $\mathscr{T}$ festgelegt. Danach wird eine Koordination gewählt, die zulässig ist, d.h. unabhängig vom Längsschnittdesign müssen die gewählten Querschnittdesigns (nach Möglichkeit) erhalten bleiben.

Typ B. Es werden die Längsschnittdesigns $p_k(\cdot)$ für alle Elemente in $\mathscr{U}$ festgelegt. Die direkte Gestaltung der Querschnittdesigns entfällt, denn diese ergibt sich aus der Umsetzung der einzelnen Längsschnittdesigns.

Bei Verfahren vom Typ A muss ein gemeinsames Design $p(\cdot)$ gefunden werden, das mit den Querschnittdesigns vereinbar ist, d.h.

$$p^t(\vec{s}^t) = \sum_{\vec{s}^1 \in \mathscr{S}^1} \cdots \sum_{\vec{s}^{t-1} \in \mathscr{S}^{t-1}} \sum_{\vec{s}^{t+1} \in \mathscr{S}^{t+1}} \cdots \sum_{\vec{s}^T \in \mathscr{S}^T} p(\mathbf{S}) \quad \forall t \in \mathscr{T}.$$

Zur Ziehung von $\vec{\mathfrak{s}}^1$ kann jedes beliebige Querschnittdesign $p^1(\cdot)$ umgesetzt werden. Für alle folgenden Stichproben $\vec{\mathfrak{s}}^t$ mit $t = 2, \ldots, T$ sind für einen strikt sequentiellen Ziehungsalgorithmus die bedingten Querschnittdesigns

$$p^t(\vec{s}^t | \vec{s}^{t-1}, \ldots, \vec{s}^1) = \frac{Pr(\vec{s}^1, \ldots, \vec{s}^t)}{Pr(\vec{s}^1, \ldots, \vec{s}^{t-1})},$$

zu bestimmen.

Ein Verfahren vom Typ A findet sich in Matei & Tillé (2005b) und Matei & Skinner (2009), die eine Methode zur Bestimmung eines gemeinsamen Designs unter Vorgabe zweier Querschnittdesigns $p^t(.)$ und $p^u(.)$ beschreiben. Die Koordination besteht hier in der Maximierung bzw. Minimierung des Erwartungswerts von $\mathfrak{n}^{t,u}$. Es wird mit Hilfe eines *Iterativ Proportional Fitting* (IPF) Algorithmus versucht, ein gemeinsames Design zu finden, welches eine negative oder eine positive Koordination einschließt. Den Ausgangspunkt bildet hierbei die Matrix $\mathbf{A} = (a_{ij})$ der Dimension $m^t \times m^u$, mit $a_{ij} = p^t(\vec{s}_i^t) p^u(\vec{s}_j^u)$, wobei $\vec{s}_i^t$ hier für die i-te Stichprobe aus dem Träger $\mathscr{S}^t$ steht und $\vec{s}_j^u$ entsprechend. Matei & Tillé (2005b) geben eine Regel an, nach welcher entschieden wird, welche gemeinsame Stichproben $\mathbf{S} = (\vec{s}^t, \vec{s}^u)$ zugelassen sind, wenn der Erwartungswert von $\mathfrak{n}^{t,u}$ maximiert bzw. minimiert werden soll. Für alle Kombinationen $\vec{s}_i^t$ und $\vec{s}_j^u$, die verworfen werden, wird a_{ij} gleich Null gesetzt. Danach wird auf $\mathbf{A}$ ein IPF Algorithmus angewendet, um die Werte der Zellen mit $a_{ij} > 0$ derart zu verändern, dass die Reihen- und Spaltensummen mit den Querschnittdesigns $p^t(\cdot)$ bzw. $p^u(\cdot)$ übereinstimmen. Des Weiteren geben Matei & Tillé (2005b) noch Bedingungen an, unter denen das Maximum bzw. Minimum von $\mathfrak{n}^{t,u}$ erreicht werden kann. Die Ziehung der Stichproben kann wie folgt durchgeführt werden: $\vec{\mathfrak{s}}^t$ wird nach $p^t(.)$ ausgewählt und $\vec{\mathfrak{s}}^u$ wird für $\vec{\mathfrak{s}}^t = \vec{s}^t$ nach dem Design

$$p(\vec{s}^u | \vec{s}^t) = \frac{p(\mathbf{S})}{\sum_{\mathbf{S} | \bigcap_{k \in \mathscr{U}} (\mathfrak{I}_k^t = I_k^t)} p(\mathbf{S})}$$

gezogen.

Die Bestimmung des gemeinsamen Designs mittels IPF hat den Vorteil, dass der Algorithmus einfach zu implementieren ist und schnell konvergiert, im Gegensatz zur linearen Programmierung, einem ebenfalls in der Literatur diskutierten Ansatz zur Lösung des Koordinationsproblems. Zudem sind auch andere Koordinationen, die zwischen der Maximierung und Minimierung des Erwartungswerts von $\mathfrak{n}^{t,u}$ liegen, umsetzbar. Wenn beispielsweise eine feste Überlappung gewünscht ist, werden nur gemeinsame Stichproben zugelassen, die diese Überlappung aufweisen. Für alle anderen wird $a_{ij} = 0$ gesetzt. Voraussetzung für die Durchführung des IPF ist, dass vor dem Start des Algorithmus keine Reihen oder Spaltensumme in $\mathbf{A}$ Null sein dürfen, was dem Ausschluss einer Querschnittstichprobe gleichkommen würde.

Das durch den IPF erzeugte gemeinsame Design ist jedoch weder stickt noch schwach sequenziell, da für $t < u$ zur Ziehung von $\vec{\mathfrak{s}}^t$ das Design $p^u(.)$ bekannt sein muss und umgekehrt. Ein weiterer Nachteil des Verfahrens lieget bei der Notwendigkeit, alle Stichproben der Querschnittdesigns zu nummerieren, was die Anwendbarkeit auf kleine Populationen bzw. kleine Träger der Querschnittdesigns beschränkt. Zudem konvergiert der IPF Algorithmus bei einer dünn besetzten Matrix $\mathbf{A}$ möglicherweise nicht. In diesem Fall müssen um von Null verschiedene Werte für die unzulässigen Kombinationen in $\mathbf{A}$ verwendet werden, wobei diese so nahe an Null wie möglich gewählt werden sollten.

Verfahren vom Typ B bieten die Möglichkeit, ein Rotationsschema direkt und individuell für jede Beobachtung in $\mathscr{U}$ festzulegen. Die allgemeine Form eines Algorithmus zur Umsetzung von strikt sequenziellen Längsschnittdesigns wurde von Nedyalkova et al. (2009, S. 275), beschrieben. Die erste Querschnittstichprobe $\vec{\mathfrak{s}}^1$ wird als Poisson Stichprobe (Hájek, 1964, S. 1493) gezogen. Für alle folgenden Erhebungszeitpunkte werden jeweils die bedingten Inklusionswahrscheinlichkeiten $\pi_k^{t|t-1,\dots,1} = Pr(\mathfrak{I}_k^t = 1|\mathfrak{I}_k^{t-1} = I_k^{t-1},\dots,\mathfrak{I}_k^1 = I_k^1)$ für alle $k \in \mathscr{U}^t$ berechnet. Die Ziehung erfolgt wiederum als Poisson Stichprobe, mit $\mathfrak{I}_k^t = 1$ wenn $u_k^t \leq \pi_k^{t|t-1,\dots,1}$ und $\mathfrak{I}_k^t = 0$ sonst und $u_k^t \sim \text{Unif}(0, 1)$. Insbesondere systematische Längsschnittdesigns[3] lassen sich so umsetzten, dass u_k^t sich als eine lineare Transformation von u_k^1 darstellen lässt (Nedyalkova et al., 2009).

Ein systematisches Design eignet sich oft besonderes gut als Längsschnittdesign, wenn eine negative Koordination für eine Reihe aufeinander folgende Beobachtungszeitpunkte erwünscht ist. Diese haben aufgrund ihres, im Vergleich zur einfachen Zufallsstichprobe oder Poisson Stichprobe, besonders kleinen Trägers die Eigenschaft den Abstand zur nächstmöglichen Ziehung $\min(\phi_k^t)$ zu maximieren. In diesem Zusammenhang wird ein systematisches Design auch als *minimales Träger* Design bezeichnet. Dies bedeutet, dass unter der Annahme eines fixen Stichprobenumfangs es unter allen Designs mit gleichen Inklusionswahrscheinlichkeiten erster Ordnung den kleinsten möglichen Träger hat (Qualité, 2009, Kapitel 2). Nedyalkova et al. (2009, S. 278) gegeben Algorithmus 1 an als eine mögliche Selektionsmethode für ein strikt sequenzielles Längsschnittdesign, das sich als geordnetes systematisches Design darstellt. Dabei folgt Algorithmus 1 den gleichen Regeln wie sie Tillé (2006, S. 124) zur Umsetzung eines geordneten systematischem Designs angibt. Der Unterschied ist, dass hier Zeitpunkte ausgewählt werden und nicht Elemente.

Für $V_k^t = \sum_{i=1}^t \pi_k^i$ hat das Längsschnittdesign, wie es von Algorithmus 1 umgesetzt wird, nur im Fall von $V_k^T = n_k$, mit $n_k \in \mathbb{N}$, einen festen Stichprobenumfang. Ansonsten gilt $n_k = \lfloor V_k^T \rfloor$ oder $n_k = \lceil V_k^T \rceil$. Zudem gilt für jede ganze Zahl j mit $j \leq V_k^t$, dass $0 \leq j < \sum_{i=1}^t \mathfrak{I}_k^t$, womit n_k für beliebig lange Ziehungszeiträume eine kontrollierbare Obergrenze hat (Nedyalkova et al., 2009, S. 278). In Algorithmus 1 gilt durch die

---

[3] Für eine Übersicht zur systematischen Ziehung von Stichproben siehe Särndal et al. (1992)

Algorithmus 1 Strikt sequenzielle systematische Stichprobenziehung

```
1:  for $k = 1, \ldots, N$ do
2:      #Kumulative Summe der Inklusionswahrscheinlichkeiten bis Zeitpunkt $t$
3:      $V_k^t \leftarrow f(k,t) := \begin{cases} \sum_{i=1}^{t} \pi_k^i & \text{für } t > 0 \\ 0 & \text{sonst} \end{cases}$
4:      $u_k \leftarrow \mathrm{Unif}(0, 1)$
5:      for $t = 1, \ldots, T$ do
6:          if $\exists\, j \geq 0$, mit $j \in \mathbb{N}_0$, so dass, $V_k^{t-1} \leq u_k + j < V_k^t$ then
7:              $I_k^t \leftarrow 1$
8:          else
9:              $I_k^t \leftarrow 0$
10:         end if
11:     end for
12: end for
```

Verwendung einer sog. *Permanent Random Number* (PRN) u_k in Zeile 1.4[4], für jedes Element $k \in \mathscr{U}$, $u_k = u_k^t \; \forall \, t \in \mathscr{T}$. Somit sind durch die Bestimmung von u_k alle Selektionen des k-ten Elements im Beobachtungszeitraum festgelegt. Die Mächtigkeit des Trägers ist über $card(\{v_k^t\}_{t=1,\ldots,T}) = m$, mit $v_k^t = V_k^t \mod 1$, d.h. also die Anzahl wohl unterscheidbarer Werte v_k^t bestimmbar. Wenn $v_k^{(t)}$ die t-te geordnete Statistik von $v_k^t, t = 1, \ldots, T$, ist, so kann das Intervall $[0,1)$ in $m+1$ nicht leere und nicht überlappende Intervalle

$$[0, v_k^{(1)}), [v_k^{(1)}, v_k^{(2)}), \ldots, [v_k^{(m-1)}, v_k^{(m)}), [v_k^{(m)}, 1) \tag{2.13}$$

unterteilt werden. Jedes Intervall mit $u_k \in [v_k^{(t-1)}, v_k^{(t)})$ korrespondiert eindeutig mit einer Stichprobe $\vec{s}_k$, und seine Länge entspricht $p_k(\vec{s}_k)$ (Qualité, 2009, Kapitel 2). Somit ist im Falle eines systematischen Längsschnittdesigns dessen Träger m_k gegeben mit $m_k = m + 1$. Für ein konstantes n_k ist $m_k \leq T$, da $v_k^T \mod 1 = 0 = v_k^{(1)}$. Bei variablen n_k ist $m_k \leq T + 1$.

Viele der in der Praxis durchgeführten Rotationsstichproben verwenden ein Verfahren, das sich am besten durch eine Koordination vom Typ A beschreiben lässt. Dies ist oft dadurch begründet, dass bei den meisten wiederholt stattfindenden Erhebungen die Querschnittanalyse der Population den höchsten Stellenwert einnimmt und eine mögliche Längsschnittanalyse diesem Aspekt untergeordnet wird. Dies führt meist dazu, dass bei der Gestaltung einer Erhebung die Querschnittdesigns im Vordergrund stehen und die Koordination der Querschnittstichproben dies respektieren soll.

2.4 Algorithmen zur Koordination von Querschnittstichproben

Im Folgenden werden einige Algorithmen zur Ziehung koordinierter Querschnittstichproben sowie die sich daraus ergebenden Längsschnittdesigns und Inklusionswahrscheinlichkeiten beschrieben.

---

[4] Zeile n.m, bezieht sich auf die m-te Zeile, des n-ten Algorithmus.

2.4.1 Koordinierte Poisson Stichproben

Brewer et al. (1972) präsentieren ein Verfahren zur Koordination von Poisson Stichproben wie sie von Hájek (1964, S. 1493) beschrieben werden. Die Vorgehensweise ist leicht zu implementieren. Aufgrund der Verwendung eines Poisson Designs lassen sich auch ungleiche Inklusionswahrscheinlichkeiten einfach realisieren. Zunächst wird eine gleichverteilte Zufallsvariable $u_k^1 \sim \mathrm{Unif}(0,1)$ für jedes Element in der Population unabhängig voneinander erstellt. Die erste Poisson Stichprobe wird nach der Regel $\mathfrak{I}_k^1 = 1 \; \forall \; u_k^1 \leq \pi_k^1$, sonst $\mathfrak{I}_k^1 = 0$, gezogen. Alle folgenden Ziehungen werden über die Konstruktion einer neuen Variable $u_k^t \; \forall \; k \in \mathscr{U}$ koordiniert. Die Ziehung von $\vec{\mathfrak{s}}^t$ erfolgt wiederum nach einem Poisson Design, d.h. $\mathfrak{I}_k^t = 1 \; \forall \; u_k^t \leq \pi_k^t$ sonst $\mathfrak{I}_k^t = 0$. Eine positive Koordination für Element k zwischen Zeitpunkt t und u kann durch eine positive Korrelation zwischen u_k^t und u_k^u erreicht werden. Entsprechendes gilt für eine negative Koordination. Beispielsweise würde $u_k^{t+1} = (a + u_k^t) \bmod 1$, mit $a + u_k^t > 1$ zu einer negativen und für $a + u_k^t < 1$ zu einer positiven Koordination führen. Algorithmus 2 implementiert eine negative Koordination für aufeinander folgende Ziehungen von Poisson Stichproben, (Nedyalkova et al., 2009, S. 283). Dabei wird dieser für alle Elemente im Ziehungsrahmen unabhängig voneinander angewendet. Die Koordinationsvariable u_k^t in Zeile 2.8 stellt sich dabei als eine lineare Transformation von u_k^1 in Zeile 2.1 dar. Beispielsweise gilt für $u_k^1 \leq \pi_k^1$, sowie $\pi_k^1 + \pi_k^2 < 1$ und $\pi_k^1 + \pi_k^2 + \pi_k^3 > 1$, dass $u_k^2 = 1 + u_k^1 - \pi_k^1$ und $u_k^3 = 1 + u_k^1 - \pi_k^1 - \pi_k^2$. Somit ist $I_k^2 = 0$ und $I_k^3 = 1$, wenn $u_k^1 \leq \pi_k^1 + \pi_k^2 + \pi_k^3 - 1$. In der Tat ist das Längsschnittdesign, welches sich aus Algorithmus 2 ableitet, ein systematisches. Die beiden Algorithmen 1 und 2 setzen demnach für gleiche Inklusionswahrscheinlichkeiten π_k^t $(k = 1, \ldots, N\,; t = 1, \ldots, T)$ das identische gemeinsame Design um.

Algorithmus 2 Koordination von Poisson Stichproben

```
 1: $u_k^1 \leftarrow \mathrm{Unif}(0,1)$
 2: if $u_k^1 \leq \pi_k^1$ then
 3:     $I_k^1 \leftarrow 1$
 4: else
 5:     $I_k^1 \leftarrow 0$
 6: end if
 7: for $t = 2, \ldots, T$ do
 8:     $u_k^t \leftarrow (u_k^{t-1} - \pi_k^{t-1}) \bmod 1$
 9:     if $u_k^t \leq \pi_k^t$ then
10:         $I_k^t \leftarrow 1$
11:     else
12:         $I_k^t \leftarrow 0$
13:     end if
14: end for
```

Zur Berechnung der Inklusionswahrscheinlichkeit $\pi_k^{t,u}$ kann es wegen des kleinen Trägers des Längsschnittdesigns praktikabel sein, dies direkt über die Nummerierung aller m_k Stichproben zu tun, so dass

$$\pi_k^{t,u} = \sum_{\vec{s}_k \in \mathscr{S}_k} I_k^t I_k^u p_k(\vec{s}_k) \, .$$

Die Identifizierung aller Stichproben kann dabei durch die in Gleichung (2.13) beschriebenen Intervalle erfolgen (Qualité, 2009, S. 38). Aufgrund des kleinen Träger bzw. der negativen Koordination ist nicht auszuschließen, dass einige $\pi_k^{t,u}$ gleich Null sind, was jedoch mit Blick auf das Rotationsschema durchaus erwünscht sein kann. Wegen der Unabhängigkeit der Längsschnittdesigns $p_k(.)$ und $p_l(.)$ gilt zudem,

$$\pi_{k,l}^t = \pi_k^t \pi_l^t \,,$$
$$\pi_{k,l}^{t,u} = \pi_k^t \pi_l^u \,.$$

Um den Wertebereich von ϕ_k^t anzugeben, muss der erste mögliche Ziehungszeitpunkt, v_k^t und der letzte mögliche, w_k^t bestimmt werden. Hierzu werden zunächst die Mengen $\mathscr{V}_k^t$ und $\mathscr{W}_k^t$ bestimmt mit

$$\mathscr{V}_k^t = \{a_k^t \,|\, a_k^t \in \mathbb{N}, V_k^{t-1} + 1 \leq V_k^{a_k^t}, t < a_k^t \leq T\}$$

und

$$\mathscr{W}_k^t = \{a_k^t \,|\, a_k^t \in \mathbb{N}, V_k^t + 1 \leq V_k^{a_k^t}, t < a_k^t \leq T\} \,.$$

Somit ist

$$v_k^t = \begin{cases} min(\mathscr{V}_k^t) & \text{für } \mathscr{V}_k^t \neq \emptyset \\ T & \text{sonst} \end{cases} \,, \quad w_k^t = \begin{cases} min(\mathscr{W}_k^t) & \text{für } \mathscr{W}_k^t \neq \emptyset \\ T & \text{sonst} \end{cases} \,.$$

Da es zwischen den Zeitpunkten v_k^t und w_k^t höchstens eine Ziehung geben kann, ist die Wahrscheinlichkeitsverteilung von ϕ_k^t für $v_k^t \neq w_k^t$ gegeben durch:

$$w_{\phi_k^t}(x) = \begin{cases} \dfrac{\pi_k^{t,v_k^t}}{\pi_k^t} & \text{für } x = v_k^t - t \\ \dfrac{\pi_k^{t,v_k^t+1}}{\pi_k^t} & \text{für } x = v_k^t - t + 1 \\ \vdots & \vdots \\ 1 - \sum_{i=w_k^t}^{w_k^t-1} \dfrac{\pi_k^{t,i}}{\pi_k^t} & \text{für } w_k^t - t \end{cases} \,.$$

Bei gegebenem u_k^1 in Zeile 2.1 ist ϕ_k^t deterministisch.

Aufgrund der voneinander unabhängigen Längsschnittdesigns kann Algorithmus 2 einfach auf den Fall einer sich verändernden Population angepasst werden. Hierzu wird für jedes zum Zeitpunkt α geborene Element k ein $u_k^\alpha \sim \text{Unif}(0,1)$ bestimmt. Der Algorithmus wird für dieses Element mit einem Startzeitpunkt $t = \alpha$ ausgeführt. Für alle Elemente k, die in ω die Population verlassen, ist $\mathfrak{I}_k^t = 0$ für alle $t \geq \omega$. Zu- und Abgänge der Population können somit zwar zu Veränderungen der zu erwartenden Stichprobenumfänge $E(\mathfrak{n}_k)$ und $E(\mathfrak{n}^t)$ führen, jedoch haben diese darüber hinaus keinen Einfluss auf die Eigenschaft des Querschnitt- und Längsschnittdesigns.

2.4.2 Koordinierte einfache Zufallsstichproben bei nicht veränderlichen Populationen

Algorithmus 3 stellt ein Verfahren zur Koordination von einfachen Zufallsstichproben als Querschnittdesign dar. Die Koordination zu jedem Zeitpunkt erfolgt über die Ordnung aller Elemente in der Population mittels der Ordnungsstatistik $O_{(1)}^t, O_{(2)}^t, \ldots, O_{(N)}^t$

Algorithmus 3 Koordination mittels Ordnungsstatistik

```
1:  $o_k^0 \leftarrow 0 \,\forall\, k \in \mathscr{U}$ #Initialisierung der Koordinationsvariable
2:  $\vec{o}^{\,0} \leftarrow (o_1^0, \dots, o_k^0, \dots, o_N^0)^\top$
3:  for $t \in \mathscr{T}$ do
4:      $I_k^t \leftarrow 0 \,\forall\, k \in \mathscr{U}$
5:      $K^{t-1} \leftarrow card(\{o_k^{t-1} | k \in \mathscr{U}\})$
6:      $\mathscr{U}_{(i)}^{t-1} \leftarrow \{k | k \in \mathscr{U}, O_{(k)}^{t-1} = i\}\ (i = 1, \dots, K^{t-1})$
7:      $f^t(x) \leftarrow \sum_{i=1}^{x} card(\mathscr{U}_{(i)}^{t-1})$
8:      $g^t(y) \leftarrow \max(\{x | x \in \mathbb{N}, f^t(x) < y\})$
9:      $\vec{s}^{\,t*} \leftarrow$ SRS aus $\mathscr{U}_{(g^t(n^t)+1)}^{t-1}$ der Größe $n^t - f^t(g^t(n^t))$
10:     $I_k^t \leftarrow 1 \,\forall\, k \in \bigcup_{i=1}^{g^t(n^t)} \mathscr{U}_{(i)}^{t-1}$
11:     $\vec{s}^{\,t} \leftarrow (I_1^t, \dots, I_k^t, \dots, I_N^t)^\top$
12:     $\vec{s}^{\,t} \leftarrow \vec{s}^{\,t} + \vec{s}^{\,t*\top}$
13:     $\vec{o}^{\,t} \leftarrow h((\vec{s}^{\,1}, \dots, \vec{s}^{\,t}), \vec{o}^{\,t-1}, \boldsymbol{\Theta})$ #Aktualisierung der Koordinationsvariable
14: end for
```

einer sog. Koordinationsvariable $\mathfrak{o}_1^t, \mathfrak{o}_2^t, \dots, \mathfrak{o}_N^t$. Für $\vec{\mathfrak{s}}^{\,t} = \vec{s}^{\,t}$ ist $\mathfrak{o}_1^t = o_k^t$ für alle $k \in \mathscr{U}$. Vor der Ziehung in $t = 1$ besitzen alle Elemente die gleich Ordnung, bzw. der Wert der Koordinationsvariablen ist gleich für alle Elemente in der Population.

Zu jedem Zeitpunkt t lässt sich die Ziehung der Stichprobe $\vec{\mathfrak{s}}^{\,t}$ wie folgt beschreiben: Ist $\mathscr{U}_{(i)}^{t-1} = \{k | k \in \mathscr{U}, O_{(k)}^{t-1} = i\}$, d.h. die Menge von Elementen mit der i-t höchsten Ausprägung der Koordinationsvariable vor der Ziehung der Stichprobe in t,

> werden n^t Elemente aus $\mathscr{U}_{(1)}^{t-1}$ mittels einer einfache Zufallsstichprobe ausgewählt,
>
> ist $\mathscr{U}_{(1)}^{t-1}$ zu klein, wird der Rest aus $\mathscr{U}_{(2)}^{t-1}$ gezogen,
>
> reichen $\mathscr{U}_{(1)}^{t-1}$ und $\mathscr{U}_{(2)}^{t-1}$ nicht aus, wird der Rest aus der $\mathscr{U}_{(3)}^{t}$ gezogen, usw..
>
> Diese Vorgehensweise wird entsprechend fortgesetzt, bis n^t Elemente entnommen sind.

Es werden demnach alle $k \in \bigcup_{i=1}^{v} \mathscr{U}_{(i)}^{t-1}$ mit Sicherheit zum Zeitpunkt t ausgewählt werden, wobei v die kleinste natürliche Zahl ist, für die gilt $\sum_{i=1}^{v} card(\mathscr{U}_{(i)}^{t-1}) \leq n^t$. Des Weiteren werden $n^t - card(\bigcup_{i=1}^{v} \mathscr{U}_{(i)}^{t-1})$ Elemente mittels einer einfachen Zufallsstichprobe aus der Menge $\mathscr{U}_{(v+1)}^{t-1}$ gezogen. Ist die Stichprobe gezogen, werden in Zeile 3.13 die Werte der Koordinationsvariable neu bestimmt. Die Funktion h mit $h((\vec{s}^{\,1}, \dots, \vec{s}^{\,t}), \vec{o}^{\,t-1}, \boldsymbol{\Theta}) : (\mathbb{R}^{N \times t}, \mathbb{R}^{N}, \mathbb{R}^{N \times t}) \mapsto \mathbb{R}^{N}$, beschreibt dabei den Zusammenhang zwischen den Werten der Koordinationsvariable in $t-1$, der Stichprobe in t und einem frei wählbaren Parameter $\boldsymbol{\Theta}$.

Koordination über die Zeit außerhalb der Stichprobe

Algorithmus 3 eignet sich insbesondere zur Umsetzung einer negativen Koordination zwischen benachbarten Beobachtungszeitpunkten mit einem größtmöglichen Mindestabstand. Die negative Koordination kommt dadurch zum Ausdruck, dass $\pi_k^{t,u} = 0$ für $|t-u| \in [\min(t,u) - v_k; \min(t,u) + w_k]$, mit $w_k = \min(\phi_k^{\min(t,u)}) - 1$ und $v_k = \min(\chi_k^{\min(t,u)}) + 1$.

Hierzu werden beispielsweise allen Elementen k, die zum Zeitpunkt t gezogen wurden, entsprechend kleine Werte $\mathfrak{o}_k^t$ zugeordnet. Für jeden Zeitpunkt, in dem ein Element nicht ausgewählt wird, steigt der Wert der Koordinationsvariable wieder an. Für eine Koordination mit $\mathfrak{o}_k^t = \mathfrak{p}_k^t$, wobei $\mathfrak{p}_k^t$ die Zeit ist, die ein Element k nach der Ziehung zum Zeitpunkt t außerhalb der Stichprobe verbracht hat, wird Funktion h in Zeile 3.13 definiert wie in (2.14) beschrieben.

$$\vec{o}^t = h((\vec{s}^1, \ldots, \vec{s}^t), \vec{o}^{t-1}, \mathbf{\Theta}) = h(\vec{s}^t, \vec{o}^{t-1}) := diag(1 - \vec{s}^t)(\vec{o}^{t-1} + 1) \tag{2.14}$$

Algorithmus 3 setzt als Querschnittdesign eine einfache Zufallsstichprobe für jeden Zeitpunkt $t \in \mathscr{T}$ um. Dementsprechend ist $\pi_k^t = \frac{n^t}{N}$ und $\pi_{k,l}^t = \frac{n^t(n^t-1)}{N(N-1)}$. Wenn die Menge unterscheidbarer Werte der Koordinationsvariablen zu jedem Zeitpunkt t $\mathscr{K}^t = \{\mathfrak{o}_k^t | k \in \mathscr{U}\}$ ist, sowie $\mathscr{U}_{(i)}^t = \left\{k | k \in \mathscr{U}, O_{(k)}^t = i\right\}$, dann gilt $\bigcup_i^{K^t} \mathscr{U}_{(i)}^t = \mathscr{U}$, mit $card(\mathscr{K}^t) = K^t$. Des Weiteren ist $Pr(\{k \in \mathscr{U}_i^t\}) = \frac{N_{(i)}^t}{N}$, mit $N_{(i)}^t = card(\mathscr{U}_{(i)}^t)$. Hieraus folgt, dass

$$\begin{aligned}
\pi_k^t &= \sum_i^{K^t} Pr(\mathfrak{I}_k^t = 1 | k \in \mathscr{U}_{(i)}^t) Pr(\{k \in \mathscr{U}_{(i)}^t\}) \\
&= \sum_i^{K^t} \frac{card(\mathscr{s}^t \cap \mathscr{U}_{(i)}^t)}{N_{(i)}^t} \frac{N_{(i)}^t}{N} \\
&= \frac{n^t}{N} ,
\end{aligned} \tag{2.15}$$

als auch, dass

$$\begin{aligned}
\pi_{k,l}^t &= \sum_{i=1}^{K^t} \sum_{j=1}^{K^t} Pr\left(\mathfrak{I}_k^t \mathfrak{I}_l^t = 1 | \{k \in \mathscr{U}_{(i)}^t\} \cap \{l \in \mathscr{U}_{(j)}^t\}\right) Pr\left(\{k \in \mathscr{U}_{(i)}^t\} \cap \{l \in \mathscr{U}_{(j)}^t\}\right) \\
&= 2 \sum_{\substack{j=1 \\ i>j}}^{K^t} \sum_{j=1}^{K^t} \frac{N_{(i)}^t N_{(j)}^t}{N(N-1)} \frac{card(\mathscr{s}^t \cap \mathscr{U}_{(i)}^t) card(\mathscr{s}^t \cap \mathscr{U}_{(j)}^t)}{N_{(i)}^t N_{(j)}^t} \\
&\quad + \sum_i^{K^t} \frac{N_{(i)}^t (N_{(i)}^t - 1)}{N(N-1)} \frac{card(\mathscr{s}^t \cap \mathscr{U}_{(i)}^t)(card(\mathscr{s}^t \cap \mathscr{U}_{(i)}^t) - 1)}{N_{(i)}^t (N_{(i)}^t - 1)} \\
&= \frac{n^t(n^t-1)}{N(N-1)} .
\end{aligned} \tag{2.16}$$

Das Längsschnittdesign von Algorithmus 3 mit (2.14) stellt eine Art von systematischem Design dar, das jedoch in Abhängigkeit von T einen größeren Träger aufweisen kann als ein geordnetes systematisches Design, wie es durch Algorithmus 1 umgesetzt ist. Algorithmus 4 setzt dieses Längsschnittdesign unabhängig für N Elemente um. Im Unterschied zu dem systematischem Längsschnittdesign liegen hier nach der ersten Ziehung eines Elements nicht gleichzeitig alle weiteren Ziehungen fest. Zwar wird auch hier das k-te Element zum ersten Mal genau zum Zeitpunkt t ausgewählt, wenn $V_k^{t-1} \leq u_1 < V_k^t$ für $u_1 \sim \text{Unif}(0, 1)$ gilt. Die nächste Ziehung erfolgt jedoch zum Zeitpunkt $t + a$ wenn $V_k^{t+a-1} \leq u_2 < V_k^{t+a}$ mit $u_2 \sim \text{Unif}(V_k^{t-1} + 1; V_k^t + 1)$. Dieser Vorgang wird, mit jeweils neuen Grenzen für die gleichverteilte Zufallsvariable wiederholt, bis $u_{n_k+1} \geq V_k^T$. Im Allgemeinen ist der Stichprobenumfang $\mathfrak{n}_k$ des Designs eine Zufallsvariable mit $\text{E}(\mathfrak{n}_k) = \sum_{t=1}^T \pi_k^t$.

Algorithmus 4 Längsschnittdesign für Koordination über die Zeit außerhalb der Stichprobe

```
for $k = 1, \ldots, N$ do
    $V_k^t \leftarrow f(k,t) := \begin{cases} \sum_{i=1}^t \pi_k^i & \text{für } t > 0 \\ 0 & \text{sonst} \end{cases}$
    $I_k^t \leftarrow 0 \; \forall t \in \mathscr{T}$
    $u \leftarrow \text{Unif}(0, 1)$
    $i \leftarrow \{i | i \in \mathbb{N}, V_k^{i-1} \leq u < V_k^i\}$
    if $i \neq \emptyset$ then
        $I_k^i \leftarrow 1$
    end if
    NEXT ← TRUE
    while NEXT do
        $v \leftarrow V_k^{i-1} + 1 \; ; w \leftarrow V_k^i + 1$
        $u \leftarrow \text{Unif}(v, w)$
        $j \leftarrow \{j | j \in \mathbb{N}, V_k^{j-1} \leq u < V_k^j\}$
        if $j \neq \emptyset$ then
            $I_k^j \leftarrow 1$
        end if
        if $u \geq V_k^T$ then
            NEXT ← FALSE
        end if
        $i \leftarrow j$
    end while
    $\vec{s}_k \leftarrow (I_k^1, \ldots, I_k^t, \ldots, I_k^T)$
end for
```

Beispiel 2.11. Um die Algorithmen 3 mit (2.14) und 4 zu illustrieren, wird im Folgenden eine Population mit $N = 16$ Elementen betrachtet, die zu $T = 15$ Zeitpunkten beobachtet werden soll. Die Stichprobenumfänge der Querschnittstichproben n^t und die kumulierten Inklusionswahrscheinlichkeiten $V_k^t = V^t = \sum_i^t n^t / N$ sind in Tabelle 2.1 dargestellt.

Tabelle 2.1: n^t und V_k^t

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
n^t	2.00	4.00	3.00	2.00	4.00	4.00	4.00	3.00	2.00	3.00	4.00	3.00	4.00	4.00	3.00
V^t	0.12	0.38	0.56	0.69	0.94	1.19	1.44	1.62	1.75	1.94	2.19	2.38	2.62	2.88	3.06

Tabelle 2.2 stellt die Werte der Koordinationsvariable für alle 16 Elemente der Population, in einer absteigenden Ordnung, vor der Stichprobenziehung zu den jeweiligen Zeitpunkten dar. Der Rang blau umrahmter Werte entspricht dem Stichprobenumfang in der jeweiligen Periode. Kommt der blau umrahmte Wert mehr als einmal in der Periode vor, so wird eine Stichprobe aus der Menge der Elemente, die zu diesen Werten korrespondieren, gezogen. Der Umfang dieser Stichprobe entspricht in den ersten fünf Perioden n^t $(t = 1, 2, 3, 4, 5)$. In der sechsten Periode $n^6 - 3$ und in der siebten $n^7 - 3$, usw.. Die Häufigkeiten der Koordinationsvariablen in Tabelle 2.2 sind unabhängig von der Stichprobenziehung, d.h. Tabelle 2.2 hat für alle gemeinsamen Stichproben die gleiche Form.

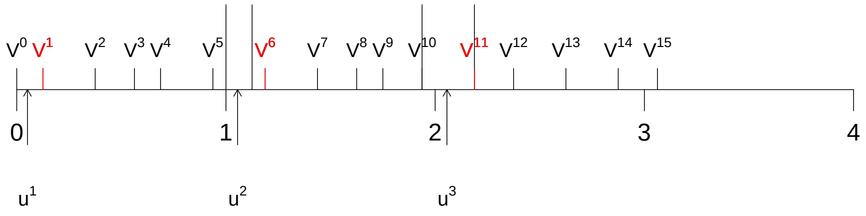
Tabelle 2.2: Werte der Koordinationsvariable für $\mathfrak{o}_k^t = \mathfrak{p}_k^t$

	t														
$O_{(k)}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0	1	2	3	4	5	4	4	4	4	4	4	5	4	4
2	0	1	2	3	4	4	4	4	3	4	4	4	4	4	3
3	0	1	2	3	4	4	4	3	3	4	4	4	4	3	3
4	0	1	2	3	4	3	3	3	3	3	4	4	4	3	3
5	0	1	2	3	4	3	3	2	3	3	3	3	3	3	3
6	0	1	2	3	3	3	3	2	2	3	3	3	3	2	2
7	0	1	2	3	3	3	2	2	2	3	3	3	2	2	2
8	0	1	2	2	2	2	2	2	2	2	3	2	2	2	2
9	0	1	2	2	2	2	1	1	2	2	2	2	2	2	1
10	0	1	2	1	2	2	1	1	1	2	2	1	1	1	1
11	0	1	1	1	2	1	1	1	1	2	2	1	1	1	1
12	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	0	1	0	1	1	0	0	0	1	1	1	0	1	0	0
14	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Ein möglicher Ablauf von Algorithmus 4 kann wie folgt beschrieben werden. Wenn die gleichverteilte Zufallsvariable in Zeile 4.4 den Wert $u^1 = 0{,}0513$ annimmt, können die Auswahlregeln, wie sie Abbildung 2.1 dargestellt, wie folgt aussehen:

- $V_k^0 \leq u^1 < V_k^1$, Auswahl in $t = 1$;
- $u \sim \text{Unif}(1, V_k^1 + 1)$ in Zeile 4.12;
- für $u = u^2 = 1{,}055629$, Auswahl in $t = 6$, da $V_k^5 \leq u^2 < V_k^6$;
- $u \sim \text{Unif}(V_k^5 + 1, V_k^6 + 1)$ in Zeile 4.12;
- für $u = u^3 = 2{,}056221$, Auswahl in $t = 11$, da $V_k^{10} \leq u_3 < V_k^{11}$;
- $u \sim \text{Unif}(V_k^{10} + 1, V_k^{11} + 1)$ in Zeile 4.12;
- für $u = u^4 = 3{,}148989$, keine weitere Auswahl, da $u^4 \geq V_k^T$.

Abbildung 2.1: Längsschnittstichprobe mit Koordination über die Zeit außerhalb der Stichprobe



△

Es kann angemerkt werden, dass Algorithmus 4 für beliebige Inklusionswahrscheinlichkeiten angewendet werden kann, wobei sich als Querschnittdesigns immer ein Poisson Design mit $\mathrm{E}\left(\mathrm{n}^t\right) = \sum_k^N \pi_k^i$ und $\mathrm{V}\left(\mathrm{n}^t\right) = \sum_{k=1}^N \pi_k^t(1-\pi_k^t)$, ergibt. Des Weiteren ist aufgrund der voneinander unabhängigen Ziehungen der einzelnen Elemente $\pi_{k,l}^t = \pi_k^t \pi_l^t$ und $\pi_{k,l}^{t,u} = \pi_k^t \pi_l^u$.

Die Inklusionswahrscheinlichkeiten $\pi_k^{t,u}$ unter Algorithmus 3 mit (2.14) lassen sich analytisch in keiner geschlossenen Form darstellen. Für konstante Stichprobenumfänge der Querschnittstichproben über den Beobachtungszeitraum ist dies jedoch in einer einfachen Form möglich. Darum soll im Folgenden zur weiteren Beschreibung des durch Algorithmus 3 und (2.14) implementieren Designs zunächst von folgendem Spezialfall ausgegangen werden, dass

$$n^t = n \quad \forall\, t \in \mathscr{T} \; . \tag{2.17}$$

Unter der Annahme von (2.17) gilt, $\pi_k^t = \pi = n/N$. Des Weiteren sei $r = \lceil N/n \rceil$, also die kleinste natürliche Zahl, für die gilt $r\pi > 1$. Die Zufallsvariable ϕ_k^t, für $t \leq T - r$, hat somit eine Bernoulli-Verteilung und ihre Wahrscheinlichkeitsfunktion $w_{\phi_k^t}(x|\rho)$ ist gegeben durch

$$w_{\phi_k^t}(x|\rho) = \begin{cases} \rho & \text{für } x = r-1 \\ 1-\rho & \text{für } x = r \\ 0 & \text{sonst} \end{cases} \; , \tag{2.18}$$

$$\text{mit } \frac{rn - N}{n} = \rho \; .$$

Für $t > T - r$ ist $Pr(\phi_k^t = T - t) = 1$. Daraus folgt, dass nach der Selektion eines Elements dieses mindestens in $r-1$ und höchstens in r Zeitpunkten wieder in die Stichprobe gelangen kann bzw. muss. Nach jeder Ziehung von k in t, für $T - t \geq r$ ergeben sich jeweils zwei weitere mögliche Längsschnittstichproben. Dies bedeutet, dass der Träger des Designs in Algorithmus 4 unter der Annahme von (2.17) für hinreichend große T größer als T bzw. $T+1$ werden kann. Nur für $T < r$ ist $card(\mathscr{S}_k) = T$ sicher. So wird sich $card(\{\vec{s}_k \in \mathscr{S}_k | I_k^1 = 1\})$ mit steigendem T alle r Zeitpunkte verdoppeln. Entsprechendes gilt auch für die Menge der Stichproben, welche k zum ersten Mal in

$t = 2, \dots, r$ einschließen. Für den Fall, dass N durch n teilbar ist, gilt $Pr(\phi_k^t = N/n) = 1$ und damit $card(\mathscr{S}_k) = N/n$ wenn $T \geq r$, sonst ist ebenfalls $card(\mathscr{S}_k) = T$.

Unter (2.17) kann die daraus folgende Verteilung von $\phi_k^t \; \forall \, k \in \mathscr{U}$ in (2.18) zur Bestimmung von $\pi_k^{t,u} \; \forall \, t{,}u \in \mathscr{T}$ genutzt werden. Hierzu wird zunächst $\pi_k^{1,u} \; \forall \, u \in \mathscr{T}$ bestimmt. So ist

$$\pi_k^{1,u} = \begin{cases} \sum\limits_{i \in \mathscr{A}} \binom{i}{x_i} \rho^{x_i} (1-\rho)^{i-x_i} \dfrac{n}{N} & \text{für } \mathscr{A} \neq \emptyset \\ 0 & \text{sonst} \end{cases} \quad , \tag{2.19}$$

mit $\mathscr{A} = \left\{ i \in \mathbb{N} | \left\lceil \frac{u-1}{r} \right\rceil \leq i \leq \left\lfloor \frac{u-1}{r-1} \right\rfloor \right\}$ und $x_i = -(u-1) + ir$.

Dabei lässt sich das Ergebnis in (2.19) wie folgt demonstrieren. Für alle $u \geq 2$ mit $\pi_k^{1,u} > 0$ existiert ein $(x,y) \in \mathbb{N}_0^2$, so dass

$$c = xa + yb \, , \tag{2.20}$$

mit $c = u - 1$, $a = r - 1$ und $b = r$. Da (2.20) eine lineare diophantische Gleichung ist, hat sie eine Lösung für alle

$$(x,y) \in \left\{ \left(x_0 + i \frac{b}{ggT(a,b)}, y_0 - i \frac{a}{ggT(a,b)} \right) | i \in \mathbb{Z} \right\} \, ,$$

wobei (x_0, y_0) eine Partikularlösung von (2.20) ist (siehe hierzu Lemma von Bézout in Meyberg, 1980, S. 43). Da a und b teilerfremd sind, kann $x_0 = -c$, $y_0 = c$ verwendet werden. Somit ist die Gesamtheit aller Lösungen von (2.20)

$$\mathscr{L} = \{(-c + ib, c - ia) | i \in \mathbb{Z}\}$$

Die Gesamtheit aller Lösungen von (2.20) in $\mathbb{N}_0^2$ ist

$$\mathscr{L} \cap \mathbb{N}_0^2 = \left\{ (-c + ib, c - ia) | i \in \mathbb{N}, \left\lceil \frac{u-1}{r} \right\rceil \leq i \leq \left\lfloor \frac{u-1}{r-1} \right\rfloor \right\} \, .$$

Sei nun c in (2.20) gleich

$$c = \mathfrak{x} a + \mathfrak{y} b \, , \tag{2.21}$$

wobei $\mathfrak{x}$ und $\mathfrak{y}$ Zufallsvariablen sind und a, b wie in (2.20). Die Zufallsvariablen $\mathfrak{x}$ und $\mathfrak{y}$ beschreibt, wie häufig ein Element k, gegeben seiner Ziehung in $t = 1$, bis zum Zeitpunkt u nach jeweils $r - 1$, bzw. nach jeweils r Zeitpunkten, gezogen wurde. Somit ist nach (2.18)

$$Pr((\mathfrak{x}, \mathfrak{y}) = (x,y)) = \binom{x+y}{x} \rho^x (1-\rho)^y \, ,$$

woraus das Resultat in (2.19) folgt.

Alle weiteren $\pi_k^{t,u}$, mit $t \neq 1$, sind durch $\pi_k^{1,u}$ gegeben, da $\pi^{t,u} = \pi^{1,u-t+1}$ für $u = 2, \dots, T$ gilt. Zudem ist $\pi_k^{t,u}$ eine Konstante für alle $k \in \mathscr{U}$.

In Anlehnung an Wood (2008) gilt für die Inklusionswahrscheinlichkeiten $\pi_{k,l}^{t,u}$, die sich aus Anwendung von Algorithmus 3 mit (2.14) ergeben

$$\begin{aligned}\pi_{k,l}^{t,u} =& Pr(\mathfrak{I}_k^t = 1, \mathfrak{I}_l^t = 1) Pr(\mathfrak{I}_l^u = 1 | \mathfrak{I}_l^t = 1) \\ &+ Pr(\mathfrak{I}_k^t = 1, \mathfrak{I}_l^t = 0) Pr(\mathfrak{I}_l^u = 1 | \mathfrak{I}_l^t = 0) \text{ und} \qquad (2.22a)\\ \pi_{k,l}^{t,u} =& Pr(\mathfrak{I}_k^u = 1, \mathfrak{I}_l^u = 1) Pr(\mathfrak{I}_k^t = 1 | \mathfrak{I}_k^u = 1) \\ &+ Pr(\mathfrak{I}_k^u = 1, \mathfrak{I}_l^u = 0) Pr(\mathfrak{I}_k^t = 1 | \mathfrak{I}_k^u = 0)\ , \qquad (2.22b)\end{aligned}$$

Die Darstellungen in (2.22) basieren für $u > t \geq 1$ auf den Annahmen, dass

$$\begin{aligned} Pr(\mathfrak{I}_k^u = 1 | \mathfrak{I}_k^t,\ \mathfrak{I}_l^t) = Pr(\mathfrak{I}_k^u = 1 | \mathfrak{I}_k^t) \text{ und} \qquad (2.23a)\\ Pr(\mathfrak{I}_k^t = 1 | \mathfrak{I}_k^u,\ \mathfrak{I}_l^u) = Pr(\mathfrak{I}_k^t = 1 | \mathfrak{I}_k^u)\ . \qquad (2.23b)\end{aligned}$$

Dabei folgt (2.22a) aus (2.23a) und (2.22b) aus (2.23b). Insbesondere ist unter der Gültigkeit von (2.23) $\pi_{k,l}^{t,u} = \pi_{k,l}^{u,t}$, wie in (2.22) dargestellt.

Ein gemeinsames Design erfüllt (2.23a), wenn die bedingte Wahrscheinlichkeit, dass ein Element in Stichprobe $\vec{\mathfrak{s}}^u$ gelangt, geben $\vec{\mathfrak{s}}^t$, unabhängig davon ist, ob ein anderes Element zum Zeitpunkt t gezogen wird oder nicht. (2.23b) ist erfüllt, wenn für $\vec{\mathfrak{s}}^u$ gegeben, im Rückschluss sich für ein Element keine Einschränkungen seiner möglichen Inklusion in $\vec{\mathfrak{s}}^t$ ergeben, in Abhängigkeit von der Ziehung eines anderen Elements in u. Wood (2008) führt nur Bedingung (2.23a) an, jedoch ist Bedingung (2.23b) in Einklang mit (2.23a). Denn hängt die Selektion eines Element nur von dessen eigener Ziehungshistorie ab, so besteht im Rückblick auch keine Einschränkung bezüglich der Selektion eines anderen Elementes. Somit lassen sich die Bedingungen in (2.23) auch interpretieren als

$$Pr(\mathfrak{I}_k^u = I_k^u | \vec{\mathfrak{s}}^t = \vec{s}^t) = Pr(\mathfrak{I}_k^u = I_k^u | \mathfrak{I}_k^t = I_k^t)\ ,$$

bzw.

$$Pr(\mathfrak{I}_k^t = I_k^t | \vec{\mathfrak{s}}^u = \vec{s}^u) = Pr(\mathfrak{I}_l^t = I_k^t | \mathfrak{I}_k^u = I_k^u)\ .$$

Wood (2008) betrachtet den Fall von zwei aufeinander folgenden Beobachtungszeitpunkten. Aus der Verallgemeinerung dieser Eigenschaft, dass die Ziehung eines Elements nur von der eigenen Ziehungshistorie, aber nicht von der eines anderen Elements abhängig ist, folgt, dass die nachstehenden Eigenschaften ebenfalls, für $t < u$ gelten:

$$\begin{aligned} Pr(\mathfrak{I}_k^u = I_k^u | \mathfrak{I}_k^{t+1},\ \mathfrak{I}_k^t,\ \mathfrak{I}_l^{t+1},\ \mathfrak{I}_l^t)\ , &= Pr(\mathfrak{I}_k^u = I_k^u | \mathfrak{I}_k^{t+1},\ \mathfrak{I}_k^t)\ , \\ Pr(\mathfrak{I}_k^u = I_k^u | \mathfrak{I}_k^{t+2},\ \mathfrak{I}_k^{t+1},\ \mathfrak{I}_k^t,\ \mathfrak{I}_l^{t+2},\ \mathfrak{I}_l^{t+1},\ \mathfrak{I}_l^t) &= Pr(\mathfrak{I}_k^u = I_k^u | \mathfrak{I}_k^{t+2},\ \mathfrak{I}_k^{t+1},\ \mathfrak{I}_k^t)\ , \qquad (2.24a)\\ \vdots \qquad & = \qquad \vdots \\ Pr(\mathfrak{I}_k^u = 1 | \mathfrak{I}_k^{u-1},\ \ldots,\ \mathfrak{I}_k^t,\ \mathfrak{I}_l^{u-1},\ \ldots,\ \mathfrak{I}_l^t) &= Pr(\mathfrak{I}_k^u = I_k^u | \mathfrak{I}_k^{u-1},\ \ldots,\ \mathfrak{I}_k^t)\ ,\end{aligned}$$

$$\begin{aligned} Pr(\mathfrak{I}_k^t = I_k^t | \mathfrak{I}_k^{t+1},\ \mathfrak{I}_l^{t+1}) &= Pr(\mathfrak{I}_k^t = I_k^t | \mathfrak{I}_k^{t+1})\ , \\ Pr(\mathfrak{I}_k^t = I_k^l | \mathfrak{I}_k^{t+1},\ \mathfrak{I}_k^{t+2},\ \mathfrak{I}_l^{t+1},\ \mathfrak{I}_l^{t+2}) &= Pr(\mathfrak{I}_k^t = I_k^t | \mathfrak{I}_k^{t+1},\ \mathfrak{I}_k^{t+2})\ , \qquad (2.24b)\\ \vdots \qquad & = \qquad \vdots \\ Pr(\mathfrak{I}_k^t = I_k^t | \mathfrak{I}_k^{t+1},\ \ldots,\ \mathfrak{I}_k^u,\ \mathfrak{I}_l^{t+1},\ \ldots,\ \mathfrak{I}_l^u) &= Pr(\mathfrak{I}_k^t = I_k^t | \mathfrak{I}_k^{t+1},\ \ldots,\ \mathfrak{I}_k^u)\ .\end{aligned}$$

Ob Eigenschaften (2.23) erfüllt sind, lässt sich aus der Verteilung von ϕ_k^t und χ_k^t ableiten. Für Algorithmus 3 mit (2.14) trifft dies zu, da $\phi_k^t = \phi^t$ und $\chi_k^t = \chi^t$, d.h. ϕ_k^t und χ_k^t sind iid für alle $k \in \mathscr{U}$. (Die Annahme von (2.17) führt dazu, dass ϕ_k^t diese Eigenschaft auch für alle $t \in \mathscr{T}$ besitzt.) Wood (2008) zeigt, das sich unter der Gültigkeit von (2.23) $\pi_{k,l}^{t,u}$ für $t < u$ darstellen lässt als

$$\pi_{k,l}^{t,u} = \begin{cases} \pi_k^t \pi_l^u & \text{wenn } \pi_l^t = 1 \\ \dfrac{\pi_l^{t,u}}{\pi_l^t} \pi_{k,l}^t + \dfrac{(\pi_l^u - \pi_l^{t,u})(\pi_l^t - \pi_{k,l}^t)}{(1 - \pi_l^t)} & \text{sonst} \end{cases} , \tag{2.25}$$

$\forall k,l \in \mathscr{U}$ und $\forall t,u \in \mathscr{T}$. Für $t > u$ sind die Indices k und l in (2.25) auf der rechten Seite zu vertauschen.

Unter der Annahme von (2.17) ist somit

$$\begin{aligned} \pi_{k,l}^{t,u} &= \pi_l^{t,u} \frac{n-1}{N-1} + \left(\frac{n}{N} - \pi_l^{t,u} \right) \frac{n}{N-1} \\ &= \frac{n^2 - N\pi_l^{t,u}}{N(N-1)} . \end{aligned} \tag{2.26}$$

Da hier $N\pi_l^{t,u} = E(\mathfrak{n}^{t,u})$ gilt, ist auch $\pi_{k,l}^{t,u}$ eine Konstante für alle $k,l \in \mathscr{U}$ und die Inklusionswahrscheinlichkeiten in (2.26) und (2.19) lassen sich auch schreiben als

$$\pi_k^{t,u} = \pi_l^{t,u} = \frac{E(\mathfrak{n}^{t,u})}{N} \tag{2.27}$$

$$\pi_{k,l}^{t,u} = \frac{n^2 - E(\mathfrak{n}^{t,u})}{N(N-1)} . \tag{2.28}$$

Die Darstellungen in (2.27) und (2.28) entsprechen auch jenen von Tam (1984), welche dieser für seinen sog. *Sampling Plan A* angibt.

Wird nun Annahme (2.17) fallen gelassen, hat dies zur Folge, dass der Wertebereich von ϕ_k^t nicht weiter auf nur maximal zwei Werte beschränkt ist. Die kleinstmögliche Ausprägung von ϕ_k^t, v_k^t und die größtmögliche w_k^t sind gegeben durch

$$v_k^t = \begin{cases} \min(a \in \mathscr{V}^t) & \text{wenn } \mathscr{V}^t \neq \emptyset \\ T - t & \text{sonst} \end{cases} , \tag{2.29}$$

$$\text{mit } \mathscr{V}^t = \left\{ a | \forall a \in \{1, \ldots, T-t\}, \sum_{i=0}^{a} n^{t+i} > N \right\}$$

und

$$w_k^t = \begin{cases} \min(a \in \mathscr{W}^t) & \text{wenn } \mathscr{W}^t \neq \emptyset \\ T - t & \text{sonst} \end{cases} , \tag{2.30}$$

$$\text{mit } \mathscr{W}^t = \left\{ a | \forall a \in \{1, \ldots, T-t\}, \sum_{i=1}^{a} n^{t+i} \geq N \right\} .$$

Die Wahrscheinlichkeitsfunktion von ϕ_k^t, $w_{\phi_k^t}(x)$ ist

$$w_{\phi_k^t}(x) = \begin{cases} \min(1, \frac{\sum_{i=0}^{v_k^t} n^{t+i} - N}{n^t}) & \text{für } x = v_k^t \\ \max(0, 1 - \frac{\sum_{i=0}^{v_k^t} n^{t+i} - N}{n^t}) & \text{für } x = v_k^t + 1 \\ 0 & \text{sonst} \end{cases} , \tag{2.31}$$
$$\text{wenn } w_k^t - v_k^t \leq 1$$

und

$$w_{\phi_k^t}(x) = \begin{cases} \frac{\sum_{i=0}^{v_k^t} n^{t+i} - N}{n^t} & \text{für } x = v_k^t \\ \frac{n^{t+v_k^t+1}}{n^t} & \text{für } x = v_k^t + 1 \\ \vdots & \\ \frac{n^{t+v_k^t+j}}{n^t} & \text{für } x = v_k^t + j \\ \vdots & \\ \frac{N - \sum_{i=1}^{w_k^t - 1} n^{t+i}}{n^t} & \text{für } x = w_k^t \\ 0 & \text{sonst} \end{cases} , \tag{2.32}$$
$$\text{wenn } w_k^t - v_k^t > 1 \, .$$

Unter Verwendung der in (2.10) beschriebenen Beziehung lässt sich $\pi_k^{t,u}$, für $u < T$ bestimmen als

$$\begin{aligned} \pi_k^{t,u} &= \pi_k^t Pr(\phi_k^t = u - t) \\ &\quad + \mathbb{1}((u-t) > 1) \sum_{i=2}^{u-t} \left(\sum_{\mathbf{a} \in \mathscr{A}^i} \pi_k^t Pr(\phi_k^t = a_1) \prod_{j=2}^{i} Pr(\phi_k^{t + \sum_{v=0}^{j-1} a_v} = a_j) \right) , \end{aligned} \tag{2.33}$$

falls $\mathscr{V}^t \neq \emptyset$ in (2.29), sonst $\pi_k^{t,u} = 0$. Die Vereinfachung des Zusammenhangs in (2.10) durch die Verwendung der unbedingten Wahrscheinlichkeiten ist möglich, da in Algorithmus 3 mit (2.14) die Verteilung von ϕ_k^t unabhängig ist von allen $\mathfrak{I}_k^1, \ldots, \mathfrak{I}_k^{t-1}$. So ist hier beispielsweise

$$\begin{aligned} Pr(\phi_k^{t+a_1} = a_2 | \phi_k^t = a_1) = Pr \Bigg(\mathfrak{I}_k^{t+a_1+a_2} = 1, \sum_{j=1}^{a_1+a_2-1} \mathfrak{I}_k^{t+j} = 0 \Bigg| & \\ \mathfrak{I}_k^{t+a_1} = 1, \sum_{j=1}^{a_1-1} \mathfrak{I}_k^{t+j} = 0, \mathfrak{I}_k^t = 1 \Bigg) & \\ = Pr \left(\mathfrak{I}_k^{t+a_1+a_2} = 1, \sum_{j=1}^{a_1+a_2-1} \mathfrak{I}_k^{t+j} = 0 \middle| \mathfrak{I}_k^{t+a_1} = 1 \right) & \\ = Pr(\phi_k^{t+a_1} = a_2) \, . & \end{aligned}$$

Die Koordination richtet sich nur nach der Zeit, die seit der letzten Ziehung von k vergangen ist. Es wird beispielsweise nicht berücksichtigt, wie häufig k vor dem Zeitpunkt

t schon gezogen wurde oder wie die Verteilung der Zeitabstände zwischen diesen Ziehungen ausfällt.

Die Effizienz der Berechnung von $\pi_k^{t,u}$ lässt sich steigern, indem nur die Tupel $\mathbf{a} \in \mathscr{A}^i$, $i = 2, \ldots, u-t$, betrachtet werden, für welche der zweite Term auf der rechten Seite von (2.33) positiv ist. Ist $\mathscr{B}_{\phi_k^t}$ der Wertebereich von ϕ_k^t, mit $\mathscr{B}_{\phi_k^t} = \{v_k^t, v_k^t + 1, \ldots, w_k^t\}$, lassen sich auch die Wertebereiche von $\phi_k^{t+v_k^t}$, $\phi_k^{t+v_k^t+1}$, …, $\phi_k^{t+w_k^t}$ bestimmen. Entsprechendes gilt für

$$\begin{array}{llll}
\phi_k^{t+v_k^t+v_k^{t+v_k^t}}, & \phi_k^{t+v_k^t+v_k^{t+v_k^t}+1}, & \ldots, & \phi_k^{t+v_k^t+w_k^{t+v_k^t}}, \\
\phi_k^{t+v_k^t+1+v_k^{t+v_k^t+1}}, & \phi_k^{t+v_k^t+1+v_k^{t+v_k^t+1}+1}, & \ldots, & \phi_k^{t+v_k^t+1+w_k^{t+v_k^t+1}}, \\
\vdots & \vdots & \vdots & \vdots \\
\phi_k^{t+w_k^t+v_k^{t+w_k^t}}, & \phi_k^{t+w_k^t+v_k^{t+w_k^t}+1}, & \ldots, & \phi_k^{t+w_k^t+w_k^{t+w_k^t}}.
\end{array}$$

Algorithmus 5 bestimmt auf diese Weise die Wahrscheinlichkeiten aller möglichen Teillängsschnittstichproben $(\mathfrak{I}_k^t = 1, \mathfrak{I}_k^{t+1}, \ldots, \mathfrak{I}_k^{u-1}, \mathfrak{I}_k^u = 1)^\top$, sowie aus deren Summe $\pi_k^{t,u}$.

Algorithmus 5 Algorithmus zur Bestimmung von $\pi_k^{t,u}$

```
$\vec{A}^1 \leftarrow t; \vec{P}^1 \leftarrow \pi_k^t$
$\pi_k^{t,u} \leftarrow 0; i \leftarrow 1$
while any$(\vec{A}^i < u)$ do
    for $j = 1, \ldots, length(\vec{A}^i)$ do
        $o \leftarrow \vec{A}_j^i$
        $\vec{a}_j^{i+1} \leftarrow$ if $v_k^o \neq w_k^o$ then $(v_k^o, v_k^o + 1, \ldots, w_k^o)$ else $v_k^o$
        $\vec{p}_j^{i+1} \leftarrow w_{\phi^o}(\vec{a}_j^{i+1})\vec{P}_j^i$
        $\vec{a}_j^{i+1} \leftarrow \vec{a}_j^{i+1} + o$
    end for
    $\vec{A}^{i+1} \leftarrow \left(\vec{a}_j^{i+1}\right)_{j=1,\ldots,length(\vec{A}^i)}$
    $\vec{P}^{i+1} \leftarrow \left(\vec{p}_j^{i+1}\right)_{j=1,\ldots,length(\vec{A}^i)}$
    $i \leftarrow i + 1$
    for $j = 1, \ldots, length(\vec{A}^i)$ do
        $\pi_k^{t,u} \leftarrow$ if $\vec{A}_j^i = u$ then $\pi_k^{t,u} + \vec{P}_j^i$ else $\pi_k^{t,u}$
    end for
end while
```

Da auch bei ungleichen Stichprobenumfängen der Querschnittstichproben ϕ_k^t iid für alle $k \in \mathscr{U}$ ist, bleiben die Eigenschaften (2.23) für Algorithmus 3 mit (2.14) auch erhalten. Somit ist nach (2.25)

$$\pi_{k,l}^{t,u} = \frac{n^t n^u - N\pi_k^{t,u}}{N(N-1)} . \tag{2.34}$$

Beispiel 2.12. Zu Illustration der Inklusionswahrscheinlichkeiten des gemeinsamen Designs unter Algorithmus 3 mit (2.14) sind diese für die gleichen Stichprobenumfänge wie in Beispiel 2.11 (gegeben in Tabelle 2.1) hier berechnet. Tabellen 2.3 und 2.4 zeigen die Wahrscheinlichkeiten $\mathrm{E}\left(\mathfrak{I}_k^t\mathfrak{I}_k^u\right)$ $(t,u=1,\ldots,T)$ $\forall\, k\in\mathscr{U}$ bzw. $\mathrm{E}\left(\mathfrak{I}_k^t\mathfrak{I}_l^u\right)$ $(t,u=1,\ldots,T)$ $\forall\, k,l\in\mathscr{U}$ $k\neq l$. Dabei wurde $\mathrm{E}\left(\mathfrak{I}_k^t\mathfrak{I}_k^u\right)$ nach Algorithmus 5 und $\mathrm{E}\left(\mathfrak{I}_k^t\mathfrak{I}_l^u\right)$ entsprechend (2.34) berechnet. Die Ergebnisse in beiden Tabellen 2.3 und 2.4 sind auf die zweite Nachkommastelle gerundet.

Tabelle 2.3: $\mathrm{E}\left(\mathfrak{I}_k^t\mathfrak{I}_k^u\right)$ für Algorithmus 3 mit (2.14) und n^t nach Tabelle 2.1

$\mathrm{E}\left(\mathfrak{I}_k^t\mathfrak{I}_k^u\right)$									u							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	1	0,12	0,00	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,06
	2	0,00	0,25	0,00	0,00	0,00	0,06	0,19	0,00	0,00	0,00	0,06	0,14	0,05	0,00	0,03
	3	0,00	0,00	0,19	0,00	0,00	0,00	0,06	0,12	0,00	0,00	0,00	0,05	0,14	0,00	0,00
	4	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,06	0,06	0,00	0,00	0,00	0,06	0,06	0,00
	5	0,00	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,06	0,19	0,00	0,00	0,00	0,19	0,06
	6	0,12	0,06	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,12
	7	0,00	0,19	0,06	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,00	0,19	0,06	0,00	0,00
t	8	0,00	0,00	0,12	0,06	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,00	0,19	0,00	0,00
	9	0,00	0,00	0,00	0,06	0,06	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,00	0,12	0,00
	10	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,12	0,06
	11	0,12	0,06	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,12
	12	0,00	0,14	0,05	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,00
	13	0,00	0,05	0,14	0,06	0,00	0,00	0,06	0,19	0,00	0,00	0,00	0,00	0,25	0,00	0,00
	14	0,00	0,00	0,00	0,06	0,19	0,00	0,00	0,00	0,12	0,12	0,00	0,00	0,00	0,25	0,00
	15	0,06	0,03	0,00	0,00	0,06	0,12	0,00	0,00	0,00	0,06	0,12	0,00	0,00	0,00	0,19

Tabelle 2.4: $\mathrm{E}\left(\mathfrak{I}_k^t\mathfrak{I}_l^u\right)$ $(k\neq l)$ für Algorithmus 3 mit (2.14) und n^t nach Tabelle 2.1

$\mathrm{E}\left(\mathfrak{I}_k^t\mathfrak{I}_l^u\right)$									u							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	1	0,01	0,03	0,03	0,02	0,03	0,03	0,03	0,03	0,02	0,03	0,03	0,03	0,03	0,03	0,02
	2	0,03	0,05	0,05	0,03	0,07	0,06	0,05	0,05	0,03	0,05	0,06	0,04	0,06	0,07	0,05
	3	0,03	0,05	0,03	0,03	0,05	0,05	0,05	0,03	0,03	0,04	0,05	0,03	0,04	0,05	0,04
	4	0,02	0,03	0,03	0,01	0,03	0,03	0,03	0,02	0,01	0,03	0,03	0,03	0,03	0,03	0,03
	5	0,03	0,07	0,05	0,03	0,05	0,07	0,07	0,05	0,03	0,04	0,07	0,05	0,07	0,05	0,05
	6	0,03	0,06	0,05	0,03	0,07	0,05	0,07	0,05	0,03	0,05	0,05	0,05	0,07	0,07	0,04
	7	0,03	0,05	0,05	0,03	0,07	0,07	0,05	0,05	0,03	0,05	0,07	0,04	0,06	0,07	0,05
t	8	0,03	0,05	0,03	0,02	0,05	0,05	0,05	0,03	0,03	0,04	0,05	0,04	0,04	0,05	0,04
	9	0,02	0,03	0,03	0,01	0,03	0,03	0,03	0,03	0,01	0,03	0,03	0,03	0,03	0,03	0,03
	10	0,03	0,05	0,04	0,03	0,04	0,05	0,05	0,04	0,03	0,03	0,05	0,04	0,05	0,04	0,03
	11	0,03	0,06	0,05	0,03	0,07	0,05	0,07	0,05	0,03	0,05	0,05	0,05	0,07	0,07	0,04
	12	0,03	0,04	0,03	0,03	0,05	0,05	0,04	0,04	0,03	0,04	0,05	0,03	0,05	0,05	0,04
	13	0,03	0,06	0,04	0,03	0,07	0,07	0,06	0,04	0,03	0,05	0,07	0,05	0,05	0,07	0,05
	14	0,03	0,07	0,05	0,03	0,05	0,07	0,07	0,05	0,03	0,04	0,07	0,05	0,07	0,05	0,05
	15	0,02	0,05	0,04	0,03	0,05	0,04	0,05	0,04	0,03	0,03	0,04	0,04	0,05	0,05	0,03

△

Koordination über die Zeit außerhalb der Stichprobe und Belastung

Bei gegebenen Stichprobenumfängen der Querschnittstichproben maximiert Algorithmus 3 mit (2.14) den zu erwartenden Abstand zwischen zwei Selektionen eines Elements, d.h. $E(\mathfrak{n}^{t,u})$ wird minimiert. Es wird jedoch nicht die Belastung eines Elements zum jeweiligen Ziehungszeitpunkt mit b_k^t , wie in (2.5) beschrieben, berücksichtigt.

Bei Algorithmus 3 mit (2.14) ist es möglich, dass für $k \neq l$, $\mathfrak{o}_k^t = \mathfrak{o}_l^t$ und $\mathfrak{b}_k^t < \mathfrak{b}_l^t$ dennoch der Fall eintreten kann, dass $Pr(\mathfrak{I}_l^t = 1 | \mathfrak{I}_l^t = I_l^{t-1}, \ldots, \mathfrak{I}_l^1 = I_l^1) > 0$ sowie $Pr(\mathfrak{I}_k^t = 1 | \mathfrak{I}_k^t = I_k^{t-1}, \ldots, \mathfrak{I}_k^1 = I_k^1) < 1$ ist. Dies bedeutet, dass trotz der höheren Belastung des l-ten Elements dieses in t ausgewählt werden kann, wenn gleichzeitig das k-te Element, trotz dessen niedrigerer Belastung, nicht ausgewählt werden muss. So ist bei Algorithmus 3 mit (2.14) ein höheres T auch mit einer potentiellen Vergrößerung des Abstandes zwischen maximaler Belastung $\mathfrak{b}_{\max}^T$ und minimaler Belastung $\mathfrak{b}_{\min}^T$ eines Elements verbunden, wobei

$$\mathfrak{b}_{\max}^t = \max_{k \in \mathscr{U}} \mathfrak{b}_k^t \qquad\qquad \mathfrak{b}_{\min}^t = \min_{k \in \mathscr{U}} \mathfrak{b}_k^t \,. \tag{2.35}$$

Eine Möglichkeit, die unterschiedliche Belastung der Elemente bei der Koordination zu berücksichtigen, ist, diese zu jedem Zeitpunkt aufsteigend nach $\mathfrak{p}_k^t$ zu ordnen und gegeben dieser Ordnung absteigend nach $\mathfrak{b}_k^t$ zu sortieren. Die Koordination erfolgt so primär über die Zeit außerhalb der Stichprobe und gegeben dieser Zeit über die Belastung der Elemente. Der Wert der Koordinationsvariable $\mathfrak{o}_k^t \in \mathbb{Z}$ kann bei dieser Ordnung durch

$$\mathfrak{o}_k^t = (\mathfrak{b}_{\max}^t - \mathfrak{b}_{\min}^t + 1)\mathfrak{p}_k^t - \mathfrak{b}_k^t$$

bestimmt werden. Diese Art von Koordination lässt sich umsetzen, indem Funktion h in Zeile 3.13 definiert wird wie in (2.36) beschrieben.

$$\begin{aligned} h((\vec{s}^1 \ldots, \vec{s}^t), \vec{o}^{\,t-1}, \mathbf{\Theta}) =& h((\vec{s}^1 \ldots, \vec{s}^t), \vec{o}^{\,t-1}, \bar{\mathbf{G}}) \\ &:= \left(b_{\max}^t - b_{\min}^t + 1\right) diag(1 - \vec{s}^t) \left(\frac{\vec{o}^{\,t-1} + \vec{b}^{\,t-1}}{b_{\max}^{t-1} - b_{\min}^{t-1} + 1} + 1 \right) - \vec{b}^t \\ &\text{wobei} \quad \vec{b}^t \text{ die Hauptdiagonale der Matrix } \bar{\mathbf{S}}\bar{\mathbf{G}}^\top \text{ ist,} \\ &\text{mit} \quad \bar{\mathbf{S}} = (\vec{s}^1 \ldots, \vec{s}^t). \end{aligned} \tag{2.36}$$

Dabei ist $\mathbf{\Theta}$ die $N \times t$ Matrix der Erhebungsbelastungen $\bar{\mathbf{G}} = [g_k^i]_{\substack{k=1,\ldots,N \\ i=1,\ldots,t}}$.

Für den weiteren Verlauf der Arbeit wird davon ausgegangen, dass die zusätzliche Belastung für jedes Element durch eine Erhebung zu allen Zeitpunkten die gleiche ist. So wird festgelegt, dass

$$g_k^t = 1 \quad \forall\, k \in \mathscr{U} \;\forall\, t \in \mathscr{T} \,. \tag{2.37}$$

Somit gilt für Algorithmus 3 mit (2.36)

$$\mathfrak{b}_{\max}^t - \mathfrak{b}_{\min}^t \leq 1 \;\forall t \in \mathscr{T} \,. \tag{2.38}$$

Dies bedeutet, dass zu jedem Zeitpunkt ein Teil der Population höchstens einmal mehr gezogen würde als der Rest. Zum Vergleich, für h nach (2.14) und unter der Annahme von (2.17) sowie (2.37) ist

$$\mathfrak{b}_{\max}^T \leq \left\lfloor \frac{T-1}{\lceil \frac{N}{n} \rceil - 1} \right\rfloor + 1 \quad \text{und} \quad \mathfrak{b}_{\min}^T \geq \left\lfloor \frac{T}{\lceil \frac{N}{n} \rceil} \right\rfloor \,.$$

Gilt Bedingung (2.37), ist das durch Algorithmus 3 mit (2.36) umgesetzte Querschnittdesign ebenfalls eine einfache Zufallsstichprobe und π_k^t und $\pi_{k,l}^t$ können dargestellt werden wie in (2.15) und (2.16) beschrieben. Das Längsschnittdesign von Algorithmus 3 mit (2.36) stellt sich in diesem Fall auch als eine Art von systematischem Design dar, jedoch mit einem potentiell kleineren Träger als im Falle der Koordination nach (2.14). So schränkt die Beachtung der Belastung bei der Koordination die minimale und maximale Anzahl an Ziehungen über den Beobachtungszeitraum ein. Algorithmus 6 setzt dieses Längsschnittdesign um. Im Unterschied zu Algorithmus 4 wird hier eine Fallunterscheidung getroffen, bevor der Wertebereich für die nächste zu ziehende gleichverteilte Zufallsvariable bestimmt wird. Für den Fall, dass das zuletzt verwendete Intervall $\left[V_k^{i-1}, V_k^i\right]$ eine natürliche Zahl einschließt, wird geprüft, ob Element k zur Gruppe mit der höheren oder niedrigeren Belastung gehört. Für Elemente aus der Gruppe mit der höheren Belastung bildet diese natürliche Zahl plus eins die Untergrenze und für die Elemente aus der Gruppe mit der geringeren Belastung die Obergrenze des Wertebereichs für die nächste zu ziehende gleichverteilte Zufallsvariable. Dieses Vorgehen stellt sicher, dass die Breite des Intervalls, das den Wertebereich der jeweils nächsten gleichverteilten Zufallsvariable einschließt, der Anzahl der Elemente entspricht, die den gleichen Rang aufweisen, wie Element k nach seiner letzten Ziehung.

Beispiel 2.13. Unter Verwendung der gleichen Population und Stichprobenumfänge im Querschnitt wie in in Beispiel 2.11 (gegeben in Tabelle 2.1) soll eine Auswahl unter den Algorithmen 3 mit (2.36) und 6 demonstriert werden.

Tabelle 2.5 ist das Äquivalent zu Tabelle 2.2 unter Algorithmus 3 mit (2.36). Auch hier gilt: die Häufigkeitsverteilung der Koordinationsvariablen ist nicht abhängig von gemeinsamen der Stichprobe.

Tabelle 2.5: Werte der Koordinationsvariable für $\mathfrak{o}_k^t = (\mathfrak{b}_{\max}^t - \mathfrak{b}_{\min}^t + 1)\mathfrak{p}_k^t - \mathfrak{b}_k^t$

	t														
$O_{(k)}$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0	2	4	6	8	10	7	7	7	7	7	6	8	6	6
2	0	2	4	6	8	7	7	7	5	7	6	6	6	6	4
3	0	2	4	6	8	7	7	5	5	7	6	6	6	4	3
4	0	2	4	6	8	5	5	5	5	5	6	6	6	4	3
5	0	2	4	6	8	5	5	3	5	4	4	4	4	4	3
6	0	2	4	6	5	5	5	3	3	4	4	4	4	2	1
7	0	2	4	6	5	5	3	3	2	4	4	4	2	1	1
8	0	2	4	3	3	3	3	3	2	2	4	2	2	1	1
9	0	2	4	3	3	3	1	1	2	2	2	2	2	1	-1
10	0	2	4	1	3	3	1	0	0	2	2	0	0	-1	-1
11	0	2	1	1	3	1	1	0	0	2	2	0	-1	-1	-1
12	0	2	1	1	1	1	1	0	0	0	0	0	-1	-1	-1
13	0	2	-1	1	1	-1	-1	-2	0	0	0	-2	-1	-3	-3
14	0	2	-1	-1	1	-1	-2	-2	-2	0	-2	-3	-3	-3	-3
15	0	-1	-1	-1	-1	-1	-2	-2	-2	-2	-2	-3	-3	-3	-3
16	0	-1	-1	-1	-1	-1	-2	-2	-2	-2	-2	-3	-3	-3	-3

Zur Veranschaulichung des Ablaufs der Auswahlregeln von Algorithmus 3 mit (2.36) stellt Abbildung A.1 in Appendix A diese anhand einer konkreten Stichprobenziehung dar.

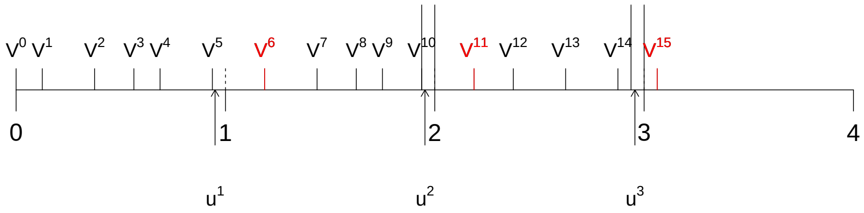
Abbildung 2.2 verdeutlicht die Ziehung einer Stichprobe durch Algorithmus 6. Wenn die gleichverteilte Zufallsvariable in Zeile 6.4 den Wert $u^1 = 0{,}9502325$ annimmt, lassen sich die Auswahlregeln, wie Abbildung 2.2 dargestellt, wie folgt beschreiben :

Algorithmus 6 Längsschnittdesign für Koordination über die Zeit außerhalb der Stichprobe und Belastung.

```
for $k = 1, \dots, N$ do
    $V_k^t \leftarrow f(k,t) := \begin{cases} \sum_{i=1}^{t} \pi_k^i & \text{für } t > 0 \\ 0 & \text{sonst} \end{cases}$
    $I_k^t \leftarrow 0 \; \forall t \in \mathscr{T}$
    $u \leftarrow \text{Unif}(0,\, 1)$
    $i \leftarrow \{i | i \in \mathbb{N}, V_k^{i-1} \leq u < V_k^i\}$
    if $i \neq \emptyset$ then
        $I_k^i \leftarrow 1$
    end if
    NEXT ← TRUE
    while NEXT do
        if $\lfloor V_k^{i-1} \rfloor < \lfloor V_k^i \rfloor$ then
            if $\sum_t^T I_k^t > \lfloor V_k^i \rfloor$ then
                $v \leftarrow \lfloor V_k^i \rfloor + 1 \; ; w \leftarrow V_k^i + 1$
            else
                $v \leftarrow V_k^{i-1} + 1 \; ; w \leftarrow \lfloor V_k^i \rfloor + 1$
            end if
        else
            $v \leftarrow V_k^{i-1} + 1 \; ; w \leftarrow V_k^i + 1$
        end if
        $u \leftarrow \text{Unif}(v,\, w)$
        $j \leftarrow \{j | j \in \mathbb{N}, V_k^{j-1} \leq u < V_k^j\}$
        if $j \neq \emptyset$ then
            $I_k^j \leftarrow 1$
        end if
        if $u \geq V_k^T$ then
            NEXT ← FALSE
        end if
        $i \leftarrow j$
    end while
    $\vec{s}_k \leftarrow (I_k^1, \dots, I_k^t, \dots, I_k^T)$
end for
```

- $V_k^5 \leq u^1 < V_k^6$, Auswahl in $t = 6$;
- $u \sim \text{Unif}(V_k^5 + 1, 2)$ in Zeile 6.20;
- für $u = u^2 = 1{,}953046$, Auswahl in $t = 11$, da $V_k^{10} \leq u^2 < V_k^{11}$;
- $u \sim \text{Unif}(V_k^{10} + 1, 3)$ in Zeile 6.20;
- für $u = u^3 = 2{,}95612$, Auswahl in $t = 15$, da $V_k^{14} \leq u^3 < V_k^{15}$;
- $u \sim \text{Unif}(V_k^{14} + 1, 4)$ in Zeile 6.20;
- für $u = u^4 = 3{,}909049$, keine weitere Auswahl, da $u^4 \geq V_k^T$.

Abbildung 2.2: Längsschnittstichprobe mit Koordination über die Zeit außerhalb der Stichprobe und Belastung



△

Ist $O_{(k)}^t = i$, d.h. Element k hat den i-ten Rang, bzw. o_k^t hat den i-t größten Wert $\forall k \in \mathcal{U}$, so lässt sich die minimale Zeit v_k^t und die maximale Zeit w_k^t bis zur nächsten Selektion nach einer Ziehung von k in t wie folgt bestimmen:

$$v_k^t = \begin{cases} \min(a \in \mathcal{V}_{(i)}^t) & \text{wenn } \mathcal{V}_{(i)}^t \neq \emptyset \\ T - t & \text{sonst} \end{cases} ,$$

$$\text{mit } \mathcal{V}_{(i)}^t = \left\{ a | a \in \{1, \dots, T-t\}, \sum_{i=1}^{a} n^{t+i} > N - N^v \right\} , \tag{2.39}$$

$$N^v = \sum_{v=i}^{K^t} N_{(v)}^t \text{ und } O_{(k)}^t = i \ .$$

$$w_k^t = \begin{cases} \min(a \in \mathcal{W}_{(i)}^t) & \text{wenn } \mathcal{W}_{(i)}^t \neq \emptyset \\ T - t & \text{sonst} \end{cases} ,$$

$$\text{mit } \mathcal{W}_{(i)}^t = \left\{ a | a \in \{1, \dots, T-t\}, \sum_{i=1}^{a} n^{t+i} \geq N - (N^v - N^w) \right\} \tag{2.40}$$

$$\text{und } N^w = \mathbb{1}(O_{(k)}^t \neq K^t) \sum_{v=O_{(k)}^t+1}^{K^t} N_{(v)}^t + \left(1 - \mathbb{1}(O_{(k)}^t \neq K^t)\right) N_{(K^t)}^t$$

Dabei ist N^v die Anzahl der Elemente, die einen gleich hohen oder niedrigeren Rang haben als Element k, und N^w die Anzahl der Elemente mit einem niedrigeren Rang als Element k. Falls k den höchsten Rang hat, ist $N^v = N^w$. Im Vergleich zu (2.29) bzw. (2.30) gilt die obige Darstellung von v_k^t und w_k^t allgemein für Algorithmus 3, unabhängig von der gewählten Definition für die Koordinationsvariable.

So lässt sich die Wahrscheinlichkeitsverteilung von ϕ_k^t gegeben $O_{(k)}^t$, für $w_k^t - v_k^t \leq 1$, schreiben als

$$w_{\phi_k^t|O_{(k)}^t}(x) = \begin{cases} \min\left(1, \dfrac{\sum_{i=1}^{v_k^t} n^{t+i} - (N - N^v)}{N_{(i)}^t}\right) & \text{für } x = v_k^t \\ \max\left(0, 1 - \dfrac{\sum_{i=0}^{v_k^t} n^{t+i} - (N - N^v)}{N_{(i)}^t}\right) & \text{für } x = v_k^t + 1 \\ 0 & \text{sonst} \end{cases} \tag{2.41}$$

und für $w_k^t - v_k^t > 1$

$$w_{\phi_k^t|O_{(k)}^t}(x) = \begin{cases} \dfrac{\sum_{i=1}^{v_k^t} n^{t+i} - (N - N^v)}{N_{(i)}^t} & \text{für } x = v_k^t \\ \dfrac{n^{t+v_k^t+1}}{N_{(i)}^t} & \text{für } x = v_k^t + 1 \\ \vdots & \\ \dfrac{n^{t+v_k^t+j}}{N_{(i)}^t} & \text{für } x = v_k^t + j \\ \vdots & \\ \dfrac{(N - N^v) - \sum_{i=1}^{w_k^t-1} n^{t+i}}{N_{(i)}^t} & \text{für } x = w_k^t \\ 0 & \text{sonst} \end{cases} . \tag{2.42}$$

In Anlehnung an (2.33) lässt sich für Algorithmus (3) mit Koordination über die Zeit außerhalb der Stichprobe und Belastung $\pi_k^{t,u}$ für $u < T$ bestimmen als

$$\begin{aligned} \pi_k^{t,u} = &\sum_{\kappa=x}^{\iota^t} \frac{N_{(\kappa)}^t}{N} Pr(\phi_k^t = u - t | O_{(k)}^t = x) \\ &+ \mathbb{1}((u-t) > 1) \sum_{i=2}^{u-t} \left(\sum_{\vec{a} \in \mathscr{A}^i} \sum_{\kappa=x}^{\iota^t} \frac{N_{(\kappa)}^t}{N} Pr(\phi_k^t = a_1 | O_{(k)}^t = x) \right. \\ &\left. \prod_{j=2}^{i} Pr(\phi_k^{t+\sum_{\nu=0}^{j-1} a_\nu} = a_j | O_{(k)}^t = x) \right) , \end{aligned} \tag{2.43}$$

falls $\mathscr{V}_{(i)}^t \neq \emptyset$ in (2.39) für wenigstens ein $O_{(k)}^t = i \in \{1, \ldots, \iota^t\}$, sonst $\pi_k^{t,u} = 0$. Dabei ist

$$\iota^t = \begin{cases} \max(\mathscr{Z}^t) + 1 & \text{für} \quad \mathscr{Z}^t \neq \emptyset \\ 1 & \text{sonst} \end{cases} ,$$

mit $\mathscr{X}^t = \left\{ j \in \{1, \ldots, K^t\} | \sum_{i=1}^{j} N_{(i)}^t < n^t \right\}$.

Im Vergleich zur Koordination, die nur die Zeit außerhalb der Stichprobe beachtet, ist die Verteilung von ϕ_k^t nicht unabhängig von $\mathfrak{I}_k^1, \ldots, \mathfrak{I}_k^{t-1}$. Ob Element k ein weiteres Mal nach Zeitpunkt t ausgewählt werden kann, hängt von der Belastung des Elements nach der Ziehung in Zeitpunkt t ab. Unter Bedingung (2.37) ist die Verteilung von ϕ_k^t beispielsweise abhängig von der Anzahl der erfolgten Ziehungen bis zum Zeitpunkt t.

Da für Algorithmus 3 mit (2.36) ϕ_k^t weiterhin iid für alle $k \in \mathscr{U}$ mit $t = 1, \ldots, T$ ist, bleiben die Eigenschaften (2.23) erhalten und $\pi_{k,l}^{t\,;u}$ hat folglich die gleiche Form wie in (2.34) beschrieben. Wie bereits angemerkt, erfüllt Algorithmus 3, mit (2.14) oder (2.36), die Eigenschaft einer negativen Koordination zwischen Beobachtungszeitpunkt in einer gewissen Nachbarschaft. So ist $\pi_k^{t\,;u} = 0$ für alle $|t-u| \in [\min(t,u) - v; \min(t,u) + w]$ mit

$$w = \max(\{w | w \in \{0, \ldots, T-t\}, \sum_{i=0}^{w} n^{\min(t,u)+i} \leq N\}) \text{ und}$$

$$v = \max(\{v | v \in \{0, \ldots, t-1\}, \sum_{i=0}^{v} n^{\min(t,u)-i} \leq N\}) \, .$$

Für alle $\pi_k^{t\,;u} > 0$ ist die Koordination positiv, wenn $n^t n^u / N < \mathrm{E}\left(\mathfrak{n}^{t\,;u}\right)$, und sie ist negativ, wenn $n^t n^u / N > \mathrm{E}\left(\mathfrak{n}^{t\,;u}\right)$.

Beispiel 2.14. In Analogie zu Beispiel 2.12 sollen hier für Algorithmus 3 mit (2.36) ebenfalls die Inklusionswahrscheinlichkeiten des gemeinsamen Designs dargestellt werden. Tabellen 2.6 und 2.7 sind dabei die Entsprechungen der Tabellen 2.3 und 2.4. Die Ergebnisse sind ebenfalls auf die zweite Nachkommastelle gerundet. Es zeigt sich, dass Tabellen 2.3 und 2.6 sich nur hinsichtlich $\mathrm{E}\left(\mathfrak{I}_k^1 \mathfrak{I}_k^{15}\right)$ und $\mathrm{E}\left(\mathfrak{I}_k^2 \mathfrak{I}_k^{15}\right)$ bzw. $\mathrm{E}\left(\mathfrak{I}_k^{15} \mathfrak{I}_k^1\right)$ und $\mathrm{E}\left(\mathfrak{I}_k^{15} \mathfrak{I}_k^2\right)$ unterscheiden. Erst bei einem längeren Beobachtungszeitraum würden sich die Unterschiede in den Längsschnittdesigns der beiden Varianten von Algorithmus 3 stärker abzeichnen.

Tabelle 2.6: $\mathrm{E}\left(\mathfrak{I}_k^t \mathfrak{I}_k^u\right)$ für Algorithmus 3 mit (2.36) und n^t nach Tabelle 2.1

$\pi_k^{t\,;u}$									u							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	1	0,12	0,00	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,04
	2	0,00	0,25	0,00	0,00	0,00	0,06	0,19	0,00	0,00	0,00	0,06	0,14	0,05	0,00	0,02
	3	0,00	0,00	0,19	0,00	0,00	0,00	0,06	0,12	0,00	0,00	0,00	0,05	0,14	0,00	0,00
	4	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,06	0,06	0,00	0,00	0,00	0,06	0,06	0,00
	5	0,00	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,06	0,19	0,00	0,00	0,00	0,19	0,06
	6	0,12	0,06	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,12
	7	0,00	0,19	0,06	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,00	0,19	0,06	0,00	0,00
t	8	0,00	0,00	0,12	0,06	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,00	0,19	0,00	0,00
	9	0,00	0,00	0,00	0,06	0,06	0,00	0,00	0,00	0,12	0,00	0,00	0,00	0,00	0,12	0,00
	10	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,12	0,06
	11	0,12	0,06	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,00	0,25	0,00	0,00	0,00	0,12
	12	0,00	0,14	0,05	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,00
	13	0,00	0,05	0,14	0,06	0,00	0,00	0,06	0,19	0,00	0,00	0,00	0,00	0,25	0,00	0,00
	14	0,00	0,00	0,00	0,06	0,19	0,00	0,00	0,00	0,12	0,12	0,00	0,00	0,00	0,25	0,00
	15	0,04	0,02	0,00	0,00	0,06	0,12	0,00	0,00	0,00	0,06	0,12	0,00	0,00	0,00	0,19

Tabelle 2.7: $\mathrm{E}\left(\mathfrak{I}_k^t \mathfrak{I}_l^u\right)$ $(k \neq l)$ für Algorithmus 3 mit (2.36) und n^t nach Tabelle 2.1

	$\pi_{k,l}^{t,u}$	u														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	1	0,01	0,03	0,03	0,02	0,03	0,03	0,03	0,03	0,02	0,03	0,03	0,03	0,03	0,03	0,02
	2	0,03	0,05	0,05	0,03	0,07	0,06	0,05	0,05	0,03	0,05	0,06	0,04	0,06	0,07	0,05
	3	0,03	0,05	0,03	0,03	0,05	0,05	0,05	0,03	0,03	0,04	0,05	0,03	0,04	0,05	0,04
	4	0,02	0,03	0,03	0,01	0,03	0,03	0,03	0,02	0,01	0,03	0,03	0,03	0,03	0,03	0,03
	5	0,03	0,07	0,05	0,03	0,05	0,07	0,07	0,05	0,03	0,04	0,07	0,05	0,07	0,05	0,05
	6	0,03	0,06	0,05	0,03	0,07	0,05	0,07	0,05	0,03	0,05	0,05	0,05	0,07	0,07	0,04
	7	0,03	0,05	0,05	0,03	0,07	0,07	0,05	0,05	0,03	0,05	0,07	0,04	0,06	0,07	0,05
t	8	0,03	0,05	0,03	0,02	0,05	0,05	0,05	0,03	0,03	0,04	0,05	0,04	0,04	0,05	0,04
	9	0,02	0,03	0,03	0,01	0,03	0,03	0,03	0,03	0,01	0,03	0,03	0,03	0,03	0,03	0,03
	10	0,03	0,05	0,04	0,03	0,04	0,05	0,05	0,04	0,03	0,03	0,05	0,04	0,05	0,04	0,03
	11	0,03	0,06	0,05	0,03	0,07	0,05	0,07	0,05	0,03	0,05	0,05	0,05	0,07	0,07	0,04
	12	0,03	0,04	0,03	0,03	0,05	0,05	0,04	0,04	0,03	0,04	0,05	0,03	0,05	0,05	0,04
	13	0,03	0,06	0,04	0,03	0,07	0,07	0,06	0,04	0,03	0,05	0,07	0,05	0,05	0,07	0,05
	14	0,03	0,07	0,05	0,03	0,05	0,07	0,07	0,05	0,03	0,04	0,07	0,05	0,07	0,05	0,05
	15	0,02	0,05	0,04	0,03	0,05	0,04	0,05	0,04	0,03	0,03	0,04	0,04	0,05	0,05	0,03

△

PRN Koordination von einfachen Zufallsstichproben

Im Folgenden soll Algorithmus 3 mit der Koordination mittels PRN verglichen werden. Zunächst wird ein PRN Verfahren mit zufälligen Stichprobenumfängen im Querschnitt beschrieben. Algorithmus 7 weist jedem Element in der Population eine PRN u_k als unabhängige Ziehung aus einer Gleichverteilung auf dem Intervall $[0,1]$ zu, d.h. $u_k \sim \mathrm{Unif}(0,1)$ für alle $k \in \mathscr{U}$. Zu jedem Zeitpunkt t wird ein Intervall $(a^t, b^t]$, wie in (2.44) beschrieben, bestimmt und alle Elemente, deren PRN innerhalb dieses Intervalls liegen, gelangen in die Stichprobe $\vec{\mathfrak{s}}^{\,t}$.

$$\begin{aligned} (a^t, b^t] &= \begin{cases} (a^{*t}, b^{*t}] & \text{für } a^{*t} < b^{*t} \\ (a^{*t}, 1] \cup (0, b^{*t}] & \text{sonst} \end{cases} \\ &\text{mit } V^{t-1} \bmod 1 = a^{*t}, \qquad V^t \bmod 1 = b^{*t} \\ &\text{und } V^t = \sum_{i=1}^{t} \frac{n^i}{N}, \qquad \text{wobei } V^0 = 0\,. \end{aligned} \tag{2.44}$$

Dabei stellen die n^i für $i = 1, \ldots, t$ in (2.44) die geplanten Erwartungswerte der Stichprobenumfänge der Querschnittdesigns dar.

Für $\pi_k^t = n^t/N \;\forall\, k \in \mathscr{U}$ setzen Algorithmus 6 und 7 das identische Querschnittdesign um. Die Längsschnittdesigns sind jedoch verschieden. Insbesondere ist der Träger des Längsschnittdesigns von Algorithmus 6 größer als der von Algorithmus 7. Dies ist darauf zurückzuführen, dass das Längsschnittdesign von Algorithmus 7 ein geordnetes systematisches Design ist, wie es durch Algorithmus 1 umgesetzt wird. So ist

$$p_k(\vec{s}_k) = \left| \bigcap_{t \mid I_k^t = 1}^{T} (a^t, b^t] \right| . \tag{2.45}$$

Dabei beschreibt (2.45) die Länge der Schnittmengen aller Intervalle, für die gilt, $u_k \in (a^t, b^t]$ (Nedyalkova et al., 2009). Diese entspricht genau einem der Intervalle in (2.13). Damit sind die Wahrscheinlichkeiten $\pi_k^{t,u}$ für Algorithmus 7 gegeben durch

$$\pi_k^{t,u} = \left|(a^t, b^t] \cap (a^u, b^u]\right| \,, \tag{2.46}$$

d.h. $\pi_k^{t,u}$ entspricht der Länge der Schnittmenge der Intervalle in (2.44) zu den Zeitpunkten t und u (Nedyalkova et al., 2009).

Sowohl bei Algorithmus 6 als auch 7 ist $Pr(\mathfrak{n}^t = 0) > 0$. Asymptotisch mag dies zu vernachlässigen sein, wenn die geplanten Stichprobenumfänge n^t , $t \in \mathscr{T}$, und N simultan gegen unendlich streben, die Auswahlsätze n^t/N hingegen konstant bleiben (Stenger, 1989). Für Algorithmus 7 bedeutet dies, dass die empirische Verteilung der PRN sich immer mehr der einer stetigen Gleichverteilung annähert, die Selektionsintervalle $(a^t, b^t]$ in (2.44) aber die gleiche Länge beibehalten. Somit geht die Wahrscheinlichkeit, kein Element auszuwählen, gegen Null. Dennoch ist $Pr(\mathfrak{n}^t = 0) > 0$ bei niedrigen Auswahlsätzen und kleinen N nicht vernachlässigbar.

Algorithmus 7 PRN Koordination einfacher Zufallsstichproben mit zufälligen Stichprobenumfängen

```
$u_k \leftarrow \mathrm{Unif}(0, 1)\ \forall\, k \in \mathscr{U}$ #Festlegung der PRN
$\vec{u} \leftarrow (u_1, \dots, u_k, \dots, u_N)^\top$
$V^0 \leftarrow 0$
for $t = 1, \dots, T$ do
    $I_k^t \leftarrow 0\ \forall\, k \in \mathscr{U}$
    $V^t \leftarrow \sum_{i=1}^{t} \frac{n^i}{N}$ mod 1
    $a \leftarrow V^{t-1}$
    $b \leftarrow V^t$
    if $b > a$ then
        $I_k^t \leftarrow 1\ \forall\, k \in \mathscr{U}$ mit $a < u_k \leq b$
    else
        $I_k^t \leftarrow 1\ \forall\, k \in \mathscr{U}$ mit $a < u_k \leq 1 \wedge 0 < u_k \leq b$
    end if
    $\vec{s}^t \leftarrow (I_1^t, \dots, I_k^t, \dots, I_N^t)^\top$
end for
```

Algorithmus 8 setzt ebenfalls ein Design mit sequentiellen disjunkten Querschnittstichproben um, falls $N \leq n^t + n^{t+1}$ für alle $t \in \mathscr{T}$ gilt. Es wird jedoch die von Ohlsson (1992, 1995) beschriebene PRN Selektion verwendet, um konstante Stichprobenumfänge im Querschnitt zu erhalten.

Die Selektionsintervalle $[a^t, b^t]$ für Algorithmus 8 sind in (2.47) gegeben. Ein Element k wird für Stichprobe $\vec{\mathfrak{s}}^t$ ausgewählt, wenn $O_{(k)} \in [a^t, b^t]$. Dabei ist $O_{(k)}$ die Ordnungsstatistik der PRN des k-ten Elements, die in Zeile 8.3 gebildet wird.

Algorithmus 8 PRN Koordination einfacher Zufallsstichproben mit konstanten Stichprobenumfängen

```
$u_k \leftarrow \mathrm{Unif}(0,1)\ \forall\, k \in \mathscr{U}$ #Festlegung der PRN
$\vec{u} \leftarrow (u_1, \ldots, u_k, \ldots, u_N)^\top$
$\vec{o} \leftarrow (O_{(1)}, \ldots, O_{(k)}, \ldots, O_{(N)})$ #Ordnungsstatistik der PRN
$V^0 \leftarrow 0$
for $t = 1, \ldots, T$ do
    $I_k^t \leftarrow 0\ \forall\, k \in \mathscr{U}$
    $V^t \leftarrow \sum_{i=1}^{t} n^i \bmod N$
    $a \leftarrow V^{t-1} + 1$
    $b \leftarrow V^t$
    if $b > a$ then
        $I_k^t \leftarrow 1\ \forall\, k \in \mathscr{U}$ mit $a \leq O_{(k)} \leq b$
    else
        $I_k^t \leftarrow 1\ \forall\, k \in \mathscr{U}$ mit $a \leq O_{(k)} \leq N \wedge 0 \leq O_{(k)} \leq b$
    end if
    $\vec{s}^{\,t} \leftarrow (I_1^t, \ldots, I_k^t, \ldots, I_N^t)^\top$
end for
```

$$[a^t, b^t] = \begin{cases} [a^{*t}, b^{*t}] & \text{für } a^{*t} < b^{*t} \\ [a^{*t}, N] \cup [0, b^{*t}] & \text{sonst} \end{cases} \tag{2.47}$$

$$\text{mit } V^{t-1} \bmod N + 1 = a^{*t}, \qquad V^t \bmod N = b^{*t}$$

$$\text{und } V^t = \sum_{i=1}^{t} n^i, \qquad \text{wobei } V^0 = 0\,.$$

$\pi_k^{t,u}$ entspricht unter Algorithmus 8 dem von 7, da die beiden Algorithmen das gleiche Längsschnittdesign besitzen. Im Unterschied zu den Algorithmen 6 und 7 ist $\mathfrak{n}^{t,u}$ in Algorithmus 8 keine Zufallsvariable. Mit der Zuweisung der PRN in $t = 1$ ist die gemeinsame Stichprobe, für gegebene n^t, ebenfalls bestimmt, und es gilt für alle $t \in \mathscr{T}$

$$\mathrm{V}\left(\sum_{k \in \mathscr{U}} \mathfrak{I}_k^t \mathfrak{I}_k^u\right) = 0\,.$$

Dies stellt auch den größten Unterschied zu Algorithmus 3 mit (2.36) dar. Zwar haben die Algorithmen 3 mit (2.36) und 8 beide das gleiche Querschnittdesign, jedoch unterscheiden sich, wie schon verdeutlicht, ihre Längsschnittdesigns. Der Träger des Längsschnittdesigns von Algorithmus 8 ist in dem von Algorithmus 3 mit (2.36) enthalten. Das Längsschnittdesign von Algorithmus 3 mit (2.36) ermöglicht Kombinationen von Beobachtungszeitpunkten für ein Element, die unter dem geordneten systematischen Längsschnittdesign von 8 nie möglich sind. Dieser Unterschied ist von besonderer Tragweite bei langen Beobachtungszeiträumen mit hohen Auswahlsätzen. In diesem Fall entstehen schneller Überlappungen zwischen weiter auseinander liegenden Querschnittstichproben, die unter Algorithmus 7 und 8 ausgeschlossen sind.

2.4.3 Koordinierte einfache Zufallsstichproben bei veränderlichen Populationen

Im Falle einer sich verändernden Population ändert sich der Ziehungsrahmen über den Beobachtungszeitraum hinweg. Es können Elemente in die Population eintreten, sog. Geburten, und es können Elemente die Population verlassen, sog. Sterbefälle. Die Menge der Elemente, welche in t in die Population eintreten, wird mit $\mathscr{B}^t$ bezeichnet und die Menge der Elemente, welche die Population in t verlassen, wird mit $\mathscr{D}^t$ bezeichnet. Somit ist der Ziehungsrahmen zum Zeitpunkt t gegeben durch

$$\mathscr{U}^t = \left\{\mathscr{U}^{t-1} \setminus \mathscr{D}^t\right\} \cup \mathscr{B}^t \,, \tag{2.48}$$

bzw. ist

$$\mathscr{B}^t = \mathscr{U}^t \setminus \left\{\mathscr{U}^t \cap \mathscr{U}^{t-1}\right\} \,, \tag{2.49}$$

$$\mathscr{D}^t = \mathscr{U}^{t-1} \setminus \left\{\mathscr{U}^t \cap \mathscr{U}^{t-1}\right\} \,, \tag{2.50}$$

wobei $\mathscr{U}^0 = \emptyset$, $\mathscr{B}^1 = \mathscr{U}^1$ und $\mathscr{D}^1 = \emptyset$.

Algorithmus 9 Koordination mittels Ordnungsstatistik bei veränderlichen Populationen

```
$o_k^0 \leftarrow 0 \;\forall\, k \in \mathscr{U}$ #Initialisierung der Koordinationsvariable
$\vec{o}^{\,0} \leftarrow (o_1^0, \ldots, o_k^0, \ldots, o_N^0)^\top$
for $t \in \mathscr{T}$ do
    $I_k^t \leftarrow 0 \;\forall\, k \in \mathscr{U}^t$
    #Koordinationsvariable für Geburten
    if $(card(\mathscr{B}^t) > 0 \;\&\; t > 1)$ then
        $\mathscr{U}_{l\,(i)}^{t-1} \leftarrow \{k \in \mathscr{U}^{t-1} \setminus \mathscr{D}^t | O_{(k)}^{t-1} = i\}$
        $K_l^{t-1} \leftarrow card(\{o_k^{t-1} | k \in \mathscr{U}^{t-1} \setminus \mathscr{D}^t\})$
        $\vec{\theta} \leftarrow \left(\theta_{(1)}, \ldots, \theta_{(i)}, \ldots, \theta_{(K_l^{t-1})}\right)$ mit $\theta_{(i)} = \frac{card(\mathscr{U}_{l\,(i)}^{t-1})}{N^{t-1} - card(\mathscr{D}^t)}$
        for $k \in \mathscr{B}^t$ do
            $\vec{x}_k \leftarrow multinom\left(m = 1; \vec{p} = \vec{\theta}\right)$
            $o_k^{t-1} \leftarrow \vec{x}_k (o_{(1)}^{t-1}, \ldots, o_{(i)}^{t-1}, \ldots, o_{(K_l^{t-1})}^{t-1})^\top$
        end for
    end if
    $K^{t-1} \leftarrow card(\{\mathfrak{o}_k^{t-1} | k \in \mathscr{U}^t\})$
    $\mathscr{U}_{(i)}^{t-1} \leftarrow \{k \in \mathscr{U}^t | O_{(k)}^{t-1} = i\} \; (i = 1, \ldots, K^{t-1})$
    $f^t(x) \leftarrow \sum_{i=1}^{x} card(\mathscr{U}_{(i)}^{t-1})$
    $g^t(y) \leftarrow \max(\{x | x \in \mathbb{N}, f^t(x) < y\})$
    $\vec{s}^{\,t*} \leftarrow$ SRS aus $\mathscr{U}_{(g^t(n^t)+1)}^{t-1}$ der Größe $n^t - f^t(g^t(n^t))$
    $I_k^t \leftarrow 1 \;\forall\, k \in \bigcup_{i=1}^{g^t(n^t)} \mathscr{U}_{(i)}^{t-1}$
    $\vec{s}^{\,t} \leftarrow (I_1^t, \ldots, I_k^t, \ldots, I_N^t)^\top$
    $\vec{s}^{\,t} \leftarrow \vec{s}^{\,t} + \vec{s}^{\,t*}$
    $\vec{o}^{\,t} \leftarrow h(\vec{s}^{\,t}, \vec{o}^{\,t-1}, \Theta)$#Aktualisierung der Koordinationsvariablen
end for
```

Algorithmus 9 stellt eine Anpassung von Algorithmus 3 auf eine sich verändernde Population dar. Hierzu muss allen Elementen $k \in \mathscr{B}^t$ ein Werte $\mathfrak{o}_k^{t-1}$ zugewiesen werden, bevor die Stichprobe $\vec{\mathfrak{s}}^t$ gezogen werden kann. Dies geschieht wie in Zeile 9.6 bis 9.14 beschrieben. Für jedes Element $k \in \mathscr{B}^t$ wird ein Vektor aus einer Multinomialverteilung mit der Wahrscheinlichkeitsfunktion

$$w(m, \vec{p} = (p_1 \ldots, p_i)) = \begin{cases} \dfrac{m}{m_1! \ldots m_i!} p_1^{m_1} \ldots p_i^{m_i} & \text{für} (m_1, \ldots, m_i) \in \mathbb{N}_0^i \\ & \sum_{j=1}^i m_i = m \text{ und } \sum_{j=1}^i p_i = 1 \\ 0 & \text{sonst} \end{cases},$$

bestimmt. In Algorithmus 9 ist $m = 1$ und $\vec{p}$ der Vektor der Eintrittswahrscheinlichkeiten der Ereignisse $\mathfrak{o}_k^{t-1} = o_{(i)}^{t-1}$ mit $i = 1, \ldots, K_l^{t-1}$. Dabei ist $K_l^{t-1} = card(\mathscr{K}_l^{t-1})$, mit $\mathscr{K}_l^{t-1} = \{o_k^{t-1} | k \in \mathscr{U}^{t-1} \setminus \mathscr{D}^t\}$, die Anzahl möglicher Werte, die $\mathfrak{o}_k^{t-1}$ für jedes Element $k \in \mathscr{B}^t$ annehmen kann. Weiter ist $Pr(o_k^{t-1} = o_{(i)}^{t-1})$ gleich dem Anteil von Elementen, welche den i-ten Rang in der Menge $\mathscr{U}_l^{t-1} = \mathscr{U}^{t-1} \setminus \mathscr{D}^t$ der Überlebenden bis Periode t haben. Demnach ist

$$Pr(\mathfrak{o}_k^{t-1} = o_{(i)}^{t-1}) = \frac{card(\mathscr{U}_{l\,(i)}^{t-1})}{card(\mathscr{U}_l^{t-1})}, \tag{2.51}$$

mit $\mathscr{U}_{l\,(i)}^{t-1} = \{k \in \mathscr{U}^{t-1} \setminus \mathscr{D}^t | O_{(k)}^{t-1} = i\}$. Damit folgt die Verteilung der Koordinationsvariable für die Geburten der Häufigkeitsverteilung der Koordinationsvariable nach der Ziehung in der vorangegangen Periode abzüglich der Sterbefälle. Eine Anpassung von Algorithmus 3 an Sterbefälle geschieht durch die Verwendung von $\mathscr{U}^t$ anstelle von $\mathscr{U}$, d.h. es wird jeweils nur der aktuelle Ziehungsrahmen bei der Entnahme der Elemente berücksichtigt. Für alle $k \in \mathscr{D}^t$ wird somit der zuletzt vergebene Wert der Koordinationsvariable $\mathfrak{o}_k^{t-1}$ bei der Bestimmung der Ordnung der Elemente in $\mathscr{U}_l^t$ für die Stichprobenkoordination in t nicht mehr berücksichtigt. Die Abfolge der Ereignisse in Algorithmus 9 lassen sich wie folgt zusammenfassen:

- Stichprobenziehung in Periode $t-1$
- Abzug der Sterbefälle in Periode t
- Zuteilung der Geburten in Periode t
- Stichprobenziehung in Periode t
- Abzug der Sterbefälle in Periode $t+1$
- usw.

Um die Auswahlregeln von Algorithmus 9 zu veranschaulichen stellt Abbildung A.2 in Appendix A den oben beschrieben Ablauf anhand einer konkreten Stichprobenziehung dar.

Das Vorhandensein von Zu- und Abgängen zur Population über die Zeit bringt es mit sich, dass die Häufigkeitsverteilung der Koordinationsvariablen für $t > 1$, gegeben durch $(\mathfrak{N}_{(i)}^t)_{1,\ldots,K^t}$, mit $\mathfrak{N}_{(i)}^t = card(\mathscr{U}_{(i)}^t)$ und $\mathscr{U}_{(i)}^t$ nach Zeile 9.16, im Gegensatz zu einer sich nicht veränderlichen Population ein Zufallsvektor ist. Dies erschwert die Bestimmung der Inklusionswahrscheinlichkeiten erheblich, unter anderem dadurch, dass sich wesentlich komplexere Abhängigkeiten bei der Stichprobenkoordination ergeben. Algorithmus 9 mit (2.36) erfüllt nicht mehr die Eigenschaften (2.23). So hängt die bedingte Wahrscheinlichkeit, dass ein Element k zum Zeitpunkt u gezogen wird, gegeben $\vec{\mathfrak{s}}^t$, mit $t < u$, unter anderem auch von der Ziehung jener Elemente in t ab, welche die Population nach dem Zeitpunkt t bis einschließlich zum Zeitpunkt u verlassen. Dies hat Auswirkungen auf die Folgen

$$\{O^{t+j}_{(k)}\}_{j=1,\dots,u} \text{ und } \{\mathfrak{N}^{t+j}_{(O^{t+j}_{(k)})}\}_{j=1,\dots,u} \,.$$

Das gilt ebenfalls für die Generierung der Koordinationsvariablen der Geburten in t, was vor der Ziehung einer neuen Stichprobe zum jeweiligen Zeitpunkt geschieht. Die Existenz von Geburten hat so unter anderem zur Folge, dass bei Algorithmus 9 für $t > 1$, π^t_k nicht notwendigerweise gleich n^t/N^t für alle $k \in \mathscr{U}^t$ ist und somit keine einfache Zufallsstichprobe mehr für Zeitpunkt $t > 1$ realisiert wird.

Bei der Bestimmung der Inklusionswahrscheinlichkeiten, die sich durch Algorithmus 9 ergeben, ist es von Vorteil, eine Einteilung von $\mathscr{U}$ nach sog. Kohorten vorzunehmen. Eine Kohorte bezeichnet dabei eine Menge von Elementen, welche zum gleichen Zeitpunkt α in die Population eintreten und sie zum gleichen Zeitpunkt ω wieder verlassen, bzw. bis zum Ende des Beobachtungszeitraums in der Population verweilen. Im Folgenden wird mit $\mathscr{C}^{\alpha\,\omega}$ eine Kohorte bezeichnet, welche zum Zeitpunkt α in die Population eintritt und diese zum Zeitpunkt ω verlässt. Für $\omega = T+1$ wird der Fall beschrieben dass die Elemente einer Kohorte bis zum Ende des Beobachtungszeitraums in der Population verweilen. Somit ist

$$\begin{aligned} \mathscr{C}^{\alpha\,\omega} &= \{k | k \in \mathscr{B}^{\alpha} \cap \mathscr{D}^{\omega}\}\,, &&\text{für } 2 \leq \omega \leq T\,, \\ \mathscr{C}^{\alpha\,T+1} &= \left\{k | k \in \mathscr{B}^{\alpha} \cap \mathscr{U}^{T}\right\}, &&\text{für } \omega = T+1\,. \end{aligned} \tag{2.52}$$

Die Anzahl von Elementen, welche zu einer Kohorte gehören, ist über den gesamten Beobachtungszeitraums konstant. Die maximale Anzahl von Kohorten beträgt $\frac{T(T+1)}{2}$.

Die Inklusionswahrscheinlichkeiten π^t_k, $\pi^{t,u}_k$, $\pi^t_{k,l}$ und $\pi^{t,u}_{k,l}$ von Algorithmus 9 sind bestimmt durch die gemeinsame Wahrscheinlichkeitsverteilung der zweidimensionalen Zufallsvariablen

$$\left(\mathfrak{N}^{\alpha-1}_{c^{\alpha\,\omega}(i)}, \mathfrak{n}^{t}_{c^{\alpha\,\omega}(i)}\right)_{\substack{t=1,\dots,T \\ \alpha<\omega,\ \alpha=1,\dots,\max(t,u),\ \omega=2,\dots,\max(t,u)+1 \\ i=1,\dots,K^{t-1}_l}}\,, \tag{2.53}$$

wobei

$$\mathfrak{N}^{t}_{c^{\alpha\,\omega}(i)} = card(\mathscr{U}^t_{(i)} \cap \mathscr{C}^{\alpha\,\omega})\,, \tag{2.54}$$

$$\mathfrak{n}^{t}_{c^{\alpha\,\omega}(i)} = card(\mathscr{s}^t \cap \mathscr{U}^{t-1}_{(i)} \cap \mathscr{C}^{\alpha\,\omega})\,. \tag{2.55}$$

Die Zufallsvariable $\mathfrak{N}^{\alpha-1}_{c^{\alpha\,\omega}(i)}$ gibt somit an, wie viele Elemente in Kohorte $\mathscr{C}^{\alpha\,\omega}$ zu ihrer Geburt den i-t höchsten Wert der Koordinationsvariable des Zeitpunktes $\alpha - 1$ zugewiesen bekommen. Entsprechend ist $\mathfrak{n}^{t}_{c^{\alpha\,\omega}(i)}$ die Anzahl der Elemente aus Kohorte $\mathscr{C}^{\alpha\,\omega}$, die aus der Gruppe der Elemente mit dem i-ten Rang vor der Stichprobenziehung in t in Stichprobe $\vec{\mathfrak{s}}^t$ gelangen. Es ist anzumerken, dass mit $(\mathfrak{N}^{\alpha-1}_{c^{\alpha\,\omega}(i)})_{k=1,\dots,K^{\alpha-1}_l}$

gegeben, d.h. der Häufigkeitsverteilung der Koordinationsvariablen aller Elemente in Kohorte $\mathscr{C}^{\alpha\,\omega}$ zu deren Geburt, sowie $\mathfrak{n}^t_{c\alpha\omega}$ gegeben, mit

$$\mathfrak{n}^t_{c\alpha\omega} = \sum_i^{K_l^{t-1}} \mathfrak{n}^t_{c\alpha\omega(i)} \ ,$$

$(\mathfrak{N}^{\alpha}_{c\alpha\omega\,(i)})_{k=1,\ldots,K^\alpha}$ eindeutig bestimmt ist. Da sich $\mathfrak{N}^t_{c\alpha\omega\,(i)}$ als Funktion von $\mathfrak{N}^{t-1}_{c\alpha\omega\,(i)}$ und $\mathfrak{n}^t_{c\alpha\omega}$ darstellen lässt.

Zur Vereinfachung der Schreibweise werden folgende Zusammenfassungen gemacht:

$$\begin{aligned}
\vec{\mathfrak{N}}^t_{c\alpha\omega} &= \left(\mathfrak{N}^t_{c\alpha\omega\,(i)}\right)_{k=1,\ldots,K^t} \ , \\
\vec{\mathfrak{n}}^t_{c\alpha\omega} &= \left(\mathfrak{n}^t_{c\alpha\omega\,(i)}\right)_{k=1,\ldots,K_l^{t-1}} \ , \\
\vec{\mathfrak{N}}^t_{c} &= \left(\mathfrak{N}^t_{c\alpha\omega}\right)_{\alpha<\omega,\ \alpha=1,\ldots,t,\ \omega=2,\ldots,t+1} \ , \\
\vec{\mathfrak{n}}^t_{c} &= \left(\mathfrak{n}^t_{c\alpha\omega}\right)_{\alpha<\omega,\ \alpha=1,\ldots,t,\ \omega=2,\ldots,t+1} \ , \\
\vec{\mathfrak{N}}_{c\alpha\omega} &= (\mathfrak{N}^{t-1}_{c\alpha\omega})_{t=1,\ldots,T} \ , \\
\vec{\mathfrak{n}}_{c\alpha\omega} &= (\mathfrak{n}^{t}_{c\alpha\omega})_{t=1,\ldots,T} \ .
\end{aligned}$$

Die Inklusionswahrscheinlichkeiten der Querschnittstichprobe sind für Algorithmus 9 gegeben durch,

$$\pi^t_k = \frac{\mathrm{E}\left(\mathfrak{n}^t_{c\alpha\omega}\right)}{N_{c\alpha\omega}} \qquad \forall\, k \in \mathscr{C}^{\alpha\,\omega} \ , \tag{2.56}$$

wobei $N_{c\alpha\omega} = card(\mathscr{C}^{\alpha\,\omega})$ und $\mathfrak{n}^t_{c\alpha\omega} = \sum_i^{K_l^{t-1}} \mathfrak{n}^t_{c\alpha\omega(i)}$. Die Inklusionswahrscheinlichkeiten zweiter Ordnung sind gegeben durch

$$\pi^t_{k,l} = \begin{cases} \dfrac{\mathrm{E}\left(\mathfrak{n}^t_{c\alpha\omega}\left(\mathfrak{n}^t_{c\alpha\omega}-1\right)\right)}{N_{c\alpha\omega}(N_{c\alpha\omega}-1)} & \forall\, k,l \in \mathscr{C}^{\alpha\,\omega}, \text{für } N_{c\alpha\omega} > 1 \\ 0 & \text{sonst} \end{cases} \ . \tag{2.57}$$

Beispiel 2.15. Um die Inklusionswahrscheinlichkeiten des Querschnittdesigns von Algorithmus 9 zu illustrieren, soll davon ausgegangen werden, dass für $T = 2$ die Ziehungsrahmen $\mathscr{U}^1 = \{1,2,3,4,5\}$ und $\mathscr{U}^2 = \{2,3,4,5,7,8\}$ bestehen. Damit ergeben sich aus den beiden Ziehungsrahmen die folgenden drei Kohorten $\mathscr{C}^{1\,2} = \{1\}$, $\mathscr{C}^{1\,3} = \{2,3,4,5\}$ sowie $\mathscr{C}^{2\,3} = \{7,8\}$. Die Stichprobenumfänge sind gegeben durch $n^1 = n^2 = 2$.

Vor der Ziehung von Stichprobe $\vec{\mathfrak{s}}^{\,1}$ gibt es nur eine mögliche Häufigkeitsverteilung der Koordinationsvariablen der Elemente in $\mathscr{C}^{1\,2}$ und $\mathscr{C}^{1\,3}$, nämlich

$$\begin{aligned}
Pr\left(\vec{\mathfrak{N}}^0_{c^{12}} = \left(\mathfrak{N}^0_{c^{12}\,(1)} = 1\right)\right) &= 1 \ , \\
Pr\left(\vec{\mathfrak{N}}^0_{c^{13}} = \left(\mathfrak{N}^0_{c^{13}\,(1)} = 4\right)\right) &= 1 \ .
\end{aligned}$$

Mit Hinblick auf die Koordination der Stichproben $\vec{\mathfrak{s}}^{\,1}$ und $\vec{\mathfrak{s}}^{\,2}$ sind die folgenden Ereignisse bei der Ziehung von $\vec{\mathfrak{s}}^{\,1}$ von Interesse: Es wird jeweils ein Element aus $\mathscr{C}^{1\,2}$ und

$\mathscr{C}^{1\,3}$ entnommen oder es werden zwei Elemente aus $\mathscr{C}^{1\,3}$ gezogen. Damit ergeben sich nach der Ziehung von $\vec{\mathfrak{s}}^{\,1}$ zwei mögliche Ausprägungen von $\vec{\mathfrak{n}}^{t}_{c1\,3}$ mit den Wahrscheinlichkeiten,

$$
\begin{aligned}
Pr\left(\vec{\mathfrak{n}}^{1}_{c1\,3} = \left(\mathfrak{n}^{1}_{c1\,3\,(1)} = 1,\ \mathfrak{n}^{1}_{c1\,3\,(2)} = 0\right)\right) &= 0{,}4\ ,\\
Pr\left(\vec{\mathfrak{n}}^{1}_{c1\,3} = \left(\mathfrak{n}^{1}_{c1\,3\,(1)} = 2,\ \mathfrak{n}^{1}_{c1\,3\,(2)} = 0\right)\right) &= 0{,}6\ .
\end{aligned}
$$

Folglich bestehen für die Kohorte $\mathscr{C}^{2\,3}$ drei mögliche Ausprägungen von $\vec{\mathfrak{N}}^{1}_{c2\,3}$ mit den folgenden Wahrscheinlichkeiten,

$$
\begin{aligned}
Pr\left(\vec{\mathfrak{N}}^{1}_{c2\,3} = \left(\mathfrak{N}^{1}_{c2\,3\,(1)} = 2,\ \mathfrak{N}^{1}_{c2\,3\,(2)} = 0\right)\right) &= 0,375\ ,\\
Pr\left(\vec{\mathfrak{N}}^{1}_{c2\,3} = \left(\mathfrak{N}^{1}_{c2\,3\,(1)} = 1,\ \mathfrak{N}^{1}_{c2\,3\,(2)} = 1\right)\right) &= 0,450\ ,\\
Pr\left(\vec{\mathfrak{N}}^{1}_{c2\,3} = \left(\mathfrak{N}^{1}_{c2\,3\,(1)} = 0,\ \mathfrak{N}^{1}_{c2\,3\,(2)} = 2\right)\right) &= 0,175\ .
\end{aligned}
$$

Tabelle 2.8 stellt alle Ausprägungen von (2.53) sowie deren jeweilige Eintrittswahrscheinlichkeiten dar. So existieren, wie zuvor beschrieben, zwei mögliche Ausprägungen von $(\vec{\mathfrak{N}}^{0}_{c}, \vec{\mathfrak{n}}^{1}_{c})$, aber bereits zwölf mögliche Ausprägungen von $(\vec{\mathfrak{N}}^{1}_{c}, \vec{\mathfrak{n}}^{2}_{c})$.

Tabelle 2.8: Beispiel 1 zu Algorithmus 9

		$t=1$									
	Pr	$\mathfrak{p}_k^{t-1}$	$\mathfrak{b}_k^{t-1}$	$\mathfrak{o}_k^{t-1}$	$\mathfrak{N}^{t-1}_{c1\,2\,(i)}$	$\mathfrak{n}^{t}_{c1\,2\,(i)}$	$\mathfrak{p}_k^{t-1}$	$\mathfrak{b}_k^{t-1}$	$\mathfrak{o}_k^{t-1}$	$\mathfrak{N}^{t-1}_{c1\,3\,(i)}$	$\mathfrak{n}^{t}_{c1\,3\,(i)}$
1	0,4	0	0	0	1	1	0	0	0	4	1
2	0,6	0	0	0	1	0	0	0	0	4	2
		$t=2$									
		$\mathfrak{p}_k^{t-1}$	$\mathfrak{b}_k^{t-1}$	$\mathfrak{o}_k^{t-1}$	$\mathfrak{N}^{t-1}_{c1\,3\,(i)}$	$\mathfrak{n}^{t}_{c1\,3\,(i)}$	$\mathfrak{p}_k^{t-1}$	$\mathfrak{b}_k^{t-1}$	$\mathfrak{o}_k^{t-1}$	$\mathfrak{N}^{t-1}_{c2\,3\,(i)}$	$\mathfrak{n}^{t}_{c2\,3\,(i)}$
1.1	0,0675	1	0	2	3	2	1	0	2	2	0
		0	1	-1	1	0	0	1	-1	0	0
1.2	0,1350	1	0	2	3	1	1	0	2	2	1
		0	1	-1	1	0	0	1	-1	0	0
1.3	0,0225	1	0	2	3	0	1	0	2	2	2
		0	1	-1	1	0	0	1	-1	0	0
1.4	0,0750	1	0	2	3	2	1	0	2	1	0
		0	1	-1	1	0	0	1	-1	1	0
1.5	0,0750	1	0	2	3	1	1	0	2	1	1
		0	1	-1	1	0	0	1	-1	1	0
1.6	0,0250	1	0	2	3	2	1	0	2	0	0
		0	1	-1	1	0	0	1	-1	2	0
2.7	0,0250	1	0	2	2	2	1	0	2	2	0
		0	1	-1	2	0	0	1	-1	0	0
2.8	0,1000	1	0	2	2	1	1	0	2	2	1
		0	1	-1	2	0	0	1	-1	0	0
2.9	0,0250	1	0	2	2	0	1	0	2	2	2
		0	1	-1	2	0	0	1	-1	0	0
2.10	0,1000	1	0	2	2	2	1	0	2	1	0
		0	1	-1	2	0	0	1	-1	1	0
2.11	0,2000	1	0	2	2	1	1	0	2	1	1
		0	1	-1	2	0	0	1	-1	1	0
2.12	0,1500	1	0	2	2	2	1	0	2	0	0
		0	1	-1	2	0	0	1	-1	2	0

Unter Verwendung von Tabelle 2.8 lassen sich die Inklusionswahrscheinlichkeiten wie folgt bestimmen:

$$
\begin{aligned}
\pi_k^1 &= \frac{0,4}{1} = 0,4 && \forall\, k \in \mathscr{C}^{1\,2}, & \pi_k^1 &= \frac{1,6}{4} = 0,4 && \forall\, k \in \mathscr{C}^{1\,3}, \\
\pi_k^2 &= \frac{1,395}{4} = 0,34875 && \forall\, k \in \mathscr{C}^{1\,3}, & \pi_k^2 &= \frac{0,605}{2} = 0,3025 && \forall\, k \in \mathscr{C}^{2\,3}, \\
\pi_{k,l}^1 &= \frac{1,2}{12} = 0,1 && \forall\, \{k,l\} \in \mathscr{C}^{1\,3}, & \pi_{k,l}^1 &= \frac{0,4}{4} = 0,1 && \forall\, k \in \mathscr{C}^{1\,3}\; l \in \mathscr{C}^{1\,2}, \\
\pi_{k,l}^2 &= \frac{0,885}{12} = 0,07375 && \forall\, \{k,l\} \in \mathscr{C}^{1\,3}\; k \neq l, & \pi_{k,l}^2 &= \frac{0,095}{2} = 0,0475 && \forall\, \{k,l\} \in \mathscr{C}^{2\,3}\; k \neq l, \\
\pi_{k,l}^2 &= \frac{0,51}{8} = 0,06375 && \forall\, k \in \mathscr{C}^{1\,3}\; l \in \mathscr{C}^{2\,3}.
\end{aligned}
$$

△

Unter Algorithmus 9 ist zudem

$$
\begin{aligned}
\pi_k^t &= \pi_{c^\alpha}^t && \forall\, k \in \bigcup_{\omega=2}^{T+1} \mathscr{C}^{\alpha\,\omega} \\
\pi_{k,l}^t &= \pi_{c^\alpha,\,c^{\bar\alpha}}^t && \forall\, k \neq l,\, k \in \bigcup_{\omega=2}^{T+1} \mathscr{C}^{\alpha\,\omega},\, l \in \bigcup_{\bar\omega=2}^{T+1} \mathscr{C}^{\bar\alpha\,\bar\omega}\,.
\end{aligned}
$$

Das heißt, alle Elemente, die zum gleichen Zeitpunkt α geborenen wurden, haben die gleiche Inklusionswahrscheinlichkeit $\pi_{c^\alpha}^t\ \forall\, t \in \mathscr{T}$. Alle Elemente k, welche zum Zeitpunkt α geboren wurden, sowie alle Elemente k, welche zum Zeitpunkt $\bar\alpha$ geboren wurden, haben die gleiche Inklusionswahrscheinlichkeit zweiter Ordnung im Querschnitt $\pi_{c^\alpha,\,c^{\bar\alpha}}^t\ \forall\, t \in \mathscr{T}$, was jedoch nicht bedeutet, dass $\pi_{c^\alpha,\,c^{\bar\alpha}}^t = \pi_{c^\alpha,\,c^\alpha}^t$ für $\alpha \neq \bar\alpha$ gilt.

Trotz der ungleichen Inklusionswahrscheinlichkeiten für Elemente, welche zu unterschiedlichen Zeitpunkten geboren werden, ist das arithmetische Mittel von π_k^t und $\pi_{k,l}^t$ über alle $k,l \in \mathscr{U}^t$ gegeben durch

$$
\frac{1}{N^t} \sum_{k \in \mathscr{U}^t} \pi_k^t = \frac{n^t}{N^t}\,\text{, bzw.}
$$

$$
\frac{1}{N^t(N^t-1)} \sum_{k \in \mathscr{U}^t} \sum_{\substack{l \in \mathscr{U}^t \\ l \neq k}} \pi_{k,l}^t = \frac{n^t(n^t-1)}{N^t(N^t-1)}\,.
$$

Dies gilt, da unter Algorithmus 9 π_k^t und $\pi_{k,l}^t$ folgende Eigenschaften erfüllen:

$$
\begin{aligned}
\sum_{k \in \mathscr{U}^t} \pi_k^t &= \sum_{\alpha\omega} \sum_{k \in \mathscr{C}^{\alpha\,\omega}} \frac{E\left(\mathfrak{n}_{c^{\alpha\,\omega}}^t\right)}{N_{c^{\alpha\,\omega}}} \\
&= \sum_{\alpha\omega} E\left(\mathfrak{n}_{c^{\alpha\,\omega}}^t\right) \\
&= \sum_{\alpha\omega} \sum_{v} Pr\left(\mathfrak{n}_{c^{\alpha\,\omega}}^t = n_{c^{\alpha\,\omega}_v}^t\right) n_{c^{\alpha\,\omega}_v}^t = \sum_{v} Pr\left(\mathfrak{n}_{c^{\alpha\,\omega}}^t = n_{c^{\alpha\,\omega}_v}^t\right) \sum_{\alpha\omega} n_{c^{\alpha\,\omega}_v}^t \\
&= \sum_{v} Pr\left(\mathfrak{n}_{c^{\alpha\,\omega}}^t = n_{c^{\alpha\,\omega}_v}^t\right) n^t \\
&= n^t\,,
\end{aligned}
$$

wobei $n^t_{c\alpha\omega_\nu}$ die ν-te mögliche Ausprägung von $\mathfrak{n}^t_{c\alpha\omega}$ ist. Des Weiteren ist

$$\begin{aligned}
\sum_{k\in\mathscr{U}^t}\sum_{\substack{l\in\mathscr{U}^t\\k\neq l}}\pi^t_{k,l} &= \sum_{\alpha\omega}\sum_{k\in\mathscr{C}^{\alpha\omega}}\sum_{\substack{l\in\mathscr{C}^{\alpha\omega}\\k\neq l}}\frac{E(\mathfrak{n}^t_{c\alpha\omega}(\mathfrak{n}^t_{c\alpha\omega}-1))}{N_{c\alpha\omega}(N_{c\alpha\omega}-1)}+\sum_{\alpha\omega}\sum_{\substack{\bar\alpha\bar\omega\\\bar\alpha\bar\omega\neq\alpha\omega}}\sum_{k\in\mathscr{C}^{\alpha\omega}}\sum_{l\in\mathscr{C}^{\bar\alpha\bar\omega}}\frac{E(\mathfrak{n}^t_{c\alpha\omega}\mathfrak{n}^t_{c\bar\alpha\bar\omega})}{N_{c\alpha\omega}N_{c\bar\alpha\bar\omega}}\\
&= \sum_{\alpha\omega}\sum_{\bar\alpha\bar\omega}E(\mathfrak{n}^t_{c\alpha\omega}\mathfrak{n}^t_{c\bar\alpha\bar\omega})+\sum_{\alpha\omega}E(\mathfrak{n}^t_{c\alpha\omega})\\
&= \sum_{\alpha\omega}\sum_{\bar\alpha\bar\omega}\sum_{\nu}Pr\left(\mathfrak{n}^t_{c\alpha\omega}\mathfrak{n}^t_{c\bar\alpha\bar\omega}=n^t_{c\alpha\omega_\nu}n^t_{c\bar\alpha\bar\omega_\nu}\right)n^t_{c\alpha\omega_\nu}n^t_{c\bar\alpha\bar\omega_\nu}+n^t\\
&= \sum_{\nu}Pr\left(\mathfrak{n}^t_{c\alpha\omega}\mathfrak{n}^t_{c\bar\alpha\bar\omega}=n^t_{c\alpha\omega_\nu}n^t_{c\bar\alpha\bar\omega_\nu}\right)\sum_{\alpha\omega}n^t_{c\alpha\omega_\nu}\sum_{\bar\alpha\bar\omega}n^t_{c\bar\alpha\bar\omega_\nu}+n^t\\
&= \sum_{\nu}Pr\left(\mathfrak{n}^t_{c\alpha\omega}\mathfrak{n}^t_{c\bar\alpha\bar\omega}=n^t_{c\alpha\omega_\nu}n^t_{c\bar\alpha\bar\omega_\nu}\right){n^t}^2+n^t\\
&= n^t(n^t-1)\,.
\end{aligned}$$

Die Darstellung von $\pi^{t,u}_k$ ist weit komplexer als im Falle einer invarianten Population. Für $k\in\mathscr{C}^{\alpha\omega}$ und $(\vec{\mathfrak{N}}_{c\alpha\omega},\vec{\mathfrak{n}}_{c\alpha\omega})$ gegeben, kann die Wahrscheinlichkeitsverteilung von $\phi^t_k|O^t_{(k)}$ analog zu (2.42) bzw. (2.41) bestimmt werden, wenn N, N^v, N^w und n^t durch ihre Kohorten spezifischen Analogons $N_{c\alpha\omega}$, $\mathfrak{N}^v_{c\alpha\omega}$ $\mathfrak{N}^w_{c\alpha\omega}$, und $\mathfrak{n}^t_{c\alpha\omega}$, ersetzt werden, sowie K^t durch K^{t-1}_l. Dabei sind $N^v_{c\alpha\omega}$ und $N^w_{c\alpha\omega}$ gegeben durch (2.39) und (2.40), mit dem Unterschied, dass $\mathfrak{N}^t_{c\alpha\omega\,(\nu)}$ anstatt $N^t_{(\nu)}$ zu verwenden ist. Zudem ist anzumerken, dass $O^t_{(k)}$ sowohl von $\vec{\mathfrak{s}}^{\,1},\ldots,\vec{\mathfrak{s}}^{\,t}$ sowie auch von der anfänglichen Verteilung der Koordinationsvariablen der Geburten bis t abhängt.

Als Alternative lässt sich $\pi^{t,u}_k$ für $k\in\mathscr{C}^{\alpha\omega}$, gegeben

$$\vec{\Xi}^t_{\alpha\omega}=\left(\vec{\mathfrak{N}}^{\alpha-1}_{c\alpha\omega},(\vec{\mathfrak{n}}^{\nu}_{c\alpha\omega})_{\nu=\alpha,\ldots,t}\right)\,, \tag{2.58}$$

schreiben als die Summe der Eintrittswahrscheinlichkeiten aller möglichen Ausprägungen von $(\mathfrak{o}^{\nu-1}_k)_{\nu=\alpha,\ldots,\max(t,u)}$, gegeben $\vec{\Xi}^{\max(t,u)}_{\alpha\omega}$, welche die Bedingungen $o^{t-1}_k>o^t_k$ und $o^{u-1}_k>o^u_k$ erfüllen, denn aus $(\mathfrak{o}^{\nu-1}_k)_{\nu=\alpha,\ldots,\max(t,u)}$ kann auch eindeutig $(\mathfrak{I}^{\nu-1}_k)_{\nu=\alpha,\ldots,\max(t,u)}$ bestimmt werden, da für $\mathfrak{o}^{t-1}_k>\mathfrak{o}^t_k$, $\mathfrak{I}^t_k=1$ und $\mathfrak{I}^t_k=0$ sonst, für alle $t\in\mathscr{T}$. So ist

$$\begin{aligned}
&\mathrm{E}\left(\mathfrak{I}^t_k\mathfrak{I}^u_k|\vec{\Xi}^{\max(t,u)}_{\alpha\omega}\right)=\\
&\qquad\sum_{\vec{o}\in\mathscr{S}^{\max(t,u)}_{\alpha\omega}}Pr\left((\mathfrak{o}^{\nu-1}_k)_{\nu=\alpha,\ldots,\max(t,u)}=\vec{o}|\vec{\Xi}^{\max(t,u)}_{\alpha\omega}\right)\,.
\end{aligned} \tag{2.59}$$

Dabei ist

$$\begin{aligned}
&Pr\left((\mathfrak{o}^{\nu-1}_k)_{\nu=\alpha,\ldots,\max(t,u)}=\vec{o}|\vec{\Xi}^{\max(t,u)}_{\alpha\omega}\right)=\\
&\qquad Pr(\mathfrak{o}^{\alpha-1}_k=o^{\alpha-1}_k|\vec{\mathfrak{N}}^{\alpha-1}_{c\alpha\omega})\prod_{\nu=\alpha}^{\max(t,u)}\pi^{\nu}_{k(\iota^{\nu}_{c\alpha\omega})}\,,
\end{aligned}$$

mit

$$\pi^{\nu}_{k(\iota^{\nu}_{c\alpha\omega})}=\binom{\mathfrak{N}^{\nu-1}_{c\alpha\omega\,(\iota^{\nu}_{c\alpha\omega})}}{\mathfrak{n}^{\nu}_{c\alpha\omega\,(\iota^{\nu}_{c\alpha\omega})}}^{-1}\binom{\mathfrak{N}^{\nu-1}_{c\alpha\omega\,(\iota^{\nu}_{c\alpha\omega})}-\mathbb{1}^{\nu}_{\mathscr{U}_{(\iota^{\nu}_{c\alpha\omega})}}(k)}{\mathfrak{n}^{\nu}_{c\alpha\omega\,(\iota^{\nu}_{c\alpha\omega})}-\mathbb{1}^{\nu}_{\mathfrak{d},\mathscr{U}_{(\iota^{\nu}_{c\alpha\omega})}}(k)}\,,$$

$$\mathbb{1}^t_{\mathscr{U}_{(i)}}(k) = \begin{cases} 1 & \text{für} \quad k \in \mathscr{U}^{t-1}_{(i)} \\ 0 & \text{sonst} \end{cases} ,$$

$$\mathbb{1}^t_{\mathscr{s},\mathscr{U}_{(i)}}(k) = \begin{cases} 1 & \text{für} \quad k \in \{\mathscr{s}^t \cap \mathscr{U}^{t-1}_{(i)}\} \\ 0 & \text{sonst} \end{cases}$$

sowie $\mathscr{U}^{t-1}_{(i)} = \{k \in \mathscr{U}^t | O^{t-1}_{(k)} = i\}$ wie in Zeile 9.16 und

$$\imath^t_{c\alpha\omega} = \begin{cases} \max\left(\mathscr{Z}^t_{c\alpha\omega}\right) + 1 & \text{für} \quad \mathscr{Z}^t_{c\alpha\omega} \neq \emptyset \\ 1 & \text{sonst} \end{cases} ,$$

wobei $\mathscr{Z}^t_{c\alpha\omega} = \left\{ j \in \{1, \ldots, K^{t-1}_l\} | \sum_{i=1}^{j} \mathfrak{N}^{t-1}_{c\alpha\omega\,(i)} < \mathfrak{n}^t_{c\alpha\omega} \right\}$.

Zudem ist

$$\mathscr{S}^{\max(t,u)}_{\alpha\omega} = \left\{ \vec{o} | \vec{o} \in \mathscr{O}^{\max(t,u)}_{\alpha\omega},\ (\vec{\mathfrak{N}}^{\nu-1}_{c\alpha\omega}, \vec{\mathfrak{n}}^{\nu}_{c\alpha\omega})_{\nu=\alpha,\ldots,\max(t,u)},\ o^{t-1}_k > o^t_k,\ o^{u-1}_k > o^u_k \right\} ,$$

wobei $\mathscr{O}^{\max(t,u)}_{\alpha\omega}$ die Menge aller möglichen Ausprägungen von $(o^{\nu-1}_k)_{\nu=\alpha,\ldots,\max(t,u)}$ ist. $Pr(\mathfrak{o}^{\alpha-1}_k = o^{\alpha-1} | \vec{\mathfrak{N}}^{\alpha-1}_{c\alpha\omega})$ ist die Wahrscheinlichkeit, dass Element k zu seiner Geburt den Wert $o^{\alpha-1}_k$ zugewiesen bekommt, gegeben der Häufigkeitsverteilung der Koordinationsvariablen von Kohorte $\mathscr{C}^{\alpha,\omega}$ zu deren Geburt. Für $O^{\alpha-1}_{(k)} = i$ ist

$$Pr(\mathfrak{o}^{\alpha-1}_k = o^{\alpha-1}_{(i)} | \vec{\mathfrak{N}}^{\alpha-1}_{c\alpha\omega}) = \frac{\mathfrak{N}^{\alpha-1}_{c\alpha\omega\,(i)}}{N_{c\alpha\omega}} .$$

Die Bestimmung von $\pi^{t,u}_{k,l}$ für $k \in \mathscr{C}^{\alpha\omega}$ und $l \in \mathscr{C}^{\bar{\alpha}\bar{\omega}}$ gegeben $\vec{\Xi}^{\max(t,u)}_{\alpha\omega}$ und $\vec{\Xi}^{\max(t,u)}_{\bar{\alpha}\bar{\omega}}$ ist analog zu Algorithmus 3 im Falle einer sich nicht verändernden Population, d.h.

$$\begin{aligned} &\mathrm{E}\left(\mathfrak{I}^t_k \mathfrak{I}^u_l | \vec{\Xi}^{\max(t,u)}_{\alpha\omega}, \vec{\Xi}^{\max(t,u)}_{\bar{\alpha}\bar{\omega}}\right) = \\ &\begin{cases} \dfrac{\mathfrak{n}^t_{c\alpha\omega}\mathfrak{n}^u_{c\alpha\omega} - N_{c\alpha\omega} E\left(\mathfrak{I}^t_k \mathfrak{I}^u_k | \vec{\Xi}^{\max(t,u)}_{\alpha\omega}\right)}{N_{c\alpha\omega}(N_{c\alpha\omega} - 1)} & \text{für} \quad (\alpha,\omega) = (\bar{\alpha},\bar{\omega}) \\ & \qquad\quad \text{und } N_{c\alpha\omega} > 1 , \\ \dfrac{\mathfrak{n}^t_{c\alpha\omega}}{N_{c\alpha\omega}} \dfrac{\mathfrak{n}^u_{c\bar{\alpha}\bar{\omega}}}{N_{c\bar{\alpha}\bar{\omega}}} & \text{für} \quad (\alpha,\omega) \neq (\bar{\alpha},\bar{\omega}) \\ & \qquad\quad N_{c\alpha\omega} > 0 \text{ und } N_{c\bar{\alpha}\bar{\omega}} > 0 , \\ 0 & \text{sonst} . \end{cases} \end{aligned} \tag{2.60}$$

Beispiel 2.16. Zur Bestimmung von $\mathrm{E}\left(\mathfrak{I}^t_k \mathfrak{I}^u_k | \vec{\Xi}^{\max(t,u)}_{\alpha\omega}\right)$ sei folgendes Beispiel gegeben. Der Beobachtungszeitraum von 2.15 wird um zwei Perioden erweitert, mit den Ziehungsrahmen $\mathscr{U}^3 = \{3,4,5,7,8\}$ und $\mathscr{U}^4 = \{3,4,5,7,8,9,10\}$, d.h. es existieren die folgenden Kohorten, $\mathscr{C}^{12} = \{1\}$, $\mathscr{C}^{13} = \{2\}$, $\mathscr{C}^{23} = \{6\}$, $\mathscr{C}^{15} = \{3,4,5\}$, $\mathscr{C}^{25} = \{7\}$, $\mathscr{C}^{35} = \{8\}$ und $\mathscr{C}^{45} = \{9,10\}$. Die Stichprobenumfänge sind gegeben durch $n^1 = n^2 = n^3 = 2$ und $n^4 = 3$. Tabelle 2.9 stellt eine möglich Ausprägung von $(\vec{\mathfrak{N}}^{t-1}_c, \vec{\mathfrak{n}}^t_c)_{t=1,2,3,4}$ dar.

△

Tabelle 2.9: Beispiel 2 zu Algorithmus 9

α	ω	$\mathfrak{p}_k^{t-1}$	$\mathfrak{b}_k^{t-1}$	$\mathfrak{o}_k^{t-1}$	$\mathfrak{N}_{c\alpha\omega\,(i)}^{t-1}$	$\mathfrak{n}_{c\alpha\omega\,(i)}^{t}$
		$t=1$				
1	2	0	0	0	1	1
1	3	0	0	0	1	1
1	5	0	0	0	3	0
		$t=2$				
1	3	1	0	2	0	0
		0	1	-1	1	0
2	3	1	0	2	1	1
		0	1	-1	0	0
1	5	1	0	2	3	0
		0	1	-1	0	0
2	5	1	0	2	1	1
		0	1	-1	0	0
		$t=3$				
1	5	2	0	4	3	2
		0	1	-1	0	0
2	5	2	0	4	0	0
		0	1	-1	1	0
3	5	2	0	4	1	0
		0	1	-1	0	0
		$t=4$				
1	5	3	0	6	1	1
		1	1	1	0	0
		0	1	-1	2	0
2	5	3	0	6	0	0
		1	1	1	1	1
		0	1	-1	0	0
3	5	3	0	6	1	1
		1	1	1	0	0
		0	1	-1	0	0
4	5	3	0	6	0	0
		1	1	1	2	0
		0	1	-1	0	0

2.5 Rotationspanel

Algorithmus 3 mit (2.14) oder (2.36) setzen eine negative Koordination der Elemente zwischen verschiedenen Zeitpunkten um. Dies hat zur Folge, dass ψ_k^t, die Verweildauer nach der Ziehung von Element k zum Zeitpunkt t, minimiert wird, bei gegebenen Stichprobenumfängen der Querschnittdesigns. Beispielsweise ist für $N^t > \mathfrak{n}^{t+1} + n^t$ $Pr(\psi_k^t = 1) = 0$ und somit auch $n^{t,t+1} = 0$ für alle $t \in \mathscr{T}$.

Für viele Erhebungen ist es jedoch gerade von besonderem Interesse, dass sich aufeinanderfolgende Stichproben zu einem bestimmten Grad überlappen, d.h. auch einen Panelteil beinhalten. Rotationspanels bieten hier einen Mittelweg zwischen einer klassischen Panelstudie, bei welcher nach einmaliger Ziehung einer Querschnittstichprobe in $t = 1$ alle Elemente bis zum Ende des Beobachtungszeitraums erhoben werden und einer Querschnittstudie, die zu jedem Beobachtungszeitpunkt eine neue Querschnittstichprobe zieht. Die Vorteile überlappender Stichproben bestehen darin, dass sich Veränderungsmaße oftmals effizienter schätzen lassen und Zeitreihenanalysen möglich sind, wohingegen es bei Querschnittstudien einfacher ist, unverzerrte Statistiken für Querschnitte zu schätzen, insbesondere, wenn die Population nicht konstant über die Zeit hinweg ist. Bei einem klassischen Panel ist dies schwieriger, da die vorhandene Stichprobe nur eine Teilmenge des Ziehungsrahmens $\mathscr{U}^1$ sein kann. Andere Aspekte eines Panels, die von Nachteil sein können, sind die ungleiche Verteilung der Erhebungslast und dem damit möglichen einhergehenden Anstieg der Teilnahmeverweigerung im Zeitverlauf sowie die natürliche Verkleinerung des Stichprobenumfangs durch Sterbefälle. Rotationspanels vereinen nun Querschnitt- und Panelstudien, indem in festgelegten Zeitintervallen ein Teil der Stichprobe durch eine neu gezogene Querschnittstichprobe ersetzt wird und diese neugezogenen Elemente wiederum nach einem festlegten Rotationsschema erhoben werden. Ein Rotationsschema sollte nach Möglichkeit unverzerrte Schätzungen im Querschnitt ermöglichen, eine im Erwartungswert gleiche Belastung der Elemente haben und möglichst effiziente Schätzungen von Veränderungsmaßen ermöglichen (Steel & McLaren, 2009, Kap. 3).

In der Praxis kommen eine Vielzahl von Rotationsschemata zur Anwendung. Eine der einfachsten Formen ist ein Längsschnittdesign mit einer fixen Verweildauer. Beispielsweise wird festgelegt, dass ein gezogenes Element für exakt d aufeinander folgende Erhebungszeitpunkte in der Stichprobe verbleiben soll. Ein weiteres weit verbreitetes Rotationsschema stellt eine fixe Verweildauer mit Unterbrechungen dar. Bei diesen sog. *in-out-in* Rotationen bleibt ein gezogenes Element zunächst für d_1 Perioden in der Stichprobe, verlässt sie dann für d_2 Perioden und kehrt danach für d_1 Perioden wieder zurück. Dieser Vorgang kann auch entsprechend häufig wiederholt werden, bis das Element eine gewisse Belastung erreicht hat. Die Verallgemeinerung dieser Rotationsschemata soll mit $d_1^m\_d_2^{m-1}$ bezeichnen werden. Dies bedeutet, dass ein Element nach seiner Ziehung alternierend für d_1 Perioden in der Stichprobe verweilt und sie für d_2 Perioden verlässt. Dabei wird der Beobachtungszyklus d_1 m-mal wiederholt und der Pausenzyklus $(m-1)$-mal. Das Rotationsschema sollte nach Möglichkeit sicherstellen, dass die gemeinsame Stichprobe sowohl im Längsschnitt, als auch im Querschnitt ausbalanciert ist. Dies ist sichergestellt, wenn für jeden beliebigen Beobachtungszeitpunkt t die gleiche Anzahl von Elementen zum ersten Mal (im Erhebungszyklus), zum zweiten Mal, usw., zum md_1-ten Mal erhoben wird (Park et al., 2001, S. 1485).

Für beliebige Rotationsschemata lässt sich die gemeinsame Stichprobe des Rotationspanels, schreiben als

$$ {}^*\mathfrak{S} = \mathfrak{S}\mathbf{R}\,. \tag{2.61}$$

Für $d_1^m\_d_2^{m-1}$ Rotationsschemata hat die $T \times T$ Matrix $\mathbf{R}$ die folgende Form

$$\mathbf{R} = \begin{pmatrix} r_1 & r_2 & \cdots & \cdots & r_M & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & r_1 & r_2 & \cdots & \cdots & r_M & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & r_1 & r_2 & \cdots & \cdots & r_M \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & r_1 & r_2 & \cdots & r_{M-1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & r_1 \end{pmatrix}. \tag{2.62}$$

Dabei ist r_1 das erste, r_2 das zweite, usw., Element des Vektors

$$\vec{r} = (\vec{d}_{1,1}, \vec{d}_{2,1}, \vec{d}_{1,2}, \vec{d}_{2,2}, \ldots, \vec{d}_{2,m-1}, \vec{d}_{1,m})$$

mit $m_1 + (m-1)d_2 = M$. Des Weiteren ist $\vec{d}_{1,i} = \vec{1}_{d_1}$, mit $\vec{1}_{d_1}$ als $1 \times d_1$ Einsvektor für $i = 1, \ldots, m$ und $\vec{d}_{2,i} = \vec{0}_{d_2}$, mit $\vec{0}_{d_2}$ als ein $1 \times d_2$ Vektor mit Nullen, für $i = 1, \ldots, m-1$. Die Querschnittstichprobe des Rotationspanels ${}^{*}\vec{\mathfrak{s}}^{t}$, besteht somit aus einer Summe von Querschnittstichproben aus $\mathfrak{S}$.

Aus (2.61) ist auch leicht ersichtlich, dass die Matrix ${}^{*}\mathbf{\Pi}_k^{t,u}$ der Inklusionswahrscheinlichkeiten des Längsschnittdesign ${}^{*}p_k$ des k-ten Elements der Rotationsstichprobe, gegeben ist durch

$${}^{*}\mathbf{\Pi}_k^{t,u} = \mathbf{R}^{\top}\mathbf{\Pi}_k^{t,u}\mathbf{R}\,.$$

Folglich ist für gemeinsame Designs $p(.)$, mit $\mathbf{\Pi}_k^{t,u} = \mathbf{\Pi}^{t,u}\ \forall\, k \in \mathscr{U}$, die Matrix der erwarteten Überlappung zwischen allen Paaren von Querschnittstichproben in ${}^{*}\mathfrak{S}$, $\mathrm{E}\left({}^{*}\mathbf{n}^{t,u}\right)$, gegeben durch

$$\mathrm{E}\left({}^{*}\mathbf{n}^{t,u}\right) = N\mathbf{R}^{\top}\mathbf{\Pi}^{t,u}\mathbf{R}\,.$$

Für die gemeinsame Stichprobe $\mathfrak{S}$ in (2.61) kann ein beliebiges Design $p(.)$ verwendet werden. Es eignen sich jedoch hierzu vor allem Designs, die eine negative Koordination aufweisen. Mittels einer negativen Koordination kann so der Abstand zwischen den Selektionszeitpunkten der Elemente maximiert und daher die Ziehungen jedes Elements möglichst gleich über den Beobachtungszeitraum verteilt werden. Die Zeit zwischen den Ziehungen lässt dann Raum für das gewünschte Rotationsschema. So ist beispielsweise Ziehen ohne Zurücklegen für Stichprobe ${}^{*}\mathfrak{S}^{t}$ immer möglich, wenn es zu keinen Überlagerungen von Erhebungszyklen eines Elements kommen kann. Dies ist gegeben, falls

$$\phi_k^t \geq M \quad \forall\, k \in \mathscr{U} \text{und}\ t \in \mathscr{T}\ ,$$

mit der Ausnahme von $\phi_k^t = T - t$ in Verbindung mit $\mathfrak{I}_k^t = 0$, also für den Fall, dass Element k nicht mehr ausgewählt wird. Eine Voraussetzung für ein Rotationspanel, wie in (2.61) beschrieben, balanciert zu sein ist, dass die Querschnittstichproben in $\mathfrak{S}$ zu jedem Zeitpunkt den gleichen Stichprobenumfang haben und das Rotationsschema nach den ersten M Ziehungen voll implementiert ist. Folglich kann nur das Rotationspanel $({}^{*}\vec{\mathfrak{s}}^{t}, \ldots, {}^{*}\vec{\mathfrak{s}}^{T})$ mit $t \geq M$ balanciert sein.

Für veränderliche Populationen müssen bei der Bildung des Rotationspanels auch Geburten und Sterbefälle beachtet werden. Hierzu wird die Definition des Rotationspanels ${}^{*}\mathfrak{S}$ in (2.61) wie folgt angepasst

$${}^{*}\mathfrak{S} = (\mathfrak{S}\mathbf{R}) \circ \mathbf{U}\,. \tag{2.63}$$

Dabei ist $\mathbf{U}$ eine $N \times T$ Indikatormatrix der Form

$$\mathbf{U} = \left[\mathbb{1}(k, \mathscr{U}^{t})\right]_{\substack{k=1,\ldots,N\\ t=1,\ldots,T}}\ , \tag{2.64}$$

´mit

$$\mathbb{1}(k, \mathscr{U}^{t}) := \begin{cases} 1 & \text{für } k \in \mathscr{U}^{t} \\ 0 & \text{sonst} \end{cases}\ .$$

Da die Querschnittstichproben in ${}^{*}\mathfrak{S}$ nach (2.63) keine festen Stichprobenumfänge haben, ist das Rotationspanel nicht über den gesamten Beobachtungszeitraum hinweg balanciert.

Beispiel 2.17. Für Algorithmus 3 mit h wie in (2.14) und unter der Annahme von (2.17) ist $\pi_k^{t,u} = \pi^{t,u}\ \forall\, k \in \mathscr{U}$, sowie $n^{t,u} = 0$ für alle $|u-t| < r-1$, mit $r = \lceil N/n \rceil$. Um eine konstante Überlappung für eine Anzahl von aufeinanderfolgenden Stichproben zu erhalten, soll ein $d_1^m\_d_2^{m-1}$ Rotationsschema mit $m = 1$ und $d_1 = 3$ umgesetzt werden. Matrix $\mathbf{R}$ in (2.61) hätte somit die Form

$$\mathbf{R} = \begin{pmatrix} 1 & 1 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Diese *d-in* Rotation kann auch durch die Anwendung von Algorithmus 10 erzeugt werden. ${}^*\vec{\mathfrak{s}}^t$ besteht somit aus der Summe der Teilquerschnittstichproben $\vec{\mathfrak{s}}^t$, $\vec{\mathfrak{s}}^{t-1}$, $\ldots, \vec{\mathfrak{s}}^{t-\min(t-1,d-1)}$ und hat den Stichprobenumfang ${}^*n^t = \sum_{i=0}^{\min(t-1,d-1)} n^{t-i}$. Daher verlassen zu jedem Zeitpunkt $t > d$, $n^{t-d} = 3$ Elemente ${}^*\vec{\mathfrak{s}}^t$ und $n^t = 3$ Elemente werden hinzugefügt. Für $t \leq d$ befindet sich ${}^*\vec{\mathfrak{s}}^t$ im Aufbau, d.h. das Rotationsschema ist noch nicht voll implementiert und es werden in jeder Periode n^t neue Elemente aufgenommen. Die Querschnittstichproben ${}^*\vec{\mathfrak{s}}^t$ können immer ohne Zurücklegen gezogen werden, wenn $3 \leq r-1$.

Algorithmus 10 Rotationsschema d-in

```
for $t \in \mathscr{T}$ do
    if $t = 1$ then $^*\vec{s}^t \leftarrow \vec{s}^t$
    else
        if $t \leq d$ then $^*\vec{s}^t \leftarrow {}^*\vec{s}^{t-1} + \vec{s}^t$
        else$^*\vec{s}^t \leftarrow {}^*\vec{s}^{t-1} + \vec{s}^t - \vec{s}^{t-d}$
        end if
    end if
end for
```

Die Inklusionswahrscheinlichkeiten, die zu dem Rotationsschema in Algorithmus 10 korrespondieren, lassen sich wie folgt bestimmen,

$${}^*\pi_k^{t,u} = \sum_{i=0}^{\min(t-1,d-1)} \sum_{j=0}^{\min(u-1,d-1)} \pi_k^{t-i,u-i}, \tag{2.65}$$

$${}^*\pi_{k,l}^{t,u} = \sum_{i=0}^{\min(t-1,d-1)} \sum_{j=0}^{\min(u-1,d-1)} \pi_{k,l}^{t-i,u-i}. \tag{2.66}$$

${}^*\pi_k^{t,u}$ ist die Wahrscheinlichkeit, dass ein Element k sowohl in ${}^*\vec{\mathfrak{s}}^t$ als auch in ${}^*\vec{\mathfrak{s}}^u$ enthalten ist. ${}^*\pi_{k,l}^{t,u}$ ist die Wahrscheinlichkeit, dass k in ${}^*\vec{\mathfrak{s}}^t$ und l in ${}^*\vec{\mathfrak{s}}^u$ enthalten ist. Zudem sind ${}^*\pi_k^t$ und ${}^*\pi_{k,l}^t$ durch (2.65) bzw. (2.66) mit $t = u$ gegeben.

Die Erwartungswerte der Überlappungen lassen sich analog zur Inklusionswahrscheinlichkeit in (2.65) im Allgemeinen wie folgt darstellen,

$$\mathrm{E}\left({}^*\mathfrak{n}^{t,u}\right) = \sum_{i=0}^{\min(t-1,d-1)} \sum_{j=0}^{\min(u-1,d-1)} E\left(\mathfrak{n}^{t-i,u-i}\right). \tag{2.67}$$

Tabelle 2.10 enthält $\mathrm{E}(\mathfrak{n}^{t,u})$ $\forall\, t,u \in \mathscr{T}$ für den Fall, dass $n^t = n = 3$, $N^t = N = 16$ und $T = 14$. Alle $\mathfrak{n}^{1,u}$ mit $u = 1, \ldots, 10$ sind hier deterministisch. Zeitpunkt $t = 6$ ist hier der erst mögliche Zeitpunkt für ein Element, das in $t = 1$ gezogen wurde, wieder in die Stichprobe zu gelangen. Bei der Ziehung der 3 Elemente in $t = 6$ muss das Element gezogen werden, welches noch nie gezogen wurde. Die übrigen zwei Elemente werden aus der Gruppe jener Elemente entnommen, die zum Zeitpunkt $t = 1$ gezogen wurden. Folglich ist $\mathfrak{n}^{1,6} = 2$ nicht variabel. Entsprechend ist $\mathfrak{n}^{1,7} = 1$, denn in $t = 7$ muss das Element entnommen werden, das in $t = 1$ aber nicht in $t = 6$ gezogen wurde. Der Zeitpunkt $t = 11$ ist für ein Element, das in $t = 1$ gezogen wurde, der erste mögliche Zeitpunkt, zum dritten Mal gezogen zu werden. Die Überlappung $\mathfrak{n}^{1,11}$ ist nun variablel, da in $\vec{\mathfrak{s}}^{11}$ sowohl ein Element enthalten sein kann, welches in $t = 6$ zum ersten Mal gezogen wurde, als auch Elemente, die in $t = 1$ und $t = 6$ gezogen wurden.

Tabelle 2.10: $\mathrm{E}(\mathfrak{n}^{t,u})$ für Algorithmus 3 mit (2.14) sowie $n^t = n = 3$ und $N^t = N = 16$

$\mathrm{E}(\mathfrak{n}^{t,u})$		u													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
t	1	3,00	0,00	0,00	0,00	0,00	2,00	1,00	0,00	0,00	0,00	1,33	1,33	0,33	0,00
	2	0,00	3,00	0,00	0,00	0,00	0,00	2,00	1,00	0,00	0,00	0,00	1,33	1,33	0,33
	3	0,00	0,00	3,00	0,00	0,00	0,00	0,00	2,00	1,00	0,00	0,00	0,00	1,33	1,33
	4	0,00	0,00	0,00	3,00	0,00	0,00	0,00	0,00	2,00	1,00	0,00	0,00	0,00	1,33
	5	0,00	0,00	0,00	0,00	3,00	0,00	0,00	0,00	0,00	2,00	1,00	0,00	0,00	0,00
	6	2,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00	0,00	0,00	2,00	1,00	0,00	0,00
	7	1,00	2,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00	0,00	0,00	2,00	1,00	0,00
	8	0,00	1,00	2,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00	0,00	0,00	2,00	1,00
	9	0,00	0,00	1,00	2,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00	0,00	0,00	2,00
	10	0,00	0,00	0,00	1,00	2,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00	0,00	0,00
	11	1,33	0,00	0,00	0,00	1,00	2,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00	0,00
	12	1,33	1,33	0,00	0,00	0,00	1,00	2,00	0,00	0,00	0,00	0,00	3,00	0,00	0,00
	13	0,33	1,33	1,33	0,00	0,00	0,00	1,00	2,00	0,00	0,00	0,00	0,00	3,00	0,00
	14	0,00	0,33	1,33	1,33	0,00	0,00	0,00	1,00	2,00	0,00	0,00	0,00	0,00	3,00

Tabelle 2.11 enthält die Werte von Matrix $\mathrm{E}({}^{*}\boldsymbol{\mathfrak{n}}^{t,u})$. Entsprechend den obigen Ausführungen zu Tabelle 2.10 ist ${}^{*}\mathfrak{n}^{t,u}$ konstant für alle $t,u \leq 10$, sowie auch für $t,u \in \mathscr{T}$ mit $|u-t| \leq 2$.

△

Wie Beispiel 2.17 deutlich zeigt, ist es einfach möglich (für eine unveränderliche Population), ein Design für Rotationsstichproben zu finden, welches balanciert für den Erhebungszyklus ist. Wird die gemeinsame Stichprobe durch Algorithmus 3 mit (2.36) gezogen, ist eine Balancierung über einen beliebig langen Beobachtungszeitraum jedoch nur möglich für $n^t = n\ \forall\, t \in \mathscr{T}$ und zusätzlich $N \mod n = 0$. Nur dann ist $\mathfrak{n}^{t,u}$ für alle $t,u \in \mathscr{T}$ keine Zufallsvariable und somit ${}^{*}\mathfrak{n}^{t,u}$ auch nicht.

2.6 Zielkonflikte zwischen Quer- und Längsschnittdesign

Wie Nedyalkova et al. (2009) anmerken, besteht offensichtlich ein Konflikt zwischen der Wahl des Längsschnitt- und des Querschnittdesigns. Die Festlegung des Längsschnittdesigns ermöglicht es, die gewünschte Koordination zu implementieren, jedoch

Tabelle 2.11: $\mathrm{E}\left(^*\mathfrak{n}^{t,u}\right)$ für Rotationsschema $d_1^m\_d_2^{m-1}$ mit $m = 1$, $d_1 = 3$ in Verbindung mit Algorithmus 3 mit (2.14) sowie $n^t = n = 3$ und $N^t = N = 16$

$\mathrm{E}\left(^*\mathfrak{n}^{t,u}\right)$		u													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
t	1	3,00	3,00	3,00	0,00	0,00	2,00	3,00	3,00	1,00	0,00	1,33	2,67	3,00	1,67
	2	3,00	6,00	6,00	3,00	0,00	2,00	5,00	6,00	4,00	1,00	1,33	4,00	5,67	4,67
	3	3,00	6,00	9,00	6,00	3,00	2,00	5,00	8,00	7,00	4,00	2,33	4,00	7,00	7,33
	4	0,00	3,00	6,00	9,00	6,00	3,00	2,00	5,00	8,00	7,00	4,00	2,33	4,00	7,00
	5	0,00	0,00	3,00	6,00	9,00	6,00	3,00	2,00	5,00	8,00	7,00	4,00	2,33	4,00
	6	2,00	2,00	2,00	3,00	6,00	9,00	6,00	3,00	2,00	5,00	8,00	7,00	4,00	2,33
	7	3,00	5,00	5,00	2,00	3,00	6,00	9,00	6,00	3,00	2,00	5,00	8,00	7,00	4,00
	8	3,00	6,00	8,00	5,00	2,00	3,00	6,00	9,00	6,00	3,00	2,00	5,00	8,00	7,00
	9	1,00	4,00	7,00	8,00	5,00	2,00	3,00	6,00	9,00	6,00	3,00	2,00	5,00	8,00
	10	0,00	1,00	4,00	7,00	8,00	5,00	2,00	3,00	6,00	9,00	6,00	3,00	2,00	5,00
	11	1,33	1,33	2,33	4,00	7,00	8,00	5,00	2,00	3,00	6,00	9,00	6,00	3,00	2,00
	12	2,67	4,00	4,00	2,33	4,00	7,00	8,00	5,00	2,00	3,00	6,00	9,00	6,00	3,00
	13	3,00	5,67	7,00	4,00	2,33	4,00	7,00	8,00	5,00	2,00	3,00	6,00	9,00	6,00
	14	1,67	4,67	7,33	7,00	4,00	2,33	4,00	7,00	8,00	5,00	2,00	3,00	6,00	9,00

müssen die sich hieraus ergebenden Querschnittdesigns akzeptiert werden. Umgekehrt schränkt die Wahl des Querschnittdesigns die Möglichkeit zur Koordination ein. Insbesondere gilt dies für Designs mit festen Stichprobenumfängen. Dies wird besonders bei geschichteten Querschnittdesigns deutlich. Die Schichtung stellt dabei eine Art Koordination im Querschnitt dar, die nur schwer mit der im Längsschnitt vereinbar ist (Rivière, 2001, Nedyalkova, 2009). Dieses Problem besteht für Querschnittdesigns mit ungleichen Inklusionswahrscheinlichkeiten und festen Stichprobenumfängen allgemein. So lassen sich zwar für $t < u$ bedingte Querschnittdesigns $p^u(\vec{s}^u|\vec{s}^t)$ frei wählen, das Problem besteht aber darin, hieraus das Querschnittdesign $p^u(.)$ abzuleiten (Deville & Tillé, 2000, Tillé & Favre, 2004, Nedyalkova et al., 2009). Bekannte Designs für die dies möglich ist, sind nicht sequenziell (Matei & Tillé, 2005b, Matei & Skinner, 2009).

Aus diesem Grund wird die folgende Vermutung aufgestellt:

Vermutung 2.18. Ein Algorithmus für ein strikt sequenzielles gemeinsames Design mit beliebigen Inklusionswahrscheinlichkeiten erster Ordnung und festen Stichprobenumfängen im Querschnitt, sowie negativer oder positiver Koordination, existiert nicht.

Auch wenn Vermutung 2.18 nicht allgemein zutreffen sollte, gilt sie doch für die Menge der bekannten Algorithmen zur Implementierung sequenzieller Designs (siehe hierzu z.B. Rosén, 1997a,b, Ohlsson, 1998, Deville & Tillé, 2000, Kröger et al., 2003).

Kapitel 3

Schätzung von Statistiken im Querschnitt

Dieses Kapitel widmet sich der Schätzung von Statistiken einer interessierenden Variable zu einem einzigen Beobachtungszeitpunkt, d.h. es kann von $T = 1$ ausgegangen werden. Aus diesem Grund wird bis auf Weiteres auf Indizes zur Unterscheidung von verschiedenen Beobachtungszeitpunkten abgesehen und es wird auf die Notation aus Abschnitt 2.1.1 zurückgegriffen. Es sei y die interessierende Variable, auch Untersuchungsvariable genannt, für die $y(k) = y_k$ der Wert des k-ten Elements ist. $\vec{y} = (y_1, \ldots, y_N)^\top$ bezeichnet den (unbekannten) Parameter der endlichen Population, beschrieben durch $\mathcal{U}$. Jede reelle Funktion von $\vec{y}$ wird Statistik genannt. Das Ziel einer Stichprobenerhebung ist es, eine Inferenz bezüglich einer, oder mehrerer Statistiken θ von $\vec{y}$ zu erstellen unter Verwendung der bekannten Stichprobendaten $\{k, y_k | k \in s\}$. Dabei lässt sich θ in der Regel als eine Funktion von $\vec{y}$ darstellen.

Statistiken von Interesse können beispielsweise der Total- oder Mittelwert von $\vec{y}$ sein oder komplexere Funktionen, z.B. jene Maße für Armut und Einkommensungleichheit, wie sie in Abschnitt 3.4 beschrieben sind. Ein Schätzer für eine solche Statistik von $\vec{y}$ ist eine reelle Funktion $Q(\vec{\mathfrak{s}}, \vec{y})$, wobei $\mathfrak{q} = Q(\vec{\mathfrak{s}}, \vec{y})$ für den Schätzer und $q = Q(\vec{s}, \vec{y})$ für dessen realisierten Wert bezüglich der gezogenen Stichprobe $\vec{s}$ steht.

Der Erwartungswert und die Varianz von $\mathfrak{q}$ sind für Design $p(.)$ und dessen Träger $\mathcal{S}$ gegeben durch

$$\mathrm{E}(\mathfrak{q}) = \sum_{\vec{s} \in \mathcal{S}} p(\vec{s}) q \,,$$

bzw.

$$\begin{aligned} \mathrm{V}(\mathfrak{q}) &= \mathrm{E}\left([\mathfrak{q} - \mathrm{E}(\mathfrak{q})]^2 \right) \\ &= \sum_{\vec{s} \in \mathcal{S}} p(\vec{s}) \left[\mathfrak{q} - \mathrm{E}(\mathfrak{q})\right]^2 \\ &= \mathrm{E}\left(\mathfrak{q}^2\right) - \left[\mathrm{E}(\mathfrak{q})\right]^2 \,. \end{aligned}$$

Die Kovarianz zwischen zwei Schätzern $\mathfrak{q}_1 = Q_1(\vec{\mathfrak{s}}, \vec{y})$ und $\mathfrak{q}_2 = Q_2(\vec{\mathfrak{s}}, \vec{y})$ ist gegeben durch

$$\begin{aligned}
\operatorname{COV}(\mathfrak{q}_1, \mathfrak{q}_2) &= \operatorname{E}([\mathfrak{q}_1 - \operatorname{E}(\mathfrak{q}_1)][\mathfrak{q}_2 - \operatorname{E}(\mathfrak{q}_2)]) \\
&= \sum_{\vec{s} \in \mathscr{G}} p(\vec{s})\,[q_1 - \operatorname{E}(\mathfrak{q}_1)]\,[q_2 - \operatorname{E}(\mathfrak{q}_2)] \\
&= \operatorname{E}([\mathfrak{q}_1 \mathfrak{q}_2]) - \operatorname{E}(\mathfrak{q}_1)\operatorname{E}(\mathfrak{q}_2)\ .
\end{aligned}$$

Die obigen Erwartungswerte werden über alle möglichen Stichproben des Trägers $\mathscr{G}$ mit Design $p(.)$ gebildet. In diesem Zusammenhang wird darum auch von Designerwartungswert, (bzw. Designvarianz und Designkovarianz) gesprochen. Im weiteren Verlauf wird auf diese explizite Hervorhebung, falls nichts anderweitig ausdrücklich erwähnt, verzichtet, da im Fokus des Interesses maßgeblich die designbasierte Inferenz steht (Särndal et al., 1992, S. 34-35).

3.1 Horvitz-Thompson Schätzer

Zunächst wird der Fokus der Untersuchung auf den Totalwert von $\vec{y}$ gelegt, der mit τ bezeichnet wird, d.h.

$$\tau = \sum_{k \in \mathscr{U}} y_k\ . \tag{3.1}$$

Der *Horvitz-Thompson* Schätzer (Horvitz & Thompson, 1952) gehört zur Klasse der linearen Schätzer und ist als Schätzer für τ wie folgt definiert:

$$\begin{aligned}
\hat{\tau}_\pi &= \sum_{k \in \mathscr{U}} \mathfrak{I}_k \frac{y_k}{\pi_k} \\
&= \vec{\mathfrak{s}}^\top \check{\vec{y}}\,,
\end{aligned} \tag{3.2}$$

wobei $\check{\vec{y}} = (\frac{y_1}{\pi_1}, \ldots, \frac{y_N}{\pi_N})^\top$. Aufgrund der Gewichtung der Beobachtung y_k in der Stichprobe mit der Inversen der Inklusionswahrscheinlichkeit erster Ordnung π_k wird der Schätzer in (3.2) im Folgenden π-Schätzer genannt.

Der Schätzer $\hat{\tau}_\pi$ ist erwartungstreu für τ und Design $p(.)$ mit dem Träger $\mathscr{G}$, falls $\pi_k > 0 \; \forall \; k \in \mathscr{U}$, d.h.

$$\operatorname{E}(\hat{\tau}_\pi) = \tau\ ,$$

da

$$\begin{aligned}
\sum_{\vec{s} \in \mathscr{G}} p(\vec{s})\,\hat{\tau}_\pi(\vec{s}) &= \sum_{\vec{s} \in \mathscr{G}} p(\vec{s}) \sum_{k \in \mathscr{U}} I_k \frac{y_k}{\pi_k} \\
&= \sum_{k \in \mathscr{U}} \sum_{\vec{s} \in \mathscr{G}} p(\vec{s}) I_k \frac{y_k}{\pi_k} \\
&= \sum_{k \in \mathscr{U}} y_k\ .
\end{aligned}$$

Die Varianz von $\hat{\tau}_\pi$, hat die folgende Form:

$$\begin{aligned} \mathrm{V}\left(\hat{\tau}_\pi\right) &= \sum_{k \in \mathscr{U}} \sum_{l \in \mathscr{U}} \left(\pi_{k\,l} - \pi_k \pi_l\right) \frac{y_k}{\pi_k} \frac{y_l}{\pi_l} \\ &= \sum_{k \in \mathscr{U}} \pi_k (1 - \pi_k) \left(\frac{y_k}{\pi_k}\right)^2 + 2 \sum_{k \in \mathscr{U}} \sum_{\substack{l \in \mathscr{U} \\ l < k}} \left(\pi_{k\,l} - \pi_k \pi_l\right) \frac{y_k}{\pi_k} \frac{y_l}{\pi_l} \, . \end{aligned} \tag{3.3}$$

Für die Herleitung der Varianz in (3.3) siehe Särndal et al. (1992, S. 43f.) bzw. Cochran (1977, S. 260). In Anlehnung an die Schreibweise $\hat{\tau}_\pi = \vec{\mathfrak{s}}^\top \check{\vec{y}}$ lässt sich $\mathrm{V}\left(\hat{\tau}_\pi\right)$ alternativ in der quadratischen Form $\check{\vec{y}}^\top \boldsymbol{\Sigma}_\pi \check{\vec{y}}$ darstellen. Dabei ist

$$\boldsymbol{\Sigma}_\pi = \left[\mathrm{COV}\left(\mathfrak{I}_k, \mathfrak{I}_l\right)\right]_{\substack{k=1,\dots,N \\ l=1,\dots,N}} = \left[\pi_{k\,l} - \pi_k \pi_l\right]_{\substack{k=1,\dots,N \\ l=1,\dots,N}} \tag{3.4}$$

Unter der Bedingung, dass $\pi_{k\,l} > 0 \; \forall \, k, l \in \mathscr{U}$, ist ein unverzerrter Schätzer für $\mathrm{V}\left(\hat{\tau}_\pi\right)$ gegeben durch

$$\widehat{\mathrm{V}}\left(\hat{\tau}_\pi\right) = \sum_{k \in \mathscr{s}} (1 - \pi_k) \left(\frac{y_k}{\pi_k}\right)^2 + 2 \sum_{k \in \mathscr{s}} \sum_{\substack{l \in \mathscr{s} \\ l < k}} (1 - \frac{\pi_k \pi_k}{\pi_{k\,l}}) \frac{y_k}{\pi_k} \frac{y_l}{\pi_l} \, , \tag{3.5}$$

(Särndal et al., 1992, S. 44).

Für den Fall, dass $p(.)$ ein Design mit festem Stichprobenumfang ist, lässt sich die Varianz von $\hat{\tau}_\pi$ alternativ auch schreiben als

$$\begin{aligned} \mathrm{V}\left(\hat{\tau}_\pi\right) &= -\frac{1}{2} \sum_{k \in \mathscr{U}} \sum_{l \in \mathscr{U}} \left(\pi_{k\,l} - \pi_k \pi_l\right) \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l}\right)^2 \\ &= \sum_{k \in \mathscr{U}} \sum_{\substack{l \in \mathscr{U} \\ l < k}} \left(\pi_k \pi_l - \pi_{k\,l}\right) \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l}\right)^2 , \end{aligned} \tag{3.6}$$

(Yates & Grundy, 1953, S. 257) . Für $\pi_{k\,l} > 0 \; \forall \, k, l \in \mathscr{U}$ ist ein unverzerrter Schätzer von (3.6) durch

$$\widehat{\mathrm{V}}\left(\hat{\tau}_\pi\right) = \sum_{k \in \mathscr{U}} \sum_{\substack{l \in \mathscr{U} \\ l < k}} \frac{\pi_k \pi_l - \pi_{k\,l}}{\pi_{k\,l}} \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l}\right)^2 , \tag{3.7}$$

(Cochran, 1977, S. 260f.) gegeben. Damit Varianzschätzer (3.7) strikt größer gleich Null ist, ist eine hinreichende Bedingung, dass $\pi_k \pi_l \geq \pi_{k\,l}, \forall \, k,l \in \mathscr{U}, k \neq l$ (Raj, 1968, S. 55).

Der Vorteil der Varianzschätzer in (3.7) und (3.5) liegt in ihrer generischen Form. In der Anwendung erweisen sie sich jedoch für eine Vielzahl von Stichprobendesigns als wenig praktikabel. Zunächst bringt die doppelte Summe in beiden Varianzschätzern eine schnell wachsende Anzahl von Ausdrücken mit sich, die berechnet werden müssen. Zum anderen kann es aufwändig sein, alle Inklusionswahrscheinlichkeiten zweiter Ordnung zu bestimmen. Dies gilt vor allem für komplexe Designs. Dabei handelt es sich insbesondere um Designs mit ungleichen Inklusionswahrscheinlichkeiten und festen Stichprobenumfängen, eine Kombination, wie sie in der Praxis oftmals zur Anwendung kommt. Nur für einfache Designs mit festen Stichprobenumfängen, wie der einfachen Zufallsstichprobe, sind die Inklusionswahrscheinlichkeiten zweiter Ordnung einfach zu bestimmen bzw. konstant für alle $k, l \in \mathscr{U}$.

3.1.1 Einfache Zufallsstichprobe

Für eine einfache Zufallsstichprobe (SRS) mit einem Stichprobenumfang von n ist

$$p(\vec{s}) = \frac{1}{\binom{N}{n}}$$

für alle $\vec{s} \in \mathscr{S}_n = \binom{N}{n}$. Dieses Design wird einfach genannt, da jede mögliche Stichprobe die gleiche Wahrscheinlichkeit hat, ausgewählt zu werden (Tillé, 2006). Die Inklusionswahrscheinlichkeiten erster und zweiter Ordnung sind folglich

$$\pi_k = \frac{n}{N} \quad \forall\, k \in \mathscr{U}$$

bzw.

$$\pi_{k\,l} = \frac{n(n-1)}{N(N-1)} \quad \forall\, k,l \in \mathscr{U} \text{ mit } k \neq l\,.$$

Unter SRS lässt sich $\hat{\tau}_\pi$ schreiben als

$$\hat{\tau}_\pi = N\bar{y}\,,$$

mit $\bar{y} = \dfrac{1}{n} \sum_{k \in \mathscr{U}} \mathfrak{I}_k y_k$ als Stichprobenmittel.

Werden die Inklusionswahrscheinlichkeiten in (3.7) eingesetzt, ergibt sich

$$\mathrm{V}\left(\hat{\tau}_\pi\right) = N^2 \frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)\,, \tag{3.8}$$

wobei $\sigma^2 = \frac{\sum_{k=1}^{N} (y_k - \frac{1}{N}\tau)^2}{N-1}$ als Varianz von $\vec{y}$ bezeichnet wird (Särndal et al., 1992, S. 46). Durch Ersetzen von σ^2 durch s^2, wobei

$$s^2 = \frac{1}{n-1} \sum_{k \in \mathscr{s}} (y_k - \bar{y})^2$$

die Stichprobenvarianz von $\vec{y}$ ist, ergibt sich ein unverzerrter Schätzer für (3.8).

3.1.2 Poisson Design

Ein *Poisson* Design hat einen zufälligen Stichprobenumfang und ungleiche Inklusionswahrscheinlichkeiten, mit $0 < \pi_k < 1\ \forall\ k \in \mathscr{U}$[1]. Das Stichprobendesign ist gegeben durch

$$p(\vec{s}) = \prod_{k \in \mathscr{U}} \pi_k^{I_k} (1 - \pi_k)^{(1 - I_k)}\,.$$

Da die Ziehungen der Elemente unabhängig voneinander erfolgen, ist $\pi_{k\,l} = \pi_k \pi_l$ und somit $\mathrm{COV}\left(\mathfrak{I}_k, \mathfrak{I}_l\right) = 0\ \forall\ k \neq l \in \mathscr{U}$. Entsprechend lässt sich die Varianz von $\hat{\tau}_\pi$ unter einem Poisson Design schreiben als

$$\mathrm{V}\left(\hat{\tau}_\pi\right) = \sum_{k \in \mathscr{U}} \pi_k (1 - \pi_k) \left(\frac{y_k}{\pi_k}\right)^2\,. \tag{3.9}$$

---

[1] Der Fall mit gleichen Inklusionswahrscheinlichkeiten wird als *Bernoulli* Design bezeichnet.

Ein unverzerter Schätzer für die Varianz in (3.9) ist

$$\widehat{\mathrm{V}}\left(\hat{\tau}_\pi\right) = \sum_{k\in\mathscr{U}} \mathfrak{I}_k^t (1-\pi_k) \left(\frac{y_k}{\pi_k}\right)^2 . \tag{3.10}$$

An dieser Stelle sei angemerkt, dass der Schätzer $\hat{\tau}_\pi$ eine relativ hohe Varianz haben kann im Vergleich zu Designs mit einem festen Stichprobenumfang, der dem erwarteten Stichprobenumfang des Poisson Designs entspricht. Ein hoher Teil dieser Varianz ist auf den zufälligen Stichprobenumfang zurückzuführen, dessen Erwartungswert durch $\sum_{k\in\mathscr{U}} \pi_k$ gegeben ist. Folglich stellt sich die Frage, wie die Inklusionswahrscheinlichkeiten zu wählen sind, um die Varianz in (3.9) zu minimieren. Unter der Bedingung, dass $y_k > 0 \;\forall\; k \in \mathscr{U}$, kann dies erreicht werden, indem π_k proportional zum Anteil von y_k an τ gewählt wird, d.h.

$$\pi_k = \frac{n y_k}{\tau} \text{ für } \quad k = 1, \ldots, N$$

vorausgesetzt, dass $n y_k \leq \tau \;\forall k \in \mathscr{U}$. Da $\vec{y}$ jedoch unbekannt ist, wird zur Planung des Stichprobendesigns anstelle der Untersuchungsvariable oftmals eine mit y hoch korrelierte Hilfsvariable χ verwendet. Hierzu soll x_k die Ausprägung von χ für das k-te Element in $\mathscr{U}$ sein. Falls $\frac{y_k}{x_k}$ konstant oder nahezu konstant für alle $k \in \mathscr{U}$ ist, wird die Varianz von $\hat{\tau}_\pi$ nicht weit von ihrem Minimum abweichen (Särndal et al., 1992, S.,86f).

Interessanterweise ist für eine feste Relation zwischen Inklusionswahrscheinlichkeit und Untersuchungsvariable, d.h. $\pi_k = y_k \lambda$, mit $\lambda \in \mathbb{R}$ unter der Bedingung $0 < y_k \lambda \leq 1$ für alle $k \in \mathscr{U}$, bei allen Designs mit festem Stichprobenumfang $\mathrm{V}\left(\hat{\tau}_\pi\right) = 0$. Dieser Umstand ist direkt ersichtlich aus der Darstellung der Varianz in (3.6), bei welcher in diesem Fall der quadratische Faktor für jeden Summanden Null ist.

3.2 Varianzschätzung bei Ziehen mit Zurücklegen

Um die Kalkulation der doppelten Summe in (3.7) bzw. (3.5) im Falle von komplexen Designs zu umgehen, finden sich in der Literatur verschiedene Approximationen, die ausschließlich die Kenntnis der π_k voraussetzen, nicht jedoch die der π_{kl}, (Berger & Skinner, 2004). Ein anderer Weg, diese aufwändigen Berechnungen zu umgehen, ist die Verwendung von Varianzschätzern für Designs mit Zurücklegen. Für Ziehen ohne Zurücklegen bezeichnet p_k die Wahrscheinlichkeit, dass ein bestimmtes Element k, bei einem beliebigen Zug der n_{wr} Züge entnommen wird. Diese Selektionswahrscheinlichkeiten p_k, $k = 1, \ldots, N$, können frei gewählt werden, jedoch soll $p_k > 0 \;\forall\; k \in \mathscr{U}$ sein und $\sum_{k\in\mathscr{U}} p_k = 1$. Sind alle p_k gleich, wird von einer einfachen Zufallsstichprobe mit Zurücklegen gesprochen (SRSWR). Der sog. *Hansen-Hurwitz* Schätzer $\hat{\tau}_{pwr}$ für τ bei Ziehen mit Zurücklegen, vorgeschlagen von Hansen & Hurwitz (1943), gewichtet die beobachteten Werte mit den Selektionswahrscheinlichkeiten der jeweiligen Elemente. So ist

$$\hat{\tau}_{pwr} = \frac{1}{n_{wr}} \sum_{i=1}^{n_{wr}} \frac{y_{k_i}}{p_{k_i}} . \tag{3.11}$$

Dabei ist $k_i = l$, wenn im i-ten Zug das l-te Element gezogen wurde (Särndal et al., 1992, S. 51f.). Der Index *pwr* steht für das Expandieren der y_k mit p_k^{-1} beim Ziehen mit Zurücklegen. Es besteht der folgende Zusammenhang zwischen π_k und p_k,

$$\begin{aligned}\pi_k &= 1-(1-p_k)^{n_{wr}}\\ &= n_{wr}p_k + \sum_{l=2}^{\infty}\binom{n_{wr}}{l}(-p_k)^l \, .\end{aligned}$$

Sollte p_k klein sein, was bei großen N anzunehmen ist, dann ist $\pi_k \approx n_{rw}p_k$.

Wegen

$$Pr\left(\frac{y_{k_i}}{p_{k_i}} = \frac{y_k}{p_k}\right) = p_k; \qquad k = 1,\ldots,N\,,$$

ist $\hat{\tau}_{pwr}$ unverzerrt für τ, da $\mathrm{E}\left(\frac{y_{k_i}}{p_{k_i}}\right) = \sum_{k\in\mathscr{U}} \frac{y_k}{p_k} p_k = \tau$ und somit ist

$$\begin{aligned}\mathrm{E}\left(\hat{\tau}_{pwr}\right) &= \frac{1}{n_{wr}}\sum_{i=1}^{n_{wr}}\mathrm{E}\left(\frac{y_{k_i}}{p_{k_i}}\right)\\ &= \tau \, .\end{aligned}$$

Die Varianz von $\hat{\tau}_{pwr}$ ist

$$\mathrm{V}\left(\hat{\tau}_{pwr}\right) = \sum_{k\in\mathscr{U}}\frac{1}{n_{wr}}\left(\frac{y_k}{p_k}-\tau\right)^2 p_k \, . \tag{3.12}$$

Ein unverzerrter Schätzer für die Varianz von $\hat{\tau}_{pwr}$ ist gegeben durch

$$\widehat{\mathrm{V}}\left(\hat{\tau}_{pwr}\right) = \frac{1}{n_{wr}(n_{wr}-1)}\sum_{i=1}^{n_{wr}}\left(\frac{y_{k_i}}{p_{k_i}}-\hat{\tau}_{pwr}\right)^2 . \tag{3.13}$$

Beweise für (3.12) und (3.13) finden sich in Särndal et al. (1992, S. 52).

Ein Varianzschätzer der Form (3.13) kann nun alternativ als Varianzschätzer für $\hat{\tau}_\pi$ verwendet werden (Särndal et al., 1992, S. 422). So lässt sich der folgende Varianzschätzer für $\hat{\tau}_\pi$ unter einem hypothetischen Design mit Zurücklegen aufstellen:

$$\hat{\mathrm{V}}_0 = \frac{1}{n(n-1)}\sum_{k\in\mathscr{S}}\left(\frac{y_k}{p_k}-\hat{\tau}\right)^2 . \tag{3.14}$$

Dabei wird p_k mit $p_k = \frac{\pi_k}{n}$ angenommen für alle $k \in \mathscr{S}$ und das Design hat $n = n_{wr}$ Züge. Durch die Verwendung von (3.14) anstelle von (3.7) oder (3.5) wird ein Varianzschätzer für SRSWR verwendet, obwohl das eigentliche Design SRS ist. Der Preis dieses rechnerisch einfacheren Varianzschätzers, welcher ohne die Verwendung der Inklusionswahrscheinlichkeiten zweiter Ordnung auskommt, ist, dass dieser nicht länger unverzerrt für $\mathrm{V}\left(\hat{\tau}_\pi\right)$ ist. Diese Verzerrung ist für jedes Design mit einem festen Stichprobenumfang gegeben durch

$$\mathrm{E}_{wr}\left(\hat{V}_0\right) - V\left(\hat{\tau}_\pi\right) = \frac{n}{n-1}\left(V_0 - V\left(\hat{\tau}_\pi\right)\right) \, , \tag{3.15}$$

mit $V_0 = \frac{1}{n}\sum_{k\in\mathscr{U}}\left(\frac{y_k}{p_k} - \tau\right)^2 p_k$ und E_{wr} als der Erwartungswert in Bezug auf das Desgin mit Zurücklegen (siehe Särndal et al., 1992, S. 422 bzw. Durbin, 1953). Die Verzerrung in (3.15) ist positiv, wenn $\hat{\tau}_\pi$ für ein Design ohne Zurücklegen mit einem Stichprobenumfang von n effizienter ist als $\hat{\tau}_{pwr}$ für ein Design mit Zurücklegen mit n Zügen und $p_k = \pi_k/n$. In diesen Fällen führt die Verwendung von $\hat{V}_0$ zur Konstruktion von Konfidenzintervallen, die konservativ, d.h. zu lang sind.

Wie Raj (1968, S. 55f) zeigt, ist eine hinreichende Bedingung für $V_0(\hat{\tau}) > V(\hat{\tau})$ bei Designs mit festen Stichprobenumfängen gegeben durch

$$\pi_{kl} \geq \pi_k\pi_l(1-\frac{1}{n})\ \forall\, k,l \in \mathscr{U}, k \neq l\,, \tag{3.16}$$

was ersichtlich wird durch das Umschreiben der Varianz in (3.12) in der folgenden Weise

$$\mathrm{V}\left(\hat{\tau}_{pwr}\right) = \sum_{k=1}^{N}\frac{y_k^2}{n_{wr}p_k} - \frac{\tau^2}{n_{wr}} \tag{3.17}$$

$$= \frac{1}{n_{wr}}\sum_{k=1}^{N}\sum_{\substack{l=1\\l<k}}^{N}\left(\frac{y_k}{p_k} - \frac{y_l}{p_l}\right)^2 p_k p_l\,. \tag{3.18}$$

Da $p_k = \frac{\pi_k}{n_{wr}}$ ist folgt für $n_{wr} = n$, dass (3.17) immer größer als (3.6) ist, wenn die Bedingung in (3.16) erfüllt ist, wie dies z.B. für SRS der Fall ist. Die Bedingung in (3.16) ist hinreichend, aber nicht notwendig und für viele andere Designs ohne Zurücklegen als SRS schon nicht gegeben. Für sog. *connected* Designs[2] gibt Gabler (1984) die weniger strenge, jedoch immer noch nicht notwendige Bedingung

$$\sum_{k=1}^{N}\min_{1\leq l\leq N}\frac{\pi_{kl}}{\pi_l} \geq n-1\,,$$

für $V_0 \geq V(\hat{\tau}_\pi)$ an. Eine Demonstration der obigen Bedingung für den Ziehungsalgorithmus von Sampford (Sampford, 1967) findet sich ebenfalls in Gabler (1984).

Es lässt sich zusammenfassen, dass für manche Designs mit Zurücklegen der Varianzschätzer in (3.13) als eine einfach zu berechnende obere Abschätzung für die Varianz unter dem tatsächlichen Design dienen kann. Da

$$\mathrm{E}\left(\hat{V}_0\right) = \frac{1}{(n-1)}\sum_{k\in\mathscr{U}}\sum_{\substack{l\in\mathscr{U}\\l<k}}\left(\frac{y_k}{\pi_k} - \frac{y_k}{\pi_k}\right)\pi_{k,l}\,,$$

und somit $\mathrm{E}\left(\hat{V}_0\right)$ größer gleich (3.6) ist, wenn (3.16) gilt. Die Nutzbarkeit einer solchen Vereinfachung ist jedoch beschränkt durch den Umstand, dass die Inklusionswahrscheinlichkeiten zweiter Ordnung bekannt sein müssen.

---

[2] Ein Design wird als *connected* bezeichnet, wenn für alle $k,l \in \mathscr{U}\ k \neq l$ eine Folge von Elemten des Ziehungsrahmes $(k_i)_{i=1\ldots,m}$ existiert, für die gilt $\pi_{k,k_1}\pi_{k_1,k_2}\ldots\pi_{k_m,l} > 0$.

3.3 Varianz Approximationen

Durch eine alternative Formulierung der Varianz in (3.6), soll für Designs mit einem festen Stichprobenumfang im Folgenden eine erste Approximation für die Varianz von $\hat{\tau}_\pi$ gefunden werden.

Zuerst sollen einige allgemeine Eigenschaften der Inklusionswahrscheinlichkeit bei Designs mit festen Stichprobenumfängen aufgelistet werden, (Tillé, 1996, p. 184):

$$\sum_{\substack{k\in\mathcal{U}\\ k\neq l}} \pi_{k,l} = \pi_l(n-1) \tag{3.19a}$$

$$\sum_{\substack{k\in\mathcal{U}\\ k\neq l}} \pi_k\pi_l = \pi_l(n-\pi_l) \tag{3.19b}$$

$$\sum_{k\in\mathcal{U}} \sum_{\substack{l\in\mathcal{U}\\ l\neq k}} \pi_{k,l} = n(n-1) \tag{3.19c}$$

$$\sum_{k\in\mathcal{U}} \sum_{\substack{l\in\mathcal{U}\\ l\neq k}} \pi_l\pi_k = n^2 - \sum_{k\in\mathcal{U}} \pi_k^2 \tag{3.19d}$$

$$\sum_{\substack{l\in\mathcal{U}\\ l\neq k}} (\pi_k\pi_l - \pi_{k,l}) = \pi_k(1-\pi_k) \tag{3.19e}$$

Jetzt wird (3.6) umformuliert mit

$$\begin{aligned}
\mathrm{V}(\hat{\tau}_\pi) =& \frac{1}{2}\sum_{k=1}^{N}\sum_{l\neq k}^{N}(\pi_k\pi_j - \pi_{k,l})\left(\frac{y_k}{\pi_k} - \frac{\tau}{n} - \frac{y_l}{\pi_l} - \frac{\tau}{n}\right)^2 \\
=& \sum_{k=1}^{N}\sum_{l\neq k}^{N}(\pi_k\pi_l - \pi_{k,l})\left(\frac{y_k}{\pi_k} - \frac{\tau}{n}\right)^2 \\
& - \sum_{k=1}^{N}\sum_{l\neq k}^{N}(\pi_k\pi_l - \pi_{k,l})\left(\frac{y_k}{\pi_k} - \frac{\tau}{n}\right)\left(\frac{y_l}{\pi_l} - \frac{\tau}{n}\right),
\end{aligned} \tag{3.20}$$

und durch Verwendung von Beziehung (3.19e)

$$\begin{aligned}
\mathrm{V}(\hat{\tau}_\pi) =& \sum_{k=1}^{N}\pi_i(1-\pi_k)\left(\frac{y_k}{\pi_k} - \frac{\tau}{n}\right)^2 \\
& - \sum_{k=1}^{N}\sum_{l\neq k}^{N}(\pi_k\pi_l - \pi_{k,l})\left(\frac{y_k}{\pi_k} - \frac{\tau}{n}\right)\left(\frac{y_l}{\pi_l} - \frac{\tau}{n}\right).
\end{aligned} \tag{3.21}$$

Für $p_k = n\pi_k$ und $n = n_{wr}$ ist der erste Term in (3.21) gleich der Varianz von $\hat{\tau}_{pwr}$ multipliziert mit einem Endlichkeitskorrekturfaktor $(1-\pi_k)$. Somit kann dieser Term als eine erste Approximation zur Varianz in (3.21) verwendet werden, die ohne die Inklusionswahrscheinlichkeiten zweiter Ordnung bestimmt werden kann (Brewer, 2002, S. 149f.). Damit der Einfluss des zweiten Terms in (3.21), im Vergleich zum ersten, vernachlässigbar ist, muss das Design die Eingenschaft haben, dass $\pi_{k,l} \approx \pi_k\pi_l\ \forall k,l \in \mathcal{U}\ k \neq l$. Dies ist beispielsweise bei hinreichend großem N für SRS der Fall. Ist jedoch davon auszugehen, dass der zweite Term nicht zu vernachlässigen ist, müssen weitere, geeignetere Approximationen für die Inklusionswahrscheinlichkeiten zweiter Ordnung gefunden werden, d.h. Ausdrücke, die der Varianz von $\hat{t}_\pi$ näherungsweise entsprechen und die gleiche, einfache Form haben wie der erste Term in (3.21).

Um eine solche Approximation zu finden, wird zunächst das sog. bedingte Poisson Design betrachtet. Dabei handelt es sich um ein Poisson Design, dass auf einen festen Stichprobenumfang n konditioniert wird. Eine heuristische Beschreibung eines solchen Designs wäre

$$p_{\text{poiss}}(\vec{s}^*|\mathfrak{n}=n) = \frac{p_{\text{poiss}}(\vec{s}^*)}{P(\vec{s}^* \in \mathscr{S}_n)} \ .$$

Dabei ist $p_{\text{poiss}}(.)$ das unbedingte Poisson Design mit einem variablen Stichprobenumfang von $\mathfrak{n}$, den Inklusionswahrscheinlichkeiten π_k^* $(k = 1, \ldots, N)$, mit $\sum_{k=1}^N \pi_k^* = n$, und einer nach $p_{\text{poiss}}(.)$ ausgewählten Stichprobe $\vec{s}^*$. Des Weiteren ist $p_{\text{poiss}}(\vec{s}^*|\mathfrak{n} = n) = 0$ falls $\vec{s}^* \notin \mathscr{S}_n$ und $Pr(\vec{s}^* \in \mathscr{S}_n)$ bezeichnet die Wahrscheinlichkeit, unter Design $p_{\text{poiss}}(.)$ eine Stichprobe vom Umfang $\mathfrak{n} = n$ zu ziehen (Berger, 2004b, S. 454). Die Konditionierung auf einen festen Stichprobenumfang hat jedoch zur Folge, dass die Inklusionswahrscheinlichkeiten π_k^*, für das bedingte Poisson Design, neu evaluiert werden müssen.

Wie Matei & Tillé (2005a) darstellen, kann ein Stichprobendesign $p(\vec{s})$ mit festem Stichprobenumfang n und Inklusionswahrscheinlichkeiten π_k, das die Entropie[3]

$$-\sum_{\vec{s}\in\mathscr{S}} p(\vec{s}) \log p(\vec{s}) \tag{3.22}$$

maximiert, als ein bedingtes Poisson Design mit dem gleichen Stichprobenumfang angesehen werden. Es ist dann möglich, die Varianz von $\hat{t}_\pi$ für ein Design $p(.)$ mit maximaler Entropie, darzustellen als

$$\mathrm{V}\left(\hat{t}_\pi\right) = \mathrm{V}_{\text{poiss}}\left(\hat{t}_\pi|\mathfrak{n}=n\right) \ .$$

Dabei ist $\mathrm{V}_{\text{poiss}}$ die Varianz in Bezug auf das Poisson Design mit den gleichen Inklusionswahrscheinlichkeiten erster Ordnung wie $p(.)$ (Matei & Tillé, 2005a, S. 548).

Es wird davon ausgegangen, dass das Tupel $(\hat{t}_\pi, \mathfrak{n})$ unter dem Poisson Design normalverteilt ist (siehe Hájek, 1964 und Berger, 1998). Folglich kann die lineare Beziehung zwischen $\hat{t}_\pi$ und $\mathfrak{n}$ ausgenutzt werden, um folgende Aussage zu treffen

$$\mathrm{V}_{\text{poiss}}\left(\hat{t}_\pi|\mathfrak{n}=n\right) = \mathrm{V}_{\text{poiss}}\left(\hat{t}_\pi + (n-\mathfrak{n})\beta\right) \ , \tag{3.23}$$

wobei

$$\beta = \frac{\mathrm{COV}_{\text{poiss}}(\mathfrak{n}, \hat{t}_\pi)}{\mathrm{V}_{\text{poiss}}(\mathfrak{n})} \ ,$$

und

$$\mathrm{COV}_{\text{poiss}}(\mathfrak{n}, \hat{t}_\pi) = \sum_{k\in\mathscr{U}} \pi_k^*(1-\pi_k^*)\frac{y_k}{\pi_k} \ ,$$
$$\mathrm{V}_{\text{poiss}}(\mathfrak{n}) = \sum_{k\in\mathscr{U}} \pi_k^*(1-\pi_k^*) \ .$$

---

[3] Designs mit einer maximalen oder hohen Entropie sind Designs mit möglichst großem Träger in Bezug auf den Ziehungsrahmen sowie mit $p(.)$ möglichst gleich für alle Stichproben des Trägers. Das Design soll also möglichst viele unterschiedliche Stichproben haben und diese mit möglichst gleichen Wahrscheinlichkeiten ziehen.

Für $b_k = \pi_k^*(1-\pi_k^*)$ ergibt sich hieraus die folgende Approximation für die Varianz von $\mathrm{V}(\hat{\tau}_\pi)$

$$\mathrm{V}_{\text{approx}}(\hat{\tau}_\pi) = \sum_{k\in\mathscr{U}} b_k \varepsilon_k^2 \,, \tag{3.24}$$

mit

$$\begin{aligned}\varepsilon_k &= \frac{y_k}{\pi_k} - \beta \\ &= \frac{y_k}{\pi_k} - \frac{\sum_{k\in\mathscr{U}} b_k \frac{y_k}{\pi_k}}{\sum_{k\in\mathscr{U}} b_k}\end{aligned}$$

(Matei & Tillé, 2005a). Die b_k in (3.24), bzw. die π_k^*, sind nicht bekannt. Jedoch findet sich in der Literatur eine Vielzahl von Vorschlägen zu deren Approximation, wie z.B. bei Hájek (1964, S. 1508ff). Eine Übersicht und empirische Analyse der in der Literatur zu findenden Alternativen für b_k findet sich in Matei & Tillé (2005a) (siehe auch, Deville & Tillé, 2005).

Es lässt sich anmerken, dass die Varianzen in (3.23) bzw. (3.24), der approximativen Varianz eines Regressionsschätzers für τ unter Verwendung des Stichprobenindikators $\mathfrak{I}_k$ als Hilfsvariable bezüglich des Designs $p_{\text{poiss}}(.)$ entspricht (siehe Deville & Tillé, 2005 bzw. Särndal et al., 1992, S. 235).

3.3.1 Approximation von Hájek

Für Designs mit festem Stichprobenumfang n stellt Hájek (1964, S. 1511) folgende Beziehung zwischen π_k,π_l und $\pi_{k,l}$ auf

$$\pi_k\pi_l - \pi_{k,l} = d^{-1}\pi_k(1-\pi_k)\pi_l(1-\pi_l)[1+o(1)] \,, \tag{3.25}$$

wobei $d = \sum_{k\in\mathscr{U}} \pi_k(1-\pi_k)$ und $o(1) \to 0$ wenn $d \to \infty$. Die obige Beziehung ist gültig für *rejective sampling*, welche er als bedingtes Poisson Design bzw. bedingtes Ziehen mit Zurücklegen definiert, (Hájek, 1981, p. 66f.). Somit lässt sich schreiben:

$$\pi_{k,l} \approx \pi_k\pi_l\left(1-(1-\pi_k)(1-\pi_l)d^{-1}\right) \qquad 1 \le k \ne l \le N\,. \tag{3.26}$$

Durch das Einsetzen von (3.26) in (3.6) ergibt sich die folgende Approximation für $\mathrm{V}(\hat{\tau}_\pi)$

$$\mathrm{V}_{\text{Haj}}(\hat{\tau}_\pi) = \sum_{k\in\mathscr{U}} \pi_k(1-\pi_k)\left(\frac{y_k}{\pi_k} - B\right)^2 \,, \tag{3.27}$$

mit

$$B = \frac{\sum_{k\in\mathscr{U}} \frac{y_k}{\pi_k}\pi_k(1-\pi_k)}{\sum_{k\in\mathscr{U}} \pi_k(1-\pi_k)}\,.$$

Die Verwendung der Approximation in (3.26) genügt jedoch nicht der Bedingung in (3.19e). Aus diesem Grund schlägt Hájek (1981) vor, $\pi_k\pi_l - \pi_{k,l}$ durch das Produkt c_kc_l, mit $c_k = \pi_k(1-\lambda_k)d_\lambda^{-1/2}$ und $d_\lambda = \sum_{k=1}^N \pi_k(1-\lambda_k)$, darzustellen. Zudem ist

$$1-\lambda_k = (1-\pi_k)\left[1 - \frac{\pi_k(1-\lambda_k)}{d_\lambda}\right]^{-1} \tag{3.28}$$

(Hájek, 1981, S. 27).

Die Terme $(1-\lambda_k)$ sind unbekannt, können aber über eine Iteration der obigen Beziehung approximiert werden. Hierzu wird $1-\pi_k$ als Startwert für $1-\lambda_k$ verwendet und auf der rechten Seite von (3.28) eingesetzt. Die sich daraus ergebende erste Approximation von $1-\lambda_k$ wird dann wiederum in die rechte Seite von (3.28) eingesetzt. Dieser Vorgang wird wiederholt bis sich die resultierenden Approximationen für $(1-\lambda_k)$ stabilisieren, bzw. gegen $\pi_i(1-\pi_i)$ konvergieren (Hájek, 1981, S. 76). Wenn $(1-\lambda_k^*)$ das Ergebnis dieser iterativen Approximation von$(1-\lambda_k)$ ist, dann kann b_k in (3.24) durch $\pi_k(1-\lambda_k^*)$ ersetzt werden, was eine weitere Approximationsmöglichkeit der Varianz von $\hat{\tau}_\pi$ darstellt.

Als einen Kompromiss zwischen Einfachheit und Präzision schlägt Hájek (1981) für kleine π_k , bzw. hinreichend große N, den folgenden Wert

$$
\begin{aligned}
{}_{\text{HAJEK}}b_k &= \pi_k(1-\lambda_k) && (3.29)\\
&\approx \pi_k(1-\pi_k)\frac{n}{n-1} && (3.30)
\end{aligned}
$$

für b_k in (3.24) vor. Dabei ergibt sich die Approximation von $(1-\lambda_k)$ in (3.29) aus (3.28) unter der Annahme, dass $\pi_k(1-\pi_k)\approx\pi_k$ ist.

Falls $\pi_k=\frac{n}{N}\ \forall\, k\in\mathscr{U}$ gilt, ist (3.27) identisch zu (3.8), der Varianz von $\hat{\tau}_\pi$ unter SRS ist (Berger, 2003, S. 9). Zudem ist $B\approx\dfrac{\tau}{n}$ für kleine π_k in (3.27), und somit ist (3.27) gleich dem ersten Term in (3.21).

Des Weiteren zeigt Berger (1998), dass die Approximation in (3.26) ebenfalls für die Klasse von hoch randomisierten Designs bzw. Designs mit einer hohen Entropie (siehe auch, Berger, 2004a, S. 307) verwendet werden kann. Diese schließt das Sampford Design mit ein (Sampford, 1967).

3.3.2 Fixpunktiteration

Das Analogon von Beziehung (3.19e) der Approximation $\pi_k\pi_k-\pi_{k,l}\approx b_kb_l(\sum_{k\in\mathscr{U}}b_k)^{-1}$ ist

$$
\sum_{\substack{l\in\mathscr{U}\\ l\neq k}}\frac{b_kb_l}{\sum_{k\in\mathscr{U}}b_k}=b_k-\frac{b_k^2}{\sum_{k\in\mathscr{U}}b_k}\,.
$$

Folglich lässt sich die allgemeine Approximation in (3.24) auch schreiben als

$$
\begin{aligned}
\mathrm{V}_{\text{approx}}\left(\hat{\tau}_\pi\right) &= \frac{1}{2}\sum_{k\in\mathscr{U}}\sum_{\substack{l\in\mathscr{U}\\ l\neq k}}\frac{b_kb_l}{\sum_{k\in\mathscr{U}}b_k}\left(\frac{y_k}{\pi_k}-\frac{y_l}{\pi_l}\right)^2\\
&= \sum_{k\in\mathscr{U}}\frac{y_k^2}{\pi_k^2}\left(b_k-\frac{b_k^2}{\sum_{l\in\mathscr{U}}b_l}\right)-\frac{1}{\sum_{k\in\mathscr{U}}b_k}\sum_{k\in\mathscr{U}}\sum_{\substack{l\in\mathscr{U}\\ l\neq k}}\frac{y_ky_l}{\pi_k\pi_l}b_kb_l\,.
\end{aligned}
\qquad (3.31)
$$

Durch den Vergleich von (3.31) mit (3.3), schlagen Deville & Tillé (2005) vor, die genauest mögliche Approximation für die Varianz von $\hat{\tau}_\pi$ durch die Lösung des folgenden Gleichungssystems zu finden

$$b_k - \frac{b_k^2}{\sum_{l\in\mathscr{U}} b_l} = \pi_k(1-\pi_k) \qquad k = 1,\ldots,N\,. \tag{3.32}$$

Weil (3.32) ein nichtlineares Gleichungssystem ist, werden die b_k durch Iteration approximiert (Tillé, 2006, p. 139f.). Um (3.32) zu lösen wird von Deville & Tillé (2005) eine Fixpunktiteration angewendet durch die Verwendung der rekursiven Gleichung

$$b_k^{(i)} = \frac{\left[b_k^{(i-1)}\right]^2}{\sum_{k\in\mathscr{U}} b_k^{(i-1)}} + \pi_k\left(1-\pi_k\right) \qquad \text{für} \quad i = 0,1,2,3,\ldots \tag{3.33}$$

bis eine Konvergenz eintritt. Als Startwert wird $b_k^{(0)} = \pi_k(1-\pi_k)\dfrac{N}{N-1}$ verwendet. Um eine eindeutige Lösung für (3.33) zu finden, muss die folgende Bedingung erfüllt sein,

$$\max_{1\le k\le N} \frac{\pi_k(1-\pi_k)}{\sum_{l\in\mathscr{U}} \pi_l(1-\pi_l)} < \frac{1}{2}\,,$$

(Deville & Tillé, 2005, S. 575). Sollte der Prozess nicht konvergieren, schlägt Tillé (2006) die Verwendung des Wertes $b_i^{(1)}$ vor, der sich nach der ersten Iteration ergibt, d.h.

$$b_k^{(1)} = \pi_k(1-\pi_k)\left[\frac{N\pi_k(1-\pi_k)}{(N-1)\sum_{l\in\mathscr{U}} \pi_l(1-\pi_l)} + 1\right]\,.$$

Die obige Vorgehensweise ist sehr ähnlich zu der Iteration, die Hájek (1981) zur Approximation von $\pi_k(1-\lambda_k)$ anführt.

3.3.3 Brewer Approximation

Eine andere Klasse von approximativen Ausdrücken für $\mathrm{V}\left(\hat{\tau}_\pi\right)$, bei Designs mit hoher Entropie, stellen Brewer (2002) und Brewer & Donadio (2003) dar. Diese basiert auf einer Approximation für die $\pi_{k,l}$, welche von Hartley & Rao (1962) für eine randomisierte, systematische Stichprobenziehung mit ungleichen Inklusionswahrscheinlichkeiten hergeleitet wurde. Die Approximation für $\pi_{k,l}$ hat die folgende Form

$$\pi_{k,l} \approx \frac{1}{2}\pi_k\pi_l\left(b_k^* + b_l^*\right)\,, \tag{3.34}$$

mit

$$b_k^* = \frac{(n-1)}{n}\left(1 - n^{-2}\sum_{l\in\mathscr{U}} \pi_l^2 + 2\frac{\pi_k}{n}\right). \tag{3.35}$$

Aus (3.34) folgt, dass

$$\pi_k\pi_l - \pi_{k,l} \approx \frac{1}{2}\pi_k\pi_l\left(2 - b_k^* - b_l^*\right)\,. \tag{3.36}$$

Eine Approximation für die Varianz in (3.6) ergibt sich durch die Entwicklung des zweiten Terms in (3.21) mit (3.36), so dass

$$\sum_{k=1}^{N} \pi_k^2 (1 - b_k^*) \left(\frac{y_k}{\pi_k} - \frac{\tau}{n} \right)^2 .$$

Addieren des ersten Terms in (3.21) zum obigen ergibt die folgende Approximation

$$\mathrm{V}_{\mathrm{Brew}}(\hat{\tau}_\pi) = \sum_{k=1}^{N} \pi_k (1 - b_k^* \pi_k) \left(\frac{y_k}{\pi_k} - \frac{\tau}{n} \right)^2 \tag{3.37}$$

(Brewer, 2002, S. 151f). Durch Ersetzen der b_k^* in (3.37) mit (3.35) ergibt sich schließlich die eigentliche Approximation.

3.3.4 Schätzer für Varianzapproximationen

Nachdem im vorangegangen Abschnitt eine Übersicht zu möglichen Approximationen für die Varianz in (3.6) gegeben wurde, widmet sich der nächste Abschnitt der Schätzung dieser Approximationen.

Für die allgemeinen Varianzapproximation in (3.24) wird der folgende Schätzer formuliert

$$\hat{\mathrm{V}}_{\mathrm{approx}}(\hat{\tau}_\pi) = \sum_{k \in \mathcal{U}} \mathfrak{I}_k \hat{b}_k \hat{e}_k^2 , \tag{3.38}$$

wobei

$$\hat{e}_k = \frac{y_k}{\pi_k} - \hat{B} \tag{3.39}$$

$$\text{und} \quad \hat{B} = \frac{\sum_{k \in \mathcal{S}} \frac{y_k}{\pi_k} \hat{b}_k}{\sum_{k \in \mathcal{S}} \hat{b}_k} \pi_k . \tag{3.40}$$

In Abhängigkeit von der Wahl der $\hat{b}_k$ ergibt sich so eine Vielzahl von möglichen Schätzern der Form (3.38) (Matei & Tillé, 2005a, Kapitel 4). Als einen einfachen Wert für $\hat{b}_k$ kann

$${}_1\hat{b}_k = \frac{{}_{\mathrm{HAJEK}}b_k}{\pi_k} , \tag{3.41}$$

verwendet werden, wobei ${}_{\mathrm{HAJEK}}b_k$ wie in (3.30) approximiert wird. Für ein SRS Design entspricht dies dem Varianzschätzer in (3.8).

Eine von Deville (1999) entwickelte komplexere Wahl für $\hat{b}_k$ lautet

$${}_2\hat{b}_k = (1 - \pi_k) \left[1 - \sum_{l \in \mathcal{S}} \left(\frac{1 - \pi_l}{\sum_{l \in \mathcal{S}} (1 - \pi_l)} \right)^2 \right] . \tag{3.42}$$

Ebenso kann eine Fixpunktiteration verwendet werden, um einen Wert für $\hat{b}_k$ zu erhalten. Es wird der gleiche Algorithmus wie in Abschnitt (3.3.2) genommen, mit dem

Unterschied, dass die rechte Seite des Gleichungssystems in (3.32) sowie der zweite Term auf der rechten Seite von (3.33) mit π_k^{-1} multipliziert werden.

Der Startwert des Algorithmus lautet

$$ {}_3\hat{b}_k^{(0)} = (1-\pi_k)\frac{n}{n-1} \; . $$

Eine notwendige Bedingung für eine Lösung ist

$$ \max_{1 \le k \le n} \frac{(1-\pi_k)}{\sum_{l \in \mathcal{s}}(1-\pi_l)} < \frac{1}{2} \; , $$

(Tillé, 2006, S. 141f). Sollte der Prozess nicht konvergieren, schlägt Tillé (2006) die Verwendung des Wertes ${}_3\hat{b}_k^{(1)}$ vor, der nach der ersten Iteration

$$ {}_3\hat{b}_k^{(1)} = (1-\pi_k)\left[\frac{n(1-\pi_k)}{(n-1)\sum_{l \in \mathcal{s}}(1-\pi_l)} + 1\right] $$

vorliegt.

In Übereinstimmung mit der Varianzapproximation in (3.37) stellen Brewer (2002) und Brewer & Donadio (2003) einen Varianzschätzer vor, der unter SRS als unverzerrt angegeben wird. Der Varianzschätzer ist gegeben durch

$$ \hat{\mathrm{V}}_{\mathrm{Brew}}(\hat{\tau}_\pi) = \sum_{k=1}^{N} \mathfrak{I}_k \left(\frac{1}{b_k^*} - \pi_k\right)\left(\frac{y_k}{\pi_k} - \frac{\hat{\tau}_\pi}{n}\right)^2 \; , \tag{3.43} $$

wobei (3.43) der korrespondierende π-Schätzer zu der Summe in (3.37) ist, korrigiert um den Faktor b_k^{*-1} (Brewer & Donadio, 2003). Brewer (2002) stellt eine Auswahl von möglichen Werten für b_k^* vor. Zwei Vorschläge für b_k^* sind

$$ \begin{aligned} {}_1\hat{b}_k^* &= \frac{n-1}{n-\pi_k} \\ {}_2\hat{b}_k^* &= \frac{n-1}{n-n^{-1}\sum_{k \in \mathcal{U}} \pi_k^2} \; . \end{aligned} $$

Die Motivation für die Wahl von ${}_1\hat{b}_i^*$ ist gegeben durch die Relation zwischen Beziehung (3.19a) und (3.19b) und für ${}_2\hat{b}_i^*$ durch die Relation zwischen (3.19c) und (3.19d) (Brewer & Donadio, 2003).

3.4 Nichtlineare Statistiken

Im Rahmen von Stichprobenerhebungen besteht oftmals die Notwendig nicht lineare Schätzer zu verwenden. Leider führt dies dazu, dass in den meisten Fällen die Varianz eines nichtlinearen Schätzers nicht mehr in einer geschlossenen Form dargestellt werden kann. Des Weiteren existieren zudem meist auch keine unverzerrten Varianzschätzer. Die zwei häufigsten Ansätze, die zur Handhabung dieses Problems zur Anwendung kommen sind die sog. *Resampling* Verfahren und Methoden der Linearisierung.

Zu den Resampling Verfahren zählen die verschiedenen *Jackknife*- *Bootstrap*- oder auch *Balanced Repeated Replication* Methoden. Deren Anwendung hängt, mehr noch als vom verwendeten Punktschätzer selbst, vom vorliegenden Stichprobendesign ab. Eine Übersicht über die verschiedenen Resampling Verfahren und mögliche Modifikationen im Rahmen komplexer Stichprobedesigns findet sich beispielsweise in Wolter (2007) und Bruch, Münnich & Zins (2011, Kapitel 3).

Bei der Linearisierung wird ein nicht linearer Schätzer durch eine lineare Funktion approximiert. Auf diese Weise kann die Varianz geschätzt werden, wie dies in den Abschnitten 3.1 bzw. 3.2 und 3.3 dargestellt wurde. Dieser indirekte Ansatz über die Schätzung der approximativen Varianz eines Schätzers liefert keine unverzerrten Schätzungen. Im Idealfall sind die Ergebnisse aber konsistent (Wolter, 2007, Kapitel 6).

Ist θ eine Statistik mit $\theta = f(\vec{y})$, und $f : \mathbb{R}^N \mapsto \mathbb{R}$ und $\hat{\theta} = Q(\vec{\mathfrak{s}}, \vec{y})$ ein konsistenter Schätzer für θ, d.h. $\hat{\theta}_0 \xrightarrow{p} \theta_0$ für $n \to \infty$, so beschreiben Dell & d'Haultfoeuille (2008) θ als linearisierbar, wenn ein $\vec{z} = (z_k)_{k=1,\ldots,N}$ existiert, so dass

$$\left(\sqrt{\mathrm{V}\left(\sum_{k\in\mathscr{U}} w_k z_k\right)}\right)^{-1} \left(\hat{\theta} - \theta\right) \xrightarrow{d} NV(0,1) \quad \text{für } N \to \infty. \tag{3.44}$$

Dabei gilt die Annahme, dass gleichzeitig $\frac{n}{N} = o(\log(\log(N))^{-1})$, d.h. der Auswahlsatz gegen Null konvergiert, falls die Population unendlich große wird. $\sum_k^N w_k z_k = \hat{\tau}_z$, ist ein Totalwertschätzer der Werte z_k, im Folgenden auch als linearisierter Werte bezeichnet. Die in Schätzer $\hat{\tau}_z$ verwendeten Gewichte w_k werden als Erhebungsgewichte bezeichnet. Diese sind frei bestimmbar, mit

$$w_k = \begin{cases} w_k & \text{für } k \in \mathscr{s} \\ 0 & \text{sonst} \end{cases} , \tag{3.45}$$

jedoch soll $\hat{\tau}_z$ konsistent für $\tau_z = \sum_k^N z_k$ sein.

Zunächst soll der Fall einer linearen Funktion f untersucht werden, da es das Ziel ist, das Problem der Varianzschätzung für $\hat{\theta}$ auf den Fall eines linearen Schätzers zu reduzieren. Angenommen, es liegen d verschiedene Untersuchungsvariablen vor und die Statistik θ der Population $\mathscr{U}$ soll geschätzt werden. Dabei hat θ die folgende Form

$$\theta = f(\vec{\tau}) , \tag{3.46}$$

mit $\vec{\tau} = (\tau_1, \ldots, \tau_k, \ldots, \tau_d)^\top$ und $\tau_i = \sum_{k\in\mathscr{U}} y_{ik}$, wobei y_{ik} die Beobachtung des k-ten Elements der i-ten Untersuchungsvariable in $\mathscr{U}$ ist. Zur Schätzung von θ werden die unbekannten Statistiken in $\vec{\tau}$ durch $\hat{\vec{\tau}} = (\hat{\tau}_1, \ldots, \hat{\tau}_i, \ldots, \hat{\tau}_d)$ ersetzt,

$$\hat{\theta} = f\left(\hat{\vec{\tau}}\right) ,$$

mit $\hat{\tau}_i = \sum_{k\in\mathscr{s}} y_{ik} w_k$ als konsistenter Schätzer für den Totalwert der i-ten Untersuchungsvariable und w_k wie in (3.45).

Da f eine lineare Funktion ist, lässt sich $\hat{\theta}$ darstellen als

$$\hat{\theta} = f\left(\hat{\vec{\tau}}\right) = a_0 + \sum_{i=1}^{d} a_i \hat{\tau}_i ,$$

wobei $a_0, a_i \in \mathbb{R}$ mit $i = 1, \ldots, d$ als beliebige Konstanten. Dann gilt für die Varianz von $\hat{\theta}$

$$\mathrm{V}(\hat{\theta}) = f(\vec{\tau}) = \sum_{i=1}^{d} a_i^2 \mathrm{V}(\hat{\tau}_i) + 2 \sum_{i=1}^{d} \sum_{\substack{j=1 \\ i<j}}^{d} a_i a_j \operatorname{Cov}(\hat{\tau}_i, \hat{\tau}_j) \ . \tag{3.47}$$

Ausdruck (3.47) beinhaltet d Varianzen und $d(d-1)2^{-1}$ Kovarianzen, die geschätzt werden müssen. Falls $w_k = \pi_k^{-1}$, d.h. der π-Schätzer wird zur Schätzung von τ_i verwendet, kann $\mathrm{V}(\hat{\tau}_i)$, wie in Abschnitt 3.1 beschrieben, geschätzt werden. Die Kovarianz $\operatorname{Cov}(\hat{\tau}_i, \hat{\tau}_j)$ lässt sich schätzen durch

$$\widehat{\operatorname{Cov}}(\hat{\tau}_i, \hat{\tau}_j) = \sum_{k \in \mathscr{s}} \sum_{l \in \mathscr{s}} \frac{\pi_{k,l} - \pi_k \pi_l}{\pi_{k,l}} \frac{y_{ik}}{\pi_k} \frac{y_{jl}}{\pi_l} \ . \tag{3.48}$$

Für den Fall, dass f eine nicht lineare Funktion ist, soll auch eine Ausdruck für $V(\hat{\theta})$ in ähnlicher Form wie (3.47) gefunden werden. Die Idee ist, die Funktion f in der Umgebung ihres tatsächlichen Wertes durch eine bekannte lineare Funktion zu approximieren. Die hierzu verwendete Methode ist als Taylor-Linearisierung bekannt. Die Approximation wird durch eine Expansion mittels einer Taylorreihe erster Ordnung vorgenommen. Zum Theorem der Taylorreihe siehe beispielsweise Serfling (1980, p. 43f).

Ist die Funktion f stetig differenzierbar bis zweiter Ordnung an jedem Punkt einer offenen Menge, welche $\vec{\tau}$ und $\hat{\vec{\tau}}$ einschließt so, ist

$$\hat{\theta} - \theta = \sum_{i=1}^{d} \left[\frac{\partial f(p_1, \ldots, p_d)}{\partial p_i} \right]_{\mathbf{p} = \vec{\tau}} (\hat{\tau}_i - \tau_i) + R(\hat{\vec{\tau}}, \vec{\tau}) \ , \tag{3.49}$$

wobei

$$R(\hat{\vec{\tau}}, \vec{\tau}) = \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \left[\frac{\partial^2 f(p_1, \ldots, p_d)}{\partial p_i \partial p_j} \right]_{\mathbf{p} = \ddot{\vec{\tau}}} (\hat{\tau}_i - \tau_i)(\hat{\tau}_j - \tau_j) \ .$$

Dabei liegt $\ddot{\vec{\tau}}$ auf der Geraden $L(\vec{\tau}, \hat{\vec{\tau}})$, die $\vec{\tau}$ und $\hat{\vec{\tau}}$ verbindet . Gilt

$$\hat{\vec{\tau}} - \vec{\tau} = O_p(r_n) \tag{3.50}$$

wobei $(r)_n$ eine Folge von reellen Zahlen ist und $r_n \to 0$ für $n \to \infty$, dann folgt für den zweiten Term auf der rechten Seiten in (3.49), den sog. Rest $R(\hat{\vec{\tau}}, \vec{\tau}) = O_p(r_n^2)$. Dabei bedeutet Bedingung (3.50), dass $\hat{\vec{\tau}}$ konsistent für $\vec{\tau}$ ist.

In den meisten Anwendungen wird für genügend große Stichprobenumfänge davon ausgegangen, dass R einen zu vernachlässigenden Teil von $\hat{\theta} - \theta$ im Vergleich zu den linearen Termen in (3.49) darstellt. Dies rechtfertigt die Verwendung der folgenden Approximation:

$$\hat{\theta} - \theta \approx \sum_{i=1}^{d} \left[\frac{\partial f(p_1, \ldots, p_d)}{\partial \tau_i} \right]_{\mathbf{p} = \vec{\tau}} (\hat{\tau}_i - \tau_i) \ . \tag{3.51}$$

Es wird also nur der lineare Teil der Taylorreihe verwendet.

Wenn $\mathrm{E}\left(\hat{\vec{\tau}}\right) = \vec{\tau}$ gilt kann (3.51) dazu verwendet werden, eine Approximation für die mittlere quadratische Abweichung (MSE) von $\hat{\theta}$ abzuleiten

$$\begin{aligned} \mathrm{MSE}\left(\hat{\theta}\right) &\approx \mathrm{V}\left(\sum_{i=1}^{d}\left[\frac{\partial f(p_1,\dots,p_d)}{\partial p_i}\right]_{\mathbf{p}=\vec{\tau}}\hat{\tau}_i\right) \qquad (3.52) \\ &= \sum_{i=1}^{d} a_i^2 \mathrm{V}(\hat{\tau}_i) + 2\sum_{i=1}^{d}\sum_{\substack{j=1\\ i<j}}^{d} a_i a_j \,\mathrm{Cov}\left(\hat{\tau}_i,\hat{\tau}_j\right) \ , \end{aligned}$$

mit $a_i = [\frac{\partial f(p_1,\dots,p_d)}{\partial p_i}]_{\mathbf{p}=\vec{\tau}}$. Wegen $\mathrm{MSE}(\hat{\theta}) = \mathrm{V}(\hat{\theta}) + \mathrm{Bias}(\hat{\theta})^2$, wobei $\mathrm{Bias}(\hat{\theta}) = \hat{\theta} - \theta$, kann die Varianz von $\hat{\theta}$ durch $\mathrm{MSE}(\hat{\theta})$ approximiert werden. $\mathrm{V}(\hat{\theta})$ ist, für unverzerrte oder zumindest konsistente Schätzer, von höherer Ordnung als $\mathrm{Bias}(\hat{\theta})^2$ (Wolter, 2007).

Zur Schätzung von (3.52) sind die unbekannten Varianzen und Kovarianzen durch entsprechende Schätzer zu ersetzten. Für den Fall, dass d groß ist, kann dies allerdings impraktikabel sein. Hier kann die folgende von Woodruff (1971) vorgeschlagene Transformation von y_{ik} verwendet werden, mit

$$z_k = \sum_{i=1}^{d} a_i y_{ik} \ . \qquad (3.53)$$

Entsprechend lässt sich dann schreiben

$$\mathrm{V}\left(\hat{\theta}\right) \approx \mathrm{V}\left(\sum_{k\in\mathcal{s}} w_k z_k\right) \ .$$

Die Transformation in (3.53) ist anwendbar, wenn die einzelnen Schätzer in $\hat{\vec{\tau}}$ linear sind. Es ist dann möglich, die Rangfolge der Summierungen in (3.52) zu ändern (Andersson & Nordberg, 1994). Um $\mathrm{MSE}(\hat{\theta})$ zu schätzen, müssen dann die unbekannten a_i in (3.53) durch einen Schätzer $\hat{a}_i = [\frac{\partial f(p_1,\dots,p_d)}{\partial \tau_i}]_{\mathbf{p}=\hat{\vec{\tau}}}$ ersetzt werden. Eingesetzt in (3.53) ergibt sich ein Schätzwert $\hat{z}_k$ für z_k. Schließlich ist es möglich, die approximative Varianz von $\hat{\theta}$ durch $\widehat{\mathrm{V}}\left(\sum_{k\in\mathcal{s}} w_k \hat{z}_k\right)$ zu schätzen.

3.5 Einflussfunktion

Der in Abschnitt 3.4 dargestellte Ansatz zur Varianzschätzung ist nur anwendbar, falls der Schätzer als eine bis zweiter Ordnung stetig differenzierbare und asymptotisch normal verteilte Funktion linearer Schätzer darstellbar ist. Für Schätzer, die diesen Anforderungen nicht entsprechen, kann es dennoch möglich sein, eine lineare Approximation mit Hilfe der sog. Einflussfunktion zu finden, wie es von Hampel (1974) beschrieben wird. Dabei kann eine Einflussfunktion als ein Messinstrument für die asymptotische Verzerrung eines Schätzers verstanden werden, die durch eine Kontamination in den beobachteten Daten hervorgerufen wird. Daher findet das Konzept der Einflussfunktion vor allem im Bereich der robusten Statistik Anwendung.

In dem gegebenen Kontext ist es geboten, zunächst das Konzept *statistischer Funktionale*, im weiteren Verlauf Funktionale genannt, einzuführen. Wenn die interessierende Variable y ist eine reellwertige Zufallsvariable und F ihre Verteilungsfunktion ist, dann sei

$$F \in \mathscr{F} = \{F_\theta | \theta \in \Theta \subseteq \mathbb{R}^d\}$$

mit $d \in \mathbb{N}$ und Θ als ein d-dimensionaler Parameterraum. Parameter θ kann nun als (statistisches) Funktional T dargestellt werden, d.h. als eine Abbildung der Menge von Verteilungsfunktionen $\mathscr{F}$ nach $\mathbb{R}^d$. Dies lässt sich schreiben als

$$\theta = T(F)\ .$$

Ist F_n die empirische Verteilungsfunktion, bezüglich der beobachteten Werte $(y_k)_{k\in\mathscr{s}}$ von y aus der Stichprobe $\mathscr{s}$ mit dem Unfang n, kann ein Schätzer für θ definiert werden als $T(F_n)$ mit

$$F_n(y) = \frac{1}{n}\sum_{k\in\mathscr{s}} \mathbb{1}\,(y_k \leq y)\ \ y \in \mathbb{R}.$$

Ein Moment der Variable y kann somit dargestellt werden als Funktional $T(F) = \int_{\mathbb{R}} h(y)dF(y)$ mit h als integrierbarer Funktion. Der dazugehörige Schätzer ließe sich schreiben als $T(F_n) = \int h(y)dF_n(y) = n^{-1}\sum_{k\in\mathscr{s}} h(y_k)$, wobei $T(F_n)$ in diesem Fall auch Stichprobenmoment genannt wird (Shao, 2003, S. 338).

Für den Stichprobenvektor $(y_k)_{k\in\mathscr{s}}$, dessen Elemente nach der Verteilungsfunktion F unabhängig identisch verteilt sind, ist die Einflussfunktion IF einer Statistik $T = T(F)$ am Punkt y

$$IF(T,F,y) = \lim_{\varepsilon\to 0} \frac{T\left((1-\varepsilon)F + \varepsilon\delta_y^\bullet\right) - T(F)}{\varepsilon}, \tag{3.54}$$

Dabei ist $\delta_y^\bullet = \mathbb{1}_{[y,\infty)}$ die Dirac-Verteilung am Punkt $y \in \mathbb{R}$ (siehe, Shao, 2003, S. 339 und S. 19).

Zur Herleitung von Einflussfunktionen wird ein Differenzierungsbegriff für Funktionale benötigt. Es existieren hier verschiedene Arten von Differenzialen, wie das Gâteaux-Differential, das ρ-Hadamard-Differential und das ρ-Fréchet-Differential, wobei das Gâteaux-Differential hier im Vordergrund steht. Zu deren Definition siehe (Shao, 2003, S. 338f).

Im Folgenden sei: $T : \mathscr{F}_0 \to \mathbb{R}$ ein reellwertiges Funktional über der Menge aller absolut stetigen Verteilungsfunktionen $\mathscr{F}_0$ über $\mathbb{R}^d$. Des Weiteren sei
$\mathscr{D} := \{c(F-G)|F,G \in \mathscr{F}_0,\, c \in \mathbb{R}\}$.

$T : \mathscr{F}_0 \to \mathbb{R}$ ist Gâteaux-differenzierbar an der Stelle $F \in \mathscr{F}_0$, falls eine lineare Funktion L_F existiert, mit $L_F : \mathscr{D} \to \mathbb{R}$ (d.h. $L_F(c_1\Delta_1 + c_2\Delta_2) = c_1L_F(\Delta_1) + c_2L_F(\Delta_2)\ \forall\, \Delta_j \in \mathscr{D}, c_j \in \mathbb{R},\ j = 1,2$), so dass für alle $\Delta \in \mathscr{D}$ und $F + t\Delta \in \mathscr{F}_0$:

$$\lim_{t\to 0}\left(\frac{T(F+t\Delta) - T(F)}{t} - L_F(\Delta)\right) = 0.$$

Wird nun eine Funktion h definiert als $h : \mathbb{R} \to \mathbb{R} = T(F + t\Delta)$, dann ist Gâteaux-Differenzierbarkeit äquivalent zur Differenzierbarkeit der Funktion h an der Stelle $t = 0$, d.h. $L_F(\Delta) = h^{'}(0)$ (siehe, Shao, 2003, S. 339).

Die Einflussfunktion von $T(F)$ an der Stelle y ist gegeben durch $L_F(\delta_y^{\bullet} - F) = IF(T, F, y)$. Nun kann das asymptotische Verhalten von $T(F_n)$ aufgestellt werden. Falls T an der Stelle F Gâteaux-differenzierbar ist, dann ist für $t = \frac{1}{\sqrt{n}}$, und $\Delta = \sqrt{n}(F_n - F)$

$$\sqrt{n}(T(F_n) - T(F)) = L_F(\sqrt{n}(F_n - F)) + R_n \, , \tag{3.55}$$

mit R_n als einem stochastischen Rest. Für $n \to \infty$ gilt nach dem Zentralen Grenzwertsatz

$$L_F(\sqrt{n}(F_n - F)) = \frac{1}{\sqrt{n}} \sum_{k \in \mathscr{S}} IF(T, F, y_k) \xrightarrow{d} N(0, \sigma_F^2) \, ,$$

wenn $E\left(IF(T, F, y_k)\right) = 0$ und $\sigma_F^2 = E\left(IF(T, F, y_k)^2\right) < \infty$. Somit ist $T(F_n)$ asymptotisch normal verteilt, wenn $R_n = O_p(1)$, d.h. $R_n \xrightarrow{p} 0$, was nicht durch die Gâteaux-Differenzierbarkeit alleine gegeben ist. Hierzu wird zudem die ρ-Hadamard-Differenzierbarkeit oder ρ-Fréchet-Differenzierbarkeit benötigt. Dies bedeutet, dass zunächst L_F ermittelt wird durch das Differenzieren von $h(t) = T(F + t\Delta)$ an der Stelle $t = 0$. Dann wird überprüft ob T ρ-Hadamard' oder ρ-Fréchet-differenzierbar ist mit einer gegeben Distanz ρ über $\mathscr{F}_0$ (siehe, Shao, 2003, p.340).

In den meisten Anwendungsfällen ist jedoch für $T(G) = \int f(y) dF(G)$ mit $G \in \mathscr{F}$ das Funktional T linear und somit ρ-Fréchet-differenzierbar für beliebige ρ. Des Weiteren lässt sich für eindimensionale F und $F'(y) > 0$ für alle $y \in \mathbb{R}$ zeigen, dass die Quantilfunktion $T(G) = G^{-1}$ ρ_∞-Hadamard-differenzierbar ist an der Stelle F, für $F \in \mathscr{F}$ und

$$\rho_\infty = \|F_1 - F_2\|_\infty = \sup_{t \in \mathbb{R}^d} |F_1(t) - F_2(t)| \, .$$

Die Distanz ρ_∞ ist bestimmt durch die Supremumsnorm (siehe, Shao, 2003, S. 321 und S. 341 sowie Serfling, 1980, S. 216). Schließlich lässt sich mittels des Totalwerts der Werte der Einflussfunktion $z_k = IF(T, F, y_k)$ die Varianz des Schätzers $T(F_n)$ approximieren. In diesem Zusammenhang wird von den z_k auch als linearisieren Werten gesprochen.

3.6 Schätzgleichungen

Ein alternativer Ansatz zur Linearisierung nicht stetiger Funktionen ist die Verwendung von Schätzgleichungen. Schätzgleichungen können verwendet werden, um sowohl Punktschätzer als auch deren linearisierte Werte z_k zur Varianzschätzung zu bestimmen (Binder & Patak, 1994). Insbesondere verwenden Kovacevic & Binder (1997) und Binder & Kovacevic (1993) diese Methode zur Schätzung von Disparitätsmaßen und deren Varianzen. Im Folgenden wird zunächst ein kurzer Überblick über die Grundstruktur des Ansatzes gegeben.

Die interessierende Variable y sei wiederum eine Zufallsvariable mit Verteilungsfunktion F und differenzierbarer Wahrscheinlichkeitsfunktion f. Es wird davon ausgegangen, dass die interessierende Statistik bzw. der Parameter θ die Lösung θ_0 zu

$$U(\theta) = \int_{\mathbb{R}} u(y,\theta) dF(y) = 0 \tag{3.56}$$

ist (Binder & Patak, 1994). Weiter sei $u(y,\theta) = g'(y,\theta)$ definiert mit $g(y,\theta) = \log f(y,\theta)$. Dann ist

$$u(y,\theta) = \frac{\partial \log f(y,\theta)}{\partial \theta} \ ,$$

wobei u Schätzgleichung genannt wird. So ist zum Beispiel für $\theta = \mu_y = \int_{\mathbb{R}} y dF(y)$, dem Erwartungswert von y, die Schätzgleichung gegeben durch

$$u(y,\theta) = y - \theta \ ,$$

(Binder & Patak, 1994).

Für endliche Populationen ist θ eine Statistik von $\vec{y} = (y_1, \dots, y_N)^\top$ dem Parameter der endlichen Population. θ ist dann eine Lösung θ_0 zur Gleichung

$$U(\theta) = \sum_{k=1}^{N} u(y_k,\theta) = 0 \ . \tag{3.57}$$

Ein Schätzer $\hat{\theta}_0$ für θ_0 ist die Lösung zu Gleichung

$$\hat{U}(\theta) = \sum_{k \in \mathcal{S}} w_k u(y_k,\theta) = 0 \ ,$$

wobei w_k ein zu bestimmendes Erhebungsgewicht ist wie in (3.45) beschrieben. Die Varianz von $\hat{\theta}_0$ kann wie folgt geschätzt werden. Sei $\hat{U}(\hat{\theta}_0) = \sum_{k \in \mathcal{S}} u(y_k, \hat{\theta}_0) = 0$, dann ist

$$\begin{aligned} \hat{U}(\hat{\theta}_0) = & \sum_{k=1}^{N} \left(u(y_k,\hat{\theta}_0) - u(y_k,\theta_0) \right) \\ & + \sum_{k=1}^{N} w_k u(y_k,\theta_0) \\ & + \sum_{k=1}^{N} \left(u(y_k,\hat{\theta}_0) - u(y_k,\theta_0) \right) (w_k - 1) \ . \end{aligned}$$

Unter Verwendung der Taylorreihe ist zudem

$$\begin{aligned} \hat{U}(\hat{\theta}_0) = & \sum_{k=1}^{N} \left(\frac{\partial u(y_k,\theta)}{\partial \theta} \right)_{\theta=\theta_0} \left(\hat{\theta}_0 - \theta_0 \right) + R(\hat{\theta}_0,\theta_0) \\ & + \sum_{k=1}^{N} w_k u(y_k,\theta_0) \\ & + \sum_{k=1}^{N} \left(u(y_k,\hat{\theta}_0) - u(y_k,\theta_0) \right) (w_k - 1) \ . \end{aligned} \tag{3.58}$$

mit R wie in (3.49), wenn d=1. Wenn $\hat{\theta}_0$ ein konsistenter Schätzer für θ_0 ist, so ist der letzte Term in (3.58) vernachlässigbar für genügend große Stichprobenumfänge. Wird zudem auch R in (3.58) vernachlässigt ergibt sich die Approximation

$$\hat{\theta}_0 - \theta_0 \approx \sum_{k=1}^{N} w_k z_k \ , \tag{3.59}$$

mit

$$z_k = -\left[\left(\frac{\partial u(y_k, \theta)}{\partial \theta}\right)_{\theta=\theta_0}\right]^{-1} u(y_k, \theta_0) \ ,$$

(Kovacevic & Binder, 1997). Schließlich kann auch hier die Varianz von $\hat{\theta}_0$ approximiert werden, durch die Varianz des Totalwerts der linearisierten Werte z_k, da

$$\mathrm{V}(\hat{\theta}_0 - \theta_0) = \mathrm{V}(\hat{\theta}_0) \approx \mathrm{V}\left(\sum_{k=1}^{N} w_k z_k\right) \ . \tag{3.60}$$

3.7 Linearisierung von Armuts- und Disparitätsmaßen

In diesem Abschnitt sollen die Einflussfunktionen bzw. die linearisierten Werte für einige Armuts- und Disparitätsmaße hergeleitet werden. Hierzu wird der Ansatz von Deville (1999) vorgestellt, der Einflussfunktionen verwendet, die etwas von der Beschreibung in (3.54) abweichen. Diese Einflussfunktionen sind auf einem endlichen und diskreten Maße M für die Größe der Population definiert und nicht auf einer Verteilungsfunktion F. Die interessierende Statistik wird dabei dargestellt als Funktional von M, d.h. als $T(M)$. Als Schätzer für $T(M)$ wird $T(\hat{M})$ verwendet, einem Funktional von dem stochastischem Maß $\hat{M}$, welches sich aus den Erhebungsgewichten w_k ergibt und nahe an N liegt (Goga, Deville & Ruiz-Gazen, 2009). Die Einflussfunktion einer Statistik $T = T(M)$ am Punkt y wird dabei definiert als

$$IF\left(T, M, y\right) = \lim_{\varepsilon \to 0} \frac{T\left(M + \varepsilon \delta_k^\circ(y)\right) - T(M)}{\varepsilon} \ , \tag{3.61}$$

wobei $\delta_k^\circ(y)$ das Dirac-Maß am Punkt $y \in \mathbb{R}$ ist, mit

$$\delta_k^\circ(y) := \begin{cases} 1 & \text{falls } y = y_k \text{ und } k \in \mathcal{U} \\ 0 & \text{sonst} \end{cases}$$

und $M = \sum_{k \in \mathcal{U}} \delta_k^\circ(y_k)$.

Ein Schätzer für den Totalwert τ von $\vec{y}$ kann geschrieben werden als Schätzer $\hat{M}$ von M. Da $\tau = \sum_{k \in \mathcal{U}} y_k = \int y dM = \sum_{k \in \mathcal{U}} y_k \delta_k^\circ(y_k)$ ein Funktional von M ist. Ein naheliegender Schätzer für M wäre $\hat{M} = \sum_{k \in \mathcal{U}} w_k \delta_k^\circ(y_k)$ mit $\hat{M}$ als einem Maß, das die Werte w_k kumuliert (Deville, 1999).

Beispielsweise lässt sich das Verhältnis zwischen den Totalwerten von $\vec{y}$ und $\vec{x}$, mit $\vec{y}, \vec{x} \in \mathbb{R}^N$ darstellen als Funktional $T(M) = \frac{\int y dM}{\int x dM} = \frac{\tau_y}{\tau_x}$. Der Wert der Einflussfunktion von $T(M)$ an der Stelle $(y,x) \in \mathbb{R}^2$ ist gegeben durch

$$\begin{aligned} IF\,(T,M,(y,x)) &= \frac{1}{\tau_x} IF(\tau_y,M,y) + \tau_y IF(\frac{1}{\tau_x},M,x) \\ &= \frac{y}{\tau_x} - T\frac{x}{\tau_x}\,, \end{aligned}$$

(siehe Regel 2 Deville, 1999).

In der Praxis ist es oftmals nicht notwendig, die Einflussfunktion aufzustellen wie in (3.61) beschrieben, was möglicherweise komplexe Grenzwertbestimmungen mit sich bringen würde. Es ist möglich, Regeln zur Ableitung von Einflussfunktionen anzuwenden, wie sie auch zur Bestimmung der Ableitungen von differenzierbaren Funktionen verwendet werden (Deville, 1999, S. 197).

3.7.1 Armutsgefährdungsquote

Die Armutsgefährdungsquote (ARPR) ist ein Armutsmaß und definiert als der Anteil einer Population deren Wohlfahrtsvariable, üblicherweise eine Art von Einkommen, unterhalb einer definierten Armutsschranke liegt. Die Armutsschranke kann exogen sein, wird jedoch auch über die Verteilung der Variable bestimmt, bezüglich derer Armut gemessen werden soll. Hier soll Definition 1.1 aus Abschnitt 1.1 verwendet werden. Das heißt, die Armutsschranke ist als 60% des Medians der Einkommensvariable y definiert.

Im Folgenden wird nun eine approximative Varianz eines Schätzers der ARPR abgeleitet. Die folgenden Ausführungen lehnen sich an die Arbeit von Osier (2009) an, in welcher die Einflussfunktion für verschiedene nicht lineare Schätzer mit Hilfe der Ableitungsregel von Deville (1999) aufstellt werden.

Da die ARPR unmittelbar von der Armutsschranke (ARPT) abhängt, wird zunächst die Einflussfunktion der ARPT bestimmt. Hierzu wird die Armutsschranke geschrieben als

$$\mathrm{ARPT} = 0{,}6\,\mathrm{MED}(M) = T(M)\,,$$

mit dem Funktional $\mathrm{MED}(M)$ als Median von $\vec{y}$. Des Weiteren soll $F_N(M,y_0)$ die empirische Verteilungsfunktion von $\vec{y}$ an der Stelle y_0 bezeichnen, mit

$$F_N(M,y_0) = \frac{\int \mathbb{1}[y \leq y_0]dM}{\int dM} \tag{3.62}$$

$$= \frac{\sum_{k\in\mathscr{U}} \mathbb{1}[y_k \leq y]}{N}\,. \tag{3.63}$$

Somit ist $F_N(M,\mathrm{MED}(M)) = 0{,}5$. Folglich sind die Werte der Einflussfunktion $IF(F_N[M,\mathrm{MED}(M)],M,y)$ gleich Null für alle $y = y_k$ mit $k \in \mathscr{U}$.

Für ein Funktional der Form $F_N(M,\mathrm{MED}(M))$ kann zur Bestimmung seiner Einflussfunktion *Regel 7* von Deville (1999) verwendet werden. Diese besagt, dass wenn $S(M)$ ein

Funktional in $\mathbb{R}^d$ und $T(M,\lambda)$ eine Klasse von Funktionalen mit $\lambda \in \mathbb{R}^d$ ist, die Einflussfunktion von $T_S = T(M,S(M))$ gegeben ist durch

$$IF(T_S) = IF(T(\lambda,M),M,y_k|\lambda = S(M)) + \left(\frac{\partial T(\lambda,M)}{\partial \lambda}\right)_{\lambda=S(M)} IF(S(M),M,y_k)\ . \tag{3.64}$$

Aus (3.64) folgt

$$\begin{aligned} 0 =& IF(F_N(M,y),M,y_k|y = \mathrm{MED}(M)) \\ &+ \left[\left(\frac{\partial F_N(M,y)}{\partial y}\right)_{y=\mathrm{MED}(M)}\right] IF(\mathrm{MED}(M),M,y_k)\ . \end{aligned} \tag{3.65}$$

Der erste Term in (3.65) ist die Einflussfunktion von $F_N(M,MED(M))$ mit $MED(M)$ als Konstante. Der zweite Term bezieht sich auf den Einfluss von $\mathrm{MED}(M)$. In (3.62) wird der Wert der Verteilungsfunktion an einer bestimmten Stelle als der Quotient zweier Totalwerte dargestellt. Demnach ist der erste Termin in (3.65) gegeben durch

$$IF(F_N(M,y),M,y_k|y = \mathrm{MED}(M)) = \frac{1}{N}\left(\mathbb{1}[y_k \leq \mathrm{MED}(M)] - 0{,}5\right)\ .$$

Daher kann die Einflussfunktion des Median dargestellt werden als

$$\begin{aligned} IF(\mathrm{MED},M,y_k) &= IF\left(\mathrm{MED}(M),M,y_k\right) \\ &= -\frac{1}{NF_N'[\mathrm{MED}(M)]}\left(\mathbb{1}[y_k \leq \mathrm{MED}(M)] - 0{,}5\right)\ , \end{aligned} \tag{3.66}$$

mit F_N' als der ersten Ableitung von F_N nach y. Unter Verwendung von (3.66) kann die Einflussfunktion der ARPT angegeben werden als

$$\begin{aligned} IF(\mathrm{ARPT},M,y_k) &= IF\left(0{,}6\mathrm{MED}(M),M,y_k\right) \\ &= -\frac{0{,}6}{NF_N'(\mathrm{MED}(M))}\left(\mathbb{1}[y_k \leq \mathrm{MED}(M)] - 0{,}5\right)\ . \end{aligned} \tag{3.67}$$

Nach der Herleitung der Einflussfunktion der ARPT kann nun in einem letzten Schritt auch die Einflussfunktion für die ARPR hergeleitet werden. Die ARPR ist der Anteil der Elemente in der Population für die $y_k \leq \mathrm{ARPT}$ ist. Folglich ist die ARPR gegeben durch das Funktional

$$\mathrm{ARPR} = F_N\left(M,\mathrm{ARPT}(M)\right)\ .$$

Durch die erneute Anwendung von Regel (3.64) ergibt sich

$$\begin{aligned} IF(\mathrm{ARPR},M,y_k) =& IF\left(F_N(M,\mathrm{ARPT}(M)),M,y_k\right) \\ =& IF\left(F_N(M,y),M,y_k|y = \mathrm{ARPT(M)}\right) \\ &+ \left[\frac{\partial F_N(M,y)}{\partial y}|_{y=\mathrm{ARPT}(M)}\right] IF(\mathrm{ARPT},M,y_k)\ . \end{aligned} \tag{3.68}$$

Der erste Term in (3.68) ist gegeben durch

$$IF\left(F_N(M,y),M,y_k|y = \mathrm{ARPT(M)}\right) = \frac{\mathbb{1}[y_k \leq \mathrm{ARPT}(M)] - \mathrm{ARPR}(M)}{N}$$

und der zweite Term ist durch

$$F_N'(\mathrm{ARPT}(M))IF(\mathrm{ARPT},M,y_k)$$

mit $IF(\mathrm{ARPT},M,y_k)$ wie in (3.67) gegeben.

Wenn $F_N'(y) > 0$ für alle $y \in \mathbb{R}$ ist, lässt sich die Einflussfunktion der ARPR schließlich schreiben als

$$\begin{aligned} IF(\mathrm{ARPR},M,y_k) =& \frac{1}{N}\left(\mathbb{1}[y_k \leq \mathrm{ARPT}(M)] - \mathrm{ARPR}(M)\right) \\ & - \frac{0,6F_N'[\mathrm{ARPT}(M)]}{F_N'[MED(M)]}\left(\frac{\mathbb{1}[y_k \leq \mathrm{MED}(M)] - 0{,}5}{N}\right) . \end{aligned}$$

3.7.2 Quintile Share Ratio

Die *Quintile Share Ratio* (QSR) stellt ein einfach konstruiertes und aus diesem Grund auch leicht zu verstehendes Disparitätsmaß dar. Für ein endliche Population ist die QSR definiert als das Verhältnis zwischen dem Totalwert der 20% höchsten Werte in $\vec{y}$ und dem Totalwert der 20% kleinsten Werte in $\vec{y}$. Ist die Untersuchungsvariable ein Einkommen, lässt sich die QSR einfach interpretieren als das Verhältnis des Einkommens der 20% reichsten zu dem der 20% ärmsten Personen in der Population. Somit kann die QSR wie folgt festgelegt werden

$$\begin{aligned} \mathrm{QSR} &= \frac{\int y dM - \int y\mathbb{1}[y \leq q_{0,8}]dM}{\int y\mathbb{1}\left(y \leq q_{0,2}\right)dM} && (3.69) \\ &= \frac{\sum_{k\in\mathscr{U}}\left(y_k - y_k\mathbb{1}\left(y_k \leq q_{0.8}\right)\right)}{\sum_{k\in\mathscr{U}} y_k\mathbb{1}[y_k \leq q_{0,2}]} , && (3.70) \end{aligned}$$

mit q_p als dem p%-Quantil definiert als

$$\begin{aligned} q_p = F_N^{-1}(p) : [0,1] &\mapsto \mathbb{R} \quad \text{mit} && (3.71) \\ F_N^{-1}(p) := \inf\{y \in \mathbb{R}&|F(y) \geq p\} , && (3.72) \end{aligned}$$

wobei F_N^{-1} auch Quantilfunktion genannt wird.

Alternativ zu der Darstellung in (3.69) kann die QSR auch als das Verhältnis zwischen dem Mittelwert des Einkommens der reichsten 20%, (μ_r), zu dem der ärmsten 20%, (μ_a), definiert werden, mit

$$\begin{aligned} \mu_\mathrm{r} &= \frac{\sum_{l\in\mathscr{U}}\left(y_l - y_l\mathbb{1}[y_l \leq q_{0,8}]\right)}{\sum_{k\in\mathscr{U}}(1-0{,}8)} \quad \text{und} \\ \mu_\mathrm{a} &= \frac{\sum_{l\in\mathscr{U}} y_l\mathbb{1}[y_l \leq q_{0,2}]}{\sum_{k\in\mathscr{U}} 0{,}2} , \end{aligned}$$

(Hulliger & Münnich, 2006). Diese lässt sich wiederum schreiben als Funktion von vier Totalwerten, d.h.

$$\mathrm{QSR} = \frac{\mu_\mathrm{r}}{\mu_\mathrm{a}} = \frac{\tau_1}{\tau_2}\Big/\frac{\tau_3}{\tau_4} , \qquad (3.73)$$

dabei sind τ_1 und τ_3 die Zähler und τ_2 sowie τ_4 die Nenner in μ_r bzw. in μ_a.

Es ist anzumerken, dass die Definitionen der QSR in (3.69) und (3.73) nicht identisch sind. In (3.73) wird davon ausgegangen, dass $\tau_2 = \tau_4 = N0,2$. Dies stellt eine Vereinfachung gegenüber (3.69) dar, übernommen aus der Betrachtung einer unendlichen Population. Es ist aber davon auszugehen, dass (3.69) und (3.73) für genügend große N approximativ gleich sind.

Die Einflussfunktion für die QSR kann über die Ermittlung der Einflussfunktion für jeden der vier Totalwerte in (3.73) aufgestellt werden. Da $\tau_2 = \tau_4$ ist, lassen sich beide Totalwerte durch das Funktional $T(M) = \int 0.2dM$ darstellen. Folglich ist $IF(\tau_2, M, y_k) = IF(\tau_4, M, y_k) = 0.2, \forall\, k \in \mathscr{U}$.

Der Totalwert τ_1 ist gegeben mit

$$\tau_1 = \int ydM - \int y\mathbb{1}\,(y \leq q_{0,8})\,dM\,.$$

Die Einflussfunktion des ersten Terms von τ_1 ist gegeben durch $IF(\tau_y, M, y_k) = y_k$. Die des zweiten Terms kann bestimmt werden unter Verwendung von Regel (3.64). Demnach ist für $Y(M,q) = \int y\mathbb{1}\,[y_k \leq q]\,dM$

$$\begin{aligned} IF(Y(M,q),M,y_k) =& IF\,(Y(M,q),M,y_k|q = q_p(M)) \\ &+ \left[\frac{\partial Y(M,q)}{\partial q}|_{q=q_p(M)}\right] IF(q_p(M),M,y_k)\,. \end{aligned} \tag{3.74}$$

wobei das Funktional $q_p(M)$ dem p%-Quantil entspricht.

In Analogie zu (3.66) folgt, dass

$$IF(q_p(M),M,y_k) = -\frac{1}{NF'_N[q_p(M)]}\,(\mathbb{1}[y_k \leq q_p(M)] - p) \tag{3.75}$$

mit F_N wie in (3.62). Für den ersten Term in (3.74) zeigt Osier (2009, S. 184f), dass

$$IF\,(Y(M,q),M,y_k|q = q_p(M)) = y_k\mathbb{1}[y_k \leq q_p(M)]\,. \tag{3.76}$$

Durch das Einsetzen von (3.75) und (3.76) in (3.74) ergibt sich die Einflussfunktion von $Y(M,q)$. Diese Darstellung der Einflussfunktion verlangt jedoch die Ableitung zweier nicht stetiger Treppenfunktionen F_N und Y. Um diese zu vermeiden, geben Langel & Tillé (2011) eine alternative Form für die Einflussfunktion von $Y(M,q)$, mit

$$IF(Y(M,q_p),M,y_k) = (y_k - q_p)\mathbb{1}[y_k \leq q_p] + pq_p\,, \tag{3.77}$$

an. Siehe hierzu auch Hulliger & Münnich (2006) und Hulliger & Schoch (2014).

Die Einflussfunktionen der vier Totalwerte in (3.73) lassen sich dann wie folgt darstellen

$$\begin{aligned} &IF(\tau_1,M,y_k) = y_k - \left((y_k - q_{0,8})\mathbb{1}\,[y_k \leq q_{0,8}] + 0.8q_{0,8}\right)\,, \\ &IF(\tau_2,M,y_k) = 0{,}2\,, \\ &IF(\tau_3,M,y_k) = (y_k - q_{0,2})\mathbb{1}\,[y_k \leq q_{0,2}] + 0{,}2q_{0,2}\,, \\ &IF(\tau_4,M,y_k) = 0{,}2\,. \end{aligned}$$

Somit lassen sich die Einflussfunktionen von μ_r und μ_a darstellen als

$$IF(\mu_\mathrm{r}, M, y_k) = (IF(\tau_1, M, y_k) - \mu_\mathrm{r}\, 0,2)\, \frac{1}{\tau_2}\,,$$
$$IF(\mu_\mathrm{p}, M, y_k) = (IF(\tau_3, M, y_k) - \mu_\mathrm{a}\, 0,2)\, \frac{1}{\tau_4}\,.$$

Schließlich ist

$$IF(\mathrm{QSR}, M, y_k) = (IF(\mu_\mathrm{r}, M, y_k) - \mathrm{QSR}\, IF(\mu_\mathrm{a}, M, y_k))\, \frac{\tau_4}{\tau_3}\,.$$

3.7.3 Gini-Koeffizient

Der Gini-Koeffizient (GINI) ist ein Disparitätsmaß für die Verteilung einer Variable. Als solches kann er über die Verteilungsfunktion der betreffenden Variable definiert werden. Ist die Untersuchungsvariable y strikt nicht negativ und F ihre Verteilungsfunktion, dann ist der GINI definiert als

$$GINI = \frac{1}{2\mu_y} \int\limits_0^\infty \int\limits_0^\infty |x-y| dF(y) dF(x)\,,$$

mit $\mu_y = \int\limits_0^\infty y dF(y)$ als Erwartungswert von y. Für nicht strikt nicht negative Untersuchungsvariablen wird üblicherweise nicht der GINI als Disparitätsmaß verwendet. Entsprechend der obigen Darstellung steht der GINI unmittelbar im Verhältnis zum erwarteten Abstand zwischen zwei unabhängigen Ziehungen aus der Verteilung (bzw. Population) der Untersuchungsvariable (Binder & Kovacevic, 1993).

In Anlehnung an die Ausführungen von Binder & Kovacevic (1993) beruht der im Folgenden vorgestellte Ansatz zur Linearisierung des GINI auf der Anwendung einer Schätzgleichung. Hierzu wird zunächst eine alternative Darstellung des GINI verwendet,

$$\mathrm{GINI} = \frac{1}{\mu_y} \int\limits_0^\infty [2F(y) - 1)\, y dF(y)\,. \tag{3.78}$$

Als Schätzfunktion $u(y, \mathrm{GINI})$ für den GINI kann folglich

$$u(y, \mathrm{GINI}) = (2F(y) - 1)\, y - \mathrm{GINI}\, y$$

verwendet werden. Sei nun J eine Funktion in $\mathbb{R}$ mit $J(p) = 2p - 1$, so lässt sich schreiben

$$\hat{U}(\widehat{\mathrm{GINI}}) = \int\limits_0^\infty \left(J\left[\hat{F}(y)\right] y - \mathrm{GINI} y\right) d\hat{F}\,, \tag{3.79}$$

mit $\hat{F}$ als einem Schätzer für die Verteilungsfunktion F.

Entsprechend der Darstellung in (3.58) kann (3.79) approximiert durch

$$\begin{aligned}&\int_0^\infty \left(J\left[\hat{F}(y)\right] - J\left[F(y)\right]\right) y dF(y) - (\widehat{\text{GINI}} - \text{GINI}) \int_0^\infty y dF(y) \\ &\quad + \int_0^\infty \left(J\left[\hat{F}(y)\right] y - \text{GINI} y\right) d\hat{F}(y)\ ,\end{aligned} \tag{3.80}$$

wozu die Entsprechung des dritten Terms in (3.58) vernachlässigt wird. Der erste Term in (3.80) lässt sich dann schreiben als

$$\begin{aligned}\int_0^\infty \left(J\left[\hat{F}(y)\right] - J\left[F(y)\right]\right) y dF(y) &\approx \int_0^\infty \left(\hat{F}(y) - F(y)\right) J'\left[F(y)\right] y dF(y) \\ &= \int_0^\infty \hat{F}(y) J'\left[F(y)\right] y dF(y) - \mathrm{E}\left(F(y) J'\left[F(y)\right] y\right)\end{aligned}$$

und

$$\begin{aligned}\int_0^\infty \hat{F}(y) J'\left[F(y)\right] y dF(y) &= \int_0^\infty \int_0^y J'\left[F(y)\right] y d\hat{F}(x) dF(y) \\ &= \int_0^\infty \left[\int_y^\infty J'\left[F(x)\right] x dF(x)\right] d\hat{F}(y)\ .\end{aligned}$$

Durch das Verschieben der Integrationsgrenzen des inneren Integrals lässt sich dieses schreiben als $\int_y^\infty J'\left[F(x)\right] x dF(x) = \int_{F(y)}^1 J'\left[p\right] F^{-1}(p) dp$. Dabei ist F^{-1} die Quantilfunktion bezüglich F wie in (3.71) dargestellt, jedoch beschränkt auf den Wertebereich $\mathbb{R}^+$. So zeigt sich, dass

$$\widehat{\text{GINI}} - \text{GINI} \approx \int_0^\infty z(y) d\hat{F}(y)\ . \tag{3.81}$$

Dabei ist $z(y)$

$$z(y) = \frac{1}{\int_0^\infty y dF(y)} \left(\int_{F(y)}^1 J'\left[p\right] F^{-1}(p) dp + J\left[F(y)\right] y - \text{GINI} y - \mathrm{E}\left[F(y) J'\left[F(y)\right] y\right]\right)\ .$$

Für den Fall einer endlichen Population kann (3.79) formuliert werden als

$$\hat{U}(\widehat{\text{GINI}}) = \frac{1}{\hat{N}\hat{\mu}_y} \sum_{k \in \mathcal{s}} w_k \left(2\hat{F}_N(y_k) - 1\right) y_k\ , \tag{3.82}$$

mit

$$\hat{F}_N(y) = \frac{\sum_{k \in \mathcal{s}} w_k \mathbb{1}\left[y_k \leq y\right]}{\sum_{k \in \mathcal{s}} w_k} \qquad \text{und} \qquad \hat{\mu}_y = \frac{\sum_{k \in \mathcal{s}} w_k y_k}{\sum_{k \in \mathcal{s}} w_k}\ . \tag{3.83}$$

Zudem ist

$$\int_{F(x)}^{1} F^{-1}(y)dy = \sum_{i=F_N(x)}^{N} y_i/N = \sum_{i=1}^{N} \mathbb{1}(y_i \geq x)x/N \, .$$

Das diskrete Analogon zu $z(y)$ an der Stelle y_k ist gegeben durch

$$z(y_k) = \frac{1}{\overline{y}} \left[\sum_{i=1}^{N} 2 \frac{\mathbb{1}(y_i \geq y_k)y_k}{N} + (2F_N(y_k) - 1 - \mathrm{GINI})y_k - \frac{2\sum_{k=1}^{N} F_N(y_k)y_k}{N} \right] = z_k^* \, ,$$

mit $\mathrm{GINI} = \frac{1}{N\overline{y}} \sum_{k=1}^{N} (2F_N(y_k) - 1)\, y_k$ und $\overline{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$. Folglich ist

$$\frac{2\sum_{k}^{N} F_N(y_k)y_k}{N} = \overline{y}(\mathrm{GINI} + 1) \, ,$$

was zu

$$z_k^* = \frac{2}{\overline{y}} \left[\sum_{i=1}^{N} \frac{\mathbb{1}[y_i \geq y_k]y_k}{N} + \left(F_N(y_k) - \frac{\mathrm{GINI} + 1}{2} \right) y_k - \frac{\overline{y}}{2}(\mathrm{GINI} + 1) \right] \, ,$$

führt.

Die approximative Varianz des Schätzers in (3.79) entspricht der Varianz des Ausdrucks auf der rechten Seite von (3.81), der sich im diskreten Fall als Mittelwert der z_k^* darstellt. Somit ist $z_k = z_k^*/N$ der linearisierte Wert für den Schätzer des GINI in (3.82). Das gleiche Resultat findet sich in Kovacevic & Binder (1997), jedoch unter Verwendung eines allgemeineren Ansatzes.

3.8 Varianzschätzung für Armuts- und Disparitätsmaße

Nach der Herleitung der linearisierten Werte z_k für die ARPR, die QSR, und den GINI in Abschnitt 3.7 gilt der folgende Abschnitt der Schätzung der approximativen Varianz dieser Statistiken.

Die asymptotische Varianz eines Schätzers $\hat{\theta}$ für θ, dessen linearisierte Werte bestimmbar sind, ist gleich der Varianz des Totalwertschätzers dieser linearisierten Werte

$$\mathrm{V}\left(\sum_{k \in \mathscr{U}} \mathfrak{I}_k w_k z_k \right) \, . \qquad (3.84)$$

Für den Fall $w_k = \pi_k^{-1} \; \forall \, k \in \mathscr{U}$ wurde in Abschnitt 3.3 gezeigt, wie (3.84) bei komplexen Designs approximiert werden kann.

Da die Werte z_k selbst von θ abhängen, sind diese unbekannt. Deswegen werden zur Schätzung von (3.84) zudem auch Schätzer $\hat{z}_k$ für die linearisierte Werte benötigt. Mit dem Ersetzen der unbekannten Statistiken in z_k durch Schätzer, sowie θ durch $\hat{\theta}$, kann ein konsistenter Schätzer für $\mathrm{V}\left(\hat{\theta}\right)$ aufgestellt werden (Dell & d'Haultfoeuille, 2008).

Tabelle 3.1 gibt hier einen Überblick über die linearisierten Werte der in Abschnitt 3.7 behandelten Statistiken und deren Schätzer. Für manche Schätzer wird $\hat{F}_N'$ benötigt. Dabei handelt es sich um die Ableitung des Schätzers für die empirische Verteilungsfunktion $\hat{F}_N$ nach (3.83). Für den gesamten Definitionsbereich von $\hat{F}_N$ ist $\hat{F}_N'$ jedoch 0 oder nicht definiert. Dieses Problem kann durch die Verwendung von $\widetilde{\hat{F}}_N'$, der Ableitung von $\widetilde{\hat{F}}_N$, eines stetig differenzierbaren Schätzers für die empirische Verteilungsfunktion F_N umgangen werden. Hierzu können Kerndichteschätzer verwendet werden. Für die Wahl des Gaußkerns gilt für alle $y \in \mathbb{R}$:

$$\widetilde{\hat{F}}_N'(y) = \frac{1}{\hat{N}h\sqrt{2\pi}} \sum_{k \in \mathcal{U}} \mathfrak{I}_k w_k \exp\left[-\frac{(y-y_k)^2}{2h^2}\right] , \tag{3.85}$$

mit $\widetilde{\hat{F}}_n{}'(y) > 0$ für alle $y \in \mathbb{R}$ und $\hat{N} = \sum_{k \in \mathcal{S}} w_k$. Die Wahl der Bandbreite $h > 0$ in (3.85) ist entscheidend für den Fehler von $\widetilde{\hat{F}}_N'(y)$ gemessen an dem mittleren integrierten quadratischen Fehler

$$\mathrm{E}\left(\int \left(\widetilde{\hat{F}}_N'(y) - F_N'(y)\right)^2 dy\right) .$$

Die Wahl eines geeigneten Kerns und einer geeigneten Bandbreite hängt von den Stichprobendaten ab. Oftmals kann hier eine Analyse der Daten mit graphischen Methoden hilfreich sein, um eine Entscheidung zu treffen. Es findet sich auch eine Vielzahl von datenbasierten Ansätzen zur Bestimmung einer Bandbreite in der Literatur (siehe z.B. Jones, Marron & Sheather, 1996 für einen Überblick zu einigen dieser Methoden).

3.9 Kalibrierungsgewichte

Ein weitere Wahl für der Erhebungsgewichte w_k in (3.45) sind sog. Kalibrierungsgewichte. Ein allgemeiner Ansatz zur Kalibrierung der Designgewichte π_k^{-1} auf exogene, d.h. nicht durch die Stichprobe erhobene Daten wurde von Deville & Särndal (1992) und Deville, Särndal & Sautory (1993) eingeführt. Aufgrund seiner breiten Anwendbarkeit in vielen Problemfeldern einer Stichprobenerhebung[4] hat die Bereitstellung und Verwendung von Kalibrierungsgewichten eine weite Verbreitung erfahren.

Die Gewichte w_k werden dabei so gewählt, dass sie die Kalibrierungsgleichung

$$\sum_{k \in \mathcal{S}} w_k \vec{x}_k = \sum_{k \in \mathcal{U}} \vec{x}_k . \tag{3.86}$$

erfüllen, dabei ist $\vec{x}_k \in \mathbb{R}^d$ ein Spaltenvektor von Hilfsvariablen dessen Totalwert $\vec{\tau}_x = \sum_{k \in \mathcal{U}} \vec{x}_k$ bekannt ist. Ist $\hat{\tau}_{\mathrm{cal}}$, ein Schätzer für τ unter Verwendung der Kalibrierungsgewichte, dann lässt sich dieser schreiben als

$$\begin{aligned} \hat{\tau}_{\mathrm{cal}} &= \sum_{k \in \mathcal{U}} \mathfrak{I}_k w_k y_k \\ &= \hat{\tau}_\pi + \sum_{k \in \mathcal{U}} \mathfrak{I}_k (w_k - \pi_k^{-1}) y_k . \end{aligned} \tag{3.87}$$

---

[4] Z.B. um Verzerrungen, hervorgerufen durch Non-Response oder Fehler im Ziehungsrahmen, verringern zu können oder die Varianz von Schätzern zu senken.

Durch die Darstellung von $\hat{\tau}_{\text{cal}}$ als π-Schätzer plus einem Rest wird klar, dass $\hat{\tau}_{\text{cal}}$ nur approximativ unverzerrt sein kann, wenn $\mathrm{E}\left(\sum_{k\in\mathfrak{s}}(w_k - \pi_k^{-1})y_k\right) \approx 0$, (Särndal, 2007). Es ist somit wünschenswert, dass die Kalibrierung den Abstand zwischen w_k und π_k^{-1} unter der Bedingung (3.86) minimiert.

Für ein geeignetes Distanzmaß $G_k(w,\pi,q)$ (siehe hierzu Särndal, 2007) zwischen w_k und π_k^{-1} stellt sich das Kalibrierungsproblem als die Minimierung des Erwartungswerts

$$\mathrm{E}\left(\sum_{k\in\mathscr{U}} \mathfrak{I}_k G_k(w_k,\pi_k,q_k)\right) \tag{3.88}$$

unter Nebenbedingung (3.86) dar. Der Faktor q_k ist ein frei wählbares positives und von π_k unabhängiges Gewicht für Element k und erlaubt ein gewisses Maß an Flexibilität bei der Verwendung der obigen Metrik. Für ein gegebenes Design $p(.)$ und $\vec{\mathfrak{s}} = \vec{s}$ stellt sich das Kalibrierungsproblem folglich als die Minimierung von

$$\sum_{k\in\mathscr{U}} I_k G_k(w_k,\pi_k,q_k)$$

dar (Deville & Särndal, 1992).

Wird als Distanzmaß eine der Chi-Quadrat-Statistik ähnlichen Metrik

$$G_k(w,\pi,q) = G(w_k,\pi_k,q_k) = (w_k - \pi_k^{-1})^2/(2\pi_k^{-1}q_k)$$

verwendet, ergibt sich das folgende Gewicht für das k-te Element

$$w_k = \pi_k^{-1}(1 + q_k\vec{x}_k^\top\vec{\lambda})\,, \tag{3.89}$$

mit $\vec{\lambda}$ als Lösung der Gleichung $\sum_{k\in\mathscr{U}} I_k\pi_k^{-1}\vec{x}_k(1+q_k\vec{x}_k^\top\vec{\lambda}) = \vec{\tau}_x$, d.h.

$$\vec{\lambda} = \hat{\mathbf{T}}_x^{-1}\left(\vec{\tau}_x - \sum_{k\in\mathscr{U}} \mathfrak{I}_k\pi_k^{-1}\vec{x}_k\right)$$

unter der Annahme, dass die Inverse von $\hat{\mathbf{T}}_x = \sum_{k\in\mathscr{U}} I_k\pi_k^{-1}q_k\vec{x}_k\vec{x}_k^\top$ existiert (Särndal, 2007).

Es ist anzumerken, dass $\hat{\tau}_{\text{cal}}$ für G_k wie oben beschrieben identisch ist zu dem sog. linearen generalisierten Regressionsschätzer (lGREG). Der lGREG ist ein Schätzer, der die nicht beobachteten Elemente einer Population mit Hilfe eines linearen Regressionsmodells mit fixen Effekten schätzt, um so eine effizientere Schätzung zu ermöglichen, (Särndal et al., 1992, S. 225ff).

Die Varianz von $\hat{\tau}_{\text{cal}}$ mit Gewichten wie in (3.89) lässt sich approximieren durch

$$\mathrm{V}\left(\hat{\tau}_{\text{cal}}\right) \approx \mathrm{V}\left(\sum_{k\in\mathscr{U}} \mathfrak{I}_k\pi_k^{-1}e_k\right)\,. \tag{3.90}$$

Dabei ist e_k das Residuum der Regressionsgleichung $y_k = \vec{x}_k^\top\vec{\beta}$ mit

$$\vec{\beta} = \mathbf{T}_x^{-1}\left(\sum_{k\in\mathscr{U}} q_k\vec{x}_k y_k\right)$$

und $\mathbf{T}_x = \sum_{k\in\mathcal{U}} q_k\vec{x}_k\vec{x}_k^\top$. Die Herleitung für die asymptotische Varianz in (3.90) findet sich in Särndal et al. (1992, S. 234ff) bzw. in Deville (1999) unter Verwendung der Einflussfunktion des Kalibrierungsschätzers, wenn dieser als Funktional geschrieben wird.

Zur Schätzung der Varianzapproximation in (3.90) werden die geschätzten Residuen $\hat{e}_k$ benötigt. Hierzu wird der Schätzer $\hat{\vec{\beta}}$ für $\vec{\beta}$ mit

$$\hat{\vec{\beta}} = \hat{\mathbf{T}}_x^{-1}\left(\sum_{k\in\mathcal{U}} \mathfrak{I}_k\pi_k^{-1}q_k\vec{x}_k y_k\right)$$

verwendet. Dann ist $\hat{e}_k = y_k - \vec{x}_k^\top\hat{\vec{\beta}}$. Folglich kann die Varianz in (3.84) für w_k wie in (3.89) geschätzt werden durch

$$\widehat{\mathrm{V}}\left(\sum_{k\in\mathcal{U}} \mathfrak{I}_k w_k z_k\right) = \sum_{k\in\mathcal{U}}\sum_{l\in\mathcal{U}} \frac{(\pi_{k,l} - \pi_k\pi_l)}{\pi_{k\,l}}\frac{\hat{e}_k^*}{\pi_k}\frac{\hat{e}_l^*}{\pi_l}\,. \tag{3.91}$$

mit $\hat{e}_k^* = \hat{z}_k - \vec{x}_k^\top\hat{\vec{\beta}}$ (Deville, 1999). Hulliger, Alfons, Bruch et al. (2011, Kapitel 9) stellen eine Simulationstudie vor, welche die Varianzschätzung unter Verwendung der in Tabelle 3.1 darzustellen geschätzten linearisierten Werte für verschiedene Stichprobedesigns untersucht. Zudem findet sich in Hulliger, Alfons, Bruch et al. (2011, Kapitel 9) auch ein Vergleich zwischen den hier vorgestellten Verfahren der Linearisierung und Resampling Methoden.

Tabelle 3.1: Schätzer der linearisierten Werte

Statistik θ	Punktschätzer $\hat{\theta}$	Estimated linearized value ${}_{\hat{\theta}}\hat{z}_k$
ARPT	$0,6\hat{F}_N^{-1}(0,5)$	$-\frac{0,6}{\hat{N}\hat{F}_N'(\hat{F}_N^{-1}(0,5))}\left(\mathbb{1}[y_k \le \hat{F}_N^{-1}(0,5)] - 0,5\right)$
ARPR	$\hat{F}_N(\widehat{\mathrm{ARPT}})$	$\frac{1}{\hat{N}}\left(\mathbb{1}[y_k \le \widehat{\mathrm{ARPT}}] - \widehat{\mathrm{ARPR}}\right) -$ $\frac{0,6\hat{F}_N'[\widehat{\mathrm{ARPT}}]}{\hat{F}_N'(\hat{F}_N^{-1}(0,5))}\left(\frac{\mathbb{1}[y_k \le \hat{F}_N^{-1}(0,5)] - 0,5}{\hat{N}}\right)$
QSR	$\frac{\hat{\mu}_\mathrm{r}}{\hat{\mu}_\mathrm{a}}$	$\frac{y_k - [(y_k - \hat{F}_N^{-1}(0,8))\mathbb{1}[y_k \le \hat{F}_N^{-1}(0,8)] + 0,8\hat{F}_N^{-1}(0,8)]}{\sum_{k\in\mathscr{s}} w_k y_k \mathbb{1}[y_k \le \hat{F}_N^{-1}(0,2)]} -$ $\frac{\widehat{\mathrm{QSR}}\,[(y_k - \hat{F}_N^{-1}(0,2))\mathbb{1}[y_k \le \hat{F}_N^{-1}(0,2)] + 0,2\hat{F}_N^{-1}(0,2)]}{\sum_{k\in\mathscr{s}} w_k y_k \mathbb{1}[y_k \le \hat{F}_N^{-1}(0,2)]}$
GINI	$\frac{1}{\hat{N}\hat{\mu}}\sum_{k\in\mathscr{s}} w_k(2\hat{F}_N(y_k) - 1)y_k$	$\frac{2}{\hat{N}\hat{\mu}}\left[\sum_{k\in\mathscr{s}} w_k\frac{\mathbb{1}[y_k \ge y_k]y_k}{\hat{N}} + \left(\hat{F}_N(y_k) - \frac{\widehat{\mathrm{GINI}}+1}{2}\right)y_k - \frac{\hat{\mu}}{2}(\widehat{\mathrm{GINI}}+1)\right]$

$$\hat{F}_N(y) = \sum_{k\in\mathscr{s}} w_k\mathbb{1}(y_k \le y)(\sum_{k\in\mathscr{s}} w_k)^{-1} \qquad \hat{\mu}_\mathrm{p} = \sum_{k\in\mathscr{s}} w_k y_k\mathbb{1}[y_k \le \hat{F}_N^{-1}(0,2)]\frac{1}{\sum_{k\in\mathscr{s}} w_k 0,2}$$

$$\hat{F}_N^{-1}(p) = \inf\{y\in\mathbb{R}|p \le \hat{F}_N(y)\} \qquad \hat{\mu}_\mathrm{r} = \sum_{k\in\mathscr{s}} w_k\left(y_k - y_k\cdot\mathbb{1}[y_k \le \hat{F}_N^{-1}(0,8)]\right)\frac{1}{\sum_{k\in s} w_k(1-0,8)}$$

$$\hat{N} = \sum_{k\in s} w_k$$

Kapitel 4

Schätzung von Veränderungen in Querschnitten über die Zeit

4.1 Schätzung von Querschnitten im Zeitverlauf

In Abschnitt 2.1 Definition 2.5 wurde festlegt, dass $\mathfrak{I}_k^t = 0$ für alle $k \notin \mathscr{U}^t$. Eine ähnliche Festlegung wird für die Werte der Untersuchungsvariable getroffen. Ist die interessierende Variable y, so sind zu den verschiedenen Untersuchungszeitpunkten ihre Ausprägungen gegeben durch $\vec{y}^t = (y_1^t, \ldots, y_N^t)^\top$ für alle $t \in \mathscr{T}$. Dabei gilt

Definition 4.1.

$$y(k,t) = \begin{cases} y_k^t & \text{für } k \in \mathscr{U}^t \\ 0 & \text{für } k \notin \mathscr{U}^t \end{cases}$$

So bezeichnet die $N \times T$ Matrix $\boldsymbol{Y} = (\vec{y}^1, \ldots, \vec{y}^T)$ den (unbekannten) Parameter der endlichen Population beschrieben durch $\mathscr{U} = \bigcup_{t=1}^T \mathscr{U}^t$.

Da Ziehungsrahmen und Population als identisch zu verstehen sind, wäre ein strikterer Ansatz $y(k,t)$ für $k \notin \mathscr{U}^t$ als nicht definiert zu behandeln. Denn Element k existiert nicht zum Zeitpunkt t, womit ihm auch kein Merkmalsbetrag y_k^t zugeschrieben werden kann. Somit ist für $k \notin \mathscr{U}^t$ die Definition von $y(k,t)$ eine weitaus folgenreichere Entscheidung als die Festlegung $\mathfrak{I}_k^t = 0$, die keinerlei Einfluss auf das Stichprobendesign hat. Auch wird Definition 4.1 wieder nach rein praktischen Überlegungen getroffen und kann vor allem bei der Betrachtung linearer Statistiken von Vorteil sein. So ist der Totalwert von $\vec{y}^t$ gegeben durch

$$\tau^t = \sum_{k \in \mathscr{U}} y_k^t = \sum_{k \in \mathscr{U}^t} y_k^t \,.$$

Bei anderen Statistiken, die z.B. auf der empirischen Verteilungsfunktion beruhen, ist eine derart allgemeine Darstellung nicht möglich. Hier sollte eine Formulierung verwendet werden, die explizit nur von Parametern $y(k,t)$ mit $k \in \mathscr{U}^t$ abhängt. Unabhängig von der Wahl eines Wertes für $y(k,t)$, wenn $k \notin \mathscr{U}^t$, ermöglicht eine Definition wie 4.1 die Verwendung bekannter Rechenregeln und eine vereinfachte Schreibweise.

Zur Veranschaulichung des Problems soll exemplarisch am Beispiel des π-Schätzers für τ^t eine gültige Definition für diesen in $\mathbb{R}$ gefunden werden. Zwar ist $\tau^t = \vec{y}^{t\top}\vec{1}_N$, jedoch lässt sich der π-Schätzer $\hat{\tau}^t \in \mathbb{R}$ für τ^t allgemein nicht darstellen als $\hat{\tau}^t = \check{\vec{y}}^{t\top}\vec{\mathfrak{s}}^t$, mit $\check{\vec{y}}^t = (y_1^t/\pi_1^t, \ldots, y_N^t/\pi_N^t)^\top$, denn für alle $k \notin \mathscr{U}^t$ ist $1/\pi_k^t \notin \mathbb{R}$. Um dieses Problem zu umgehen, wird eine Funktion E^t definiert, mit $E^t : \mathbb{R}^N \mapsto \mathbb{R}^{N^t}$. E^t an der Stelle $x \in \mathbb{R}^N$, mit x als Spaltenvektor, ist gegeben mit

$$x_e = E^t(x) := (x^\top \boldsymbol{E}^t)^\top ,$$

dabei ist $\boldsymbol{E}^t$ eine $N \times N^t$ Matrix, der Form $\boldsymbol{E}^t = (\vec{I}_l)_{l \in \mathscr{U}^t}$ und $\vec{I}_l$ als der l-te Spaltenvektor der $N \times N$ Einheitsmatrix $\boldsymbol{I}_N$. Eine zulässige Definition für $\hat{\tau}^t \in \mathbb{R}$ ist dann

$$\begin{aligned} \hat{\tau}^t &= \vec{\mathfrak{s}}_e^{t\top}\check{\vec{y}}_e^t \\ &= \sum_{k \in \mathscr{U}^t} \frac{y_k^t}{\pi_k^t} , \end{aligned} \tag{4.1}$$

mit $\vec{\mathfrak{s}}_e^t = E^t(\vec{\mathfrak{s}}^t)$ und $\check{\vec{y}}_e^t = inv(diag(E^t(\vec{\pi}^t)))E^t(\vec{y}^t)$.

4.2 Schätzung von Veränderungen von Querschnitten im Zeitverlauf

Ein Zeitreihe von Querschnitten der gleichen Statistik θ über den Beobachtungszeitraum wird dargestellt als

$$\vec{\theta} = (\theta^1, \ldots, \theta^T)^\top .$$

Ziel ist es, die Veränderungen von θ zwischen allen Zeitpunkten in $\mathscr{T}$ zu schätzen.

Ein allgemeines Veränderungsmaß G lässt sich beschreiben als

$$G : \mathbb{R}^T \times \mathbb{R}^T \mapsto \mathbb{R}^{T \times T}$$

mit

$$G(\vec{x},\vec{y}) := \left[g^{t,u}(x^t,y^u)\right]_{\substack{t=1,\ldots,T \\ u=1,\ldots,T}} \tag{4.2}$$

Die Funktion $g^{t,u}$ in (4.2), mit $g^{t,u} : \mathbb{R}^2 \mapsto \mathbb{R}$, ist das Maß für die Veränderung zwischen den Zeitpunkten t und u. Symmetrie für G an der Stelle $(\vec{x},\vec{y})$ ist nicht zwingend gegeben, d.h. $g^{t,u}(x,y) = g^{u,t}(y,x)$ liegt allgemein nicht vor. Wird das identische Veränderungsmaß zwischen allen Beobachtungszeitpunkten gewählt, ist $g^{t,u} = g$ für alle $t,u \in \mathscr{T}$.

Es sei $\boldsymbol{H} = G(\vec{\theta},\vec{\theta})$, die Matrix der Veränderungen zwischen allen Elementen von $\vec{\theta}$. $\boldsymbol{H}$ wird durch $\widehat{\boldsymbol{H}}$ geschätzt, mit $\widehat{\boldsymbol{H}} = \hat{G}(\mathfrak{S},\boldsymbol{Y})$. $\boldsymbol{\Sigma}_{\hat{H}}$ bezeichnet die Matrix der Varianzen aller geschätzten Veränderungen in $\widehat{\boldsymbol{H}}$, d.h.

$$\boldsymbol{\Sigma}_{\hat{H}} = \left[\mathrm{V}\left(\hat{g}^{t,u}(\mathfrak{S},\boldsymbol{Y})\right)\right]_{\substack{t=1,\ldots,T \\ u=1,\ldots,T}} . \tag{4.3}$$

Es ist zu beachten, dass die Festlegung von $\hat{g}^{t,u}: \mathbb{R}^{N\times T} \times \mathbb{R}^{N\times T} \mapsto \mathbb{R}$ in (4.3), mehr Flexibilität bei der Wahl eines Schätzers für Veränderungen erlaubt. Für die weitere Betrachtung steht jedoch der Fall im Vordergrund, dass $g^{t,u}(\theta^t, \theta^u)$ geschätzt wird aus der Veränderung von Querschnittschätzern. Hier ist $\hat{g}^{t,u}(\mathfrak{S}, \boldsymbol{Y}) = g^{t,u}(\hat{\theta}^t, \hat{\theta}^u)$. Aus Gründen der Kohärenz kann eine solche Wahl für $\hat{g}^{t,u}$ sinnvoll sein. So fallen die Schätzungen der Veränderungen mit der der Querschnitte zusammen.

4.2.1 Varianz von Veränderungsmaßen

Zunächst soll davon ausgegangen werden, dass $\vec{\theta}$ ein Vektor von Totalwerten ist, d.h. $\vec{\theta} = \vec{\tau}$, mit $\vec{\tau} = (\tau^1, \ldots, \tau^T)^\top$. Dabei ist $\theta^t = \tau^t$ der Totalwert der Variable $\vec{y}^t = (y_1^t, \ldots, y_N^t)^\top$. Der π-Schätzer für $\vec{\tau}$ ist $\vec{\hat{\tau}} = (\hat{\tau}^t)_{t\in\mathscr{T}}^\top$, mit $\hat{\tau}^t$ nach (4.1). Somit ist $\boldsymbol{\Sigma}_{\hat{H}_{t\,u}}$, das (t,u)-te Element in $\boldsymbol{\Sigma}_{\hat{H}}$, für lineare g gegeben durch

$$\begin{aligned} \mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) = & \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} \sum_{k\in\mathscr{U}^i} \sum_{l\in\mathscr{U}^j} (\pi_{k,l}^{i,j} - \pi_k^i \pi_l^j) \frac{z_k^i}{\pi_k^i} \frac{z_l^j}{\pi_l^j} \\ = & \nabla^{t,u\top} \boldsymbol{\Sigma}_{(\hat{\tau}^t, \hat{\tau}^u)} \nabla^{t,u} \,. \end{aligned} \tag{4.4}$$

Dabei sind

$$\begin{aligned} z_k^t &= \frac{\partial g^{t,u}(\tau^t, \tau^u)}{\partial \tau^t} y_k^t \,, \\ \nabla^{t,u} &= \left(\frac{\partial g^{t,u}(\tau^t, \tau^u)}{\partial \tau^i} \right)_{i\in\{t,u\}}^\top \quad \text{und} \\ \boldsymbol{\Sigma}_{(\hat{\tau}^t, \hat{\tau}^u)} &= \begin{pmatrix} \mathrm{V}(\hat{\tau}^t) & \mathrm{COV}(\hat{\tau}^u, \hat{\tau}^t) \\ \mathrm{COV}(\hat{\tau}^t, \hat{\tau}^u) & \mathrm{V}(\hat{\tau}^u) \end{pmatrix} \,. \end{aligned}$$

Für nicht lineare $g^{t,u}$ kann (4.4) als Approximation zu $\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right)$ verwendetet werden, sofern $g^{t,u}$ stetig differenzierbar bis zweiter Ordnung an jedem Punkt der offenen Menge ist, die (τ^t, τ^u) und $(\hat{\tau}^t, \hat{\tau}^u)$ einschließt (siehe, Abschnitt 3.4).

Wird beispielsweise die Differenz zwischen den Querschnitten als Veränderungsmaß gewählt, ist $g^{t,u} = g \; \forall t,u \in \mathscr{T}$ und $g(x,y) = y - x$. Da g eine lineare Funktion ist, gilt

$$g(\hat{\tau}^t, \hat{\tau}^u) = \begin{pmatrix} \mathfrak{s}_e^t \\ \mathfrak{s}_e^u \end{pmatrix}^\top \begin{pmatrix} \check{z}_e^t \\ \check{z}_e^u \end{pmatrix} \,, \tag{4.5}$$

mit $\check{z}_e^t = \check{E}^t(\vec{z}^t)$ und $\vec{z}^t = (z_1^t, \ldots, z_N^t)$. Zudem ist

$$\begin{aligned} z_k^t &= \frac{\partial g(\tau^t, \tau^u)}{\partial \tau^t} y_k^t = -y_k^t \text{ und} \\ z_k^u &= \frac{\partial g(\tau^t, \tau^u)}{\partial \tau^u} y_k^u = y_k^u \,. \end{aligned}$$

Die Varianz von $g(\hat{\tau}^t, \hat{\tau}^u)$ lässt sich schreiben als

$$\begin{aligned}
\mathrm{V}\left(g(\hat{\tau}^t, \hat{\tau}^u)\right) =& \mathrm{V}\left(\hat{\tau}^u\right) + \mathrm{V}\left(\hat{\tau}^u\right) - 2\mathrm{COV}\left(\hat{\tau}^t, \hat{\tau}^u\right) \\
=& \sum_{k \in \mathscr{U}^t} \pi_k^t (1 - \pi_k^t) \left(\frac{y_k^t}{\pi_k^t}\right)^2 + \sum_{k \in \mathscr{U}^t} \sum_{\substack{l \in \mathscr{U}^t \\ l \neq k}} \left(\pi_{k,l}^t - \pi_k^t \pi_l^t\right) \frac{y_k^t}{\pi_k^t} \frac{y_l^t}{\pi_l^t} \\
&+ \sum_{k \in \mathscr{U}^u} \pi_k^u (1 - \pi_k^u) \left(\frac{y_k^u}{\pi_k^u}\right)^2 + \sum_{k \in \mathscr{U}^u} \sum_{\substack{l \in \mathscr{U}^u \\ l \neq k}} \left(\pi_{k,l}^u - \pi_k^u \pi_k^u\right) \frac{y_k^u}{\pi_k^u} \frac{y_l^u}{\pi_l^u} \\
&- 2 \left[\sum_{k \in \mathscr{U}^{t,u}} \left(\pi_k^{t,u} - \pi_k^t \pi_k^u\right) \frac{y_k^t}{\pi_k^t} \frac{y_k^u}{\pi_k^u} + \sum_{k \in \mathscr{U}^t} \sum_{\substack{l \in \mathscr{U}^u \\ l \neq k}} \left(\pi_{k,l}^{t,u} - \pi_k^t \pi_l^u\right) \frac{y_k^t}{\pi_k^t} \frac{y_l^u}{\pi_l^u} \right],
\end{aligned} \tag{4.6}$$

mit $\mathscr{U}^{t,u} = \{\mathscr{U}^t \cap \mathscr{U}^u\}$.

In Anlehnung an Theorem 3.3 von Hájek (1981, S. 24) bei einmaliger Stichprobenziehung ist zu vermuten, dass es für gemeinsame Designs mit beliebigen Inklusionswahrscheinlichkeiten erster Ordnung, festen Stichprobenumfängen im Querschnitt und positiver oder negativer Koordination, keine einfache Form von $\mathrm{COV}\left(\mathfrak{I}_k^t, \mathfrak{I}_k^t\right)$, der gestalt $c_k^t c_l^u$ gibt. Auch sind Approximationen für $\pi_k^{t,u}$, wie sie in Abschnitt 3.3 für $\pi_{k,l}^t$ geschildert wurden, gerade bei systematischen Längsschnittdesigns nur schwer möglich. Es handelt sich hier gerade nicht um Designs mit maximaler oder hoher Entropie, sondern um Designs mit einer minimalen oder geringen Entropie. Unter der Annahme, dass Vermutung 2.18 zutrifft, scheint eine Einschränkung der Betrachtung auf einfache Querschnittdesigns, wie in Algorithmus 3, oder voneinander unabhängige Längsschnittdesigns wie in den Algorithmen 1 und 2, jedoch gerechtfertigt. Demnach stellt sich zum Teil die Problematik einer einfachen Form für $\mathrm{COV}\left(\mathfrak{I}_k^t, \mathfrak{I}_l^u\right)$ in der Praxis gar nicht. Für Designs mit $\pi_{k,l}^{t,u} \neq \pi_k^t \pi_l^u$ besteht die Schwierigkeit zudem meist in der Bestimmung der $\pi_k^{t,u}$, wie dies für die Algorithmen 3 und 9 in Kapitel 2 geschildert wurde.

Für die Klasse von gemeinsamen Designs mit festen Stichprobenumfängen im Querschnitt stellt Tam (1984) das Analogon der Beziehung in (3.19e) auf mit

$$\sum_{\substack{l \in \mathscr{U}^u \\ l \neq k}} \left(\pi_k^t \pi_l^u - \pi_{k,l}^{t,u}\right) = \pi_k^{t,u} - \pi_k^t \pi_k^u . \tag{4.7}$$

Zwar ist (4.7) auch gültig für veränderliche Populationen, jedoch lässt sich die Varianz in (4.4) allgemein nicht in einer quadratischen Form darstellen. Dies ist nur möglich wenn (4.4) konditioniert wird auf die Anzahl der Elemente, die in $\mathfrak{s}^t$ und $\mathfrak{s}^u$ aus $\mathscr{U}^{t,u}$ gezogen werden. So ist

$$\begin{aligned}
&\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | (\mathfrak{n}_{\mathscr{U}^{t,u}}^t, \mathfrak{n}_{\mathscr{U}^{t,u}}^u)\right) = \\
&\quad \mathrm{V}\left(\hat{\tau}_z^t | (\mathfrak{n}_{\mathscr{U}^{t,u}}^t, \mathfrak{n}_{\mathscr{U}^{t,u}}^u)\right) + \mathrm{V}\left(\hat{\tau}_z^u | (\mathfrak{n}_{\mathscr{U}^{t,u}}^t, \mathfrak{n}_{\mathscr{U}^{t,u}}^u)\right) + 2\mathrm{COV}\left(\hat{\tau}_z^t, \hat{\tau}_z^u | (\mathfrak{n}_{\mathscr{U}^{t,u}}^t, \mathfrak{n}_{\mathscr{U}^{t,u}}^u)\right),
\end{aligned} \tag{4.8}$$

mit

$$\hat{\tau}_z^t = \sum_{k \in \mathscr{U}^t} \frac{z_k^t}{\pi_k^t},$$

$\mathfrak{n}^t_{\mathcal{U}^{t,u}} = card(\mathscr{s}^t \cap \mathcal{U}^{t,u})$ und $\mathfrak{n}^u_{\mathcal{U}^{t,u}} = card(\mathscr{s}^u \cap \mathcal{U}^{t,u})$. Die beiden Varianz-Terme in (4.8) lassen sich in der gleichen Form darstellen wie (3.6). Die Kovarianz lässt sich schreiben als

$$\begin{aligned} &\mathrm{COV}\left(\hat{\tau}^t_z, \hat{\tau}^u_z | (\mathfrak{n}^t_{\mathcal{U}^{t,u}}, \mathfrak{n}^u_{\mathcal{U}^{t,u}})\right) = \\ &-\frac{1}{2}\left(\sum_{k\in\mathcal{U}^{u\setminus t}}\sum_{l\in\mathcal{U}^{t,u}} \delta^{t,u}_{k,l}\gamma^{t,u}_{k,l} + \sum_{k\in\mathcal{U}^{t,u}}\sum_{l\in\mathcal{U}^{t,u}} \delta^{t,u}_{k,l}\gamma^{t,u}_{k,l} + \sum_{k\in\mathcal{U}^{t,u}}\sum_{l\in\mathcal{U}^{t\setminus u}} \delta^{t,u}_{k,l}\gamma^{t,u}_{k,l}\right), \end{aligned} \tag{4.9}$$

wobei $\gamma^{t,u}_{k,l} = \left(\frac{z^t_k}{\pi^t_k} - \frac{z^t_l}{\pi^t_l}\right)\left(\frac{z^u_k}{\pi^u_k} - \frac{z^u_l}{\pi^u_l}\right)$, $\delta^{t,u}_{k,l} = \mathrm{COV}\left(\mathfrak{I}^t_k \mathfrak{I}^u_l | (\mathfrak{n}^t_{\mathcal{U}^{t,u}}, \mathfrak{n}^u_{\mathcal{U}^{t,u}})\right)$, $\mathcal{U}^{u\setminus t} = \mathcal{U}^u \setminus \mathcal{U}^{t,u}$ und $\mathcal{U}^{t\setminus u} = \mathcal{U}^t \setminus \mathcal{U}^{t,u}$. Es sei angemerkt, dass

$$\begin{aligned} \sum_{k\in\mathcal{U}^{t,u}}\sum_{l\in\mathcal{U}^{t,u}} \delta^{t,u}_{k,l}\gamma^{t,u}_{k,l} &= \sum_{k\in\mathcal{U}^{t,u}}\sum_{l\in\mathcal{U}^{t,u}} \delta^{u,t}_{k,l}\gamma^{u,t}_{k,l} \text{ und} \\ \sum_{k\in\mathcal{U}^{u\setminus t}}\sum_{l\in\mathcal{U}^{t,u}} \delta^{t,u}_{k,l}\gamma^{t,u}_{k,l} &= \sum_{k\in\mathcal{U}^{t,u}}\sum_{l\in\mathcal{U}^{u\setminus t}} \delta^{u,t}_{k,l}\gamma^{u,t}_{k,l}\,. \end{aligned}$$

Letzteres gilt auch analog für den dritten Term auf der rechten Seite von (4.9).

Die Form der Kovarianz (4.9) ist eine Verallgemeinerung der Darstellung von Laniel (1987). Im Speziellen ist zudem (4.8) gleich (4.4) für eine unveränderliche Population, sowie $p^t(.)$ und $p^u(.)$ mit festen Stichprobenumfängen. Weitere Darstellungen der Kovarianz (4.9) im Falle Koordinierter SRS, finden sich auch in Hulliger (1995) und Nordberg (2000).

Unter der Bedingungen (2.23) ist eine Darstellung für $\mathrm{COV}\left(\hat{\tau}^t_z, \hat{\tau}^u_z\right)$ ohne Verwendung der $\pi^{t,u}_{k,l}$ möglich. So ist in diesem Fall für $t < u$, $k \in \mathcal{U}^t$ und $l \in \mathcal{U}^u$

$$\pi^{t,u}_{k,l} = \begin{cases} \dfrac{\pi^{t,u}_l}{\pi^t_l}\pi^t_{k,l} + \dfrac{(\pi^u_l - \pi^{t,u}_l)(\pi^t_k - \pi^t_{k,l})}{(1-\pi^t_l)} & \text{für } l,k \notin \mathcal{U}^{t,u} \wedge \pi^t_l, \pi^u_k < 1 \\ \pi^t_k\pi^u_l & \text{sonst}\,. \end{cases} \tag{4.10}$$

Unter der Annahme, dass $\pi^u_k < 1$ und $\pi^t_l < 1$ $\forall k,l \in \mathcal{U}^{t,u}$ ist, ergibt sich durch das Einsetzten von (4.10) in (4.4), für $i = t$, $j = u$ und $t < u$,

$$\mathrm{COV}\left(\hat{\tau}^t_z, \hat{\tau}^u_z\right) = \sum_{k\in\mathcal{U}^{t,u}}\sum_{l\in\mathcal{U}^{t,u}} \Delta^{t,u}_{k,l}\frac{z^t_k}{\pi^t_k}\frac{z^u_l}{\pi^u_l}\,. \tag{4.11}$$

Dabei ist $\Delta^{t,u}_{k,l}$ die Kovarianz der Stichprobenindikatoren, die gegeben ist durch

$$\Delta^{t,u}_{k,l} = \frac{(\pi^t_{k,l} - \pi^t_k\pi^t_l)(\pi^{t,u}_l - \pi^t_l\pi^u_l)}{\pi^t_l(1-\pi^t_l)}\,, \tag{4.12}$$

(Wood, 2008). Es ist zu beachten, dass die Kovarianz in (4.11) nur von den Elementen in der gemeinsamen Population $\mathcal{U}^{t,u}$ abhängt, da für $l \notin \mathcal{U}^t$ oder $k \notin \mathcal{U}^u$ $\mathrm{COV}\left(\mathfrak{I}^t_k, \mathfrak{I}^u_l\right) = 0$, wenn (2.23) gilt. Für eine unveränderliche Population stellt damit (4.11) die Kovarianz zwischen zwei π-Schätzern $\hat{\tau}^t$ und $\hat{\tau}^u$ unter Algorithmus 3 mit der Ordnungsvariable gebildet nach h in (2.14) bzw. (2.36) dar.

4.2.2 Varianzschätzung für Veränderungsmaße

Unveränderliche Populationen

Zuerst soll der Fall einer unveränderlichen Population mit Querschnittdesigns $p^t(.)$ als SRS für alle $t \in \mathscr{T}$ sowie dem identischen Längsschnittdesign für alle $k \in \mathscr{U}$ betrachtet werden. Es wird festgelegt, dass $\pi_k^t = \pi^t$, $\pi_k^u = \pi^u$, $\pi_k^{t,u} = \pi^{t,u}$ und $\Delta_{k,l}^{t,u} = \Delta^{t,u} \; \forall \, k,l \in \mathscr{U}$ ist. Durch das Einsetzten von (4.12) in (4.4) lässt sich $\widehat{\mathrm{V}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right)$ dann schreiben als

$$\begin{aligned}
\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) &= \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} \frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{N} \frac{(\pi^i \pi^j - \pi^{i,j})}{\pi^i \pi^j} \left(z_k^i - z_l^i\right)\left(z_k^j - z_l^j\right) \\
&= \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} \Delta^{i,j} N \sum_{k=1}^{N} \left(z_k^i - \mu_z^i\right)\left(z_k^j - \mu_z^j\right) \\
&= \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} \Delta^{i,j} N(N-1)\sigma_z^{i,j} \\
&= \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} N^2 \left(\frac{\mathrm{E}\left(\mathfrak{n}^{i,j}\right)}{n^i n^j} - \frac{1}{N}\right) \sigma_z^{i,j}
\end{aligned} \tag{4.13}$$

mit $\Delta^{t,u} = (\pi^t \pi^u - \pi^{t,u})/(\pi^t \pi^u)$, $\pi^t = n^t/N$, $\pi^u = n^u/N$, $\pi^{t,u} = \pi_k^{t,u}$ und $\forall \, k \in \mathscr{U}$, sowie $\mathrm{E}\left(\mathfrak{n}^{t,u}\right) = N\pi^{t,u}$. Des Weiteren sind $\mu_z^t = \sum_{k=1}^N z_k^t/N$ und $\mu_z^u = \sum_{k=1}^N z_k^u/N$ die Mittelwerte der Variablen $\vec{z}^t$ bzw. $\vec{z}^u$ und $\sigma_z^{t,u}$ deren Populationskovarianz

$$\sigma_z^{t,u} = \frac{1}{N-1} \sum_{k=1}^{N} (z_k^t - \mu_z^t)(z_k^u - \mu_z^u) \, .$$

Da die Erwartungstreue von $\hat{\tau}_z^t$ bzw. $\hat{\tau}_z^u$ nicht von $\mathfrak{n}^{t,u}$ abhängt, ist eine naheliegende Annahme, dass auch

$$\mathrm{E}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | \mathfrak{n}^{t,u}\right) = g^{t,u}(\tau_z^t, \tau_z^u) \tag{4.14}$$

gilt. Für diesen Fall folgt, dass

$$\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) = \mathrm{E}_{\mathfrak{n}^{t,u}}\left(\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | \mathfrak{n}^{t,u}\right)\right) \, .$$

Dabei ist $\mathrm{E}_{\mathfrak{n}^{t,u}}$ der Erwartungswert in Bezug auf die Überlappung $\mathfrak{n}^{t,u}$. Somit lässt sich ein unverzerrter Schätzer für (4.13), unter der Annahme von (4.14), beschreiben als $\hat{\mathrm{V}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | \mathfrak{n}^{t,u}\right)$, mit

$$\mathrm{E}\left(\hat{\mathrm{V}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | \mathfrak{n}^{t,u}\right)\right) = \mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | \mathfrak{n}^{t,u}\right) \, .$$

Da $\mathrm{E}_{\mathfrak{n}^{t,u}}\left(\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | \mathfrak{n}^{t,u}\right)\right) = \mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right)$, ist $\hat{\mathrm{V}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | \mathfrak{n}^{t,u}\right)$ ein unverzerrter Schätzer für $\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right)$. Folglich kann (4.13), unter der Annahme von (4.14), unverzerrt geschätzt werden durch

$$\widehat{\mathrm{V}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) = \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} N^2 \left(\frac{\mathfrak{n}^{i,j}}{n^i n^j} - \frac{1}{N}\right) \hat{\sigma}_z^{i,j} \, , \tag{4.15}$$

dabei ist $\hat{\sigma}_z^{t,u}$ die Stichprobenkovarianz

$$\hat{\sigma}_z^{t,u} = \frac{1}{\mathfrak{n}^{t,u} - 1} \sum_{k=1}^{N} \mathfrak{I}_k^t \mathfrak{I}_k^u (z_k^t - \vec{z}^t)(z_k^u - \vec{z}^u) \,,$$

mit $\vec{z}^t_{t,u} = \sum_{k=1}^N \mathfrak{I}_k^t \mathfrak{I}_k^u y_k^t / \mathfrak{n}^{t,u}$ und $\vec{z}^u_{t,u} = \sum_{k=1}^N \mathfrak{I}_k^t \mathfrak{I}_k^u y_k^u / \mathfrak{n}^{t,u}$, (Qualité & Tillé, 2008). Es ist zu beachten, dass zur Anwendung von 4.15 das Längsschnittdesigns nicht bekannt sein muss. Die Kenntnis der beiden SRS Querschnittdesigns genügt in diesem Fall.

Gilt (4.14) nicht, so kann (4.14) auch unverzerrt geschätzt werden durch

$$\widehat{\mathrm{V}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) = \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} N^2 \left(\frac{\mathrm{E}\left(\mathfrak{n}^{t,u}\right)}{n^t n^u} - \frac{1}{N}\right) \hat{\sigma}_z^{t,u} \,, \tag{4.16}$$

da (4.13) die unbedingte Varianz von $g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)$ ist und

$$\mathrm{E}_{\mathfrak{n}^{t,u}}\left(\mathrm{E}\left(\hat{\sigma}_z^{t,u} | \mathfrak{n}^{t,u}\right)\right) = \sigma_z^{t,u} \,,$$

für $\pi^{t,u} > 0$ (Qualité & Tillé, 2008, Knottnerus & van Delden, 2012).

Der Nachteil von (4.16) gegenüber (4.15) ist, dass $\mathrm{E}\left(\mathfrak{n}^{t,u}\right)$ bekannt sein muss. Ist dies nicht der Fall, kann $\mathrm{E}\left(\mathfrak{n}^{t,u}\right)$ unverzerrt geschätzt werden durch $\mathfrak{n}^{t,u}/N$. Mit anderen Worten, es muss Schätzer (4.15) verwendet werden. Eine Bedingung für die Anwendbarkeit beider Schätzer ist, dass $\mathfrak{n}^{t,u} > 1$ gilt. Es sei denn, bei der Anwendung von (4.16) sind die Stichproben $\vec{\mathfrak{s}}^t$ und $\vec{\mathfrak{s}}^u$ abhängig von einander oder für (4.15) gilt $n^t n^u = N$, (Qualité & Tillé, 2008).

Es ist zu beachten, dass $\hat{\sigma}_z^{t,u}$, wenn $\pi^{t,u} = 0$, nicht unverzerrt für $\sigma_z^{t,u}$ ist, da hier $\hat{\sigma}_z^{t,u} = -1$ und somit $\widehat{\mathrm{COV}}\left(\hat{\tau}_z^t, \hat{\tau}_z^u\right) = N$, aber $\mathrm{COV}\left(\hat{\tau}_z^t, \hat{\tau}_z^u\right) = -N\sigma_z^{t,u}$, es sei denn $\vec{z}^t = -\vec{z}^u$, was eine für die Praxis irrelevante Ausnahme darstellen dürfte. In diesem Fall wären sowohl (4.15) als (4.16) nicht unverzerrt für (4.13). Beispielsweise würde dies für $g^{t,u}(x,y) = y - x$ zu einer Unterschätzung der Varianz in (4.6) führen. Falls $\pi_k^{t,u} > 0$, eignen sich somit (4.15) und (4.16) bedingt als Varianzschätzer für (4.4) unter dem gemeinsamen Design von Algorithmus 3 mit (2.14) oder (2.36).

Für ein gemeinsames Design mit einer fixen Überlappung $\mathfrak{n}^{t,u}$ und festen Stichprobenumfängen im Querschnitt präsentiert Berger (2004b) einen Schätzer für (4.4), der auf einer Verallgemeinerung des in Abschnitt 3.3 vorgestellten Ansatzes zur Approximation der Varianz eines π-Schätzers beruht. Analog zum Fall der einmaligen Ziehung wird angenommen, dass das gemeinsame Design eine maximale oder zumindest hohe Entropie (3.22) besitzt, so dass

$$\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) = \mathrm{V}_{\text{poiss}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | (\mathfrak{n}^t, \mathfrak{n}^u, \mathfrak{n}^{t,u})\right) \,.$$

Dabei ist die rechte Seite die Varianz unter Design

$$p_{\text{poiss}}\left(\mathbf{S} | \mathfrak{n}^t, \mathfrak{n}^u, \mathfrak{n}^{t,u}\right) \,, \tag{4.17}$$

einem hypothetischen gemeinsamen Poisson Design $p_{\text{poiss}}(.)$, konditioniert auf $(\mathfrak{n}^t, \mathfrak{n}^u, \mathfrak{n}^{t,u})$. Berger (2004b) verwendet keine Approximation für die Inklusionswahrscheinlichkeiten des Designs in (4.17), wie dies beispielsweise in Abschnitt 3.3.1 getan wurde. Statt dessen wird angenommen, dass diese den Wahrscheinlichkeiten π_k^t, π_k^u, und $\pi_k^{t,u}$ des tatsächlich verwendeten Designs $p(.)$ entsprechen. Es wird davon ausgegangen, dass unter einem gemeinsamen Poisson Design der Vektor

$$\vec{\vartheta}^{\,t,u} = (\hat{\tau}_z^t, \hat{\tau}_z^u, \mathfrak{n}^t, \mathfrak{n}^u, \mathfrak{n}^{t,u})^\top$$

multivariat normalverteilt ist, d.h.

$$\vec{\vartheta}^{\,t,u} \sim NV\left(\mathrm{E}\left(\vec{\vartheta}^{\,t,u}\right), \varSigma_{\vec{\vartheta}^{\,t,u}}\right) .$$

Dabei ist $\varSigma_{\vec{\vartheta}^{\,t,u}}$ die 5×5 Kovarianzmatrix

$$\varSigma_{\vec{\vartheta}^{\,t,u}} = \begin{pmatrix} \varSigma^{t,u}_{\vec{\hat{\tau}}_z, \vec{\hat{\tau}}_z} & \varSigma^{t,u}_{\vec{\hat{\tau}}_z, \vec{\mathfrak{n}}} \\ \varSigma^{t,u\,\top}_{\vec{\hat{\tau}}_z, \vec{\mathfrak{n}}} & \varSigma^{t,u}_{\vec{n}, \vec{\mathfrak{n}}} \end{pmatrix} ,$$

mit

$$\varSigma^{t,u}_{\vec{\hat{\tau}}_z, \vec{\hat{\tau}}_z} = \begin{pmatrix} \mathrm{V}_{\text{poiss}}\left(\hat{\tau}_z^t\right) & \mathrm{COV}_{\text{poiss}}\left(\hat{\tau}_z^t,\, \hat{\tau}_z^u\right) \\ \mathrm{COV}_{\text{poiss}}\left(\hat{\tau}_z^u,\, \hat{\tau}_z^t\right) & \mathrm{V}_{\text{poiss}}\left(\hat{\tau}_z^u\right) \end{pmatrix}$$

$$\varSigma^{t,u}_{\vec{\hat{\tau}}_z, \vec{\mathfrak{n}}} = \begin{pmatrix} \mathrm{COV}_{\text{poiss}}\left(\hat{\tau}_z^t,\, \mathfrak{n}^t\right) & \mathrm{COV}_{\text{poiss}}\left(\hat{\tau}_z^t,\, \mathfrak{n}^u\right) & \mathrm{COV}_{\text{poiss}}\left(\hat{\tau}_z^t,\, \mathfrak{n}^{t,u}\right) \\ \mathrm{COV}_{\text{poiss}}\left(\hat{\tau}_z^u,\, \mathfrak{n}^t\right) & \mathrm{COV}_{\text{poiss}}\left(\hat{\tau}_z^u,\, \mathfrak{n}^u\right) & \mathrm{COV}_{\text{poiss}}\left(\hat{\tau}_z^u,\, \mathfrak{n}^{t,u}\right) \end{pmatrix}$$

$$\varSigma^{t,u}_{\vec{\mathfrak{n}}, \vec{\mathfrak{n}}} = \begin{pmatrix} \mathrm{V}_{\text{poiss}}\left(\mathfrak{n}^t\right)\mathfrak{n}^t & \mathrm{COV}_{\text{poiss}}\left(\mathfrak{n}^t,\, \mathfrak{n}^u\right) & \mathrm{COV}_{\text{poiss}}\left(\mathfrak{n}^t,\, \mathfrak{n}^{t,u}\right) \\ \mathrm{COV}_{\text{poiss}}\left(\mathfrak{n}^t,\, \mathfrak{n}^u\right) & \mathrm{V}_{\text{poiss}}\left(\mathfrak{n}^u\right) & \mathrm{COV}_{\text{poiss}}\left(\mathfrak{n}^t,\, \mathfrak{n}^{t,u}\right) \\ \mathrm{COV}_{\text{poiss}}\left(\mathfrak{n}^t,\, \mathfrak{n}^{t,u}\right) & \mathrm{COV}_{\text{poiss}}\left(\mathfrak{n}^t,\, \mathfrak{n}^u\right) & \mathrm{V}_{\text{poiss}}\left(\mathfrak{n}^{t,u}\right) \end{pmatrix}$$

Somit ist

$$\mathrm{V}_{\text{poiss}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u) | (\mathfrak{n}^t, \mathfrak{n}^u, \mathfrak{n}^{t,u})\right) = \vec{1}_2^\top \Sigma^{t,u}_{\vec{\hat{\tau}}_z | \vec{\mathfrak{n}}} \vec{1}_2 ,$$

mit

$$\Sigma^{t,u}_{\vec{\hat{\tau}}_z | \vec{\mathfrak{n}}} = \varSigma^{t,u}_{\vec{\hat{\tau}}_z, \vec{\hat{\tau}}_z} - \varSigma^{t,u}_{\vec{\hat{\tau}}_z, \vec{\mathfrak{n}}} \varSigma^{t,u}_{\vec{\mathfrak{n}}, \vec{\mathfrak{n}}} \varSigma^{t,u\,\top}_{\vec{\hat{\tau}}_z, \vec{\mathfrak{n}}} \tag{4.18}$$

und $\vec{1}_2 = (1,1)^\top$. Um die Terme auf der rechten Seite von (4.18) erwartungstreu zu schätzen, wird ein Schätzer $\widehat{\varSigma}_{\vec{\vartheta}^{\,t,u}}$ benötigt, der unverzerrt für $\varSigma_{\vec{\vartheta}^{\,t,u}}$ ist. Ein solcher Schätzer lässt sich mit Hilfe der $3N \times 3N$ Kovarianzmatrix

$$\varDelta^{t,u}_{\text{poiss}} = [c_{i,j}]_{\substack{i=1,\dots,3N \\ j=1,\dots,3N}}$$

des Zufallsvektors $(\vec{\mathfrak{s}}_k^{t\top}, \vec{\mathfrak{s}}_k^{u\top}, \vec{\mathfrak{s}}_k^{t\top} \vec{\mathfrak{s}}_k^{u\top})^\top$ konstruieren. Da unter $\mathrm{COV}_{\text{poiss}}\left(I_k^t, I_l^u\right) = 0$ für $k \neq l$ oder $t \neq u$, sind nur die Elemente $c_{j,i}$ in $\varDelta^{t,u}_{\text{poiss}}$ von Null verschieden, mit $|j - i| = 0 \dot{\vee} N \dot{\vee} 2N$. Folglich sind die einzigen Inklusionswahrscheinlichkeiten, die für den Schätzer $\widehat{\varSigma}_{\vec{\vartheta}^{\,t,u}}$ bekannt sein müssen, π_k^t, π_k^u, und $\pi_k^{t,u}$. Der von Berger (2004b) entwickelte Varianzschätzer für (4.4) eignet sich für gemeinsame Designs mit komplexen Querschnittdesigns, wie beispielsweise das nicht sequenziellen Design, vorgeschlagen von Matei & Tillé (2005b), Matei & Skinner (2009), zur Maximierung oder Minimierung von $\mathfrak{n}^{t,u}$, das auch, wie in Abschnitt 2.3 geschildert, auf fixe $\mathfrak{n}^{t,u}$ angepasst angewendet werden kann. Eine Verwendung unter dem strikt sequenziellen Design, vorgeschlagen von Deville & Tillé (2000), mit ungleichen Inklusionswahrscheinlichkeiten erster Ordnung, ist ebenfalls denkbar. Eine einfache Form für $\pi_k^{t,u} - \pi_k^t \pi_k^u$ wird jedoch nicht angegeben. Für systematische Längsschnittdesigns ist dies mit der gewählten Methode, wie bereits erwähnt, auch kaum möglich, da es sich bei $p_k(.)$ gerade um ein Design mit einer niedrigen Entropie handelt. Es sei jedoch erwähnt, dass Berger (2003, 2005) einen Varianzschätzer für ein geordnetes systematisches Design entwickelt, indem er die Varianz in (3.23) nicht nur auf einen fixen Stichprobenumfang konditioniert, sondern auch auf die systematische Verteilung der Ziehungen über die geordnete Liste (z.B. Beobachtungszeitpunkte) hinweg.

Unter der Annahme vernachlässigbarer Auswahlsätze, d.h.

$$\pi_k^t \approx 0\,, \pi_k^u \approx 0 \text{ und } (\pi_k^t \pi_k^u)/\pi_k^{t,u} \approx 0\,, \tag{4.19}$$

schlagen Berger & Priam (2015) zudem einen vereinfachten Schätzer für $\mathrm{V}_{\text{poiss}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right)$ vor. Hierzu wird $\Sigma^{t,u}_{\vec{\hat{\tau}}_z|\vec{\mathfrak{n}}}$ geschätzt aus dem (ungewichteten) kleinste Quadrate Schätzer der Kovarianzmatrix der Residuen der multivariaten Regression von $(y_k^t/\pi_k^t \mathfrak{I}_k^t, y_k^u/\pi_k^u \mathfrak{I}_k^u)^\top$ auf $(\mathfrak{I}_k^t, \mathfrak{I}_k^u, \mathfrak{I}_k^t \mathfrak{I}_k^u)^\top$.

Veränderliche Populationen

Bei der Betrachtung veränderlicher Populationen besteht das Problem, dass Bedingung (2.23) und feste Stichprobenumfänge im Querschnitt unvereinbar scheinen. Grund hierfür ist das zufällige Auswählen von Elementen in $\vec{\mathfrak{s}}^t$, die nicht in $\mathscr{U}^u$ enthalten sind, oder von Elementen, die in $\vec{\mathfrak{s}}^u$ aber nicht in $\mathscr{U}^t$ enthalten sind. Dies hat zur Folge, dass ϕ_k^t nicht länger iid für alle $k \in \mathscr{U}^t$ ist. Wie in Abschnitt 2.4.3 beschrieben, trifft dies auch für Algorithmus 9 zu, wobei hier ϕ_k^t iid für alle $k \in \mathscr{C}^{\alpha\omega}$ ist. Somit ist die Darstellung der Kovarianz nach Gleichung (4.11) (bzw. der Varianz, wenn $t = u$, für Algorithmus 9 nicht möglich.

Ist $\mathscr{G}^t$ die Indexmenge der Kohorten zum Zeitpunkt t mit

$$\mathscr{G}^t = \{(\alpha,\omega) | (\alpha,\omega) \in \{(1,2), (1,3), \ldots, (1,t+1), (2,3), \ldots, (t,t+1)\}, \mathscr{C}^{\alpha\omega} \neq \emptyset\}\ ,$$

so kann (4.4) unter Algorithmus 9 dargestellt werden als

$$\begin{aligned}\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) &= \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} \sum_{\kappa\in\mathscr{G}^i} \sum_{\upsilon\in\mathscr{G}^j} \sum_{k\in\mathscr{C}^\kappa} \sum_{k\in\mathscr{C}^\upsilon} (\pi_{k,l}^{i,j} - \pi_k^i \pi_l^j) \frac{z_k^i}{\pi_k^i} \frac{z_l^j}{\pi_l^j} \\ &= \sum_{i\in\{t,u\}} \sum_{j\in\{t,u\}} \sum_{\kappa\in\mathscr{G}^i} \sum_{\upsilon\in\mathscr{G}^i} \Delta_{\kappa,\upsilon}^{t,u} \sum_{k\in\mathscr{C}^\kappa} \sum_{k\in\mathscr{C}^\upsilon} \frac{z_k^i}{\pi_k^i} \frac{z_l^j}{\pi_l^j}\ .\end{aligned} \tag{4.20}$$

Dabei ist $(\pi_{k,l}^{i,j} - \pi_k^i \pi_l^j) = \Delta_{\alpha\omega,\bar{\alpha}\bar{\omega}}^{t,u} \forall\, k \in \mathscr{C}^{\alpha\omega}$ und $\forall\, l \in \mathscr{C}^{\bar{\alpha}\bar{\omega}}$.

$$\mathrm{V}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) =$$

$$\mathrm{E}_{\vec{\Xi}^{\max(t,u)}}\left(\mathrm{V}\left(g^{t,u}(\hat{\tau}_z^t, \hat{\tau}_z^u) | \vec{\Xi}^{\max(t,u)}\right)\right) \tag{4.21a}$$

$$+\,\mathrm{V}_{\vec{\Xi}^{\max(t,u)}}\left(\mathrm{E}\left(g^{t,u}(\hat{\tau}_z^t, \hat{\tau}_z^u) | \vec{\Xi}^{\max(t,u)}\right)\right) \tag{4.21b}$$

mit $\vec{\Xi}^t = \left(\vec{\Xi}^t_c\right)_{c \in \mathscr{G}^t}$ und $\vec{\Xi}^t_c$ nach (2.58).

Unter Verwendung von (2.56) ist

$$\mathrm{E}(\hat{\tau}^t_z | \vec{\Xi}^t) = \sum_{\kappa \in \mathscr{G}^t} \frac{\mathfrak{n}^t_\kappa}{\mathrm{E}\,(\mathfrak{n}^t_\kappa)} \sum_{k \in \mathscr{C}^\kappa} z^t_k \, . \tag{4.22}$$

Somit ist $\mathrm{E}(\hat{\tau}^t_z | \vec{\Xi}^t) \neq \tau^t_z$, wenn $\exists\, \kappa \in \mathscr{G}^t$ mit $\mathfrak{n}^t_\kappa \neq \mathrm{E}\,(\mathfrak{n}^t_\kappa)$ und $\sum_{k \in \mathscr{C}^\kappa} z^t_k \neq 0$. Folglich ist (4.21b) nicht notwendigerweise Null. Ein Varianzschätzer

$$\widehat{\mathrm{V}}\left(g^{t,u}(\hat{\tau}^t_z, \hat{\tau}^u_z) | \vec{\Xi}^{\max(t,u)}\right) \, ,$$

der unverzerrt für (4.21a) ist, wäre in diesem Fall nicht zwingend unverzerrt für (4.20). Der Wert von (4.21b) hängt ab von den Parametern der multivariat hypergeometrischen Verteilung von $\vec{\mathfrak{n}}^t_c$, sowie von den $N_{c^{\alpha\omega}}$ und den kohortenspezifischen Mittelwerten $\bar{z}^t_\kappa = \sum_{k \in \mathscr{C}^{\alpha\omega}} z^t_k / N_{c^{\alpha\omega}}$, $\forall\ \kappa \in \mathscr{G}^t$. Eine analytische Darstellung von (4.21b) erscheint kaum praktikabel aufgrund ihrer komplexen Form, zudem müsste $\vec{\Xi}^{\max(t,u)}$ bekannt sein. Da ohnehin zur Verwendung des π-Schätzers für $\hat{\tau}^t_z$ zumindest $\mathrm{E}\,(\vec{\mathfrak{n}}^t_c)$ bekannt sein muss, was ebenfalls analytisch komplex zu beschreiben ist, wird vorgeschlagen, $\Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}}$ in (4.20) mittels einer Monte-Carlo Integration zu bestimmen. Nordberg (2000) verwendet aus dem gleichen Grund einen solchen Ansatz zur Approximation von (4.21b).

$\Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}}$ lässt sich darstellen als

$$\begin{aligned} \Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}} &= \int_{\mathscr{S}} \Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}} | \vec{\Xi}^T\,(\mathfrak{S})\, dF(\mathfrak{S}) \\ &= \sum_{\nu=1}^{card(\mathscr{S})} \Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}} | \vec{\Xi}^T\,(\mathfrak{S}_\nu)\, p(\mathfrak{S}_\nu) \, . \end{aligned}$$

Dabei ist F ist die Verteilungsfunktion des gemeinsamen Designs $p(.)$, $\mathfrak{S}_\nu$ das ν-te Element in $\mathscr{S}$ und $\vec{\Xi}^T\,(\mathfrak{S}_\nu)$ die Ausprägungen von $\vec{\Xi}^T$ für Stichprobe $\mathfrak{S}_\nu$. Die bedingte Kovarianz $\Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}} | \vec{\Xi}^T\,(\mathfrak{S}_\nu)$ kann, wie in (2.60) beschrieben, unter Anwendungen von (2.59) bestimmt werden. Eine Approximation $\Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}}$ wäre somit gegeben durch

$$\Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}} \approx \Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}} = \frac{1}{M} \sum_{i=1}^{M} \Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\varpi}} | \vec{\Xi}^T\,(\mathfrak{S}_i) \, , \tag{4.23}$$

mit $\mathfrak{S}_1, \ldots, \mathfrak{S}_i, \ldots, \mathfrak{S}_M \in \mathscr{S}$ als M unabhängige Ziehungen aus F. Dies kann durch das M-malige Wiederholen von Algorithmus 9 erfolgen. Da die M Ziehungen unabhängig voneinander erfolgen, lässt sich das computerintensive Verfahren zur Berechnung der Summe in (4.23) gut parallelisieren, was den Zeitaufwand für dieses Verfahren potentiell verringert.

Schließlich ist

$$\widehat{\mathrm{V}}\left(g^{t,u}(\hat{\tau}^t, \hat{\tau}^u)\right) \approx \sum_{i \in \{t,u\}} \sum_{j \in \{t,u\}} \sum_{\kappa \in \mathscr{G}^i} \sum_{\upsilon \in \mathscr{G}^i} \check{\Delta}^{t,u}_{\kappa,\upsilon} \sum_{k \in \mathscr{C}^\kappa} \sum_{k \in \mathscr{C}^\upsilon} \frac{z^i_k}{\pi^i_k} \frac{z^j_l}{\pi^j_l} \tag{4.24}$$

mit

$$\check{\Delta}^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\omega}} = \frac{\Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\omega}}}{\mathrm{AE}\left(\pi^{t,u}_{k,l}\right)} .$$

Dabei ist AE $\left(\mathfrak{I}^t_k\mathfrak{I}^t_l\right)$ eine Approximation zu $\pi^{t,u}_{k,l}$ und entspricht dem ersten Term in $\Delta^{t,u}_{\alpha\omega,\bar{\alpha}\bar{\omega}}$, bestimmt nach (4.23).

4.2.3 Schätzung von Veränderungen nicht linearer Statistiken

Im Folgenden soll die Linearisierung von $g^{t,u}(\hat{\theta}^u,\hat{\theta}^t)$ dargestellt werden für den Fall, dass $g^{t,u}$ stetig differenzierbar bis zweiter Ordnung an jedem Punkt in $\mathbb{R}^2$ ist, nicht jedoch $\hat{\theta}^t$ und $\hat{\theta}^u$.

Zu diesem Zweck wird das Konzept der Einflussfunktion, welches in den Abschnitten 3.5 und 3.7 vorgestellt wurde, auf den Fall mehrdimensionaler Funktionalen erweitert. Ein mehrdimensionales statistisches Funktional wird zu diesem Zweck bezeichnet mit $T(\vec{M})$, wobei $\vec{M} = \left(M^1,\ldots,M^T\right)$. Dabei ist $M^t = \sum_{i=1}^N \delta^{\circ t}_y$ mit $\delta^{\circ t}_y$ das Dirac-Maß zum Zeitpunkt t am Punkt $y \in \mathbb{R}$, mit

$$\delta^{\circ t}_k(y) := \begin{cases} 1 & \text{falls } y = y^t_k \text{ und } \ k \in \mathscr{U}^t \\ 0 & \text{sonst} \end{cases} ,$$

(Goga et al., 2009). Ein Schätzer für M^t ist gegeben durch $\hat{M}^t = \sum_{k\in\mathscr{U}^t} w^t_k \delta^{\circ t}_k(y_k)$, mit w^t_k als dem Erhebungsgewicht des k-ten Elementes zum Zeitpunkt t, gegeben durch

$$w^t_k = \begin{cases} w^t_k & \text{für } k \in \mathscr{S}^t \\ 0 & \text{sonst} \end{cases} . \tag{4.25}$$

Es wird angenommen, dass $\hat{M}^t$ konsistent für M^t ist. Ist $\theta^t = T(M^t)$, so wird die Veränderung in θ mit $T(\vec{M}) = g^{t,u}(T(M^t),T(M^u))$ beschrieben. Folglich hängt T nur von $\vec{M} = (M^t,M^u)$ ab. Als Schätzer für die Veränderung wird wiederum die Veränderung der Querschnittschätzer verwendet, d.h. $T(\vec{\hat{M}}) = g^{t,u}\left(T(\hat{M}^t),T(\hat{M}^u)\right)$ mit $\vec{\hat{M}} = (\hat{M}^t,\hat{M}^u)$.

Der Wert der Einflussfunktion der Statistik $T = T(\vec{M})$ am Punkt (y^t_k,y^u_k) ist gegeben durch die Summe der Werte der partiellen Einflussfunktionen von $T(\vec{M})$ an den Stellen y^t_k, bzw. y^u_k, (Goga et al., 2009). Die partielle Einflussfunktion von $T(\vec{M})$, $IF^t(T,\vec{M},y)$, ist gegeben durch

$$IF^t\left(T,\vec{M},y\right) = \lim_{\varepsilon\to 0} \frac{T\left(M^t+\varepsilon\delta^{\circ t}_k(y),M^u\right) - T(\vec{M})}{\varepsilon} , \tag{4.26}$$

unter der Vorrausetzung, dass der Grenzwert in (4.26) existiert, (Pires & Branco, 2002). Unter der Annahme der Asymptotik, beschrieben durch Goga et al. (2009), ist eine Approximation der Varianz von $T(\vec{\hat{M}})$ gegeben durch

$$\mathrm{V}\left(T(\vec{\hat{M}})\right) \approx \mathrm{V}\left(\sum_{i=\{t,u\}}\sum_{k\in\mathscr{U}^t} \mathfrak{I}^i_k w^i_k z^i_k\right) , \tag{4.27}$$

mit z_k^t als Wert der partiellen Einflussfunktion IF^t an der der Stelle y_k^t, d.h.

$$z_k^t = IF^t\left(T, \vec{M}, y_k^t\right) .$$

Ein Schätzer für (4.27) lässt sich beschreiben als

$$\widehat{\mathrm{V}}\left(\sum_{i=\{t,u\}}\sum_{k\in\mathscr{U}^i} \mathfrak{I}_k^i w_k^i \hat{z}_k^i\right) , \tag{4.28}$$

mit $\hat{z}_k^t = IF^t\left(T, \hat{\vec{M}}, y_k^t\right)$ als geschätzter Wert der partiellen Einflussfunktion IF^t an der Stelle y_k^t. Ist $w_k^t = 1/\pi_k^1$, nimmt (4.27) die Form (4.4) an. In Abhängigkeit von dem gemeinsamen Design können für (4.28) Schätzer gewählt werden, wie sie in Abschnitt 4.2.2 beschrieben sind. Für w_k^t als Kalibrierungsgewicht kann, entsprechend den Ausführungen in Abschnitt 3.8, ein Schätzer der folgenden Form verwendet werden

$$\widehat{\mathrm{V}}\left(\sum_{i=\{t,u\}}\sum_{k\in\mathscr{U}^i} \mathfrak{I}_k^i \frac{\hat{e}_k^i}{\pi_k^i}\right) .$$

Dabei ist $\hat{e}_k^t$ das Residuum der Regression von $\hat{z}_k^t$ auf die verwendeten Variablen zur Kalibrierung zum Zeitpunkt t.

Beispiel 4.2. Es soll die Veränderung des Indikators für von Armut oder sozialer Ausgrenzung bedrohte Personen, AROPE, beschrieben in Abschnitt 1.1 Definition 1.4, mittels dessen Differenz zwischen verschiedenen Beobachtungszeitpunkten gemessen werden. Das heißt, die interessierende Statistik ist

$$\Delta^{(l)}\left(\theta^t\right)$$

mit $\Delta^{(l)}\left(\hat{\theta}^t\right) = \hat{\theta}^t - \hat{\theta}^{t-l}$, $\theta^t = \mathrm{AROPE}^t$ oder $\theta^t = \mathrm{AROPER}^t$. Dabei ist AROPE^t der Wert des Indikators AROPE zum Zeitpunkt t und $\mathrm{AROPER}^t = \mathrm{AROPE}^t/N^t$ als der Anteil der von Armut oder sozialer Ausgrenzung bedrohten Personen in t.

AROPE^t ist ein zusammengesetzter Indikator aus den Indikatoren ARP^t, LWI^t und DEP^t definiert nach 1.1, 1.2 bzw. 1.3. Diese lassen sich darstellen als Totalwerte einer dichotomen Variablen. Stellt $\mathscr{I}_A^t$ die Menge der Personen zum Zeitpunkt t beschrieben in der entsprechenden Definitionen des Indikators A dar, so ist $\tau_A^t = card(\mathscr{I}_A^t)$ der Wert des Indikators in t, mit

$$\tau_A^t = \sum_{k\in\mathscr{U}^t} \mathbb{y}_{A,k}^t$$

und

$$\mathbb{y}_{A,k}^t = \mathbb{1}(k, \mathscr{I}_A^t)$$

mit

$$\mathbb{1}(k, \mathscr{I}_A^t) := \begin{cases} 1 & \text{für } k \in \mathscr{I}_A^t \\ 0 & \text{sonst} \end{cases} .$$

Entsprechend werden die Werte der Indikatoren APR, LWI und DEP in t mit τ_{ARP}^t, τ_{LWI}^t, sowie τ_{DEP}^t bezeichnet. Die zu den Indikatoren korrespondierenden Indikatorvariablen $\mathbb{y}_{\mathrm{ARP},k}^t = \mathbb{1}(k, \mathscr{I}_{\mathrm{APR}}^t)$, $\mathbb{y}_{\mathrm{LWI},k}^t = \mathbb{1}(k, \mathscr{I}_{\mathrm{LWI}}^t)$ und $\mathbb{y}_{\mathrm{DEP},k}^t = \mathbb{1}(k, \mathscr{I}_{\mathrm{DEP}}^t)$ sind Parameter der Population in t, jedoch lassen sich nur $\mathbb{y}_{\mathrm{LWI},k}^t$ und $\mathbb{y}_{\mathrm{DEP},k}^t$ beobachten. Wie in

Abschnitt 3.7.1 beschrieben, muss $\mathrm{y}^t_{\mathrm{ARP},k}$ geschätzt werden, da der Median des Äquivalenzeinkommens nicht bekannt ist. Die Indikatorvariable von AROPE^t ist gegeben durch

$$\mathrm{y}^t_{\mathrm{AROPE},k} = \mathrm{y}^t_{\mathrm{ARP},k} \vee \mathrm{y}^t_{\mathrm{LWI},k} \vee \mathrm{y}^t_{\mathrm{DEP},k} \,.$$

Somit lässt sich $\mathrm{y}^t_{\mathrm{AROPE},k}$ darstellen als

$$\mathrm{y}^t_{\mathrm{AROPE},k} = 1 - (1 - \mathrm{y}^t_{\mathrm{ARP},k})(1 - \mathrm{y}^t_{\mathrm{LWI},k})(1 - \mathrm{y}^t_{\mathrm{DEP},k}) \,. \tag{4.29}$$

Damit sind τ^t_{AROPE} und τ^t_{AROPER} gegeben durch

$$\begin{aligned} \tau^t_{AROPE} &= \sum_{k \in \mathscr{U}^t} \mathrm{g}^t_k \mathrm{y}^t_{\mathrm{ARP},k} + (1 - \mathrm{g}^t_k) \,, \\ \tau^t_{AROPER} &= \sum_{k \in \mathscr{U}^t} \frac{\mathrm{g}^t_k \mathrm{y}^t_{\mathrm{ARP},k} + (1 - \mathrm{g}^t_k)}{N^t} \,, \end{aligned}$$

mit

$$\mathrm{g}^t_k = 1 - \mathrm{y}^t_{\mathrm{LWI},k} - \mathrm{y}^t_{\mathrm{DEP},k} + \mathrm{y}^t_{\mathrm{LWI},k} \mathrm{y}^t_{\mathrm{DEP},k} \,.$$

Als statistische Funktionale lassen sich AROPE^t und AROPER^t darstellen als

$$\begin{aligned} \mathrm{AROPE}^t(M^t) &= \int^{0{,}6\mathrm{MED}^t(M^t)} \mathrm{g}^t dM^t + \int (1 - \mathrm{g}^t) \\ &= \sum_{k \in \mathscr{U}^t} \mathrm{g}^t_k \mathbb{1}[y^t_k \leq 0{,}6\mathrm{MED}^t(M^t)] + \sum_{k \in \mathscr{U}^t} (1 - \mathrm{g}^t_k) dM^t \,, \\ \mathrm{AROPER}^t(M^t) &= \frac{\mathrm{AROPE}^t(M^t)}{\int dM^t} \\ &= \frac{\mathrm{AROPE}^{\,t}(M^t)}{N^t} \,, \end{aligned}$$

mit y^t_k als dem verfügbaren Äquivalenzeinkommen (nach Sozialtransfers) des k-ten Elements zum Zeitpunkt t[1] und $\mathrm{MED}^t(M^t)$ als dem Median des Äquivalenzeinkommens zum Zeitpunkt t.

Des Weiteren soll $T(\vec{M}) = \Delta^{(l)}(\mathrm{AROPE}^t)$ und $R(\vec{M}) = \Delta^{(l)}(\mathrm{AROPER}^t)$, mit $\vec{M} = (M^{t-l}, M^t)$, definiert sein. Zudem sei $T(\vec{\hat{M}})$ ein Schätzer für $T(\vec{M})$, mit $\vec{\hat{M}} = (\hat{M}^{t-l}, \hat{M}^t)$ und $\hat{M}^i = \sum_{k \in \mathscr{U}^i} 1/\pi^i_k$, für $i = t, t-l$. Entsprechendes gilt für $R(\vec{M})$. Für die beiden partiellen Einflussfunktionen von $T(\vec{M})$, $IF^t(T, \vec{M}, y)$ und $IF^{t-l}(T, \vec{M}, y)$, gilt $-IF^t(T, \vec{M}, y)$ $= IF^{t-l}(T, \vec{M}, y)$ und IF^t ist die Einflussfunktion von $\mathrm{AROPE}^t(M^t)$. Gleiches gilt auch für $R(\vec{M})$.

Es wird vorgeschlagen die Einflussfunktion von $\mathrm{AROPE}^t(M^t)$ durch eine Modifikation des von Langel & Tillé (2011) verwendeten Ansatzes zur Herleitung der Einflussfunktion der QSR aufstellen, was sich wie folgt darstellt

$$\begin{aligned} IF(\mathrm{AROPE}^t, M^t, y^t_k) = & \mathrm{g}^t_k \mathbb{1}[y^t_k \leq 0{,}6\mathrm{MED}^t(M^t)] - \\ & (0{,}6)^2 \mathrm{MED}^t(M^t)(\mathbb{1}[y^t_k \leq \mathrm{MED}^t(M^t)] - 0{,}5) + (1 - \mathrm{g}^t_k) \,, \end{aligned}$$

---

[1] Eurostat, 2015, `http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Equivalised_disposable_income/de`

entsprechend ist

$$IF(\text{AROPER}^t, M^t, y_k^t) = \left(IF(\text{AROPE}^t, M^t, y_k^t) - \text{ARPER}^t\right) \frac{1}{N^t} \ .$$

Schließlich lässt sich die approximative Varianz der Schätzer für Veränderung von AROPE^t und AROPER^t zwischen den Zeitpunkten $t-l$ und t als

$$\begin{aligned} \mathrm{V}\left(T(\vec{M})\right) \approx & \mathrm{V}\left(\sum_{i=\{t,t-l\}} \sum_{k\in\mathscr{U}^i} \mathfrak{I}_k^i \frac{z_k^i}{\pi_k^i}\right) \text{ bzw.} \\ \mathrm{V}\left(R(\vec{M})\right) \approx & \mathrm{V}\left(\sum_{i=\{t,t-l\}} \sum_{k\in\mathscr{U}^i} \mathfrak{I}_k^i \frac{u_k^i}{\pi_k^i}\right) \end{aligned}$$

schreiben, mit

$$\begin{aligned} z_k^t =& IF^t(T,\vec{M},y_k^t) = & IF(\text{AROPE}^t, M^t, y_k^t) \ , \\ z_k^{t-l} =& -IF^t(T,\vec{M},y_k^{t-l}) = & -IF(\text{AROPE}^t, M^t, y_k^{t-l}) \ , \\ u_k^t =& IF^t(R,\vec{M},y_k^t) = & IF(\text{AROPER}^t, M^t, y_k^t) \ , \\ u_k^{t-l} =& -IF^t(R,\vec{M},y_k^{t-l}) = & -IF(\text{AROPER}^t, M^t, y_k^{t-l}) \ . \end{aligned}$$

$\triangle$

Kapitel 5

Abschließende Bemerkungen

5.1 Stichproben Designs

5.1.1 Unveränderliche Populationen

Der in Abschnitt 2.4.2 vorgestellte Algorithmus 3 zur Koordination von SRS hat, in Abhängigkeit von der Wahl der Koordinationsvariable, ähnliche Eigenschaften wie die Koordination von SRS mittels PRN. So kann in beiden Fällen dasselbe Querschnittdesign und die gleiche Verteilung der Erhebungslast erzielt werden. Ein Unterschied zwischen den beiden Verfahren findet sich bezüglich ihrer Längsschnittdesigns. So ist ϕ_k^t, gegeben der PRN von Element k, keine Zufallsvariable. Dies war nicht der Fall für Algorithmus 3 unter den beiden untersuchten Varianten. Eine höhere Varianz von ϕ_k^t kann Vor- und Nachteile mit sich bringen. Da $\mathfrak{n}^{t,u} = \sum_{k\mathscr{U}} \mathfrak{I}_k^t \mathfrak{I}_k^u$ für die PRN Koordination von SRS keine Zufallsvariable ist, sind die $\pi_k^{t,u}$ einfach zu bestimmen als $\pi_k^{t,u} = \mathfrak{n}^{t,u}/N$, womit auch $\pi_{k,l}^{t,u}$, nach (2.34), gegeben ist. Dies gilt nicht für Algorithmus 3. Jedoch bietet dieser die Möglichkeit, Elemente in Zeitpunktkombinationen zu beobachten, die unter der PRN Koordination ausgeschlossen sind. Dies kann ein wünschenswerter Effekt sein, gerade wenn es von Interesse ist, dass Veränderungen zwischen weit auseinanderliegenden Zeitpunkten gemessen werden sollen, deren Querschnittstichproben unter der PRN Koordination keine Überlappung aufweisen können. Dieser Effekt kommt schneller zum Tragen bei hohen Auswahlsätzen, was die Methode möglicherweise interessant macht für die Koordination von Querschnittstichproben, sog. primären Ziehungseinheiten eines mehrstufigen Stichprobendesigns.

Eine weitere Variante für eine Koordinationsvariable von Algorithmus 3 wäre, nur die Belastungen $\mathfrak{b}_k^t$ zu verwenden. Unter der Annahme einer gleichen Erhebungslast für jede Ziehung würde sich das Längsschnittdesign von Algorithmus 3 als Devilles' Systematisches Design (Tillé, 2006, S. 128ff) darstellen. Eine solcher Koordination über die Erhebungslast wird in Nedyalkova et al. (2009) beschrieben. Obwohl aufeinanderfolgende Querchnittstichproben bei dieser Koordination ebenfalls negativ koordiniert sind, kann ein Element, auch für $N > n^t + n^{t+1}$, zu den Zeitpunkten t und $t+1$ gezogen werden, da bei dieser Variante $\pi_k^{t,u} > 0$ für alle $t,u \in \mathscr{T}$ ist. Wird die Koordinationsvariable in Algorithmus 3 bestimmt als $\mathfrak{o}_k^t = \mathfrak{o}_k^{t-1} - \mathfrak{I}_k^t$ und $\mathfrak{o}_k^0 = u_k$ mit $u_k \sim \mathrm{Unif}(0,1)$, setzten Algorithmus 3 und 8 das identische gemeinsame Design um. Tabelle 5.1 stellt die Rangfolge der verschiedenen Varianten der Koordinationsvariable bezüglich der Größe des Trägers des Längsschnittdesign von Algorithmus 3 dar. Diese Rangfolge korrespondiert auch zur Höhe der Varianz von ϕ_k^t.

Tabelle 5.1: Trägergrößen der Längsschnittdesigns verschiedener Varianten von Algorithmus 3

Rang	$\mathfrak{o}_k^t$	Variante
1.	$\mathfrak{b}_k^t$	Erhebungslast
2.	$\mathfrak{p}_k^t$	Zeit außerhalb der Stichprobe
3.	$(\mathfrak{b}_{\max}^t - \mathfrak{b}_{\min}^t + 1)\mathfrak{p}_k^t - \mathfrak{b}_k^t$	Zeit außerhalb der Stichprobe und Erhebungslast
4.	$u_k - \sum_{i=1}^t \mathfrak{I}_k^i$	PRN

Weitere Varianten für die Koordinationsvariable von Algorithmus 3 sind denkbar. Damit aber die Bedingungen (2.23) erfüllt bleiben, darf die Häufigkeitsverteilung der $\mathfrak{o}_k^t$ $(k = 1,\ldots,N)$ nicht von der gemeinsamen Stichprobe abhängen. Gelten die Bedingungen 2.23, sind die ϕ_k^t iid für $(k = 1,\ldots,N)$, was wiederum die Berechnung der $\pi_k^{t,u}$ unter Verwendung der Verteilungen von $\phi_k^i(i = \min(t,u),\ldots,\max(t,u))$, wie durch Algorithmus 5 beschrieben, stark erleichtert.

5.1.2 Veränderliche Populationen

Für veränderliche Populationen sind die π_k^t von Algorithmus 9 nicht gleich für alle $k \in \mathscr{U}^t$. Unter der Annahme, dass n^t und N^t simultan gegen unendlich streben und $N_{c^{\alpha\omega}}/N^t$ für alle Kohorten $\mathscr{C}^{\alpha\omega}$ sowie n^t/N^t konstant bleiben (siehe z.B. Stenger, 1989 für eine solche Asymptotik) ist

$$\mathrm{E}\left(\mathfrak{n}_{c^{\alpha,\omega}}^t\right) = \frac{N_{c^{\alpha,\omega}}}{N^t} n^t$$

und somit ist auch

$$\pi_k^t = \frac{n^t}{N^t} \; .$$

Dies bedeutet, dass sich asymptotisch das Querschnittdesign von Algorithmus 9 dem einer einfachen Zufallsstichprobe annähert. Die Geschwindigkeit der Konvergenz der π_k^t, für $k \in \mathscr{C}^{\alpha\omega}$ ist noch zu untersuchen.

Allgemein ist bei einer veränderlichen Population abzuwägen zwischen festen Stichprobenumfängen im Querschnitt und der Komplexität des gemeinsamen Stichprobendesigns. So können auch die Längsschnittdesigns von Algorithmus 3 mit (2.14) oder (2.36) bei veränderlichen Populationen unabhängig voneinander implementiert werden. Hierzu könnten die Algorithmen 4 bzw. 6 verwendet werden, mit $\pi_k^t = n^t/N^t$ für alle $k \in \mathscr{U}$. Die gemeinsame Stichprobe wäre dann gegeben durch $\mathfrak{S} \circ \mathbf{U}$. Dabei ist $\mathbf{U}$ die Indikatormatrix nach (2.64). Die Bestimmung der $\pi_k^{t,u}$ und $\pi_{k,l}^{t,u}$ würde sich dann gestalten wie im Falle einer unveränderlichen Population. Zufällige Stichprobenumfänge sind auch bei Erhebungen der Sozialwissenschaften anzutreffen. So werden beispielsweise für die EU-SILC Erhebung Daten zu allen Personen in den gezogenen Haushalten gesammelt, einer der Gründe, warum die Stichprobenumfänge auf Personenebene variabel sind.

5.2 Schätzung

Brewer et al. (1972) schlagen bei variablen Stichprobenumfängen im Querschnitt, für $\mathfrak{n}^t > 0$, die Verwendung des Verhältnisschätzers

$$\hat{\tau}^t_r = \sum_{k=\mathcal{U}^t} \mathfrak{I}^t_k \frac{\mathrm{E}\left(\mathfrak{n}^t\right)}{\mathfrak{n}^t} \frac{y^t_k}{\pi^t_k} ,$$

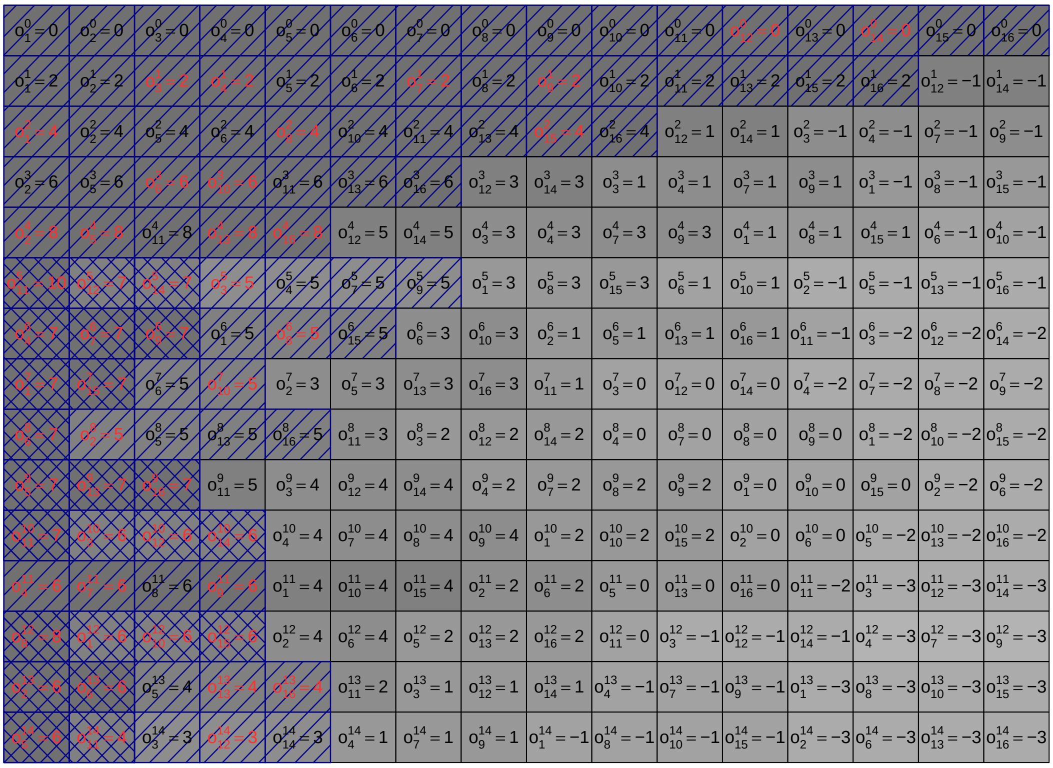
anstelle des π-Schätzers vor.

Unabhängig von variablen oder fixen Stichprobenumfängen ist es möglich, die vorhandenen Informationen aus der gemeinsamen Stichprobe besser zu nutzten, um effizientere Schätzer für Querschnitte zu erhalten. In den Kapitel 3 und 4 wurden nur separierte Querschnittschätzer betrachtet, d.h. Schätzer, die für sich genommen nur auf den Daten einer Querschnittstichprobe beruhen. Es ist aber möglich, die Korrelationsstruktur des Populationsparameter $\boldsymbol{Y}$ auszunutzen, um so effizientere Querschnittschätzer zu erhalten. Goga (2003) macht hierzu ausführliche Angaben für die Schätzung linearer Statistiken (hierzu siehe auch Särndal et al., 1992, Abschnitt 9.9), insbesondere auch zur Schätzung statistischer Funktionale (siehe auch Goga et al., 2009).

Ein gewichtiges Problem bei der Varianzschätzung von Veränderungen mit negativ koordinierten Stichproben besteht in nicht überlappenden Querschnittstichproben. Resampling Verfahren könnten hier möglicherweise zur Approximation der Varianz eines Schätzers für Veränderungen verwendet werden. Chipperfield & Preston (2007) verwenden eine Bootstrap Methode zur Schätzung der Varianz einer Differenz zwischen Totalwertschätzern. Das betrachtete gemeinsame Design sind im Querschnitt PRN koordinierte SRS, die jedoch eine Überlappung aufweisen. Einige experimentelle Stimulationsstudien mit der Bootstrap Methode deuten daraufhin, dass zumindest für nicht überlappende SRS eine Approximation der Varianz einer Differenz zwischen linearen Schätzern möglich sein könnte. Die Asymptotik eines solchen Ansatzes bedarf aber weiterer Untersuchungen. Alternativ sind auch modellbasierte Ansätze zur Schätzung der Korrelationsstruktur von $\boldsymbol{Y}$ möglich, wie sie zum Beispiel von Skinner & Holmes (2003) verwendet werden.

Anhang A

Abbildung A.1: Werte der Koordinationsvariable für $\mathfrak{o}_k^t = (\mathfrak{b}_{\max}^t - \mathfrak{b}_{\min}^t + 1)\mathfrak{p}_k^t - \mathfrak{b}_k^t$

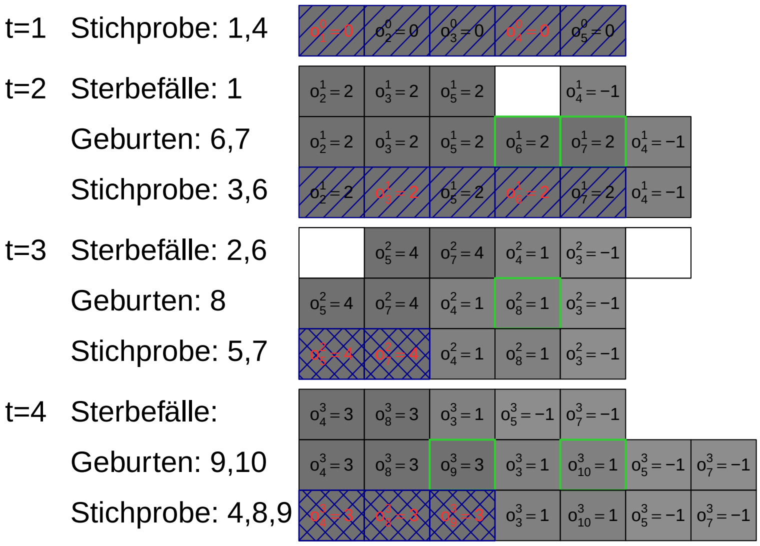


$o_{1}^{0}=0$	$o_{2}^{0}=0$	$o_{3}^{0}=0$	$o_{4}^{0}=0$	$o_{5}^{0}=0$	$o_{6}^{0}=0$	$o_{7}^{0}=0$	$o_{8}^{0}=0$	$o_{9}^{0}=0$	$o_{10}^{0}=0$	$o_{11}^{0}=0$	$o_{12}^{0}=0$	$o_{13}^{0}=0$	$o_{14}^{0}=0$	$o_{15}^{0}=0$	$o_{16}^{0}=0$
$o_{1}^{1}=2$	$o_{2}^{1}=2$	$o_{3}^{1}=2$	$o_{4}^{1}=2$	$o_{5}^{1}=2$	$o_{6}^{1}=2$	$o_{7}^{1}=2$	$o_{8}^{1}=2$	$o_{9}^{1}=2$	$o_{10}^{1}=2$	$o_{11}^{1}=2$	$o_{13}^{1}=2$	$o_{15}^{1}=2$	$o_{16}^{1}=2$	$o_{12}^{1}=-1$	$o_{14}^{1}=-1$
$o_{1}^{2}=4$	$o_{2}^{2}=4$	$o_{5}^{2}=4$	$o_{6}^{2}=4$	$o_{8}^{2}=4$	$o_{10}^{2}=4$	$o_{11}^{2}=4$	$o_{13}^{2}=4$	$o_{15}^{2}=4$	$o_{16}^{2}=4$	$o_{12}^{2}=1$	$o_{14}^{2}=1$	$o_{3}^{2}=-1$	$o_{4}^{2}=-1$	$o_{7}^{2}=-1$	$o_{9}^{2}=-1$
$o_{2}^{3}=6$	$o_{5}^{3}=6$	$o_{6}^{3}=6$	$o_{10}^{3}=6$	$o_{11}^{3}=6$	$o_{13}^{3}=6$	$o_{16}^{3}=6$	$o_{12}^{3}=3$	$o_{14}^{3}=3$	$o_{3}^{3}=1$	$o_{4}^{3}=1$	$o_{7}^{3}=1$	$o_{9}^{3}=1$	$o_{1}^{3}=-1$	$o_{8}^{3}=-1$	$o_{15}^{3}=-1$
$o_{2}^{4}=8$	$o_{5}^{4}=8$	$o_{11}^{4}=8$	$o_{13}^{4}=8$	$o_{16}^{4}=8$	$o_{12}^{4}=5$	$o_{14}^{4}=5$	$o_{3}^{4}=3$	$o_{4}^{4}=3$	$o_{7}^{4}=3$	$o_{9}^{4}=3$	$o_{1}^{4}=1$	$o_{8}^{4}=1$	$o_{15}^{4}=1$	$o_{6}^{4}=-1$	$o_{10}^{4}=-1$
$o_{11}^{5}=10$	$o_{12}^{5}=7$	$o_{14}^{5}=7$	$o_{3}^{5}=5$	$o_{4}^{5}=5$	$o_{7}^{5}=5$	$o_{9}^{5}=5$	$o_{1}^{5}=3$	$o_{8}^{5}=3$	$o_{15}^{5}=3$	$o_{6}^{5}=1$	$o_{10}^{5}=1$	$o_{2}^{5}=-1$	$o_{5}^{5}=-1$	$o_{13}^{5}=-1$	$o_{16}^{5}=-1$
$o_{4}^{6}=7$	$o_{7}^{6}=7$	$o_{9}^{6}=7$	$o_{1}^{6}=5$	$o_{8}^{6}=5$	$o_{15}^{6}=5$	$o_{6}^{6}=3$	$o_{10}^{6}=3$	$o_{2}^{6}=1$	$o_{5}^{6}=1$	$o_{13}^{6}=1$	$o_{16}^{6}=1$	$o_{11}^{6}=-1$	$o_{3}^{6}=-2$	$o_{12}^{6}=-2$	$o_{14}^{6}=-2$
$o_{1}^{7}=7$	$o_{15}^{7}=7$	$o_{6}^{7}=5$	$o_{10}^{7}=5$	$o_{2}^{7}=3$	$o_{5}^{7}=3$	$o_{13}^{7}=3$	$o_{16}^{7}=3$	$o_{11}^{7}=1$	$o_{3}^{7}=0$	$o_{12}^{7}=0$	$o_{14}^{7}=0$	$o_{4}^{7}=-2$	$o_{7}^{7}=-2$	$o_{8}^{7}=-2$	$o_{9}^{7}=-2$
$o_{6}^{8}=7$	$o_{2}^{8}=5$	$o_{5}^{8}=5$	$o_{13}^{8}=5$	$o_{16}^{8}=5$	$o_{11}^{8}=3$	$o_{3}^{8}=2$	$o_{12}^{8}=2$	$o_{14}^{8}=2$	$o_{4}^{8}=0$	$o_{7}^{8}=0$	$o_{8}^{8}=0$	$o_{9}^{8}=0$	$o_{1}^{8}=-2$	$o_{10}^{8}=-2$	$o_{15}^{8}=-2$
$o_{5}^{9}=7$	$o_{13}^{9}=7$	$o_{16}^{9}=7$	$o_{11}^{9}=5$	$o_{3}^{9}=4$	$o_{12}^{9}=4$	$o_{14}^{9}=4$	$o_{4}^{9}=2$	$o_{7}^{9}=2$	$o_{8}^{9}=2$	$o_{9}^{9}=2$	$o_{1}^{9}=0$	$o_{10}^{9}=0$	$o_{15}^{9}=0$	$o_{2}^{9}=-2$	$o_{6}^{9}=-2$
$o_{11}^{10}=7$	$o_{3}^{10}=6$	$o_{12}^{10}=6$	$o_{14}^{10}=6$	$o_{4}^{10}=4$	$o_{7}^{10}=4$	$o_{8}^{10}=4$	$o_{9}^{10}=4$	$o_{1}^{10}=2$	$o_{10}^{10}=2$	$o_{15}^{10}=2$	$o_{2}^{10}=0$	$o_{6}^{10}=0$	$o_{5}^{10}=-2$	$o_{13}^{10}=-2$	$o_{16}^{10}=-2$
$o_{4}^{11}=6$	$o_{7}^{11}=6$	$o_{8}^{11}=6$	$o_{9}^{11}=6$	$o_{1}^{11}=4$	$o_{10}^{11}=4$	$o_{15}^{11}=4$	$o_{2}^{11}=2$	$o_{6}^{11}=2$	$o_{5}^{11}=0$	$o_{13}^{11}=0$	$o_{16}^{11}=0$	$o_{11}^{11}=-2$	$o_{3}^{11}=-3$	$o_{12}^{11}=-3$	$o_{14}^{11}=-3$
$o_{8}^{12}=8$	$o_{1}^{12}=6$	$o_{10}^{12}=6$	$o_{15}^{12}=6$	$o_{2}^{12}=4$	$o_{6}^{12}=4$	$o_{5}^{12}=2$	$o_{13}^{12}=2$	$o_{16}^{12}=2$	$o_{11}^{12}=0$	$o_{3}^{12}=-1$	$o_{12}^{12}=-1$	$o_{14}^{12}=-1$	$o_{4}^{12}=-3$	$o_{7}^{12}=-3$	$o_{9}^{12}=-3$
$o_{2}^{13}=6$	$o_{6}^{13}=6$	$o_{5}^{13}=4$	$o_{13}^{13}=4$	$o_{16}^{13}=4$	$o_{11}^{13}=2$	$o_{3}^{13}=1$	$o_{12}^{13}=1$	$o_{14}^{13}=1$	$o_{4}^{13}=-1$	$o_{7}^{13}=-1$	$o_{9}^{13}=-1$	$o_{1}^{13}=-3$	$o_{8}^{13}=-3$	$o_{10}^{13}=-3$	$o_{15}^{13}=-3$
$o_{5}^{14}=6$	$o_{11}^{14}=4$	$o_{3}^{14}=3$	$o_{12}^{14}=3$	$o_{14}^{14}=3$	$o_{4}^{14}=1$	$o_{7}^{14}=1$	$o_{9}^{14}=1$	$o_{1}^{14}=-1$	$o_{8}^{14}=-1$	$o_{10}^{14}=-1$	$o_{15}^{14}=-1$	$o_{2}^{14}=-3$	$o_{6}^{14}=-3$	$o_{13}^{14}=-3$	$o_{16}^{14}=-3$

Abbildung A.1 soll die Auswahlregeln von Algorithmus 3 mit (2.36) verdeutlichen. Die einzelnen Zellen enthalten die Werte der Koordinationvariable, für jedes Element in der Population, vor der Stichprobenziehung zu den jeweiligen Beobachtungszeitpunkten. Dabei steht eine Reihe für einen Beobachtungszeitpunkt. Die Elemente sind Absteigend nach dem Wert ihrer Koordinationvariable von Rechts nach Links geordnet. Einfach blau staffierte Zellen markieren Gruppen von Elementen mit gleichen Werten der Koordinationvariable, aus welchen zufällig Elemente gezogen werden. Zweifach blau staffierte Zellen markieren Gruppen von Elementen mit gleichen Werten der Koordinationvariable, die vollständig ausgewählt werden. Mit Rot sind jeweils die Element gekennzeichnet, die in die Stichprobe gelangen.

Abbildung A.2 soll die Auswahlregeln von Algorithmus 9 mittels einer konkreten Stichprobenziehung verdeutlichen. Es werden zu vier Beobachtungszeitpunkten Querschnittstichproben mit den Stichprobenumfängen $n^1 = n^2 = n^3 = 2$ und $n^4 = 3$ gezogen. Die beobachtet Population ist die gleiche wie in Beispiel 2.16. Die Darstellungsweise des Algorithmus ist identisch zu jener in Abbildung A.1, mit dem Unterschied, dass vor der Markierung der gezogenen Elemente und Ziehungsgruppen zuerst die Sterbefälle entfernt werden und Geburten hinzugefügt werden. Die Zellen von Geburten sind Grün umrandet.

Abbildung A.2: Werte der Koordinationsvariable für $\mathfrak{o}_k^t = (\mathfrak{b}^t_{\max} - \mathfrak{b}^t_{\min} + 1)\mathfrak{p}^t_k - \mathfrak{b}^t_k$ bei einer veränderlichen Population



In Zeitpunkt $t = 1$ wird eine einfache Zufallsstichprobe aus den fünf vorhandenen Elementen gezogen. Vor der Ziehung in Zeitpunkt $t = 2$ wird ein Sterbefall, dass Element mit $k = 1$, entfernt. Die Elemente $k = 6$ und $k = 7$ treten in die Population zum Zeitpunkt $t = 2$ ein und ihnen wird jeweils als Wert für die Koordinationsvariable $o_k^1 = 2$ zugewiesen. Damit befinden sich die Geburten in der selben Gruppe wie die Elemente $k = 2, 3, 5$, aus welcher zwei Element zufällige entnommen werden. Dieser Vorgang wird weiter zwei Mal wiederholt für die entsprechende Anzahl von Geburten und Sterbefälle, sowie Stichprobenumfänge, wobei es in $t = 4$ keine Sterbefälle gibt.

Literaturverzeichnis

Andersson, C. & Nordberg, L. (1994). A Method for Variance Estimation of Non-Linear Functions of Totals in Surveys - Theory and Software Implementation. *Journal of Official Statistics*, *10* (4), 395–405.

Berger, Y. (1998). Rate of convergence for asymptotic variance of the Horvitz-Thompson estimator. *Journal of Statistical Planning and Inference*, *74*, 149–168.

Berger, Y. (2003). A Modified Hájek Variance Estimator for Systematic Sampling. *Statistics in Transition*, *6* (1), 5–21.

Berger, Y. (2004a). A Simple Variance Estimator for Unequal Probability Sampling without Replacement. *Journal of Applied Statistics*, *31* (3), 305–315.

Berger, Y. (2004b). Variance estimation for measures of change in probability sampling. *The Canadian Journal of Statistics. La Revue Canadienne de Statistique*, *32* (4), 451–467.

Berger, Y. (2005). A Variance Estimator for Systematic Sampling from a Deliberately Ordered Population. *Communications in Statistics: Theory and Methods*, *34* (7), 1533–1541.

Berger, Y. & Priam, R. (2015). *A simple variance estimator of change for rotating repeated surveys: an application to the EU-SILC household surveys.* Zugriff auf `http://eprints.soton.ac.uk/347142/`

Berger, Y. & Skinner, C. (2004). *Variance Estimation for Unequal Probability Designs* (Research Project Report Nr. WP6 – D6.1). DACSEIS deliverable D6.1. Zugriff auf `http://www.dacseis.de`

Binder, D. A. & Kovacevic, M. S. (1993). Estimating Some Measures of Income Inequality from Survey Data: An Application of the Estimating Equation Approach. *Journal of the American Statistical Association* (550), 550–555.

Binder, D. A. & Patak, Z. (1994). Use of estimating functions for estimation from complex surveys. *Journal of the American Statistical Association*, *89* (427), 1035–1043.

Brewer, K. (2002). *Combined Survey Sampling Inference*. New York: Arnold.

Brewer, K. & Donadio, M. (2003). The High Entropy Variance of the Horvitz-Thompson Estimator. *Survey Methodology*, *29* (2), 189–196.

Brewer, K., Early, L. & Joyce, S. (1972). Selecting Several Samples from a Single Population. *Australian Journal of Statistics*, *14* (3), 231–239.

Bruch, C., Münnich, R. & Zins, S. (2011). *Variance Estimation for Complex Surveys* (Research Project Report Nr. WP3 – D3.1). FP7-SSH-2007-217322 AMELI. Zugriff auf `http://ameli.surveystatistics.net`

Chipperfield, J. & Preston, J. (2007). Efficient bootstrap for business surveys. *Survey Methodology*, *33* (2), 167–172.

Cochran, W. G. (1977). *Sampling Techniques*. New York: Wiley.

Dell, F. & d'Haultfoeuille, X. (2008). Measuring the Evolution of Complex Indicators: Theory and Application to the Poverty Rate in France. *Annals of Economics and Statistics*, *90*, 259–290.

Deville, J.-C. (1999). Variance estimation for complex statistics and estimators: Linearization and residual techniques. *Survey Methodology*, *25* (2), 193–203.

Deville, J.-C. & Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, *87* (418), 376–382.

Deville, J.-C., Särndal, C.-E. & Sautory, O. (1993). Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, *88* (423), 1013–1020.

Deville, J.-C. & Tillé, Y. (2000). Selection of several unequal probability samples from the same population. *Journal of Statistical Planning and Inference*, *86* (2), 215–227.

Deville, J.-C. & Tillé, Y. (2005). Variance approximation under balanced sampling. *Journal of Statistical Planning and Inference*, *128* (2), 569–591.

Durbin, J. (1953). Some results in sampling theory when the units are selected with unequal probabilities. *Journal of the Royal Statistical Society, Series B.*, *15* (2), 262–269.

Europäische Kommission. (2011). *Europa-2020-Ziele.* `http://ec.europa.eu/europe2020/targets/eu-targets/index_de.htm`. ([Online; zugegriffen 29.05.2015])

Eurostat. (2014a). *Glossary: At risk of poverty or social exclusion (AROPE)* . `http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:At_risk_of_poverty_or_social_exclusion_%28AROPE%29`. ([Online; zugegriffen 29.05.2015])

Eurostat. (2014b). *Glossary: At-risk-of-poverty rate.* `http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:At-risk-of-poverty_rate`. ([Online; zugegriffen 29.05.2015])

Eurostat. (2014c). *Glossary: Material deprivation.* `http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Severe_material_deprivation_rate`. ([Online; zugegriffen 29.05.2015])

Eurostat. (2014d). *Glossary: Persons living in households with low work intensity.* `http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Persons_living_in_households_with_low_work_intensity`. ([Online; zugegriffen 29.05.2015])

Eurostat. (2015). *EU statistics on income and living conditions (EU-SILC) methodology* . `http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_%28EU-SILC%29_methodology`. ([Online; zugegriffen 29.05.2015])

Forsman, G. & Gåras. (1982). Optimal estimation of change in sample surveys. In *Proceedings of the survey research methods section.*

Gabler, S. (1984). On unequal probability sampling: sufficient conditions for the superiority of sampling without replacement. *Biometrika*, *71* (1), 171–175.

Goga, C. (2003). *Estimation de la variance dans les sondages à plusieurs échantillons et prise en compte de l'information auxiliaire par des modèles nonparamétriques* (Unveröffentlichte Dissertation). Université de Rennes II, Haute Bretagne, France.

Goga, C., Deville, J.-C. & Ruiz-Gazen, A. (2009). Use of functionals in linearization and composite estimation with application to two-sample survey data. *Biometrika*, *96* (3), 691–709.

Hájek, J. (1964). Asymptotic Theory of Rejective Sampling with Varying Probabilities from a Finite Population. *The Annals of Mathematical Statistics*, *35* (4), 1491–1523.

Hájek, J. (1981). *Sampling from a finite population.* New York: Dekker.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, *69*, 383–393.

Hansen, M. H. & Hurwitz, W. N. (1943). On the Theory of Sampling from Finite Populations. *The Annals of Mathematical Statistics*, *14* (4), 333–362.

Hartley, H. O. & Rao, J. N. K. (1962). Sampling with unequal probabilities and without replacement. *Ann. Math. Statist.*, *33*, 350–374.

Hesse, C. (1999). *Sampling Co-ordination: A Review by country* (Methodologies and Working Papers Nr. E9908). Institut national de la statistique et des études économiques. Zugriff auf `http://www.insee.fr/fr/publications-et-services/docs_doc_travail/e9908.pdf`

Hidiroglou, M. A., Särndal, C.-E. & Binder, D. A. (1995). Weighting and Estimation in BusinessSurveys. In B. G. Cox, D. A. Binder, B. N. Chinnappa, A. Christianson, M. J. Colledge & P. S. Kott (Hrsg.), *Business survey methods* (S. 477–502). Wiley.

Horvitz, D. G. & Thompson, D. J. (1952). A Generalization of Sampling Without Replacement From a Finite Universe. *Journal of the American Statistical Association*, *47* (260), 663–685.

Hulliger, B. (1995). *Konjunkturelle Mietpreiserhebung: Stichprobenplan und Schätzverfahren* (Methodenbericht). Bundesamt für Statistik, Schweiz.

Hulliger, B., Alfons, A., Bruch, C., Filzmoser, P., Graf, M., Kolb, J.-P., … Zins, S. (2011). *Report on the Simulation Results* (Research Project Report Nr. WP7 – D7.1). FP7-SSH-2007-217322 AMELI. Zugriff auf `http://ameli.surveystatistics.net`

Hulliger, B., Alfons, A., Filzmoser, P., Meraner, A., Schoch, T. & Templ, M. (2011). *Robust Methodology for Laeken Indicators* (Research Project Report Nr. WP4 – D4.2). FP7-SSH-2007-217322 AMELI. Zugriff auf `http://ameli.surveystatistics.net`

Hulliger, B. & Münnich, R. (2006). Variance Estimation for Complex Surveys in the Presence of Outliers. In *Proceedings of the American Statistical Association, Survey Research Methods Section* (S. 3153–3161).

Hulliger, B. & Schoch, T. (2014). Robust, distribution-free inference for income share ratios under complex sampling. *Advances in Statistical Analysis*, *98* (1), 63–85.

Jones, M. C., Marron, J. S. & Sheather, S. J. (1996). A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, *91* (433), 401–407.

Knottnerus, P. & van Delden, A. (2012). On Variances of Change estimated from Rotating Panels and Dynamic Strata. *Survey Methodology*, *38* (1), 43–52.

Kovacevic, M. S. & Binder, D. A. (1997). Variance Estimation for Measures of Income Inequality and Polarization - The Estimating Equations Approach. *Journal of Official Statistics*, *13* (1), 41–58.

Kröger, H., Särndal, C.-E. & Teikari, I. (1999). Poisson Mixture Sampling: A family of Designs for Coordinated Selection using Permanent Random Numbers. *Survey Methodology*, *25* (1), 3–11.

Kröger, H., Särndal, C.-E. & Teikari, I. (2003). Poisson Mixture Sampling Combined with Order Sampling. *Journal of Offical Statistics*, *19* (1), 59–70.

Langel, M. & Tillé, Y. (2011). Statistical inference for the quintile share ratio. *Journal of Statistical Planning and Inference*, *141* (8), 2976–2985.

Laniel, N. (1987). Variances for a Rotating Sample from a Changing Population. In *Proceedings of the American Statistical Association, Survey Research Methods Section.*

Matei, A. & Skinner, C. (2009). Optimal Sample Coordination Using Controlled Selection. *Journal of Statistical Planningand Inference*, *139*, 3112–3121.

Matei, A. & Tillé, Y. (2005a). Evaluation of Variance Approximations and Estimators in Maximum Entropy Sampling with Unequal Probability and Fixed Sample Size. *Journal of Official Statistics*, *21* (4), 543–570.

Matei, A. & Tillé, Y. (2005b). Maximal and Minimal Sample Co-ordination. *The Indian Journal of Statistics*, *67* (3), 590–512.

Matei, A. & Tillé, Y. (2007). Computational aspects of order πps sampling schemes. *Computational Statistics & Data Analysis*, *51*, 3703–3717.

Meyberg, K. (1980). *Algebra – Teil 1* (2. Aufl.). Carl Hanser Verlag.

Münnich, R. & Zins, S. (2011). *Variance Estimation for Indicators of Poverty and Social Exclusion* (Research Project Report Nr. WP3 – D3.2). FP7-SSH-2007-217322 AMELI. Zugriff auf `http://ameli.surveystatistics.net`

Nedyalkova, D. (2009). *Evaluation and Development of Strategies for Sampling Coordination and Statistical and Statistical Inference in Finite Population Sampling* (Unveröffentlichte Dissertation). Université de Neuchâtel, Switzerland.

Nedyalkova, D., Qualité, L. & Tillé, Y. (2009). General framework for the rotation of units in repeated survey sampling. *Statistica Neerlandica. Journal of the Netherlands Society for Statistics and Operations Research*, *63* (3), 269–293.

Nordberg, L. (2000). On Variance Estimation for Measures of Change When samples are Coordinated by the Use of Permanent Random Numbers. *Journal of Official Statistics*, *16* (4), 363–378.

Ohlsson, E. (1992). *SAMU - The System for Co-ordination of Samples from the Business Register at Statistics Sweden - A Methodological Description* (Research Report). Statistics Sweden. Zugriff auf `http://gauss.stat.su.se/master/es/SAMU-2.pdf`

Ohlsson, E. (1995). Coordination of Samples Using Permanent Random Numbers. In B. G. Cox, D. A. Binder, B. N. Chinnappa, A. Christianson, M. J. Colledge & P. S. Kott (Hrsg.), *Business survey methods* (S. 153–169). Wiley.

Ohlsson, E. (1998). Sequential Poisson Sampling. *Journal of Official Statistics*, *14* (2), 149–162.

Osier, G. (2009). Variance estimation for complex indicators of poverty and inequality using linearization techniques. *Survey Research Methods*, *3* (3), 167–195.

Osier, G., Berger, Y. & Goedeme, T. (2013). *Variance Estimation for Complex Surveys* (Methodologies and Working Papers Nr. KS-RA-13-024-EN). Eurostat. Zugriff auf `http://ec.europa.eu/eurostat/documents/3888793/5855973/KS-RA-13-024-EN.PDF`

Park, Y. S., Kim, K. W. & Choi, J. W. (2001). One-Level Rotation Design Balanced on Time in Monthly Sample in Rotation Group. *Journal of the American Statistical Association*, *96* (456), 1483–1496.

Pires, A. M. & Branco, J. A. (2002). Partial influence functions. *Journal of Multivariate Analysis*, *83* (2), 451–468.

Qualité, L. (2009). *Unequal Probability Sampling and Repeated Surveys* (Unveröffentlichte Dissertation). Université de Neuchâtel, Switzerland.

Qualité, L. (2011). Developments on coordinated poisson sampling. In *Proccedings of the 2011 world statistics congress.*

Qualité, L. & Tillé, Y. (2008). Variance Estimation of changes in repeated surveys and its application to the Swiss survey of value added. *Survey Methodology*, *34* (2), 173–181.

Raj, D. (1968). *Sampling theory*. New York: McGraw-Hill.

Rivière, P. (2001). Coordinating samples using the microstrata methodology. In *Proceedings of statistics canada symposium.*

Rosén, B. (1997a). Asymptotic Theory for Order Sampling. *Journal of Statistical Planning and Inference*, *62* (2), 135–158.

Rosén, B. (1997b). On Sampling with Probability Proportional to Size. *Journal of Statistical Planning and Inference*, *62* (2), 159–191. doi: 10.1016/S0378-3758(96)00186-3

Salamin, P.-A. (2009). *Measuring the performance of a sample coordination system.* Zugriff auf `http://enbes.wikispaces.com/file/view/S3_2_Salamin.pdf/130772511/S3_2_Salamin.pdf`

Sampford, M. R. (1967). On Sampling Without Replacement with Unequal Probabilities of Selection. *Biometrika*, *54* (3/4), 499–513.

Särndal, C.-E. (2007). The Calibration Approach in Survey Theory and Practice. *Survey Methodology*, *33* (2), 99-119.

Särndal, C.-E., Swensson, B. & Wretman, J. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Serfling, R. J. (1980). *Approximation theorems of mathematical statistics*. New York: Wiley. (Wiley Series in Probability and Mathematical Statistics)

Shao, J. (2003). *Mathematical Statistics* (2. Aufl.). New York: Springer.

Sinn, H.-W. (2012). *Kasino Kapitalismus* (3. Aufl.). Ullstein.

Skinner, C. J. & Holmes, D. J. (2003). Random Effects Models for Longitudinal Survey Data. In R. L. Chambers & S. C. J. (Hrsg.), *Analysis of survey data* (S. 205–219). Wiley.

Steel, D. & McLaren, C. (2009). Design and Analysis of Survey Repeated over Time. In C. Rao (Hrsg.), *Handbook of Statistics - Sample Surveys: Inference and Analysis* (Bde. 29, Part 2, S. 289–249). Elsevier.

Stenger, H. (1989). Asymptotic Analysis of Minimax Strategies in Survey Sampling. *The Annals of Statistics*, *17* (3), 1301–1314.

Tam, S. M. (1984). On covariances from overlapping samples. *The American Statistician*, *38* (4), 288–289.

Tillé, Y. (1996). Some Remarks on Unequal Proability Sampling Designs Without Replacement. *Annales D'Économie et de Statistique*, *44*, 177–189.

Tillé, Y. (2006). *Sampling Algorithms*. New York: Springer.

Tillé, Y. & Favre, A.-C. (2004). Coordination, Combination and Extension of Balanced Samples. *Biometrika*, *91* (4), 691–709.

Valliant, R., Dever, J. A. & Kreuter, F. (2013). *Practical Tool for Designing and Weighting Survey Samples*. Springer-Verlag New York.

Verma, V., Betti, G. & Ghellini, G. (2007). Cross-sectional and Longitudinal Weighting in a Rotational Houshold Panel: Application to EU-SILC. *Statistics in Transition*, *8* (1), 5–50.

Verma, V., Betti, G. & Ghellini, G. (2011). Taylor linearization sampling errors and design effects for poverty measures and other complex statistics. *Journal of Applied Statistics*, *38* (8), 1549–1576.

Wolter, K. M. (1985). *Introduction to variance estimation*. New York: Springer.

Wolter, K. M. (2007). *Introduction to variance estimation*. New York: Springer.

Wood, J. (2008). On the Covariance Between Related Horvitz-Thompson Estimators. *Journal of Official Statistics*, *24* (1), 53–78.

Woodruff, R. S. (1971). A Simple Method for Approximating the Variance of a Complicated Estimate. *Journal of the American Statistical Association*, *66* (334), 411–414.

Yates, F. & Grundy, P. M. (1953). Selection Without Replacement from Within Strata with Probability Proportional to Size. *Journal of the Royal Statistical Society. Series B (Methodological)*, *15* (2), 253–261. Zugriff auf `http://www.jstor.org/stable/2983772`