

# MAGISTERARBEIT

Titel der Magisterarbeit

## „Statistische Modelle zur Schätzung von Nährstoffaufnahmeverteilungen

Gegenüberstellung der Modelle und Anwendung auf die Daten  
des österreichischen Ernährungsberichts 2012“

Verfasst von

Dominik Hlauschek Bakk. rer. soc. oec.

angestrebter akademischer Grad

Magister der Sozial- und Wirtschaftswissenschaften (Mag. rer. soc. oec.)

Wien, 2015

Studienkennzahl lt. Studienblatt:  
Studienrichtung lt. Studienblatt:  
Betreuer / Betreuerin:  
Mitbetreut von:

A 066 951  
Magisterstudium Statistik  
ao. Univ.-Prof. Mag. Dr. Marcus Hudec  
Mag. Dr. Andreas Baierl

## **Danksagung**

An dieser Stelle möchte ich mich bei meinem Betreuerteam Ao. Univ.-Prof. Dr. Marcus Hudec und Mag. Dr. Andreas Baierl bedanken, dass Sie diese Magisterarbeit möglich gemacht haben. Vor allem gilt mein Dank Mag. Dr. Andreas Baierl für die Unterstützung und fachliche Beratung bezüglich der Fragestellung dieser Arbeit.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
1.1	Nährstoffbedarf . . . . .	4
1.1.1	Bedarfsermittlung . . . . .	5
1.1.2	Bedarfsempfehlungen . . . . .	5
1.2	Datenerhebung in den Ernährungswissenschaften . . . . .	9
1.2.1	Datenerhebungsmethoden . . . . .	9
1.2.2	Zufalls- oder systematische Stichprobe? . . . . .	10
1.3	Österreichischer Ernährungsbericht . . . . .	10
<b>2</b>	<b>Statistische Modelle zur Schätzung einer Aufnahmeverteilung</b>	<b>12</b>
2.1	Mathematischer Rahmen . . . . .	12
2.2	Basismodelle . . . . .	17
2.2.1	National Research Council Methode (NRC) . . . . .	17
2.2.2	ISU Best Power (BP) Method . . . . .	19
2.2.3	Bisherige Annahmen . . . . .	21
2.3	Hauptmethoden dieser Arbeit . . . . .	22
2.3.1	National Cancer Institute Methode (NCI) . . . . .	22
2.3.2	Multiple Source Methode (MSM) . . . . .	24
2.3.3	Statistical Program to Assess Dietary Exposure (SPADE) . . . . .	26
2.3.4	Iowa State University Methode (ISU) . . . . .	28
2.4	Inhaltlicher Vergleich der Hauptmethoden . . . . .	30
2.4.1	Schlussfolgerungen . . . . .	33
2.5	Schätzung des Anteils: Probability Approach . . . . .	36
2.6	Schätzung des Anteils: EAR cut-point Methode . . . . .	37
2.6.1	Andere cut-points als der EAR . . . . .	38
2.7	Auswertungen in dieser Arbeit . . . . .	39
<b>3</b>	<b>Daten</b>	<b>41</b>
3.1	Analyse der Personen mit einem und zwei Recalls . . . . .	41
3.1.1	BMI . . . . .	41
3.1.2	Indizes . . . . .	42
3.2	Weitere Dropoutanalyse . . . . .	44
3.3	Vergleiche der Zeitpunkte . . . . .	46
3.4	Verteilungen der Nährstoffaufnahmen . . . . .	46
3.4.1	Episodisch konsumierte Nährstoffe . . . . .	46
<b>4</b>	<b>Vergleiche der Hauptmethoden</b>	<b>48</b>
4.1	Vergleich der Performance anhand von früheren Studien . . . . .	48
4.1.1	Weitere Vergleiche . . . . .	49
4.2	Vergleich anhand der Daten des Ernährungsberichtes . . . . .	49
4.2.1	Verwendung für Subgruppen . . . . .	53
4.3	Schlussfolgerungen . . . . .	55

<b>5</b>	<b>Ernährungsbericht Neuauswertung</b>	<b>57</b>
5.1	Energie und Makronährstoffe . . . . .	59
5.1.1	Energie . . . . .	59
5.1.2	Kohlenhydrate . . . . .	60
5.1.3	Eiweiß . . . . .	61
5.1.4	Fett . . . . .	63
5.2	Vitamine . . . . .	63
5.2.1	Vitamin A . . . . .	63
5.2.2	Vitamin B1 . . . . .	64
5.2.3	Vitamin B2 . . . . .	64
5.2.4	Vitamin B3 . . . . .	65
5.2.5	Vitamin B5 . . . . .	65
5.2.6	Vitamin B6 . . . . .	66
5.2.7	Vitamin B7 . . . . .	67
5.2.8	Vitamin B9 . . . . .	68
5.2.9	Vitamin B12 . . . . .	70
5.2.10	Vitamin C . . . . .	70
5.2.11	Vitamin D . . . . .	71
5.2.12	Vitamin E . . . . .	72
5.2.13	Vitamin K . . . . .	73
5.2.14	Zusammenfassung bezüglich Vitamine . . . . .	73
5.3	Mineralstoffe . . . . .	76
5.3.1	Kalzium . . . . .	76
5.3.2	Magnesium . . . . .	76
5.3.3	Eisen . . . . .	77
5.3.4	Kupfer . . . . .	77
5.3.5	Phosphor . . . . .	78
5.3.6	Zink . . . . .	78
5.3.7	Zusammenfassung der Mineralstoffe . . . . .	78
<b>6</b>	<b>Diskussion und Schlussfolgerung</b>	<b>80</b>
6.1	Verzerrte Daten? . . . . .	80
6.1.1	Goldberg Cut-Off . . . . .	81
6.1.2	Resultate . . . . .	82
6.2	Diskussion bezüglich der statistischen Methoden . . . . .	84
6.3	Diskussion bezüglich der Neuauswertung . . . . .	85
6.4	Schlussfolgerungen . . . . .	86
	<b>Appendices</b>	<b>95</b>
	<b>A Tabellen</b>	<b>96</b>
	<b>B Dietary Reference Intakes</b>	<b>98</b>
	<b>C DACH Referenzwerte</b>	<b>107</b>
	<b>D Lebenslauf</b>	<b>108</b>
	<b>E Zusammenfassung/Abstract</b>	<b>109</b>
E.1	Zusammenfassung . . . . .	109
E.2	Abstract . . . . .	109

## Abkürzungsverzeichnis

<b>AD-Test</b>	Anderson Darling Test
<b>AI</b>	Adequate Intake
<b>DACH</b>	Verband von österreichischer, deutscher und schweizer Ernährungsgesellschaften
<b>DRI</b>	Dietary Reference Intakes
<b>EAR</b>	Estimated Average Requirement
<b>ISU</b>	Iowa State University (Methode)
<b>MSM</b>	Multiple Source Method
<b>NCI</b>	National Cancer Institute (Methode)
<b>RDA</b>	Recommended Daily Allowance
<b>SD</b>	Standardabweichung
<b>SPADE</b>	Statistical Program to Assess Dietary Exposure
<b>SW-Test</b>	Shapiro Wilk Test
<b>WPM</b>	within-person Mittelwert (Methode)

# Kapitel 1

## Einleitung

47,1% der Frauen und 19,8% der Männer haben, laut Ernährungsbericht 2012, eine “deutlich erniedrigte“ Vitamin B12 Aufnahme. Auf diese Aussage hin können viele Fragen gestellt werden. Wie kommen solche Schätzungen zustande? Wie genau sind sie? Machen solche Aussagen Sinn? Ist dies ein Zeichen, dass viele Österreicher einen Vitamin B12 Mangel vorweisen? Um diese Fragen zu beantworten, ist klarerweise eine gewisse Kenntnis von Ernährungswissenschaften notwendig. Allerdings ist es auch notwendig die statistische Theorie die sich hinter dieser Schätzung befindet zu verstehen. Die Hauptaufgabe dieser Arbeit ist es einen Einblick in moderne statistische Methoden, welche zur Schätzung solcher Anteile verwendet werden, zu geben. Ebenso werden in dieser Arbeit die Methoden auf die Daten der Erwachsenen vom Ernährungsbericht 2012 (Elmadfa et al., 2012) angewandt. Als erstes wird jetzt dem Leser im 1. Kapitel eine kurze Einführung in die Ernährungswissenschaften rund um das Thema Nährstoffbedarf, Referenzwerte und Datenerhebung gegeben, welche auch in gängigen Ernährungswissenschaftsbüchern nachgelesen werden kann. Dieses Kapitel richtet sich weitgehend nach Schek (2011). Ebenso wird kurz auf den Ernährungsbericht der Österreicher eingegangen. Als nächstes wird dann in Kapitel 2 die statistische Theorie erläutert. Im 3. Kapitel werden die Daten vom Ernährungsbericht näher betrachtet. Kapitel 4 stellt die Hauptmethoden der Arbeit, sowohl inhaltlich als auch angewandt auf die Daten, gegenüber. Im 5. Kapitel werden die Ernährungsbericht Aufnahmedaten der Erwachsenen neu ausgewertet und die Ergebnisse verglichen. Im letzten Kapitel wird dann eine Diskussion und Schlussfolgerung gegeben.

### 1.1 Nährstoffbedarf

Nährstoffe, wie Kohlenhydrate, Fette, Proteine, Vitamine aber auch Mineralstoffe und viele mehr spielen eine wichtige Rolle in unserem Leben. Sie sind notwendig um uns am Leben und gesund zu halten. Daher ist es nicht verwunderlich, dass in den Ernährungswissenschaften der Versorgungsstatus der Bevölkerung mit diesen wichtigen Nährstoffen eine zentrale Rolle spielt. Der Versorgungsstatus stellt dar, ob von einem Nährstoff ein Mangel <sup>1</sup> vorliegt. Ein Mangel kann über mehrere Möglichkeiten festgestellt werden. Eine relativ direkte Methode stellen Blut oder Harnmarker dar. Dabei wird das Wissen über den menschlichen Stoffwechsel ausgenutzt. Die über die Nahrung aufgenommenen Nährstoffe werden im Magen und Darm verdaut bzw. absorbiert (aufgenommen) und anschließend übers Blut (und Lymphsystem) zu den einzelnen Organen oder Zellen befördert. Daher können die Nährstoffmengen (manchmal auch Stoffwechselprodukte) im Blut bestimmt werden und mit ermittelten Grenzwerten, welche bei gesunden Personen ermittelt worden sind, abgeglichen werden. Ebenso können manche Nährstoffe bzw. Abbauprodukte im Harn bestimmt werden um einen Eindruck vom Versorgungsstatus zu bekommen. Weiter wird hier allerdings auf dieses Thema nicht eingegangen, da diese Blutmarkerwerte nicht für die späteren Auswertungen herangezogen werden.

Eine weitere (kostengünstigere) Möglichkeit den Versorgungsstatus zu ermitteln stellt die

---

<sup>1</sup>Ein Mangel eines Nährstoffes bedeutet, dass dem Körper zu wenig von dem Nährstoff vorliegt und er dadurch nicht mehr alle Körperfunktionen (welche mit diesem Nährstoff in Verbindung stehen) optimal ausführen kann. Siehe auch Nährstoffbedarfsdefinition im weiteren Textverlauf.

Nahrungsaufnahme dar, da ein Nährstoffmangel vor allem durch eine zu geringe Aufnahmemenge entstehen kann. Um beurteilen zu können ob eine zu geringe Aufnahme vorliegt, muss zuerst einmal der Bedarf der Nährstoffe bekannt sein. Der Nährstoffbedarf ist definiert als die aufzunehmende Menge, die notwendig ist, um die Funktionen des Nährstoffes im Körper aufrechtzuerhalten. Da jeder Mensch unterschiedlich groß, muskulös und alt ist oder sich in verschiedenen Zuständen wie Krankheit, Schwangerschaft etc. befindet, benötigt auch jeder unterschiedliche Mengen desselben Nährstoffes um seine Funktionen aufrechtzuerhalten. Um den Nährstoffbedarf zu bestimmen müssen also auch diese Faktoren berücksichtigt werden.

### 1.1.1 Bedarfsermittlung

Ermittelt kann der Bedarf eines Nährstoffes meist nur sehr schwer und umständlich. Eine Möglichkeit sind Mangelexperimente. Dabei wird entweder am Menschen oder am Tier durch das Weglassen bzw. Reduzieren des Nährstoffes versucht herauszufinden, welche Mangelsymptome auftreten. Danach wird die Dosis wieder erhöht um festzustellen, ab welcher Menge die Symptome wieder abklingen bzw. der Körper keine Probleme mehr andeutet. Ein vereinfachtes Beispiel zur Verdeutlichung wäre hier z.B. die Nachtblindheit bei Vitamin A. Wird zu wenig Vitamin A aufgenommen, so kann Nachtblindheit die Folge sein. Dabei entsteht eine erhöhte Reizschwelle der Netzhaut und eine verlangsamte Hell-Dunkel Adaption<sup>2</sup> ist die Folge. Jetzt kann z.B. bei Personen die dieses Symptom haben Vitamin A zusätzlich zugeführt und beobachtet werden, wann der Effekt nicht mehr auftritt. In der Praxis ist dies wesentlich schwieriger, da Vitamin A im Körper für mehrere Funktionen verwendet wird und es auch gespeichert werden kann, was es noch komplexer macht eine Bedarfsgrenze zu ermitteln. Ein Kritikpunkt der Mangelexperimente ist, dass diese meist nur an sehr kleinen Personengruppen durchgeführt werden können und somit die Aussagekraft nicht sehr hoch ist. Bei Tierversuchen hingegen ist immer die Übertragbarkeit der Resultate auf den Menschen ein Kritikpunkt.

Eine andere Möglichkeit ist zum Beispiel Bilanzen des Nährstoffes aufzustellen. Dabei wird herausgefunden, ab welcher Menge eine ausgeglichene Bilanz (Zufuhr = Ausscheidung) vorliegt. Dies liefert wichtige Erkenntnisse über den Bedarf, da die Körperreserven gesättigt sind, wenn das meiste der aufgenommenen Nährstoffe wieder ausgeschieden wird (Ausnahmen sind Probleme bei der Absorption der Nährstoffe z.B. durch Krankheit etc.).

Schwieriger wird es den Bedarf durch epidemiologische Studien zu ermitteln, da hierbei oft kein extremer Mangelzustand eines einzelnen Nährstoffes vorliegt. Ebenso müssen Wechselwirkungen unter den Nährstoffen, sowie andere mögliche Einflussfaktoren (z.B. Umwelt) berücksichtigt werden.

Der tatsächliche Bedarf eines Nährstoffes ist also nur sehr schwer festzustellen und für manche Gruppen kaum möglich. Vor allem Säuglinge oder Kinder stellen hier keine leichte Aufgabe dar. Kann der Bedarf nicht so direkt ermittelt werden, so wird er geschätzt. Bei Säuglingen wird z.B. die Nährstoffzusammensetzung der Muttermilch für die Bedarfsschätzung einiger Nährstoffe herangezogen. Für Kinder wird oft der Bedarf der Erwachsenen extrapoliert (bezüglich Körpergröße usw.) um Schätzungen zu bekommen. Was der Leser an dieser Stelle mitnehmen soll, ist, dass die meisten Schätzungen und Empfehlungen nicht als hundertprozentig genau angesehen werden dürfen. Es sind immer Unsicherheiten gegeben und auch der Bedarf von Person zu Person kann teils stark schwanken. Eine Empfehlung kann daher für die eine Person sinnvoll sein und für die andere als weniger streng angesehen werden. Weiter wird hier auf die Komplexität der Thematik nicht eingegangen.

### 1.1.2 Bedarfsempfehlungen

Wenn Empfehlungen ermittelt oder geschätzt werden, so werden diese als Referenzwerte bezeichnet bzw. angegeben. Dabei gibt es für viele Länder unterschiedliche Referenzwerte, sowie unterschiedliche Definitionen. Als nächstes sollen die nordamerikanischen, sowie die Österreichischen Werte näher betrachtet werden, da diese für die Auswertungen verwendet wurden.

Referenzwerte haben meist eine sehr ähnliche Herleitung. Dabei wird versucht den mittleren

---

<sup>2</sup>typisches Problem beim Autofahren in der Nacht. Die Scheinwerfer der entgegenkommenden Fahrzeuge blendet und es dauert unterschiedlich lang, bis sich die Augen wieder an die Dunkelheit gewöhnt haben.

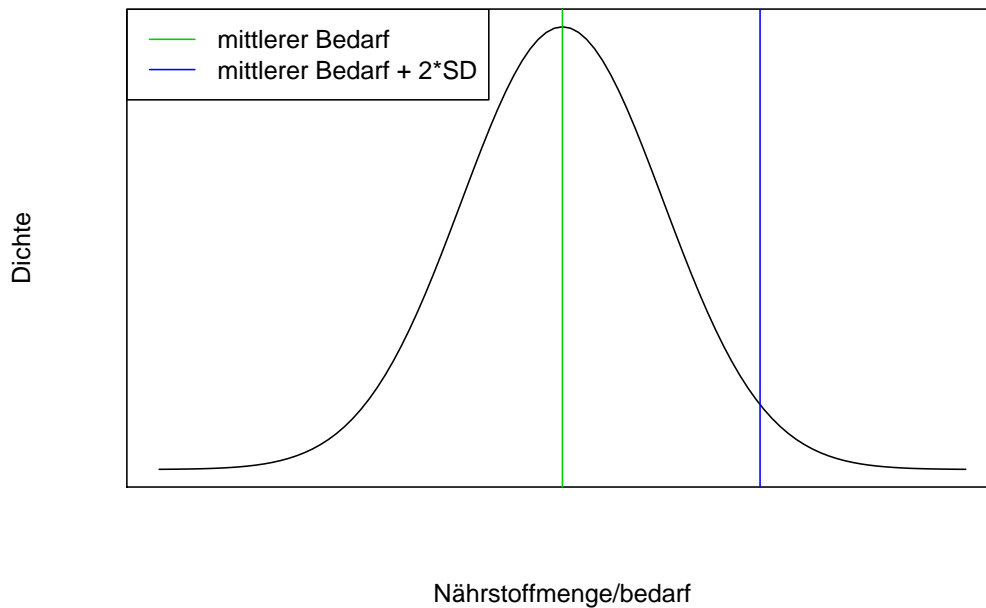


Abbildung 1.1: Nährstoffbedarfsbestimmung: normalverteilte Bedarfsverteilung einer Gruppe

Bedarf einer Gruppe (z.B. alle Männer zwischen 18-25 Jahre) zu bestimmen und ausgehend von diesem dann Empfehlungen für Einzelpersonen anzugeben. Dies kann besser mit einer Grafik erklärt werden. In Abbildung 1.1 betrachten wir eine Bedarfsverteilung. Angenommen es wäre aus Vorwissen die Form der Verteilung bekannt. In diesem Fall wäre dies eine Normalverteilung. In Experimenten wird nun der mittlere Bedarf des untersuchten Kollektivs, sowie die Streuung ermittelt. Mit dieser Information kann nun einerseits der mittlere geschätzte Bedarf des Kollektivs als Referenzwerte angegeben werden (grüne Linie). 50% der Personen in diesem Kollektiv hätten einen niedrigeren Bedarf und 50% einen höheren. Wenn eine Einzelperson also ihre Aufnahmemenge mit diesem Referenzwert vergleichen würde, so kann sie noch nicht viel aussagen, ob sie wirklich ausreichend versorgt ist oder nicht (auch wenn sie über dem mittleren Bedarf liegt). Daher wird meist auch ein Referenzwert, welcher für die meisten Personen in diesem Kollektiv ausreichend ist angegeben. Im Fall einer Normalverteilung würde zum mittleren Bedarf zweimal die geschätzte Standardabweichung dazugerechnet werden um diesen Referenzwert zu bekommen (blaue Linie). 97-98% der Personen in dieser Gruppe hätten also einen niedrigeren Bedarf. Eine Einzelperson kann also bei einer höheren Aufnahme davon ausgehen, dass sie recht sicher ausreichend viel von diesem Nährstoff aufnimmt.

Anmerkungen: Da die Bestimmung der Referenzwerte meist mit Unsicherheiten zu kämpfen hat, werden zusätzlich zu der zweifachen Standardabweichung noch Sicherheitszuschläge (z.B. 10-20%) hinzugegeben. Weiters darf nicht vergessen werden, dass der Bedarf meist keine direkt messbare Größe ist. Sprich es kann nicht einfach ein Wert - wie auch immer - gemessen werden und dann ausgesagt werden, dass eine Person unterversorgt ist oder nicht. Dies beruht immer alles auf Überlegungen und Zusammenhängen zwischen mehreren Parametern (siehe Vitamin A Beispiel vom vorigen Abschnitt). Es kommen also auch hier Unsicherheiten hinzu.

### Dietary Reference Intakes (DRI)

Vom Institute of Medicine wurden die sogenannten Dietary Reference Intakes herausgegeben. Vor allem 4 Arten von Referenzwerten spielen dabei eine wichtige Rolle. Diese genannten Re-



ferenzwerte können unter anderem in National Research Council (2006) nachgelesen werden. Da für die Arbeit diese Werte ebenfalls wichtig sind, ist im Anhang eine Zusammenfassung aller Werte gegeben.

Einerseits wird vom Institute of Medicine ein Estimated Average Requirement (mittlerer Bedarf), sowie eine Recommended Dietary Allowance (entspricht ca. dem mittleren Bedarf plus zweimal der Standardabweichung) angegeben. Andererseits wird für weniger sichere Bedarfschätzungen ein Adequate Intake angegeben. Dieser hat verschiedene Definitionen für verschiedene Nährstoffe und kann daher auf den mittleren Bedarf oder den mittleren Bedarf plus Zuschläge abzielen. Für manche Nährstoffe wird auch Tolerable Upper Intake Level angegeben, welche als obere Grenze angesehen werden kann.

**Estimated Average Requirement (EAR)** Der EAR Wert stellt das durchschnittliche tägliche Aufnahmelevel eines Nährstoffes, der für die Hälfte einer gesunden -nach Geschlecht und Alter - bestimmten Personengruppe ausreichend ist, dar. Der EAR stellt also eine Zufuhrmenge für ein bestimmtes Kollektiv dar, welche für 50% ausreichend sein sollte (siehe nochmals Abbildung 1.1). Der EAR Wert wird experimentell ermittelt und gilt daher als sicherer als manch andere Referenzwerte. Verwendet wird der EAR Wert wenn die Zufuhr von Gruppen betrachtet wird. Später in Kapitel 2 Abschnitt 2.6 wird noch weiter darauf eingegangen.

**Recommended Dietary Allowance (RDA)** Der RDA stellt das Nährstofflevel dar, welches ausreichend für fast alle (97 bis 98%) gesunden Personen eines bestimmten Kollektivs ist. Der Wert wird vom EAR abgeleitet und kann auch Sicherheitszuschläge beinhalten. Verwendet wird der RDA, wenn die Aufnahme einzelner Individuen betrachtet wird. Liegt die Aufnahme einer Person über dem RDA, so kann geschlussfolgert werden, dass die Person recht sicher ausreichend von diesem Nährstoff zu sich nimmt. Für Aufnahmen unter dem RDA muss zwar noch kein Mangel oder eine zu niedrige Aufnahme vorliegen, es wird aber wahrscheinlicher.

**Adequate Intake (AI)** Der AI hat verschiedene Definitionen und Herleitungen. Er wird für Nährstoffe definiert, wo kein EAR bzw. RDA ermittelt werden konnte. Hergeleitet wird er z.B. aus epidemiologischen Untersuchungen. Definiert ist er manchmal als beobachteter Gruppenmittelwert der geschätzten Aufnahme eines gesunden Kollektivs. Die Schätzung ist daher weitaus unsicherer als beim experimentell ermittelten EAR. Trotzdem kann der AI als grobe Schätzung oder Richtlinie für den Bedarf von Einzelpersonen oder Gruppen verwendet werden.

**Tolerable Upper Intake Level (UL)** Der UL Wert wird als Obergrenze festgelegt. Dieser sollte möglichst nicht überschritten werden, um keine negativen Effekte einer Nährstoffüberdosis befürchten zu müssen. Je höher die Aufnahme über dem UL liegt, desto höher ist das Risiko die Gesundheit zu schädigen bzw. zu gefährden.

## DACH Referenzwerte

Für Österreich werden meist die DACH (Deutschland-Österreich-Schweiz) Referenzwerte verwendet (Deutsche Gesellschaft für Ernährung et al., 2015). Diese sind ähnlich definiert wie die DRIs allerdings nicht komplett gleich. Sie beinhalten allerdings keine mittleren Bedarfsschätzungen, sondern nur Werte, welche schon die Zuschläge beinhalten. Dabei werden "Empfehlungen" (mittlerer Bedarf plus zweimal SD), "Schätzungen" (unsicherer Wert), Richtwerte und obere Grenzwerte angegeben. Im Anhang sind die verwendeten<sup>3</sup> DACH Referenzwerte in einer Tabelle zusammengefasst.

**Empfehlungen** Die "Empfehlungen" sind DACH Werte die den Durchschnittsbedarf plus Sicherheitszuschläge beinhalten. Zu beachten ist, dass die "Empfehlungen" als ausreichend für fast die ganze Population (97.5%) definiert sind (siehe Abbildung 1.1). Sollte die

<sup>3</sup>es handeltet sich hierbei nicht um die aktuellsten Referenzwerte, sondern um dieselben, die auch im Ernährungsbericht 2012 verwendet wurden.

Einnahme eines Nährstoffes für ein Individuum also über der Empfehlung liegen, so ist das ein Hinweis dafür, dass die Person mit hoher Wahrscheinlichkeit ausreichend versorgt ist. Doch auch wenn die Person eine Aufnahme unter dem Referenzwert hat kann es sein, dass sie ausreichend versorgt ist. Die "Empfehlungen" entsprechen also nicht dem Bedarfsmedian eines Kollektivs wie die Estimated Average Requirements, sondern kommen den RDA Werten des Institute of Medicine am nächsten und sollten daher nur für einzelne Individuen und nicht auf Gruppenebene verwendet werden (wird in Kapitel 2 Abschnitt 2.6.1 noch weiter erläutert).

**Schätzungen** Bei den DACH Werten sind die "Schätzungen" ähnlich dem AI der Amerikaner und dadurch als unsicherer als die "Empfehlungen" anzusehen. Sollte die geschätzte mittlere Aufnahme des Nährstoffes einer Gruppe größer oder gleich der "Schätzung" des Bedarfs sein, so kann zumindest vorsichtig (wie auch beim AI) argumentiert werden, dass nur wenige Leute unterversorgt sind. Ebenso kann dieser Wert verwendet werden, wenn einzelne Individuen betrachtet werden.

**Richtwerte** Diese sollten, wie der Name schon sagt, nur als Richtwerte verwendet werden und nur zur vorsichtigen qualitativen Beurteilung einer ausreichenden Aufnahme verwendet werden.

**Obere Grenzwerte** Die oberen Grenzwerte sind wie die ULs als Höchstwerte zu verstehen. Sie sollten nicht überschritten werden.

Zusammenfassend lässt sich sagen, dass die DACH Referenzwerte keine Angaben zu dem mittleren (oder medianen) Bedarf eines Kollektivs und somit keine ähnlichen Werte, wie die EAR vom Institute of Medicine haben (grüne Linie in Abbildung 1.1). Die "Empfehlungen" der DACH Referenzwerte sind wie die RDA Werte vom Institute of Medicine zu interpretieren (blaue Linie in Abb.1.1). Die DACH "Schätzungen", sowie die AI Werte sind in Abbildung 1.1 nicht direkt darstellbar, da sie unterschiedliche Definitionen und Herleitungen haben können. Die DACH "Schätzungen" beinhalten allerdings oft schon Sicherheitszuschläge und sind daher eher im Bereich der blauen Linie anzuordnen. Die AI Werte können hingegen durchaus auch im Bereich der grünen Linie liegen.

Die besprochenen Referenzwerte, sowohl die DRI als auch die DACH Werte, werden für Männer und Frauen und für verschiedene Altersgruppen separat angegeben. Ebenso gibt es manchmal auch extra Empfehlungen für Schwangere. Verwendet werden sie oft um Ernährungspläne zu erstellen, aber auch um Mängel in Bevölkerungsgruppen oder einzelnen Individuen festzustellen. Dies stellt allerdings nur eine "indirekte" Methode dar, da wie schon erwähnt, nicht alle Personen tatsächlich dieselbe Menge aufnehmen müssen, um ihren Bedarf zu decken. Also selbst wenn der ermittelte Referenzwert perfekt wäre, so kann es immer noch sein, dass eine Person die eine Aufnahme unter dem Wert hat, ausreichend versorgt ist mit dem Nährstoff. Umgekehrt könnte es genauso sein, dass eine Person die weit über dem Wert liegt, noch immer eine zu niedrige Aufnahme des Nährstoffes hat, da ihr Bedarf weit aus höher liegt. Wünschenswert wäre daher eine Bedarfsverteilung, welche auch die Streuung des Bedarfs berücksichtigt. Mit so einer Bedarfsverteilung könnte dann für jede Person eine Wahrscheinlichkeit für den Grad der Unterversorgung<sup>4</sup> des Nährstoffes ermittelt werden. Es gibt allerdings kaum wirkliche Verteilungsschätzungen, da dazu die Datenlage zu ungewiss oder zu klein ist. Für manche Nährstoffe werden allerdings trotzdem Variationskoeffizienten bestimmt und Verteilungsannahmen (meist Normalverteilung) getroffen, um solche Wahrscheinlichkeitsaussagen machen zu können. Im zweiten Kapitel Abschnitt 2.5 wird dann noch mehr auf Bedarfsverteilungen eingegangen.

Weiters soll noch erwähnt werden, dass ein Referenzwert nicht für jeden Tag genau eingehalten werden muss, um einen Mangel vorzubeugen. Der Körper weist für die meisten Nährstoffe einen kleinen oder größeren Gewebespeicher (vor allem fettlösliche Vitamine können lang gespeichert werden) auf und kann daher bei zu niedriger Aufnahme darauf zugreifen. Daher ist man meist nicht daran interessiert, ob ein Referenzwert täglich eingehalten wird, sondern

---

<sup>4</sup>Unterversorgung im Sinne von zu niedriger Aufnahmemenge des Nährstoffes und dadurch bestehender Gefahr eines Nährstoffmangels

die Langzeitaufnahme (mittlere Aufnahme über längere Periode) dem Referenzwert entsprechend ist. Am Anfang von Kapitel 2 wird darauf noch weiter eingegangen, da dies wichtig für die Schätzungen von Aufnahmeverteilungen ist.

## 1.2 Datenerhebung in den Ernährungswissenschaften

Um für ein Kollektiv den Anteil der Personen, welche zu wenig von einem Nährstoff aufnehmen, festzustellen zu können, werden Nahrungsaufnahmedaten benötigt. Da die Aufnahme eines Nährstoffes nicht (oder nur umständlich) direkt (am Körper) gemessen werden kann, wird sie über Befragungen des Ernährungsverhaltens ermittelt. Mehrere Möglichkeiten ergeben sich hierbei. Barrett-Connor (1991) zeigt eine Auflistung einiger dieser Methoden auf, jedoch können die meisten auch in gängiger Einführungsliteratur zu Ernährungswissenschaften nachgelesen werden. Alle Methoden haben dabei Vor- und Nachteile.

### 1.2.1 Datenerhebungsmethoden

**Verzehrshäufigkeitsfragebogen** Eine Möglichkeit dafür sind sogenannte Verzehrshäufigkeitsfragebogen (engl. food frequency questionnaires). Dabei wird allgemein abgefragt, welche Lebensmittel in einer gewissen Periode (letzten Tage, Wochen oder Monate) wie häufig zugeführt wurden. Darüber kann dann versucht werden zu schätzen wie viel von welchem Nährstoff im Schnitt (über lange Zeit) aufgenommen wurde/wird. Diese Methode gilt aber als unsichere Schätzung, da einige Probleme mit dieser Art der Datenerhebung vorliegen. Das Erinnerungsvermögen, welche Lebensmittel und wie oft sie konsumiert wurden, sowie die ungenaue Abschätzung der Menge stellen nur zwei solcher Probleme dar.

**Wiegemethode** Eine andere Methode wäre die konsumierten Lebensmittel von den Personen dokumentieren und abwiegen zu lassen. Während dies eine der genauesten Möglichkeiten zur Bestimmung der aufgenommenen Nährstoffe darstellt, so hat es doch zwei große Nachteile. Einerseits sind solche Erhebungen teuer und daher für große epidemiologische Untersuchungen meist ungeeignet. Andererseits entsteht eine Veränderung des Konsumverhaltens der Personen, wenn sie alles aufschreiben und abwiegen müssen, was sie essen. Daher stellt diese Methode dann nicht mehr die sonst übliche Aufnahme dar.

**Diäthistorie/Ernährungsprotokoll** Weniger teuer und aufwendig für die Studienteilnehmer wäre die Diäthistorie Methode. Dabei muss der Proband die konsumierten Lebensmittel "nur" mitschreiben und die Menge abschätzen. Trotzdem kann auch hier eine Veränderung des Ernährungsverhalten auftreten.

**24 Stunden Erinnerungsprotokolle** Für die Schätzung von Aufnahmeverteilungen und Populationsmittelwerten wird laut Brussaard et al. (2002) die Befragung über die Nahrungsaufnahme der letzten 24 Stunden als passendste Methode angesehen (24 Stunden Erinnerungsprotokolle; engl. 24 hour Recall). Es wird dabei meist nach einem standardisierten Protokoll abgefragt, zu welcher Mahlzeit welche Lebensmittel zu sich genommen wurden. Mit einem Fotobuch, welches verschiedene Lebensmittel und Portionsgrößen beinhaltet, wird versucht, die ungefähre Menge bestmöglich zu bestimmen. Die angegebenen Lebensmittel(-mengen) werden nun anhand bestimmter Nährstoffdatenbanken (z.B. dem Bundeslebensmittelschlüssel) auf die Zusammensetzung hin ausgewertet. Somit kann für jeden untersuchten Tag und (fast) jeden Nährstoff die verzerrte Menge bestimmt werden. Klarerweise entstehen dadurch Fehler, da die tatsächlich aufgenommene Menge nicht mit der geschätzten übereinstimmen wird. Solange die Schätzungen aber erwartungstreu sind, spielt dies eine untergeordnete Rolle. Das größere Problem stellt hier die vom Individuum angegebenen Lebensmittel dar. Werden systematisch Lebensmittel vergessen oder absichtlich nicht angegeben, kommt es zu verzerrten Schätzungen. Im Abschlusskapitel wird weiter auf dieses Problem eingegangen.

Aus einem wichtigen Grund (Varianzschätzung) ist es erforderlich mehrere 24 Stunden Protokolle je Person zu erheben. Auf diesen Grund wird später (in Kapitel 2) noch genauer eingegangen. Als Minimum für die Schätzung der Aufnahmeverteilung werden zwei 24 Stunden Erinnerungsprotokolle an nicht aufeinanderfolgenden Tagen angesehen (Brussaard et al.,

2002). Das erste 24 Stunden Erinnerungsprotokoll wird dabei oft durch einen Interviewer vor Ort erhoben. Dieser fragt (im Idealfall) nach einem standardisierten validierten Protokoll die konsumierten Lebensmittel ab. Nachfolgende Erinnerungsprotokolle können auch telefonisch abgefragt werden.

### 1.2.2 Zufalls- oder systematische Stichprobe?

Ein wichtiger Punkt der Datenerhebung ist auch die Auswahl der Stichprobe. Während für große Populationen einfache Zufallsstichproben nur schwer möglich sind, wird hier gern auf andere Arten der Erhebung zurückgegriffen wie z.B. auf Quotenstichproben. Bei diesen wird nach einem System - anstatt zufällig - aus der Grundgesamtheit gezogen. Es wird versucht anhand bestimmter Merkmale und Zusammensetzungen der Grundgesamtheit (z.B. Alter, Geschlecht) eine möglichst ähnliche Zusammensetzung der Stichprobe zu erhalten. Ein Vorteil gegenüber der einfachen Zufallsstichprobe ist, dass für Ausfälle versucht wird einen möglichst gleichen Fall neu zu ziehen. Gerade bei solchen Erhebungen auf großer Populationsebene kann es passieren, dass bei einer einfachen Zufallsstichprobe viele Personen wegfallen und es dadurch zu einem Bias in der Stichprobe kommt. Bei der Quotenstichprobe wird versucht diesen Bias zu reduzieren in dem eben eine möglichst ähnliche Person nachgezogen wird.

Während die Quotenstichprobe zwar diesen Vorteil gegenüber der Zufallsstichprobe besitzt, so hat sie jedoch auch einige Nachteile und Kritikpunkte. Einerseits sind die Daten nicht zufällig gezogen worden, weshalb die meisten statistischen Methoden (welche für Zufallsdaten konzipiert wurden) nicht angewendet werden sollten. Auch die Wahl der Quotenkriterien stellt einen Kritikpunkt dar. Obwohl eigentlich versucht wird weniger bis gar keine Verzerrungen zu erhalten, so kann doch durch die Wahl der Kriterien ebenfalls eine (unbewusste) Verzerrung entstehen. Ebenso wird der "direkte" Einfluss auf die Auswahl der Personen kritisiert. Zum Beispiel können immer noch bestimmte Personen systematisch wegfallen und die Auswahl der "ähnlichen" Personen nicht objektiv genug erfolgen (z.B. können eher Individuen nachgezogen werden, welche leichter erreichbar sind als andere passendere Personen). Siehe zur Quotenstichprobe und diesen Kritikpunkten auch Schnell et al. (2005) ab Seite 294.

## 1.3 Österreichischer Ernährungsbericht

Bevor jetzt auf die statistischen Modelle zur Aufnahmeschätzung eingegangen wird, wird noch kurz die verwendete Datenquelle - der österreichische Ernährungsbericht 2012 - beschrieben.

In Österreich wird zur Zeit alle 4 Jahre der Versorgungsstatus der Bevölkerung ermittelt. Nahrungsaufnahmedaten werden erhoben und anschließend die Verteilung der Aufnahme sowie der Anteil der Unterversorgten bestimmt. Ein Ernährungsbericht wird erstellt, in dem die Ergebnisse der Auswertung dokumentiert und festgehalten werden. Der letzte Bericht stammt aus dem Jahre 2012 (Elmadfa et al., 2012). 2016 soll wieder ein neuer Ernährungsbericht veröffentlicht werden. 2012 wurden erstmals auch Blut- und Harnmarker bestimmt. Die Ergebnisse des Versorgungsstatus wurden daher anhand der Aufnahme, sowie der Markerwerte bestimmt. Für diese Arbeit werden allerdings nur die Aufnahmedaten verwendet. Auf diese werden dann die im nächsten Kapitel beschriebenen Methoden angewandt. Im Laufe der Arbeit wird auf mögliche Fehlerquellen der - im Ernährungsbericht 2012 - verwendeten Methodik hingewiesen, sowie Verbesserungsvorschläge eingebracht.

**Datenerhebung** Auch für den Ernährungsbericht von 2012 wurde eine Quotenstichprobe verwendet. Es wurde dabei versucht, eine möglichst repräsentative Stichprobe stratifiziert nach Geschlecht, Alter und Region zu erheben. Mögliche Verzerrungen durch die Stichprobenrekrutierung wurden durch Gewichtungsfaktoren (bezüglich der genannten drei Faktoren) versucht auszugleichen (Elmadfa et al., 2012). Die Wahl der Methodik zur Nährstoffaufnahmebestimmung fiel auf zwei 24 Stunden Erinnerungsprotokolle, wobei die erste Befragung vor Ort und die zweite telefonisch durchgeführt wurde.

Der Hauptteil der Arbeit beschäftigt sich nun, wie schon erwähnt, mit der Schätzung der

Aufnahmeverteilung einer Gruppe. Dabei wird jetzt in Kapitel 2 näher darauf eingegangen, was überhaupt eine Aufnahmeverteilung ist und was genau zur Schätzung eines Anteils notwendig ist.

## Kapitel 2

# Statistische Modelle zur Schätzung einer Aufnahmeverteilung

Im ersten Kapitel wurde jetzt öfter mal die Aufnahmeverteilung erwähnt. In diesem Kapitel geht es nun darum, wie man zu so einer Aufnahmeverteilung kommt. Wir betrachten hier den Fall, dass die Aufnahme einer gesamten Gruppe bzw. eines Kollektivs geschätzt werden soll. Dafür wird in der Literatur von einer üblichen oder gewöhnlichen Aufnahmeverteilung (engl. usual/habitual intake distribution) gesprochen. Die übliche Aufnahme ist definiert als mittlere Langzeitaufnahme eines Nährstoffes. Dies ist im wesentlichen nichts anderes als der Erwartungswert der täglichen Aufnahme. Manchmal wird die übliche Aufnahme eines Nährstoffes auch für eine gewisse Periode, wie z.B. ein Jahr, definiert. Für *eine* Person wäre daher der Mittelwert der aufgenommenen Menge von den 365 Tagen die übliche Aufnahme. Die übliche Aufnahmeverteilung eines Kollektivs wäre die Verteilung der zusammengenommenen Mittelwerte. Im Rahmen dieser Arbeit wird keine fixe Periode für die Langzeitaufnahme definiert und daher wird die übliche Aufnahme einer Person als der Erwartungswert festgelegt. Die Verteilung der Erwartungswerte stellt also die übliche Aufnahmeverteilung dar. Geschätzt werden solche Aufnahmeverteilungen mit 24 Stunden Erinnerungsprotokollen.

### 2.1 Mathematischer Rahmen

Hat man jetzt für ein repräsentatives Stichprobenkollektiv 24 Stunden Protokolle (für manche Personen mehrere), so kann eine Verteilung für die Gruppennährstoffaufnahme geschätzt werden. Als nächstes betrachten wir diese Problemstellung in einem formaleren Rahmen. Beobachtungen über die Aufnahme eines Nährstoffes der letzten 24 Stunden bezeichnen wir mit  $y_{ij}$ . Der Index  $i$  steht dabei für das  $i$ -te Individuum und  $j$  für den  $j$ -ten Tag. Interessiert sind wir an der üblichen Aufnahme eines Nährstoffes, sprich dem Mittel bzw. Erwartungswert  $y_i = E(y_{ij}|i)$ . Ist man nun an der Schätzung der üblichen Aufnahme pro Individuum interessiert, ist die beste Möglichkeit diese zu schätzen in dem der Mittelwert über alle Beobachtungen  $y_{ij}$  für  $j = 1, \dots, m$  genommen wird ( $m$  steht hierbei für die Anzahl der 24 Stunden Befragungen). Hat man nun eine Schätzung  $\bar{y}_i = \frac{1}{m} \sum_{j=1}^m y_{ij}$  für das  $i$ -te Individuum, so kann diese einfach mit dem Referenzwert (Bedarfwert)  $r_i$  verglichen werden um einen Eindruck davon zu bekommen, ob die Person mit diesem Nährstoff gut versorgt ist. Möchte man nun für ein Gruppe die Unterversorgung abschätzen, so könnte man als erste Idee darauf kommen, für jede Person in der Stichprobe den Mittelwert, über die beobachteten Tage, mit dem Bedarfwert zu vergleichen. Den Anteil der Personen der unter dem Grenzwert liegt, könnte man als Schätzung für den Anteil der Unterversorgten in der betrachteten Population ansehen. Diese Art der Schätzung wurde auch im Ernährungsbericht 2012 verwendet. Das diese einfache Methode jedoch zu verzerrten Schätzungen führen kann, zeigt die kleine Simulationsstudie im nächsten Abschnitt. Tabelle 2.1 fasst nochmal die verwendete Notation der Fragestellung zusammen.

Variable	Bezeichnung
$y_{ij}$	beobachtete Aufnahme am j-ten Tag der i-ten Person
$F_i$	Verteilung der Tagesaufnahmen für Individuum i
$y_i = E(y_{ij} i)$	übliche (Langzeit) Aufnahme der i-ten Person
$F_Y$	übliche (Langzeit) Aufnahmeverteilung des Kollektivs
$r_i$	Bedarf (Referenzwert) der i-ten Person
$y_i < r_i$	Aufnahme i-ter Person unbefriedigend

Tabelle 2.1: Notation

### Simulationsstudie

Unsere übliche Aufnahmeverteilung  $F_Y$  soll für einen hypothetischen Nährstoff einer Normalverteilung mit Mittelwert 100mg und Standardabweichung 10mg folgen. Diese Verteilung ist das Ziel unserer Schätzung.

Um nun Daten, wie sie bei 24 Stunden Befragungen der Fall sind, zu simulieren, werden 500 Zufallszahlen ( $y_i, i = 1, \dots, 500$ ) aus der gerade beschriebenen Aufnahmeverteilung gezogen. Dies stellt nun unsere repräsentative Stichprobe von 500 Personen des zu untersuchenden Kollektivs dar. Da jedes Individuum nicht genau an jedem Tag die gleiche Menge eines Nährstoffes konsumiert, sondern eine Tag-zu-Tag Variabilität hinzukommt, nehmen wir eine weitere hypothetische Verteilung  $F_i$  der täglichen Aufnahme an. Diese folgt einer weiteren Normalverteilung mit der vorher gezogenen Zufallszahl  $y_i$  als Mittelwert und einer Standardabweichung von 15mg. Diese Streuung wird auch within-person (intra-individual) Variabilität oder Tag-zu-Tag Variabilität genannt. Sie ist im Gegensatz zur between-person Variabilität (Standardabweichung der Aufnahmeverteilung der gesamten Gruppe, in dem Fall 10mg) ein bisschen größer, was ein plausibles Szenario wiedergibt (später mehr dazu). Aus diesen 500 individuellen Verteilungen soll nun jeweils 2 mal ( $y_{ij}, j = 1, 2$ ) gezogen werden. Dies stellt die zwei "erhobenen" 24 Stunden Protokolle pro Person dar. Die Abbildung 2.1 soll das eben beschriebene besser darstellen. Die schwarze Linie stellt die übliche Langzeit Aufnahme der Gruppe dar ( $N(100mg, 10mg)$ ). Zuerst wird aus dieser Verteilung für jede Person eine Zufallszahl gezogen, welche dessen mittlere Langzeiteinnahme des Nährstoffes darstellt.

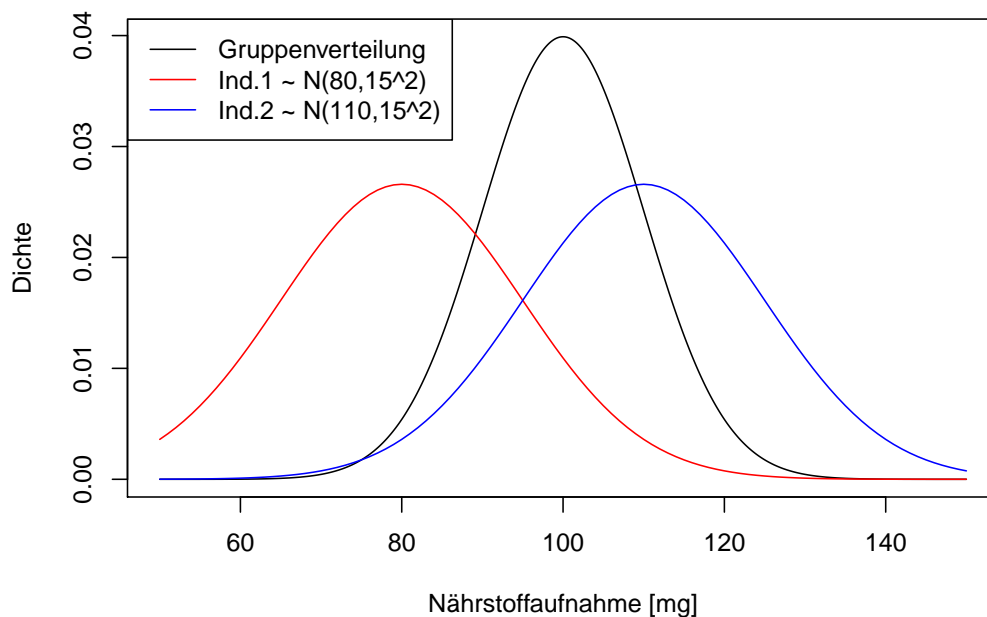


Abbildung 2.1: Aufnahmeverteilung Gruppe vs Individuen

Als nächstes werden auf der Individualebene 2 weitere Ziehungen gemacht, um die tatsächliche Datenlage herzustellen. Die blaue und rote Linie stellen jeweils die Verteilung eines Individuums dar. Die rote Person hat eine mittlere Aufnahme von 80mg und die blaue eine Aufnahme von 110mg und beide dieselbe Standardabweichung von 15mg. Die einfachste Variante um die gewöhnliche Aufnahmeverteilung zu schätzen, wäre jetzt die Mittelwerte über die 2 Beobachtungen je Person zu berechnen und diese für alle Leute zusammen zu nehmen. Das dies allerdings keine gute Variante ist zeigt die nächste Abbildung (2.2). Die schwarze Linie stellt wieder die tatsächliche Aufnahmeverteilung des Kollektivs dar und die rote Linie die bezüglich der Mittelwerte geschätzte. Bei einer normalverteilten üblichen Aufnahme ( $N(100, 10^2)$ ) und normalverteilten Tagesaufnahmen ( $N(y_i, 15^2)$ ) ergibt sich für den Mittelwerte die Mischverteilung  $N(100, 10^2 + 15^2/2)$ . In der Praxis ist man oft an dem Anteil der Personen, welche weniger als einen gewissen Grenzwert aufnehmen, interessiert. Die senkrechte blaue Linie soll hier einen hypothetischen Referenzwert von 80mg darstellen. Dieser Referenzwert könnte z.B. der EAR sein. Der tatsächliche (in der Praxis unbekannt) Anteil des Kollektivs der weniger als diesen Referenzwert aufnimmt beträgt 2.3%. Nimmt man nun als Schätzung für diesen Anteil den Prozentsatz der Personen in der Stichprobe, welche eine geringere geschätzte mittlere Aufnahme als 80mg haben, so erhält man eine (verzerrte) Schätzung von 8.5%. Die Wahrscheinlichkeit kann unter anderem mittels Statistikprogrammen berechnet werden <sup>1</sup>.

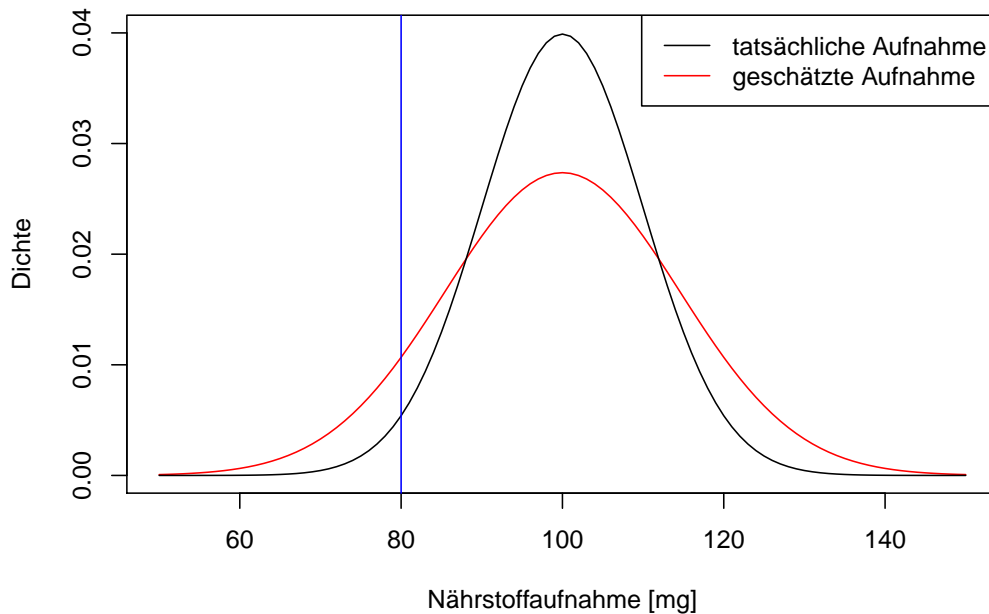


Abbildung 2.2: tatsächliche vs geschätzte Aufnahmedichte

Der Grund für dieses schlechte Abschneiden dieser Variante liegt in der Tag-zu-Tag Variabilität. Wie man an Abbildung 2.1 erkennen kann, streuen die Einnahmen der roten und blauen Person so stark, dass es sein kann, dass die Schätzungen für diese Personen doch deutlich entfernt von der Gruppenverteilung liegen. Dies führt dann zu einer breiteren geschätzten Verteilung und somit zu schlechten Anteilsschätzern. Werden mehr 24 Stunden Befragungen durchgeführt verbessert sich klarerweise das Bild, da immer bessere Schätzungen pro Person möglich sind.

Dieses Problem tritt nicht nur bei Normalverteilungen auf. Auch bei rechtsschiefen Verteilungen, wie sie in der Ernährung häufiger auftreten, werden breitere Verteilungen mit der

<sup>1</sup> z.B. im R mit `pnorm(80, 100, sqrt(102 + 152/2))`



Mittelwertmethode geschätzt als sie tatsächlich sind. Es braucht also statistische Methoden, welche diese Tag-zu-Tag Variabilität abschätzen und aus dem Modell nehmen. Mit dieser Problemstellung und statistischen Methoden, die besseres schätzen von Anteilen ermöglichen, beschäftigt sich diese Arbeit in diesem Kapitel.

In Tabelle 2.2 sind die Ergebnisse von 2 theoretischen Resultaten bezüglich der Normalverteilung und von 2 Simulationen bezüglich einer Gammaverteilung (je 10000 Wiederholungen) dargestellt. Die erste Spalte stellt dabei die übliche Aufnahmeverteilung und die zweite Spalte die Aufnahmeverteilung jeder Person dar. Sowohl symmetrische Verteilungen sowie schiefe Verteilungen wurden gewählt. Die Schätzung bei den 2 Simulationen erfolgte, wie oben schon angesprochen, durch Mittelwertbildung über die beobachteten Tage und anschließender Anteilsberechnung. Für steigende Anzahl an 24-Stunden Befragungen (=m) wird die Schätzung genauer.

Die tatsächliche Problemstellung ist natürlich weitaus komplexer als solche Szenarien. Weitere Probleme stellen die Saisonalität, welche in der Ernährung eine wichtige Rolle spielt oder z.B. die Aufnahme von Lebensmitteln, welche nicht täglich eingenommen werden, dar. Auch die Messung der individuellen Aufnahme eines Nährstoffes mittels der 24 Stunden Befragung steht in Kritik nicht valide zu sein (siehe Kapitel 6). Trotz der Einfachheit dieser Simulationen lässt sich das Problem gut aufzeigen und wurde deshalb gewählt.

$F_Y$	$F_{y_i}$	m	Schätzung	wahrer Anteil
$y_i \sim N(100,10)$	$y_{ij} \sim N(y_i,15)$	2	8.5%	2.3%
$y_i \sim N(100,10)$	$y_{ij} \sim N(y_i,15)$	5	4.8%	2.3%
$y_i \sim \text{Gamma}(2,1)$	$y_{ij} \sim \chi^2(\text{df}=y_i)$	2	34%	19.1%
$y_i \sim \text{Gamma}(2,1)$	$y_{ij} \sim \chi^2(\text{df}=y_i)$	5	26.4%	19.1%

Tabelle 2.2: Simulationsstudie: Gammaverteilungsparameter (shape,scale),  $\chi^2$ ...Chi-Quadrat Verteilung. Als Referenzwerte für die Anteile wurden für die Normalverteilungssimulationen 80 und für die Gamma Verteilungen 0.8 gewählt

## Modellierung

Um zu zeigen, warum die Mittelwertmethodik schlecht abschneidet müssen wir ihre statistischen Eigenschaften betrachten. Jede Beobachtung  $y_{ij}$  kann mit der folgenden Modellierungsstruktur beschrieben werden. Wir zerlegen die Beobachtungen dazu in drei Komponenten. Erste Komponente stellt die mittlere übliche Aufnahme  $\mu$  des Kollektivs dar. Weiters kann von dieser mittleren Kollektivaufnahme eine Abweichung  $\mu_i$  für jede Person  $i$  vorliegen. Diese stellt die typische Abweichung der mittleren Kollektivaufnahme zur mittleren Aufnahme der Person dar. Als letzte Komponente kann noch eine spezifische Abweichung  $\epsilon_{ij}$  für jeden Tag festgelegt werden. Dies stellt die Streuung der Aufnahme jeder Person über die verschiedenen Tage dar. Zusammengefasst sieht dies folgendermaßen aus:

$$y_{ij} = \mu + \mu_i + \epsilon_{ij}. \quad (2.1)$$

Eine Reihe von Annahmen kann jetzt für diese Darstellung getroffen werden. Diese sind noch sehr allgemein und ohne bestimmte Verteilungsannahmen, allerdings doch mit kleinen Vereinfachungen, die später wieder aufgehoben werden können.

1.  $\mu_i$  iid mit  $E(\mu_i) = 0$   $Var(\mu_i) = \sigma_b^2$
2.  $\epsilon_{ij}$  iid mit  $E(\epsilon_{ij}) = 0$   $Var(\epsilon_{ij}) = \sigma_w^2$
3.  $\mu_i$  unabhängig von  $\epsilon_{ij}$  für  $i = 1, \dots, n$   $j = 1, \dots, m$

Die erste Annahme stellt keine großen Einschränkungen dar. Die Variablen  $\mu_i$   $i = 1, \dots, n$  werden als unabhängig und identisch verteilt (iid) mit Erwartungswert 0 und Varianz  $\sigma_b^2$  angenommen.  $\sigma_b^2$  stellt hier die Varianz zwischen unabhängigen Individuen dar (between-person Varianz). Annahme 2 ist sehr ähnlich und legt für den Abweichungsterm der Tag-zu-Tag Veränderung ebenfalls unabhängig und identisch verteilte Zufallsvariablen fest. Die Varianz  $\sigma_w^2$  bezeichnet die Streuung innerhalb der Person (within-person Varianz). Diese Annahme ist auf jeden Fall eine Einschränkung, da so jede Person dieselbe Variabilität

in ihrer Ernährung hat, was natürlich nicht stimmt, da es Leute mit einseitigen Diäten und welche mit sehr variablen Ernährungsweisen gibt. Fürs erste wird mal diese Annahme getroffen um das Problem einfacher darzustellen. Später kann diese aufgelockert werden. Die letzte Annahme legt Unabhängigkeit zwischen den beiden Zufallsvariablen  $\mu_i$  und  $\epsilon_{ij}$  fest. Auch dies könnte eine Einschränkung sein, da womöglich Personen mit einer üblichen Aufnahme, welche höher als die mittlere Kollektivaufnahme liegt, auch eine höhere Tag-zu-Tag Variabilität haben, als Personen mit niedrigerer Aufnahme.

Erste Folgerungen dieser Annahmen sind in der nächsten Aufzählung zu sehen.

- $y_{ij}$  *id* mit  $E(y_{ij}) = \mu$   $Var(y_{ij}) = \sigma_b^2 + \sigma_w^2$
- $y_{ij}|i$  *unabhängig* mit  $E(y_{ij}|i) = \mu + \mu_i$   $Var(y_{ij}|i) = \sigma_w^2$
- $y_i \equiv E(y_{ij}|i) = \mu + \mu_i$   $y_i$  *iid Kopien von*  $Y \sim F_Y$ 
  - $E(Y) = E(\mu + \mu_i) = \mu$   $Var(Y) = Var(\mu + \mu_i) = \sigma_b^2$

Die einzelnen unbedingten  $y_{ij}$  für  $i = 1, \dots, n$  und  $j = 1, \dots, m$  sind identisch verteilt (id) mit Erwartungswert  $\mu$  und Varianz  $\sigma_b^2 + \sigma_w^2$ . Konditioniert auf das Individuum  $i$  sind die Zufallsvariablen für alle  $i$  und  $j$  unabhängig voneinander. Der Erwartungswert setzt sich aus der Summe der mittleren Kollektivaufnahme  $\mu$  und der individuellen Abweichung je Person zusammen. Die Varianz der Beobachtungen auf Individualebene betrachtet enthält natürlich nur mehr die within-person Variabilität  $\sigma_w^2$ . Der letzte Aufzählungspunkt zeigt den Teil der für unsere Problemstellung interessant ist, nämlich den bedingten Erwartungswert pro Person als eigene Zufallsvariable betrachtet. Die übliche Langzeitaufnahme  $E(y_{ij}|i)$  je Person kann als unabhängig identisch verteilte Kopie einer Zufallsvariable  $Y$  betrachtet werden. Die Verteilung  $F_Y$  dieser Zufallsvariable ist das Ziel unserer Modellierung. Sie stellt die übliche Aufnahmeverteilung der untersuchten Gruppe dar. Aus ihr soll dann der Anteil der unterversorgten Personen abgeleitet bzw. geschätzt werden. Die Eigenschaften dieser Verteilung hängen von der Verteilung der  $\mu_i$  ab. Nachdem diese, wie in der ersten Aufzählung zu sehen, Erwartungswert 0 und die between person Varianz hat, hat auch die Verteilung von  $Y$  Erwartungswert  $\mu$  und Varianz  $\sigma_b^2$ . Für spätere Zwecke hängt es also stark von der gewählten Verteilungsart von  $Y$  ab, ob der Anteil der Unterversorgten gut geschätzt wird.

Fürs Erste bleiben wir mal ohne Verteilungsannahme und zeigen weiter auf, warum die Anteilsschätzung bei der Simulationsstudie so schlecht abgeschnitten hat. Angenommen wir hätten nur einen Tag pro Person betrachtet, so hätten wir nur  $n$  Beobachtungen  $y_{i1}$  ( $i = 1, \dots, n$ ). Diese Beobachtungen haben eine Varianz von  $\sigma_b^2 + \sigma_w^2$  (siehe Folgerungen oben). Es sollte also klar sein, dass der einfache Anteil (Anzahl der Personen mit zu geringer Aufnahme) als Schätzung unbrauchbar ist, da die Varianz der Beobachtungen größer als die der gesuchten Verteilung  $F_Y$  ist.

Im Fall der Simulationsstudie wurde der Mittelwert von zwei Tagesbeobachtungen berechnet und anhand dieser dann der Anteil geschätzt. Wie die nächste Berechnung (für den allgemeinen Fall von  $m$  Tagesbeobachtungen) zeigt, ist auch dies keine gute Idee. Beachte hier, dass  $y_{ij}$  und  $y_{ij'}$  für  $j \neq j'$  abhängig sind ( $Cov(y_{ij}, y_{ij'}) = \sigma_b^2$ ).

$$Var(\bar{y}_i) = Var\left(\frac{1}{m} \sum_{j=1}^m y_{ij}\right) = Var(\mu_i) + \frac{1}{m^2} \sum_{j=1}^m Var(\epsilon_{ij}) = \sigma_b^2 + \frac{\sigma_w^2}{m} \quad (2.2)$$

Die Varianz erklärt warum die Schätzungen der Simulationsstudie besser mit mehr Beobachtungen pro Person werden. Der Anteil der Variabilität innerhalb der Personen  $\sigma_w^2$  kann somit mit steigender Beobachtungszahl  $m$  verringert werden. Beachte, dass die Anzahl der Personen in der Stichprobe  $n$  hingegen die Schätzung in diesem Sinne nicht verbessern würde.

### Varianzverringern

Eine Lösung dieses Problems kann mit einer geeigneten Transformation der Mittelwerte  $\bar{y}_i$  erreicht werden. Bei dieser Transformation werden die Mittelwerte der Individuen näher

an die übliche mittlere Aufnahme  $\mu$  der Gruppe gebracht (shrinkage). Man erhält dadurch Zwischenwerte, anhand derer die Anteilsschätzung besser vollzogen werden kann.

$$\tilde{y}_i = (1 - \omega)\mu + \omega\bar{y}_i$$

$$\text{mit } \omega = \sqrt{\frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2/m}}$$

Der Parameter  $\omega$  stellt dabei einen sogenannten Shrinkage Faktor dar. Berechnen des Erwartungswertes und der Varianz der Zwischenwerte zeigt, dass diese nun die gewünschten Eigenschaften haben.

$$E(\tilde{y}_i) = (1 - \omega)\mu + \omega E(\bar{y}_i) = \mu$$

$$\text{Var}(\tilde{y}_i) = \text{Var}(\omega\bar{y}_i) = \omega^2(\sigma_b^2 + \sigma_w^2/n) = \sigma_b^2$$

Natürlich sind in der Praxis  $\mu$  und die zwei Varianzen unbekannt, weshalb diese mit Schätzungen ersetzt werden müssen. Für  $\mu$  kann einfach der Gesamtmittelwert der Beobachtungen genommen werden. Für die Schätzung der Varianzen gibt es mehrere Möglichkeiten. Diese werden in den nächsten Abschnitten dieses Kapitels näher erläutert.

Das Problem ist mit dieser Transformation jedoch noch nicht ganz gelöst. Da die vollständige Verteilung eine Rolle bei der Anteilsschätzung spielt und nicht nur deren Erwartungswert und Varianz, kann es immer noch zu verzerrten Schätzungen kommen. Dies zeigt sich auch bei Anwenden der Transformation auf die Simulationsstudie vom Beginn. Die Schätzungen werden zwar besser für alle 4 Fälle, aber nur für die Normalverteilungen werden die Schätzungen unverzerrt. Dies liegt daran, dass die Normalverteilung durch die Transformation nicht "zerstört" wird. Bei der Gammaverteilung zum Beispiel, kommt noch die besondere Form zur Geltung. Die Mittelwerte haben schon nicht mehr eine Gammaverteilung und durch die Transformation entsteht ebenfalls eine neue Verteilung. Daher kann es hier auch weiterhin zu verzerrten Schätzungen kommen. Bei der Simulationsstudie waren es immerhin noch Abweichungen vom echten Anteil von 5.1% für  $m=2$  und 3.4% für  $m=5$  (wieder 1000 Wiederholungen). Da für die Aufnahme der meisten Nährstoffe keine Normalverteilung vorliegt, muss also auch hier eine Lösung gefunden werden. Meist ist dies eine Transformation der Daten um eine möglichst gute Approximation an die Normalverteilung zu erhalten. Auch dies wird in den nächsten Abschnitten noch behandelt.

Nicht unerwähnt soll hier der Fakt bleiben, dass die Mittelwertmethode zwar für die Quantilschätzungen zu Verzerrungen führen kann, jedoch ist sie für den Erwartungswert des Kollektivs ein unverzerrter guter Schätzer. Sollte also nur die mittlere Aufnahme des Kollektivs von Interesse sein, so stellt die Mittelwertmethode eine gute und zuverlässige Variante der Schätzung dar.

## 2.2 Basismodelle

### 2.2.1 National Research Council Methode (NRC)

Auf die Lösung der eben beschriebenen Problemstellung hatte vor allem das National Research Council (1986) einen großen Einfluss. Sie stellten einen ersten kompletten methodischen Ansatz der Schätzung einer üblichen Aufnahmeverteilung, welcher die oben genannten Probleme berücksichtigt, vor. Daher soll auch in dieser Arbeit mit dieser noch recht einfachen Methode begonnen werden.

Beobachtet werden also die Aufnahmen  $y_{ij}$  eines Nährstoffes für  $i = 1, \dots, n$  Personen an  $j = 1, \dots, m$  Tagen. Diese setzen sich aus der tatsächlichen Aufnahme  $T_{ij}$  und einem Mess- bzw. Reportingfehler  $\epsilon'_{ij}$  zusammen (Gl:2.3). Dies kann weiter beschrieben werden als die übliche Aufnahme eines Nährstoffes  $y_i$  plus der Abweichung der tatsächlichen Tagesaufnahme  $t_{ij}$  und dem Messfehler. Die Abweichung  $t_{ij}$  beinhaltet die within-person Varianz  $\sigma_w^2$  und der Messfehler  $\epsilon'_{ij}$  die Streuung der 24 Stunden Befragungen vom tatsächlichen Aufnahmewert. Da der Messfehler nur schwer bis gar nicht ermittelt werden kann, werden die zwei letztgenannten Terme zusammengefasst zu  $\epsilon_{ij}$ . Wir landen also bei folgender Gleichung.

$$y_{ij} = T_{ij} + \epsilon'_{ij} = y_i + t_{ij} + \epsilon'_{ij} = y_i + \epsilon_{ij} \quad (2.3)$$

Beachte, dass diese Gleichung dieselbe wie aus dem vorigen Abschnitt (Gl: 2.1) ist. Die Terme  $\mu$  und  $\mu_i$  in der Gl:2.1 stellen den bedingten Erwartungswert  $E(y_{ij}|i)$  dar und somit  $y_i$ . Die Verteilung der  $y_i$  gilt es zu schätzen. Die NRC Methode beinhaltet 4 wesentliche Schritte dafür:

1. Datentransformation
2. Varianzschätzung
3. Varianzeliminierung
4. Rücktransformation

Diese 4 Schritte werden auch in den anderen Methoden vorkommen, allerdings werden für die einzelnen Punkte verschiedene mathematische Anwendungen verwendet.

**1. Datentransformation:** Da die Aufnahmedaten meist schief sind und es daher nicht zu guten Schätzungen kommt, ist der erste Schritt in der NRC Methode eine Datentransformation um eine approximative Normalverteilung der Daten zu bekommen. Vorgeschlagen wurde damals eine einfache Logarithmustransformation. Klarerweise zeigt sich schon hier das erste Problem. Nicht alle schiefen Verteilungen lassen sich mittels Logarithmus zu einer approximativen Normalverteilung transformieren. Bessere und weitaus variabelere Transformationen werden in diesem Schritt bei den anderen Methoden verwendet. Allgemein geschrieben erhalten wir eine neue Variable  $y_{ij}^* = g(y_{ij})$ , wobei hier  $g(\cdot) = \ln(\cdot)$  (ln...natürlicher Logarithmus). Da diese neue Variable nun einer approximativen Normalverteilung folgen soll, wird wieder eine Varianzzerlegung wie in Gl 2.3 gewählt:

$$y_{ij}^* = y_i^* + \epsilon_{ij}^* \quad (2.4)$$

wobei  $y_i^* \sim N(\mu^*, \sigma_{b^*}^2)$ ,  $\epsilon_{ij}^* \sim N(0, \sigma_{w^*}^2)$  und beide unabhängig voneinander sein sollen. Da, zwar die Verhältnisgröße der Varianzterme denen im nicht transformierten Fall ähneln, aber nicht dieselben Werte vorliegen, werden diese mit einem "Stern" versehen. Die Unabhängigkeitsannahme schränkt hier die Möglichkeit ein, dass Leute mit höherer üblicher Aufnahme auch andere (höhere) Streuung der Tag-zu-Tag Varianz haben.

**2. Varianzschätzung:** Dieser Schritt erfolgte mittels einfacher ANOVA (Personen als "Faktor"), wobei die Varianzen mittels den Quadratsummen geschätzt wurden.  $\sigma_{w^*}^2$  wird mittels der mittleren Fehlerquadratsumme (MSE) geschätzt und  $\sigma_{b^*}^2$  durch die Differenz der mittleren Modellquadratsumme und der MSE ( $\sigma_{b^*}^2 = \frac{1}{m}[MSM - MSE]$ ). Auch hier werden in den anderen Verfahren andere Arten der Varianzschätzung vorkommen. Vor allem wird hier angenommen, dass alle Personen dieselbe Tag-zu-Tag Variabilität aufweisen, was in der Praxis nicht immer der Fall ist.

**3. Varianzeliminierung:** In diesem Schritt erfolgt nun die Anwendung der Shrinkage Schätzung auf die transformierten Mittelwerte um Zwischenwerte

$$\tilde{y}_i^* = (1 - \omega^*)\bar{y}^* + \omega^* y_i^* \quad (2.5)$$

zu erhalten, welche die Verteilung der  $y_i^*$  approximieren sollen. Als Shrinkagefaktor wird hier derselbe wie im vorigen Abschnitt, allerdings mit den eingesetzten Schätzungen der Parameter, verwendet.

$$\omega^* = \sqrt{\frac{\hat{\sigma}_{b^*}^2}{\hat{\sigma}_{b^*}^2 + \hat{\sigma}_{w^*}^2/m}}. \quad (2.6)$$

**4. Rücktransformation:** Diese Zwischenwerte werden nun im 4. und letzten Schritt zurück auf die originale Skalierung transformiert. In diesem Fall mittels der Funktion  $h(\cdot) = g^{-1}(\cdot) = \exp(\cdot)$ . Diese Rücktransformation erfolgt unter einer wesentlichen Annahme, nämlich Unverzerrtheit auf der transformierten Skala. Es wird also angenommen, dass die transformierten 24-Stunden Befragungen unverzerrte Aufnahmemessungen für die transformierten üblichen Aufnahmen darstellen. Mathematisch gesehen zeigt sich dies im letzten Schritt der folgenden Gleichung:

$$E(y_{ij}^*|i) = E(y_i^*|i) = y_i^* \doteq g(y_i) \quad (2.7)$$

Da die Zwischenwerte  $\tilde{y}_i^*$  nun die  $y_i^*$  approximieren und diese wiederum die transformierte übliche Aufnahme, kann durch die einfache Rücktransformation unter Annahme 2.7 eine unverzerrte Schätzung der üblichen Verteilung erfolgen ( $h(\tilde{y}_i^*) \sim y_i$ ). Die Schätzung der Verteilung erfolgt über die empirische Verteilungsfunktion der rücktransformierten Werte. Eine weitere Konsequenz der letzten Annahme stellt die Abweichung zwischen dem nicht transformierten Stichprobenmittelwert  $\bar{y}$  und dem Mittelwert der rücktransformierten Zwischenwerte ( $= \frac{1}{m} \sum h(\tilde{y}_i^*)$ ) dar. Je nach Art (konvex, konkav) der Rücktransformation liegen die Zwischenwerte im Schnitt unter oder über dem tatsächlichen Mittelwert. Ob diese Annahme gut oder schlecht ist, liegt an der Verzerrung der 24-Stunden Befragungswerte. Da in der Praxis häufiger zu niedrig (underreporting) gemessene Werte der Fall sind (siehe Kapitel 6) ist eine konvexe Rücktransformation, wie es die Exponentialfunktion ist, eine schlechte. Sie würde in stärker verzerrten Schätzungen resultieren, als Rücktransformationen, welche durch die Annahme von unverzerrten 24-Stunden Befragungen (auf originaler Skalierung) zustande gekommen sind. Darauf wird am Ende der Beschreibung der nächsten Methode weiter eingegangen.

## 2.2.2 ISU Best Power (BP) Method

Die nächste Methode ist eine Erweiterung der NRC Methode und wurde von Nusser et al. (1996) an der Iowa State University (ISU) vorgestellt. Sie beinhaltet dieselben 4 Schritte wie vorher beschrieben.

**1. Datentransformation:** Anstatt der einfachen Log-Transformation wird hier eine Power Transformation  $g(x) = x^\alpha$  durchgeführt. Der Parameter  $\alpha$  wird dabei durch Minimierung der Anderson-Darling Statistik über eine vorbestimmte Auswahl, gewählt ( $\alpha \in [1, 1.5^{-1}, 2^{-1}, \dots, 10^{-1}]$ ). Die Anderson-Darling Statistik ist ein Anpassungstest an beliebige Verteilungen unter anderem an die Normalverteilung (Anderson and Darling (1952)). Die Teststatistikberechnung erfolgt unter anderem durch die Abweichung der empirischen Verteilung zur getesteten. Je größer die Werte der Anderson-Darling Statistik desto weniger wahrscheinlich stammen die Daten aus der getesteten Verteilung. An dieser Stelle kann natürlich darüber diskutiert werden, ob dies der beste Normalverteilungstest ist. Da dies den Umfang der Arbeit überschreiten würde, wird hier nur kurz erwähnt, dass der Shapiro-Wilk Test auf Normalität auch eine Möglichkeit wäre, da dieser in diversen Simulationen bessere Power gezeigt hat (Razali and Wah (2011)). So oder so stellt diese Art der Power Transformation auf jeden Fall eine flexiblere Variante als die der NRC Methode dar. Die Annahmen des transformierten Modells (Gl: 2.4) werden hier genauso wie in der NRC Methode getroffen.

**2. Varianzschätzung:** Die Berechnung der between-subject Varianz erfolgt über die geschätzte Varianz der transformierten Mittelwerte  $\bar{y}_i^*$

$$\widehat{Var}(\bar{y}_i^*) = \frac{1}{n-1} \sum_{i=1}^n (\bar{y}_i^* - \bar{y}^*)^2 \quad (2.8)$$

und über die Schätzung der within-person Varianz

$$\hat{\sigma}_{w^*}^2 = \frac{1}{n(m-1)} \sum_{i=1}^n \sum_{j=1}^m (y_{ij}^* - \bar{y}_i^*)^2. \quad (2.9)$$

Mit diesen zwei Varianzschätzern lässt sich nun durch Einsetzen in die Gl:2.2  $\hat{\sigma}_{b^*}^2$  ermitteln. Auch hier wird wieder angenommen, dass alle Personen dieselbe Tag-zu-Tag Variabilität aufweisen. Beachte, dass diese Schätzungen dieselben Ergebnisse bringen, wie die ANOVA Methode bei NRC.

**3. Varianzeliminierung:** Dieser Schritt ist wiederum gleich wie in der NRC Methode über die Verwendung der Shrinkageschätzung (Gl: 2.5 und 2.6).

**4. Rücktransformation:** Die Rücktransformation läuft zwar auch über die Inverse der Powertransformation  $h(x) = g^{-1}(x) = x^{1/\alpha}$  aber unter einer anderen Annahme als bei der NRC Methode. Angenommen wird hier, dass die 24-Stunden Befragungen erwartungstreue Schätzungen auf der nicht transformierten Skala liefern ( $E(y_{ij}|i) = y_i$ ). Im Grunde genommen wird für die Gleichung 2.3 angenommen, dass die Fehlerterme  $\epsilon_{ij}$  Erwartungswert 0 haben (Beachte: dies wurde für die NRC Methode nicht angenommen).

Da hier nun nicht wie in der vorigen Methode angenommen wird, dass  $y_i^* = g(y_i)$  gilt, muss die Rücktransformation über einen anderen Ansatz gemacht werden. Dazu wird von der angenommenen Erwartungstreue auf der nicht transformierten Skala ausgegangen und über die Taylorreihenapproximation ( $h(x) \approx h(a) + h'(a)(x - a) + 1/2 h''(a)(x - a)^2$  mit  $x = y_{ij}^*$  und  $a = E(y_{ij}^*|i)$ ) abgeschätzt.

$$y_i = E(y_{ij}|i) = E(h(y_{ij}^*)|i) \approx E \left( h(E(y_{ij}^*|i)) + h'(E(y_{ij}^*|i))(y_{ij}^* - E(y_{ij}^*|i)) + \frac{1}{2} h''(E(y_{ij}^*|i))(y_{ij}^* - E(y_{ij}^*|i))^2 | i \right)$$

Dies kann natürlich noch weiter vereinfacht werden. Beachte, dass im nächsten Schritt der mittlere Summenterm verschwindet, da der bedingte Erwartungswert hineingezogen wird und dadurch der mittlere Term 0 wird. Ebenso gilt  $E(y_{ij}^*|i) = y_i^*$  was zu folgender Formel führt.

$$\dots = h(y_i^*) + 0 + \frac{1}{2} h''(y_i^*) E((y_{ij}^* - E(y_{ij}^*|i))^2 | i) = h(y_i^*) + \frac{1}{2} h''(y_i^*) \sigma_{w^*}^2 \quad (2.10)$$

Die letzte Gleichung gilt, da  $Var(y_{ij}^*|i) = Var(\epsilon_{ij}^*|i) = \sigma_{w^*}^2$ . Der Ergebnisterm verhält sich also approximativ wie die  $y_i$ . Im Gegensatz zur NRC Methode ist ein Bias-Anpassungsterm hinzugekommen, welcher berücksichtigt werden kann, solange die Inverse der Transformation eine zweifache Ableitung besitzt. Im Fall der Best Power Methode ist  $h(x)'' = \frac{1}{\alpha}(\frac{1}{\alpha} - 1)x^{1/\alpha - 2}$  und somit immer existent (Beachte:  $x = \tilde{y}_i^* \geq 0$ ). Um die Rücktransformation nun zu ermöglichen wird für die within-person Varianz in Gleichung 2.10 ihre Schätzung und für die  $y_i^*$  die aus dem vorigen Schritt berechneten Zwischenwerte  $\tilde{y}_i^*$  eingesetzt. Mit diesen rücktransformierten Werten kann nun eine empirische Verteilung geschätzt werden, welche die übliche Aufnahmeverteilung approximiert.

Der große Unterschied der Best Power Methode zur NRC Methode liegt einerseits im variablen Transformationsschritt und im letzten Teil durch den Bias-Anpassungsterm. Dieser Term ist für  $\alpha \leq 1$  immer positiv und produziert dadurch höhere Werte als die NRC Methode, was im Angesicht von verzerrten 24-Stunden Befragungen (underreporting) für die Best Power Methode spricht. (Dodd et al., 2006, Figure 3) erwähnt, dass die Werte für  $\alpha$ , durch die rechtsschiefen Daten, meist kleiner gleich 1 sein werden und dies daher oft erfüllt ist (BP Methode hat sowieso nur Werte kleiner 1 möglich). Die Best Power Methode der Iowa State University wäre also meist gegenüber der NRC Methode zu bevorzugen.

Es soll noch erwähnt werden, dass die Best Power Methode nur eine Vereinfachung der komplexeren Methode von Nusser et al. (1996) darstellt. Diese komplexere Variante wird später noch vorgestellt.

## S-Nusser Methode

Da die ISU Best Power Methode nur eine einfache Powertransformation, sowie eine Approximation der Rücktransformation (Gl: 2.10) verwendet, kann diese sehr einfach erweitert und dadurch verbessert werden. Hoffmann et al. (2002) hat eine solche Erweiterung der Best Power Methode vorgeschlagen und diese "simplified Nusser" Methode genannt. Diese beinhaltet statt der Powertransformation eine zwei-Parameter Box Cox Transformation der Daten (Box and Cox (1964))

$$g(x) = \begin{cases} \frac{(x+\lambda_2)^{\lambda_1}}{\lambda_1} & \text{für } \lambda_1 \neq 0 \\ \ln(x + \lambda_2) & \text{für } \lambda_1 = 0, \end{cases} \quad (2.11)$$

welche eine flexiblere Variante der Normalverteilungstransformation darstellt. Die Parameter  $\lambda_1$  und  $\lambda_2$  werden dabei über die Maximierung der Shapiro-Wilk Statistik gewählt. Die Auswahl der Parameter umfasst hier ein sehr ähnliches Raster wie die BP Methode

( $\lambda_1 \in [1, 1/2, 1/3, \dots, 1/10, 0]$ ,  $\lambda_2$  selbes Raster multipliziert mit dem Mittelwert der originalen Daten), allerdings mit weniger Möglichkeiten. Dies hat den Grund, dass für die Rücktransformation nicht die Taylorreihenapproximation (2.10) genutzt wird, sondern eine exakte Berechnung, welche nur für gewisse  $\lambda_1$  Werte analytisch bestimmt werden kann. Die Schritte 2. und 3. sind allerdings gleich wie bei der BP Methode.

Da der Schritt der Rücktransformation eine wichtige Rolle für die Schätzung der üblichen Aufnahmeverteilung spielt, soll dies hier genauer betrachtet werden. Wichtig ist hierbei die Berechnung des bedingten Erwartungswertes  $E(y_{ij}|i) = E(h(y_{ij})|i)$ . Mathematisch gesehen, kann dies als Integral über die Fehlerterm ( $= \epsilon$ ) Verteilung gesehen werden:

$$y_i = E(h(y_i^* + \epsilon_{ij}^*)|i) = \int_{-\infty}^{\infty} h(y_i^* + \epsilon)\varphi(\epsilon)d\epsilon, \quad (2.12)$$

wobei  $\varphi$  die Dichte der  $N(0, \sigma_w^2)$  darstellt. Während Gl: 2.12 für den Fall  $\lambda_1 = 0$  ( $g(\cdot) = \ln(\cdot)$ ) noch einfacher zu berechnen ist (Gl:2.12 wäre hier  $\exp(y_i^* + \frac{1}{2}\sigma_w^2) - \lambda_2$ ), so stellt die Berechnung des Integral für allgemeinere  $\lambda_1 > 0$  ein größeres Problem dar. Analytisch kann es nur für die Spezialfälle  $\lambda_1 = \frac{1}{\delta}$  für  $\delta = 1, 2, \dots$  bestimmt werden und ergibt:

$$\int_{-\infty}^{\infty} h(y_i^* + \epsilon)\varphi(\epsilon)d\epsilon = \delta^{-\delta} \sum_{s=0}^{int(\delta/2)} \binom{\delta}{2s} (y_i^* + \delta)^{\delta-2s} \sigma_w^{2s} (2s-1)!! - \lambda_2, \quad (2.13)$$

wobei hier *int* für Integer und  $x!!$  für das Produkt aller ungeraden Zahlen von 1 bis  $x$  stehen. Die Parameter  $\delta$  und  $\sigma_w^2$  müssen natürlich noch durch ihre Schätzungen  $\hat{\delta} = \hat{\lambda}_1^{-1}$  sowie  $\hat{\sigma}_w^2$  ersetzt werden um approximative  $\tilde{y}_i$  zu erhalten.

Für die Berechnung der beiden Integrale ( $\lambda_1 = 0$  und  $\lambda_1 > 0$ ) siehe (Hoffmann et al., 2002, Appendix). Für allgemeinere Fälle muss das Integral 2.12 numerisch gelöst werden, worauf wir später noch näher eingehen werden.

### 2.2.3 Bisherige Annahmen

Die für die Praxis einschränkenden Annahmen, die diese “noch einfachen“ Methoden treffen, sollen hier nochmal kurz zusammengefasst werden (Die Zahlen in den runden Klammern bezeichnen den Schritt, in dem die Annahme notwendig ist):

- Transformation zu approximativer Normalität ist mit Log, Power oder Box Cox Transformation möglich (1)
- Unabhängigkeit der Terme  $y_i^*$  und  $\epsilon_{ij}^*$  (2)
- gleiche Varianzen  $\sigma_w^2$  für alle Personen (2,4)
- Erwartungstreue auf transformierter oder originaler Skalierung (4)

Diese Annahmen stellen nicht immer plausible Anforderungen dar und Methoden, die diese auflockern wären zu bevorzugen. Vor allem eine breitere Transformationsklasse wäre wünschenswert, um die Annahme der Normalität noch plausibler zu machen (nicht nur für eingeschränkte Parameter  $\alpha$  und  $\lambda_1$ ). Erwähnt soll noch werden, dass die Unabhängigkeit der Terme  $y_i^*$  und  $\epsilon_{ij}^*$  und die Annahme der gleichen Varianzen für alle Personen nur für die transformierten Daten gelten müssen. Sollten die Annahmen, für die “originalen“ Daten, verletzt sein, so kann es sein, dass die Verletzungen der Annahmen weniger drastisch im transformierten Rahmen sind.

#### weitere Probleme

Die zwei genannten Methoden haben noch eine weitere Einschränkung, die bis jetzt außer Acht gelassen wurde. Bei Ernährungsbefragungen kommt es auch vor, dass Personen gewisse Lebensmittel(gruppen) oder Nährstoffe in dem Beobachtungszeitraum nicht konsumiert haben und es deswegen zu vielen Aufnahmewerten gleich null kommt. Sowohl Personen, die tatsächlich Nicht-Konsumenten dieses Lebensmittel oder Nährstoffes sind, sowie Personen die nur gerade an den beobachteten Tagen diese nicht zu sich genommen haben, werden somit als Nicht-Konsumenten eingestuft und es kommt zu groben Fehlschätzungen (bzw.

ist die Transformation der Daten nicht mal möglich). Die 3 bisher vorgestellten Methoden berücksichtigen diese Problematik in der Modellierung nicht und liefern daher für Nährstoffe oder Lebensmittel, die nicht täglich konsumiert werden keine guten Schätzungen der üblichen Aufnahmeverteilung. Es werden also Methoden gebraucht, welche dieses Problem berücksichtigen und zu modellieren versuchen. Wie wir sehen werden, haben die nächsten Methoden solche Erweiterungen, jedoch mit verschiedenen Ansätzen. Der Fokus darauf wird allerdings klein gehalten, da wie im 3. Kapitel zu sehen sein wird, für die hier ausgewerteten Nährstoffe das Problem nicht vorkommt und diese extra Modellierung somit wegfällt. Eine weitere wichtige Anpassung und in diesem Fall wichtige für den hier untersuchten Datensatz, sollte ebenfalls beachtet werden. Da Ernährungserhebungen nicht immer einfache Zufallsstichproben sind, sondern komplexe Surveys, werden oft Gewichtungsfaktoren zur Korrektur der Dropoutrate oder Verzerrungen in der Stichprobe verwendet. Diese sollen die untersuchte Population besser modellieren und daher ebenfalls in die Schätzung mitaufgenommen werden können. Beim Ernährungsbericht 2012 wurde mit Gewichten gearbeitet um eine möglichst gute Angleichung der Stichprobe an die Gesamtbevölkerung Österreichs zu erhalten (Elmadfa et al. (2012)). Es wird in dieser Arbeit also großer Wert darauf gelegt Methoden zu verwenden, welche die Behandlung von Survey Gewichten ermöglicht.

## 2.3 Hauptmethoden dieser Arbeit

In diesem Abschnitt sollen die 4 Hauptmethoden, die auch heute in der Praxis verwendet werden, erklärt und verglichen werden. Diese 4 Methoden haben alle die Erweiterung zu einem 2 Teile Modell (2-part model), wobei ein Teil die Schätzung der Aufnahmewahrscheinlichkeit des Nahrungsbestandteils und ein Teil die Schätzung der üblichen Aufnahmeverteilung beinhaltet. Dabei wird zuerst, entweder unabhängig oder abhängig vom 2. Teil, die Aufnahmewahrscheinlichkeit  $p_i$  für jedes Individuum mittels Kovariaten und der vorhandenen Daten geschätzt und anschließend mit der Schätzung der üblichen Aufnahme  $y_i$  multipliziert. Dadurch erhält man eine "angepasste" übliche Aufnahmeverteilung, welche die nicht tägliche Aufnahme des Nahrungsbestandteiles berücksichtigt. Da für die, in dieser Arbeit, untersuchten Nährstoffe (Makro und Mikronährstoffe) keine Modellierung der Aufnahmewahrscheinlichkeit notwendig ist, wird dieser Teil nur als Zusatz erwähnt und der Fokus auf den Teil der Schätzung der Aufnahmeverteilung gelegt.

### 2.3.1 National Cancer Institute Methode (NCI)

Die erste betrachtete Methode kommt vom National Cancer Institute und wurde erstmals von Tooze et al. (2006) vorgestellt. Sie besteht aus einem 2 Teile Modell und kann damit sowohl für Nahrungsstoffe, welche täglich oder nur hin und wieder konsumiert werden, verwendet werden. Die Modellierung für täglich aufgenommene Nährstoffe wurde in einem extra Paper nochmals beschrieben, siehe Tooze et al. (2010). Es handelt sich hierbei um eine gemischte Effekte Modellierung der transformierten Variablen, mit möglicher Berücksichtigung von Kovariaten. Es wird nun zuerst der Part dieser Schätzung beschrieben werden und anschließend die mögliche Erweiterung in ein 2 Teile Modell.

**1. Datentransformation:** Der erste Schritt besteht auch hier aus einer Transformation zu approximativer Normalität der Tagesbeobachtungen. Dieser Schritt wird ähnlich wie in der S-Nusser Methode mittels Box Cox Transformation durchgeführt. Anstatt der 2 Parameter Box Cox Variante (Gl:2.11) wird nur die einfache Box Cox Transformation verwendet.

$$g(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{für } \lambda \neq 0 \\ \ln(x) & \text{für } \lambda = 0, \end{cases} \quad (2.14)$$

Der Parameter  $\lambda$  wird hier allerdings nicht mittels Rastersuche und vor den Schritten der Varianzzerlegung durchgeführt, sondern erst später im Schritt der Maximierung der Likelihood für das gemischte Effekte Modell. Dies hat den Vorteil, dass einerseits der Parameter  $\lambda$  alle möglichen Werte (im Gegensatz zu einer Vorauswahl) annehmen und somit besser transformiert werden kann und andererseits die Parameterwahl auch an die verwendeten Kovariaten angepasst wird. Mathematisch betrachtet ist die NCI Methode ähnlich darstellbar



wie vorher in Gl:2.1. Hier wird der Gesamtmittelwert  $\mu$  allerdings durch einen zusätzlichen Kovariatenvektor  $X_i = [Kovariate_1, Kovariate_2, \dots]'$  modelliert.

$$g(y_{ij}) = X_i' \beta + \mu_i^* + \epsilon_{ij}^* = y_i^* + \epsilon_{ij}^* \quad (2.15)$$

Hier werden die Parameter  $\mu_i^*$  und  $\epsilon_{ij}^*$  als normalverteilte Zufallsvariablen mit Erwartungswert 0 und Varianzen  $\sigma_{b^*}^2$  und  $\sigma_{w^*}^2$  betrachtet. Die mittlere Aufnahme  $y_i^*$  jeder Person ist nun ähnlich wie bei der BP und NRC Methode normalverteilt, allerdings mit einem Erwartungswert der nicht nur aus einem Intercept besteht ( $y_i^* \sim N(X_i' \beta, \sigma_{b^*}^2)$ ).

Als mögliche Kovariaten können alle Variablen verwendet werden, welche einen möglichen (plausiblen) Einfluss auf die Nährstoffaufnahme haben, ebenso wie Variablen anhand denen eine Analyse von Subgruppen interessant ist (z.B. Geschlecht, Alter, usw.). Als mögliche Kovariaten können auch Variablen von einem Verzehrshäufigkeitenfragebogens (Food Frequency Questionare), der zusätzlich zu den 24 Stunden Erinnerungsfragebögen ausgefüllt wurde, verwendet werden. Solche Informationen können unter anderem die Verwendung von Nahrungsergänzungsmitteln oder bestimmte Ernährungsweisen (vegan, vegetarisch, usw.) sein. Ein weiterer Vorteil in der Verwendung von Kovariaten ist, dass die within-person Varianzen dabei anhand von Kovariatleveln variabel modelliert werden können.

**2. Varianzschätzung:** Das Modell in Gleichung 2.15 kann auch als nichtlineares gemischte Effekte Modell betrachtet werden. Obwohl die fixen Effekte  $\beta$  und zufälligen Effekte  $\mu_i^*$  sowie  $\epsilon_{ij}^*$  als lineare Prädiktoren ins Modell aufgenommen werden, so ergibt sich doch durch die Box Cox Transformation ein nicht lineares Modell. Geschätzt werden nun alle Parameter gleichzeitig mittels Likelihood Optimierung. Dies erfolgt in der NCI Methode mittels Quasi-Newton Algorithmus.

**3. Varianzeliminierung:** Im Gegensatz zu den vorher schon beschriebenen Methoden verwendet diese nicht einen Shrinkage Faktor um die übliche Aufnahme  $y_i^*$  auf der transformierten Skala zu approximieren, sondern simuliert direkt Werte aus einer Verteilung mit gewünschter between-person Varianz. Es werden zuerst für jedes Individuum  $i = 1, \dots, n$  100 Datenpunkte  $\tilde{\mu}_i$  aus einer Normalverteilung mit Erwartungswert 0 und Varianz  $\hat{\sigma}_{b^*}^2$  simuliert. Diese werden dann anschließend in die Zwischenwerte  $\tilde{y}_i^* = X_i' \hat{\beta} + \tilde{\mu}_i$  umgeformt. Diese  $100n$  Zwischenwerte stellen nun eine Approximation der üblichen Aufnahmeverteilung auf transformierter Skala dar. Im Gegensatz zur Shrinkage Methode mittels 2.6 wird hier die Normalverteilungsannahme stärker strapaziert. Während der Shrinkage Faktor noch die ursprüngliche Verteilungsform der Daten zum Teil erhalten kann, nimmt die NCI Methode stattdessen direkten Gebrauch von der Normalverteilungsannahme und kann somit bei schlechter Datentransformation zu noch schlechteren Ergebnissen führen.

**4. Rücktransformation:** Die Zwischenwerte müssen jetzt noch auf die ursprüngliche Skalierung zurückgebracht werden, was auch hier unter der Annahme von unverzerrten Daten auf originaler Skala stattfindet (Gl:2.12). Da allgemeine  $\lambda$  Parameter geschätzt werden, kann nicht dieselbe Rücktransformation wie in Gl:2.13 verwendet werden. Ursprünglich wurde die Taylorreihenapproximation (Gl:2.10) im Paper von Tooze et al. (2010) vorgeschlagen. Diese wurde inzwischen schon durch eine genauere 9 Punkt Approximationsmethode (wie auch in der allgemeinen Methode von Nusser et al. (1996)) ersetzt. Dabei werden Gewichte  $w_k$  und Knoten  $c_k$  für  $k = 1, \dots, 9$  so gewählt, dass die diskrete "9 Punktverteilung" dieselben ersten fünf Momente, wie die Verteilung der  $\tilde{y}_i^*$  bedingt auf  $X_i' \beta$  hat. Die Gewichte haben ebenso noch die Restriktion, dass ihre Summe gleich 1 sein muss. Die Rücktransformation erfolgt dann mittels der Formel:

$$y_i \approx \sum_{k=1}^9 w_k g^{-1}(\tilde{y}_i^* + c_k \sigma_{w^*}) \quad (2.16)$$

Natürlich müssen auch hier noch die geschätzten Parameter  $\hat{\lambda}$  und  $\hat{\sigma}_{w^*}$  eingesetzt werden um die Rücktransformation möglich zu machen. Mit diesen  $100n$  Werten  $\tilde{y}_i$  kann nun die übliche Aufnahmeverteilung mittels der empirischen Verteilungsfunktion geschätzt werden. Bemerkung: Die Inverse der Funktion  $g$  ist nur für nicht negative Werte möglich, daher setzt diese Methode das Gewicht  $w_k = 0$  für negative Werte innerhalb der Klammer.

## Schätzung von nicht täglich konsumierten Nahrungsbestandteilen

Wie schon erwähnt besitzt die NCI Methode auch eine Erweiterung zur Modellierung der Aufnahmewahrscheinlichkeit (Tooze et al. (2006)). Dies erfolgt mittels logistischer Regression mit einem personen-spezifischem Zufallseffekt. Ebenso können hier Kovariaten verwendet werden, welche nicht zwingend dieselben sein müssen wie im vorigen Teil. Die Modellierung und Schätzung der Aufnahmewahrscheinlichkeit und der üblichen Aufnahmeverteilung erfolgt simultan und erlaubt auch Korrelationsstrukturen zwischen den beiden Parts. Einerseits durch Verwenden gleicher Kovariaten in beiden Teilen oder durch Modellierung von Korrelation zwischen den Zufallseffekten der beiden Teile. Somit muss keine Unabhängigkeit der  $p_i$  und  $y_i$  angenommen werden, was in der Praxis wichtig ist, da die Aufnahmehäufigkeit eines Nahrungsmittels auch mit der Aufnahmemenge korrelieren kann. Die Schätzung der üblichen Aufnahmeverteilung erfolgt dann durch Multiplizieren der simulierten rücktransformierten Zwischenwerte  $\tilde{y}_i$  mit den geschätzten  $\tilde{p}_i$  aus dem logistischen Modell.

### weitere Möglichkeiten

Die NCI Methode ermöglicht ebenfalls die Verwendung von Populationsgewichten zur Anpassung an das untersuchte Kollektiv. Es können also auch mit Daten, die nicht aus einfachen Zufallsstichproben kommen, repräsentative Auswertungen und Schätzungen vorgenommen werden. Die Gewichte werden in den einzelnen Schritten und in der Modellierung mit berücksichtigt.

Ein weiterer Vorteil der NCI Methode ist die Auswertung für Subgruppen. Dies erfolgt durch eine zusätzliche Kovariate, die die Subgruppen bestimmt. Die Schätzung der Aufnahmeverteilung erfolgt dann wie oben beschrieben, nur mit den richtigen Kovariatwerten für die Subgruppe. Hier kann auch die Varianzschätzung für die unterschiedlichen Subgruppen angepasst werden, indem für die Gruppen ungleiche Varianzen im nicht linearen Modell zugelassen werden.

### Implementierung der Methode

Die NCI Methode kann mittels eines vom National Cancer Institute zur Verfügung gestellten SAS Makro ausgeführt werden. Zusätzlich dazu gibt es noch ein User Manual auf der Homepage mit Erklärungen zu den Spezifikationen der Funktionen im SAS Makro.

### 2.3.2 Multiple Source Methode (MSM)

Einen etwas anderen Ansatz der Schätzung ermöglicht die Multiple Source Methode von Harttig et al. (2011). Es handelt sich aber auch hier um ein 2-Teile Modell für den Fall von vielen Null-Messwerten. Der erste Part dient wieder zur Schätzung der Aufnahmewahrscheinlichkeit und wird deswegen nach dem 2. Teil der Methode kurz erklärt. Die Erklärung zur Modellierung findet man in (Haubrock et al., 2011, Online Supporting Material).

**0. Mittelwertmodellierung:** Im ersten Schritt noch vor einer Datentransformation wendet die MSM ein lineares Regressionsmodell auf die Tagesbeobachtungen an. Mögliche unabhängige Variablen sind dabei das Alter, Geschlecht, eine mögliche Interaktion der beiden Variablen oder Informationen von einem Verzehrshäufigkeitsfragebogens. Es können auch andere Kovariablen wie BMI oder ähnliches genutzt werden. Ziel ist es den Gruppenmittelwert  $\mu$  bzw. einen Teil der between-person Varianz anhand von Kovariaten zu modellieren. Anschließend kann jede Beobachtung als Schätzung des Modells plus einem Residuenterm gesehen werden  $y_{ij} = X_i' \hat{\beta} + e_{ij}$ . Die Residuen beinhalten nun noch den Teil  $\mu_i + \epsilon_{ij}$  aus der Gleichung 2.1 und werden daher zur weiteren Modellierung verwendet, um den within-person und (nicht erklärten) between-person Anteil zu bestimmen.

**1. Datentransformation:** Die Residuen werden nun mittels 2 Parameter Box Cox Transformation (wie in Gl:2.11) transformiert. Als mögliche Parameterwerte für  $\lambda_1$  kommen wieder nur Kehrwerte von positiven ganzen Zahlen (bis 10) zur Verwendung. Ebenso wie bei der S-Nusser Methode wird die Shapiro-Wilk Statistik zur optimalen Parameterwahl

verwendet. Die transformierten Residuen werden dann wieder in einem klassischem Messfehler Modell dargestellt:

$$e_{ij}^* = g(e_{ij}) = \mu_i^* + \epsilon_{ij}^*, \quad (2.17)$$

wobei  $\mu_i$  mit  $E(\mu_i^*) = 0$ ,  $Var(\mu_i^*) = \sigma_{b^*}^2$  und  $\epsilon_{ij}^* \sim N(0, \sigma_{w^*}^2)$  unabhängig voneinander sind.  $\sigma_{b^*}^2$  stellt hier wieder die Varianz zwischen den Personen und  $\sigma_{w^*}^2$  die Tag-zu-Tag Variabilität dar.

**2. Varianzschätzung:** Ähnlich wie schon bei der BP-Methode werden auch hier die Varianzen des Modells 2.17 mit denselben Formeln 2.8, 2.9 und 2.2 berechnet.

**3. Varianzeliminierung:** Hier wird nun wieder die Shrinkagemethode (Gl:2.6 und Gl:2.5) auf die  $\mu_i^*$  angewendet um Zwischenwerte  $\tilde{\mu}_i^*$  mit gewünschter Varianz  $\sigma_{b^*}^2$  zu bekommen.

**4. Rücktransformation:** Diese erfolgt wie bei der S-Nusser Methode mit der exakten Formel 2.13. Statt den  $y_i^*$  Werten werden hier die  $\tilde{\mu}_i^*$  Werte eingesetzt. Für unbalancierte Daten (nicht alle Personen haben gleich viele Tagesbeobachtungen) muss noch ein Korrekturfaktor für die Transformation verwendet werden. Die nun erhaltenen Werte  $\tilde{\mu}_i$  werden nun zurück in die ursprüngliche Regressionsgleichung für  $e_{ij}$  eingesetzt. Man erhält nun für jede Person die Modellgleichung:

$$\tilde{y}_i = X_i' \hat{\beta} + \tilde{\mu}_i$$

als Schätzung für die übliche Aufnahmemenge pro Person. Die übliche Aufnahmeverteilung kann wieder über die empirische Verteilungsfunktion geschätzt werden.

Bemerkung: Sollte es das Ziel sein für jedes Individuum eine bestmögliche Schätzung der Langzeitaufnahme zu bestimmen, so sollten nicht die Werte  $\tilde{y}_i$  dafür verwendet werden. Diese dienen nur um die übliche Aufnahmeverteilung der gesamten Gruppe zu schätzen, da sie in Richtung Gesamtgruppenmittelwert "gezogen" wurden. Als Schätzung auf Individual-ebene ist der Mittelwert der Tagesbeobachtungen je Person eine viel bessere Methode zur Bestimmung der Langzeitaufnahme.

### Schätzung von nicht täglich konsumierten Nahrungsbestandteilen

Wie schon erwähnt hat auch die Multiple Source Methode eine Erweiterung für dieses Problem. Wie die NCI Methode verwendet die MSM auch ein logistisches Regressionsmodell dafür, allerdings ohne Zufallseffekt. Als Kovariaten können wieder dieselben wie im 2. Teil der Methode verwendet werden (müssen aber nicht die gleichen sein). Auch hier wird eine ähnliche, wie gerade für den 2. Teil beschriebene Vorgehensweise angewandt. Die beobachteten  $P_{ij} \in \{0, 1\}$  (Person  $i$  konsumierte am Tag  $j$  den Nahrungsbestandteil) werden als Gleichung  $P_{ij} = M_i + e_{ij}$  betrachtet.  $M_i$  stellt die Modellschätzung der logistischen Regression für Person  $i$  dar. Auf die Residuen  $e_{ij} \in [-1, 1]$  soll nun die Transformation  $g(e) = \ln((1+e)/(1-e))$  angewandt werden um reelle Zahlen von  $-\infty$  bis  $\infty$  zu erhalten. Nun wird wieder ein Messfehler Modell angewandt und anschließend die Varianzkomponenten (within, between) geschätzt. Die mittels Shrinkagefaktor transformierten Zwischenwerte werden dann wieder für die Gleichung  $p_i = M_i + h(\tilde{b}_{i(1)})$  verwendet um Aufnahmewahrscheinlichkeitsschätzungen für jede Person zu bekommen.  $h$  stellt dabei die Inverse der Funktion  $g$  plus einen Korrekturfaktor dar.

Wie in der NCI Methode erhält man nun die übliche Aufnahmeverteilung durch Multiplizieren der Aufnahmewahrscheinlichkeiten mit den Schätzungen  $\tilde{y}_i$  aus Part 2 der Methode. Anders als bei der NCI Methode kann hier allerdings keine bestimmte Korrelationsstruktur der beiden Teile definiert werden. Eine gewisse Abhängigkeit kann aber trotzdem, durch verwenden derselben Kovariaten in beiden Teilen, entstehen.

### weitere Möglichkeiten

MSM ermöglicht nicht die Verwendung von Datengewichten. Dies stellt einen großen Nachteil der Methode dar, da prinzipiell keine repräsentativen Schätzungen für nicht einfache Zufallsstichproben möglich sind. Wir werden später allerdings sehen, dass ein "neuer" Datensatz aus

den nicht gewichteten Daten in Kombination mit den Gewichtungsfaktoren erstellt werden kann, welcher sich verhält wie die gewichteten Daten. Dieser neue Datensatz kann allerdings nun wie ein nicht gewichteter Datensatz verwendet werden und liefert trotzdem Ergebnisse, die ähnlich denen einer gewichteten Analyse sind.

### Implementierung der Methode

Die MSM ist zum Großteil mittels der Software R implementiert worden und für Anwender auf einer Homepage zur Verfügung gestellt (Harttig et al. (2011)). Es ist auch ein User Manual vorhanden, welches dem Verwender das Hochladen der Daten und die Anwendung der Methode erklärt.

### 2.3.3 Statistical Program to Assess Dietary Exposure (SPADE)

Die Schätzung der üblichen Aufnahmeverteilung von Nahrungsstoffen kann bei der SPADE (Statistical Program to Assess Dietary Exposure) Methode in verschiedenen Modellen erfolgen (Dekkers et al. (2014)). Je nachdem ob einfache tägliche Tagesbeobachtungen (1-Part Modell), hin und wieder konsumierte Nahrungsinhalte (2-Part Modell), Nahrungsergänzungsmittel (2-Part Modell) oder die Kombination der letzten beiden (3-Part Modell) modelliert werden sollen, kann die SPADE Methode angepasst werden. Es gibt auch ein Multipartmodell für die Modellierung der Aufnahme von mehreren Essensquellen in einem. Das 1-Teile Modell wird hier wieder ausführlich erklärt und anschließend das 2-Teile Modell für nicht täglich konsumierte Nahrungsinhalte. Für die anderen Modelle wird auf Dekkers et al. (2014) verwiesen.

**1. Datentransformation:** Wie die anderen bisher gezeigten Methoden hat auch SPADE einen Transformationsschritt der Daten. Eine 1-Parameter Box Cox Transformation soll die Tagesbeobachtungen  $y_{ij}$  symmetrisch und annähernd normal verteilt hinbekommen. Der Parameter  $\lambda$  wird mittels Maximum Likelihood Methode geschätzt und kann alle möglichen Werte größer gleich 0 annehmen. Für Werte kleiner gleich 0 wird automatisch die Logtransformation verwendet.

Im nächsten Schritt wird ein fraktionales Polynom bezüglich des Alters der Personen für die transformierten Daten geschätzt. Dies setzt sich wie folgt zusammen:

$$y_{ij}^* = \begin{cases} a + bz_i^p + cz_i^q + e_{ij} & \text{für } p \neq q \\ a + bz_i^p + cz_i^q \ln(z_i) + e_{ij} & \text{für } p = q, \end{cases} \quad (2.18)$$

wobei  $z_i$  die Variable Alter der Person  $i$  darstellt. Beachte, dass hier, ähnlich wie bei der MSM Methode, die Residuen  $e_{ij}$  noch den between- und within-person Anteil beinhalten. Die Parameter  $p$  und  $q$  können die Werte  $\{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$  annehmen. Die Werte werden dabei wie in Royston and Sauerbrei (2008) (ab Seite 74) beschrieben geschätzt und ausgewählt: Unter der Normalverteilung der transformierten Variable wird das beste Polynom durch Maximum Likelihood Schätzung gefunden. Dabei werden für alle 36 (alle Powerkombinationen) möglichen Modelle die Parameter  $a$ ,  $b$  und  $c$  mittels Maximum Likelihood Methode geschätzt. Anschließend wird das Polynom mit dem höchsten Likelihood Wert als bestes gewählt. Ob das Modell schlussendlich verwendet wird, wird noch mittels Testprozedur beschlossen. Diese ist ebenfalls in SPADE implementiert und wird in Royston and Sauerbrei (2008)(Kapitel 4.10.2) beschrieben. Das beste Modell wird dabei gegen ein einfaches Regressionsmodell (Intercept + Steigung) sowie gegen ein einfaches Polynom getestet (Likelihoodratiotest). Sollte das durch Maximum Likelihood bestimmte Polynom jedes mal signifikant "besser" sein, so wird es verwendet. Ansonsten wird eines der zwei einfacheren Modelle gewählt. Die dabei geschätzten Parameter  $\hat{p}$  und  $\hat{q}$ , welche die Form der Kurve festlegen, werden für die nächsten Schritte weiter verwendet. Die anderen Parameter werden (nochmal) im Kontext eines linearen gemischten Modell geschätzt. Die Residuen  $e_{ij}$  von Gleichung 2.3.3 werden nun noch in den gewünschten between-person  $\mu_i$  und within-person Anteil  $\epsilon_{ij}$  aufgeteilt.

$$y_{ij}^* = a + bz_i^{\hat{p}} + cz_i^{\hat{q}} + \mu_i + \epsilon_{ij} \quad (2.19)$$

$\mu_i$  ist dabei ein zusätzlicher Zufallseffekt mit Normalverteilung  $N(0, \sigma_{b^*}^2)$ . Die Fehler  $\epsilon_{ij}$  folgen auch einer Normalverteilung mit Erwartungswert 0 und konstanter within-person Varianz  $\sigma_w^2$  (auch übers Alter hinweg konstant).

**2. Varianzschätzung:** Die Parameter werden nun mittels “normaler“ (für die fixen Effekte) und für die Varianzen über restriktierte Maximum Likelihood Schätzung (REML) ermittelt. Bei der REML Schätzung wird die Likelihood faktorisiert. Dabei erhält man einen Teil der nicht von den Varianzkomponenten und einem Teil der davon abhängt. Letzterer wird dann für die Schätzung der Varianzen verwendet und resultiert in weniger verzerrten Schätzungen als die normale ML Methode.

**3. Varianzeliminierung:** Im Gegensatz zu den anderen Methoden werden bei SPADE keine Simulationen oder Shrinkagefaktoren angewandt. Stattdessen werden direkt aus der Normalverteilung  $N(\hat{a} + \hat{b}z^{\hat{p}} + \hat{c}z^{\hat{q}}, \sigma_{b^*}^2)$  für die gesuchte Altergruppe (=z) die gesuchten Quantile, bzw. der Mittelwert für die Rücktransformation auf die originale Skalierung verwendet.

**4. Rücktransformation:** Da durch die Box Cox Transformation allgemeine  $\lambda$  Werte und nicht nur die spezifischen, wie in der S-Nusser und MSM Methode, möglich sind, kann keine exakte (analytische) Berechnung des Integrals 2.12 stattfinden. Stattdessen wird in der SPADE Methode der Ausdruck mittels modifizierter Gauss Hermite Quadratur approximiert. Dieser Ansatz wurde auch in Dekkers and Slob (2012) mit anderen Methoden (unter anderem Taylor Approximation und exakte Rücktransformation wie in S-Nusser) verglichen und als sehr gute Methode für solche Rücktransformationen angesehen. Gauss Hermite Quadratur ist eine Approximation für Integrale der Form  $\int e^{-x^2} f(x) dx$ . Dabei wird das Integral mit dieser Form durch eine Summe aus Gewichten und Knoten abgeschätzt ( $\int \dots dx = \sum w_k f(c_k)$ ). Die Gewichte sind dabei vorgegeben und die Knoten die Nullstellen eines Hermite Polynoms vom Grade  $p$ . Mathematisch erhält man folgendermaßen eine vereinfachte Darstellung des Integrals von Gl:2.12.

$$\begin{aligned} y_i &= \int_{-\infty}^{\infty} h(y_i^* + \epsilon) \varphi(\epsilon) d\epsilon = \int_{-\infty}^{\infty} h(y_i^* + \epsilon) \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left(-\frac{\epsilon^2}{2\sigma_w^2}\right) d\epsilon \\ &= \int_{-\infty}^{\infty} h(y_i^* + \sqrt{2\sigma_w^2} x) \frac{1}{\sqrt{\pi}} e^{-x^2} dx \\ &\approx \frac{1}{\sqrt{\pi}} \sum_{k=1}^p w_k h(\tilde{y}_i^* + c_k \sqrt{2}\sigma_w) \end{aligned} \quad (2.20)$$

Nach einsetzen der Dichte für  $\varphi(\epsilon)$  erhält man einen Ausdruck, der einfach durch Integration per Substitution in ein Integral der gewünschten Form zu bringen ist. Danach kann die Approximation durch die Gauss Quadratur erfolgen und man erhält die Summe in 2.20, welche ähnlich der 9 Punkte Approximation (Gl: 2.16) ist. Beachte, dass aber im Allgemeinen die Gewichte und Knoten nicht gleich jener der NCI Methode sind und somit unterschiedliche Resultate zustande kommen. Für den Grad  $p$  des Polynoms wird standardmäßig 9 in SPADE verwendet. Die standardmäßige Gauss Hermite Quadratur muss allerdings noch angepasst werden, da die Funktion  $h$  hier im Fall die Inverse der einparametrischen Box Cox Transformation ist und somit nicht für negative Werte definiert. Dies erfolgt in SPADE durch eine Anpassung, in der 3 separate Berechnungen für das Integral gemacht werden, wenn negative Inputwerte für die Inverse auftauchen. 1. wird in 2.20 für negative Inputwerte das Gewicht  $w_k$  gleich 0 gesetzt und die Summe damit berechnet. 2. wird die Anzahl der Knoten iterativ um 1 gesenkt bis das Problem nicht mehr auftaucht und die Summe berechenbar ist. 3. wird die Taylorapproximation verwendet. Das Maximum der 3 Werte wird dann für die modifizierte Gauss Quadratur verwendet.

In die Formel 2.20 muss nun noch die geschätzte within-person Varianz eingesetzt werden. Anstatt der  $y_i^*$  können nun die gesuchten Quantile der (im Schritt der Varianzeliminierung beschriebenen) Normalverteilung oder der Mittelwert eingesetzt werden, um Schätzungen für die Quantile oder den Mittelwert der üblichen Aufnahmeverteilung auf originaler Skalierung zu bekommen. Falls ein Anteil unter einem Grenzwert gesucht wird, wird ein Algorithmus zum Finden des passenden Quantils verwendet um so eine Schätzung zu ermöglichen.

## Schätzung von nicht täglich konsumierten Nahrungsbestandteilen

SPADE bietet hier die Möglichkeit die Aufnahmewahrscheinlichkeit als Funktion des Alters mittels Beta Binomial Modell zu schätzen. Zusatzinformationen, ob ein Individuum Nichtkonsument eines Nahrungsstoffes ist, können falls vorhanden verwendet werden. Ansonsten werden alle Personen als potentielle Konsumenten eingestuft, auch wenn an allen Beobachtungstagen nichts konsumiert wurde. Das Beta-Binomial Modell wurde verwendet, da es flexibler als die einfache logistische Regression ist, da es einen Streuungsparameter zulässt. Die übliche Aufnahmeverteilung wird hier wieder durch Multiplizieren der Aufnahmewahrscheinlichkeit mit der Menge erreicht. Es werden allerdings kleine Änderungen bei der Schätzung der Aufnahmemenge im Vergleich zum 1-Part Modell verwendet (siehe Dekkers et al. (2014)). Eine wichtige Änderung ist die Rücktransformation, die im 2-Part Modell nicht mittels Gauss Quadratur sondern Monte Carlo Simulation durchgeführt wird.

### weitere Möglichkeiten

SPADE bietet die Nutzung von Populationsgewichten und die direkte Berechnung der Anteilsschätzung für gewünschte Grenzwerte (auch mittels Probability Approach). Ebenso bietet SPADE Abschätzungen zur Unsicherheit der berechneten Perzentile und Anteile durch Bootstrap Konfidenzintervalle. Überprüfungsplots der Modellannahmen (normalverteilte Residuen usw.) werden ebenfalls ausgegeben. Weiters kann die Aufnahme der Nährstoffe auch unabhängig vom Alter modelliert werden.

### Implementierung der Methode

Verwendet kann SPADE mittels dem zur Verfügung gestellten R-Package "SPADE.RIVM" werden. Dieses, sowie ein passendes User Manual, kann über die Homepage beantragt werden.

## 2.3.4 Iowa State University Methode (ISU)

Die ursprüngliche Methode der Iowa State University (Nusser et al. (1996)) stellt eine aufwendigere Methode als die simple Best Power Methode dar. Sie stellt auch von den bisher gezeigten Methoden die komplexeste dar, die schon vor dem 1. Schritt der Datentransformation einige Datenanpassungen vornimmt. Eine Erweiterung der Methode in 2 Teile, für die Schätzung von nicht täglichen Nahrungsinhalten, kam erst später dazu (Nusser et al. (1997)). Zuerst wird wieder die 1 Part Methode von 1996 vorgestellt und anschließend kurz die Erweiterung.

**0. Anfangsanpassungen:** Die ISU Methode ermöglicht schon als ersten Schritt die Auswahl von mehreren Datenanpassungen zur Verringerung von unerwünschten Effekten wie Tag der Woche, des Monats, Interview Reihenfolge und Interview Art. Dazu werden die Daten zuerst mittels einer Powertransformation an Normalscores angepasst. Um für strikt positive Daten zu sorgen, werden sie zuerst mit einer kleinen Konstante addiert um folgende Werte  $W_{ij} = y_{ij} + 0.0001\bar{y}$  zu erhalten. Die Ermittlung der besten Power erfolgt mittels Kleinstquadrat Rastersuche. Dabei wird die gewichtete Summe  $\sum w_i \sum (U_{ij} - \beta_0 - \beta_1 W_{ij}^\gamma)^2$  für jede Power  $\gamma \in [1, 1.5^{-1}, 2^{-1}, \dots, 10^{-1}, 0]$  anhand der Parameter  $\beta_0$  und  $\beta_1$  minimiert ( $\gamma = 0$  bedeutet, dass der natürliche Logarithmus auf die Daten angewandt wird). Jene Power mit kleinster gewichteter Summe wird ausgewählt. Die  $w_i$  stellen die Populationsgewichte (falls gewichtete Daten verwendet werden) und die  $U_{ij}$  die Normalscores (Blom's Scores) dar. Diese berechnen sich nach  $\Phi^{-1}[(s_{ij} - 3/8)/(nm - 1/4)]$ , wobei  $s_{ij}$  den Rang der Beobachtung  $y_{ij}$  und  $\Phi$  die Standardnormalverteilung bezeichnet. Nach dieser Powertransformation werden die  $W_{ij}^\gamma$  für eine Regressionsanalyse weiter verwendet. Hier sollen nun die oben genannten Effekte als additive Klassifikationsvariablen verwendet und getestet werden. Bei signifikanten Ergebnissen können die Daten angepasst werden. Für eine Anpassung bezüglich dem Tageffekt (sprich z.B. liefert der erste Tag signifikant andere Werte als die nachfolgenden), sieht dies wie folgt aus. Die neuen adjustierten Werte setzen sich dann als  $Z_{ij} = \bar{Z}_1 W_{ij}^\gamma / \hat{Z}_{ij}$  zusammen.  $\bar{Z}_1$  stellt dabei den Mittelwert

der powertransformierten Variablen für den ersten Beobachtungstag dar.  $\hat{Z}_{ij}$  stellt hingegen die Modellschätzung der Regression dar. Die Werte werden somit an den Mittelwert des ersten Beobachtungstag angepasst, da dies als genaueste Messung angenommen werden kann. Die Anpassung kann allerdings auch an den Gesamtmittelwert für alle Beobachtungstage erfolgen. Ebenso können noch die Varianzen der unterschiedlichen Tage an den ersten Tag angepasst werden. Sollte jedoch kein Grund für die Annahme, dass der erste Tag am genauesten aufgezeichnet worden ist, vorliegen, so können diese Anpassungen ausgelassen werden. Für die anderen möglichen Variablen, die oben erwähnt wurden, plus weitere die einem wichtig vorkommen, können ebenfalls ähnliche Anpassungen wie gerade beschrieben vorgenommen werden.

Bevor nun die eigentlichen 4 Hauptschritte ausgeführt werden können, besteht noch die Möglichkeit für gewichtete Daten einen "neuen" Datensatz zu generieren, welcher in den weiteren Schritten wie ein nicht gewichteter Datensatz verwendet werden kann. Dies erfolgt über die empirische kumulative Verteilungsfunktion  $\hat{F}_Z(x) = \sum w_i \sum I_{Z_{ij}}(x)$ , wobei  $I_Z(x)$  die Indikatorfunktion darstellt ( $I_Z(x) = 1$  für  $Z \leq x$  und 0 sonst). Diese stufenweise Funktion wird nun noch geglättet bzw. stetig, durch verbinden der Mittelpunkte jeder Stufe, gemacht. Diese Funktion  $\tilde{F}_Z$  wird jetzt verwendet, um eine gleichgewichtige Stichprobe (engl. equal weight sample) zu erhalten, die sich aber approximativ verhält wie die Datenpunkte mit ungleichen Gewichten. Diese nun erhaltenen Daten werden für den weiteren Verlauf der Methode wieder mit  $y_{ij}$  bezeichnet, auch wenn sie eigentlich schon transformiert und adjustiert wurden ( $y_{ij} = \tilde{F}_Z^{-1}[(s_{ij} - 0.5)/(nm)]$ ).

**1. Datentransformation:** Mit den angepassten Daten, die sich nun verwenden lassen wie ein nicht gewichteter Datensatz, kann nun die Transformation zur Normalität erfolgen. Im Gegensatz zu den anderen Methoden wird bei der ISU Methode ein semiparametrischer Ansatz gewählt, um die Daten an die Blom Normalscores anzupassen. Dies erfolgt in 2 Schritten. Zuerst wird eine Powertransformation wie vorher mittels Kleinstquadrate Minimierung (diesmal ungewichtet) durchgeführt. Diesmal wird die Power wieder mit  $\alpha$  bezeichnet, um den Zusammenhang zu den anderen Methoden besser darzustellen. Im zweiten Schritt wird ein zusammengesetztes kubisches Polynom (engl. grafted cubic polynomial) an die Datenpaare  $(U_{ij}, y_{ij}^\alpha)$  angepasst. Minimiert wird dabei in  $y$  Richtung. Ein zusammengesetztes Polynom bedeutet, dass die Anpassungsfunktion aus mehreren Polynomen für mehrere Intervalle besteht, welche an den Anfangs- und Endpunkten zusammengefügt werden. Es sind dazu Verbindungspunkte zu wählen, wobei der Abstand der Punkte gleich lang sein soll. Die Anzahl der Verbindungspunkte (von 3 bis 12) wird so gewählt, dass sie minimal ist, aber gleichzeitig auch einen vorgegeben Wert der Anderson-Darling Statistik erreicht. Zusätzlich sollen die Polynome ein paar Eigenschaften erfüllen. Sie müssen einerseits für das erste und letzte Intervall linear sein und außerdem kubisch für die Intervalle dazwischen, mit stetiger erster und zweiter Ableitung an den Verbindungspunkten. Für nähere Informationen siehe auch Fuller (2009) Kapitel 9.2 (grafted polynomials).

Die transformierten Daten  $y_{ij}^*$  werden nun, wieder wie in den anderen Methoden, nach Gl:2.4 betrachtet. Der Unterschied der ISU Methode ist, dass sie zusätzlich ungleiche within-person Varianzen für alle Individuen zulassen kann. Es gilt  $y_i^* \sim N(0, \sigma_{b^*}^2)$  unabhängig von  $\epsilon_{ij}^* \sim N(0, \sigma_{w_i^*}^2)$  mit Erwartungswert  $E(\sigma_{w_i^*}^2) = \mu_A$  und Varianz  $Var(\sigma_{w_i^*}^2) = \sigma_A^2$ .

**2. Varianzschätzung:** Geschätzt werden die Parameter ähnlich wie bei der Best Power Methode allerdings mit leichten Änderungen, da ja nun ungleiche Tag-zu-Tag Variabilität möglich ist. Die Schätzung der Varianz der transformierten Mittelwerte ist wieder gleich wie in Gl:2.8. Die Schätzung in Gl:2.9 stellt nun die Schätzung für den Erwartungswert  $\mu_A$  von der within-person Varianz dar. Diese beiden Varianzschätzer können nun wieder verwendet werden um mit Gleichung 2.1 die between-person Varianz zu schätzen. Um nun noch die Varianz  $\sigma_A^2$  schätzen zu können, verwendete Nusser et al. (1996) folgende Schätzung

$$\hat{\sigma}_{\mu_A}^2 = \frac{1 + \frac{2}{m-1}}{n} \sum_{i=1}^n A_i^2 - \bar{A}^2, \quad (2.21)$$

wobei  $A_i$  die unverzerzte Varianzschätzung (für die Tagesbeobachtungen) auf Individualebene darstellt und  $\bar{A}$  der Mittelwert dieser Varianzen ist.

**3. Varianzeliminierung:** In diesem Schritt wird ein Set von 400 Zwischenwerten  $\tilde{y}_i^* = \sqrt{\hat{\sigma}_{b^*}^2} U_i$  ( $U_i$  stellt wieder den Blom Normalscore dar, wobei statt  $s_{ij}$  der Rang des Zwischenwertes eingesetzt wird) erzeugt, welches sich ähnlich einer Normalverteilung mit Erwartungswert 0 und between-person Varianz verhält.

**4. Rücktransformation:** Diese 400 Zwischenwerte werden nun mit einer 9 Punkt Approximation (wie in 2.16) in die originale Skala transformiert. Die Inverse von der ursprünglichen zweistufigen Transformation kann mathematisch hergeleitet und damit auch tatsächlich verwendet werden. Die Gewichte  $w_k$  und Knoten  $c_k$  werden hier so gewählt, dass ihre 9 Punkt Verteilung die ersten 5 geschätzten Momente der bedingten Verteilung  $(\tilde{y}^* + \epsilon^*)|\tilde{y}^*$  approximieren. Zusätzlich haben die Gewichte wieder die Restriktion, dass ihre Summe gleich 1 ist. Durch diese Approximation erhalten wir nun weitere 400 Zwischenwerte  $\tilde{y}_i$  auf der ursprünglichen Skalierung. Nun wird wieder ein zusammengesetztes Polynom auf die Paare  $(\tilde{y}_i^*, \tilde{y}_i)$  angewandt um die Rücktransformation für allgemeine Werte zu schätzen. Mit dieser Rücktransformation können dann gesuchte Quantile oder der Mittelwert der üblichen Verteilung geschätzt werden.

### Schätzung von nicht täglich konsumierten Nahrungsbestandteilen

Die Erweiterung der Methode beruht auf einem 2-Teile Modell, wobei der 2. Teil dem oben beschriebenen Part gleich ist und der 1. Teil wiederum der Schätzung der Aufnahmehäufigkeit dient. Dies erfolgt allerdings nicht mittels eines klassischen Modells wie der logistischen Regression oder dem Beta-Binomial Modell, sondern erfolgt auch hier über einen eigenen Ansatz. Dabei wird eine eigene diskrete Verteilung definiert, deren Wahrscheinlichkeitsmasse durch die beobachteten Häufigkeiten an Tageskonsumationen modelliert wird. Dies erfolgt über einen modifizierten Chi-Quadrat Schätzer. Für nähere Informationen siehe Nusser et al. (1997).

### weitere Möglichkeiten

Gewichtete Datensätze können klarerweise wegen dem oben beschriebenen Ansatz verwendet werden. Ebenso können die Varianzen innerhalb der Personen als gleich angenommen werden, falls Grund zur Annahme besteht. Ebenso bietet die Methode die Möglichkeit mit dem Output mögliche Abbildungen, wo die Annahmen auf ihre Plausibilität überprüft werden können, zu erstellen. Die ISU Methode bietet weiters eine Dichteschätzung für die übliche Aufnahme, die mittels Dichtetransformationformel ermittelt wird.

### Implementierung der Methode

Die ISU Methode ist in einem eigenständigen Programm, genannt PC-Side, implementiert und kann über die Homepage beantragt werden.

Im nächsten Abschnitt werden die 4 Hauptmethoden direkt gegenübergestellt und verglichen.

## 2.4 Inhaltlicher Vergleich der Hauptmethoden

Die 4 vorgestellten Methoden haben alle im Grunde genommen sehr ähnliche Ansätze, um die übliche Aufnahmeverteilung einer Gruppe zu schätzen. Sie gehen von der Überlegung aus, dass sich die Tagesbeobachtungen  $y_{ij}$  in drei Teile zerlegen lassen. Den Gesamtmittelwert  $\mu$  der Gruppe plus der individuellen Abweichung  $\mu_i$  der Person plus der Abweichung  $\epsilon_{ij}$ , die von Tag-zu-Tag in der Aufnahme vorkommt. Mit dieser Modellierung kann nun versucht werden, die Varianz, die in den Tag-zu-Tag Abweichungen liegt, zu schätzen, um sie entfernen zu können. Da dies besser funktioniert, wenn von einer (zumindest approximativen) Normalverteilung der Daten ausgegangen werden kann, nehmen alle 4 Methoden eine Transformation vor. Während die Methoden NCI, SPADE und ISU dies direkt auf die Tagesbeobachtungen anwenden, nimmt MSM zuerst eine Schätzung des Gesamtmittelwerts



der Gruppe mittels Regression vor. Dies kann entweder anhand von Kovariaten durchgeführt werden oder aber nur mit einem Intercept, was nichts anderes als eine gewöhnliche Mittelwertschätzung wäre. Die Residuen des Modells beinhalten noch die 2 restlichen Teile  $\mu_i$  und  $\epsilon_{ij}$  und werden somit nun auch transformiert um die Schätzung der between- und within-person Varianzen vornehmen zu können.

**1. Transformation:** Den komplexesten Ansatz der Transformation hat die ISU Methode mit einem semiparametrischen 2-Schritt Verfahren. Hier wird zuerst eine einfache Powertransformation gefolgt von einer zusammengesetzten kubischen Polynomtransformation durchgeführt. Die anderen 3 Methoden bedienen sich der Box-Cox Transformation, wobei hier noch der Unterschied vorliegt, dass es sich einerseits um eine 1- oder 2-parametrische Transformation handelt, sowie um verschiedene Möglichkeiten der Parameterschätzung. Während NCI hier den Vorteil nutzt den Parameter erst im Modellschätzverfahren zu bestimmen und somit mögliche Kovariaten berücksichtigt werden können, so verwenden MSM und SPADE die Box Cox Variante schon vor der Varianzschätzung. Ebenso schränkt MSM die voreingestellte Parameterauswahl doch deutlich ein im Gegensatz zu den anderen Methoden.

**2. Varianzschätzung** Geschätzt werden die Varianzen entweder mit einfachen Schätzern (ISU, MSM) von einem klassischen Messfehlermodell oder mittels Maximum Likelihood Methode (NCI, SPADE) im Rahmen eines komplexeren Modells. SPADE verwendet hier die Modellierung des Gesamtmittelwerts durch ein fraktionales Polynom bezüglich des Alters und anschließender Restriktierter Maximum Likelihood Schätzung der Varianzen im Rahmen eines linearen gemischten Effekte Modells. NCI verwendet die "normale" ML Schätzung bezüglich eines nichtlinearen gemischte Effekte Modell.

**3. Varianzentfernung:** Für die Entfernung der ungewollten Tag-zu-Tag Variabilität bedient sich die MSM Methode der Shrinkage Schätzung, während sich NCI auf die Simulation von Zwischenwerten mit der gewünschten Varianz stützt. SPADE hingegen verwendet direkt die Quantile der auf der transformierten Skalierung geschätzten Normalverteilung. Die ISU Methode benutzt beide Elemente. Einerseits wird ein Set aus Zwischenwerten mit der richtigen Varianz erstellt und andererseits werden die gesuchten Quantile der Normalverteilung direkt verwendet für die Rücktransformation. Während die Shrinkage Schätzung den Vorteil hat, die Annahme der Normalverteilung weniger zu strapazieren, so hängen die Methoden ISU, NCI und SPADE doch deutlich stärker von der Annahme ab.

**4. Rücktransformation:** Für die anschließende Rücktransformation auf die ursprüngliche Skala verwendet SPADE den Gauss Quadratur Ansatz und ISU sowie NCI eine 9-Punkt Approximation (bzw. ISU eine weitere Approximation mittels zusammengesetzten Polynomen). Nur MSM ermöglicht hier eine exakte Rücktransformationsformel, wobei dieser Bonus auf Kosten der eingeschränkten Parameterauswahl beruht.

**Schätzung der üblichen Aufnahmeverteilung bzw. Anteile:** Während die NCI und MSM Methode die empirische Verteilung der rücktransformierten Werte für die Schätzung der Quantile der üblichen Aufnahmeverteilung verwendet, so werden bei SPADE und ISU die Quantile direkt durch die Rücktransformation bestimmt. Falls Anteile unter gewissen Schranken gesucht werden, verwendet NCI und MSM die empirischen Anteile der rücktransformierten Werte und SPADE einen Algorithmus um das passende Quantil zu finden. Bei der ISU Methode wird nicht weiter erklärt wie dieser Vorgang durchgeführt wird.

**Kovariaten:** Während die ISU Methode zwar keine Kovariaten im Modell erlaubt, so kann sie doch in einem Schritt vor der Transformation schon versuchen, unerwünschte Störfaktoren zu beseitigen. Die NCI Methode lässt die Modellierung mittels (beliebige) Kovariaten auf transformierter Skala zu. Somit kann die Varianz innerhalb der Personen in einen aufgeklärten und einen nicht aufgeklärten Teil zerlegt werden, vorausgesetzt die Kovariaten erklären einen Teil der Varianz. SPADE versucht einen Teil der between-person Varianz zumindest mit dem Alter der Personen zu modellieren. MSM lässt ebenfalls die Modellierung

mittels beliebigen Kovariaten zu. Allerdings muss hier noch angemerkt werden, dass die Autoren der MSM Methode auf der Homepage bekanntgegeben haben, dass bei Verwendung von Kovariaten Vorsicht geboten ist. Die Autoren meinen weiter, dass bei Simulationstudien mit Kovariaten nicht die erwarteten Ergebnisse zustande kamen. Das Problem taucht auf, wenn die Residuen des Regressionsmodells nicht normalverteilt seien. Da dies so gut wie immer der Fall ist bei Aufnahmezeiten (meist rechtsschief), sollte MSM mit Kovariaten zum jetzigen Zeitpunkt nur mit äußerster Vorsicht verwendet werden.

**Analyse von Subgruppen** Für manche Fragestellungen ist es interessant, die übliche Aufnahmeverteilungen für Subgruppen (z.B. nach Geschlecht oder Altersgruppen) zu bestimmen. Während es immer möglich ist, die Methoden separat für jede Subgruppe anzuwenden, so kann es manchmal von Vorteil sein, Informationen der gesamten Stichprobe auch für bestimmte Subgruppen zu verwenden. Daher ist es auch interessant, ob die 4 Methoden diese Analyse zulassen. Die ISU Methode ermöglicht dies nicht, da keine Kovariaten verwendet werden können. Die MSM Methode lässt zwar Kovariaten zu, aber keine Subgruppenanalyse, da sie alle Personenwerte der Stichprobe mit der Shrinkage Methode und Rücktransformation auf die originale Skalierung bringt und nicht nur für die gewünschte Subgruppe. SPADE lässt zumindest die Modellierung anhand des Alters zu und kann dadurch für gewisse Altersgruppen (oder jedes beliebige Alter) übliche Aufnahmeverteilungen schätzen. Wenn nach Geschlecht getrennt analysiert werden soll, so muss dies bei SPADE auch separat durchgeführt werden. Die NCI Methode stellt von diesem Standpunkt aus gesehen das beste Paket zur Verfügung. Sie erlaubt auch beliebige Kovariaten im Modell und auch die Analyse für die gewünschten Subgruppen. Es kann sogar für die Subgruppen (solange es nicht zu viele sind) variable within-person Varianz zugelassen werden.

**Modellannahmen:** Wie zu Beginn des Kapitels schon erwähnt, werden 4 wichtige Annahmen getroffen, welche in der Praxis zu Problemen führen können. Dass sich die Daten in normalverteilte Zufallsvariablen transformieren lassen, ist immer noch eine starke Annahme, die nicht immer mit Box-Cox Transformation gelöst werden kann. Grundsätzlich hat hier die ISU Methode einen Vorteil durch ihre komplexere 2-Schritt-Transformation, doch auch diese wird bei multimodalen Verteilungen an ihre Grenzen stoßen. Hier ist immer ein strenger Blick auf die originalen Daten von hoher Priorität, wenn zuverlässige Schätzungen das Ziel sind. Sollten viele große "Ausreißer" vorhanden sein, so kann auch überlegt werden, ob die SPADE Modellierung mittels dem Supplemente Part sinnvoll wäre.

Die Annahme der Unabhängigkeit der between-person und within-person Effekte spielt auch für diese Methoden eine wichtige Rolle bei der Varianzschätzung. Sie wird für alle 4 benötigt und auch nicht vereinfacht. Nusser et al. (1996) hat damals bei der Verwendung der ISU Methode an Daten einen Test diesbezüglich für Kalzium, Energie, Eisen, Protein, Vitamin A und C durchgeführt und sich für eine Modellierung mit dieser Annahme entschieden, da keine groben Zusammenhänge sichtbar waren. Für die Praxis wäre trotzdem immer eine Kontrolle dieser Annahme sinnvoll, was durch einen Testoutput in der ISU Methode möglich ist. Vergessen darf allerdings nicht werden, dass diese Annahme nur auf die transformierten Daten zutrifft und somit weniger problematisch sein könnte.

Eine weitere wichtige Annahme muss bezüglich der within-person Varianz getroffen werden. Während ISU hier als einzige Methode ungleiche Varianzen zulässt, so sind doch keine Unterschiede in der between und within-person Varianz Schätzung zu den anderen Methoden. Die gelockerte Annahme macht sich erst bei der Rücktransformation bemerkbar (durch die Anpassung an die ersten 5 Momente der Fehlertermverteilung). Wie groß dieser Effekt ist, hängt stark von der Ungleichheit der Varianzen ab. Während eines Vergleichs der Best Power und ISU Methode konnten zwar (schwache) Unterschiede in den Enden der geschätzten üblichen Aufnahmeverteilungen festgestellt werden, jedoch könnte dies auf die einfachere Transformationsart der BP Methode zurückführbar sein (Nusser et al. (1996)). Weitere Vergleiche bezüglich der ungleichen Varianzen wären notwendig. Eine leichte Auflockerung der Annahme kann auch die NCI Methode vornehmen, wobei dann die Varianz zumindest anhand der Kovariatlevel variieren darf.

Die wohl ausschlaggebendste Annahme ist die Erwartungstreue der Daten auf der ursprünglichen Skalierung, wobei ohne sie keine der 4 Methoden auskommt. Ob diese Annahme plausibel ist, wird in Kapitel 4 noch weiter erläutert. Sollte sie verletzt sein, so wirkt sich

dies natürlich auf die komplette, geschätzte übliche Aufnahmeverteilung, sowie die Anteilsschätzung aus.

Zu diesen 4 Annahmen kommen noch kleinere, weniger ausschlaggebendere Annahmen, die für die jeweiligen Schritte und Modellierungen der Methoden gebraucht werden. Diese sind z.B. die Linearität für das Regressionsmodell in der MSM Methode oder die Annahme, dass die Fehlerterme  $\epsilon_{ij}$  für  $j = 1, \dots, n$  nicht autokorreliert sind, sowie die Annahme, dass die mittlere Aufnahme einer Person sich durch bestimmte Kovariaten modellieren lässt usw. Die meisten Annahmen lassen sich gut berücksichtigen bzw. überprüfen. Die Autokorrelation der Fehlerterme z.B. kann durch 24-Stunden Protokolle, die nicht an aufeinanderfolgenden Tagen stattfinden, gut vermieden werden.

**Modellüberprüfungen:** Hier kann vor allem die SPADE Methode punkten, da sie automatisch Überprüfungsabbildungen als Outputs liefert. Diese werden in ein PDF gespeichert und schließen Residuenplots, sowie Histogramme und Vergleiche mit der Normalitätsannahme ein. Die anderen Methoden liefern standardmäßig keine Überprüfungsplots, jedoch können mit dem Output der ISU oder NCI Methode selbst Überprüfungen durchgeführt werden. Die MSM Methode hingegen liefert nur ein Log-File sowie die Schätzungen der individuellen Aufnahmen, aus denen sich nicht mehr die Überprüfung der Residuen auf Normalität ermöglichen lässt. Logdateien für die verwendeten statistischen Modelle innerhalb der Methoden liefern alle 4 Methoden. Die ISU Methode kann hier noch mit zwei weiteren Testüberprüfungen der Hauptannahmen punkten. Dabei wird für die Gleichheit der within-person Varianzen, sowie für die Unabhängigkeit der within-person Varianz zu der mittleren Aufnahme ein Testoutput angegeben.

**Abschätzungen der Unsicherheiten:** SPADE bietet hier die Möglichkeit Konfidenzintervalle für die Schätzungen mittels Bootstrap zu ermitteln. ISU bietet die Möglichkeit der Verwendung von Jackknife oder Balanced Repeated Replication zur Unsicherheitsabschätzung. NCI kann selbstständig (mit SAS Programmierkenntnissen) erweitert werden um Balanced Repeated Replication durchzuführen. Nur die MSM Methode bietet keine Möglichkeit die Unsicherheit abzuschätzen.

**Erweiterungen für nicht täglich konsumierte Nahrungsinhalte:** Während alle 4 Methoden eine Erweiterung zu einem 2-Part Modell haben, so bestehen doch deutliche Unterschiede in der Modellierung der Aufnahmewahrscheinlichkeiten. Dies zeigt sich auch bei Anwendung auf "echte" Daten, siehe Souverein et al. (2011). Weitere Vergleiche der vier Methoden bezüglich dieses Aspekts sind notwendig um bessere Erkenntnisse zu bekommen, welche Methode besser abschneidet. Die Möglichkeit der Verwendung von Verzehrshäufigkeiten als Kovariaten in den Methoden MSM und NCI sind hier auf jeden Fall von Vorteil.

**Softwareimplementierung** Am einfachsten zu bedienen ist hier die MSM Methode über die Homepage. Es wird ziemlich genau erklärt, welche Daten hochgeladen werden können und wie dann die weiteren Schritte sind. Die ISU Methode bietet mit dem eigenen Programm PC-Side ebenfalls eine relativ leichte "Klick" Anwendung. Das Programm kann per Mausclicks relativ einfach bedient werden. Für die Methoden NCI und SPADE sind zumindest Basic Programmsyntaxkenntnisse in SAS bzw. R notwendig, um die Daten einlesen und auswerten zu können. Sie stellen daher von diesem Standpunkt aus gesehen die "schwieriger" ausführbaren Methoden dar.

### 2.4.1 Schlussfolgerungen

Die nachfolgende Tabelle (2.4.1) liefert nochmal eine Zusammenfassung der wichtigsten inhaltlichen Unterschiede der 4 besprochenen Methoden. Anhand der hier beschriebenen inhaltlichen Vergleiche lässt sich sagen, dass keine der Methoden gleichmäßig besser ist als die anderen. Jede hat ihre Vor- und Nachteile und kann dadurch in unterschiedlichen Situationen zu schlechteren oder besseren Ergebnissen führen. Eine Überprüfung der Modellannahmen sollte allerdings immer vorgenommen werden um jegliche Aussagen plausibel zu halten. Bei Daten die sich nur schlecht durch die Box-Cox Transformation auf approximative Normalität bringen lassen, sollte möglicherweise die ISU Methode bevorzugt werden. Ebenso bei

Verletzung der Annahme gleicher Varianzen innerhalb der Personen. Ansonsten kann die Methode verwendet werden, dessen Programm(paket) bevorzugt wird. Ein Nachteil für die Auswertungen des Ernährungsberichtes der Österreicher stellt hier die MSM Methode dar, da sie keine Verwendung von Populationsgewichten ermöglicht. Da die Datenerhebung des Ernährungsberichtes allerdings nicht auf einer einfachen Zufallsstichprobe basiert, können hier weniger repräsentative Ergebnisse die Folge sein. Daher müssten die Daten ähnlich wie bei der ISU Methode vorbearbeitet werden, um die MSM Methode auf die Ernährungsberichtsdaten anwenden zu können.

	NCI	MSM	SPADE	ISU
Vorauswahl	keine	Regression mittels Kovariaten	keine	Entfernung von Störfaktoren
Datentransform.	1-Parameter Box-Cox	2-Parameter Box-Cox auf Residuen	1-Parameter Box-Cox	Powertransf., zusammenges. Polynom
Parameterwahl	ML gemeinsam mit anderen Parametern	Rastersuche durch Shapiro-Wilk Statistik	ML separat von den anderen Parametern	Kleinstquadratmin., Rastersuche durch Anderson-Darling Stat.
Modellierung	Nichtlineares gemischtes Modell inkl. Kovariaten	einf. Messfehlermodell auf Residuen	gemischtes Modell incl. fraktionales Polynom	einfaches Messfehlermodell
Varianzschätzung	ML	klassische Schätzer	REML	klassische Schätzer
Varianzreduktion	simulierte Werte	Shrinkage Faktor	Quantile der NV mit gewünschter between Varianz	simulierte Werte und Quantile
Rücktransformation	9 Punktapproximation	exakte Formel	Gauss Quadratur	9 Punktappr., zusammenges. Polynom
within person Varianzen	Anpassung an Kovariatlevel möglich	gleich für alle Personen	gleich für alle Personen	unterschiedliche Varianzen möglich
Überprüfung der Modellannahmen	möglich mit dem Output	nicht möglich	automatisch generiert	möglich mit dem Output
Pop.gewichte	ja	nein	ja	ja
Kovariaten	ja (beliebige)	ja (beliebige)	nur Alter	nein
Abschätzung der Unsicherheiten	Balanced Repeated Replication	nein	Bootstrap	Balanced Repeated Replication oder Jackknife
Erweiterung zu 2 Part Modell	Logistische Regression mit Zufallseffekten	Logistische Regression	Beta-Binomialmodell	diskrete Verteilung und modifizierte Chi-Quadratschätzer

Tabelle 2.3: Inhaltliche Gegenüberstellung der Methoden SPADE, NCI, ISU und MSM

Wir haben jetzt Methoden kennen gelernt, mit denen es möglich ist eine übliche Aufnahmeverteilung zu schätzen. Im Eingangsbeispiel wurden Anteilsschätzungen für die Vitamin B12 Aufnahme erwähnt. Wie kommt man nun zu diesen Schätzungen? Eine einfache Möglichkeit wäre einfach den Anteil jener Personen, die unter einem gewissen Referenzwert liegen, zu berechnen. Doch kann für diesen Anteil auch inhaltlich geschlussfolgert werden, dass so viele Leute eine unzureichende Versorgung mit dem Nährstoff haben? Oder stellt dies lediglich einen "gewöhnlichen" Anteil dar, welcher nicht inhaltlich für Fragen bezüglich dem Versorgungsstatus einer Gruppe interpretiert werden darf. Diese Fragen werden in den nächsten Abschnitten geklärt.

## 2.5 Schätzung des Anteils: Probability Approach

Wie schätzt man nun den Anteil der, bezüglich eines Nährstoffes, unterversorgten Bevölkerungsgruppe anhand von Aufnahmedaten ab? Das National Research Council hat sich mit dieser Fragestellung schon 1986 beschäftigt und einen möglichen Lösungsweg präsentiert. Genannt wurde diese Methode "Probability Approach" (National Research Council, 1986, Kapitel 5).

Mathematisch gesehen sind wir daran interessiert die Wahrscheinlichkeit zu schätzen, dass die Aufnahme  $Y$  eines Kollektivs unter ihrem Bedarf  $R$  liegt. Als Formel sieht dies folgendermaßen aus  $P(\text{unzureichende Nährstoffversorgung}) = P(Y < R)$ . Notwendig wäre dafür die gemeinsame Verteilung  $F_{Y,R}(y, r)$  bzw. die Dichte  $f_{Y,R}(y, r)$ . Dann würde gelten:

$$P(Y < R) = \int_0^\infty \int_0^r f_{Y,R}(s, t) ds dt. \quad (2.22)$$

Die gemeinsame Dichte zu bestimmen ist allerdings sehr aufwendig bzw. kaum möglich und daher unpraktikabel für die Praxis. Daher kann die Wahrscheinlichkeit nicht ohne weitere Annahmen bestimmt werden. Wie wir im Anfangskapitel gehört haben, sind für verschiedene Personen eigentlich verschiedene Referenzwerte notwendig. Daher wäre es wünschenswert eine komplette Bedarfsverteilung  $F_R$  zu kennen. Diese würde die Streuung des Bedarfs beinhalten und würde damit weitaus passender sein um die Fragestellung, ob eine Person unterversorgt ist oder nicht zu beantworten. Es könnte eine Wahrscheinlichkeit für den Unterversorgungsgrad angegeben werden. Der Probability Approach geht davon aus, dass eine Bedarfsverteilung  $F_R$  für den zu untersuchenden Nährstoff bekannt ist. Sollte dies tatsächlich der Fall sein, so kann unter Unabhängigkeit der Variablen  $Y$  und  $R$  der Ausdruck 2.22 leicht bestimmt werden. Die Wahrscheinlichkeit kann auch geschrieben werden als:

$$P(Y < R) = \int_0^\infty P(R > y | Y = y) f_Y(y) dy = \int_0^\infty (1 - F_R(y)) f_Y(y) dy,$$

wobei im zweiten Schritt die Unabhängigkeit verwendet wird. Dies kann auch als Erwartungswert  $E_Y[(1 - F_R(y))]$  aufgefasst werden. Diesen Erwartungswert versucht man mit dem Probability Approach zu schätzen, was auch möglich ist, solange eine Bedarfs und Aufnahmeverteilung bekannt ist. Der Erwartungswert kann dann numerisch oder analytisch bestimmt werden.

Der Probability Approach bietet eine genaue unverzerrte Schätzung des Anteils solange die Verteilungen des Bedarfs und der Aufnahme gut genug bekannt sind und die Unabhängigkeit gilt. Die Unabhängigkeit ist in den meisten Fällen kein Problem, da die Aufnahme eines Nährstoffes kaum vom Bedarf abhängt. Einzig und allein die Kalorienaufnahme und ihr Bedarf dürfen nicht als unabhängig angenommen werden, da sie miteinander korrelieren. Durch Verwenden des Probability Approach kann es somit zu Fehlschätzungen kommen.

Die größte Einschränkung die diese Methode mit sich bringt ist die fehlende Bedarfsverteilung für viele Nährstoffe. Oft ist nur der mediane oder mittlere Bedarf einer Gruppe bekannt. Selbst wenn noch die Varianz bekannt ist, so kann nicht immer die Form der Verteilung gut genug abgeschätzt werden. Daher werden auch andere Ansätze gebraucht, welche nicht zu stark von der Form der Bedarfsverteilung abhängen.

## 2.6 Schätzung des Anteils: EAR cut-point Methode

Eine Vereinfachung des Probability Approachs stellt die EAR cut-point Methode (Beaton, 1994) dar. Diese benötigt nicht mehr die ganze Verteilung des Bedarfs, sondern nur die mittlere Bedarfsschätzung für ein Kollektiv. Allerdings müssen dafür weitere Annahmen (zusätzlich zur Unabhängigkeit) getroffen werden:

- Die Bedarfsverteilung ist angemessen symmetrisch und
- die Varianz der Aufnahme ist größer als die des Bedarfs des Nährstoffes.

Um die Idee hinter der Methode gut verstehen zu können, ist eine Abbildung sinnvoll (Abb.2.3). Zu sehen ist die gemeinsame Verteilung von Bedarf und Aufnahme jeder Person. Betrachten wir zuerst das linke Bild. Zu sehen sind für jede Person (ist gleich ein Punkt) der Aufnahme- und Bedarfswert. Liegt eine Person unter der Diagonale  $45^\circ$  (Bedarfsgrenze), so ist ihre Aufnahme niedriger als der Bedarf und sie ist somit unterversorgt. Diesen Anteil gilt es zu schätzen.

Da der Bedarf für jede einzelne Person aber nur in den seltensten Fällen bekannt ist, kann so nicht herausgefunden werden wer unterversorgt ist und wer nicht. Bekannt (bzw. geschätzt) sind meist nur die üblichen Aufnahmen der Nährstoffe und deren mittlerer Bedarf. Zusätzlich zur Bedarfsgrenze ist auch der EAR für den Bedarf (y-Achse) eingezeichnet, sowie der EAR als cut-point für die Aufnahme (x-Achse). Personen die also links vom cut-point liegen werden als unterversorgt eingestuft. Dadurch erhalten wir vier Kategorien. Einerseits klassifizieren wir Personen richtig als unterversorgt, wenn ihre Aufnahme höher ist als ihr Bedarf und niedriger als der cut-point (richtig positiv). Ebenso werden Personen, deren Aufnahme höher als ihr Bedarf und höher als der cut-point liegt richtigerweise als nicht unterversorgt eingestuft (richtig negativ). Jene Personen deren Aufnahme niedriger als ihr Bedarf ist und die einerseits kleiner als der cut-point (falsch positiv) oder größer als der cut-point ist (falsch negativ) werden falsch klassifiziert.

Unter den oben genannten Voraussetzungen ist nun jener Anteil links vom cut-point (klassifizierte Unterversorgte) sehr ähnlich dem Anteil über der Diagonale (wahre Unterversorgte). Der Grund dafür ist, dass der Anteil im Dreieck "falsch negativ" ungefähr gleich dem Anteil im Dreieck "falsch positiv" ist (linkes Bild). Es werden also Personen die zwar ihren Bedarf erfüllen als unterversorgt mitgezählt, aber ungefähr derselbe Anteil an Leuten die nicht gut versorgt sind als versorgt bestimmt. Somit gleicht sich dies aus und jener Prozentsatz unter dem EAR cut-point ist dem zu schätzenden sehr ähnlich. Mathematisch notiert gilt also:

$$P(Y < R) = \int_0^\infty \int_0^r f_Y(s) f_R(t) ds dt = E_R(F_Y(R)) \approx F_Y(EAR).$$

Sollte von der Bedarfsverteilung nur der Median bekannt sein und nicht die ganze Verteilung, so ist dieser Ansatz dem Probability Approach vorzuziehen. Die 3 wesentlichen Annahmen der EAR cut-point Methode sollten allerdings nicht außer Acht gelassen werden, wenn die zuverlässige Schätzung des Anteils das Ziel ist. In Carriquiry (1999) und National Research Council (2000) werden die Szenarien, wenn die Annahmen verletzt sind behandelt und erklärt. Da diese Arbeit dieses Thema nicht als Fokus hat, soll hier nur darauf verwiesen und kurz die Eckpunkte zusammengefasst werden.

- Unabhängigkeit nicht erfüllt:  
Sollten die Aufnahme und der Bedarf korrelieren (z.B. Nahrungsenergie), so können der Anteil im Dreieck 1 und 2 nicht mehr ähnlich sein und es kommt zur verzerrten Schätzung. Je nach Korrelationsstärke variiert auch die Stärke der Verzerrung. Ob unter- oder überschätzt wird, hängt von der Lage des EAR und der mittleren Aufnahme, sowie von der Korrelationsart (positiv oder negativ) ab.
- Symmetrie der Bedarfsverteilung nicht erfüllt:  
Sollte der Bedarf als nicht symmetrisch angesehen werden können (z.B. Eisen bei Frauen), so ist Vorsicht geboten mit der cut-point Methode. Auch hier kommt es wieder auf die Stärke und Art der Schiefe (rechts- oder linksschief) an, ob unter- oder überschätzt wird.

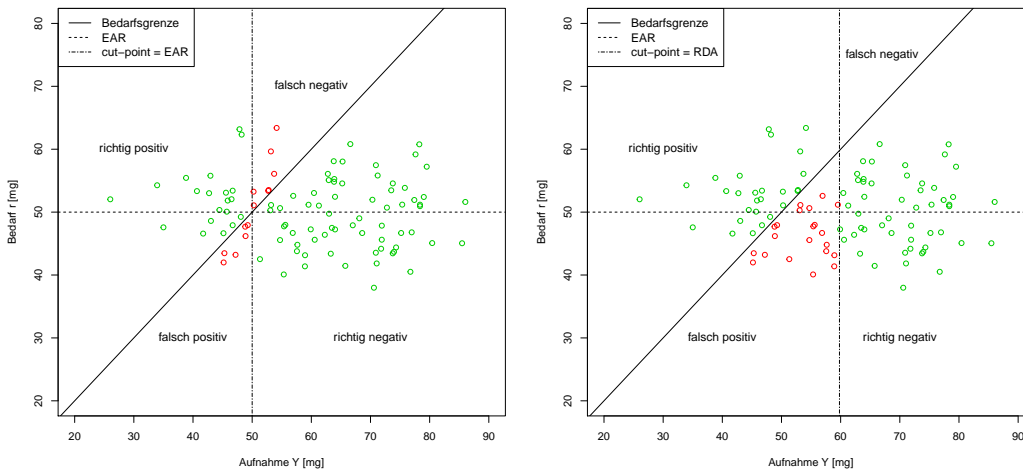


Abbildung 2.3: Gemeinsame Verteilung von Bedarf und Aufnahme. Links mit dem EAR und rechts mit dem  $RDA = EAR + 1.96 * SD$  als cut-point

- Varianz der Bedarfsverteilung ist größer als die der Aufnahme:  
 In der Praxis kann dies z.B. bei Betrachtung eines Kollektivs aus einem Gefängnis der Fall sein. Durch die sehr ähnliche Ernährung kann es zu kleiner Streuung der Aufnahmeverteilung und somit zu Verzerrungen der Schätzung kommen. Je ähnlicher die mittlere Aufnahme und der mediane Bedarf desto weniger schlecht fällt die Schätzung aus (siehe Carriquiry (1999) Monte Carlo Simulationen).

Sollte der EAR (ungefähr) gleich der mittleren Aufnahme sein, so sind die Annahmen weniger bis gar nicht einflussreich auf die Schätzung. Zusammenfassend lässt sich sagen, dass die EAR cut-point Methode in vielen Situationen gut funktioniert, allerdings immer überprüft werden sollte, ob nicht eine der 3 Annahmen verletzt ist.

### 2.6.1 Andere cut-points als der EAR

Wenn nicht der Bedarfsmedian als Grenzwert genommen wird, muss von Verzerrungen der Anteilsschätzung ausgegangen werden. Dies gilt auch, wenn die obigen 3 Annahmen gelten sollten. National Research Council (2000)(Seite 102) empfiehlt ausdrücklich als cut-point nicht Referenzwerte wie den RDA zu verwenden, da dadurch eine Überschätzung des Anteils der Unterversorgten auftritt. Dies ist wenig verwunderlich, da per Definition dieser Referenzwerte nicht jede Person so einen hohen und die meisten einen niedrigeren Bedarf haben. Klarerweise wird der wahre Anteil überschätzt, wenn jede Person aus der Stichprobe mit dem Referenzwert verglichen wird. Ein kleines Zahlenbeispiel soll dies verdeutlichen:

Angenommen die 3 Annahmen der cut-point Methode gelten. Die Bedarfsverteilung soll eine Normalverteilung mit Erwartungswert 8 und Standardabweichung 1.5 sein. Die übliche Aufnahmeverteilung folgt einer Gammaverteilung mit "shape" Parameter 10 und "scale" gleich 1. Erwartungswert und Varianz der Aufnahme sind somit gleich 10. Der EAR (= 8) liegt also unter der mittleren Aufnahme. Die Varianz der Aufnahme ist weitaus größer als die des Bedarfs und auch die Unabhängigkeit soll gelten. Wie im Abschnitt über den Probability Approach schon besprochen wurde, liefert dieser Ansatz eine unverzerrte Schätzung. In unserem Fall (beide Verteilungen vollständig bekannt) können wir mit dem Ansatz auch das "exakte" Ergebnis numerisch berechnen <sup>2</sup>. Die EAR cut-point Methode würde Schätzungen rund um 28,3% <sup>3</sup> liefern und wäre somit ein recht guter Schätzer, wenn nicht die ganze Bedarfsverteilung bekannt wäre. Würde z.B. der RDA (in diesem Fall  $RDA = 8 + 1.96 * 1.5 = 10.94$ ) als cut-point verwendet werden, so würden wir Schätzungen rund um 65,3% bekommen und somit deutlich überschätzen. Es macht also keinen Sinn Referenzwerte zu verwenden, die für

<sup>2</sup> in R z.B. mit `mean(1-pnorm(rgamma(1000000, shape=10, rate=1), mean=8, sd=1.5)) ≈ 29,9%`

<sup>3</sup> in R `pgamma(8, shape=10, scale=1)`



fast alle Personen als ausreichend definiert sind.

Um dies verständlicher zu machen, kann auch wieder die Abbildung 2.3 betrachtet werden. Links haben wir den Fall, wo der EAR als cut-point verwendet wird (normalverteilter Bedarf und normalverteilte Aufnahme). Rechts haben wir dieselbe Grafik mit dem RDA als cut-point. Durch die höhere (weiter rechts liegende) Grenze ist der Anteil der richtig positiv klassifizierten Personen gestiegen und der Anteil der falsch negativen gesunken (in der Abbildung sogar auf gleich 0). Dies bedeutet eine höhere Sensitivität<sup>4</sup> gegenüber dem EAR als cut-point. Wenn das Ziel ist, auf Individualebene möglichst viele Personen richtig als unterversorgt zu identifizieren (hohe Sensitivität), so ist der RDA als cut-point zulässig und sinnvoll. Allerdings ist nebenbei auch die Spezifität<sup>5</sup> gesunken, da nun deutlich mehr Personen als falsch positiv und weniger als richtig negativ eingestuft werden. Wenn also die Frage besteht, möglichst gut den Anteil der Unterversorgten (jener Anteil über der Bedarfsgrenze) abzuschätzen, so müssten beide Kennzahlen maximiert werden. Anders ausgedrückt möchten wir die Anzahl der richtig klassifizierten Personen maximieren. Wenn der RDA allerdings als cut-point verwendet wird, so werden verhältnismäßig mehr Personen, welche nicht unterversorgt sind falsch klassifiziert (falsch positiv), als die falsch Negativen verringert und es kommt dadurch zu deutlichen Überschätzungen des wahren Anteils.

Laut National Research Council (2000)(Seite 103) ist es ebenso nicht sinnvoll den Gruppenmittelwert der Aufnahmeverteilung mit dem RDA zu vergleichen und bei Überschreitung automatisch zu schließen, dass das Kollektiv ausreichend versorgt ist. Auch der Vergleich mit einem Prozentsatz des RDAs kann irreführend sein und zu Fehlschätzungen führen (siehe auch de Lauzon et al. (2004)). In vielen Fällen (Standardabweichung der Aufnahmeverteilung ist größer als die des Bedarfs) lässt sich sagen, dass die mittlere Aufnahme deutlich über dem Referenzwert liegen muss, damit fast alle Personen ausreichend versorgt sind. Für die Auswertungen im 5. Kapitel sollten also Werte ähnlich dem EAR zum Einsatz kommen, um möglichst verlässliche Schätzungen zu bekommen.

Für manche Nährstoffe gibt es weder eine Bedarfsverteilung noch eine Schätzung des EAR oder RDA. Es wird dann auch nicht empfohlen, andere cut-points wie z.B. den AI zu verwenden. Als cut-point verwendet, kann dies zu klaren Fehlschätzungen führen. Stattdessen soll mit dem AI nur qualitativ beurteilt werden. Sollte dieser unter der mittleren Aufnahme des Kollektivs liegen, so ist dies ein gutes Zeichen, dass die meisten Leute ausreichend versorgt sind. Im umgekehrten Fall kann eher davon ausgegangen werden, dass nur wenig Leute ausreichend versorgt sind.

## 2.7 Auswertungen in dieser Arbeit

Obwohl bekannt ist, dass die Empfehlungen als cut-points zu Überschätzungen der Unterversorgung führen, wurden sie dennoch im Ernährungsbericht verwendet. Sogar manche DACH-“Schätzungen“ wurden als Grenzwerte verwendet (z.B. bei Vitamin D). Das eingangs erwähnte Zahlenbeispiel mit Vitamin B12 (=Empfehlung) zeigt also deutlich auf, dass hier schon eine Überschätzung zustande kommt. Der verwendete Grenzwert (Empfehlung= $3\mu\text{g}$ ) für die cut-point Methode sollte eher durch den EAR (=2,4 $\mu\text{g}$ ) ersetzt werden, wenn die Prävalenz der Unterversorgung das Ziel der Schätzung ist. Dies stellt allerdings nur eine Fehlerquelle dar. Wie schon erwähnt, wurde im Ernährungsbericht die Mittelwertmethode (welche zu hohe Varianz aufweist) als Aufnahmeverteilung verwendet. Dies führt ebenfalls zu Fehleinschätzungen.

Die Erkenntnisse die nun über die Schätzung des Anteils der Unterversorgten gewonnen wurden, werden für die Neuauswertung der Ernährungsbericht Daten verwendet. Wenn möglich werden die EAR Werte des Institute of Medicine für die Anteilsschätzungen verwendet. Sollte eine Bedarfsverteilung bekannt sein, so wird diese Information ebenfalls verwendet und zusätzlich ein Vergleich des Probability Approach und der EAR cut-point Methode gegeben. Für Nährstoffe, wo nur ein “Schätzwert“ oder “Richtwert“ bzw. ein AI vom Institute of Medicine bekannt ist, wird nur eine qualitative Auswertung durchgeführt. Zusätzlich sollen

<sup>4</sup>Sensitivität: Jener Anteil der unterversorgten Personen, die richtig klassifiziert werden. richtig positiv / (richtig positiv+falsch negativ)

<sup>5</sup>Spezifität: Jener Anteil der nicht unterversorgten Personen, die richtig klassifiziert werden. richtig negativ / (richtig negativ+falsch positiv)

immer wieder die Ergebnisse des Ernährungsberichtes von 2012 mit den neu ausgewerteten Erkenntnissen verglichen werden.

Bemerkung: Der Anteil jener Personen, welcher weniger zu sich nimmt als ein bestimmter Grenzwert darf sehr wohl berechnet werden. Nur die mögliche inhaltliche Interpretation muss berücksichtigt werden. Wenn Werte wie der RDA oder DACH Empfehlungen verwendet werden, so darf dieser Anteil nicht als jener Anteil der Personen die von einem Mangel gefährdet sind beurteilt werden. Er bedeutet inhaltlich lediglich, dass so und so viele Personen weniger konsumieren. Wird der EAR Wert verwendet, so kann auch zusätzlich inhaltlich argumentiert werden, dass dieser Anteil die Prävalenz der Personen mit Mangel widerspiegelt.

# Kapitel 3

## Daten

In diesem Kapitel wird auf die Daten des Ernährungsberichtes von 2012 näher eingegangen und eine deskriptive Analyse gegeben. Die Daten liegen hauptsächlich in SPSS Dateien vor. Diese wurden für die weiteren Auswertungen in das Statistikprogramm R eingelesen. Für die Auswertungen in R wurde die zur Zeit neueste Version von R (Version 3.2.2, (R Core Team, 2015)) und R-Studio (Version 0.99.465, (RStudio Team, 2015)) verwendet. Für die Tests auf Unterschiede der Gruppen, wurde der Wilcoxon-Rangsummentest mit Stetigkeitskorrektur (R Funktion `shapiro.test`) verwendet, da meist nicht von normalverteilten Daten ausgegangen werden konnte.

Die Daten des Ernährungsberichts 2012 beinhalten Daten für Kinder, Erwachsene und Senioren. Für diese Arbeit wurden nur die Daten von den Erwachsenen herangezogen. Diese beinhalten Einträge bezüglich Nahrungsaufnahmen von 419 Personen. Zu einer Person sind kaum Angaben und keine Einträge bezüglich der 24 Stunden Befragungen vorhanden. Für die Datenauswertung können also nur 418 Personen verwendet werden. Davon sind 168 männlich und 250 weiblich. Das Alter verteilt sich von 18 bis 64 Jahre (Mittel: 42 für Männer, 38.8 für Frauen). Daten zu Körpergröße und Körpergewicht finden sich in Tabelle 3.1. Dabei handelt es sich um gemessene Werte, die nicht für alle Personen vorhanden sind.

Geschlecht	#Personen	Körpergröße	Körpergewicht	BMI
Weiblich	192	165.3	65.4	24
Männlich	121	178.7	82.6	25.9

Tabelle 3.1: Körperabmessungen: Die Angaben entsprechen den gewichteten Mittelwerten und sind nur für abgemessene Personen angegeben.

Neben den üblichen Personenangaben gibt es auch noch zu den meisten Erwachsenen Messungen bezüglich der Körperzusammensetzung (Körpermasse, Wasseranteil, Fettanteil, usw.). Diese kommen allerdings bei den statistischen Auswertungen dieser Arbeit nicht zum Einsatz. Wichtiger hingegen sind die Angaben zum persönlichen Lebensstil, Rauchverhalten, Nahrungsergänzungsmittel usw. Sie können helfen, die Datenlage besser zu analysieren und auch mögliche Informationen für die Verbesserung der statistischen Schätzungen zu gewinnen.

Der eigentliche Kernpunkt sind die Daten der 24 Stunden Protokolle. Diese beinhalten für 110 Personen die Aufnahmen eines 24 Stunden Recalls und die Aufnahmen von zwei Befragungen für 308 Erwachsene. Die erste Auswertung erfolgte mittels Indizes und den BMI Angaben um mögliche Unterschiede in den zwei Gruppen zu finden.

### 3.1 Analyse der Personen mit einem und zwei Recalls

#### 3.1.1 BMI

Vergleicht man die BMI Werte der beiden Gruppen, so lässt sich ein kleiner Unterschied aufzeigen. Die Gruppe mit nur einem Recall hat etwas höhere Werte beim Minimum, ersten Quantil, dem Median und Mittelwert, sowie dem 3. Quantil. Einzig das Maximum ist in der Gruppe mit 2 Recalls deutlich höher. Es muss aber auch erwähnt werden, dass es sich hier

nicht um gemessene, sondern um berichtete Werte handelt und wir dadurch mit Ungenauigkeiten zu kämpfen haben. Ein Vergleich der gemessenen Werte kann nicht durchgeführt werden, da diese für die Personen mit nur einer 24 Stunden Befragung kaum vorliegen (gerade mal 5 gemessene Werte).

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
1 Recall [n=110]	17.5	22.3	24.6	24.9	27.5	34.4	2
2 Recalls [n=308]	16.3	21.5	23.3	24.2	26.4	39.3	24

Tabelle 3.2: BMI Vergleich der Gruppen mit einem und zwei Befragungen (nur geschätzte Körpermaße der Personen). Nicht gewichtete Daten.

Dass die geschätzten Körpermaße der Personen tatsächlich weit von den gemessenen abweichen können, zeigt ein Vergleich bei jenen Personen, wo beide Angaben vorhanden sind. Die Abweichungen bei den Erwachsenen zwischen geschätzten und gemessenen Werten liegt für die Körpergröße zwischen  $-10.5$  und  $4.7\text{cm}$  und für das Körpergewicht zwischen  $-9.2$  und  $11.3\text{kg}$ . Die Unterschiede zwischen den zwei Gruppen sind also nur schwer zu interpretieren. Es liegt auf jeden Fall kein extremer Unterschied vor ( $p$ -Wert= 0.03).

### 3.1.2 Indizes

Für weitere Analysen wurden 3 Indizes herangezogen. Dabei handelt es sich um einen Lifestyle, einem Nährstoff und einem Nahrungsmittel Index. Als erstes wurde der Lifestyle der zwei Gruppen verglichen.

#### Lifestyle Index

Der Lifestyle Index lehnt stark an jenen von Kwaśniewska et al. (2007) an und kann Werte von 0 bis 4 annehmen. Je höher der Index, desto besser soll der Lebensstil sein. Zur Bewertung werden 4 Kategorien (Rauchen, Gemüse/Obst Konsum, BMI, Bewegung) herangezogen. Für das Erreichen der Bedingung erhielt man einen Punkt und umgekehrt keinen. Die Nichtraucher und ehemaligen Raucher erhielten einen Punkt. Leute mit einem Obst und Gemüse Konsum von mindestens 5 Portionen ( $>400\text{g}$ ) pro Tag (Mittelwert der Beobachtungen wurde verwendet) erhielten ebenfalls einen Punkt. Die BMI Kategorie wurde mit 1 benotet für Personen zwischen  $18.5$  und  $25\text{kg}/\text{m}^2$ . Auch hier wurden wieder die selbst berichteten Werte genommen, da sonst zu wenig Daten für einen Vergleich vorliegen würden. Als letztes wurde noch die Bewegung benotet. Personen mit mindestens 2 Stunden Bewegung (sowohl anstrengende und moderate Aktivität) pro Woche wurden als gut eingestuft. Die Bewegung der Personen wurde mittels Fragebogen ermittelt, was natürlich auf der Einschätzung beruht und daher Fehler beinhalten kann. Für eine weitere Erklärung wie diese Daten zustande gekommen sind, wird hier auf den Österreichischen Ernährungsbericht 2012 verwiesen.

Die Aufteilung der Personen in die Klassen 0 bis 4 sind in Tabelle 3.3 zu sehen. Es zeigt sich kein eindeutiger Unterschied der zwei Gruppen ( $p$ -Wert 0.674). Während bei den Personen mit einer Befragung ein paar Prozent mehr die beste Benotungsklasse (4) bekommen haben, so sind doch deutlich mehr Personen mit Klasse 3 von der Gruppe mit 2 Befragungen. Auch bei den schlechten Klassen zeigt sich kein großer Unterschied.

	0	1	2	3	4
1 Recall [%] n=110	4.2	14.7	37.9	27.4	15.8
2 Recalls [%] n=308	2.8	16.7	31.5	35.5	13.5

Tabelle 3.3: Lifestyle Index

## Food Based Quality Index

Der nächste Index bewertet die Personen anhand ihrer Essgewohnheiten. 7 Nahrungsgruppen wurden dazu betrachtet und bei nicht Einhalten einer gewissen Schranke an Portionen wurde dem Individuum kein Punkt und sonst ein Punkt gegeben. Der Index setzt sich wieder aus der Summe aller 7 Kategorien zusammen. Dies lehnt an dem Index von Löwik et al. (1999) an, allerdings mit der Veränderung, dass die Portionsgrößen an die Empfehlungen der Österreichischen Gesellschaft für Ernährung angelehnt wurden. Dadurch wurden folgende Grenzwerte für die 7 Nahrungsgruppen ermittelt. Die Aufnahme von Nudeln, Brot und Cerealien sollte gesamt über 200g betragen. Die von Reis und Kartoffeln sollte 150 oder 200g überschreiten, wobei hier ein Punkt vergeben wurde, wenn entweder 150g Reis oder 200g Kartoffeln oder eine Kombination aus beiden konsumiert wurde. Gemüse plus Hülsenfrüchte sollten nicht weniger als 330g betragen, ebenso wie für Obst 330g. Milch und Milchprodukte können wieder als Mischung betrachtet werden, wobei 200ml Milch oder 180g Milchprodukte oder eine Kombination aus beiden als ausreichend angesehen wurde. Käse sollte mehr als 50g und Fleisch, Fisch und Eier zusammen mehr als 150g konsumiert werden. Der Index zielt darauf ab, dass eine große Variabilität in der Ernährung als wichtig angesehen wird, weshalb von diesen 7 Kategorien die gerade beschriebene Menge konsumiert werden sollte. Die Ergebnisse der beiden Gruppen können in Tabelle 3.4 betrachtet werden.

	0	1	2	3	4	5
1 Recall [%] n=110	1.8	13.6	39.1	32.7	10.9	1.8
2 Recalls [%] n=308	7.5	29.2	40.9	14.9	6.5	1.0

Tabelle 3.4: Food Based Quality Index

Auffällig ist in erster Linie, dass für beide Gruppen nur Indexwerte bis maximal 5 erreicht wurden. Als zweites fällt auf, dass die Gruppe mit einem Recall doch besser abschneidet (p-Wert < 0.0001). Vor allem in den Kategorien 3 und 4 halten sie einen höheren Prozentsatz. Beachtet werden muss hier, dass für die Leute mit 2 Befragungen das Mittel der Aufnahme über die zwei beobachteten Tage genommen wurde. Da die Daten zu den Nahrungsgruppen schon aggregiert für die 2 Tage vorliegen, kann hier nicht der Vergleich für die einzelnen Tage (für die Personen mit 2 Recalls) betrachtet werden.

## Nutrient Quality Index

Der letzte Index soll die Nährstoffe betrachten und die Zufuhr je Person beurteilen. Angelehnt ist die Erstellung dieses Nährstoffqualitätsindex (NQI) an Gedrich and Karg (2001). Dabei können alle Nährstoffe, die eine Referenzgrenze oder ähnliche Grenzwerte für ihre Aufnahme haben, verwendet werden. Für die Auswertung in dieser Arbeit wurden die Nährstoffe Vitamin D, Vitamin B9, Kalzium, Jod und Zink sowie die gesättigten Fettsäuren, Kochsalz und Ballaststoffe verwendet. Diese stellen zum Großteil Nährstoffe dar, welche nicht in optimaler Menge, laut Ernährungsbericht, aufgenommen werden (Zink ist hier noch eine Ausnahme). Berechnet wird der NQI nun in zwei Schritten.

Im ersten werden für alle betrachteten Nährstoffe "intake quality scores" (IQS) berechnet. Diese Scores werden für Nährstoffe mit Ober- und Untergrenze verschieden bestimmt. Für Nährstoffe mit Untergrenze (Vit D, B9, Kalzium, Jod, Ballaststoffe, Zink) wird der IQS als Verhältnis der Aufnahme zum Referenzwert mal 100 berechnet. Bei überschreiten des Referenzwertes bleibt der IQS auf 100. Somit kann der IQS für jeden Nährstoff die Werte 0 (gar keine Aufnahme) bis 100 (Referenzwert erreicht bzw. darüber) annehmen. Für Nährstoffe mit einer Aufnahmeobergrenze wird der IQS mit 100 bestimmt solange die Aufnahme unter der Grenze bleibt und mit 0, wenn sie über dem zweifachen Grenzwert liegt. Dazwischen wird der IQS mit  $(2 - \frac{\text{Aufnahme}}{\text{Referenzwert}}) * 100$  berechnet. Somit kann auch hier der IQS nur Werte von 0 (Aufnahme deutlich zu hoch) bis 100 (Aufnahme unter Grenze) annehmen.

Der zweite Schritt beinhaltet die Zusammenfassung aller 8 Scores durch das harmonische Mittel. Somit kann gewährleistet werden, dass der NQI ebenfalls Werte zwischen 0 und 100 annehmen kann und sehr niedrige Werte erzielt, wenn schon eine einzige Nährstoffversorgung schlecht aussieht. Im extremsten Fall reicht es schon aus, einen NQI Wert von 0 zu erhalten, wenn nur ein Nährstoff IQS den Wert 0 hat.

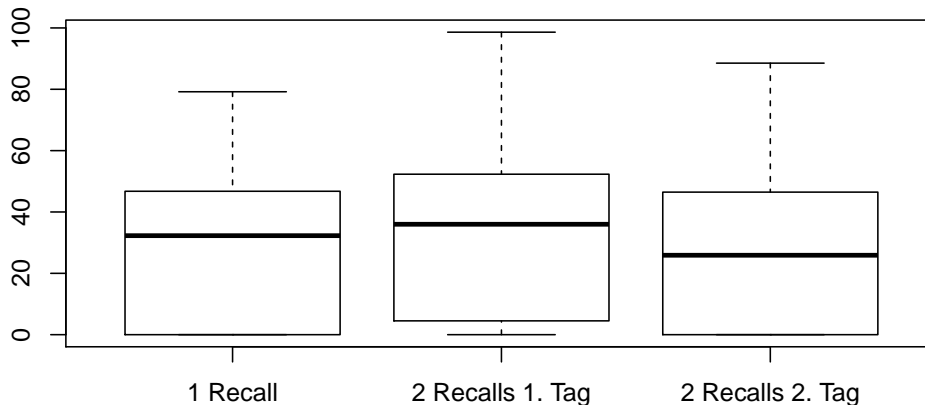


Abbildung 3.1: Nutrient Quality Index

Als Referenzwerte für die Nährstoffe mit Untergrenze wurden die DACH Werte genommen (bzw. für Vitamin D der Referenzwert der auch im Ernährungsbericht 2012 verwendet wurde). Für Salz wurde wie im Ernährungsbericht 2012 eine Obergrenze von 6g verwendet. Für die gesättigten Fettsäuren wurden 10 Energie% als Grenze genommen. Sprich bei einer Person mit 2000 kcal Energieaufnahme sollten nicht mehr als 200 kcal ( $\frac{200\text{kcal}}{9.3\text{kcal/g}} = 21.5\text{g}$ ) aus gesättigten Fettsäuren stammen.

In der nächsten Abbildung 3.1 ist ein Boxplotvergleich des NQI für die Gruppe mit einem Recall mit der Gruppe mit 2 Recalls für jeden Tag separat dargestellt. Personen mit nur einem Recall haben eine etwas kleinere Streuung als Leute mit 2 Befragungen. Auffällig ist der niedrige Median vom 2. Tag (25.9) verglichen mit einem Recall (32.3) und zwei Recalls 1. Tag (36). Da keine Unabhängigkeit zwischen den Gruppen mit zwei Recalls angenommen werden kann, wurde hier auf einen statistischen Vergleichstest der 3 Gruppen verzichtet. Stattdessen wurde der Wilcoxon Rangsummentest einmal auf die Differenz der beiden abhängigen Gruppen (p-Wert= 0.0014) angewandt und einmal bezüglich dem Unterschied, zwischen dem Mittelwert aus den NQI Werten (der zwei Tage) und den NQI Werten der Gruppe mit nur einem Recall (p-Wert= 0.297), verwendet. Der zweite Recall Tag liefert also deutlich schlechtere Werte, was - bei näherem betrachten der einzelnen IQS Werte - vor allem an den Ballaststoffen, Salz und den gesättigten Fettsäuren liegt.

Ein weiterer auffälliger Punkt sind die relativ niedrigen mittleren und medianen Werte für alle drei Gruppen. Dies liegt hauptsächlich an den niedrigen Aufnahmewerten von Vitamin D (Mittelwert=3.1 $\mu\text{g}$ ). Der Referenzwert hingegen liegt bei 5  $\mu\text{g}$  und damit deutlich darüber. Durch das harmonische Mittel werden die Werte des NQI stark in Richtung des kleinsten IQS, in dem Fall Vitamin D, gezogen.

## 3.2 Weitere Dropoutanalyse

Wie oben schon angekündigt, beschäftigt sich dieser Teil damit, die einzelnen Nährstoffe auf Hinweise von Unterschieden zwischen den Personen mit einer und zwei 24 Stunden Befragungen. Bei Betrachtung der einzelnen Tagesaufnahmen für die Nährstoffe und Energie ließen sich kaum grobe Unterschiede feststellen. Sowohl deskriptive Statistiken wie Median, Mittelwerte und Quantile wurden in Betracht gezogen. Die größten Unterschiede sind mittels Dichteplots ("density" Funktion in R) in der Abbildung 3.2 zu sehen. Sowohl die Energieaufnahme als auch die Makronährstoffe und vereinzelte Mikronährstoffe sind darin abgebildet. Jede Grafik beinhaltet drei Linien, welche die unterschiedlichen Gruppen darstellen. Die schwarze Linie stellt die Aufnahme der Personen mit zwei 24 Stunden Befragungen am 1.

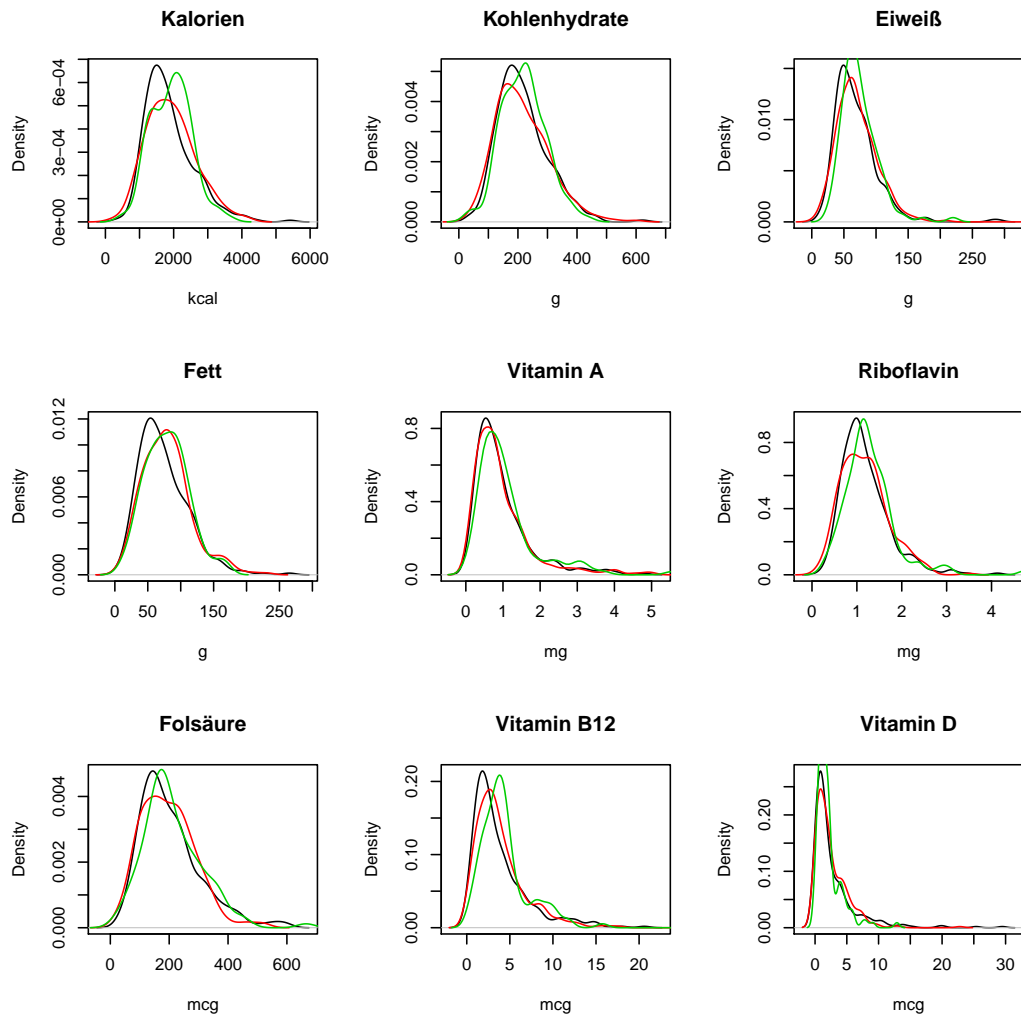


Abbildung 3.2: Dropoutanalyse: Aufnahme der Gruppe 2 Recalls 1.Tag (schwarze Linie) vs. Gruppe 2 Recalls 2.Tag (rote Linie) vs Gruppe 1 Recall (grüne Linie)

Tag dar. Die rote Linie stellt die Beobachtungen derselben Personen am zweiten Tag dar. Zuletzt stellt die grüne Linie die Tagesaufnahmen der Gruppe mit nur einer Befragung dar. Auffällig ist hier, dass die Aufnahme der letzten Gruppe meist rechts von den anderen liegt, was auf eine leichte Verschiebung der Verteilung zu höheren Aufnahmen darstellt. Individuell für jeden Nährstoff sind diese Änderungen gering, doch zeigt sich dies doch häufiger als zu erwarten wäre. Da sich die aufgenommenen Kalorien aber auch unterscheiden und die Nährstoffe damit zusammenhängen, ist dies möglicherweise auf die höhere Gesamtzufuhr an Nahrungsmitteln zurückzuführen. Die Unterschiede in der Kalorien und den Makronährstoffaufnahmen (Median) werden in Tabelle 3.5 wiedergegeben. Die erhöhte Kalorienaufnahme zeigt sich vor allem durch eine erhöhte Kohlenhydrat- und Fettaufnahme in der Gruppe mit nur einer Befragung. Da die Nährstoffmengen zum Teil eine Abhängigkeitsstruktur mit der Energieaufnahme vorweisen, wurde nur für die Energieaufnahme für Recall 1. Tag und Recall 2. Tag mittels statistischem Test verglichen ( $p$ -Wert= 0.316), wobei kein signifikanter Unterschied festgestellt werden kann. Auch zwischen der mittleren Energieaufnahme von der Gruppe mit zwei Recalls und der mit nur einem, kann kein großer Unterschied festgestellt werden ( $p$ -Wert= 0.399).

	1 Recall	2 Recalls 1.Tag	2 Recalls 2.Tag
Energie [kcal]	1952.7	1737.7	1844.8
Kohlenhydrate [g]	220.0	201.5	203.8
Eiweiß [g]	68.0	62.3	65.7
Fett [g]	77.9	66.0	75.5

Tabelle 3.5: Vergleich der medianen Energie- und Makronährstoffaufnahme anhand der Befragungen (ungewichtet). n=110 für 1 Recall und n=308 für 2 Recalls

### 3.3 Vergleich der Zeitpunkte

Eine weitere wichtige Fragestellung ergibt sich durch die Wiederholung der Befragung bei den meisten Personen. Die Lage und Streuungen zwischen den zwei Beobachtungen, sowie die Korrelationen, wurden für alle Nährstoffe betrachtet. Für die statistischen Maße wurden jeweils die auf Ausreißer weniger empfindlichen angewandt. Dies lässt sich damit begründen, dass die meisten Nährstoffaufnahmen rechtsschiefe Verteilungen aufweisen und es dadurch auch öfters zu vereinzelt großen Werten (“Ausreißern“) kommt. Da diese den Pearson Korrelationskoeffizienten und den Mittelwert, sowie die Varianz stark beeinflussen wurden stattdessen der Spearman Koeffizient, Median und Interquartilsabstand verwendet. Die Ergebnisse sind in Tabelle 3.6 ersichtlich. Die zwei Tage unterscheiden sich vor allem in einer höheren medianen Energieaufnahme, sowie klar höherer Fettaufnahme am zweiten Tag. Auch die Streuung liegt am zweiten Tag höher. Insgesamt ist von den 29 Nährstoffen und der Energie 22-mal der Median am zweiten Tag höher. Ebenso ist 21-mal der Interquartilsabstand am zweiten Tag größer. Die größten prozentuellen Unterschiede zeigen sich bei Vitamin C (19.1% höhere Aufnahme am 1.Tag) und bei Vitamin B12 (17.8% niedrigere Aufnahme am 1.Tag). Bei beiden veränderte sich der Interquartilsabstand auch kaum über die zwei Tage.

Die Korrelationskoeffizienten zeigen Werte von 0.1 (Natrium) bis 0.5 (Fluor) auf und deuten damit auf (erwartete) Abhängigkeiten der zwei beobachteten Tage auf. Für die Modellierung der Tagesaufnahmen ist es also sinnvoll Modelle zu verwenden, welche eine Abhängigkeit zulassen. Bei einer Modellierung, wie in Gl:2.1 ist dies erfüllt.

### 3.4 Verteilungen der Nährstoffaufnahmen

Werden die Dichteschätzer (unter anderem Abbildung 3.2) bzw. Histogramme der einzelnen Nährstoffe betrachtet, so kann erkannt werden, dass nur die wenigsten Nährstoffe eine symmetrische Aufnahmeverteilung und noch weniger eine möglichst normalverteilte haben. “Unverzerrte“ Schätzungen der Schiefe (Fisher Methode, Funktion “skewness“ in der Library EnvStats) ergeben Werte zwischen 0.8 (Energie, Kalium, Kohlenhydrate) und 15 (Vitamin A), was ebenfalls zeigt, dass bei keinem Nahrungsbestandteil von einer symmetrischen Verteilung ausgegangen werden kann. Es bestätigt sich also, dass es notwendig sein wird, möglichst gute Normalverteilungstransformationen zu finden, um gute Schätzungen der Anteile der Unterversorgten zu ermöglichen. Problematisch wird dies vor allem bei Nährstoffen, die sehr stark rechtsschief sind und große “Ausreißer“ beinhalten. Überlegungen müssen stattfinden, ob die Werte rausgenommen werden oder adequate Anpassungen stattfinden.

#### 3.4.1 Episodisch konsumierte Nährstoffe

Ein weiterer interessanter Punkt ist die Anzahl an Beobachtungen, wo keine Aufnahme des Nährstoffes stattfand. Von den 21054 Tagesbeobachtungen sind gerade mal 10 mit Werten gleich 0. Vitamin D stellt dabei mit 6 Beobachtungstagen, wo keine Aufnahme stattfand, das Maximum dar. Es handelt sich dabei um 6 verschiedene Personen, wobei nur eine dabei ist mit nur einer Tagesbeobachtung und damit einen Mittelwert der zwei beobachteten Tage gleich 0 hat. Es muss also für die Auswertung bzw. Schätzung einer üblichen Aufnahmeverteilung, für diese untersuchten Nährstoffe, kaum eine Berücksichtigung überlegt werden. Die Modellierung kann daher mit den “einfachen“ 1-Part Modellen der Hauptmethoden durchgeführt werden.



	Median 1	Median 2	IQR 1	IQR 2	Spearman
Energie [kcal]	1737.7	1844.8	863.5	992.3	0.39
Kohlenhydrate [g]	201.5	203.8	102.3	122.0	0.44
Eiweiß [g]	62.3	65.7	37.8	42.0	0.21
Fett [g]	66.0	75.5	46.8	47.3	0.25
Vitamin A [mg]	0.8	0.8	0.7	0.8	0.24
Vitamin B1 [mg]	1.0	1.0	0.6	0.6	0.18
Vitamin B2 [mg]	1.1	1.1	0.6	0.7	0.31
Vitamin B3 [mg]	24.5	25.8	17.1	16.3	0.24
Vitamin B5 [mg]	3.5	3.6	2.1	2.4	0.30
Vitamin B6 [mg]	1.4	1.3	0.9	0.8	0.38
Vitamin B7 [mcg]	34.0	35.7	25.1	24.3	0.35
Vitamin B9 [mcg]	183.7	186.7	120.0	127.5	0.40
Vitamin B12 [mcg]	2.7	3.1	3.2	3.2	0.26
Vitamin C [mg]	94.1	76.1	108.5	108.3	0.33
Vitamin D [mcg]	1.6	1.8	3.1	3.2	0.26
Vitamin E [mg]	10.8	12.2	10.1	12.7	0.19
Vitamin K [mcg]	74.2	75.3	75.9	72.6	0.24
Kalzium [mg]	705.1	696.5	436.7	600.0	0.25
Chlor [mg]	4381.4	4693.6	2735.2	3490.6	0.12
Kupfer [mg]	1.6	1.6	0.9	0.8	0.40
Fluor [mcg]	744.1	788.4	434.9	523.1	0.50
Eisen [mg]	9.8	10.1	6.1	6.1	0.34
Jod [mcg]	111.9	115.5	100.8	101.0	0.16
Kalium [g]	2.5	2.4	1.2	1.4	0.43
Magnesium [mg]	296.3	293.5	141.1	168.4	0.37
Mangan [mg]	3.9	4.0	3.0	3.0	0.42
Natrium [mg]	2784.6	2958.5	1916.2	2211.2	0.09
Phosphor [mg]	1091.6	1089.2	541.2	566.7	0.33
Schwefel [mg]	652.9	689.7	400.7	433.0	0.24
Zink [mg]	8.4	9.5	4.7	5.3	0.26

Tabelle 3.6: Zeitpunktvergleiche der Nährstoffe und Energie. Ungewichtete Daten mit n=308

# Kapitel 4

## Vergleiche der Hauptmethoden

In diesem Kapitel sollen die statistischen Modelle NCI, MSM, SPADE und ISU miteinander verglichen werden. Ein erster inhaltlicher Vergleich wurde schon in Kapitel 2.4 gegeben. Daher wird hier vor allem auf datenbezogene Vergleiche eingegangen. Zuerst werden Ergebnisse von Simulationsstudien aus der Fachliteratur wiederholt. Danach werden die Methoden anhand der Ernährungsberichtsdaten verglichen und Schlussfolgerungen für die Neuauswertung getroffen.

### 4.1 Vergleich der Performance anhand von früheren Studien

Die 4 Methoden wurden schon mal von Souverein et al. (2011), sowohl an echten Daten sowie an einer Simulationsstudie, verglichen. Da sich nur wirklich etwas über die Performance der Methoden aussagen lässt, wenn der wahre Wert bzw. die wahre Verteilung bekannt ist, werden hier nur die Ergebnisse der Simulationsstudie vorgestellt. 3 Aspekte wurden dabei von Souverein et al. (2011) getestet. Einerseits der Effekt der Stichprobengröße, der Effekt der Schiefe der Verteilung, sowie der Effekt des Varianzverhältnisses von within- zu between-person Varianz. Verglichen wurden die 4 Methoden noch zusätzlich mit der einfachen empirischen Verteilungsfunktion der within-person Mittelwerte (WPM). Erwarteterweise schneidete WPM für alle Quantile deutlich schlechter ab, weswegen es nur für den Vergleich des Gesamtmittelwertes verwendet wurde. 100 Simulationen für jedes Szenario wurden durchgeführt, wobei immer von einer Normalverteilung gezogen wurde (2 Tagesbeobachtungen pro Person), um die Aufnahme auf der transformierten Skala zu bekommen. Die mittels Inverse der einparametrischen Box-Cox Transformation erzeugten Daten wurden dann für die Methoden verwendet. Berechnet wurden für die 100 Simulationen die mittleren Verzerrungen sowie die Standardfehler.

Die Vergleiche der Stichprobengröße zeigten klarerweise, dass alle Methoden größere Standardfehler mit kleinerer Stichprobe produzieren. Während weder für den Mittelwert noch für die Quantile Verzerrungen festgestellt werden konnten, so konnten doch größere Standardabweichungen bei den Methoden ISU und MSM gegenüber NCI und SPADE festgestellt werden. Scheint also so, als ob NCI und SPADE etwas genauere Schätzungen liefern. Dies wurde allerdings nur für den Fall einer log-Normalverteilung mit Varianzverhältnis 1 betrachtet. Diese Aussage kann also nicht verallgemeinert werden.

Für den Effekt der Schiefe der Verteilung wurde mit verschiedenen Parametern der Box-Cox Inverse experimentiert. Je schiefere die Verteilung, desto ungenauer wurden die Schätzungen. Auch hier zeigten die ISU und MSM Methoden größere Standardabweichungen.

Die Varianzverhältnisse wurden mit 0.25, 1, 4 und 9 getestet. Je größer die within-person Varianz im Vergleich zur between-person Varianz wird, desto ungenauer werden die Schätzungen für alle 4 Methoden. Die NCI Methode zeigte sogar für hohe Varianzverhältnisse einen Bias, wobei damals noch die Taylorapproximation für die Rücktransformation verwendet wurde. Das Problem sollte nun also nicht mehr auftreten.

Die Schlüsse der Autoren waren, dass alle 4 Methoden sehr ähnliche Resultate erzielten und es nur schwer zu sagen ist, welche Methode besser für eine gewisse Situation ist. Auch wenn

MSM und ISU von den Standardfehlern her schlechter abgeschnitten haben, so waren dies doch nicht alle möglichen Szenarien, die in der Praxis vorkommen können. Vor allem der Einfluss von Kovariaten, Stör- oder Gewichtungsfaktoren müsste überprüft werden. Wichtig für Souverein et al. (2011) ist auf jeden Fall, dass die Schiefe und das Varianzverhältnis im Auge behalten wird, und für die Auswertungen und Interpretationen der Ergebnisse nicht vernachlässigt wird.

#### 4.1.1 Weitere Vergleiche

Tooze et al. (2010) verglichen ihre NCI Methode auch mit der ISU Methode und kamen zu dem Entschluss, dass diese für die Schätzung von Anteilen unter einem bestimmten Grenzwert sehr ähnliche Ergebnisse liefern. Dekkers and Slob (2012) stellten einen Vergleich der verschiedenen Rücktransformationsformeln an, wobei die Gaussquadratur mit der 9 Punktapproximation der ISU Methode und der exakten Rücktransformation wie in MSM verwendet wurden. Zusätzlich wurde noch die "schlechtere" Taylorapproximation sowie eine adaptierte Gaussquadratur getestet. Während die adaptierte Gaussquadratur zwar manchmal leicht bessere Ergebnisse als die normale Gaussquadratur bringt, so wurde auch darauf hingewiesen, dass die adaptierte Version ungefähr doppelt so lange Berechnungszeit benötigt. Weitere Ergebnisse der Studie waren, dass die Verwendung von zumindest 9 Knoten für die Gaussquadratur recht gut abschneidet und niedrigere Werte eher weniger empfehlenswert sind. Die Taylorapproximation zeigte klare Verzerrungen für stark schiefe Verteilungen sowie hohe Werte der within-person Varianz und ist daher nicht zu empfehlen (ähnliche Ergebnisse lieferte auch die Simulationsstudie von Haubrock et al. (2011)). Für die 9 Punktapproximation war das Ergebnis, dass sie recht ähnlich wie die Gaussquadratur abschneidet, jedoch leichte Verzerrungen aufzeigte. Für die exakte Rücktransformation wie in MSM wurde schlussgefolgert, dass die eingeschränkte Auswahl der  $\lambda$  Parameter kaum einen Einfluss auf die Schätzung der üblichen Aufnahmeverteilung hat. Weiters zeigten die Autoren, dass bei Transformationen die zu schlechter approximativer Normalität oder ungleicher within-person Varianz führen, auch die Ergebnisse der Rücktransformation zu verzerrten Resultaten führen können. Insgesamt gesehen wurde die Gaussquadratur als gute Möglichkeit der Rücktransformation für Box-Cox transformierte positiv schiefe Daten angesehen. Implementiert ist die Gaussquadratur nur in der SPADE Methode.

Auch anhand der hier beschriebenen Studien zeigt sich, dass keine allzu großen Unterschiede in den Ergebnissen der Methoden zu erwarten sind. Als nächstes vergleichen wir die Methoden anhand der Ernährungsberichtsdaten.

## 4.2 Vergleich anhand der Daten des Ernährungsberichtes

Die Methoden ISU, NCI, MSM und SPADE werden hier in Bezug auf die Nahrungsenergie, die Makronährstoffe und Vitamin A und Kalzium verglichen. Die Quantile (5, 25, 50, 75, 95%) der geschätzten üblichen Aufnahmeverteilung werden dazu herangezogen. Ebenfalls wird der geschätzte Mittelwert mit dem gewichteten Mittel der Mittelwertmethode verglichen um Verzerrungen festzustellen. Dies ist insbesondere interessant, da in den Vergleichsstudien die im vorigen Kapitel erwähnt wurden, nur der Fall für nicht gewichtete Daten betrachtet wurde und da keine groben Unterschiede feststellbar waren (Souverein et al., 2011). Auch mit den ungewichteten Daten des Ernährungsberichts bestätigt sich diese Aussage. Die nächsten Vergleiche werden daher nur mit den gewichteten Daten vollzogen. Ein Problem stellt hier die MSM Methode dar, da sie keine gewichtete Auswertung zulässt. Eine Möglichkeit dieses Problem zu umgehen, ist es die Daten, ähnlich dem Schritt in der ISU Methode, vorab zu "gewichten". Dazu wurden auch hier "neue" Daten mittels der Inversen der gewichteten empirischen Verteilungsfunktion gezogen. Erfolgt ist dies in dieser Arbeit mit der `wtd.quantile` Funktion aus dem R-Package "Hmisc"<sup>1</sup>. Die neuen Daten sind somit formal gesehen  $y_{ij}^{neu} = \hat{F}^{-1}(s_{ij}/(n*m))$ , wobei  $\hat{F}$  die geglättete (gewichtete) empirische Verteilungsfunktion und  $s_{ij}$  den Rang der ursprünglichen Daten  $y_{ij}$  darstellt. Die einfachen Mittelwerte und Varianzen der neuen Daten wurden mit den gewichteten Maßen der ursprünglichen Daten verglichen, um die Anpassung zu kontrollieren. Große Abweichungen konnten keine

<sup>1</sup>`wtd.quantile(data, weights=Gewichte, probs=(rank(data)/nrow(data)), type="i/n")`

festgestellt werden. Auch bei Betrachten der Quantile konnten keine groben Abweichungen festgestellt werden. Für den Fall der Kalorienaufnahme sind die Vergleichswerte in Tabelle 4.1 zu sehen. Zum besseren Vergleich wurden noch die normalen Quantilswerte der nicht gewichteten Daten angegeben. Die neuen Daten verhalten sich also approximativ wie man es von den gewichteten Daten erwarten würde. Mit diesem neuen Datensatz wurden nun die Auswertungen für MSM durchgeführt. Dies ist hier zulässig, da wir nicht auf Individual-ebene an den Aufnahmen der einzelnen Personen interessiert sind, sondern nur die “gesamte“ übliche Aufnahmeverteilung schätzen wollen. Ist man an Schätzungen für einzelne Personen interessiert, so sollte dieser Ansatz nicht gewählt werden.

	Mittelw.	Stand.abw.	5%	25%	50%	75%	95%
urspr. Daten	1903.4	707.58	964.18	1383.6	1817.9	2329.8	3201.6
gew. Kennzahlen	1962.7	715.59	986.50	1458.1	1851.1	2372.5	3305.3
gew. Daten	1962.4	714.45	985.60	1457.8	1849.6	2371.3	3302.7

Tabelle 4.1: Vergleich der ursprünglichen Daten bezüglich Kalorienaufnahme [kcal] mit den neuen “gleichgewichteten“ Daten

Die Ergebnisse der einzelnen Methoden wurde zusätzlich mit denen der einfachen Mittelwertmethodik verglichen. Sie wird in den nächsten Schritten wieder als within-person Mittelwertmethode (WPM) bezeichnet.

### Verwendete Software

Für die Auswertungen mit den angesprochenen Methoden wurden die zur Zeit aktuellsten Versionen verwendet. Für SPADE wurde das Package SPADE.RIVM Versionsnummer 3.0, für NCI die SAS Makro Versionen 2.1, für MSM die Version 1.0.1 über die Homepage (<https://msm.dife.de>) und für ISU das Program PC-SIDE Version 1.0 verwendet. Für die SAS Makros wurde die aktuellste SAS University Edition ([http://www.sas.com/en\\_us/software/university-edition.html](http://www.sas.com/en_us/software/university-edition.html)) verwendet (SAS Institute Inc, 2015). Hier kann angemerkt werden, dass die englische Version nötig ist, da es sonst zu Problemen bei der Ausführung der Makros kommt.

### Energieaufnahme

Als erstes wurde die Kalorienaufnahme betrachtet. Die Verteilung der Daten zeigt keine wirklichen extremen Ausreißer bis auf einen. Die Daten lassen sich auch mittels Box-Cox Transformation gut zu approximativer Normalität transformieren. Laut ISU Output ist auch keine Verletzung der Varianzhomogenität, sowie der Unabhängigkeit der Terme  $\mu_i$  und  $\epsilon_{ij}$  zu befürchten. Die Ergebnisse der geschätzten Quantile und Mittelwerte sind in Tabelle 4.2 zu sehen. Es können Unterschiede in den Resultaten der Methoden festgestellt werden. ISU und NCI bringen sehr ähnliche Resultate, während bei SPADE und MSM Unterschiede zu erkennen sind. Die geschätzte Aufnahmeverteilung von SPADE ist leicht nach links verschoben, im Vergleich zu den Verteilungen der anderen 3 Methoden. MSM liegt hingegen mit den Ergebnissen eher höher als die anderen. Beim Quantilvergleich mit der WPM Methode kann ganz klar die zu große Varianz der geschätzten Verteilung erkannt werden. Beim Vergleich der Mittelwerte mit dem Mittelwert der WPM Methode kann ebenfalls festgestellt werden, dass ISU und NCI am nächsten kommen. SPADE weicht hingegen am stärksten ab mit -1.5%. Betrachtet man die Logfiles genauer, so kann man erkennen, dass der Grund dafür in der Schätzung des linearen gemischten Modells mit den Gewichtungsfaktoren liegt. Der Intercept (bzw. Gesamtmittelwert) wird dabei zu niedrig geschätzt, was mit dem gewichteten Mittelwert der Daten überprüft werden kann. Wird der gewichtete Mittelwert für die Rücktransformation (kann mit Hilfe der R Funktion “f.gauss.quad“ vom SPADE Package durchgeführt werden) verwendet, so erhält man ebenfalls eine Schätzung der üblichen mittleren Aufnahme um 1960 kcal. Im Fall von nicht gewichteten Daten passiert dies nicht, daher kann schlussgefolgert werden, dass die Verwendung von Surveygewichten in SPADE nicht komplett problemlos funktioniert. Dieses Problem tritt auch bei den anderen betrachteten Nährstoffen auf, wie wir noch sehen werden. Für die Schätzung des gemischten Modells wird

	Mittel	5%	25%	50%	75%	95%	var.ratio
WPM	1959.7	1090.8	1525.6	1917.1	2296.1	3085.2	1.70
ISU	1960.3	1298.5	1653.3	1931.0	2235.6	2722.5	1.60
MSM	1968.4	1311.5	1651.3	1961.5	2247.7	2749.4	1.62
SPADE	1930.4	1240.3	1609.4	1899.3	2217.5	2726.2	1.36
NCI	1961.3	1283.9	1647.3	1936.1	2243.6	2732.4	1.56
SPADE gew.	1964.6	1287.4	1653.4	1937.4	2246.1	2733.9	1.54

Tabelle 4.2: Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik (WPM) bezüglich Kalorienaufnahme [kcal]

in SPADE die R Funktion “lmer“ aus dem Package lme4 verwendet. Näheres betrachten des “weights“ Arguments in der Funktion zeigt, dass dieses eigentlich nur als “prior weights“ für die Varianzschätzung gedacht ist. Das Problem könnte also darin liegen, dass lmer nicht für die Verwendung von Populationsgewichten gedacht ist. Da NCI (im Fall von keinen Kovariaten) eine sehr ähnliche Methode zu SPADE darstellt, kann diese als Vergleich herangezogen werden. Im NCI Fall wird ein (nicht) lineares gemischtes Modell geschätzt, wobei die Nicht-linearität nur daher kommt, da die Box-Cox Transformation im selben Schritt durchgeführt wird. Verwendet wird im SAS Makro die Funktion “NLMIXED“ für die Schätzung. Die Verwendung der Populationsgewichte (Häufigkeitsgewichte) führt dort zu besseren Intercept-schätzungen. Es bestätigt sich also, dass dieses Problem vermutlich an der lmer Funktion in SPADE liegt. Um dies bei SPADE zu lösen, wurde auch wie für MSM die Auswertung mit dem gleichgewichteten Datensatz extra durchgeführt. Die Ergebnisse sind auch in Tabelle 4.2 letzte Zeile zu sehen. Es kommen mit diesem Ansatz, im Vergleich zu den anderen Methoden ähnlichere Ergebnisse zustande. Auch der Mittelwert weicht nun nicht mehr so stark ab. Es scheint so, als ob dies möglicherweise eine bessere Art der Schätzung für die SPADE Methode darstellt.

In Tabelle 4.2 in der letzten Spalte sind noch die einzelnen Schätzungen der Varianzverhältnisse (within/between) zu sehen. Für das Varianzverhältnis der originalen Daten, wurden die Schätzungen wie in der MSM Methode (Gl:2.8,2.9 und 2.1), allerdings mit Anpassungen für den gewichteten Fall, verwendet. Abgesehen von der “normalen“ SPADE Methode liefern alle Methoden relativ ähnliche Schätzungen.

### Kohlenhydrat und Fett

Da sich diese beiden Nährstoffe gut verhalten, sprich gut transformieren lassen und keine großen Ausreißer beinhalten, sind sehr ähnliche Ergebnisse wie vorher vorzufinden. Daher werden diese Ergebnisse im Appendix der Arbeit aufgelistet und nur kurz besprochen. Auch die Testoutputs der ISU Methode zeigen keine Verletzungen der Annahmen (Varianzgleichheit, Unabhängigkeit).

Wie schon erwähnt zeigt sich wieder ein sehr ähnliches Bild (siehe Tabellen A.1 und A.2). SPADE mit den Gewichtungsfaktoren liefert wieder ein nach unten “verzerrtes“ Ergebnis, während sich die anderen Methoden sehr ähnlich verhalten. Das Problem kann wieder behoben werden, wenn SPADE auf den gleichgewichteten Datensatz angewandt wird. Die Varianzverhältnisse zeigen auch sehr ähnliche Ergebnisse für alle Methoden (abgesehen von der normalen SPADE Methode). Bei der Fettaufnahme ist ein höheres Verhältnis (um die 2.6) zu beobachten, als bei den Kohlenhydraten (ca. 1.7).

### Eiweiß

Bei der Eiweißaufnahme gehen wir wieder genauer ins Detail. Die Verteilung der Daten zeigt einen klaren “Ausreißer“ (Eiweißaufnahme= 469g) und zwei kleinere. Die Transformation der Daten könnte also zu Problemen führen. Der Testoutput zu den Annahmen zeigt keine Probleme. Allerdings muss beachtet werden, dass dies für die transformierten Daten getestet wurde und ISU die komplexeste Transformation darstellt. Die Annahmen könnte, also für die anderen Methoden womöglich verletzt sein. Wenn die Ausreißer nicht entfernt werden, sind folgende Ergebnisse und Unterschiede zwischen den Methoden erkennbar (siehe Tab. 4.3): Die normale SPADE Methode liegt beim Mittelwert wieder zu niedrig. Für die

	Mittel	5%	25%	50%	75%	95%	var.ratio
WPM	73.4	39.4	54.3	68.6	87.7	117.1	5.51
ISU	73.6	51.6	62.7	71.9	82.5	101.4	3.70
MSM	73.8	52.6	63.7	71.9	82.2	98.2	3.92
SPADE	72.2	46.9	59.9	70.6	82.7	103.0	2.39
NCI	73.2	49.4	61.6	71.8	83.0	101.8	3.10
SPADE gew.	73.2	50.1	62.1	71.8	82.8	100.9	3.29

Tabelle 4.3: Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Eiweißaufnahme [g]

gleichgewichteten Daten hingegen passen die SPADE Ergebnisse wieder gut zu den anderen Methoden. Insgesamt sind keine großen Unterschiede zwischen den Methoden zu erkennen, abgesehen von der Schätzung zum Varianzverhältnis. Hier zeigen sich größere Schwankungen als bei den vorigen Nährstoffen. Vor allem die normale SPADE Methode liegt weit unter den anderen Schätzungen. Alle zusammen liegen wiederum deutlich unter der Schätzung bezüglich der nicht transformierten Daten (Var.ratio= 5.51). Dies liegt allerdings an dem Ausreißer. Wird dieser niedriger gesetzt (auf  $y_{neu} = \text{Median} + 2 \cdot \text{IQR}$ ), so kommt ein deutlich niedrigeres Verhältnis zustande (= 3.64), welches schon eher zu den Schätzungen der Methoden passt.

Weiters fällt auf, dass die MSM Methode eine übliche Aufnahmeverteilung schätzt, die weniger streut, als die der anderen Methoden (siehe 5% und 95% Quantil). Sonst verhalten sich alle Schätzungen sehr ähnlich, trotz der Ausreißer. Es deutet darauf hin, dass es nicht viel Unterschied macht, welche Methode verwendet wird (wiederum abgesehen von der normalen SPADE Methode). Um zu überprüfen, ob die Ergebnisse auch ähnlich bleiben bei noch stärkeren Ausreißern, wurde Vitamin A als nächster Nährstoff gewählt.

## Vitamin A

Vitamin A beinhaltet zwei sehr grobe Ausreißer. Diese haben Werte von 39.5 und 29.9mg im Vergleich zu dem  $\text{Median} + 2 \cdot \text{IQR} = 2.3$ . Sie liegen also mehr als deutlich über den gewöhnlichen Aufnahmen. Laut ISU Testoutput liegt aber auch hier keine Verletzung der Annahmen vor. Betrachtet man nun die Ergebnisse in Tabelle 4.4 so fällt auf, dass die Mittelwerte etwas mehr abweichen als sonst, ebenso die 95% Quantile (maximaler Unterschied NCI und ISU mit um die 16%). Vor allem die NCI Methode liegt hier niedriger. Der Unterschied der normalen SPADE Methode kann hier fast gar nicht mehr festgestellt werden. Die Varianzverhältnisse streuen nun nochmal mehr als vorhin. Interessant ist auch das Varianzverhältnis berechnet an den ursprünglichen Daten (= 0.83). Durch die Ausreißer wirkt es so, dass die between-person Varianz hier sogar größer ist als sonst. Dies liegt vor allem daran, dass der maximale Vitamin A Wert von einer Person stammt, welche nur einen Aufnahmewert hat. Somit wird die Varianz der "Mittelwerte" der Beobachtungen riesig und auch die between-person Varianz deutlich überschätzt. Setzt man die zwei Ausreißer wieder herunter, so bekommt man ein Varianzverhältnis von 2.37. Auch wenn die Ergebnisse der Methoden sich relativ ähnlich (abgesehen vom 95% Quantil) sind, so muss für die allgemeinen Auswertungen, von denen Aussagen über die Population möglich sein sollen, Überlegungen für die Ausreißer ange-

	Mittel	5%	25%	50%	75%	95%	var.ratio
WPM	1.09	0.322	0.538	0.808	1.19	2.44	0.83
ISU	1.08	0.529	0.758	0.981	1.29	1.99	3.34
MSM	1.07	0.534	0.766	0.979	1.23	1.97	2.94
SPADE	1.03	0.516	0.747	0.964	1.24	1.77	3.12
NCI	1.02	0.548	0.767	0.970	1.22	1.68	4.14
SPADE gew.	1.03	0.530	0.761	0.972	1.23	1.72	2.83

Tabelle 4.4: Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Vitamin A Aufnahme [g]

stellt werden. Wie dies die Ergebnisse beeinflusst, wird im Kapitel der Neuauswertung noch behandelt.

## **Kalzium**

Für den letzten Vergleich wird noch der Vollständigkeit halber ein Mineralstoff betrachtet. Die Verteilung von Kalzium zeigt allerdings keine großen Ausreißer oder mögliche Probleme. Auch der Testoutput der ISU Methode liefert gute Ergebnisse bezüglich den Annahmen. Trotzdem kann eine interessante Veränderung festgestellt werden (siehe Tabelle A.3 im Anhang). Die NCI und SPADE Methode (für den gleichgewichteten Datensatz) liefern wieder sehr ähnliche Resultate, allerdings liegen sie (vom Mittelwert her betrachtet) fast gleich wie die normale SPADE Methode, welche vorher immer die niedrigsten Ergebnisse lieferte. ISU und MSM hingegen liefern die erwarteten ähnlichen Ergebnisse zu dem within-person Mittelwert. Betrachtet man die Quantile, so fällt auf, dass (abgesehen vom 5% Quantil) NCI und SPADE-gew. deutlich unter der ISU und MSM Methode liegen. Beim 95% Quantil sind die Abweichungen immerhin noch (im Vergleich zu ISU) über 2% groß. Im Vergleich zu den Unterschieden bei Vitamin A sind diese allerdings noch immer relativ kleine. Auf der nächsten Seite ist in der Abbildung 4.1 ein Dichtevergleich der 4 Hauptmethoden anhand der genannten Nährstoffe zu finden. Die Ergebnisse die oben schon anhand der Quantile erwähnt wurden, können hier nochmal bildlich betrachtet werden.

### **4.2.1 Verwendung für Subgruppen**

Für die Auswertungen im Ernährungsbericht sind immer wieder Vergleiche zwischen Männer und Frauen und bestimmter Altersgruppen relevant. Während die ISU und MSM Methode hier nur separat für jede untersuchte Gruppe angewandt werden können, so können die anderen 2 Methoden die Modellierung der Gruppenunterschiede mittels Kovariaten aufnehmen. Für NCI können sowohl das Geschlecht als auch die relevanten Altersgruppen verwendet werden. SPADE bietet hingegen nur die Modellierung des Alters und dadurch für die Altersgruppen. Die Auswertungen für die Geschlechterunterschiede müssen bei SPADE auch separat behandelt werden. In diesem Abschnitt vergleichen wir noch kurz die Ergebnisse bezüglich Kalorien-, Eiweiß- und Kohlenhydrataufnahme. Für den Ernährungsbericht waren meist nur Vergleiche der Mittelwerte für die Altersgruppen 18-24, 25-50 und 51-64 Jahre interessant. Diese wurden meist rein qualitativ mit den Referenzwerten verglichen, um ein Gefühl für den Status der Unterversorgung zu bekommen. Dazu wurde die Stichprobe für jede Altersgruppe und je Geschlecht separat ausgewertet, was in manchen Fällen in sehr kleinen Stichprobengrößen ( $n=17$ ,  $n=37$ ) resultierte. Diese Schätzungen waren klarerweise damit relativ unpräzise. Daher sind vor allem die Methoden SPADE und NCI für die späteren Auswertungen interessant. Bei SPADE nehmen wir die Variable Alter ins Modell mit auf und lassen sie anhand des fraktionalen Polynoms modellieren. Das Ergebnis für die Altersklassen entsteht dann intern in SPADE durch mitteln über die Personen in der Altersklasse. Für das Geschlecht wird auch noch separat modelliert. Bei NCI hingegen nehmen wir die Kovariablen Geschlecht und direkt die Altersklassen ins Modell mit auf und lassen dann die Ergebnisse für alle möglichen Kombinationen ausgeben. Wichtig ist dabei, dass die Altersklassen und das Geschlecht als numerische Variablen in das Modell eingehen, da sonst ein Fehler entsteht. Für die Altersklassen wurden die Zahlen 1, 2 und 3 für die aufsteigenden Klassen gewählt und beim Geschlecht 1 für männlich und 2 für weiblich. Die Varianzen werden für alle Kombinationen als gleich angenommen. Die Ergebnisse werden in den nächsten Abschnitten verglichen. Für SPADE werden zusätzlich auch wieder die Ergebnisse für den gleichgewichteten Datensatz für den Vergleich verwendet.

## **Energieaufnahme**

Tabelle 4.5 zeigt die Ergebnisse für die Energieaufnahme. Obwohl wir nur die Mittelwerte der einzelnen Altersgruppen betrachten, so sind doch relativ große Unterschiede bezüglich den Schätzungen der Methoden. Betrachten wir zuerst die NCI Methode. Egal ob das Alter oder die Altersklassen ins Modell aufgenommen werden, es wird immer ein negativer Trend für höheres Alter geschätzt. Die Ergebnisse unterscheiden sich nur wenig, daher wurde die

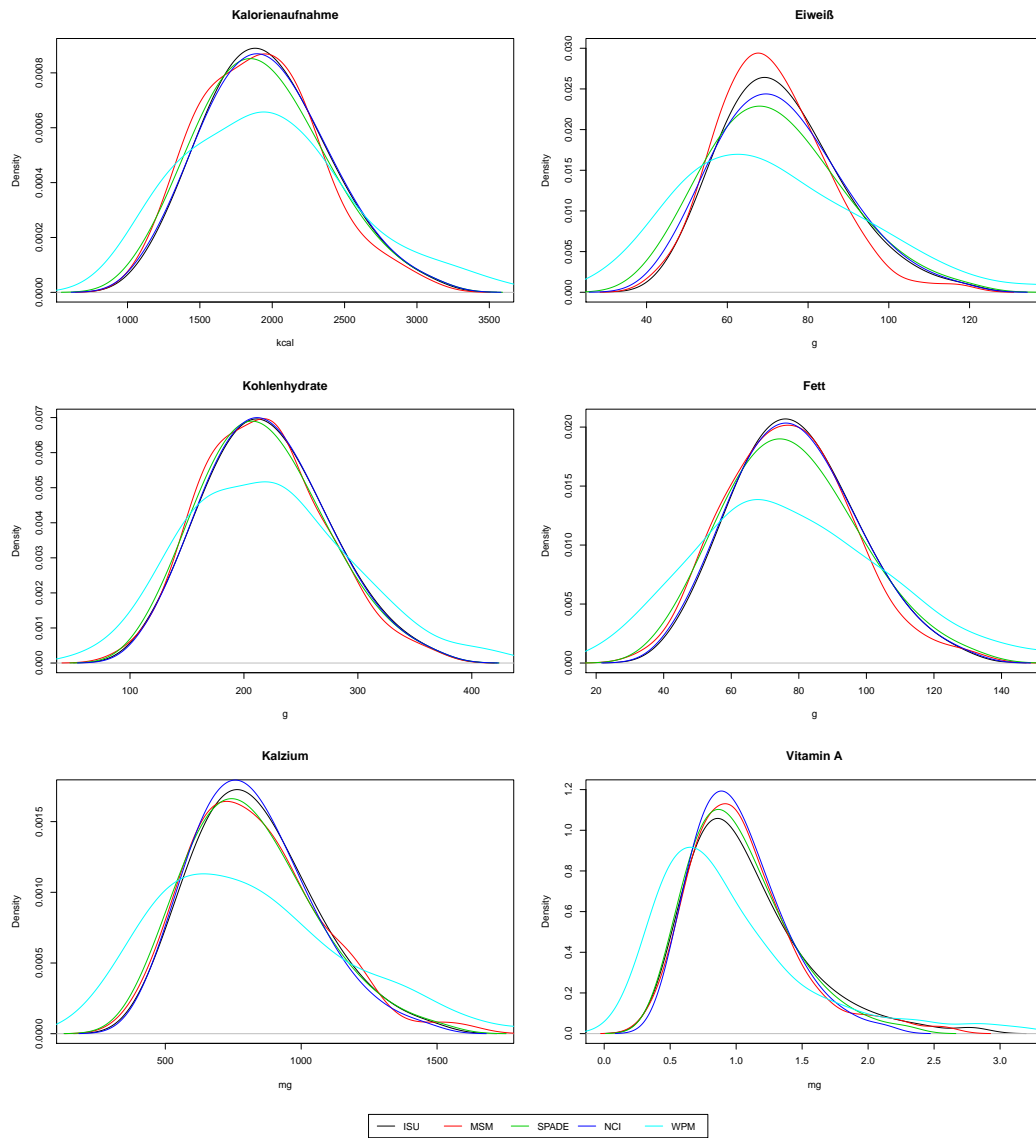


Abbildung 4.1: Vergleich der Hauptmethoden und Mittelwertmethode anhand von Dichtebildungen



	Männlich			Weiblich		
	[18,24]	[25,50]	[51,64]	[18,24]	[25,50]	[51,64]
WPM	2357	2073	2147	1880	1773	1774
SPADE	2171	2145	2112	1785	1768	1744
NCI	2176	2137	2101	1837	1790	1762
SPADE gew.	2239	2205	2161	1818	1811	1803
n	18	99	51	40	154	56

Tabelle 4.5: Vergleich der Mittelwerte bezüglich Altersgruppen und Geschlecht für die Energieaufnahme [kcal]

Modellierung mit Altersklassen verwendet. Die Effekte für Geschlecht und die Klassen wurden beide als negativ geschätzt. Für das Geschlecht macht es definitiv auch Sinn, da Männer im Schnitt mehr Kalorien aufnehmen als Frauen. Beim Alter hingegen ist es schwieriger zu sagen. Laut dem "einfachen" Mittelwert der WPM Methode sind auf jeden Fall die jungen (18-24) Leute, die mit der höchsten Aufnahme. Zwischen den anderen beiden Klassen (25-50 und 51-64) ist der Unterschied sehr gering bzw. nicht vorhanden. Ob beim Alter überhaupt ein Effekt (noch dazu ein linearer) vorhanden ist, lässt sich nur schwer aussagen. Auf jeden Fall haben die Methoden SPADE und NCI denselben Trend geschätzt (steigendes Alter - weniger Kalorien). Während bei SPADE das Alter sogar die Möglichkeit als Polynom ins Modell einzugehen, so wurde doch jedes Mal das (komplexe) Modell verworfen und auf einen einfachen linearen Effekt reduziert. Dieser scheint allerdings recht schwach zu sein und eine Modellierung ohne Alter wäre möglicherweise ebenfalls sinnvoll. Weiters zeigt sich vor allem ein großer Unterschied in der Altersklasse 25-50 bei den Männern, zwischen WPM und den anderen Methoden (trotz einer Stichprobengröße von n= 99). Dies ist vor allem auf die Restriktion eines einfachen linearen Trends des Alters, in den Methoden SPADE und NCI zurückzuführen, während die WPM Methode einen U-förmigen Trend zeigt.

Weiters zeigt sich auffällig, dass die Ergebnisse zwischen NCI und der "normalen" SPADE Methode diesmal ähnlicher sind, wobei vorher der Fall immer umgekehrt war. Die Unterschiede zwischen den 3 Methoden sind allgemein allerdings nur von kleiner Bedeutung (1-3%).

### Kohlenhydrate und Eiweiß

Die Ergebnisse für die Makronährstoffe Kohlenhydrate (Tab:A.5) und Eiweiß (Tab:A.4) sind wieder im Anhang zu finden, da die Ergebnisse ähnlich sind wie vorher. Für alle Methoden wird ein linearer Trend nach unten für steigendes Alter prognostiziert und Männer haben wieder im Schnitt eine höhere Aufnahme. Die Unterschiede zwischen den Methoden sind auch hier wieder relativ gering (1-3%). Die Vergleiche mit den einfachen Mittelwerten zeigen etwas größere Abstände als unter den Methoden. Die geringe Stichprobengröße je Alterklasse lässt die Schätzungen aber auch stärker streuen und es ist daher nur schwer zu sagen, ob wirklich signifikante Unterschiede vorliegen. Zu kritisieren bleibt auf jeden Fall, ob der lineare Effekt des Alters wirklich vorhanden ist oder ob das eine Art "overfitting" darstellt. Da der wahre Wert unbekannt ist, kann auch nicht geschlussfolgert werden, welche Methode hier am Besten abschneidet.

## 4.3 Schlussfolgerungen

Wie schon erwähnt zeigen die Ergebnisvergleiche von Simulationsstudien keine großen Unterschiede der 4 Modelle auf. Jedoch wäre ein Simulationsvergleich anhand von gewichteten Daten interessant. Beim Vergleich der 4 Methoden anhand der Ernährungsberichtsdaten konnten ebenfalls, ähnlich zur Simulationsstudie von Souverein et al. (2011), keine großen Unterschiede der Ergebnisse festgestellt werden. Einzig die "normale" SPADE Methode für Daten mit Gewichtungsfaktoren zeigt Auffälligkeiten bei den Schätzungen. Diese können aber mit einem gleichgewichteten Datensatz vermindert werden. Die Wahl der Methodik kann also, wie schon im 2. Kapitel gefolgert, eher nach beliebigen Kriterien (wie z.B. Softwarekenntnisse etc.) erfolgen. Da für den Ernährungsbericht öfter mal die Vergleiche für die Altersgruppen interessant sind, kommen eher die Methoden SPADE und NCI in Frage.

Da auch hier keine großen Unterschiede festgestellt werden konnten, ist es auch hier wenig entscheidend welche Methode verwendet wird. Für die Neuauswertung der einzelnen wird daher die Methode SPADE im R verwendet. Da die Methode mit den gleichgewichteten Daten "bessere" Ergebnisse als die normale SPADE Anwendung mit Gewichtungsfaktoren brachte, wird auch für die nächsten Auswertungen dieser Ansatz gewählt. Zusätzlich soll bei Nährstoffen, welche sehr große Ausreißer oder sonstige Auffälligkeiten zeigen, die komplexe ISU Methode verwendet werden. Dadurch besteht die Hoffnung, dass im Falle des Scheiterns der "einfachen" Box-Cox Transformation die komplexere 2-Schritt Transformation der ISU Methode bessere Ergebnisse bringt.

## Kapitel 5

# Ernährungsbericht Neuauswertung

In diesem Kapitel wird die Neuauswertung der Aufnahmedaten der Erwachsenen vom Ernährungsbericht (EB) 2012 beschrieben. Das Ziel der Neuauswertung ist zum einen die Unterschiede, die in den Anteilsschätzungen entstehen, wenn bessere Methoden als die gewöhnliche Mittelwertmethodik (WPM) für die Aufnahmeverteilung verwendet werden, aufzuzeigen. Im Ernährungsbericht von 2012 wurde zur Schätzung der Aufnahmeverteilung die empirische Verteilung der Mittelwerte der 24 Stunden Erinnerungsprotokolle verwendet. Ebenso wurden die DACH “Empfehlungen“ (Referenzwerte) als cut-points verwendet. Besser wäre die Verwendung des EAR Referenzwertes (siehe Kapitel 2 Abschnitt 2.6.1), weswegen hier auch als zweites Ziel die Unterschiede in den Anteilsschätzungen zwischen DACH und EAR Werten aufgezeigt werden soll.

Für die Neuauswertung wurden die Vitamine, sowie ein paar Mineralstoffe (für die es einen EAR Wert gibt) ausgewählt. Ebenso wurden die Makronährstoffe und die Energieaufnahme neu ausgewertet.

Für die Neuauswertung wurden die Methoden SPADE und ISU (bei Scheitern der Box-Cox Transformation) verwendet (Begründung: siehe Kapitel 4.3). Die Vergleiche der Ergebnisse sollen somit vor allem auch die Unterschiede zwischen der Mittelwertmethodik und der SPADE Methode aufzeigen. Um im weiteren Verlauf die Lesbarkeit der Arbeit hoch zu halten, wird nicht jedes Mal extra auf den Ernährungsbericht 2012 verwiesen, wenn Ergebnisse daraus verwendet wurden.

Im Ernährungsbericht wurden die Daten voranalysiert um Misreporter festzustellen. Diese wurden für die Auswertungen nicht verwendet. Im Kapitel 6 wird darauf noch weiter eingegangen. An dieser Stelle soll nur erwähnt werden, dass für die Neuauswertung ebenfalls nur mit dem aussortierten Sample ( $n=380$ ) gearbeitet wurde. Dadurch wird der Unterschied in den Ergebnissen ausschließlich durch die Methodik verursacht. Die Daten wurden zusätzlich noch, mittels der im Kapitel 4.2 beschriebenen Methodik, in einen neuen Datensatz umgewandelt, welcher sich verhält wie die gewichteten Daten, aber ohne die Gewichtungsfaktoren verwendet werden kann.

Ein weiterer Schritt, um die Lesbarkeit hoch zu halten, ist nicht alle Grafiken und Tabellen direkt im dafür vorgesehenen Abschnitt zu präsentieren, sondern erst im Anhang. Die Ergebnisse werden trotzdem, sofern relevant, beschrieben und diskutiert. Ebenso wird, so gut es geht, ein einheitliches Schema für jeden Nährstoff eingehalten. Zuerst werden die Annahmen des Modells und dann die relevanten Auswertungen diskutiert. Die Ergebnisse werden anschließend präsentiert und zuletzt die Vergleiche mit den Ergebnissen vom Ernährungsbericht 2012 durchgeführt und kommentiert.

### Annahmenüberprüfung

Um die Annahmen 1-3 (Transformierbarkeit, Varianzhomogenität, Unabhängigkeit) überprüfen zu können, werden unterschiedliche “Tests“ verwendet. Da SPADE eine einfache Box-Cox Transformation vor der Analyse durchführt, werden die Daten auch mit der Funktion “powerTransform“ (einfache Box-Cox Transformation) aus dem R-Package “car“ transformiert, bevor die Tests durchgeführt werden. Das SPADE für  $\lambda$  Werte kleiner Null automatisch die Log-Transformation durchführt, wird auch hier berücksichtigt.

Für die Überprüfung der **Normalität** wird auf die transformierten Daten der Shapiro-Wilk und auch der Anderson-Darling Test angewandt. Der Shapiro-Wilk Test (SW-Test) reagiert dabei sensitiver auf Ausreißer (Werte in den Enden der Verteilung) als der Anderson-Darling Test (AD-Test) und führt dadurch schneller zu kleinen p-Werten, weshalb der AD-Test auch betrachtet wird. Zusätzlich wird grafisch ein Histogrammvergleich mit der Normalverteilungsdichte durchgeführt.

Die **Unabhängigkeit** des Zufallseffekt  $\mu_i$  und des Fehlerterms  $\epsilon_{ij}$  wird mittels linearer Regression überprüft. Dazu werden für alle Personen mit 2 Beobachtungen der Mittelwert (als Regressor) und die geschätzte Standardabweichung (als Regressand) verwendet. Beachte, dass dies ebenfalls auf die transformierten Daten angewandt wird. Der t-Test der geschätzten Steigung wird dabei als Richtwert für die Überprüfung der Annahme verwendet. Es wird also überprüft, ob der Effekt - höhere mittlere Aufnahme führt zu höherer within-person Varianz - groß genug ist und ob Grund zur Verletzung der Annahme vorliegt.

Die **Varianzhomogenitätsüberprüfung** stellt die größte Herausforderung dar. Das Problem, welches hier auftritt, ist die hohe Anzahl an "Gruppen" (= Personen mit zwei Beobachtungen (276)) und die sehr geringe Anzahl an Werten je Gruppe für die Schätzung der Varianz ( $m = 2$ ). Die gängigen Varianztests haben daher nur sehr schwache Power bzw. ist oft das Signifikanzniveau verletzt. Ebenso sind nicht mal alle Tests anwendbar (z.B. Levene Test), da sie mehr als 2 Beobachtungen pro Gruppe benötigen. Der Test, welcher in der ISU Methode angewandt wird, ist sehr spezifisch festgelegt worden und testet ob das beobachtete vierte Moment der (mit der 2 stufigen Transformation) transformierten Daten gleich dem einer Standardnormalverteilung ist. Dieser Test kann auch nicht auf die Box-Cox transformierten Daten angewandt werden, da diese nicht in eine Standardnormalverteilung transformiert werden. Als Lösung für das Problem wird daher nur grafisch überprüft, ob die Annahme verletzt sein könnte. Dabei kann ausgenutzt werden, dass wenn die Daten normalverteilt sind und gleiche within-person Varianz vorliegt, die einzelnen Varianzschätzungen je Person gammaverteilt sind ( $\hat{\sigma}_{w_i}^2 \sim \sigma_w^2 \chi^2 \sim \text{Gamma}(\text{shape} = 1/2, \text{scale} = 2\sigma_w^2)$ ). Mit diesem Wissen können leicht zwei Überprüfungsplots erstellt werden. Einerseits kann ein QQ-Plot, wo die Quantile der beobachteten Varianzen gegen die Quantile der theoretischen Gamma Verteilung abgebildet werden. Andererseits können die beobachteten Varianzen in die Verteilungsfunktion der Gamma Verteilung eingesetzt werden und überprüft werden, ob eine Gleichverteilung vorliegt ( $F_{\text{gamma}}(\sigma_w^2 \chi^2) \sim U[0, 1]$ ). Für den Scale Parameter der Gamma Verteilung wird die Schätzung von Gl:2.9 für den unbekannt Parameter  $\sigma_w^2$  eingesetzt. Beide grafischen Überprüfungen werden für die Nährstoffauswertungen verwendet.

Zusätzlich zu diesen Überprüfungen wurden auch immer die Residuenplots (QQ-Plots bezüglich Normalverteilung) der SPADE Methodik auf Annahmeverletzungen hin kontrolliert.

## Anmerkungen zu SPADE

SPADE ist ein Programmpackage, welches eigentlich nur für Datensätze mit 2 Beobachtungen je Person vorgesehen ist. Sollten nicht balancierte Daten vorliegen, so kann es zu Fehlermeldungen kommen. Betrachtet man SPADE Schritt für Schritt, so kann man erkennen, dass es an dem Rücktransformationsschritt liegt. Der genaue Grund, warum nur exakt 2 Beobachtungen zugelassen werden, wird nicht beschrieben. Deswegen ist mit so einem Datensatz, wie im Fall des Ernährungsberichtes 2012, Vorsicht geboten. Doch solange gewisse Schritte<sup>1</sup> beachtet werden, auf die hier nicht näher eingegangen wird, stellt dies für die Neuauswertung kein Problem dar. Um diese Aussage weiter zu stützen, betrachten wir kurz die einzelnen Schritte der SPADE Funktion für das 1-Part Modell. Als erstes wird eine Box-Cox Transformation auf die Daten angewandt. Hier ist es egal, wie viele Beobachtungen pro Person vorliegen. Im nächsten Schritt wird das fraktionale Polynom geschätzt bzw. getestet. Dabei spielen die Beobachtungen pro Person auch keine Rolle, da hierfür die R Funktion "lm" im einfachen Fall oder "mfp" im anderen verwendet wird. Im Varianzschätzungsschritt wird die Funktion "lmer" aus dem Package "lme4" verwendet. Auch diese benötigt nicht für jede Person 2 Beobachtungen, was auch im Output kontrolliert werden kann. Für den letzten Schritt wird eine eigene Funktion für die Gaussquadratur verwendet. Auch diese benötigt nicht 2 Beobachtungen, da nur die einzelnen Quantile der geschätzten Normalver-

<sup>1</sup>nur Verwendung des 1-Part Modells mit Argument backtransform=1, sowie gerade Anzahl an Beobachtungen ist notwendig, damit keine Fehlermeldungen zustande kommen.

teilung rücktransformiert werden. Die Ergebnisse können auch kontrolliert werden, wenn diese Schritte selbst durchgeführt werden. Für die nachkommenden Auswertungen ist die Anwendung des SPADE Package also kein Problem. Für andere Analysen z.B. andere Rücktransformationen als die Gaussquadratur oder das 2-Part Modell (welche in dieser Arbeit nicht vorkommen) muss vorsichtig vorgegangen werden.

**Bootstrap Konfidenzintervalle (KI)** Für die Unsicherheitsabschätzungen wurde die Bootstrap Variante von SPADE gewählt. Für die meisten Nährstoffe wurden 1000 Bootstrap Ziehungen gewählt, da diese doch eine relativ lange Zeit benötigen. Für manche Nährstoffe, wo die Bootstrap KIs schlechte Werte lieferten, wurden dann 10.000 Ziehungen verwendet. Sofern im weiteren Textverlauf nicht extra angegeben wird wie viele Bootstrap Ziehungen verwendet wurden, ist immer von 1000 Ziehen die Rede.

## 5.1 Energie und Makronährstoffe

### 5.1.1 Energie

Als erstes betrachten wir die Energieaufnahme der Personen.

**Annahmen** Ein Blick auf die Energieaufnahmedaten zeigt keine Ausreißer oder Auffälligkeiten. Dies bestätigt sich in den Überprüfungen. Die Transformation führt zu guter approximativer Normalität und auch die Normalverteilungstests zeigen sehr hohe p-Werte ( $p > 0.6$ ). Die Überprüfung der Korrelation zeigt einen leichten Abfall der Standardabweichungen für höhere Mittelwerte, aber auch kein signifikantes Ergebnis ( $p > 0.2$ ). Ebenso zeigen die Varianzplots keine Annahmeverletzungen.

**Auswertung** Betrachtet man das Alter und die Energieaufnahme gemeinsam, so kann kein wirklicher Zusammenhang festgestellt werden. Die Modellierung mit dem Alter als Kovariate wird nur zu einem sehr schwachen Effekt führen und ist möglicherweise Overfitting. Für die Schätzung der üblichen Aufnahmeverteilung wird das Alter daher nicht berücksichtigt. Für die Mittelwertvergleiche der Altersklassen mit den Ergebnissen des Ernährungsberichtes wird die Kovariate allerdings wieder mit ins Modell aufgenommen, um die Unterschiede besser aufzeigen zu können.

Weiters interessant ist der Vergleich der Aufnahmeverteilungen bezüglich den verschiedenen BMI Klassen. Diese teilen sich in Untergewicht ( $BMI < 18.5$ ), Normalgewicht ( $18.5 < BMI < 25$ ), Übergewicht ( $25 < BMI < 30$ ) und Adipös ( $BMI > 30$ ). Da aber einerseits nur gemessene BMI Werte für 280 Personen vorliegen und für die Klassen Untergewicht ( $n = 7$ ) und Adipös ( $n = 27$ ) die Stichprobengrößen sehr klein sind, wird hier nur ein Vergleich zwischen der Aufnahme von normalgewichtigen ( $n = 181$ ) und übergewichtigen+adipösen ( $n = 92$ ) Personen gegeben, wobei hier nochmal in männlich und weiblich geteilt wird.

**Ergebnisse** Für den Vergleich von Männer und Frauen ist als erstes auffällig, dass Männer eine höhere Aufnahme haben. Weiterhin auffällig ist, dass Männer ein deutlich niedrigeres Varianzverhältnis ( $= 3.3$ ) als Frauen ( $= 4.5$ ) aufweisen. Tabelle 5.1 zeigt die Ergebnisse der üblichen Aufnahmeverteilungen. Es wird klar, dass es sich um relativ niedrige Werte handelt, wenn diese mit den Richtwerten (DACH) zur Energiezufuhr verglichen werden. Die DACH Richtwerte stellen hier Werte für die durchschnittliche Energiezufuhr der Personen dar. Bei geringer körperlicher Aktivität sind diese Richtwerte für Männer um die 2300 kcal und für Frauen um die 1800 kcal. Diese stimmen recht gut mit den geschätzten Mittelwerten überein.

Wirklich interessant wird es, wenn die Aufnahmeschätzungen der Normal- und Übergewichtigen betrachtet werden. Sowohl für Männer als auch für Frauen ist der Mittelwert bei den Übergewichtigen minimal kleiner als bei den Normalgewichtigen. Auch die Quantile deuten darauf hin, dass hier falsche Werte protokolliert wurden. Dieser Sachverhalt wird im 6. Kapitel noch weiter verfolgt. Dies könnte ein Zeichen von verzerrten Daten sein, was somit eine Annahmeverletzung darstellt und klarerweise zu schlechten Schätzungen führt.

	Mittel	5%	25%	50%	75%	95%
Frauen	1872	1483	1697	1859	2033	2305
Männer	2317	1792	2072	2291	2534	2928

	Mittel	5%	25%	50%	75%	95%
Männer Normalgew.	2376	1710	2061	2340	2652	3163
Männer Übergew.	2341	1912	2146	2325	2519	2826
Frauen Normalgew.	1863	1504	1703	1852	2012	2260
Frauen Übergew.	1768	1274	1530	1737	1972	2366

Tabelle 5.1: Vergleich der Kalorienaufnahme [kcal] von Männer (n=148) und Frauen (n=232) sowie von Normal (n=181) - und Übergewichtigen (+ Adipösen) (n=92).

**Vergleiche** Im Ernährungsbericht wurden hauptsächlich die Mittelwerte der verschiedenen Altersklassen betrachtet und sowohl für Männer als auch für Frauen verglichen. Dazu wurden Konfidenzintervalle, die auf theoretischen Annahmen beruhen angegeben. Diese Ergebnisse sollen hier mit den Mittelwertschätzungen der SPADE Methode, sowie mit den Bootstrap Konfidenzintervallen verglichen werden. Tabelle 5.2 zeigt die Ergebnisse. Die Bootstrap KI sind für alle Altersgruppen etwas schmaler als die des EB, trotzdem überdecken sich im Vergleich alle. Vergleicht man den Trend bezüglich den Altersklassen, so zeigt sich ein Unterschied zwischen SPADE und dem EB. Während bei SPADE für Männer als auch Frauen ein leichter Aufwärtstrend erkennbar ist, so ist beim EB eher ein Abwärtstrend erkennbar. Da sich die KI auch über die Altersklassen hinweg überdecken, lässt sich kaum aussagen, ob wirklich der eine oder andere Trend tatsächlich vorhanden ist. Vor allem die sehr geringe Stichprobengröße bei der Altersklasse (18-24) lässt nur wenig Aussagekraft zu. Welche Schätzungen als genauer angesehen werden können, lässt sich nicht aussagen. Es besteht auf jeden Fall kein großer Unterschied zwischen SPADE und dem EB.

	[18-24]	[25-50]	[51-64]
SPADE Männer	2311 [2142,2645]	2316 [2195,2422]	2323 [2201,2510]
EB Männer	2403 [2112,2693]	2172 [2066,2277]	2245 [2078,2412]
n	17	87	44
SPADE Frauen	1861 [1756,1953]	1871 [1810,1960]	1885 [1766,1973]
EB Frauen	1917 [1746,2088]	1854 [1761,1948]	1826 [1704,1948]
n	37	143	52

Tabelle 5.2: Energieaufnahmevergleiche [kcal] zwischen Ernährungsbericht und SPADE bezüglich den Alterklassen

### 5.1.2 Kohlenhydrate

Betrachtet man die Kohlenhydrataufnahme, so zeigen sich ebenfalls keine großen Unterschiede.

**Annahmen** Auch die Kohlenhydratdaten lassen sich gut zu Normalität transformieren. Der SW-Test, sowie der AD-Test zeigen keine Hinweise für eine Verletzung der Annahme. Auch der Korrelationscheck und die Varianzannahmeüberprüfung zeigen keine Auffälligkeiten. Die SPADE Methode kann angewandt werden.

**Auswertung** Das Alter zeigt auch hier keinen wirklichen Trend für die Aufnahme. Eine Modellierung mit dem Alter ist daher wenig sinnvoll. Bei den Makronährstoffen ist auch immer von Interesse, wie sich das Verhältnis der Energie aus diesen zu der Gesamtenergie ergibt. Hier gibt es zwei Ansätze zur Modellierung. Einerseits kann für jede Tagesbeobachtung das Verhältnis von Makronährstoffenergie zu Gesamtenergieaufnahme als Variable für die Methoden verwendet werden. Andererseits kann aber

auch die Energieaufnahme aus den Makronährstoffen und die Gesamtenergie als bivariate Zufallsvariable betrachtet werden. Diese Methodik wird allerdings nur von den Methoden NCI und ISU unterstützt. Welcher Ansatz der bessere ist kann nicht so einfach beantwortet werden, da dies eher eine Frage der Interpretation ist. Während der erste Ansatz interessanter wäre für die Fragestellung, wo ein bestimmtes Verhältnis von Makronährstoff zu Gesamtenergie für jeden Tag eingehalten werden soll, so wäre der zweite Ansatz interessanter, wenn ein bestimmtes Verhältnis über eine lange Periode eingehalten werden soll. Für diese Arbeit wurde nur der erste Ansatz verwendet. Für eine Verwendung des zweiten Ansatzes wird auf Freedman et al. (2010) für die NCI Methode und Carriquiry et al. (1995) für ISU verwiesen. Für NCI gibt es auch hierzu ein extra SAS Makro. Bei ISU ist der Ansatz im Programm PC-SIDE verankert und einfach durchzuführen.

**Ergebnisse** Betrachtet man das Varianzverhältnis für Männer und Frauen für die “einfache“ Kohlenhydrataufnahme, so weisen diese keine großen Unterschiede auf (Männer 1.9, Frauen 2.3). Für die Modellierung als Energieprozent kommen jedoch starke Abweichungen zustande (Männer 3.7, Frauen 11.5). Interessant ist auch, dass für Männer die Modellierung mit dem Alter als Kovariate (im Energieprozentmodell) ein leicht signifikantes Ergebnis bringt, während für Frauen kaum ein Effekt erkennbar ist. Die Modellresiduenplots weisen keine Auffälligkeiten einer Annahmenverletzung vor.

**Vergleiche** Im Ernährungsbericht wurden nur die Energieprozent der Kohlenhydrate als Ergebnis betrachtet. Daher wird auch hier nur der Vergleich anhand dieser getätigt. In Tabelle 5.3 werden die Unterschiede wiederum für Männer und Frauen separat betrachtet. Es sind wieder nur sehr geringe Unterschiede der Schätzungen zu finden. Die Konfidenzintervalle sind sich sehr ähnlich und auch relativ gleich groß (zumindest auf diese Genauigkeit betrachtet). Ob SPADE oder normale Mittelwertschätzung macht hier also keinen wirklichen Unterschied.

	[18-24]	[25-50]	[51-64]
SPADE Männer	46 [44,50]	45 [43,46]	43 [41,45]
EB Männer	47 [44,49]	44 [42,45]	42 [40,44]
n	17	87	44
SPADE Frauen	48 [47,51]	48 [46,49]	48 [46,51]
EB Frauen	47 [44,49]	47 [45,48]	48 [45,50]
n	37	143	52

Tabelle 5.3: Energieprozentvergleiche [%] für Kohlenhydrate zwischen Ernährungsbericht und SPADE bezüglich den Altersklassen

### 5.1.3 Eiweiß

Die übliche Aufnahmeschätzung zu Eiweiß stellt die größte Problemstellung der Makronährstoffe dar. Die Daten zeigen zwei größere Ausreißer ( $>290g$ ). Es handelt sich hierbei um zwei männliche Personen, welche jeweils zwei beobachtete Tagesaufnahmen vorweisen. Bei beiden ist an einem der beiden Tage dieser hohe Wert zustande gekommen. Einmal durch 1.7kg Brathähnchen und einmal durch viele eiweißreiche Quellen wie Pute, Schnitzel und auch Eiweißpulver. Ob es sich bei der einen Person mit 1.7kg Huhn um übliche Aufnahmemengen handelt, mag bezweifelt werden. Für die Modellierung stellt dies auf jeden Fall ein Problem dar, da dadurch schlechte Schätzungen zustande kommen können.

**Annahmen** Betrachtet man für den ganzen Datensatz die Normalverteilungsannahme so verwirft der SW-Test ganz klar die Annahme ( $p < 0.0004$ ). Dass dieser Test aber sensitiver auf Ausreißer reagiert, zeigt das AD-Testergebnis ( $p = 0.14$ ). Der Blick auf den Histogrammvergleich zeigt ebenfalls, dass die Normalverteilung recht gut approximiert wird, wenn die hohen Werte an den Enden der Verteilung nicht berücksichtigt werden. Die Unabhängigkeits-

und Varianzannahme scheinen hier weniger ein Problem darzustellen. Werden die zwei Ausreißer auf den niedrigeren Wert (Median+2\*IQR) gesetzt, so verbessern sich die NV-Tests deutlich. Der geschätzte Box-Cox Parameter wäre im extremen Fall -0.06 (was zu einer Logtransformation in SPADE führt) und 0.08 im anderen. Es muss für die weitere Analyse auf jeden Fall verglichen werden, wie sehr sich die Ergebnisse mit und ohne die Ausreißer verhalten.

**Auswertung** Das Alter zeigt auch hier wieder keinen wirklichen Trend für die Aufnahme. Für die Energieprozentvergleiche wird jedoch wieder das Alter ins Modell genommen. Für die anderen Vergleiche hingegen nicht. Für die Auswertungen bezüglich Eiweiß ist meist auch das Verhältnis von Aufnahme zu Körpergewicht, da darüber die Empfehlungen definiert sind, interessant. Daher wird auch dies modelliert und mittels SPADE eine Einschätzung über den Anteil der Personen die weniger als die Empfehlung aufnehmen abgegeben.

**Ergebnisse** Betrachten wir die Ergebnisse wieder separat für Männer und Frauen, so fällt der große Unterschied im Varianzverhältnis auf (Tab:5.4). Die Eiweißaufnahme streut (im Verhältnis gesehen) innerhalb der Frauen deutlich mehr als bei den Männern. Die Residuenplots zeigen auf jeden Fall keine Verletzung der Normalitätsannahme. Männer weisen wiederum, wenig verwunderlich, eine höhere Aufnahme als Frauen auf. Betrachtet man die Unterschiede der Schätzungen zwischen dem Modell mit den Ausreißern und den niedriger gesetzten Werten, so können nur geringe Unterschiede zwischen den Aufnahmeschätzungen festgestellt werden. Der größte Unterschied ist beim 95% Quantil zu erkennen und liegt bei 3.6%. Es lässt sich darüber diskutieren, ob diese Unterschiede groß genug sind um eine Modellierung ohne den Ausreißern vorzunehmen. Für die weiteren Ergebnisse wurden allerdings die "originalen" Daten verwendet.

	Mittel	5%	25%	50%	75%	95%	Var.ratio
Frauen	67.0	58.3	63.2	66.9	70.6	76.3	19.7
Männer mit Ausreißer	88.7	68.1	79.1	87.7	97.3	113.0	6.1
Männer ohne Ausreißer	87.2	68.4	78.5	86.4	95.0	109.0	6.5
relative Differenz [%]	1.7	-0.5	0.7	1.5	2.4	3.6	-6.1

Tabelle 5.4: Vergleich der Eiweißaufnahme [g] von Männer (mit und ohne Ausreißer, n=148) und Frauen (n=232). Für Männer wird zusätzlich der relative Unterschied angegeben.

**Vergleiche** Bei den Energieprozentvergleichen zeigt sich, dass auch hier zwischen der Mittelwertmethode und SPADE kaum ein Unterschied vorliegt (Tab:5.5). Bei beiden Geschlechtern ist kein wirklicher Trend bezüglich des Alters erkennbar. Im Ernährungsbericht wurde auch die mittlere Eiweißaufnahme pro kg Körpergewicht pro Tag berechnet und mit der DACH Empfehlung verglichen. Bei den Frauen und Männern waren es 1.1g/kg und somit über der Empfehlung von die 0.8g/kg. Die mittlere Aufnahme allein sagt allerdings nicht aus, ob es noch viele Personen betrifft, die vielleicht zu wenig aufnehmen oder nicht. Daher wurde mit SPADE eine Abschätzung des Anteils der Personen, die zu wenig Eiweiß aufnehmen ermittelt. Dazu wurde für alle Personen, bei denen das Körpergewicht gemessen wurde, der Quotient zwischen der Tageseiweißaufnahme und dem Körpergewicht berechnet. Für alle Personen, wo keine Messung, sondern nur das angegebene Körpergewicht bekannt war, wurde damit der Quotient ermittelt. Für eine Person (weiblich) war weder eine Messung noch eine Angabe bekannt. Für diese Person wurde das mittlere Körpergewicht aller Frauen für den Quotient verwendet. Die Abschätzung von SPADE für Leute, die unter dem DACH Referenzwert liegen, beträgt 5.3% (KI: [0.2%,10.2%]). Da dies aber, wie in Kapitel 2 Abschnitt 2.6.1 erklärt, eine Überschätzung des tatsächlichen Anteils darstellt, wurde auch der EAR Referenzwert (= 0.66g/kg) für die Ermittlung des Anteils verwendet. Die Schätzung mit dem EAR ergab einen wesentlich kleineren Anteil von 0.3% [0%,1.5%]. Daher wäre die Schlussfolgerung, dass kaum eine Proteinunterversorgung vorliegt, gerechtfertigt (Annahmen - abgesehen von Datenverzerrung - wurden auch überprüft und sind sehr wahrscheinlich nicht verletzt). Hier kann übrigens auch der Probability Approach angewandt werden. Wie



in National Research Council (1986) nehmen wir hier an, dass der Bedarf einer Normalverteilung mit Erwartungswert 0.66 und Variationskoeffizient 15% folgt. Die Schätzung des Ansatzes würde mit 0.14 % noch ein wenig niedriger ausfallen als bei dem EAR cut-point Ansatzes. Für die Aussage über den Zustand der Unterversorgung würde es allerdings keinen Unterschied machen.

	[18-24]	[25-50]	[51-64]
SPADE Männer	16 [15,17]	16 [15,17]	16 [15,17]
EB Männer	15 [14,17]	15 [14,16]	15 [14,16]
n	17	87	44
SPADE Frauen	15 [14,16]	15 [15,16]	14 [14,15]
EB Frauen	15 [13,16]	15 [14,15]	14 [13,15]
n	37	143	52

Tabelle 5.5: Energieprozentvergleiche [%] für Eiweiß zwischen Ernährungsbericht und SPADE bezüglich den Alterklassen

### 5.1.4 Fett

Den letzten untersuchten Makronährstoff stellt Fett dar. Dieser verhält sich allerdings wieder gut bezüglich den Annahmen und zeigt keine Auffälligkeiten.

**Annahmen** Die Annahmenüberprüfung zeigt nur wenig Bedenken auf. Die Normalitätstests liefern p-Werte größer als 0.18 und die anderen beiden Annahmen liefern ein noch besseres Bild.

**Auswertung** Das Alter scheint auch hier nur einen sehr kleinen Effekt auf die Fettaufnahme zu haben. Bei den Männern ist der Effekt etwas stärker, aber nicht groß genug um eine Modellierung damit zu rechtfertigen. Die Energieprozentaufnahme von Fett wurde trotzdem wieder mittels Alter modelliert um den Vergleich zum EB darstellen zu können.

**Ergebnisse und Vergleiche** Die Ergebnisse und Vergleiche zeigen wiederum sehr ähnliche Ergebnisse wie erwartet. Frauen haben eine geringere Fettaufnahme und wieder eine weitaus höhere Varianzratio (Frauen 13.7, Männer 5.1). Die Vergleiche mit den EB Ergebnissen ergeben auch wiederum sehr ähnliche Resultate, weshalb die Tabelle im Anhang zu finden ist (Tab:A.6). Einzige Auffälligkeit ist, dass SPADE bei den Männern eine leicht höhere Energieprozentaufnahme schätzt, aber keine nicht überschneidenden Konfidenzintervalle.

## 5.2 Vitamine

In diesem Abschnitt werden die Vitamine (nach Alphabet sortiert) betrachtet.

### 5.2.1 Vitamin A

Als erstes betrachten wir Vitamin A. Dieser Nährstoff erweist sich wiederum als kritisch, im Sinne der Modellierung. Ein erster Blick auf die Daten zeigt, dass ein extremer Ausreißer vorliegt (39.5 mg, Median+2\*IQR=2.8mg). Der Wert kommt bei einer weiblichen Person zustande, da sie an dem Tag gebackene Leber gegessen hat, was einen sehr hohen Vitamin A Gehalt aufweist. wie wir noch sehen werden, kommt diese Person noch häufiger als Problemfall vor.

**Annahmen** Überprüft man die Annahmen mit dem starken Ausreißer, so wären alle 3 Annahmen problematisch, bzw. würden auch die Ergebnisse verzerrt werden. Daher wird dieser Wert im weiteren Verlauf wiederum auf die Schranke (Median+2\*IQR) runter gesetzt. Die Annahmen der Varianzhomogenität und Unabhängigkeit scheinen dadurch weniger ein Problem darzustellen. Die Normalitätsannahme hingegen scheint immer noch, sowohl grafisch als auch von den Testwerten her ( $p < 0.007$ ) problematisch zu sein. Die nachfolgenden Ergebnisse müssen also mit mehr Vorsicht betrachtet werden.

**Auswertung** Im EB wurde, dadurch die mittlere Vitamin A Aufnahme für alle Altersgruppen (knapp) über den DACH Referenzwerten lag, geschlussfolgert, dass die Aufnahme zufriedenstellend ist. Da die Daten eine Modellierung mit dem Alter als Kovariate als wenig sinnvoll erscheinen lassen, wird hier nur allgemein für Männer und Frauen eine Schätzung über den Anteil der zu geringen Aufnahme modelliert.

**Ergebnisse und Vergleiche** Die Ergebnisse der Anteilsschätzung bringen recht unterschiedliche Ergebnisse, je nachdem welchen Referenzwert man verwenden möchte. Bei den Frauen und dem DACH Referenzwert (=  $0.8mg$ ) liefert die Schätzung einen Prozentsatz von 28.8%. Wird hingegen der EAR Wert (=  $0.5mg$ ) verwendet, so kommt eine deutlich niedrigere Schätzung von 1.8% zustande. Der QQ-Plot der Residuen zeigt keine Verletzung der Normalitätsannahme, was das Ergebnis sicherer macht. Bei den Männern sieht es mit dem DACH Wert (=  $1mg$ ) schlechter aus mit 46.8%. Mit dem EAR Wert (=  $0.6mg$ ) liegt die Schätzung hingegen bei nur noch 4.3%. Die Ergebnisse sind in Tabelle 5.12 am Ende der Vitaminauswertungen zu finden. Das Varianzverhältnis der Männer liegt mit 4.8 recht gleich auf mit dem der Frauen (= 4.3). Die Bootstrap Konfidenzintervalle liefern auch bei 10.000 Wiederholungen noch Verzerrungen bezüglich den EAR Werten, weshalb diese hier nicht angegeben werden. Für die DACH Referenzwerte ist das Konfidenzintervall (10.000 Wiederholungen) für Männer [20.6,62.3]. Für Frauen liegen auch bei den DACH Werten auch bei 10.000 Ziehungen noch Verzerrungen vor.

### 5.2.2 Vitamin B1

**Annahmen** Die Annahmeüberprüfungen zeigen keine Probleme. Die Normalverteilungstests liefern p-Werte größer als 0.4, der Korrelationstest einen p-Wert größer 0.2 und die Varianzüberprüfung lässt keine Auffälligkeiten erkennen.

**Auswertung** Eine Modellierung mit dem Alter scheint wiederum als wenig sinnvoll, wenn die Daten nach Geschlecht betrachtet werden. Die Mittelwerte liegen über den DACH Referenzwerten, jedoch muss wieder die Streuung mit einbezogen werden.

**Ergebnisse und Vergleiche** Die Schätzungen mit den DACH Referenzwerten liefern wieder deutlich höhere Werte für Männer (31.3%) und Frauen (31.3%) als der EAR Wert (Männer 7.3%; Frauen 12.4%). Zu beachten ist, dass bei den Männern für die verschiedenen Altersgruppen verschiedene DACH Referenzwerte vorhanden sind und die Schätzung eine Mittelung darstellt. Die Bootstrap Konfidenzintervalle weisen wieder sehr hohe Schwankungen und Verzerrungen (auch mit 10.000 Ziehungen) auf, was auch möglicherweise an den sehr hohen Varianzverhältnissen (Männer 5, Frauen 7.4) liegt und wurden deshalb hier nicht angegeben.

### 5.2.3 Vitamin B2

Vitamin B2 zeigt einen Ausreißer bei den Frauen und einen bei den Männern. Bei der Frau handelt es sich wieder um die Person, welche die gebackene Leber gegessen hat, wodurch der Ausreißer zustande kommt. Bei dem Mann liegt es vor allem an einem Eiweißpulver, welches mit Vitaminen angereichert ist.

**Annahmen** Die Annahmentests liefern trotz der "Ausreißer" gute Werte. Einzig die Korrelationsüberprüfung liefert mit den hohen Werten einen p-Wert von 0.04 (ohne Ausreißer 0.18).

**Auswertung** Eine Modellierung mit dem Alter scheint wiederum als wenig sinnvoll, wenn die Daten nach Geschlecht betrachtet werden. Die Mittelwerte liegen wieder leicht über den DACH Referenzwerten.

**Ergebnisse und Vergleiche** Die Anteilsschätzungen liefern für Männer 46.5% [31.6,58.0] (DACH Wert) und 13.2% [1.1,22.0] (EAR Wert) und für Frauen 31.7% [17.4,40.2] und 7.7% [0.5,13.8]. Das Varianzverhältnis liegt bei den Frauen wieder höher mit 3.4 als bei den Männern (=3). Die Residuen QQ-Plots zeigen (bis auf den Ausreißer bei den Männern) keine Auffälligkeiten. Die Bootstrap Konfidenzintervalle zeigen bei den EAR Werten und den DACH Werten bei den Frauen (da sie so niedrig liegen) leichte Verzerrungen und sind daher mit Vorsicht zu genießen.

#### 5.2.4 Vitamin B3

Vitamin B3 weist ein bzw. zwei große Wert bei den Männern auf. Der höchste Wert ist durch die Person zustande gekommen, welche die 1.7kg Brathähnchen gegessen hat. Da dies höchstwahrscheinlich einer sehr untypischen Aufnahme entspricht, wird dieser Wert wieder auf den niedrigeren Wert gesetzt.

**Annahmen** Die Unabhängigkeit und Varianzhomogenität stellt kein Problem dar. Die Normalitätstransformation mit dem hohen Wert führt allerdings wieder zu schlechten Testergebnissen. Vor allem der SW-Test kommt auf einen sehr kleinen p-Wert von  $5e-5$ . Auch nach dem niedriger setzten liefert ein Blick auf die transformierten Daten immer noch leichte Abweichungen von der gewünschten Normalverteilung. Die QQ-Plots der SPADE Methode müssen genau betrachtet werden.

**Auswertung** Ein Blick auf die Daten bezüglich des Alters zeigt wieder keinen wirklichen Trend. Die Auswertung wird daher wieder ohne die Modellierung der Kovariate durchgeführt.

**Ergebnisse und Vergleiche** Die Anteilsschätzungen liefern für beide Referenzwerte 0%. Die Aufnahme liegt doch deutlich höher als der empfohlene Bedarf. Die Bootstrap Konfidenzintervallschätzungen führten bei beiden Geschlechtern zu Problemen und können daher nicht angegeben werden. Die Varianzverhältnisse liefern wieder sehr hohe Werte, wobei Frauen (7.3) wieder einen höheren Wert als Männer (6.9) haben. Die Residuenplots liefern vor allem bei den Männern starke Abweichungen vom gewünschten Verhalten. Bei den Frauen sind nur die Enden der Verteilung zu hoch. Die Ergebnisse sind daher nur vorsichtig zu betrachten. Allerdings kann durch die hohe mittlere Aufnahme im Vergleich zu den Referenzwerten (Frauen MW 27, EAR 11; Männer MW 36, EAR 12) doch relativ sicher ausgesagt werden, dass die Aufnahme mit Vitamin B3 ausreichend für fast alle Leute ist.

#### 5.2.5 Vitamin B5

Pantothensäure zeigt keine vereinzelt Ausreißer, jedoch mehrere Werte außerhalb der Grenze des Medians plus 2 mal dem Interquartilsabstandes. Es ist somit schwer, nur vereinzelt Werte niedriger zu setzen. Werden alle Werte über der Grenze niedriger gesetzt, droht eine deutliche Fehleinschätzung, weshalb dies hier nicht gemacht wurde.

**Annahmen** Die Normalverteilungsüberprüfung liefert klare Abweichungen, sowohl grafisch als auch von beiden Testwerten her ( $p < 0.0007$ ). Da die Referenzwerte für Pantothensäure allerdings nur Schätzwerte bzw. AI Werte sind, sind sowieso nur Mittelwertvergleiche interessant und diese stellen kaum ein Problem - auch bei Verletzung der Annahme - dar. Die Schätzungen über die Varianzverhältnisse hingegen müssen mit äußerster Vorsicht betrachtet werden. Die anderen zwei Annahmen dürften keine Probleme darstellen.

**Auswertung** Bei den Männern zeigt diesmal der Vergleich mit dem Alter einen Trend, der möglicherweise sinnvoll modelliert werden kann. Frauen weisen hingegen keinen Trend auf.

Da die Box-Cox Transformation scheitert, die Daten an eine approximative Normalverteilung zu bringen, wird hier auch noch zusätzlich die ISU Methode angewandt.

**Ergebnisse und Vergleiche** Wie erwartet zeigen die Residuenplots an den Enden Abweichungen von der Normalverteilung. Bei den Männern wird ein linearer Trend nach unten bezüglich dem Alter geschätzt. Die Mittelwertschätzungen der SPADE Methode verglichen mit den Ergebnissen des EB weisen nur kleine Abweichungen auf (maximal 3-4%). Verglichen mit dem DACH Schätzwert (6mg) liegen bei den Männern und Frauen keine Altersklassen darüber. Betrachtet man den etwas niedrigeren AI Wert (5mg) so liegen nur bei den Männern die jüngste Alterklasse im Mittel (5.4mg) und Median (5.2mg) höher. Es kann geschlussfolgert werden, dass die Aufnahme an Pantothenensäure möglicherweise bei einigen Leuten zu niedrig liegt. Die Varianzverhältnisse zeigen Werte von 2.4 für Männer und 4.5 für Frauen, wobei diese Schätzungen nicht als genau angenommen werden dürfen, da die Normalverteilungsannahme relativ klar verletzt ist.

Die ISU Methode liefert mit ihrer 2 stufigen Transformation keine Annahmeverletzung. Die geschätzten Varianzverhältnisse liegen auch etwas niedriger als bei der SPADE Methode. Männer haben hier ein Verhältnis von 1.9 und Frauen eines von 3.7. Hier zeigt es sich, dass die Methode ISU bevorzugt werden sollte, wenn die übliche Aufnahmeverteilung geschätzt werden sollte. Tabelle 5.6 zeigt den Unterschied der Quantile für die gesamte Population. Es zeigen sich keine großen Unterschiede, trotz dieser Annahmeverletzung bei SPADE. Einzig die höheren Quantile weisen Abweichungen auf.

	Mittel	5%	25%	50%	75%	95%
SPADE	4.20	2.67	3.44	4.08	4.83	6.13
ISU	4.25	2.67	3.44	4.08	4.86	6.38
Differenz [%]	-1.18	0.00	0.00	0.00	-0.62	-4.00

Tabelle 5.6: Vergleich der SPADE und ISU Methode bezüglich der Aufnahme von Pantothenensäure [mg]

### 5.2.6 Vitamin B6

Vitamin B6 stellt sich wie Vitamin B5 als schwierig modellierbar dar. Viele Werte liegen außerhalb der Interquartilsreichweite. Betrachtet man die Werte genauer, so stellt sich heraus, dass es vor allem an dem konsumierten Energy Getränk der Marke Red Bull liegt. Diese Beobachtungen daher zu entfernen oder deren Effekt zu verringern wäre wenig gerechtfertigt, da dies ein Lebensmittel ist, dass regelmäßig von einigen Personen konsumiert wird und dadurch deutlich zur Vitamin B6 Zufuhr beiträgt.

**Annahmen** Auch hier stellt die Einfachheit der Box-Cox Transformation wieder eine klare Problematik dar. Extrem niedrige p-Werte und ein Blick aufs Histogramm der transformierten Daten zeigt dies deutlich. Es ist auch hier wieder sinnvoll die ISU Methode anzuwenden. Bezüglich den anderen 2 Annahmen muss sich weniger Gedanken gemacht werden, wie die Überprüfungen zeigen.

**Auswertung** Eine Modellierung mit dem Alter als Kovariate könnte hier bei den Männern sinnvoll sein. Der leichte Abwärtstrend erklärt sich vermutlich dadurch, dass vor allem die jüngeren Leute Red Bull konsumieren und dadurch höhere Aufnahmewerte vorweisen. Beide Modellierungsmöglichkeiten wurden in Betracht gezogen. Wie die Auswertungen aber zeigten, ist der Effekt nach der Transformation kaum mehr erkennbar. Auch die geschätzten Mittelwerte sind fast identisch, weswegen für die weiteren Vergleiche das einfache Modell genommen wurde.

**Ergebnisse und Vergleiche** In Tabelle 5.7 werden die Ergebnisse der Quantils-, sowie Anteilsschätzungen präsentiert. Da es sich bei Vitamin B6 wieder um einen Nährstoff handelt, für den ein EAR Wert, sowie eine DACH Empfehlung vorhanden ist, ist es sinnvoll auch die Anteilsschätzungen der ISU Methode mit in Betracht zu ziehen. Die ISU Methode hat wieder den Vorteil, dass ihre Transformation zu keinem Hinweis auf eine Verletzung der Normalitätsannahme führt. Diesmal liegen im Vergleich zur Pantothersäure doch stärkere Unterschiede zwischen SPADE und ISU vor, was womöglich an der noch stärker verletzten Annahme liegt. Die Residuenplots von SPADE zeigen sehr starke Abweichungen von der Normalverteilung an. Bei den Männern (obere Tabellenhälfte) sind stärkere Annahmeverletzungen als bei den Frauen zu erkennen, was sich auch im Vergleich der geschätzten Quantile mit ISU zeigt. Bei den Männern liegt die größte Abweichung beim 5% Quantil mit knapp 7.5%. Bei den Frauen sind es 2.3% beim 75% Quantil. Betrachtet man die Anteilsschätzungen der Unterversorgten, so liegt die ISU Methode fast immer höher als die SPADE Methode. Während bei den Frauen die Unterschiede nur sehr gering sind, so sind sie bei den Männern schon deutlich höher ausgefallen. Der ISU Schätzung muss hier mehr Relevanz zukommen, da sie die Annahmen weniger verletzt. Anmerkung: Die Bootstrap KI Schätzungen wurden hier weggelassen, da die Annahmen der SPADE Methode zu stark verletzt sind. Vergleicht man die Schätzungen, bezüglich der Unterversorgten, mit jenen des Ernährungsberichtes (Männer 38.4%, Frauen 38.8%), so kann wieder erkannt werden, dass diese Schätzungen zu hoch liegen. Nimmt man auch noch den EAR Wert als cut-point kommen doch deutlich niedrigere Ergebnisse und damit leicht andere Aussagen zustande.

	Mittel	5%	25%	50%	75%	95%	EAR[%]	DACH[%]	Var.ratio
SPADE M	2.01	1.08	1.50	1.90	2.39	3.33	9.48	24.8	1.66
ISU M	1.95	1.00	1.40	1.78	2.30	3.50	13.43	31.7	1.47
Differenz [%]	3.03	7.48	7.13	6.26	3.77	-4.92	-34.49	-24.5	12.14
	Mittel	5%	25%	50%	75%	95%	EAR[%]	DACH[%]	Var.ratio
SPADE F	1.47	0.94	1.20	1.42	1.69	2.16	19.83	24.80	2.91
ISU F	1.45	0.96	1.20	1.40	1.65	2.12	19.59	24.82	3.02
Differenz [%]	1.40	-1.98	0.16	1.74	2.30	1.69	1.21	-0.08	-3.71

Tabelle 5.7: Vergleich der SPADE und ISU Methode bezüglich Vitamin B6 [mg]. Oben Männer, unten Frauen

### 5.2.7 Vitamin B7

Die Aufnahmedaten zu Biotin zeigen wieder ein paar Ausreißer. Vor allem 3 Werte stechen heraus. Dabei handelt es sich um einen Mann und 2 Frauen. Einmal ist es wieder der Fall mit dem angereicherten Eiweißpulver und zweimal ein Nestle Fitness Eintrag, wobei es sich wahrscheinlich um angereicherte Frühstückscerealien handelt. Es ist also nicht gerechtfertigt die Daten zu verändern, da dies Lebensmittel sind, die regelmäßig konsumiert werden.

**Annahmen** Die Normalitätsannahme ist mit den Ausreißern stark verletzt. Beide Tests liefern p-Werte nahe 0. Ebenso ist mit den Ausreißern die Varianzannahme vermutlich verletzt, da die Überprüfungsabbildungen starke Abweichungen anzeigen. Die Unabhängigkeitsannahme scheint nicht verletzt zu sein. Setzt man die 3 Ausreißer auf den niedrigeren Wert (Median+2\*IQR) so ist die Normalverteilungsannahme zwar noch leicht problematisch, aber schon deutlich besser. Die Varianzannahme stellt auch ohne diese Werte weniger Schwierigkeiten dar.

**Auswertung** Die Modellierung mit Alter als Kovariate scheint hier eine Möglichkeit zu sein. Sie wird daher für die Analyse verwendet. Da Biotin wiederum nur AI und Schätzwerte vorweist, ist hier wiederum nur die Modellierung des Mittelwerts am wesentlichsten und daher wird der Vergleich mit ISU hier unterlassen. Ebenso werden für die Auswertung die Ausreißer nicht niedriger gesetzt.

**Ergebnisse und Vergleiche** Der AI beträgt  $30\mu\text{g}$  und der DACH Schätzwert  $30\text{-}60\mu\text{g}$ . Die mittlere und mediane Aufnahme ist für beide Geschlechter immer für jede Altersklasse, sowohl mit den Ausreißern als auch ohne, deutlich über dem AI Wert (Aufnahme um die  $40\text{-}50\mu\text{g}$ ). Bei beiden Geschlechtern zeigt die Modellierung mit dem Alter deutlich signifikante Ergebnisse. Das fraktionale Polynom wurde beide Male aber durch einen einfachen linearen Trend ersetzt. Vergleiche mit dem EB sind in Tabelle 5.8 zu sehen. Es sind doch größere Unterschiede erkennbar. Die Bootstrap KIs weisen einen leichten Bias auf, was vermutlich auf die Ausreißer zurückführbar ist. Vergleicht man die Intervalle allerdings mit denen des EB so sind sie doch relativ ähnlich groß, abgesehen vom Intervall für die Altersklasse (18-24) bei den Frauen. Hier wirkt durch die geringe Stichprobengröße der Einfluss der Ausreißer noch relativ stark. Das Konfidenzintervall ist daher auch viel größer als das Bootstrap Intervall. Womöglich hat hier SPADE einen Vorteil durch die Modellierung. Betrachtet man die Varianzverhältnisschätzungen so sind bei den Frauen wieder höhere Werte zu beobachten, egal ob mit Ausreißer (7.2) oder ohne (4.0). Männer hingegen haben ohne Ausreißer ein Verhältnis von 1.6 und mit eines von 1.4. Die Schätzungen sind allerdings, durch die Annahmenverletzungen, zu hinterfragen.

	[18-24]	[25-50]	[51-64]
SPADE Männer	51 [44,63]	46 [39,50]	39 [34,46]
EB Männer	54 [46,62]	41 [35,47]	38 [33,43]
n	17	87	44
SPADE Frauen	45 [40,49]	41 [38,44]	36 [32,40]
EB Frauen	60 [35,84]	42 [37,48]	35 [32,39]
n	37	143	52

Tabelle 5.8: Mittlere Aufnahmevergleiche von Biotin [ $\mu\text{g}$ ] zwischen Ernährungsbericht und SPADE

### 5.2.8 Vitamin B9

Vitamin B9 (hier alle Folat-Äquivalente<sup>2</sup> zusammen) stellt wieder eine einfachere Modellierung dar. Ein Wert sticht zwar ein klein wenig heraus, jedoch beeinflusst es die Auswertungen und Annahmen kaum. Da der Wert hauptsächlich durch einen Multivitaminensaft (1.7 Liter) zustande gekommen ist, wird er auch nicht aus der Auswertung genommen.

**Annahmen** Die Annahmen sind alle recht gut erfüllt.

**Auswertung** Das Alter scheint hier einen leichten Effekt zu haben, weshalb die Modellierung mit der Kovariate vorgenommen wurde. Interessanterweise wurde bei den Männern sogar ein fraktionales Polynom von SPADE als bestes Modell gewählt. Abbildung 5.1 zeigt dieses Polynom. Die roten Punkte stellen die Mittelwerte für jedes Alter dar. Die rote durchgehende Linie stellt den Mittelwert und die strichlierten schwarzen Linien die Quantile dar.

**Ergebnisse und Vergleiche** In Tabelle 5.9 ist die Gegenüberstellung der Mittelwert-schätzung vom EB und SPADE zu finden. Betrachtet man Männer zuerst, so fällt auf, dass die Schätzungen für alle Altersklassen über denen der einfachen Mittelwerte liegen. Die Bootstrap KIs sind außerdem leicht größer. Dies wird vor allem an der Modellierungsstruktur liegen, die vermutlich zu Overfitting führt. Betrachtet man hingegen Frauen, so sind die Bootstrap KIs wieder etwas kleiner als die des EB, da hier nur ein einfacher linearer Trend geschätzt wurde. Mehr Daten wären auf jeden Fall notwendig um zu sehen, ob dieser spezielle Trend bei den Männer nur Zufall war.

Betrachten wir weiters die Referenzwerte. Für Folatäquivalente gibt es EAR Werte ( $320\mu\text{g}$ ) sowie DACH Empfehlungen ( $400\mu\text{g}$ ). Im Ernährungsbericht wurden auch mit zwei weiteren

<sup>2</sup>Folat-Äquivalente bedeutet, dass es mehrere wirksame Verbindungen gibt, welche ähnlich wie Vitamin B9 wirken und daher werden alle in Betracht gezogen

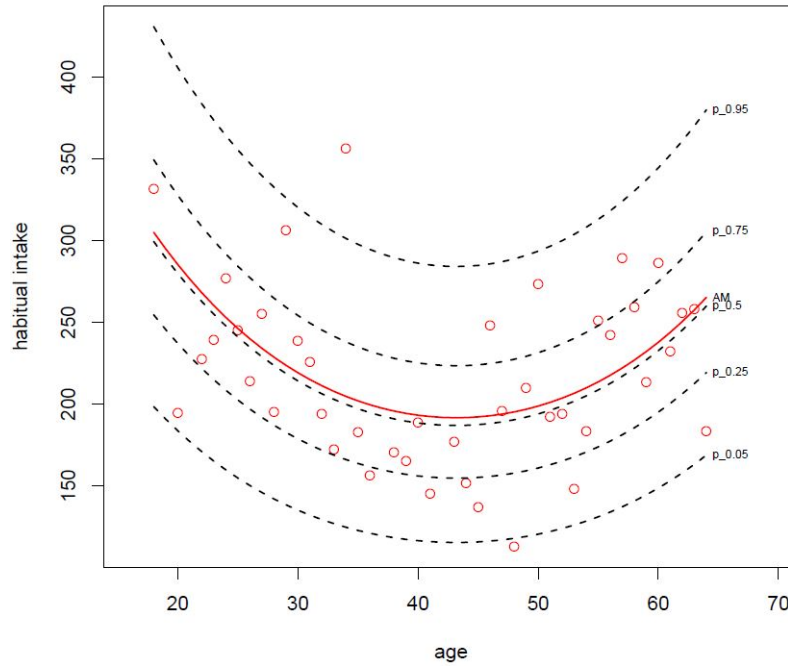


Abbildung 5.1: Vitamin B9 Aufnahmemodellierung für Männer. Geschätztes fraktionales Polynom bezüglich Alter.

Grenzen (200 und 300  $\mu\text{g}$ ) Anteile berechnet. Die SPADE Schätzung für Männer beträgt für den EAR Wert 93.3% [88.3,98.7] und für den DACH Wert 99.2% [97.5,100]. Im EB wurde mit dem DACH Wert eine Schätzung von 97.9% angegeben. Diese berücksichtigt allerdings nicht die within-person Varianz und liegt daher niedriger, da die Verteilung mehr streut. Der EAR Wert wäre allerdings angebrachter zu verwenden und würde daher doch in einer niedrigeren Schätzung resultieren.

Bei den Frauen ergab SPADE für den EAR Wert eine Schätzung von 97.3% [95.9,99.8] und für den DACH Wert 99.8% [99.6,100]. Im EB wurden 98.5% dokumentiert. Auch diese Schätzung liegt erwartungsgemäß unter jener der SPADE Methode.

Für die anderen zwei Grenzwerte wurden im EB für Frauen 47.3% und 86.6% dokumentiert. Für Männer wurden 49.7% und 84.1% geschätzt. Schätzungen mittels SPADE ergaben für diese Grenzwerte für Frauen 48.8% [39.6,58.1] und 94.9% [93.2,99.2]. Für Männer wurden die Werte 41.0% [29.0,51.1] und 89.5% [83.2,97.1] geschätzt. Erwartungsgemäß sind wieder die Schätzungen der Mittelwertmethode für Grenzwerte unterhalb dem Median größer und oberhalb des Medians kleiner, dank der within-person Varianz. Die Schätzungen der Mittelwertmethode liegen auch nur knapp innerhalb der Bootstrap Konfidenzintervalle, bis auf den Fall für Frauen und den Grenzwert 300 $\mu\text{g}$ . Es zeigt sich also hier, dass die Mittelwertmethode, für die Anteilsschätzungen der Unterversorgten, nicht geeignet ist.

Zuletzt werden hier noch die Varianzverhältnisse angegeben. Für Männer wurde ein Verhältnis von 1.8 und für Frauen wiederum das höhere Verhältnis von 2.1 geschätzt.

	[18-24]	[25-50]	[51-64]
SPADE Männer	263 [222,292]	208 [189,232]	226 [199,250]
EB Männer	255 [227,283]	197 [180,214]	222 [198,246]
n	17	87	44
SPADE Frauen	216 [199,232]	207 [196,217]	196 [177,211]
EB Frauen	229 [199,259]	216 [198,234]	193 [172,213]
n	37	143	52

Tabelle 5.9: Mittlere Aufnahmevergleiche von Folatäquivalenten [ $\mu\text{g}$ ] zwischen Ernährungsbericht und SPADE

### 5.2.9 Vitamin B12

Vitamin B12 verhält sich bis auf einen großen Ausreißer wieder recht normal. Der extreme Ausreißer kommt wieder durch die gebackene Leber zustande.

**Annahmen** Die Annahmen wären trotz dem extremen Ausreißer als wenig problematisch anzusehen. Einzig die SW-Statistik würde durch ihre Ausreißersensitivität einen schlechten Wert liefern. Entfernt man den Ausreißer, so zeigt auch diese Statistik einen deutlich besseren Wert.

**Auswertung** Die Auswertungen wurden mit und ohne Ausreißer vollzogen und die Ergebnisse der Schätzungen verglichen. Das Alter scheint wieder nur sehr wenig bis kaum einen Einfluss zu haben.

**Ergebnisse und Vergleiche** Die Ergebnisse zeigen bei den Frauen mit und ohne Ausreißer nicht allzu viele Unterschiede. Am stärksten würde sich nur das Varianzverhältnis verändern (mit Ausreißer 3.1, ohne 3.3). Männer liefern ein Varianzverhältnis von 3.8. Es werden also die Ergebnisse mit dem vollen Datensatz präsentiert. Die Anteilsschätzungen liefern für Männer mit dem EAR Wert 0.2 % [0,5,0] und mit dem DACH Referenzwert 3.9% [0,18.1]. Für Frauen kommt eine Schätzung von 5.9% [0.5,14] und 34.1% [20.8,44.8] heraus. Frauen dürften also mehr, von einer zu niedrigen Aufnahme, betroffen sein. Wie am Anfang der Arbeit schon (als Einführungsbeispiel) erwähnt, wurden im Ernährungsbericht für Vitamin B12 ebenfalls wieder Anteilsschätzungen angegeben. Für Frauen waren es mit dem DACH Wert 47.1% und für Männer 19.8%. Beide Male wiederum extrem höhere Schätzungen, durch die within-person Varianz. Es verdeutlicht hier also nochmals, dass im 2.Kapitel bereits erwähnte Problem. Einerseits ist der DACH Referenzwert ein zu hoher Wert, was zu einer Überschätzung führt und andererseits liefert die WPM Variante verzerrte Ergebnisse. Mit SPADE und den EAR Werten würde man eher zu dem Entschluss kommen, dass dies kein kritischer Nährstoff ist, während mit den DACH Werten und der WPM Methode doch eher von - vor allem bei den Frauen - einem problematischeren Ergebnis gesprochen werden würde.

### 5.2.10 Vitamin C

Vitamin C ist wieder ein Nährstoff, der sich recht gut für die Modellierung eignet. Ein extremer Ausreißer ist zu finden, welcher allerdings auf einen Eintragsfehler hinweist. Es handelt sich dabei um 150g Debreziner mit einem Vitamin C Wert von 1493mg. Zu Debreziner wird zwar Vitamin C zugesetzt, diese hohen Mengen sind aber sehr unrealistisch, weshalb der Eintrag vermutlich auf einem Fehler beruht. Die gesamte Tagesbeobachtung kann daher ohne große Bedenken für die Auswertungen herab gesetzt werden.

**Annahmen** Ohne dem klaren Ausreißer sind die Annahmen laut Testoutputs nicht verletzt. Auch die grafischen Überprüfungen zeigen keine groben Verletzungen der Annahmen.

**Auswertung** Modelliert wird für die Ergebnisse nur der Fall ohne Ausreißer, da dies gut begründbar ist. Das Alter scheint hier wieder einen Effekt zu haben, weshalb es in die Modellierung mit einbezogen wurde.

**Ergebnisse und Vergleiche** Die Varianzverhältnisse weisen Werte von 2.5 für Männer und 3.1 für Frauen auf. Die Residuenplots zeigen bei den Frauen sehr gutes Verhalten. Bei den Männern sind Abweichungen von der Normalverteilung zu erkennen. Das Ergebnis der Männer muss daher vorsichtiger interpretiert werden. Die Aufnahmewerte zeigen für die DACH Werte Schätzungen von 37.1% [21.0,46.8] für Männer und 42.8% [29.6,51.5] für Frauen. Die EAR Grenzwerte ergeben hingegen 18.1% [4.1,26.8] und 6.5% [0.4,13.6] für Frauen. Die Bootstrap Schätzungen weisen allerdings leichte Verzerrungen auf, weshalb diese nur vorsichtig interpretierbar sind. In Tabelle 5.10 können die Ergebnisse der Mittelwertsschätzung vom EB und SPADE verglichen werden. Bei den Männern zeigt vor allem SPADE für



die jüngste Altersklasse eine deutlich niedrigere Schätzung, wobei dies nicht auf den Effekt der Debreziner (kommt in Altersklasse 51-64 vor) zurückzuführen ist. Durch den linearen Trend, der bei der Modellierung vorausgesetzt wird, kommt es zu diesem deutlich niedrigeren Wert. Auffällig ist auch, dass Frauen weitaus kleinere Bootstrap Konfidenzintervalle aufweisen, was durch die niedrigere Streuung zwischen den Personen zustande kommt. Im Vergleich zum EB liegen die Schätzung von SPADE auch wieder für die niedrigste Altersklasse am weitesten entfernt. Insgesamt sind die Unterschiede nicht groß genug, um von klaren Abweichungen zu sprechen. Die Konfidenzintervalle überschneiden sich doch noch deutlich für alle Schätzungen und zeigen viel zu hohe Unsicherheit über die Punktschätzung an.

	[18-24]	[25-50]	[51-64]
SPADE Männer	141 [120,212]	127 [103,142]	110 [92,148]
EB Männer	172 [136,209]	110 [94,125]	117 [83,151]
n	17	87	44
SPADE Frauen	124 [108,142]	114 [103,123]	101 [87,115]
EB Frauen	144 [107,181]	116 [102,129]	99 [81,118]
n	37	143	52

Tabelle 5.10: Aufnahmevergleiche von Vitamin C [mg] zwischen Ernährungsbericht und SPADE

### 5.2.11 Vitamin D

Einen sehr speziellen Mikronährstoff stellt Vitamin D dar. Da dieser auch durch die Sonneneinstrahlung vom Körper selbstständig produziert werden kann, ist der Mensch nicht von der Aufnahme alleine abhängig. Daher kann (vor allem im Sommer) auch mit zu niedrigen Vitamin Aufnahmewerten ein ausreichender Status erreicht werden. Trotzdem ist man an der Aufnahmemenge interessiert. Betrachtet man die Daten des Ernährungsberichtes, so kann sofort die Schwierigkeit der Modellierung erkannt werden. Nicht nur vereinzelte hohe Werte kommen hier vor, sondern doch ein großer Anteil der über die Interquartilsreichweite hinaus reicht (knapp 6%). Diese sind wiederum zum Teil nur knapp und andererseits recht weit über dieser Grenze. Weiters kommt hinzu, dass hier 3 Tagesbeobachtungen mit einer Vitamin D Aufnahmemenge gleich 0 vorliegen. Diese Werte werden von SPADE mit der Hälfte des positiven Minimums ersetzt, um die Methode anwenden zu können.

Betrachtet man die hohen Werte genauer, so sind vor allem die Quellen Fisch und Wein als "große" Vitamin D Quellen auszumachen. Diese werden meist nicht täglich konsumiert, weshalb eine zusätzliche Schwierigkeit der Modellierung hinzu kommt. Die Werte auf den Median+2\*IQR zu reduzieren, ist ebenso nur schwer zu begründen. Mit mehr Daten wäre eine mögliche Modellierungsstrategie, die Lebensmittel die nicht täglich konsumiert werden (und Vitamin D reich sind), extra in einem 2-Teile Modell zu modellieren. Hier könnte einerseits für jede Person die Aufnahmewahrscheinlichkeit berechnet (zusätzliche Verzehrfragebögen können nützlich sein) werden und andererseits die Aufnahmemenge. Diese könnten dann, mit den täglich konsumierten Lebensmitteln, zusammengefügt werden. In dieser Arbeit wurde aber nur das einfache 1-Part Modell auf den vollen Datensatz angewandt, da für diese spezielle Modellierung deutlich zu wenig Daten vorliegen.

**Annahmen** Vitamin D stellt hier ein Problem dar. Sowohl die Normalverteilungsannahme ( $p < 3e-06$ ), sowie die Varianzhomogenität scheinen verletzt zu sein. Da beide Verletzungen für die ISU Methode weniger Probleme darstellen, wird hier auch wieder diese Methode angewandt. Die ISU Testoutputs liefern auch tatsächlich wieder keine Hinweise auf eine Annahmenverletzung. Der Korrelationscheck zeigt bei beiden Methoden keine Hinweise auf Probleme.

**Auswertung** Das Alter könnte einen Effekt haben, allerdings ist es schwierig zu sagen, da die großen Werte möglicherweise diesen beeinflussen. Nimmt man die meisten hohen

Werte hinaus, so fällt dieser Effekt nur mehr ganz gering aus, weshalb die Modellierung ohne das Alter gewählt wurde.

**Ergebnisse und Vergleiche** Wie erwartet zeigen die QQ-Plots der Residuen starke Abweichungen von der Normalverteilung. Die Ergebnisse von SPADE sind also nur vorsichtig zu betrachten. In Tabelle 5.11 wird deswegen wieder der Vergleich zwischen ISU und SPADE gegeben. Hier sind wieder große Unterschiede zwischen den Methoden zu erkennen. Vor allem bei den Männern (durch mehrere hohe Werte) ist durch die "schlechtere" Transformation der Daten bei SPADE ein wesentlicher Unterschied entstanden. Doch auch die ISU Methode hat hier Probleme, da das Varianzverhältnis extrem hoch ausgefallen ist. Dies ist durch eine sehr kleine between-person Varianzschätzung entstanden. Bei den Frauen ist der Unterschied der Schätzungen wesentlich kleiner ausgefallen. Auch die Anteilsschätzungen decken sich hier sehr gut. Bei den Männern wäre der Anteilsschätzung der ISU Methode mehr zu vertrauen als der SPADE Methode, allerdings auch nur bedingt. Bei den EB Ergebnissen wurden auch Anteile angegeben (Männer und Frauen wurden verwechselt). Für Männer wurde eine Schätzung von 82.7% und für Frauen 88.3% geschätzt. Beide Schätzungen liegen unter der ISU Schätzung (wiederum durch die within-person Varianz bedingt). Hier soll noch erwähnt werden, dass dieser Wert eigentlich nur eine DACH "Schätzung" darstellt und daher nicht argumentiert werden sollte, dass die meisten Personen unterversorgt sind. Für Vitamin D bleibt weiter zu diskutieren, wie mit den hohen Werten vorgegangen werden sollte. Ein anderer Modellierungsansatz (wie vorher beschrieben) wäre vermutlich besser. Dieser wäre allerdings nur durchführbar mit mehr Daten bzw. Informationen über die Lebensmittel mit hohen Vitamin D Werten.

	Mittel	5%	25%	50%	75%	95%	DACH[%]	Var.ratio
SPADE M	4.19	2.34	3.21	3.99	4.95	6.70	76.0	11.8
ISU M	4.43	3.80	4.15	4.41	4.73	5.13	91.5	198.0
Differenz [%]	-5.63	-47.70	-25.48	-9.94	4.49	26.55	-18.4	-177.5
	Mittel	5%	25%	50%	75%	95%	DACH[%]	Var.ratio
SPADE F	2.71	1.03	1.74	2.45	3.39	5.27	93.63	3.39
ISU F	2.67	1.07	1.69	2.35	3.32	5.32	93.47	3.94
Differenz [%]	1.39	-3.68	2.81	3.99	1.98	-0.87	0.18	-15.01

Tabelle 5.11: Vergleich der SPADE und ISU Methode bezüglich der Vitamin D Aufnahme [ $\mu\text{g}$ ]. Oben Männer, unten Frauen

### 5.2.12 Vitamin E

Vitamin E zeigt keine großen Ausreißer und auch sonst keine Hinweise auf Problemstellungen.

**Annahmen** Alle 3 Annahmen scheinen keineswegs verletzt zu sein.

**Auswertung** Die Modellierung mittels Alter dürfte hier keinen Vorteil bringen. Die Daten zeigen keinen Trend in der Aufnahme.

**Ergebnisse und Vergleiche** Im Ernährungsbericht wurde anhand der Mittelwerte geurteilt, dass die Aufnahme - da sie im Mittel passte - ausreichend gut ist. Da die DACH Werte für Vitamin E nur Schätzungen darstellen ist auch nicht mehr möglich. Das Institute of Medicine hat allerdings inzwischen schon eine EAR Schätzung und erlaubt daher die genaue Anteilsbestimmung. Die Anteilsschätzungen der Untervorgten betragen für Männer 25.8% [3.9,42.7] und für Frauen 34.7% [6.8,50.0]. Verzerrungen konnten nur für Frauen (auch bei 10.000 Ziehungen) festgestellt werden. Das KI ist daher mit Vorsicht zu betrachten. Die Varianzverhältnisse zeigen Werte von 4.5 für Männer und 5.8 für Frauen.

### 5.2.13 Vitamin K

Das letzte Vitamin, nämlich Vitamin K zeigt auf den ersten Blick keine auffälligen Schwierigkeiten für die Modellierung. Jedoch werden wir gleich sehen, dass trotzdem Probleme bei den Annahmen auftreten.

**Annahmen** Die Varianzhomogenität und Unabhängigkeitsannahme zeigen keine Probleme. Die Transformation zur Normalität wird allerdings von beiden Tests klar verworfen. Den Grund dafür zeigt die grafische Überprüfung. Es können leicht rechtsschiefe transformierte Werte festgestellt werden. Da für Vitamin K nur Schätzwerte und AI Werte vorhanden sind, ist allerdings nur der Mittelwert von größerem Interesse, weswegen die Modellierung trotzdem mit vollem Datensatz durchgeführt wird.

**Auswertung** Das Alter zeigt bei den Männern keinen Effekt, allerdings bei den Frauen einen leichten Abwärtstrend. Die Modellierung wurde trotzdem für beide Geschlechter mit dem Alter in Betracht gezogen.

**Ergebnisse und Vergleiche** Interessant ist, dass hier die AI Werte für Frauen ( $90\mu\text{g}$ ) und Männer ( $120\mu\text{g}$ ) doch deutlich höher ausgefallen sind als die DACH Schätzungen (Frauen  $60\text{-}65\mu\text{g}$ , Männer  $70\text{-}80\mu\text{g}$ ). Je nachdem welche Werte man verwenden möchte, liegen Männer über oder unter dem Referenzwert. Die Mittelwerte betragen für Männer ungefähr  $90\text{-}100\mu\text{g}$  und für Frauen  $90\text{-}110\mu\text{g}$ . Die Varianzverhältnisse zeigen für beide Geschlechter sehr ähnliche Werte. Für Frauen 3.7 und für Männer 3.6.

### 5.2.14 Zusammenfassung bezüglich Vitamine

In diesem kurzen Abschnitt werden nochmal die wichtigsten Erkenntnisse und Ergebnisse über die Vitamine zusammengefasst. In Tabelle 5.12 werden dazu nochmal alle Anteilsschätzungen bezüglich der Unterversorgten zusammengefasst. Zusätzlich wurden die, mittels WPM und DACH Referenzwert bestimmten Anteile<sup>3</sup> ebenfalls angegeben. Falls im Ernährungsbericht ebenfalls eine Anteilsschätzung angegeben wurde, ist dies mit einem Stern (\*) neben der WPM Schätzung gekennzeichnet. Weiters zeigt die Tabelle 5.12, ob die Annahmen (Unabhängigkeit, Varianzhomogenität, approx. Normalität) erfüllt sind oder ob Hinweise darauf deuten, dass sie verletzt sind. Je nachdem, ob sie wahrscheinlich verletzt sind oder nicht, müssen die Schätzungen mehr oder weniger kritisch hinterfragt werden. Die 4. Spalte zeigt außerdem an, ob die Annahmen, die im 2. Kapitel im Abschnitt 2.6 beschrieben wurden, vermutlich erfüllt sind oder nicht. Da für solche Aussagen Daten zu dem Bedarf des Kollektivs bekannt sein müssen, können diese Einschätzungen nicht mit den vorliegenden Daten getroffen werden. Stattdessen muss auf andere Quellen zurückgegriffen werden. Hier wurden die Einschätzungen des Berichtes von National Research Council (2000) (speziell Tabelle S-1, Seite 10) verwendet. Wie wir im Kapitel 2.4 der Arbeit gesehen haben, müssen auch die Annahmen für die EAR cut-point Methode erfüllt sein, damit mit den Anteilsschätzungen auch Aussagen über den tatsächlichen Versorgungsstatus des Kollektivs gemacht werden können. Wie in der Tabelle 5.12 und später auch in Tabelle 5.13 zu sehen ist, sind die Einschätzungen, ob die Annahmen der EAR cut-point Methode erfüllt sind nur für wenig Nährstoffe bekannt. Als Richtlinie können die Anteilsschätzungen für Nährstoffe, wo keine Einschätzungen vorhanden sind, jedoch trotzdem verwendet werden, um kritische Nährstoffe herauszufinden. Möchte man solche Aussagen machen, so würden vermutlich nur Vitamin D, B9 und möglicherweise Vitmamin B5 (Pantothensäure) als kritisch für die österreichische Bevölkerung eingestuft werden. Das Ziel dieser Arbeit ist es allerdings nicht herauszufinden welche Nährstoffe als kritisch angesehen werden können, sondern wo die Unterschiede in den Methoden liegen. Daher betrachten wir als nächstes nur die Ergebnisse der Schätzungen mit dem EAR und den DACH Werten, um darüber Erkenntnisse zu gewinnen. Die erste - und klarerweise logische - Erkenntnis wäre, dass die Schätzungen, welche den

<sup>3</sup>Anmerkung: Die Anteile können, je nach Art der verwendeten Verteilungsfunktion oder Umgang mit unterschiedlichen Referenzwerten je Altersklasse, auf mehrere Weisen bestimmt werden. Für die Anteilsschätzung hier wurde eine gewichtete geglättete Verteilungsfunktion (Typ  $\frac{i}{n+1}$ ) für jede Altersklasse berechnet und dann der (anhand der Anzahl an Personen je Alterklasse) gewichtete Mittelwert berechnet. R Funktion: `wtd.quantile(x,weights,type="i/(n+1)",normwt=T)`

EAR als cut-point haben immer niedriger liegen als die Anteile bezüglich den DACH Werten. Dies ist natürlich wenig verwunderlich, da sie doch beide unterschiedliche Definitionen haben. Während der EAR eher als Schätzung dient, die den Bedarf eines Nährstoffes für die Hälfte eines Kollektivs sieht, so sind die DACH Empfehlungen definiert, dass die Empfehlung für fast alle Personen ausreichend ist. Das eher der EAR Wert für die Anteilsschätzung der Unterversorgten herangezogen werden soll, wurde schon erklärt (siehe Kapitel 2.6.1). Zum Teil würden, wenn der DACH Wert verwendet wird, viele Nährstoffe viel schneller als kritisch eingestuft. Die meisten DACH Anteile liegen doch nahe den 50% und damit weit aus höher als zugleich die EAR Anteile (5-15%).

Zweite - und auch klare - Erkenntnis ist, dass die Anteilsschätzungen der Unterversorgten, die mittels der Mittelwertmethodik gemacht wurden, deutlich von der besseren Variante abweichen können. Für Vitamin D und Vitamin B9 liegen die Schätzungen weniger weit von der EAR Methode entfernt. Jedoch sind für die anderen Vitamine deutliche Abweichungen (um die 20-40%) zu erkennen. Es sollte also klar sein, dass die Mittelwertmethode auch nicht als approximative Schätzung verwendet werden sollte, da sie durch die vorhandene within-person Varianz immer zu hoch oder zu niedrig schätzt. Ebenso ist es fraglich, ob die DACH Werte als cut-points verwendet werden sollten (National Research Council, 2000), da sie den Anteil der Unterversorgten immer überschätzen.

Nächste Erkenntnis kann durch die geschätzten Varianzverhältnisse getroffen werden. Bei den 13 betrachteten Vitaminen sind 10 mal die Varianzverhältnisse bei den Frauen höher als bei den Männern. Dabei kann dies entweder auf eine höhere within-person Streuung oder eine geringere between-person Streuung der Frauen (immer relativ gesehen) deuten.

Wenn die Annahmen betrachtet werden, so kann festgestellt werden, dass meist nur die Transformation das Problem ist. Für solche Nährstoffe, wo die einfache Box-Cox Transformation nicht gut genug ist, sollte die ISU Methode zusätzlich durchgeführt werden. Solange die Annahme nicht extrem stark verletzt ist, sind allerdings ähnliche Ergebnisse zu erwarten.

Vitamin	SPADE Annahmen erfüllt	Var.ratio	cut-point Annahmen erfüllt	EAR	DACH	DACH (WPM)
A (Männer)	relativ sicher	4.8		4.3	46.8	63.0
A (Frauen)	relativ sicher	4.3		1.8	28.8	48.3
B1 (Männer)	relativ sicher	5.0	ja	7.3	31.3	51.1
B1 (Frauen)	relativ sicher	7.4	ja	12.4	31.3	44.2
B2 (Männer)	relativ sicher	3.0	ja	13.2	46.5	55.2
B2 (Frauen)	relativ sicher	3.4	ja	7.7	31.7	38.3
B3 (Männer)	unsicher	6.9	ja	0	0	
B3 (Frauen)	relativ sicher	7.3	ja	0	0	
B5 (Männer)	nein	2.4		MW<AI	MW<SW	
B5 (Frauen)	nein	4.5		MW<AI	MW<SW	
B6 (Männer)	nein	1.7	ja	9.5 (13.4)	24.8 (31.7)	38.4*
B6 (Frauen)	nein	2.9	ja	19.8	24.8	38.8*
B7 (Männer)	unsicher	1.4		MW>AI	MW>SW	
B7 (Frauen)	unsicher	4.0		MW>AI	MW>SW	
B9 (Männer)	relativ sicher	1.8	ja	93.3	99.2	97.9*
B9 (Frauen)	relativ sicher	2.1	ja	97.3	99.8	98.5*
B12 (Männer)	relativ sicher	3.8	ja	0.2	3.9	19.8*
B12 (Frauen)	relativ sicher	3.1	ja	5.9	34.1	47.1*
C (Männer)	relativ sicher	2.5	ja	18.1	37.1	53.4
C (Frauen)	relativ sicher	3.1	ja	6.5	42.8	49.9
D (Männer)	nein	11.8		MW<AI	76.1 (91.5)	82.7*
D (Frauen)	nein	3.4		MW<AI	93.6	88.3*
E (Männer)	relativ sicher	4.5	ja	25.8	MW>SW	
E (Frauen)	relativ sicher	5.8	ja	34.7	MW>SW	
K (Männer)	nein	3.6		MW<AI	MW>SW	
K (Frauen)	nein	3.7		MW<AI	MW>SW	

Tabelle 5.12: Anteilsschätzungen der Unterversorgten bezüglich der Vitaminversorgung. Vergleich zwischen den DACH und EAR Referenzwerten, ebenso wie der Vergleich zu den Ergebnissen im Ernährungsbericht 2012. Bemerkungen: In Klammern stehen die Schätzungen der ISU Methode. SW steht für den DACH Schätzwert. Der Stern (\*) in der letzten Spalte bedeutet, dass das Ergebnis auch im Ernährungsbericht veröffentlicht wurde.

## 5.3 Mineralstoffe

Bei den Mineralstoffen werden nicht alle einzeln analysiert und neu ausgewertet. Da sich die einfache Mittelwertschätzung und die SPADE Ergebnisse, wenn nur der einfache Gruppenmittelwert von Interesse ist, nur wenig unterscheiden, wird eine Auswertung nur für Mineralstoffe ausgeführt, wo EAR Werte oder DACH Empfehlungen bekannt sind. Für diese Nährstoffe kommt eine Anteilsschätzung in Frage und es ist daher sinnvoller die gesamte Aufnahmeverteilung mittels SPADE zu schätzen. Dies betrifft die Nährstoffe Kalzium, Kupfer, Eisen, Magnesium, Phosphor und Zink. Selen und Molybdän haben zwar ebenfalls einen EAR Wert, aber dazu sind keine Aufnahmedaten vorhanden. Für Jod sind die Aufnahmedaten problematisch, da eine große Jod Quelle die Salzzufuhr darstellt, welche kaum erfassbar ist. Das liegt daran, dass kaum angegeben werden kann wie viel Salz zu jedem Lebensmittelgericht hinzugefügt wurde. Daher sollten für den Jod Status andere Daten verwendet werden.

### 5.3.1 Kalzium

Der Mineralstoff Kalzium wird als erstes betrachtet. Im EB wurde dazu auch wieder ein Anteil der Leute mit zu geringer Aufnahme angegeben. Die Daten von Kalzium zeigen keine groben Ausreißer und deuten auf keine Probleme hin.

**Annahmen** Die Annahmetests sowie die grafischen Überprüfungen zeigen keine wirklichen Abweichungen vom gewünschten Zustand.

**Auswertung** Betrachtet man die Kalziumaufnahme bezüglich Alter so kann ein Effekt (Trend nach unten) erkannt werden. Die Modellierung erfolgte daher mit der Kovariate.

**Ergebnisse und Vergleiche** Die Residuenplots zeigen an den Enden der Verteilung bei Frauen leicht stärkere Abweichungen als bei den Männern. Insgesamt betrachtet aber keine groben Annahmeverletzungen. Im EB wurden für Kalzium die Anteilsschätzungen mit 70.5% für Frauen und mit 70.4% für Männer angegeben. Mit der SPADE Methode ergeben sich für die EAR Werte Schätzungen von 38.8% [23.1,49.2] für Männer und 55.4% [51.8,65.6] für Frauen. Für die DACH Empfehlungen ergeben sich 69.3% [59.6,85.7] für Männer und 86.1% [75.8,98.7] für Frauen. Vergleicht man die EAR Schätzungen mit den EB Angaben, so zeigt sich, dass bei Kalzium eine deutliche Überschätzung des Anteils durch die Mittelwertmethode zustande kommt. Auch die Bootstrap KI überdecken nicht die EB Schätzungen. Betrachten wir die Varianzverhältnisse so zeigt sich, dass Männer mit 3.0 wieder deutlich unter dem Verhältnis der Frauen liegen (5.5).

### 5.3.2 Magnesium

Magnesium zeigt ein paar erhöhte Werte auf, die allerdings nicht zu stark ins Gewicht fallen. Eine genaue Betrachtung der Einträge liefert keine wirkliche Begründung diese Werte niedriger zu setzen.

**Annahmen** Die Annahmen sind laut Tests alle plausibel. Eine grafische Überprüfung der Normalverteilungsannahme zeigt allerdings doch leichte Abweichungen. Die Residuenplots von SPADE müssen zusätzlich als Kontrolle herangezogen werden. Die Varianzannahme scheint möglicherweise verletzt zu sein. Deswegen wurde die ISU Modellierung mit ungleichen Varianzen extra angewandt um die Anteilsschätzungen bezüglich EAR mit denen von SPADE vergleichen zu können.

**Auswertung** Ein Alterseffekt ist bei den Männern möglicherweise sinnvoll. Bei den Frauen hingegen zeigt sich nur ein sehr schwacher Trend, der kaum eine Modellierung wert ist.

**Ergebnisse und Vergleiche** Für Magnesium wurden im EB keine Angaben zu den Anteilen der Unterversorgten gemacht. Nur die Mittelwerte wurde mit den DACH Empfehlungen verglichen. Die SPADE Schätzungen ergaben für Männer Anteile von 48.9% [37.0,57.9] (EAR) und 53.2% [41.8,63.5] (DACH). Für Frauen kamen 19.5% [9.2,28.6] für die EAR Werte und 42.2% [32.5,51.2] für die DACH Werte zustande. Bei den Frauen muss allerdings erwähnt werden, dass die Bootstrap Schätzungen leichte Verzerrungen aufwiesen. Die geschätzten Varianzverhältnisse ergaben für Männer 1.7 und für Frauen 2.4. Die QQ-Plots ergaben für Frauen stärkere Abweichungen als für Männer. Die Abweichungen waren allerdings nicht zu extrem um die Aussagen komplett unplausibel zu machen. Die ISU Methode erzeugt Schätzungen von 54.9% für Männer und 18.0% für Frauen (bezüglich EAR). Die Schätzung der ISU Methode bezüglich der Männer weicht mehr von der SPADE Schätzung ab, als für Frauen. Dieser größere Unterschied kann allerdings auch durch die andere Modellierungsstruktur des Mittelwertes - bei SPADE mit linearem Trend bezüglich dem Alter - erklärt werden, welcher bei den Frauen nicht vorkommt. Die mögliche Verletzung der Varianzannahme scheint also keinen großen Effekt auf die Schätzungen zu haben.

### 5.3.3 Eisen

Der Nährstoff Eisen weist ebenfalls ein unauffälliges Histogramm auf. Lediglich zwei etwas erhöhte Werte sind erkennbar, welche einmal durch "Blunzengröstl" und einmal durch Fisch, sowie Schokoladenkuchen zustande gekommen sind.

**Annahmen** Bei den Annahmen stellt die Unabhängigkeit laut Tests die größten Probleme dar ( $p=0.05$ ). Eine grafische Überprüfung zeigt keinen Hinweis, dass dieses Ergebnis durch einen Ausreißer zustande gekommen ist.

**Auswertung** Die Auswertung wurde ohne die Kovariate Alter vorgenommen, da die Daten keinen Trend aufzeigen.

**Ergebnisse und Vergleiche** Das geschätzte Varianzverhältnis beträgt für Männer 3.6 und für Frauen 2.4. Die Residuenplots zeigen keine groben Verletzungen der Normalverteilungsannahme an. Die SPADE Schätzungen der Anteile kommen für die EAR Referenzwerte auf 0% für Männer und 8.0% [0.6,12.3] für Frauen. Bei den DACH Empfehlungen liegen die Werte bei 11.7% (Männer) und 83.1% [79.0,87.9] (Frauen). Die Bootstrap KI Schätzungen für Männer führten zu Fehlern. Der Grund ist unbekannt. Im Ernährungsbericht wurden für Frauen 77.9% und für Männer 30.1% für die Anteile angegeben. Hier kann nun ein ganz extremes Beispiel betrachtet werden. Während im EB Aussagen über eine zu niedrige Aufnahme, vor allem für Frauen, geschlussfolgert wurde, so zeigten die Blutmarkerdaten, dass kaum eine Unterversorgung vorlag. Das vermutlich keine Unterversorgung vorliegt, würde auch durch die Schätzung mittels SPADE und dem EAR als cut-point, bestätigt werden.

### 5.3.4 Kupfer

Kupfer weist vor allem 3 große Ausreißer auf. Es handelt sich dabei um einen Mann und zwei Frauen. Die Lebensmittel die dazu geführt haben, sind hauptsächlich Portwein bei zwei Personen und einmal wieder die gebackene Leber.

**Annahmen** Mit den Ausreißern werden alle 3 Annahmen deutlich verletzt ( $p<0.006$  im größten Fall). Es wäre daher keine gute Idee, diese Werte mit in die Modellierung zu nehmen. Daher werden die 3 Werte herabgesetzt. Danach scheint nur noch die Normalität möglicherweise ein Problem darzustellen. Es wird wieder ein strenger Blick auf die Residuen der Modelle geworfen.

**Auswertung** Eine Modellierung mit dem Alter als Kovariate scheint für Männer eine Möglichkeit darzustellen. Für Frauen zeigt sich kein erkennbarer Trend.

**Ergebnisse und Vergleiche** Das Varianzverhältnis für Männer beträgt 2.5 und für Frauen 1.9. Die Schätzung für Männer muss allerdings mit Vorsicht betrachtet werden, da der QQ-Plot der Residuen eindeutig eine Verletzung der Normalitätsannahme anzeigt. Bei den Frauen zeigt sich nur ein Ausreißer Wert (trotz der 2 niedriger gesetzten Werte). Für Kupfer gibt es EAR Werte aber keine DACH Empfehlungen. Die EAR Werte für Frauen und Männer für die Anteilsschätzungen zeigen für beide Geschlechter 0%. Auch die Konfidenzintervalle zeigen als Obergrenzen 0% an. Ein Blick auf die geschätzten Quantile und Mittelwerte (um die 1.7mg) zeigt, dass die Aufnahme deutlich über dem EAR (0.7) liegt und es daher recht sicher keine zu niedrige Aufnahme für die Personen gibt.

### 5.3.5 Phosphor

Phosphor stellt wiederum einen recht einfachen Nährstoff für die Modellierung dar. Keine deutlichen Ausreißer zu sehen und auch das Histogramm der Daten zeigt keine Problemstellungen.

**Annahmen** Die Annahmen sind alle recht sicher erfüllt.

**Auswertung** Eine Modellierung mittels Alter scheint für keines der beiden Geschlechter sinnvoll zu sein.

**Ergebnisse und Vergleiche** Ähnlich wie bei Kupfer liegt hier keine Unterversorgung vor. Die Schätzungen für den EAR Wert beliefen sich für beide Geschlechter auf 0% durch die deutlich höhere Aufnahme der Personen. Auch die Konfidenzintervalle zeigen als Obergrenze gerade mal 0.1% an. Das Varianzverhältnis beträgt für Männer 2.9 und für Frauen 5.8.

### 5.3.6 Zink

Auch der letzte genau betrachtete Mineralstoff zeigt sich als leicht modellierbar. Keine deutlichen Ausreißer erkennbar, sowie keine Auffälligkeiten bei den Histogrammen.

**Annahmen** Die Annahmen sind alle recht sicher erfüllt.

**Auswertung** Das Alter scheint auch für die Zinkaufnahme keine Rolle zu spielen, weswegen die einfachere Modellierung gewählt wurde.

**Ergebnisse und Vergleiche** Die Residuenplots bestätigten den Verdacht. Es sind keine Auffälligkeiten zu erkennen. Das Varianzverhältnis wird für Männer auf 6.0 und für Frauen auf 6.2 geschätzt. Die Anteilsschätzungen ergeben für Männer die Werte 6.7% (EAR) und 13.5% (DACH). Für Frauen kommen Anteile von 1.3% und 2.1% heraus. Die Konfidenzintervalle wurden nicht angegeben, da sie starke Verzerrungen aufweisen.

### 5.3.7 Zusammenfassung der Mineralstoffe

In Tabelle 5.13 ist wieder ein Überblick über die Anteilsschätzungen der Unterversorgten gegeben. Die Tabelle hat dieselbe Struktur, wie Tab.5.12 für die Vitamine. Der Stern (\*) neben der WPM Anteilsschätzung zeigt wieder an, ob die Schätzung im Ernährungsbericht verwendet wurde.

Bei den betrachteten Mineralstoffen zeigt sich ein ähnliches Bild wie bei den Vitaminen. Die DACH Anteile liegen wieder immer deutlich höher als die EAR Anteile und auch die im Ernährungsbericht dokumentierten Schätzungen liegen weit neben den genaueren Schätzungen. Vor allem beim Eisen ist ein riesiger Unterschied der beiden Schätzungen für Frauen zu erkennen (8% und 78%). Einen Teil dieser Differenz macht allerdings auch die Abweichung der Bedarfswerte aus. Allerdings zeigt dies wiederum auf, wie stark unterschiedlich die Aussagen sein können. Bei Kalzium ist der Unterschied zumindest noch 15-30% groß. Auch für Zink und Magnesium liefern die Anteilsschätzungen der Unterversorgten mittels WPM deutlich höhere Werte.

Betrachtet man wieder die Varianzverhältnisse, so ist wieder 4 von 6 mal das Verhältnis der



Frauen höher als das der Männer. Wenn die Annahmenüberprüfungen betrachtet werden, zeigt sich, dass diese meist recht sicher erfüllt sind. Nur Kupfer stellt hier ein Problem dar. Eine Modellierung mit der ISU Methode ist allerdings nicht notwendig, da die Aufnahmen deutlich höher als die EAR Werte liegen und daher keine Unterschiede in den Schätzungen vorkommen. Bei Magnesium ist das "unsicher" bezüglich der Annahmeverletzung vermutlich zu streng gesehen, da die Unterschiede zwischen einer Modellierung mit ungleichen Varianzen (ISU) zur normalen Modellierung mittels SPADE gering sind.

Mineralstoff	SPADE Annahmen erfüllt	Var.ratio	cut-point Annahmen erfüllt	EAR	DACH	DACH (WPM)
Kalzium (Männer)	relativ sicher	3		38.8	69.3	70.4*
Kalzium (Frauen)	relativ sicher	5.5		55.4	86.1	70.5*
Magnesium (Männer)	unsicher	1.7	ja	48.9	53.2	61.2
Magnesium (Frauen)	unsicher	2.4	ja	19.5	42.2	45.8
Eisen (Männer)	relativ sicher	3.6		0	11.7	30.1*
Eisen (Frauen)	relativ sicher	2.4		8.0	83.1	77.9*
Kupfer (Männer)	nein	2.5		0		
Kupfer (Frauen)	unsicher	1.9		0		
Phosphor (Männer)	relativ sicher	2.9	ja	0	0	
Phosphor (Frauen)	relativ sicher	5.8	ja	0	0	
Zink (Männer)	relativ sicher	6.0		6.7	13.5	39.3
Zink (Frauen)	relativ sicher	6.2		1.3	2.1	20.2

Tabelle 5.13: Anteilsschätzungen der Mineralstoffversorgung. Vergleich zwischen den DACH und EAR Referenzwerten, ebenso wie der Vergleich zu den Ergebnissen im Ernährungsbericht 2012. Der Stern (\*) in der letzten Spalte bedeutet, dass das Ergebnis auch im Ernährungsbericht veröffentlicht wurde.

Bevor im letzten Kapitel der Arbeit noch einmal die wichtigsten Punkte zusammengefasst und Schlussfolgerungen beschrieben werden, wird im nächsten Kapitel noch auf die letzte offene Annahme (Erwartungstreue), welche in diesem Kapitel nicht kontrolliert wurde, eingegangen.

# Kapitel 6

## Diskussion und Schlussfolgerung

In diesem letzten Kapitel der Arbeit sollen nochmal die wesentlichen Erkenntnisse der Arbeit diskutiert werden. Zuerst wird über die letzte noch zu überprüfende Annahme - die Erwartungstreue der Daten - diskutiert. Im Anschluss werden dann die Erkenntnisse bezüglich den vorgestellten Modellen und der Neuauswertung der Ernährungsberichtsdaten beschrieben. Zuletzt sollen dann noch in einem kurzen Abschnitt die wichtigsten Schlussfolgerungen und Empfehlungen präsentiert werden.

### 6.1 Verzerrte Daten?

Viele Studien haben sich schon mit dem Thema "Misreporting" beschäftigt. Ein paar sollen hier auch erwähnt und deren Ergebnisse nun kurz diskutiert werden. Prinzipiell lässt sich sagen, dass es nicht ganz so einfach ist festzustellen, ob jemand zu wenig oder zu viel Lebensmittel(-menge) angegeben hat. Eine Möglichkeit ergibt sich durch Biomarker. So gibt es z.B. die Möglichkeit, die angegebene Proteinaufnahme mittels dem Stickstoff im ausgeschiedenen Urin zu validieren (Bingham, 2003). Die Studienteilnehmer müssen dazu jedes Mal über den Tag verteilt ihren Urin sammeln und für die Auswertungen abgeben. Die dadurch ermittelte Stickstoffmenge lässt darauf schließen wie viel Protein an dem Tag zu sich genommen wurde. Diese Daten können nun mit den angegebenen Proteinmengen verglichen werden. Ähnliche Möglichkeiten gibt es auch für Kalium, wobei hier die Kaliumausscheidung im Urin hochgerechnet wird auf die zugeführte Menge. Eine Studie die diese beiden Biomarker zur Validierung verwendete ist von Crispim et al. (2011). Die Aufnahme wurde mittels zwei 24 Stunden Erinnerungsprotokolle gemessen. Die Daten stammten von Männern und Frauen aus verschiedenen europäischen Ländern (Tschechische Republik, Frankreich, Niederlande, Belgien und Norwegen). Eine zu niedrig angegebene Aufnahmemenge von Protein (2-13%) und von Kalium (4-17%) wurde dabei festgestellt. Ebenso wurden die geschätzten üblichen Aufnahmeverteilungen (mittels MSM) mit den Verteilungen der Biomarker verglichen. Dabei wurde festgestellt, dass vor allem in den oberen Enden der Verteilungen unterschätzt wurde. Ebenso wurde in der OPEN Studie der Vereinigten Staaten eine Unterschätzung der Proteinaufnahme von 11-15% festgestellt (Subar et al., 2003). Bei dieser Studie wurde auch die Energieaufnahme überprüft. Dazu wurde doubly labeled water (doppelt markiertes Wasser) verwendet. Ganz vereinfacht gesagt wird das Sauerstoff und Wasseratom "markiert" um es bei der Ausscheidung über den Urin messen zu können. Dadurch kann dann der Energieverbrauch ausgerechnet und mit der angegebenen Energieaufnahme verglichen werden. Dabei wird die Gleichung  $\text{Energieaufnahme} = \text{Energieverbrauch}$  verwendet. Die verbrauchte Energie ist solange ungefähr gleich der aufgenommenen Energie solange das selbe Körpergewicht gehalten wird. Auch bei diesen Vergleichen wurde eine durchschnittlich zu niedrige Angabe von 12-20% festgestellt.

Was heißt dies für die Daten bezüglich dem Ernährungsbericht? Sind hier ebenfalls zu niedrig angegebene Daten der Fall? Um dies zu überprüfen stehen zwar keine Biomarkerdaten zur Verfügung, aber andere Möglichkeiten. So gibt es z.B. die Möglichkeit den Energieverbrauch zu schätzen und dadurch mögliche Underreporter zu identifizieren. Der sogenannte Goldberg cut-off wurde auch für den Ernährungsbericht verwendet um Leute, die womöglich zu wenig angegeben haben zu identifizieren. Auf dies wird jetzt hier noch kurz näher eingegangen.

### 6.1.1 Goldberg Cut-Off

Die Idee, die dahinter steckt, ist die Energieaufnahme gleichzusetzen mit dem Energieverbrauch. Aus dieser Gleichung kann durch den PAL<sup>1</sup> Wert (Physical Activity Level), welcher sich aus dem Quotienten Energieverbrauch/Grundumsatz<sup>2</sup> ergibt, eine neue Gleichung hergeleitet werden (Energieaufnahme/Grundumsatz=PAL). Für diese Gleichung wird nun versucht herauszufinden welche unteren (und oberen) Werte sinnvoll sind, um plausible Personenangaben zu identifizieren. Goldberg et al. (1991) hat sich damit beschäftigt und unter der Annahme einer Lognormalverteilung für die Energieaufnahme mathematische Grenzen hergeleitet (siehe Appendix des Papers). Diese sollen die möglichen Schwankungen die durch den Zufall entstehen können berücksichtigen, um so nicht jede(n) die/der an den beobachteten Tagen weniger gegessen als erwartet hat, automatisch als fälschlichen Underreporter festzustellen. Die untere 95% Konfidenzgrenze ergibt sich wie folgt:

$$\frac{EA}{GU} > PAL * exp\left(-2\frac{S/100}{\sqrt{n}}\right) \quad (6.1)$$

$$S = \sqrt{\frac{CV_{wEA}^2}{d} + CV_{wB}^2 + CV_{tP}^2}, \quad (6.2)$$

wobei EA die Energieaufnahme und GU den Grundumsatz bezeichnet. CV stellt den Koeffizient der Variation für die within-subject Energieaufnahme (wEI), den within-subject Grundumsatz (wB) und den PAL Wert dar. Die Werte hierfür wurden aus mehreren Studien zusammengetragen und folgendermaßen geschätzt: 23% für  $CV_{wEA}$ , 8.5% für  $CV_{wB}$  und 15% für  $CV_{tP}$  (Black, 2000a). Sollte der Grundumsatz gemessen anstatt geschätzt worden sein, so verringert sich der Wert für  $CV_{wB}$  auf 4%. Weiters bezeichnet der Parameter d die Anzahl an Beobachtungstagen je Person und n die Stichprobengröße. Es kann diese Untergrenze also für das mittlere Verhältnis von EA zu GU einer Stichprobe verwendet werden oder für einzelne Personen mit n=1. Ebenso kann für den Variationskoeffizienten der Energieaufnahme eine eigene Schätzung eingesetzt werden um eine bessere Anpassung der Schranke an das untersuchte Kollektiv zu gewährleisten. Dies wurde auch für die nächsten Auswertungen gemacht und anhand der folgenden, von Goldberg vorgeschlagenen, Formel getan.

$$CV_{wEA} = \sqrt{\sum_{i=1}^n CV_i^2/n}$$

Für  $CV_i$  wurde der geschätzte Variationskoeffizient für jede Person mit zwei Beobachtungen verwendet. Dadurch erhält man für den gesamten Koeffizienten einen Wert von 29%. Dieser wurde anstatt den vorgeschlagenen 23% verwendet.

Um den Ausdruck in 6.1 nun überprüfen zu können, benötigen wir eine Schätzung oder Messung des Grundumsatzes jeder Person und einen angemessenen PAL Wert. Wenn man daran interessiert ist, das mittlere EA zu GU Verhältnis und damit die Plausibilität der ganzen Stichprobe zu überprüfen, so muss ein PAL Wert gewählt werden, welcher möglichst gut das mittlere Bewegungsverhalten der ganzen Stichprobe abschätzt. Im Fall der Überprüfung jedes Individuums sollte ein geeigneter PAL für jedes einzelne Individuum gewählt werden, um die Sensitivität<sup>3</sup> der Schranke zu maximieren. Die Sensitivität leidet stark darunter, wenn nur ein einzelner PAL Wert für alle Personen genommen wird (Black, 2000b). Daher wurde der PAL Wert anhand der Informationen über das Bewegungsverhalten der Personen, welches mittels Fragebogen abgefragt wurde, gewählt. Dabei wurden Fragen zur Einschätzung der täglichen Zeit im Stehen und Sitzen usw. verwendet, um eine Klassifizierung in drei Kategorien für jede Person vornehmen zu können. Die Aktivitätsklassen sind niedrig, moderat und hoch. Für jede dieser Klassen wurde ein entsprechender PAL Wert gewählt

<sup>1</sup>Der PAL Wert gibt das Vielfache des Energieverbrauchs zum Grundumsatz an und wird oft verwendet um herauszufinden wie aktiv der Lebensstil einer Person ist. Werte von ca. 1.5 sind üblich für Personen mit leichten Berufen (wenig Bewegung) und Werte um 2 ca. sind schon für Personen mit anstrengenderen Berufen (Bauarbeiter)

<sup>2</sup>Der Grundumsatz ist jener Energieumsatz der in vollkommener Ruhe, bei angenehmer Temperatur (20-28 Grad) und nüchtern zur Aufrechterhaltung der Körperfunktionen notwendig ist. Also jene Energie, welche die Zellen und Organe ohne körperliche Tätigkeit brauchen.

<sup>3</sup>Sensitivität ist ein Maß für das korrekte Identifizieren eines Underreporters und kann Werte von 0-100% annehmen.

(niedrig = 1.55, moderat = 1.7, hoch = 2).

Für den Grundumsatz wurde in Formel 6.1 für Personen, wo aus dem Ernährungsbericht keine Messungen dazu vorlagen, eine Schätzung nach Schofield (1985) vorgenommen. Diese stellen einfache, jedoch recht gute Einschätzungen des Grundumsatzes dar. Damals wurde anhand mehrerer Datenquellen eine Regression mit dem gemessenen Grundumsatzes als abhängige und dem Gewicht der Personen als unabhängige Variable angepasst. Dies wurde separat für mehrere Altersklassen und dem Geschlecht durchgeführt. Der geschätzte Standardfehler liegt ungefähr bei 60-170 kcal. Kritik an der Schätzung kann vor allem an dem schlechteren Abschneiden der Formeln für Übergewichtige angebracht werden. Laut Black (2000a) liegt vor allem bei Frauen eine Überschätzung des Grundumsatzes für Übergewichtige vor. Dies führt beim Goldberg Cut-Off dazu, dass das Verhältnis EA zu GU niedriger ausfällt und somit Übergewichtige schneller als Underreporter eingestuft werden. Dies stellt allerdings nicht so ein großes Problem dar, da vor allem Übergewichtige Personen zu Underreporting neigen, wie auch Tabelle 5.1 zeigt. Die Übergewichtigen haben nicht mehr Lebensmittel angegeben als Personen mit normalen Gewicht. Trotzdem wurde anhand der gemessenen vorliegenden Grundumsatz Werte zusätzlich für die Stichprobe ein Regressionsmodell geschätzt (n=291). Als bestes Modell wurde das Modell mit Geschlecht, Körpergewicht und Körpergröße gewählt (siehe summary statistics in Tab.6.1). Die Underreporter wurden sowohl mit den Schofield, als auch mit diesen Modellschätzungen ermittelt.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	424.8098	118.9842	3.57	0.0004
Körpergewicht	6.5163	0.4283	15.22	0.0000
Körpergröße	3.0141	0.7428	4.06	0.0001
Geschlecht/männlich	226.1501	14.6879	15.40	0.0000

Tabelle 6.1: Summary Statistik zum Regressionsmodell bezüglich Grundumsatzschätzung

## 6.1.2 Resultate

Zuerst wurden die einzelnen Individuen betrachtet. Im Ernährungsbericht wurde ebenfalls die Goldberg Cut-Off Methode verwendet, um fälschliche Angaben zu identifizieren, allerdings wurde nicht genau auf die verwendeten Werte eingegangen. Für die Schätzung des Grundumsatzes wurden ebenfalls die Schofield Gleichungen verwendet. Im EB wurden 38 Underreporter festgestellt (9.1%). Bei Verwenden der im vorigen Abschnitt beschriebenen Werte und der Schofield Gleichung für Personen mit keinen Grundumsatzmessungen wurde ein weitaus höherer Anteil von 31.3% ermittelt. Dabei wurden alle im EB als Underreporter ermittelten Personen ebenfalls als Underreporter eingestuft. Der niedrigere Anteil im EB lässt sich vermutlich auf einen niedrigen verwendeten PAL Wert, welcher für alle Personen gleich verwendet wurde, rückschließen. Dadurch wurde eine weitaus niedrigere Sensitivität erreicht. Black (2000b) hat eine Sensitivität von um die 50% ermittelt, wenn ein einzelner PAL Wert für alle Personen verwendet wird. Durch Verwenden von drei verschiedenen PAL Werten, welche anhand von den Kategorien der Bewegungseinstufung ermittelt wurde, konnte die Sensitivität auf ungefähr 70% erhöht werden, ohne die Spezifität<sup>4</sup> (97-98%) zu erniedrigen. Es kann also davon ausgegangen werden, dass im EB zu wenig Underreporter ermittelt bzw. ausgeschlossen wurden.

Auch wenn für die Grundumsatzschätzung das eigene Regressionsmodell (Tab.6.1) verwendet wird - welches möglicherweise besser für Übergewichtige abschneidet - bleibt der Anteil recht hoch mit um die 30%.

Berechnet man das mittlere Verhältnis von EA zu GU, so erhält man ein Verhältnis von 1.27. Dieses ist niedrig und deutet auf Verzerrungen der Daten hin. Der Goldberg Cut-Off (für n=418 und einem mittleren PAL von 1.6) wäre 1.55. Somit liegt das beobachtete Verhältnis deutlich zu niedrig. Das Verhältnis würde ungefähr einem PAL Wert von 1.3 entsprechen, was eher ein Wert für bettlägerige Personen wäre und daher nicht passend für die österreichische Bevölkerung. Im Ernährungsbericht von 2012 wurden mittlere PAL Werte von 1.45-1.73 für österreichische Erwachsene angegeben (Elmadfa et al., 2012, Tab.5.6). Entfernt man die

<sup>4</sup>Spezifität ist ein Maß für die korrekte Spezifikation der Nicht-Underreporter

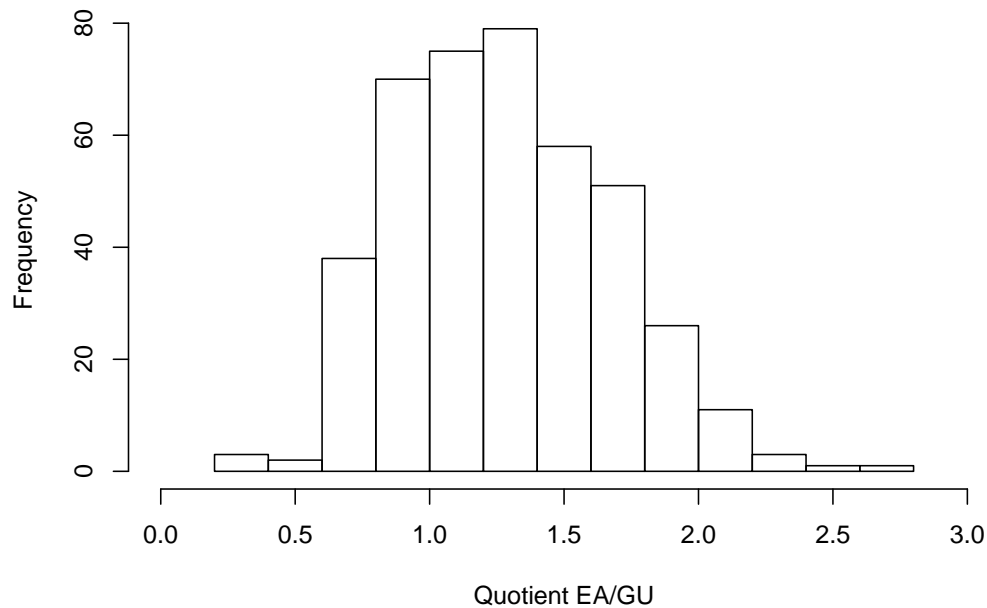


Abbildung 6.1: Histogramm der geschätzten Energieaufnahme/Grundumsatz Quotienten. Für den Grundumsatz wurde die Berechnung nach Schofield verwendet.

30% als Underreporter identifizierte Personen, so erhält man ein Verhältnis von 1.45, welches immer noch eine deutliche Verzerrung der Daten anzeigt. In Abbildung 6.1 kann die Verteilung des geschätzten Quotienten von Energieaufnahme zu Grundumsatz betrachtet werden. Für diese Grafik wurde der geschätzte Grundumsatz nach Schofield verwendet. Es ist auch hier deutlich zu sehen, dass die meisten Personen deutlich unter einem realistischen Quotienten von größer 1.6 liegen.

Was bedeuten diese Resultate für die Auswertungen in Kapitel 5, bzw. die im Ernährungsbericht? Dies lässt sich nicht allzu leicht beantworten. Das Problem liegt darin, dass nicht bekannt ist, welche Lebensmittel nicht angegeben wurden bzw. bei welchen Lebensmittel eine zu niedrige Menge angegeben wurde. Klar ist, dass eine Verzerrung nach unten für die Energieaufnahme vorliegt. Die Schätzungen sind alle vermutlich deutlich zu niedrig und keineswegs repräsentativ für die tatsächliche Energieaufnahme der österreichischen Bevölkerung. Bei den Makronährstoffen kann auch relativ sicher davon ausgegangen werden, dass alle geschätzten Werte zu niedrig liegen. Ob und bei welchen Mikronährstoffen ein Bias vorliegt ist allerdings schwer zu sagen. Wenn z.B. vor allem Lebensmittel nicht angegeben werden, welche viele Kalorien enthalten, allerdings nur wenig bis kaum Mikronährstoffe (z.B. zuckerreiche oder fettreiche Lebensmittel), so kann es durchaus sein, dass die übliche Aufnahmeschätzungen durchaus unverzerrt sind. Der Frage ist z.B. Poslusna et al. (2009) nachgegangen. Dazu wurden mehrere Studien betrachtet, welche Underreporting auch bezüglich Vitamin C, Kalzium und Eisen aber auch für die Makronährstoffe und Energie bewerteten. Prinzipiell wurden ebenso wie hier um die 30% Underreporter im Schnitt identifiziert. Auch die drei betrachteten Mikronährstoffe waren im Schnitt um 30% niedriger in den als Underreporter identifizierten Personen. Auch für die EB Daten zeigt sich ein ähnliches Bild. Tabelle 6.2 zeigt die Mittelwerte der beobachteten Daten bezüglich Underreporter und Normal kategorisierten Personen. Es ist also auch für die Mikronährstoffe zu erwarten, dass die Schätzungen alle zu niedrig liegen. Beachte, dass allerdings schon ein Teil (9%) bei den Auswertungen in Kapitel 5 ausgeschlossen wurden. Wie groß der Effekt ist, lässt sich schwer abschätzen, da auch bei den als Nicht-Underreporter eingestuften Personen relativ niedrige Energieaufnahmewerte zustande kommen und daher auch hier eine Unterschätzung der Aufnahmemengen vorliegt. Qualitativ betrachtet müssten aber alle Schätzungen weniger streng

genommen werden und auch Einschätzungen zu möglicher Unterversorgung eines Nährstoffes müssen weniger streng ausfallen als die geschätzten Werte wiedergeben. Somit wären nur noch Nährstoffe als kritisch einzustufen, welche wirklich sehr hohe Anteilsschätzungen liefern (z.B. Folsäure, Vitamin D).

	kcal	BMI	VA [mg]	VC [mg]	Kalzium [mg]
Normal	2163.8	24.31	1.22	129.4	912.4
Underreporter	1342.5	25.45	0.89	87.6	623.6
Unterschied [%]	46.9	-4.58	31.03	38.5	37.6

Tabelle 6.2: Vergleich der Underreporter mit “normalen“ Personen bezüglich Energie, Vitamin A, Vitamin C und Kalzium Aufnahme

Möglichkeiten zur “Hochrechnung“ der angegebenen Mengen existieren (Willett et al., 1997), werden allerdings noch diskutiert (Poslusna et al., 2009). Bei diesen Methoden wird versucht, anhand der Energieaufnahme und statistischen Modellen die angegebene Menge zu verbessern (engl. energy adjustment methods). Die Anwendung dieser auf die Ernährungsberichtsdaten liegt außerhalb des Ziels der Arbeit und wird deshalb hier nur ganz kurz diskutiert. Eine Möglichkeit der Energieadjustierung ist z.B. das Residual Model. Dabei wird eine einfache Regression mit Energieaufnahme als unabhängige und dem Mikronährstoff als abhängige Variable gerechnet. Die Abweichung (a) der Residuen zur Regressionsgeraden wird dann zu der - bezüglich mittlerer Energie - erwarteten Nährstoffaufnahme (b) hinzugefügt, um eine neue Schätzung der Nährstoffmenge ( $neu=a+b$ ) zu bekommen. Mit diesen Werten würde dann die übliche Aufnahmeverteilung berechnet werden. Eine andere Möglichkeit wäre die Nährstoffdichten zu betrachten ( $Nährstoffdichte = Nährstoffmenge/kcal$ ). Doch auch dies kommt nicht ohne Annahmen aus. Ebenso hat auch Lankester et al. (2014) zwei Methoden untersucht, um die Aufnahmedaten zu verbessern bzw. den Bias zu minimieren. Die Diskussion darüber, was am sinnvollsten wäre bzw. ob es überhaupt sinnvoll ist lässt sich auf jeden Fall nicht kurz halten und könnte eine eigene Arbeit füllen und wird deshalb hier nicht gegeben.

Eine andere Möglichkeit wäre die Schätzungen anhand von Biomarkerdaten (wie z.B. Urinstickstoff usw.) zu adjustieren. Hier gibt es z.B. einen Vorschlag von Yanetz et al. (2008) um die übliche Aufnahmeverteilung anhand einer modifizierten NRC Methode zu schätzen. Dabei werden extra Biomarker Daten von externen Studien verwendet um bezüglich Under- und Overreporting zu kontrollieren. Es gibt also durchaus Ansätze um den Underreporting entgegen zu treten um bessere Schätzungen zu bekommen. Ob sie sinnvoll sind und für den Ernährungsbericht verwendet werden sollten, kann nur durch diese Arbeit nicht schlussgefolgert werden. In Betracht gezogen kann und sollte es auf jeden Fall.

## 6.2 Diskussion bezüglich der statistischen Methoden

In dieser Arbeit wurden vor allem vier wesentliche Methoden zur Schätzung einer üblichen Aufnahmeverteilung vorgestellt. Diese haben alle eine sehr ähnliche Aufbaustruktur. Die Daten werden zuerst transformiert um approximative Normalität und dadurch besser Eigenschaften zu bekommen. Anhand der transformierten Daten wird dann die Varianz in eine between-person und within-person Varianz zerlegt. Bei den Methoden SPADE, MSM und NCI kann der between-person Teil auch noch mittels Kovariaten in einen erklärten und unexplorierten Varianzteil zerlegt werden. Mit der Kenntnis über die Varianzteile wird dann eine übliche Aufnahmeverteilung auf transformierter Skala geschätzt, welche anschließend zurück transformiert wird. Durch diesen ähnlichen Aufbau werden auch meist ähnliche Resultate erzielt. Dies bestätigte sich auch in diversen Simulationstudien, welche im Kapitel 4.2 erklärt worden sind. Auch die Anwendung dieser 4 Methoden auf die Ernährungsberichtsdaten zeigte dieses Bild. Zu erwähnen ist, dass die SPADE Methodik mit Gewichtungsfaktoren vermutlich leichte Probleme aufweist. Dies müsste allerdings in Simulationsstudien näher untersucht werden. Weiters kann geschlussfolgert werden, dass Unterschiede in den Ergebnissen der Methoden vor allem durch die Annahmeverletzungen auftreten. Solange die Box-Cox Transformation gute Resultate liefert, sind meist nur kleine Unterschiede in den Methoden zu erkennen. Sollte die “einfache“ Box-Cox Transformation nicht mehr ausreichen, so ist auf

jeden Fall die ISU Methode den anderen zu bevorzugen. Ihre zweistufige Transformation hat sich in dieser Arbeit bei keinem betrachteten Nährstoff als problematisch erwiesen. Jedoch sind auch multimodale Verteilungen (z.B. durch viele Nahrungsergänzungsmittel) auch bei der ISU Transformation vermutlich ein Problem. Hier kann vor allem die Methode SPADE mit ihren vielen Modellvarianten (1 bis 4 Part Modell) punkten. Sollten also Daten mit vielen hohen Werten vorhanden sein, so lohnt sich die SPADE Modellierung.

Ansonsten bleibt bezüglich den Methoden nur zu sagen, dass es von der untersuchten Fragestellung abhängt, welche Methode sinnvoller ist. Sind Aussagen bezüglich einzelner Altersklassen interessant, so hat SPADE hier den Vorteil des komplexesten Ansatzes (fraktionales Polynom oder einfacher linearer Trend). Sind hingegen Analysen bezüglich Subgruppen (z.B. BMI Klassen) relevant, so hat die NCI Methode einen klaren Vorteil gegenüber den anderen Methoden.

Die MSM Methode stellt die unsicherste Methode der 4 vorgestellten dar. Wie die Autoren auf der Homepage bekanntgegeben haben, kann die Verwendung der Kovariaten problematisch sein und zu schlechten unerwarteten Ergebnissen führen. Ebenso hat die MSM Methode keine Diagnostikplots und ermöglicht es auch nicht solche zu generieren. Einzig die sehr einfache Bedienung der Methode kommt ihr im Vergleich zu den anderen Methoden zu gute.

## Schätzung von anderen Verteilungen

Prinzipiell sind die besprochenen Methoden nicht nur für die Schätzung von Nahrungsaufnahmedaten interessant. Sie können auch für andere Fragestellungen, wo die Verteilung von Langzeitwerten (z.B. Erwartungswerte) je Person von Interesse ist und wiederholte Messungen verfügbar sind, verwendet werden. Ein Beispiel sind die Blutmarkerdaten. Wenn es sich dabei um Marker handelt, welche täglich oder sogar stündlich stark variieren, so können wiederholte Messungen und die anschließende Modellierung mittels der besprochenen Methoden durchaus sehr sinnvoll sein. Die Modelle können auch durchaus in anderen Fachbereichen verwendet werden.

## weitere Untersuchungen

In der Literatur zu den Methoden sind bis jetzt noch nicht viele Vergleiche bzw. wirklich große Untersuchungen zu finden. Weitere Überprüfungen bezüglich der Verwendung von Gewichtungsfaktoren wäre auf jeden Fall nützlich. Ebenfalls wären weiters die Auswirkungen der Annahmeverletzungen auf die Ergebnisse interessant. Dies müsste allerdings noch genau untersucht werden.

## 6.3 Diskussion bezüglich der Neuauswertung

Kommen wir nun zur Zusammenfassung der Neuauswertung der Daten des Ernährungsberichtes von 2012. Die Erkenntnisse, die aus dieser Arbeit gezogen werden können sind vor allem, dass die EAR Referenzwerte weitaus weniger problematische Aussagen über den Versorgungsstatus der österreichischen Bevölkerung bringen als die DACH Werte. Unterschiede von wenigen Prozenten bis hin zu 75% (Eisen) sind in dieser Arbeit der Fall. Mit 10 bis 30% kann im Mittel gerechnet werden. Es können also durchaus andere Aussagen und Empfehlungen für die österreichische Bevölkerung zustande kommen, wenn die DACH Referenzwerte verwendet werden. Bezüglich der Diskussion und Erläuterung, warum die EAR Werte sinnvoller als die DACH Referenzwerte - für die Anteilsschätzungen der Unterversorgten - sind, siehe Kapitel 2.6.1.

Als nächstes darf der Unterschied der durch die Mittelwertmethodik, welche im Ernährungsbericht verwendet wurde, zustande kommt nicht unerwähnt bleiben. Auch hier sind 10-20% Unterschiede in den entstehenden Anteilsschätzungen der Unterversorgten keine Seltenheit und es empfiehlt sich daher nicht die Verteilung der einfachen Mittelwerte für Aussagen heranzuziehen. Siehe dazu auch Tabelle 5.12 und Tabelle 5.13.

## Verletzung der Annahmen

Wenn die Nährstoffe bezüglich der 4 Hauptannahmen überprüft werden, so zeigt sich vor allem das Problem der verzerrten Daten als Hauptfehlerquelle. Vermutlich müssten alle Anteilsschätzungen nach unten korrigiert werden, da die Aufnahme meist unterschätzt wurde. Die anderen Annahmen sind für die meisten Nährstoffe kein Problem und wenn doch, dann können diese mit der ISU Methode gut eingehalten werden.

## kritische Nährstoffe?

Wenn Aussagen darüber gemacht werden sollen, welche Nährstoffe in Österreich bezüglich der Nahrungsaufnahme wirklich problematisch sind, so muss doch klar festgestellt werden, dass dies alles andere als sehr sicher ist. Die Unsicherheit in den Daten bezüglich einer Verzerrung sind doch sehr groß. Auch die Schlussfolgerung die man durch die beobachteten Bootstrap Konfidenzintervalle ziehen kann sind alles andere als sehr gut. Die meisten Abschätzungen der Unsicherheiten bezüglich der Anteile sind relativ groß. Es ist nicht unüblich, dass die Konfidenzintervalle um die 10-20%Punkte breit sind. Betrachtet man die Varianzverhältnisse so wird klar, dass die in Relation gesehen hohe within-person Varianz oft das Problem für diese groben Konfidenzintervalle darstellt. Es sind weitaus mehr Daten notwendig um präzisere Schätzungen zu bekommen. 418 Erwachsene, die dann noch nach Geschlecht und Alterklassen geteilt werden, sind relativ wenig um genaue Schätzungen auf Österreichebene zu bekommen. Möchte man trotzdem anhand dieser Daten Aussagen über den Versorgungszustand der Bevölkerung machen, so können vor allem Vitamin B9 und Vitamin D genannt werden (vermutlich auch die Pantothensäure=Vitamin B5). Jedoch muss hier noch erwähnt werden, dass keine Nahrungsergänzungsmittel berücksichtigt wurden. Die Anteilsschätzungen sind also nicht nur bezüglich den verzerrten Daten zu hoch, sondern auch bezüglich den nicht berücksichtigten Nahrungsergänzungsmitteln. Immerhin 113 Personen (= 27%) haben bei den Befragungen angegeben, dass sie zur Zeit Supplemente verwenden. Es kann also gut sein, dass z.B. der Anteil der Personen, welche zu wenig Vitamin D über die Nahrung aufnehmen, weitaus geringer wäre, wenn alle benutzten Vitamin D Präparate miteinbezogen werden.

## 6.4 Schlussfolgerungen

Die Arbeit hat mit folgender Behauptung begonnen:

47,1% der Frauen und 19,8% der Männer haben, laut Ernährungsbericht 2012, eine “deutlich erniedrigte“ Vitamin B12 Aufnahme.

Darauf haben wir Fragen gestellt wie: Wie kommen diese Anteile zustanden? Wie genau sind diese Schätzungen? Haben nun 20-47% der Österreicher Probleme mit der Vitamin B12 Versorgung? Anhand des theoretischen Teils dieser Arbeit, sowie der Neuauswertung, wissen wir nun wie wir diese Fragen beantworten können. Nicht nur, dass die Anteile verzerrt sind (wegen der verwendeten statistischen Methodik und der Datenqualität), sie sind auch relativ ungenau (Konfidenzintervalle von 10-20%). Wenn also anhand dieser Arbeit nun Schlussfolgerungen beziehungsweise Empfehlungen erwähnt werden sollen, dann vor allem die folgenden.

Für zukünftige Untersuchungen von Nahrungsaufnahmedaten, sollten die in dieser Arbeit erwähnten Methoden in Betracht gezogen werden. Dies gilt jedoch nicht nur für Aufnahmedaten, sondern auch für die Blutmarkerdaten, welche eine entsprechend hohe Variabilität in den Messungen aufweisen. Auf das Thema, welche diese genau sind wird hier nicht weiter eingegangen.

Sollten Anteilsschätzungen zum Versorgungstatus eines Nährstoffes für eine Gruppe von Interesse sein, so sollten die entsprechenden richtigen Grenzwerte verwendet werden. Für Aufnahmedaten sind dies mittlere bzw. mediane Bedarfsgrenzen wie die Estimated Average Requirements (EAR).

Für präzise Schätzungen wäre eine größere Stichprobengröße für Österreich von Vorteil. Die Konfidenzintervalle zeigen vor allem auf, dass die Schätzungen noch viel zu stark schwanken können, wenn so wenig Daten vorhanden sind. Wie groß genau die Stichproben sein müssen, ist schwer zu sagen. Verwendet man z.B. die Daumenregel “Stichprobengröße vervierfachen



um das Konfidenzintervall zu halbieren“ so würde ungefähr eine Stichprobengröße von 800 Männern und 800 Frauen notwendig sein um 5-10% Konfidenzintervallbreite zu bekommen. Dies ist allerdings nur eine grobe Angabe, da die verwendeten statistischen Modelle weitaus komplexer sind und die Daumenregel nur für einfache Konfidenzintervalle gilt. In Volatier et al. (2002) wurden Stichproben der Größe 2000 Erwachsene pro europäisches Land empfohlen (Ernährungsbericht 2012 n= 417), um “Trends in der Aufnahme von Nährstoffen und Nahrungsgruppen zu identifizieren“.

Wie wir im Anfangsabschnitt dieses Kapitels sehen konnten, sind die Daten sehr wahrscheinlich deutlich nach unten verzerrt. Es wäre also auch durchaus wünschenswert für zukünftige Ernährungsberichte eine Validierung mittels Biomarker durchzuführen. Ebenso sollten durchaus Modellansätze, welche diesen bekannten Fehler zu minimieren versuchen in Betracht gezogen werden. Womöglich können damit deutlich aussagekräftigere Empfehlungen für die österreichische Bevölkerung erfolgen.

Welche der 4 erwähnten Hauptmethoden allgemein gesehen die beste ist, kann mittels dieser Arbeit nicht geschlussfolgert werden. Jedoch kann eine Empfehlung für die zukünftigen Ernährungsberichte abgegeben werden. Vor allem die SPADE Methode ist hier zu nennen, da sie für verschiedene Altersklassen, welche im Ernährungsbericht immer wieder von Interesse sind, Schätzungen ermöglicht. Dabei können auch Anteilsschätzungen für die verschiedenen Altersklassen (was hier in dieser Arbeit nicht gemacht wurde) separat angegeben werden. Es können also mit entsprechender Stichprobengröße auch Aussagen darüber gemacht werden, welche Altersklasse vermutlich an welchem Nährstoff Mängel vorweisen könnte.

Zuletzt bleibt noch zu erwähnen, dass auch bei Verwendung von Nahrungsergänzungsmitteldaten die SPADE Methode durchaus Modellierungsansätze bietet und diese auch für die Aussagen im Ernährungsbericht verwendet werden sollten um ein besseres Verständnis zu bekommen, welche Nährstoffe wirklich kritisch sind und welche nicht.

# Literaturverzeichnis

- Anderson, T. W. and Darling, D. A. (1952). Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics*, 23(2):193–212.
- Barrett-Connor, E. (1991). Nutrition epidemiology: how do we know what they ate? *The American Journal of Clinical Nutrition*, 54(1):182S–187S.
- Beaton, G. H. (1994). Criteria of an adequate diet. In Shils ME, Olson JA, Shike M, editor, *Modern Nutrition in Health and Disease*, pages 1491 – 1505. Lea & Febiger, Philadelphia, 8 edition.
- Bingham, S. A. (2003). Urine Nitrogen as a Biomarker for the Validation of Dietary Protein Intake. *The Journal of Nutrition*, 133(3):921S–924S.
- Black, A. E. (2000a). Critical evaluation of energy intake using the Goldberg cut-off for energy intake: basal metabolic rate. A practical guide to its calculation, use and limitations. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*, 24(9):1119–1130.
- Black, A. E. (2000b). The sensitivity and specificity of the Goldberg cut-off for EI:BMR for identifying diet reports of poor validity. *European Journal of Clinical Nutrition*, 54(5):395–404.
- Box, G. E. P. and Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252.
- Brussaard, J. H., Löwik, M. R. H., Steingrimsdóttir, L., Møller, A., Kearney, J., De Henauw, S., Becker, W., and EFCOSUM Group (2002). A European food consumption survey method—conclusions and recommendations. *European Journal of Clinical Nutrition*, 56 Suppl 2:S89–94.
- Carriquiry, A. L. (1999). Assessing the prevalence of nutrient inadequacy. *Public health nutrition*, 2(01):23–34.
- Carriquiry, A. L., Fuller, W. A., Goyeneche, J., and Dodd, K. W. (1995). *Estimation of the usual intake distributions of ratios of dietary components*. Center for Agricultural and Rural Development, Iowa State University.
- Crispim, S. P., de Vries, J. H. M., Geelen, A., Souverein, O. W., Hulshof, P. J. M., Lafay, L., Rousseau, A.-S., Lillegaard, I. T. L., Andersen, L. F., Huybrechts, I., De Keyzer, W., Ruprich, J., Dofkova, M., Ocke, M. C., de Boer, E., Slimani, N., and van't Veer, P. (2011). Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres—results from the European Food Consumption Validation (EFCOVAL) study. *The British Journal of Nutrition*, 105(3):447–458.
- de Lauzon, B., Volatier, J. L., and Martin, A. (2004). A Monte Carlo simulation to validate the EAR cut-point method for assessing the prevalence of nutrient inadequacy at the population level. *Public Health Nutrition*, 7(7):893–900.

- Dekkers, A. L. M. and Slob, W. (2012). Gaussian Quadrature is an efficient method for the back-transformation in estimating the usual intake distribution when assessing dietary exposure. *Food and Chemical Toxicology: An International Journal Published for the British Industrial Biological Research Association*, 50(10):3853–3861.
- Dekkers, A. L. M., Verkaik-Kloosterman, J., van Rossum, C. T. M., and Ocké, M. C. (2014). SPADE, a new statistical program to estimate habitual dietary intake from multiple food sources and dietary supplements. *The Journal of Nutrition*, 144(12):2083–2091.
- Deutsche Gesellschaft für Ernährung, Österreichische Gesellschaft für Ernährung, Schweizerische Gesellschaft für Ernährungsforschung, and Schweizerische Vereinigung für Ernährung, editors (2015). *Referenzwerte für die Nährstoffzufuhr*, volume 1. Bonn, 2 edition.
- Dodd, K. W., Guenther, P. M., Freedman, L. S., Subar, A. F., Kipnis, V., Midthune, D., Tooze, J. A., and Krebs-Smith, S. M. (2006). Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *Journal of the American Dietetic Association*, 106(10):1640–1650.
- Elmadfa, I., Hasenegger, V., Wagner, K., Putz, P., Weidl, N.-M., Wottawa, D., Kuen, T., Seiringer, G., Meyer, A. L., Sturtzel, B., Kiefer, I., Zilberszac, A., Sgarabottolo, V., Meidlinger, B., and Rieder, A. (2012). *Österreichischer Ernährungsbericht 2012*. Wien, 1 edition.
- Freedman, L. S., Guenther, P. M., Dodd, K. W., Krebs-Smith, S. M., and Midthune, D. (2010). The Population Distribution of Ratios of Usual Intakes of Dietary Components That Are Consumed Every Day Can Be Estimated from Repeated 24-Hour Recalls. *The Journal of Nutrition*, 140(1):111–116.
- Fuller, W. A. (2009). *Introduction to Statistical Time Series*. John Wiley & Sons.
- Gedrich, K. and Karg, G. (2001). Dietary habits of German vs. non-German residents in Germany. In Edwards JSA and Hewedi MM, editors, *Culinary Arts and Sciences III: Global and National Perspectives*, pages 419–428. Al-Karma Press.
- Goldberg, G. R., Black, A. E., Jebb, S. A., Cole, T. J., Murgatroyd, P. R., Coward, W. A., and Prentice, A. M. (1991). Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. *European Journal of Clinical Nutrition*, 45(12):569–581.
- Harttig, U., Haubrock, J., Knüppel, S., Boeing, H., and EFCOVAL Consortium (2011). The MSM program: web-based statistics package for estimating usual dietary intake using the Multiple Source Method. *European Journal of Clinical Nutrition*, 65 Suppl 1:S87–91.
- Haubrock, J., Nöthlings, U., Volatier, J.-L., Dekkers, A., Ocké, M., Harttig, U., Illner, A.-K., Knüppel, S., Andersen, L. F., Boeing, H., and European Food Consumption Validation Consortium (2011). Estimating usual food intake distributions by using the multiple source method in the EPIC-Potsdam Calibration Study. *The Journal of Nutrition*, 141(5):914–920.
- Hoffmann, K., Boeing, H., Dufour, A., Volatier, J. L., Telman, J., Virtanen, M., Becker, W., De Henauw, S., and EFCOSUM Group (2002). Estimating the distribution of usual dietary intake by short-term measurements. *European Journal of Clinical Nutrition*, 56 Suppl 2:S53–62.
- Kwaśniewska, M., Kaleta, D., Dziańkowska-Zaborszczyk, E., Drygas, W., and Makowiec-Dabrowska, T. (2007). Lifestyle index and self-rated health status. *International Journal of Occupational Medicine and Environmental Health*, 20(4):349–356.
- Lankester, J., Perry, S., and Parsonnet, J. (2014). Comparison of Two Methods – Regression Predictive Model and Intake Shift Model – For Adjusting Self-Reported Dietary Recall of Total Energy Intake of Populations. *Frontiers in Public Health*, 2.

- Löwik, M. R., Hulshof, K. F., and Brussaard, J. H. (1999). Food-based dietary guidelines: some assumptions tested for The Netherlands. *The British Journal of Nutrition*, 81 Suppl 2:S143–149.
- National Research Council (1986). *Nutrient Adequacy: Assessment Using Food Consumption Surveys*. National Academy Press, Washington, DC.
- National Research Council (2000). *Dietary Reference Intakes: Applications in Dietary Assessment*. The National Academies Press, Washington, DC.
- National Research Council (2006). *Dietary Reference Intakes: The Essential Guide to Nutrient Requirements*. The National Academies Press, Washington, DC.
- Nusser, S. M., Carriquiry, A. L., Dodd, K. W., and Fuller, W. A. (1996). A Semiparametric Transformation Approach to Estimating Usual Daily Intake Distributions. *Journal of the American Statistical Association*, 91(436):1440–1449.
- Nusser, S. M., Fuller, W. A., and Guenther, P. M. (1997). Estimating Usual Dietary Intake Distributions: Adjusting for Measurement Error and Nonnormality in 24-Hour Food Intake Data. In Lyberg, L., Biemer, P., Collins, r., Leeuw, E. D., Dippo, C., Schwarz, N., and Trewin, D., editors, *Survey Measurement and Process Quality*, pages 689–709. John Wiley & Sons, Inc.
- Poslusna, K., Ruprich, J., de Vries, J. H. M., Jakubikova, M., and van't Veer, P. (2009). Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. *The British Journal of Nutrition*, 101 Suppl 2:S73–85.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Razali, N. M. and Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, 2(1):21–33.
- Royston, P. and Sauerbrei, W. (2008). *Multivariable Model - Building: A Pragmatic Approach to Regression Analysis based on Fractional Polynomials for Modelling Continuous Variables*. John Wiley & Sons.
- RStudio Team (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.
- SAS Institute Inc (2015). SAS® University Edition: Installation Guide for Windows.
- Schek, A. (2011). *Ernährungslehre kompakt: Kompendium der Ernährungslehre für Studierende der Ernährungswissenschaft, Medizin und Naturwissenschaften und zur Ausbildung von Ernährungsfachkräften*. Umschau Zeitschriftenverlag, 4., aktualisierte und ergänzte auflage edition.
- Schnell, R., Hill, P. B., and Esser, E. (2005). *Methoden der empirischen Sozialforschung*. Oldenbourg Verlag.
- Schofield, W. N. (1985). Predicting basal metabolic rate, new standards and review of previous work. *Human Nutrition. Clinical Nutrition*, 39 Suppl 1:5–41.
- Sovereign, O. W., Dekkers, A. L., Geelen, A., Haubrock, J., de Vries, J. H., Ocké, M. C., Harttig, U., Boeing, H., van 't Veer, P., and EFCOVAL Consortium (2011). Comparing four methods to estimate usual intake distributions. *European Journal of Clinical Nutrition*, 65 Suppl 1:S92–101.
- Subar, A. F., Kipnis, V., Troiano, R. P., Midthune, D., Schoeller, D. A., Bingham, S., Shambaugh, C. O., Trabulsi, J., Runswick, S., Ballard-Barbash, R., Sunshine, J., and Schatzkin, A. (2003). Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *American Journal of Epidemiology*, 158(1):1–13.

- Tooze, J. A., Kipnis, V., Buckman, D. W., Carroll, R. J., Freedman, L. S., Guenther, P. M., Krebs-Smith, S. M., Subar, A. F., and Dodd, K. W. (2010). A mixed-effects model approach for estimating the distribution of usual intake of nutrients: the NCI method. *Statistics in Medicine*, 29(27):2857–2868.
- Tooze, J. A., Midthune, D., Dodd, K. W., Freedman, L. S., Krebs-Smith, S. M., Subar, A. F., Guenther, P. M., Carroll, R. J., and Kipnis, V. (2006). A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *Journal of the American Dietetic Association*, 106(10):1575–1587.
- Volatier, J. L., Turrini, A., Welten, D., and EFCOSUM Group (2002). Some statistical aspects of food intake assessment. *European Journal of Clinical Nutrition*, 56 Suppl 2:S46–52.
- Willett, W. C., Howe, G. R., and Kushi, L. H. (1997). Adjustment for total energy intake in epidemiologic studies. *The American Journal of Clinical Nutrition*, 65(4 Suppl):1220S–1228S; discussion 1229S–1231S.
- Yanetz, R., Kipnis, V., Carroll, R. J., Dodd, K. W., Subar, A. F., Schatzkin, A., and Freedman, L. S. (2008). Using biomarker data to adjust estimates of the distribution of usual intakes for misreporting: application to energy intake in the US population. *Journal of the American Dietetic Association*, 108(3):455–464; discussion 464.

# Abbildungsverzeichnis

1.1	Nährstoffbedarfsbestimmung . . . . .	6
2.1	Aufnahmeverteilung Gruppe vs Individuen . . . . .	13
2.2	tatsächliche vs geschätzte Aufnahmedichte . . . . .	14
2.3	Gemeinsame Verteilung von Bedarf und Aufnahme . . . . .	38
3.1	Nutrient Quality Index . . . . .	44
3.2	Dropoutanalyse . . . . .	45
4.1	Vergleich der Hauptmethoden und Mittelwertmethodik anhand von Dichteabbildungen . . . . .	54
5.1	Vitamin B9 Aufnahmemodellierung für Männer. Geschätztes fraktionales Polynom bezüglich Alter. . . . .	69
6.1	Histogramm der geschätzten Energieaufnahme/Grundumsatz Quotienten . . . . .	83

# Tabellenverzeichnis

2.1	Notation . . . . .	13
2.2	Simulationsstudie: Gammaverteilungsparameter (shape,scale), $\chi^2$ ...Chi-Quadrat Verteilung. Als Referenzwerte für die Anteile wurden für die Normalverteilungssimulationen 80 und für die Gamma Verteilungen 0.8 gewählt . . . . .	15
2.3	Inhaltliche Gegenüberstellung der Methoden SPADE, NCI, ISU und MSM . . . . .	35
3.1	Körperabmessungen: Die Angaben entsprechen den gewichteten Mittelwerten und sind nur für abgemessene Personen angegeben. . . . .	41
3.2	BMI Vergleich der Gruppen mit einem und zwei Befragungen (nur geschätzte Körpermaße der Personen). Nicht gewichtete Daten. . . . .	42
3.3	Lifestyle Index . . . . .	42
3.4	Food Based Quality Index . . . . .	43
3.5	Vergleich der medianen Energie- und Makronährstoffaufnahme anhand der Befragungen (ungewichtet). n=110 für 1 Recall und n=308 für 2 Recalls . . . . .	46
3.6	Zeitpunktvergleiche der Nährstoffe und Energie. Ungewichtete Daten mit n=308 . . . . .	47
4.1	Vergleich der ursprünglichen Daten bezüglich Kalorienaufnahme [kcal] mit den neuen "gleichgewichteten" Daten . . . . .	50
4.2	Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik (WPM) bezüglich Kalorienaufnahme [kcal] . . . . .	51
4.3	Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Eiweißaufnahme [g] . . . . .	52
4.4	Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Vitamin A Aufnahme [g] . . . . .	52
4.5	Vergleich der Mittelwerte bezüglich Altersgruppen und Geschlecht für die Energieaufnahme [kcal] . . . . .	55
5.1	Vergleich der Kalorienaufnahme [kcal] von Männer (n=148) und Frauen (n=232) sowie von Normal (n=181) - und Übergewichtigen (+ Adipösen) (n=92) . . . . .	60
5.2	Energieaufnahmevergleiche [kcal] zwischen Ernährungsbericht und SPADE bezüglich den Alterklassen . . . . .	60
5.3	Energieprozentvergleiche [%] für Kohlenhydrate zwischen Ernährungsbericht und SPADE bezüglich den Altersklassen . . . . .	61
5.4	Vergleich der Eiweißaufnahme [g] von Männer (mit und ohne Ausreißer, n=148) und Frauen (n=232). Für Männer wird zusätzlich der relative Unterschied angegeben. . . . .	62
5.5	Energieprozentvergleiche [%] für Eiweiß zwischen Ernährungsbericht und SPADE bezüglich den Alterklassen . . . . .	63
5.6	Vergleich der SPADE und ISU Methode bezüglich der Aufnahme von Pantothenensäure [mg] . . . . .	66
5.7	Vergleich der SPADE und ISU Methode bezüglich Vitamin B6 [mg]. Oben Männer, unten Frauen . . . . .	67
5.8	Mittlere Aufnahmevergleiche von Biotin [ $\mu$ g] zwischen Ernährungsbericht und SPADE . . . . .	68
5.9	Mittlere Aufnahmevergleiche von Folatäquivalenten [ $\mu$ g] zwischen Ernährungsbericht und SPADE . . . . .	69
5.10	Aufnahmevergleiche von Vitamin C [mg] zwischen Ernährungsbericht und SPADE . . . . .	71

5.11	Vergleich der SPADE und ISU Methode bezüglich der Vitamin D Aufnahme [ $\mu\text{g}$ ]. Oben Männer, unten Frauen . . . . .	72
5.12	Anteilsschätzungen der Unterversorgten bezüglich der Vitaminversorgung. Vergleich zwischen den DACH und EAR Referenzwerten, ebenso wie der Vergleich zu den Ergebnissen im Ernährungsbericht 2012. Bemerkungen: In Klammern stehen die Schätzungen der ISU Methode. SW steht für den DACH Schätzwert. Der Stern (*) in der letzten Spalte bedeutet, dass das Ergebnis auch im Ernährungsbericht veröffentlicht wurde. . . . .	75
5.13	Anteilsschätzungen der Mineralstoffversorgung. Vergleich zwischen den DACH und EAR Referenzwerten, ebenso wie der Vergleich zu den Ergebnissen im Ernährungsbericht 2012. Der Stern (*) in der letzten Spalte bedeutet, dass das Ergebnis auch im Ernährungsbericht veröffentlicht wurde. . . . .	79
6.1	Summary Statistik zum Regressionsmodell bezüglich Grundumsatzschätzung	82
6.2	Vergleich der Underreporter mit "normalen" Personen bezüglich Energie, Vitamin A, Vitamin C und Kalzium Aufnahme . . . . .	84
A.1	Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Kohlenhydrataufnahme [g] . . . . .	96
A.2	Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Fettaufnahme [g] . . . . .	96
A.3	Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Kalziumaufnahme [g] . . . . .	96
A.4	Vergleich der Mittelwerte bezüglich Altersgruppen und Geschlecht für die Eiweißaufnahme [g] . . . . .	97
A.5	Vergleich der Mittelwerte bezüglich Altersgruppen und Geschlecht für die Kohlenhydrataufnahme [g] . . . . .	97
A.6	Energieprozentvergleiche für Fett zwischen Ernährungsbericht und SPADE . . . . .	97
C.1	DACH Referenzwerte . . . . .	107



# Appendices

# Anhang A

## Tabellen

	Mittel	5%	25%	50%	75%	95%	var.ratio
WPM	229.5	131.7	174.8	223.4	271.5	351.2	1.67
ISU	220.6	136.0	181.5	217.1	255.9	317.2	1.39
MSM	221.5	138.9	180.3	220.2	253.8	319.1	1.40
SPADE	217.8	132.9	177.9	213.7	253.1	316.5	1.29
NCI	220.6	136.9	181.4	217.2	255.6	317.0	1.47
SPADE gew.	221.0	136.4	181.8	217.4	256.3	318.3	1.40

Tabelle A.1: Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Kohlenhydrataufnahme [g]

	Mittel	5%	25%	50%	75%	95%	var.ratio
WPM	82.4	41.7	62.0	78.7	100.8	135	2.58
ISU	79.3	50.8	66.1	78.1	91.1	112	2.75
MSM	79.6	51.0	66.4	78.6	90.9	112	2.68
SPADE	78.3	47.5	63.9	76.8	91.2	114	2.15
NCI	79.2	50.3	65.7	78.0	91.2	112	2.72
SPADE gew.	79.2	50.2	65.8	78.0	91.3	113	2.62

Tabelle A.2: Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Fettaufnahme [g]

	Mittel	5%	25%	50%	75%	95%	var.ratio
WPM	859	372	579	806	1067	1519	2.82
ISU	830	497	664	805	968	1248	2.61
MSM	835	488	648	803	969	1246	2.30
SPADE	817	472	643	789	960	1256	2.16
NCI	818	499	658	795	951	1219	2.88
SPADE gew.	819	498	660	795	952	1220	2.83

Tabelle A.3: Vergleich der Methoden mit gewöhnlicher Mittelwertmethodik bezüglich Kalziumaufnahme [g]

	Männlich			Weiblich		
	[18,24]	[25,50]	[51,64]	[18,24]	[25,50]	[51,64]
WPM	91	80	81	68	66	61
SPADE	85	83	80	67	64	61
NCI	85	81	78	69	65	63
SPADE gew.	88	86	82	67	66	63
n	18	99	51	40	154	56

Tabelle A.4: Vergleich der Mittelwerte bezüglich Altersgruppen und Geschlecht für die Eiweraufnahme [g]

	Männlich			Weiblich		
	[18,24]	[25,50]	[51,64]	[18,24]	[25,50]	[51,64]
WPM	274	226	224	221	205	211
SPADE	246	235	221	210	207	203
NCI	244	233	223	222	210	201
SPADE gew.	252	240	225	213	211	208
n	18	99	51	40	154	56

Tabelle A.5: Vergleich der Mittelwerte bezüglich Altersgruppen und Geschlecht für die Kohlenhydrataufnahme [g]

	[18-24]	[25-50]	[51-64]
SPADE Männer	36 [32,39]	37 [36,39]	39 [36,41]
EB Männer	34 [32,36]	37 [36,38]	37 [35,39]
n	17	87	44
SPADE Frauen	36 [34,37]	37 [36,38]	37 [35,39]
EB Frauen	36 [33,38]	36 [35,37]	36 [34,37]
n	37	143	52

Tabelle A.6: Energieprozentvergleiche für Fett zwischen Ernährungsbericht und SPADE

Anhang B

## Dietary Reference Intakes

## Dietary Reference Intakes (DRIs): Estimated Average Requirements Food and Nutrition Board, Institute of Medicine, National Academies

Life Stage Group	Calcium (mg/d)	CHO (g/d)	Protein (g/kg/d)	Vit A (µg/d) <sup>e</sup>	Vit C (mg/d)	Vit D (µg/d)	Vit E (mg/d) <sup>f</sup>	Thiamin (mg/d)	Riboflavin (mg/d)	Niacin (mg/d) <sup>g</sup>	Vit B <sub>6</sub> (mg/d)	Folate (µg/d) <sup>d</sup>	Vit B <sub>12</sub> (µg/d)	Copper (µg/d)	Iodine (µg/d)	Iron (mg/d)	Magnesium (mg/d)	Molybdenum (µg/d)	Phosphorus (mg/d)	Selenium (µg/d)	Zinc (mg/d)	
<b>Infants</b>																						
0 to 6 mo																						
6 to 12 mo			1.0													6.9						2.5
<b>Children</b>																						
1–3 y	500	100	0.87	210	13	10	5	0.4	0.4	5	0.4	120	0.7	260	65	3.0	65	13	380	17	2.5	
4–8 y	800	100	0.76	275	22	10	6	0.5	0.5	6	0.5	160	1.0	340	65	4.1	110	17	405	23	4.0	
<b>Males</b>																						
9–13 y	1,100	100	0.76	445	39	10	9	0.7	0.8	9	0.8	250	1.5	540	73	5.9	200	26	1,055	35	7.0	
14–18 y	1,100	100	0.73	630	63	10	12	1.0	1.1	12	1.1	330	2.0	685	95	7.7	340	33	1,055	45	8.5	
19–30 y	800	100	0.66	625	75	10	12	1.0	1.1	12	1.1	320	2.0	700	95	6	330	34	580	45	9.4	
31–50 y	800	100	0.66	625	75	10	12	1.0	1.1	12	1.1	320	2.0	700	95	6	350	34	580	45	9.4	
51–70 y	800	100	0.66	625	75	10	12	1.0	1.1	12	1.4	320	2.0	700	95	6	350	34	580	45	9.4	
> 70 y	1,000	100	0.66	625	75	10	12	1.0	1.1	12	1.4	320	2.0	700	95	6	350	34	580	45	9.4	
<b>Females</b>																						
9–13 y	1,100	100	0.76	420	39	10	9	0.7	0.8	9	0.8	250	1.5	540	73	5.7	200	26	1,055	35	7.0	
14–18 y	1,100	100	0.71	485	56	10	12	0.9	0.9	11	1.0	330	2.0	685	95	7.9	300	33	1,055	45	7.3	
19–30 y	800	100	0.66	500	60	10	12	0.9	0.9	11	1.1	320	2.0	700	95	8.1	255	34	580	45	6.8	
31–50 y	800	100	0.66	500	60	10	12	0.9	0.9	11	1.1	320	2.0	700	95	8.1	265	34	580	45	6.8	
51–70 y	1,000	100	0.66	500	60	10	12	0.9	0.9	11	1.3	320	2.0	700	95	5	265	34	580	45	6.8	
> 70 y	1,000	100	0.66	500	60	10	12	0.9	0.9	11	1.3	320	2.0	700	95	5	265	34	580	45	6.8	
<b>Pregnancy</b>																						
14–18 y	1,000	135	0.88	530	66	10	12	1.2	1.2	14	1.6	520	2.2	785	160	23	335	40	1,055	49	10.5	
19–30 y	800	135	0.88	550	70	10	12	1.2	1.2	14	1.6	520	2.2	800	160	22	290	40	580	49	9.5	
31–50 y	800	135	0.88	550	70	10	12	1.2	1.2	14	1.6	520	2.2	800	160	22	300	40	580	49	9.5	
<b>Lactation</b>																						
14–18 y	1,000	160	1.05	885	96	10	16	1.2	1.3	13	1.7	450	2.4	985	209	7	300	35	1,055	59	10.9	
19–30 y	800	160	1.05	900	100	10	16	1.2	1.3	13	1.7	450	2.4	1,000	209	6.5	255	36	580	59	10.4	
31–50 y	800	160	1.05	900	100	10	16	1.2	1.3	13	1.7	450	2.4	1,000	209	6.5	265	36	580	59	10.4	

**NOTE:** An Estimated Average Requirement (EAR) is the average daily nutrient intake level estimated to meet the requirements of half of the healthy individuals in a group. EARs have not been established for vitamin K, pantothenic acid, biotin, choline, chromium, fluoride, manganese, or other nutrients not yet evaluated via the DRI process.

<sup>a</sup> As retinol activity equivalents (RAEs). 1 RAE = 1 µg retinol, 12 µg β-carotene, 24 µg α-carotene, or 24 µg β-cryptoxanthin. The RAE for dietary provitamin A carotenoids is two-fold greater than retinol equivalents (RE), whereas the RAE for preformed vitamin A is the same as RE.

<sup>b</sup> As α-tocopherol. α-Tocopherol includes *RRR*-α-tocopherol, the only form of α-tocopherol that occurs naturally in foods, and the *2R*-stereoisomeric forms of α-tocopherol (*RRR*-, *RSR*-, *RRS*-, and *RSS*-α-tocopherol) that occur in fortified foods and supplements. It does not include the *2S*-stereoisomeric forms of α-tocopherol (*SRR*-, *SSR*-, *SRS*-, and *SSS*-α-tocopherol), also found in fortified foods and supplements.

<sup>c</sup> As niacin equivalents (NE). 1 mg of niacin = 60 mg of tryptophan.

<sup>d</sup> As dietary folate equivalents (DFE). 1 DFE = 1 µg food folate = 0.6 µg of folic acid from fortified food or as a supplement consumed with food = 0.5 µg of a supplement taken on an empty stomach.

**SOURCES:** *Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride* (1997); *Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B<sub>6</sub>, Folate, Vitamin B<sub>12</sub>, Pantothenic Acid, Biotin, and Choline* (1998); *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids* (2000); *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc* (2001); *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids* (2002/2005); and *Dietary Reference Intakes for Calcium and Vitamin D* (2011). These reports may be accessed via [www.nap.edu](http://www.nap.edu).

## Dietary Reference Intakes (DRIs): Recommended Dietary Allowances and Adequate Intakes, Vitamins Food and Nutrition Board, Institute of Medicine, National Academies

Life Stage Group	Vitamin A (µg/d) <sup>e</sup>	Vitamin C (mg/d)	Vitamin D (µg/d) <sup>g,c</sup>	Vitamin E (mg/d) <sup>d</sup>	Vitamin K (µg/d)	Thiamin (mg/d)	Riboflavin (mg/d)	Niacin (mg/d) <sup>e</sup>	Vitamin B <sub>6</sub> (mg/d)	Folate (µg/d) <sup>f</sup>	Vitamin B <sub>12</sub> (µg/d)	Pantothenic Acid (mg/d)	Biotin (µg/d)	Choline (mg/d) <sup>e</sup>
<b>Infants</b>														
0 to 6 mo	400*	40*	10	4*	2.0*	0.2*	0.3*	2*	0.1*	65*	0.4*	1.7*	5*	125*
6 to 12 mo	500*	50*	10	5*	2.5*	0.3*	0.4*	4*	0.3*	80*	0.5*	1.8*	6*	150*
<b>Children</b>														
1–3 y	300	15	15	6	30*	0.5	0.5	6	0.5	150	0.9	2*	8*	200*
4–8 y	400	25	15	7	55*	0.6	0.6	8	0.6	200	1.2	3*	12*	250*
<b>Males</b>														
9–13 y	600	45	15	11	60*	0.9	0.9	12	1.0	300	1.8	4*	20*	375*
14–18 y	900	75	15	15	75*	1.2	1.3	16	1.3	400	2.4	5*	25*	550*
19–30 y	900	90	15	15	120*	1.2	1.3	16	1.3	400	2.4	5*	30*	550*
31–50 y	900	90	15	15	120*	1.2	1.3	16	1.3	400	2.4	5*	30*	550*
51–70 y	900	90	15	15	120*	1.2	1.3	16	1.7	400	2.4 <sup>h</sup>	5*	30*	550*
> 70 y	900	90	20	15	120*	1.2	1.3	16	1.7	400	2.4 <sup>h</sup>	5*	30*	550*
<b>Females</b>														
9–13 y	600	45	15	11	60*	0.9	0.9	12	1.0	300	1.8	4*	20*	375*
14–18 y	700	65	15	15	75*	1.0	1.0	14	1.2	400 <sup>i</sup>	2.4	5*	25*	400*
19–30 y	700	75	15	15	90*	1.1	1.1	14	1.3	400 <sup>i</sup>	2.4	5*	30*	425*
31–50 y	700	75	15	15	90*	1.1	1.1	14	1.3	400 <sup>i</sup>	2.4	5*	30*	425*
51–70 y	700	75	15	15	90*	1.1	1.1	14	1.5	400 <sup>i</sup>	2.4 <sup>h</sup>	5*	30*	425*
> 70 y	700	75	20	15	90*	1.1	1.1	14	1.5	400	2.4 <sup>h</sup>	5*	30*	425*
<b>Pregnancy</b>														
14–18 y	750	80	15	15	75*	1.4	1.4	18	1.9	600 <sup>j</sup>	2.6	6*	30*	450*
19–30 y	770	85	15	15	90*	1.4	1.4	18	1.9	600 <sup>j</sup>	2.6	6*	30*	450*
31–50 y	770	85	15	15	90*	1.4	1.4	18	1.9	600 <sup>j</sup>	2.6	6*	30*	450*
<b>Lactation</b>														
14–18 y	1,200	115	15	19	75*	1.4	1.6	17	2.0	500	2.8	7*	35*	550*
19–30 y	1,300	120	15	19	90*	1.4	1.6	17	2.0	500	2.8	7*	35*	550*
31–50 y	1,300	120	15	19	90*	1.4	1.6	17	2.0	500	2.8	7*	35*	550*

**NOTE:** This table (taken from the DRI reports, see [www.nap.edu](http://www.nap.edu)) presents Recommended Dietary Allowances (RDAs) in **bold type** and Adequate Intakes (AIs) in ordinary type followed by an asterisk (\*). An RDA is the average daily dietary intake level; sufficient to meet the nutrient requirements of nearly all (97–98 percent) healthy individuals in a group. It is calculated from an Estimated Average Requirement (EAR). If sufficient scientific evidence is not available to establish an EAR, and thus calculate an RDA, an AI is usually developed. For healthy breastfed infants, an AI is the mean intake. The AI for other life stage and gender groups is believed to cover the needs of all healthy individuals in the groups, but lack of data or uncertainty in the data prevent being able to specify with confidence the percentage of individuals covered by this intake.

<sup>a</sup> As retinol activity equivalents (RAEs). 1 RAE = 1 µg retinol, 12 µg β-carotene, or 24 µg β-cryptoxanthin. The RAE for dietary provitamin A carotenoids is two-fold greater than retinol equivalents (RE), whereas the RAE for preformed vitamin A is the same as RE.

<sup>b</sup> As cholecalciferol. 1 µg cholecalciferol = 40 IU vitamin D.

<sup>c</sup> Under the assumption of minimal sunlight.

<sup>d</sup> As α-tocopherol. α-Tocopherol includes *RRR*-α-tocopherol, the only form of α-tocopherol that occurs naturally in foods, and the 2*R*-stereoisomeric forms of α-tocopherol (*RRR*-, *RSR*-, *RRS*-, and *RSS*-α-tocopherol) that occur in fortified foods and supplements. It does not include the 2*S*-stereoisomeric forms of α-tocopherol (*SRR*-, *SSR*-, *SRS*-, and *SSS*-α-tocopherol), also found in fortified foods and supplements.

<sup>e</sup> As niacin equivalents (NE). 1 mg of niacin = 60 mg of tryptophan; 0–6 months = preformed niacin (not NE).

<sup>f</sup> As dietary folate equivalents (DFE). 1 DFE = 1 µg food folate = 0.6 µg of folic acid from fortified food or as a supplement consumed with food = 0.5 µg of a supplement taken on an empty stomach.

<sup>g</sup> Although AIs have been set for choline, there are few data to assess whether a dietary supply of choline is needed at all stages of the life cycle, and it may be that the choline requirement can be met by endogenous synthesis at some of these stages.

<sup>h</sup> Because 10 to 30 percent of older people may malabsorb food-bound B<sub>12</sub>, it is advisable for those older than 50 years to meet their RDA mainly by consuming foods fortified with B<sub>12</sub> or a supplement containing B<sub>12</sub>.

<sup>i</sup> In view of evidence linking folate intake with neural tube defects in the fetus, it is recommended that all women capable of becoming pregnant consume 400 µg from supplements or fortified foods in addition to intake of food folate from a varied diet.

<sup>j</sup>It is assumed that women will continue consuming 400 µg from supplements or fortified food until their pregnancy is confirmed and they enter prenatal care, which ordinarily occurs after the end of the periconceptional period—the critical time for formation of the neural tube.

**SOURCES:** *Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride* (1997); *Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B<sub>6</sub>, Folate, Vitamin B<sub>12</sub>, Pantothenic Acid, Biotin, and Choline* (1998); *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids* (2000); *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc* (2001); *Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate* (2005); and *Dietary Reference Intakes for Calcium and Vitamin D* (2011). These reports may be accessed via [www.nap.edu](http://www.nap.edu).

**Dietary Reference Intakes (DRIs): Recommended Dietary Allowances and Adequate Intakes, Elements**  
Food and Nutrition Board, Institute of Medicine, National Academies

Life Stage Group	Calcium (mg/d)	Chromium (µg/d)	Copper (µg/d)	Fluoride (mg/d)	Iodine (µg/d)	Iron (mg/d)	Magnesium (mg/d)	Manganese (mg/d)	Molybdenum (µg/d)	Phosphorus (mg/d)	Selenium (µg/d)	Zinc (mg/d)	Potassium (g/d)	Sodium (g/d)	Chloride (g/d)
<b>Infants</b>															
0 to 6 mo	200*	0.2*	200*	0.01*	110*	0.27*	30*	0.003*	2*	100*	15*	2*	0.4*	0.12*	0.18*
6 to 12 mo	260*	5.5*	220*	0.5*	130*	<b>11</b>	75*	0.6*	3*	275*	20*	<b>3</b>	0.7*	0.37*	0.57*
<b>Children</b>															
1–3 y	<b>700</b>	11*	<b>340</b>	0.7*	<b>90</b>	<b>7</b>	<b>80</b>	1.2*	<b>17</b>	<b>460</b>	<b>20</b>	<b>3</b>	3.0*	1.0*	1.5*
4–8 y	<b>1,000</b>	15*	<b>440</b>	1*	<b>90</b>	<b>10</b>	<b>130</b>	1.5*	<b>22</b>	<b>500</b>	<b>30</b>	<b>5</b>	3.8*	1.2*	1.9*
<b>Males</b>															
9–13 y	<b>1,300</b>	25*	<b>700</b>	2*	<b>120</b>	<b>8</b>	<b>240</b>	1.9*	<b>34</b>	<b>1,250</b>	<b>40</b>	<b>8</b>	4.5*	1.5*	2.3*
14–18 y	<b>1,300</b>	35*	<b>890</b>	3*	<b>150</b>	<b>11</b>	<b>410</b>	2.2*	<b>43</b>	<b>1,250</b>	<b>55</b>	<b>11</b>	4.7*	1.5*	2.3*
19–30 y	<b>1,000</b>	35*	<b>900</b>	4*	<b>150</b>	<b>8</b>	<b>400</b>	2.3*	<b>45</b>	<b>700</b>	<b>55</b>	<b>11</b>	4.7*	1.5*	2.3*
31–50 y	<b>1,000</b>	35*	<b>900</b>	4*	<b>150</b>	<b>8</b>	<b>420</b>	2.3*	<b>45</b>	<b>700</b>	<b>55</b>	<b>11</b>	4.7*	1.5*	2.3*
51–70 y	<b>1,000</b>	30*	<b>900</b>	4*	<b>150</b>	<b>8</b>	<b>420</b>	2.3*	<b>45</b>	<b>700</b>	<b>55</b>	<b>11</b>	4.7*	1.3*	2.0*
> 70 y	<b>1,200</b>	30*	<b>900</b>	4*	<b>150</b>	<b>8</b>	<b>420</b>	2.3*	<b>45</b>	<b>700</b>	<b>55</b>	<b>11</b>	4.7*	1.2*	1.8*
<b>Females</b>															
9–13 y	<b>1,300</b>	21*	<b>700</b>	2*	<b>120</b>	<b>8</b>	<b>240</b>	1.6*	<b>34</b>	<b>1,250</b>	<b>40</b>	<b>8</b>	4.5*	1.5*	2.3*
14–18 y	<b>1,300</b>	24*	<b>890</b>	3*	<b>150</b>	<b>15</b>	<b>360</b>	1.6*	<b>43</b>	<b>1,250</b>	<b>55</b>	<b>9</b>	4.7*	1.5*	2.3*
19–30 y	<b>1,000</b>	25*	<b>900</b>	3*	<b>150</b>	<b>18</b>	<b>310</b>	1.8*	<b>45</b>	<b>700</b>	<b>55</b>	<b>8</b>	4.7*	1.5*	2.3*
31–50 y	<b>1,000</b>	25*	<b>900</b>	3*	<b>150</b>	<b>18</b>	<b>320</b>	1.8*	<b>45</b>	<b>700</b>	<b>55</b>	<b>8</b>	4.7*	1.5*	2.3*
51–70 y	<b>1,200</b>	20*	<b>900</b>	3*	<b>150</b>	<b>8</b>	<b>320</b>	1.8*	<b>45</b>	<b>700</b>	<b>55</b>	<b>8</b>	4.7*	1.3*	2.0*
> 70 y	<b>1,200</b>	20*	<b>900</b>	3*	<b>150</b>	<b>8</b>	<b>320</b>	1.8*	<b>45</b>	<b>700</b>	<b>55</b>	<b>8</b>	4.7*	1.2*	1.8*
<b>Pregnancy</b>															
14–18 y	<b>1,300</b>	29*	<b>1,000</b>	3*	<b>220</b>	<b>27</b>	<b>400</b>	2.0*	<b>50</b>	<b>1,250</b>	<b>60</b>	<b>12</b>	4.7*	1.5*	2.3*
19–30 y	<b>1,000</b>	30*	<b>1,000</b>	3*	<b>220</b>	<b>27</b>	<b>350</b>	2.0*	<b>50</b>	<b>700</b>	<b>60</b>	<b>11</b>	4.7*	1.5*	2.3*
31–50 y	<b>1,000</b>	30*	<b>1,000</b>	3*	<b>220</b>	<b>27</b>	<b>360</b>	2.0*	<b>50</b>	<b>700</b>	<b>60</b>	<b>11</b>	4.7*	1.5*	2.3*
<b>Lactation</b>															
14–18 y	<b>1,300</b>	44*	<b>1,300</b>	3*	<b>290</b>	<b>10</b>	<b>360</b>	2.6*	<b>50</b>	<b>1,250</b>	<b>70</b>	<b>13</b>	5.1*	1.5*	2.3*
19–30 y	<b>1,000</b>	45*	<b>1,300</b>	3*	<b>290</b>	<b>9</b>	<b>310</b>	2.6*	<b>50</b>	<b>700</b>	<b>70</b>	<b>12</b>	5.1*	1.5*	2.3*
31–50 y	<b>1,000</b>	45*	<b>1,300</b>	3*	<b>290</b>	<b>9</b>	<b>320</b>	2.6*	<b>50</b>	<b>700</b>	<b>70</b>	<b>12</b>	5.1*	1.5*	2.3*

**NOTE:** This table (taken from the DRI reports, see www.nap.edu) presents Recommended Dietary Allowances (RDAs) in **bold type** and Adequate Intakes (AIs) in ordinary type followed by an asterisk (\*). An RDA is the average daily dietary intake level; sufficient to meet the nutrient requirements of nearly all (97–98 percent) healthy individuals in a group. It is calculated from an Estimated Average Requirement (EAR). If sufficient scientific evidence is not available to establish an EAR, and thus calculate an RDA, an AI is usually developed. For healthy breastfed infants, an AI is the mean intake. The AI for other life stage and gender groups is believed to cover the needs of all healthy individuals in the groups, but lack of data or uncertainty in the data prevent being able to specify with confidence the percentage of individuals covered by this intake.

**SOURCES:** *Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride* (1997); *Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B<sub>6</sub>, Folate, Vitamin B<sub>12</sub>, Pantothenic Acid, Biotin, and Choline* (1998); *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids* (2000); and *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc* (2001); *Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate* (2005); and *Dietary Reference Intakes for Calcium and Vitamin D* (2011). These reports may be accessed via www.nap.edu.



**Dietary Reference Intakes (DRIs): Recommended Dietary Allowances and Adequate Intakes, Total Water and Macronutrients**

Food and Nutrition Board, Institute of Medicine, National Academies

Life Stage Group	Total						
	Water <sup>a</sup> (L/d)	Carbohydrate (g/d)	Fiber (g/d)	Fat (g/d)	Linoleic Acid (g/d)	$\alpha$ -Linolenic Acid (g/d)	Protein <sup>b</sup> (g/d)
<b>Infants</b>							
0 to 6 mo	0.7*	60*	ND	31*	4.4*	0.5*	9.1*
6 to 12 mo	0.8*	95*	ND	30*	4.6*	0.5*	<b>11.0</b>
<b>Children</b>							
1–3 y	1.3*	<b>130</b>	19*	ND <sup>c</sup>	7*	0.7*	<b>13</b>
4–8 y	1.7*	<b>130</b>	25*	ND	10*	0.9*	<b>19</b>
<b>Males</b>							
9–13 y	2.4*	<b>130</b>	31*	ND	12*	1.2*	<b>34</b>
14–18 y	3.3*	<b>130</b>	38*	ND	16*	1.6*	<b>52</b>
19–30 y	3.7*	<b>130</b>	38*	ND	17*	1.6*	<b>56</b>
31–50 y	3.7*	<b>130</b>	38*	ND	17*	1.6*	<b>56</b>
51–70 y	3.7*	<b>130</b>	30*	ND	14*	1.6*	<b>56</b>
> 70 y	3.7*	<b>130</b>	30*	ND	14*	1.6*	<b>56</b>
<b>Females</b>							
9–13 y	2.1*	<b>130</b>	26*	ND	10*	1.0*	<b>34</b>
14–18 y	2.3*	<b>130</b>	26*	ND	11*	1.1*	<b>46</b>
19–30 y	2.7*	<b>130</b>	25*	ND	12*	1.1*	<b>46</b>
31–50 y	2.7*	<b>130</b>	25*	ND	12*	1.1*	<b>46</b>
51–70 y	2.7*	<b>130</b>	21*	ND	11*	1.1*	<b>46</b>
> 70 y	2.7*	<b>130</b>	21*	ND	11*	1.1*	<b>46</b>
<b>Pregnancy</b>							
14–18 y	3.0*	<b>175</b>	28*	ND	13*	1.4*	<b>71</b>
19–30 y	3.0*	<b>175</b>	28*	ND	13*	1.4*	<b>71</b>
31–50 y	3.0*	<b>175</b>	28*	ND	13*	1.4*	<b>71</b>
<b>Lactation</b>							
14–18 y	3.8*	<b>210</b>	29*	ND	13*	1.3*	<b>71</b>
19–30 y	3.8*	<b>210</b>	29*	ND	13*	1.3*	<b>71</b>
31–50 y	3.8*	<b>210</b>	29*	ND	13*	1.3*	<b>71</b>

**NOTE:** This table (take from the DRI reports, see www.nap.edu) presents Recommended Dietary Allowances (RDA) in bold type and Adequate Intakes (AI) in ordinary type followed by an asterisk (\*). An RDA is the average daily dietary intake level; sufficient to meet the nutrient requirements of nearly all (97–98 percent) healthy individuals in a group. It is calculated from an Estimated Average Requirement (EAR). If sufficient scientific evidence is not available to establish an EAR, and thus calculate an RDA, an AI is usually developed. For healthy breastfed infants, an AI is the mean intake. The AI for other life stage and gender groups is believed to cover the needs of all healthy individuals in the groups, but lack of data or uncertainty in the data prevent being able to specify with confidence the percentage of individuals covered by this intake.

<sup>a</sup> Total water includes all water contained in food, beverages, and drinking water.

<sup>b</sup> Based on g protein per kg of body weight for the reference body weight, e.g., for adults 0.8 g/kg body weight for the reference body weight.

<sup>c</sup>Not determined.

**SOURCE:** *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids* (2002/2005) and *Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate* (2005). The report may be accessed via www.nap.edu.

### Dietary Reference Intakes (DRIs): Acceptable Macronutrient Distribution Ranges

Food and Nutrition Board, Institute of Medicine, National Academies

Macronutrient	Range (percent of energy)		
	Children, 1–3 y	Children, 4–18 y	Adults
Fat	30–40	25–35	20–35
<i>n</i> -6 polyunsaturated fatty acids <sup>a</sup> (linoleic acid)	5–10	5–10	5–10
<i>n</i> -3 polyunsaturated fatty acids <sup>a</sup> ( $\alpha$ -linolenic acid)	0.6–1.2	0.6–1.2	0.6–1.2
Carbohydrate	45–65	45–65	45–65
Protein	5–20	10–30	10–35

<sup>a</sup> Approximately 10 percent of the total can come from longer-chain *n*-3 or *n*-6 fatty acids.

**SOURCE:** *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids* (2002/2005). The report may be accessed via [www.nap.edu](http://www.nap.edu).

### Dietary Reference Intakes (DRIs): Acceptable Macronutrient Distribution Ranges

Food and Nutrition Board, Institute of Medicine, National Academies

Macronutrient	Recommendation
Dietary cholesterol	As low as possible while consuming a nutritionally adequate diet
Trans fatty Acids	As low as possible while consuming a nutritionally adequate diet
Saturated fatty acids	As low as possible while consuming a nutritionally adequate diet
Added sugars <sup>a</sup>	Limit to no more than 25 % of total energy

<sup>a</sup>Not a recommended intake. A daily intake of added sugars that individuals should aim for to achieve a healthful diet was not set.

**SOURCE:** *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids* (2002/2005). The report may be accessed via [www.nap.edu](http://www.nap.edu).

**Dietary Reference Intakes (DRIs): Tolerable Upper Intake Levels, Vitamins**  
Food and Nutrition Board, Institute of Medicine, National Academies

Life Stage Group	Vitamin A (µg/d) <sup>a</sup>	Vitamin C (mg/d)	Vitamin D (µg/d)	Vitamin E (mg/d) <sup>b,c</sup>	Vitamin K	Thiamin	Riboflavin	Niacin (mg/d) <sup>f</sup>	Vitamin B <sub>6</sub> (mg/d)	Folate (µg/d) <sup>e</sup>	Vitamin B <sub>12</sub>	Pantothenic Acid	Biotin	Choline (g/d)	Carotenoids <sup>d</sup>
<b>Infants</b>															
0 to 6 mo	600	ND <sup>e</sup>	25	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
6 to 12 mo	600	ND	38	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
<b>Children</b>															
1–3 y	600	400	63	200	ND	ND	ND	10	30	300	ND	ND	ND	1.0	ND
4–8 y	900	650	75	300	ND	ND	ND	15	40	400	ND	ND	ND	1.0	ND
<b>Males</b>															
9–13 y	1,700	1,200	100	600	ND	ND	ND	20	60	600	ND	ND	ND	2.0	ND
14–18 y	2,800	1,800	100	800	ND	ND	ND	30	80	800	ND	ND	ND	3.0	ND
19–30 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
31–50 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
51–70 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
> 70 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
<b>Females</b>															
9–13 y	1,700	1,200	100	600	ND	ND	ND	20	60	600	ND	ND	ND	2.0	ND
14–18 y	2,800	1,800	100	800	ND	ND	ND	30	80	800	ND	ND	ND	3.0	ND
19–30 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
31–50 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
51–70 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
> 70 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
<b>Pregnancy</b>															
14–18 y	2,800	1,800	100	800	ND	ND	ND	30	80	800	ND	ND	ND	3.0	ND
19–30 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
31–50 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
<b>Lactation</b>															
14–18 y	2,800	1,800	100	800	ND	ND	ND	30	80	800	ND	ND	ND	3.0	ND
19–30 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND
31–50 y	3,000	2,000	100	1,000	ND	ND	ND	35	100	1,000	ND	ND	ND	3.5	ND

NOTE: A Tolerable Upper Intake Level (UL) is the highest level of daily nutrient intake that is likely to pose no risk of adverse health effects to almost all individuals in the general population. Unless otherwise specified, the UL represents total intake from food, water, and supplements. Due to a lack of suitable data, ULs could not be established for vitamin K, thiamin, riboflavin, vitamin B<sub>12</sub>, pantothenic acid, biotin, and carotenoids. In the absence of a UL, extra caution may be warranted in consuming levels above recommended intakes. Members of the general population should be advised not to routinely exceed the UL. The UL is not meant to apply to individuals who are treated with the nutrient under medical supervision or to individuals with predisposing conditions that modify their sensitivity to the nutrient.

<sup>a</sup>As preformed vitamin A only.

<sup>b</sup>As α-tocopherol; applies to any form of supplemental α-tocopherol.

<sup>c</sup>The ULs for vitamin E, niacin, and folate apply to synthetic forms obtained from supplements, fortified foods, or a combination of the two.

<sup>d</sup>β-Carotene supplements are advised only to serve as a provitamin A source for individuals at risk of vitamin A deficiency.

<sup>e</sup>ND = Not determinable due to lack of data of adverse effects in this age group and concern with regard to lack of ability to handle excess amounts. Source of intake should be from food only to prevent high levels of intake.

**SOURCES:** *Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride (1997); Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B<sub>6</sub>, Folate, Vitamin B<sub>12</sub>, Pantothenic Acid, Biotin, and Choline (1998); Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids (2000); Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc (2001); and Dietary Reference Intakes for Calcium and Vitamin D (2011).* These reports may be accessed via [www.nap.edu](http://www.nap.edu).

## Dietary Reference Intakes (DRIs): Tolerable Upper Intake Levels, Elements

Food and Nutrition Board, Institute of Medicine, National Academies

Life Stage Group	Arsenic <sup>a</sup>	Boron (mg/d)	Calcium (mg/d)	Chromium (mg/d)	Copper (µg/d)	Fluoride (mg/d)	Iodine (µg/d)	Iron (mg/d)	Magnesium (mg/d) <sup>b</sup>	Manganese (mg/d)	Molybdenum (µg/d)	Nickel (mg/d)	Phosphorus (g/d)	Selenium (µg/d)	Silicon <sup>c</sup>	Vanadium (mg/d) <sup>d</sup>	Zinc (mg/d)	Sodium (g/d)	Chloride (g/d)		
<b>Infants</b>																					
0 to 6 mo	ND <sup>e</sup>	ND	1,000	ND	ND	0.7	ND	40	ND	ND	ND	ND	ND	45	ND	ND	4	ND	ND	ND	
6 to 12 mo	ND	ND	1,500	ND	ND	0.9	ND	40	ND	ND	ND	ND	ND	60	ND	ND	5	ND	ND	ND	
<b>Children</b>																					
1–3 y	ND	3	2,500	ND	1,000	1.3	200	40	65	2	300	0.2	3	90	ND	ND	7	1.5	2.3	2.3	
4–8 y	ND	6	2,500	ND	3,000	2.2	300	40	110	3	600	0.3	3	150	ND	ND	12	1.9	2.9	2.9	
<b>Males</b>																					
9–13 y	ND	11	3,000	ND	5,000	10	600	40	350	6	1,100	0.6	4	280	ND	ND	23	2.2	3.4	3.4	
14–18 y	ND	17	3,000	ND	8,000	10	900	45	350	9	1,700	1.0	4	400	ND	ND	34	2.3	3.6	3.6	
19–30 y	ND	20	2,500	ND	10,000	10	1,100	45	350	11	2,000	1.0	4	400	ND	ND	40	2.3	3.6	3.6	
31–50 y	ND	20	2,500	ND	10,000	10	1,100	45	350	11	2,000	1.0	4	400	ND	ND	40	2.3	3.6	3.6	
51–70 y	ND	20	2,000	ND	10,000	10	1,100	45	350	11	2,000	1.0	4	400	ND	ND	40	2.3	3.6	3.6	
> 70 y	ND	20	2,000	ND	10,000	10	1,100	45	350	11	2,000	1.0	3	400	ND	ND	40	2.3	3.6	3.6	
<b>Females</b>																					
9–13 y	ND	11	3,000	ND	5,000	10	600	40	350	6	1,100	0.6	4	280	ND	ND	23	2.2	3.4	3.4	
14–18 y	ND	17	3,000	ND	8,000	10	900	45	350	9	1,700	1.0	4	400	ND	ND	34	2.3	3.6	3.6	
19–30 y	ND	20	2,500	ND	10,000	10	1,100	45	350	11	2,000	1.0	4	400	ND	ND	40	2.3	3.6	3.6	
31–50 y	ND	20	2,500	ND	10,000	10	1,100	45	350	11	2,000	1.0	4	400	ND	ND	40	2.3	3.6	3.6	
51–70 y	ND	20	2,000	ND	10,000	10	1,100	45	350	11	2,000	1.0	4	400	ND	ND	40	2.3	3.6	3.6	
> 70 y	ND	20	2,000	ND	10,000	10	1,100	45	350	11	2,000	1.0	3	400	ND	ND	40	2.3	3.6	3.6	
<b>Pregnancy</b>																					
14–18 y	ND	17	3,000	ND	8,000	10	900	45	350	9	1,700	1.0	3.5	400	ND	ND	34	2.3	3.6	3.6	
19–30 y	ND	20	2,500	ND	10,000	10	1,100	45	350	11	2,000	1.0	3.5	400	ND	ND	40	2.3	3.6	3.6	
61–50 y	ND	20	2,500	ND	10,000	10	1,100	45	350	11	2,000	1.0	3.5	400	ND	ND	40	2.3	3.6	3.6	
<b>Lactation</b>																					
14–18 y	ND	17	3,000	ND	8,000	10	900	45	350	9	1,700	1.0	4	400	ND	ND	34	2.3	3.6	3.6	
19–30 y	ND	20	2,500	ND	10,000	10	1,100	45	350	11	2,000	1.0	4	400	ND	ND	40	2.3	3.6	3.6	
31–50 y	ND	20	2,500	ND	10,000	10	1,100	45	350	11	2,000	1.0	4	400	ND	ND	40	2.3	3.6	3.6	

**NOTE:** A Tolerable Upper Intake Level (UL) is the highest level of daily nutrient intake that is likely to pose no risk of adverse health effects to almost all individuals in the general population. Unless otherwise specified, the UL represents total intake from food, water, and supplements. Due to a lack of suitable data, ULs could not be established for vitamin K, thiamin, riboflavin, vitamin B<sub>12</sub>, pantothenic acid, biotin, and carotenoids. In the absence of a UL, extra caution may be warranted in consuming levels above recommended intakes. Members of the general population should be advised not to routinely exceed the UL. The UL is not meant to apply to individuals who are treated with the nutrient under medical supervision or to individuals with predisposing conditions that modify their sensitivity to the nutrient.

<sup>a</sup>Although the UL was not determined for arsenic, there is no justification for adding arsenic to food or supplements.

<sup>b</sup>The ULs for magnesium represent intake from a pharmacological agent only and do not include intake from food and water.

<sup>c</sup>Although silicon has not been shown to cause adverse effects in humans, there is no justification for adding silicon to supplements.

<sup>d</sup>Although vanadium in food has not been shown to cause adverse effects in humans, there is no justification for adding vanadium to food and vanadium supplements should be used with caution. The UL is based on adverse effects in laboratory animals and this data could be used to set a UL for adults but not children and adolescents.

<sup>e</sup>ND = Not determinable due to lack of data of adverse effects in this age group and concern with regard to lack of ability to handle excess amounts. Source of intake should be from food only to prevent high levels of intake.

**SOURCES:** *Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride* (1997); *Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B<sub>6</sub>, Folate, Vitamin B<sub>12</sub>, Pantothenic Acid, Biotin, and Choline* (1998); *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids* (2000); *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc* (2001); *Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate* (2005); and *Dietary Reference Intakes for Calcium and Vitamin D* (2011). These reports may be accessed via [www.nap.edu](http://www.nap.edu).

## Anhang C

# DACH Referenzwerte

Hierbei handelt es sich um die verwendeten DACH Referenzwerte des Ernährungsberichtes von 2012. Diese wurden auch in dieser Arbeit verwendet.

Altersklassen	Männer			Frauen		
	18-24	25-50	51-64	18-24	25-50	61-64
Vitamin A [mg]	1	1	1	0.8	0.8	0.8
Vitamin B1 [mg]	1.3	1.2	1.1	1	1	1
Vitamin B2 [mg]	1.5	1.4	1.3	1.2	1.2	1.2
Vitamin B3 [mg]	17	16	15	13	13	13
Vitamin B5 [mg]	6	6	6	6	6	6
Vitamin B6 [mg]	1.5	1.5	1.5	1.2	1.2	1.2
Vitamin B7 [ $\mu$ g]	30-60	30-60	30-60	30-60	30-60	30-60
Vitamin B9 [ $\mu$ g]	400	400	400	400	400	400
Vitamin B12 [ $\mu$ g]	3	3	3	3	3	3
Vitamin C [mg]	100	100	100	100	100	100
Vitmain D [ $\mu$ g]	5	5	5	5	5	5
Vitmain E [mg]	15	14	13	12	12	12
Vitmain K [ $\mu$ g]	70	70	80	60	60	65
Kalzium [mg]	1000	1000	1000	1000	1000	1000
Magnesium [mg]	400	350	350	310	300	300
Eisen [mg]	10	10	10	15	15	10
Phosphor [mg]	700	700	700	700	700	700
Zink [mg]	10	10	10	7	7	7

Tabelle C.1: DACH Referenzwerte

# Anhang D

## Lebenslauf

### Persönliche Daten

Name	Dominik Hlauschek
Geburtsdatum,-ort	06.02.1990, Mödling
Staatsbürgerschaft	Österreich
E-Mail Adresse	d.hlauschek@gmx.net
Wohnhaft in	Kottingbrunn

### Ausbildung

1996-2000	Volksschule Kottingbrunn
2000-2004	Sportgymnasium Zehnergasse, Wiener Neustadt
2004-2009	Höhere Technische Bundeslehr- und Versuchsanstalt Mödling
2010-2013	Bakkalaureatsstudium Statistik an der Universität Wien
seit 2013	Magisterstudium Statistik an der Universität Wien

### Berufserfahrung

März-Juni 2015	Tutorium Inferenzstatistik an der Universität Wien
März-August 2015	Studienassistent für Statistik 1 bei Univ.-Prof. Erhard Reschenhofer

# Anhang E

## Zusammenfassung/ Abstract

### E.1 Zusammenfassung

Diese Masterarbeit stellt statistische Modelle zur Schätzung von üblichen Nährstoffaufnahmeverteilungen mit 24 Stunden Erinnerungsprotokollen vor. Dabei wird eine Übersicht über die Problemstellung, sowie mathematische Grundlagen zur Thematik gegeben. Die vorgestellten Modelle beinhalten die Methoden des National Cancer Institute (NCI), der Iowa State University (ISU), sowie die Multiple Source Methode (MSM) und die Statistical Program to Assess Dietary Exposure (SPADE) Methode. Diese Modelle werden sowohl inhaltlich, als auch im Bezug auf die Verwendung von Daten verglichen. Die Daten stammen von den 24-Stunden Erinnerungsprotokollen der Erwachsenen des österreichischen Ernährungsberichtes von 2012. Alle 4 Methoden unterscheiden sich in den Ergebnissen nur sehr wenig, solange die Annahmen der Modelle erfüllt sind. Welche Methode verwendet werden soll, hängt vor allem von der untersuchten Fragestellung ab und ob Subgruppen analysiert werden sollen.

Ein weiteres Ziel der Arbeit ist eine Neuauswertung der, im Ernährungsbericht untersuchten Makro- und Mikronährstoffe. Dabei werden die Unterschiede, der im Ernährungsbericht verwendeten statistischen Methodik (within-person Mittelwertmethode) und der SPADE Methode aufgezeigt. Vor allem auf die Anteilsschätzungen der unterversorgten Personen, bezüglich der Mikronährstoffe, wird gezielt eingegangen. Unterschiede in den Anteilen von um die 20-40% sind dabei üblich. Ein Teil davon erklärt sich durch die Unterschiede in der verwendeten Methodik, wobei hier immer mit der Mittelwertmethode systematisch unter- oder überschätzt wird (ungefähr 10-20%). Der Rest des Unterschiedes bezieht sich auf die verwendeten Grenzwerte. Hierbei kann aufgezeigt werden, dass die Verwendung der DACH Empfehlungen (Referenzwerte) eine Überschätzung des Anteils der Unterversorgten zur Folge hat. Verglichen werden die Ergebnisse mit den Anteilsschätzungen bezüglich dem Estimated Average Requirement als Grenzwert.

Im letzten Teil der Arbeit wird eine Diskussion rund um die verwendeten Methoden und Daten gegeben. Dabei wird vor allem auf die Erwartungstreue der Ernährungsberichtsdaten eingegangen. Zuletzt werden noch Schlussfolgerungen und Empfehlungen für zukünftige Ernährungsberichte gegeben.

### E.2 Abstract

This master thesis introduces statistical models for estimating usual intake distributions of nutrients and food from 24-hour Recalls. An overview of the problem and mathematical principles are given. The investigated models include the methods from the National Cancer Institute (NCI), the Iowa State University (ISU) as well as the Statistical Program to Assess Dietary Exposure (SPADE) method and the Multiple Source Method (MSM). These models are compared with regards to content and data usage. Data are provided from the "Österreichischer Ernährungsbericht (2012)" (nutrition report of the austrian population) regarding 24-hour Recalls of the adults. All 4 methods produce very similar results as long as the assumptions of the models are met. Which method to use depends mainly on the surveyed questions and whether subgroups should be analyzed or not.

Another aim of the work is a reevaluation of the macro- and micronutrients examined in the Austrian Nutrition Report. The differences in the statistical methodology used in the nutrition report (within-person mean) and the SPADE method are shown. Especially the prevalence estimates of inadequate intakes regarding the micronutrients are specifically addressed. Usual differences in the proportion estimated are between 20 and 40%. These discrepancies can be partly addressed to the used within-person mean method. This method always causes an under- or overestimation of the proportion (about 10-20%). The remaining part of the discrepancy refers to the used cut-points. It can be shown that the usage of the DACH "Empfehlungen" (DACH recommendations - reference values) results in an overestimation of the prevalence estimates. These results are compared with the Estimated Average Requirements as cut-points.

The last part of this master thesis covers a discussion of the used methods and data. Especially the problem of the unbiasedness of the data is addressed. Finally conclusions and recommendations for future nutrition reports are given.