

19. Ökologische Inferenz

Thomas Gschwend

19.1. Kurzdarstellung

Als ökologische Inferenz wird der Versuch bezeichnet, mit Daten eines höheren Aggregationsniveaus (im Folgenden verkürzend als „Aggregatebene“ bezeichnet) Rückschlüsse auf Prozesse einer niedrigeren Ebene (im Folgenden „Individualebene“) zu ziehen. Angewendet wird diese Strategie dann, wenn nur Daten auf der Aggregatebene verfügbar sind, aber Prozesse auf der Individualebene im Mittelpunkt des Analyseinteresses stehen.

Das zentrale forschungslogische Problem dieser Strategie ist, dass viele unterschiedliche Datenkonstellationen auf der Individualebene das auf der Aggregatebene beobachtete Muster erzeugen können. So passen zu einem amtlichen Endergebnis der Erst- und Zweitstimmen einer Partei bei einer Bundestagswahl sehr unterschiedliche Verteilungen des individuellen Wahlverhaltens. Aus der Information über die Verteilung der Erst- und Zweitstimmen kann daher im Allgemeinen nicht geschlossen werden, wie die einzelnen Wähler ihre Erst- und Zweitstimme vergeben haben. Mit dieser Uneindeutigkeit auf der Individual- also der analytisch interessierenden Ebene haben alle Methoden der ökologischen Inferenz zu kämpfen. Sie versuchen diesen Nachteil mit geballter Technologie und zahlreichen Annahmen über den datengenerierenden Prozess und die Aggregationslogik zu kompensieren.

19.2. Beschreibung des Verfahrens

19.2.1. Einführung

Ökologische Inferenz wird immer dann erforderlich, wenn Zusammenhänge auf der Individualebene in Ermangelung anderer Informationen mit aggregierten Daten überprüft werden sollen – beispielsweise wenn Individualdaten nicht zugänglich (Volkszählungsdaten), nicht vorhanden (historische Wahldaten) oder aber unzuverlässig (soziale Erwünschtheit) sind. Ich beschränke mich hier auf die Darstellung von Kings (1997) *Ecological Inference* (EI) Modell. EI hat unter allen neueren methodischen Weiterentwicklungen der ökologischen Inferenz die größte Breitenwirkung erzielt, viele Anwendungsbeispiele motiviert und die methodische Diskussion angeregt.

Das zentrale forschungslogische Problem jeder ökologischen Inferenz ist, dass unterschiedliche Datengenerierungsprozesse auf der Individualebene das gleiche Datenmuster auf der Aggregatebene erzeugen können. Beziehungen zwischen Variablen auf der Aggregatebene können, müssen aber nicht ähnliche Beziehungen auf der Individualebene widerspiegeln. Im Extremfall können Korrelationen zwischen Aggregatvariablen sich sogar im Vorzeichen von den Beziehungen auf der Individualebene unterscheiden (Robinson 1950). So findet man häufig eine positive Korrelationen zwischen dem Ausländeranteil in einem Wahlbezirk und dem Abschneiden rechtsextremer Parteien auf der Aggregatebene. Es ist

Tabelle 1: Das Inferenzproblem als Vierfeldertafel

	Wahl	Nichtwahl	
Afroamerikaner	β_i^b	$1 - \beta_i^b$	X_i
Weißer	β_i^w	$1 - \beta_i^w$	$1 - X_i$
	T_i	$1 - T_i$	

ein offensichtlicher *ökologischer Fehlschluss* (Pappi 1977) daraus zu folgern, dass Ausländer die Wähler solcher Parteien sind. Ausländer sind nicht wahlberechtigt. Die positive Korrelation auf der Aggregatebene kommt hier dadurch zustande, dass die wahlberechtigten Nicht-Ausländer in Wahlkreisen mit hohen Ausländeranteilen besonders häufig rechtsextreme Parteien gewählt haben.

19.2.2. Notation und Formalisierung

Zur Einführung verwende ich ein Beispiel von King (1997), in dem es um die Schätzung der Wahlbeteiligung (Turnout) ethnischer (black bzw. white) Gruppen geht. Die anhand dieses Beispiels entwickelte Notation hat sich inzwischen als Standard durchgesetzt.

Für alle Wahlkreise (i) enthalten amtliche Statistiken Informationen über den Anteil der Afroamerikaner (X_i) und der Wähler (T_i). Damit wissen wir auch, wie viele Weiße – oder genauer – Nicht-Afroamerikaner ($1 - X_i$) und wie viele Nichtwähler ($1 - T_i$) es dort gibt. Während wir aber die Randverteilung kennen, ist keine Information darüber verfügbar, wie hoch der Anteil der Wähler in den Gruppen der Afroamerikaner (β_i^b) bzw. der Weißen (β_i^w) ist. Ziel der ökologischen Inferenz ist es, die interessierenden Parameter β_i^b und β_i^w zu schätzen. Wissen wir etwa, dass der Anteil der Afroamerikaner an der Wählerschaft im Wahlkreis i 20 Prozent beträgt ($X_i = 0,2$), und dass es dort eine Wahlbeteiligung von 60 Prozent ($T_i = 0,6$) gab, dann könnten die Hälfte aller Afroamerikaner ($\beta_i^b = 0,5$) und gleichzeitig 62,5 Prozent aller Weißen ($\beta_i^w = 0,625$) zur Urne gegangen sein. Es könnte auch anders gewesen, und nur jeder vierte Afroamerikaner ($\beta_i^b = 0,25$) dafür aber 11 von 12 Weißen ($\beta_i^w \approx 0,92$) wählen gegangen sein. Beide Interpretationen sind möglich. Welche aber zutrifft, kann allein auf Grundlage der Aggregatdaten nicht entschieden werden.

Im Allgemeinen gilt nicht $\beta_i^b + \beta_i^w = 1$, da es sich um Anteilswerte aus unterschiedlichen Grundgesamtheiten handelt. Da sich die Wählerschaft im Wahlkreis i anteilig aus Afroamerikanern und Weißen zusammensetzt, ergibt sich

$$T_i = \beta_i^b \cdot X_i + \beta_i^w \cdot (1 - X_i) \quad (1)$$

Weil wir niemals von einem perfekten Zusammenhang ausgehen können, muss eine stochastische Abweichung der tatsächlichen Beobachtung von der Vorhersage in das Modell aufgenommen (ε_i) werden. Es ergibt sich daher (mit oder ohne Konstante parametrisiert):

$$T_i = \beta_i^w + (\beta_i^b - \beta_i^w) \cdot X_i + \varepsilon_i = \beta_i^b \cdot X_i + \beta_i^w \cdot (1 - X_i) + \varepsilon_i \quad (2)$$

Selbst unter der Annahme, dass der wahlkreisspezifische Fehlerterm ε_i sich wohl verhält (d.h., dass die ε_i unabhängig voneinander und unkorreliert mit X_i sind), ist die Schätzgleichung

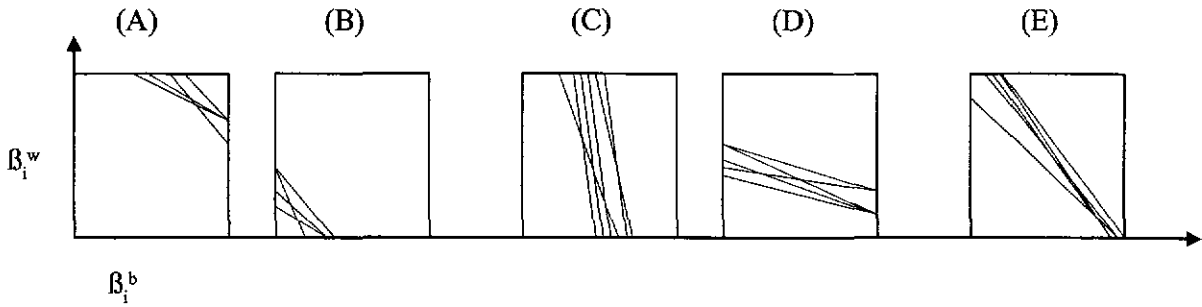


Abbildung 1: Idealtypische Muster von Tomographien

chung (2) nicht identifiziert, denn für jeden Wahlkreis müssen zwei β_i geschätzt werden, obwohl nur jeweils ein Datenpaar (X_i, T_i) zur Verfügung steht. Wir können also mit aggregierten Daten allein nicht den individuellen Zusammenhang beschreiben. Es gibt zwei Auswege aus dieser Situation: Man kann (a) angemessenere Daten benutzen oder (b) vereinfachende Annahmen treffen, die sich im günstigsten Fall auf vorhandenes Wissen gründen.

Der zweite Weg wird von der ältesten und bis vor einigen Jahren am weitesten verbreiteten Methode zur Schätzung von Gleichung (2) eingeschlagen, der *ökologischen Regression* (Ogburn und Goltra 1919, Falter 1991). Es wird die restriktive Annahme gemacht, dass die wahlkreisspezifischen Parameter im Durchschnitt gleich sind ($E(\beta_i^b) = \beta^b$ und $E(\beta_i^w) = \beta^w$). Schwankungen zwischen den Wahlkreisen werden dabei als unsystematisch aufgefasst und vom Fehlerterm ε'_i absorbiert. Mit Hilfe dieser Annahme lässt sich Gleichung (2) umformulieren als

$$T_i = \beta^b \cdot X_i + \beta^w \cdot (1 - X_i) + \varepsilon'_i \tag{3}$$

und mit einer linearen Regression unverzerrt schätzen (Goodman 1953).

Eine andere Technik, die ohne Zusatzannahmen auskommt und im Gegensatz zur ökologischen Regression für jeden Wahlkreis eine andere Schätzung für die interessierenden Parameter liefert, ist die *Methode der Ränder* (*method of bounds*, Duncan und Davis 1953). Da die Ränder T_i und X_i bekannt sind, lässt sich die deterministische Beziehung (1) umformulieren als

$$\beta_i^w = \frac{T_i}{1 - X_i} - \frac{X_i}{1 - X_i} \cdot \beta_i^b \tag{4}$$

Für jeden Wahlkreis entspricht (1') einer Geraden – β^w als Funktion von β^b – bei der Achsenabschnitt und Steigung bekannt sind. Diese Geraden können als so genannte *Tomographie* (King 1997: 81–82) dargestellt werden.

Da X_i im Einheitsintervall liegt, kann keine der durch (1') beschriebenen Geraden eine positive Steigung haben. Folglich ist β_i^b maximal (minimal), wenn $\beta_i^w = 0$ ($\beta_i^w = 1$) ist. Und da die β_i als Anteilswerte ebenfalls im Einheitsintervall liegen müssen, gilt:

$$\beta_i^b \in \left[\max \left(0, \frac{T_i - (1 - X_i)}{X_i} \right); \min \left(1, \frac{T_i}{X_i} \right) \right]$$

und

$$\beta_i^w \in \left[\max \left(0, \frac{T_i - X_i}{1 - X_i} \right); \min \left(1, \frac{T_i}{1 - X_i} \right) \right] \quad (5)$$

Generell ist die Information, die durch die Methode der Ränder aus den Daten herausgeholt wird, umso größer, je kürzer die Linien der Tomographie sind. Dementsprechend haben Datenstrukturen, die Diagonalen erzeugen (E), die also in beiden Parametern den gesamten Wertebereich von 0 bis 1 abdecken, einen sehr geringen Informationsgehalt. Ein sehr hoher Informationsgehalt für beide Parameter liegt vor, wenn die Linien über die Ecken verlaufen (A oder B). Ein hoher Informationsgehalt für die β_i^b liegt vor, wenn die Linien von oben nach unten verlaufen (C). Die Situation in (D) ist gerade umgekehrt. Hier liegt ein hoher Informationsgehalt für die β_i^w vor, wenn die Linien von rechts nach links verlaufen.

Die grundlegende Neuerung von Kings EI-Modell besteht in der Verbindung des Prinzips der ökologischen Regression mit der Methode der Ränder. Durch diese Verbindung wird die in den Daten tatsächlich vorhandene Information effektiver ausgeschöpft, indem für die Schätzung von β_i^b und β_i^w nur ein beschränkter – nämlich durch die Methode der Ränder bestimmter – Wertebereich zugelassen wird. Insbesondere müssen die interessierenden Parameter nicht gleich sein wie bei der ökologischen Regression. Im Unterschied zur Methode der Ränder schätzt EI für jeden Parameter auch Konfidenzintervalle.

Ein EI-Modell lässt sich mit einer Software schätzen, die einschließlich einer Dokumentation von Kings homepage (<http://gking.harvard.edu/stats.shtml>) als Gauss-Version („EI“) oder als stand-alone Windows Version („EzI“) kostenlos heruntergeladen werden kann. Die Software kann viele vordefinierte Graphiken zur Modelldiagnose und Veranschaulichung erstellen und bietet zudem die Möglichkeit, Statistiken und Ergebnisse auch in andere Programme zu exportieren.

19.2.3. Modellannahmen und Diagnostik

Hinter EI verbirgt sich eine Variante des ökologischen *random-effects* Regressionsmodells, bei dem die β_i innerhalb fester Grenzen zufällig um einen gemeinsamen Wert schwanken. Dabei ist die Abweichung von diesem Wert in einem bestimmten Wahlkreis bestimmbar. Die Modellannahmen sind weniger restriktiv und damit realistischer als beim herkömmlichen ökologischen Regressionsmodell. Folgende Annahmen werden getroffen: (1) Unimodalitätsannahme: β_i^b und β_i^w sind innerhalb deterministischer Grenzen, die durch die Methode der Ränder bestimmt werden, normalverteilt. Folglich ist (β_i^b, β_i^w) *trunkiert bivariat normalverteilt* (TBN-verteilt). Substanziell bedeutet das, dass die β_i von Wahlkreis zu Wahlkreis zwar nicht gleich aber ähnlich sind. (2) Die Fehlerterme sind nicht (räumlich) korreliert (siehe „Räumliche Regressionsmodelle“). (3) Es gibt keinen *Aggregationsbias*.

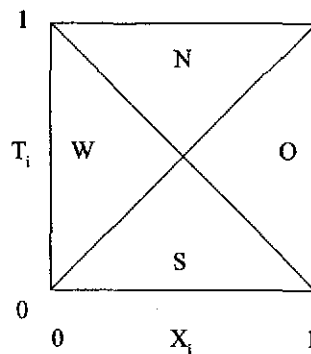


Abbildung 2: Wie informativ sind die Ränder?

Die geschätzten Koeffizienten lassen sich also nicht systematisch durch X vorhersagen. Ein Aggregationsbias würde vorliegen, wenn es zum Beispiel eine hohe Wahlbeteiligung afroamerikanischer Wähler nur in Wahlkreisen gäbe, in denen es viele afroamerikanische Wahlberechtigte gibt.

Die fünf Parameter der TBN-Verteilung (2 Mittelwerte, 2 Varianzen und eine Kovarianz) werden in einem ersten Schritt mit Maximum Likelihood geschätzt. Die substantiell interessierenden Schätzwerte ($\hat{\beta}_i^b$, $\hat{\beta}_i^w$) und ihre Konfidenzintervalle werden in einem zweiten Schritt durch (bayesianische) Simulation (siehe „Bayesianische Datenanalyse“) aus der zuvor geschätzten TBN-Verteilung ermittelt (King 1997: Kapitel 6–7).

Neben der Tomographie gibt ein einfaches Streudiagramm von T_i auf X_i bereits vor der EI-Schätzung Aufschluss darüber, wie informativ die Ränder sind. Man stelle sich ein Einheitsquadrat vor, das durch zwei Diagonalen ($T_i = X_i$ und $T_i = 1 - X_i$) in vier Dreiecke zerlegt wird (Abbildung 2).

Im westlichen Dreieck liegen die Wahlkreise (Beobachtungen), für die der Parameterraum von β_i^w durch die Methode der Ränder von oben und unten eingeschränkt ist, aber die Ränder keine Beschränkung für β_i^b liefern. Analog dazu befinden sich im östlichen Dreieck die Wahlkreise, für die der Parameterraum von β_i^b oben und unten eingeschränkt ist, die Ränder aber keine Beschränkung für β_i^w liefern. Die Wahlkreise im südlichen Dreieck haben einen von oben eingeschränkten Parameterraum für β_i^b und β_i^w . Die Wahlkreise im nördlichen Dreieck haben dagegen einen von unten eingeschränkten Parameterraum für β_i^b und β_i^w . Wenn Wahlkreise beschriftet werden, kann man eventuell Gemeinsamkeiten (d.h. räumliche Cluster) innerhalb der Punktwolke erkennen (siehe „Räumliche Regressionsmodelle“).

Neben der Abschätzung des Informationsgehaltes der Ränder kann mit einer Tomographie geprüft werden, ob die Unimodalitätsannahme verletzt ist. Für die Daten in Abbildung 3 muss davon ausgegangen werden, dass die zur Parameterschätzung genutzte ML-Funktion über mindestens zwei Maxima verfügt. Es liegt also eine Verletzung der Unimodalitätsannahme vor. Simulationen haben aber gezeigt, dass EI-Schätzergebnisse nicht stark von den Verteilungsannahmen abhängen, solange kein Aggregationsbias vorliegt.

Das Vorliegen eines Aggregationsbias kann nicht mit Sicherheit diagnostiziert werden. Erstens muss qualitatives Vorwissen genutzt werden, um die Gefahr des Vorliegens eines solchen Bias zu bestimmen. Zweitens muss das Vorliegen eines Aggregationsbias als

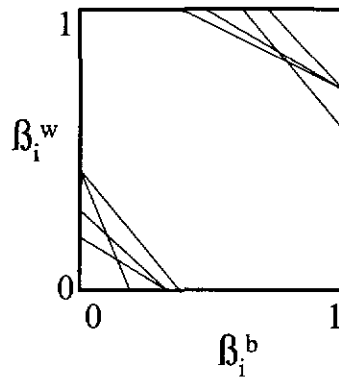


Abbildung 3: Diagnose mit Tomographien: Verletzung der Unimodalitätsannahme

wahrscheinlich angesehen werden, wenn die Regression von T auf X eine betragsmäßig große Steigung hat, wenn also die Regressionslinie das Einheitsquadrat (Abbildung 2) von Norden nach Süden schneidet. Drittens weisen hohe Korrelationen der geschätzten Parameterwerte $\hat{\beta}_i^b$ und $\hat{\beta}_i^w$ mit X_i darauf hin, dass ein Aggregationsbias vorliegen kann. Diese Parameterwerte können als Regressionskoeffizienten berechnet werden (in EzI der Menüpunkt READ *Eaggbias*). Zur grafischen Darstellung der Diagnostik werden Bias-Graphen genutzt (GRAPH *biasB*, *biasW*).

Für den Umgang mit einem Aggregationsbias gibt es mindestens vier Strategien. Man kann den Aggregationsbias (1) ignorieren, wenn der Grad der Trunkierung sehr hoch ist bzw. die Methode der Ränder eine sehr starke Einschränkung des Parameterraums bewirkt oder die Regressionsgerade eines Bias-Graphen von β_i^b auf X_i fast waagrecht ist. Man kann (2) das Vorwissen über den datengenerierenden Prozess nutzen, um mit entsprechenden A-priori-Verteilungen für die Parameter der TBN-Verteilung dieses Vorwissen in das Modell zu integrieren (MODEL *Config*). Weiterhin kann (3) im erweiterten EI-Modell (King 1997: Kapitel 9) versucht werden, die Ursachen des Bias direkt zu modellieren. Da T_i und β_i^b stark korrelieren, β_i^b aber nicht beobachtbar ist, benötigt man dazu eine Variable Z_i , die bei einer Regression von T auf X (und Z) den Koeffizienten von X praktisch 0 werden lässt. Schließlich kann man (4) die EI-Schätzung auf Teilpopulationen anwenden, in denen die Annahmen nicht verletzt sind bzw. die Parameter aufgrund der Methode der Ränder a priori nur in einem kleinen Bereich schwanken können. Teilpopulationen sollten generell getrennt geschätzt oder die Abhängigkeiten durch eine passende Variable modelliert werden, wenn unterschiedliche Datengenerierungsprozesse vermutet oder räumliche Korrelationen erwartet werden, die nicht unabhängig von X sind. Räumliche Korrelationen sind ein Indiz für das Vorliegen eines Aggregationsbias.

19.2.4. Datenstruktur und praktische Hinweise

Für die Anwendung von EI werden nur sehr geringe Anforderungen an die Daten gestellt. Ideal sind Daten, die möglichst nahe an der Analyseebene liegen. Zum einen erhöht das die Homogenität der zu untersuchenden Analyseeinheiten. Der beschränkende Einfluss der Methode der Ränder ist stärker und macht damit die Schätzung weniger anfällig für

Verletzungen der Modellannahmen. Zum anderen sorgt eine größere Nähe zur Analyseebene, die weniger Aggregation impliziert, für eine höhere Zahl analysierbarer Fälle.

Im allgemeinsten Fall muss ein für EI passendes Anwendungsproblem als Schätzung einer Vierfeldertafel (Tabelle 1) formulierbar, das heißt auf den 2×2 -Fall zurückführbar sein. Mit dem erweiterten EI-Modell gibt es zwar auch die Möglichkeit, einen $2 \times C$ -Fall zu schätzen, jedoch bereitet die praktische Anwendung für $C > 3$ bislang massive Konvergenzprobleme. Mögliche Auswege sind hier, die Schätzung verschiedener 2×3 Fälle mit dem Verfahren der bayesianischen Modelldurchschnittsbildung (Imai und King 2004) zu kombinieren. Die Forschungspraxis ist zur Beantwortung ihrer Fragen jedoch häufig auf höherdimensionale Tabellen ($R \times C$ -Fall) angewiesen. Diese Tabellen angemessen zu modellieren erweist sich bislang als schwierig (King et al. 2004). Derzeit bleiben zwei Auswege: Man kann erstens das allgemeine Schätzproblem iterativ in verschiedene Teilprobleme zerlegen, die dann sukzessive geschätzt werden. Dieses Verfahren ist allerdings ineffizient und kann Modellannahmen verletzen (Ferree 2004). Zweitens kann man sich eines anderen Verfahrens der ökologischen Inferenz bedienen, das gleichfalls die Möglichkeit bietet, deterministische Information der Ränder für die Schätzung effizient auszunutzen. Ein solches Verfahren wäre EMax (Johnston/Pattie 2000), welches allerdings höhere Ansprüche an die zugrunde liegenden Daten stellt, oder speziell für Wählerwanderungsanalysen ECOL (Thomsen 1987).

19.2.5. Innermethodische Kritik

EI wird vornehmlich wegen seiner Annahmen bzw. der (mangelnden) Robustheit der Schätzwerte kritisiert (Cho/Gaines 2004). Diese Kritik übersieht, dass im Vergleich zu anderen Verfahren durch die Methode der Ränder schon sehr viel mehr Information für die Schätzung verwendet wird. Die EI-Modellannahmen dienen lediglich dazu, die interessierenden Parameter über die Beschränkung, die durch die Methode der Ränder geliefert wird, hinaus noch weiter einzuschränken. EI ist daher a priori wesentlich unabhängiger von Verletzung der Verteilungsannahmen als die ökologische Regression. Zudem ist es möglich (Gschwend 2004: 66), durch wiederholte Schätzung und Integration der gewonnenen Ergebnisse, analog dem Verfahren der multiplen Imputation (siehe „Multiple Imputation“), den Grad der Robustheit der EI-Schätzungen in die Analyse zu integrieren.

Ein anderer Strang der Kritik hinterfragt die Konsistenz von Schätzergebnissen, die auf einer Weiterverarbeitung der EI-Schätzwerte als Variablen in Regressionsmodellen beruhen. Unter bestimmten Voraussetzungen können in einem zweiten Schritt gewichtete Regressionsmodelle geschätzt werden, wenn zuvor die abhängige Variable mittels EI geschätzt wurde (Adolph et al. 2003).

19.2.6. Anwendungsgebiete

Ökologische Inferenz wird immer dann notwendig, wenn es um den Zusammenhang von Eigenschaften geht, für die Daten auf der eigentlich interessierenden Analyseebene nicht zur Verfügung stehen und die daher mit höher aggregierten Daten bearbeitet werden müssen. In der politischen Soziologie sind das neben dem Stimmensplitting oder Wählerwanderungsanalysen Fragen nach kontextuellen Einflüssen auf Wahlverhalten oder

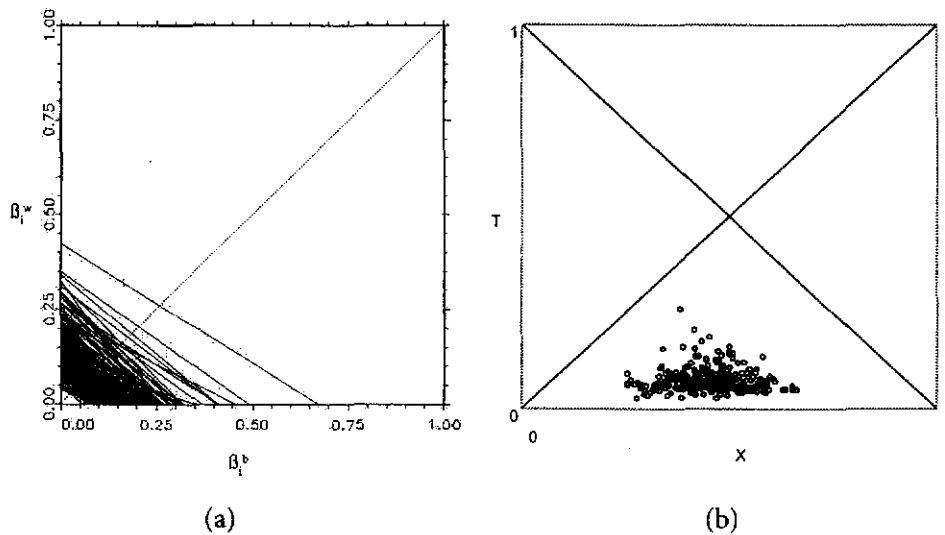


Abbildung 4: Diagnostische Graphen

-teilnahme bestimmter sozialer Gruppen. Weiterhin kann ökologische Inferenz zur Analyse von geheimem Abstimmungsverhalten in Gremien (z.B. Bundesverfassungsgericht) oder Organisationen dienen. Selbst die Umfrageforschung oder experimentelle Designs könnten von ökologischer Inferenz profitieren, um mehrere Querschnittsstudien als Quasi-Panelstudien zu analysieren, oder um dem verzerrenden Einfluss der sozialen Erwünschtheit zu begegnen.

19.3. Anwendungsbeispiel

Im Folgenden soll das Stimmensplitting zwischen SPD und den Grünen untersucht werden. Genauer soll es um die Frage gehen, in welchem Umfang Wähler, die ihre Zweitstimme den Grünen gegeben haben, sich mit ihrer Erststimme für die SPD entschieden. Gemäß der eingeführten Notation entspricht T_i dem Zweitstimmenanteil der Grünen im i -ten Wahlkreis, X_i dem Erststimmenanteil der SPD, β_i^b dem Anteil der Grünen Zweitstimmenwähler, die mit ihrer Erststimme SPD gewählt haben und β_i^w dem Anteil der Grünen Zweitstimmenwähler, die mit ihrer Erststimme etwas anderes als die SPD wählten.

Aufgrund des unterschiedlichen Parteiensystems in West und Ost ist es a priori plausibel zu erwarten, dass in einer gesamtdeutschen Betrachtung die Unimodalitätsannahme verletzt ist. Daher ist es ratsam, die Schätzungen für West und Ost getrennt durchzuführen. Exemplarisch werden hier westdeutsche Wahlkreisdaten der Bundestagswahl 1998 analysiert.

Die Tomographie wie auch das Streudiagramm zeigen übereinstimmend, dass viel Information in den Rändern vorhanden ist. Die Linien der Tomographie sind sehr kurz. β_i^b kann mit einigen wenigen Ausnahmen nur in einem kleinen Bereich schwanken. Die längste Linie gehört zum Wahlkreis Berlin-Kreuzberg (WK 255), der den größten Wert für T ($\approx 0,25$) besitzt. Dieser Wert kann als Schnittpunkte der Linien mit der ersten Winkelhalbierenden der Tomographie abgelesen werden. Alle Punkte des Streudiagramms liegen nahe an der X-Achse im südwestlichen Dreieck des Einheitsquadrates. Die Ränder schränken den möglichen Parameterraum von β_i^b also stark von oben ein.

Tabelle 2: Das Inferenzproblem im Wahlkreis 221

Erststimme	Zweitstimme		
	B90/Grüne	Andere	
SPD	β_{221}^b	$1 - \beta_{221}^b$	0,47
Andere	β_{221}^w	$1 - \beta_{221}^w$	0,53
	0,026	0,974	1

Die Instrumente zur Diagnose eines Aggregationsbias zeigen, dass ein solcher Bias höchstens geringe Auswirkungen haben kann. Der Koeffizient der Regression von T auf X , $\hat{\beta}^b - \hat{\beta}^w$, ist betragsmäßig sehr klein. Die Regressionsgerade schneidet das Einheitsquadrat also weder auf der Nord- noch auf der Süd-Seite. Weiterhin zeigen die Maximum-Likelihood-Kontourlinien in der Tomographie, dass die geschätzte bivariate Normalverteilung stark trunziert ist. Schließlich deutet auch der graphische Test (*GRAPH biasB*) nicht auf Vorliegen eines Aggregationsbias hin, denn die Regressionsgerade von $\hat{\beta}^b$ auf X verläuft fast horizontal. Die zugrundeliegenden Regressionen (*READ Eaggbias*) zeigen zwar, dass die Steigung signifikant von 0 verschieden ist. Das spielt aber substantiell nur eine geringe Rolle, da der Betrag der Steigung klein ist. Offensichtlich gibt es in SPD-Erststimmendiasporagebiete einen etwas kleineren Anteil der SPD-Erststimmenwählerschaft, die die Grünen mit der Zweitstimme wählen, als in Erststimmenhochburgen. Abschließend sei nochmals darauf hingewiesen, dass die aus dem Vorwissen um die Unterschiedlichkeit der Parteiensysteme in Ost und West durchgeführte getrennte Schätzung für beide Regionen mit großer Wahrscheinlichkeit einen erheblichen Beitrag zur Vermeidung eines Aggregationsbias geleistet hat.

Schauen wir uns nun als Beispiel den Wahlkreis Weiden ($i = 221$) bei der Bundestagswahl 1998 als Vierfeldertafel etwas genauer an. In Abbildung 4b sehen wir, dass β_{221}^b stark von oben (nicht aber von unten) eingeschränkt ist, also insbesondere kleiner als 1 sein muss. Aufgrund der Ränder und unter Nutzung der Gleichung (3) wissen wir ohne jegliche statistischen Annahmen mit Sicherheit, dass β_{221}^b nicht nur irgendwo im Einheitsintervall, sondern genau zwischen und 0 – es gibt keine Wähler, die ihre Erststimme der SPD und ihre Zweitstimme den Grünen geben – und 0,056 ($\approx 0,026/0,47$) – alle Grünen Zweitstimmenwähler (2,6 Prozent) geben ihre Erststimme der SPD – liegen muss. Mit der Ökologischen Regression ermittelt man den Anteil der SPD-Grüne-Stimmensplitter auf 3,1 Prozent ($\hat{\beta}^b = 0,065$), während EI diesen Anteil nur auf 0,4 Prozent ($\hat{\beta}_{221}^b = 0,009$) schätzt, wobei das zugehörige 95-Prozent-Konfidenzintervall von 0,03 Prozent bis 1,09 Prozent reicht.

Wie gut sind diese Schätzungen? Wenn das für jeden Wahlkreis bekannt wäre, dann hätte man für diese Fragestellung ein anderes Verfahren anwenden können. Es sollte aber immer versucht werden, die geschätzten Ergebnisse zu validieren. Im vorliegenden Fall eignet sich die repräsentative Wahlstatistik oder die regelmäßig in Umfragen auftauchende Wahlabsichtsfrage. Auf Länder- bzw. Bundesebene könnten diese Informationen aggregiert werden, um sie dann mit dem aggregierten EI-Schätzwert, als gewichtetem Mittelwert der einzelnen EI-Schätzwerte ($\hat{\beta}_i^b, \hat{\beta}_i^w$), zu vergleichen. Überschneiden sich die Konfidenz-

intervalle, dann hat man starke Evidenz dafür, dass die ökologische Inferenz tatsächliches individuelles Verhalten hinreichend gut beschreibt.

19.4. Kommentierte Literaturempfehlungen

Als Standardwerk sollte King (1997) auf jeden Fall vor einer Anwendung von EI durchgearbeitet werden. Die lautesten Kritiker von EI, die außerhalb des Methodenpublikums der Zeitschrift *Political Analysis* in der Politikwissenschaft gehört wurden, sind Cho und Gaines (2004). Inhaltlich beziehen sie sich auf die Arbeit von Burden und Kimball (1998), der ersten und damit wohl bekanntesten Anwendung von EI, die in einer politikwissenschaftlichen Fachzeitschrift veröffentlicht wurde. Als weiterführende aber auch als vertiefende Literatur ist King et al. (2004) geeignet. Das hier diskutierte Anwendungsbeispiel ist aus Gschwend (2004). Insbesondere die dortigen Kapitel 4 und 5 sind aus methodischer Sicht und für eine detaillierte Darstellung des Anwendungsbeispiels hilfreich.

19.5. Literatur

- Adolph, Christopher/King, Gary/Herron, Michael C./Shotts, Kenneth W., 2003: A Consensus on Second-Stage Analyses in Ecological Inference Models. *Political Analysis* 11: 86–94.
- Burden, Barry C./Kimball, David C., 1998: A New Approach to the Study of Ticket Splitting. *American Political Science Review* 92: 533–544.
- Cho, Wendy K. Tam/Gaines, Brian J., 2004: The Limits of Ecological Inference: The Case of Split-Ticket Voting. *American Journal of Political Science* 48: 152–171.
- Duncan, Otis Dudley/Davis, Beverly, 1953: An Alternative to Ecological Correlation. *American Sociological Review* 18: 665–666.
- Falter, Jürgen W., 1991: Hitlers Wähler. München: Beck.
- Ferree, Karen, 2004: Iterative Approaches to R x C Ecological Inference Problems: Where They Can Go Wrong and One Quick Fix. *Political Analysis* 12: 143–159.
- Goodman, Leo A., 1953: Ecological Regressions and Behavior of Individuals. *American Sociological Review* 18: 663–664.
- Gschwend, Thomas, 2004: Strategic Voting in Mixed-Electoral Systems. Reutlingen: SFG-Servicecenter Fachverlag.
- Herron, Michael C./Shotts, Kenneth W., 2004: Logical Inconsistency in EI-Based Second-Stage Regressions. *American Journal of Political Science* 48: 172–183.
- Imai, Kosuke/King, Gary, 2004: Did Illegal Overseas Absentee Ballots Decide the 2000 US Presidential Election? *Perspectives on Politics* 2: 537–549.
- Johnston, Ronald J./Pattie, Charles J., 2000: Ecological Inference and Entropy-Maximizing: An Alternative Estimation Procedure for Split-Ticket Voting. *Political Analysis* 8: 333–345.
- King, Gary, 1997: A Solution to the Ecological Inference Problem. Reconstructing Individual Behaviour from Aggregate Data. Princeton: Princeton University Press.
- King, Gary/Rosen, Ori/Tanner, Martin A. (Hg.), 2004: Ecological Inference: New Methodological Strategies. New York: Cambridge University Press.
- Ogburn, William F./Goltra, Inez, 1919: How Women Vote: A Study of an Election in Portland, Oregon. *Political Science Quarterly* 34: 413–433.
- Pappi, Franz Urban, 1977: Aggregatdatenanalyse. In Kolwijk, Jürgen van/Wieken-Mayser, Maria (Hg.), *Techniken der empirischen Sozialforschung* (Band 7), München: Oldenbourg Verlag. S. 78–110.

Robinson, William S., 1950: Ecological Correlation and the Behavior of Individuals. *American Sociological Review* 15: 351–357.

Thomsen, Søren R., 1987: Danish Elections 1920–79. A Logit Approach to Ecological Analysis and Inference .