



Wählerstromanalyse und Wahlhochrechnung

Johann Bacher Kurt Holm

Almo Statistik-System

www.almo-statistik.de

holm@almo-statistik.de

kurt.holm@jku.at

johann.bacher@jku.at

Autoren: Prof. Dr. Johann Bacher, Universität Linz, Österreich
em. Prof. Dr. Kurt Holm, Universität Linz, Österreich

Im Text wird häufig auf das Dokument **P0** Bezug genommen. Dabei handelt es sich um das Almo-Dokument "Arbeiten mit Almo.PDF" (Dokument 0).

Weitere Almo-Dokumente

Die folgenden Dokumente können alle kostenlos von der Handbuchseite in www.almo-statistik.de heruntergeladen werden

0. Arbeiten_mit_Almo.PDF (1 MB)
- 1a. Eindimensionale Tabellierung.PDF (1.8 MB)
- 1b. Zwei- und drei-dimensionale Tabellierung.PDF (1.1 MB)
2. Beliebig-dimensionale Tabellierung.PDF (1.7 MB)
3. Nicht-parametrische Verfahren.PDF (0.9 MB)
4. Kanonische Analysen.PDF (1.8 MB)
Diskriminanzanalyse.PDF (1.8 MB)
enthält: Kanonische Korrelation, Diskriminanzanalyse, bivariate Korrespondenzanalyse, optimale Skalierung
5. Korrelation.PDF (1.4 MB)
6. Allgemeine multiple Korrespondenzanalyse.PDF (1.5 MB)
7. Allgemeines ordinales Rasch-Modell.PDF (0.6 MB)
- 7a. Wie man mit Almo ein Rasch-Modell rechnet.PDF (0.2 MB)
8. Tests auf Mittelwertsdifferenz, t-Test.PDF (1,6 MB)
9. Logitanalyse.pdf (1,2MB) enthält Logit- und Probitanalyse
10. Koeffizienten der Logitanalyse.PDF (0,06 MB)
11. Daten-Fusion.PDF (1,1 MB)
12. Daten-Imputation.PDF (1,3 MB)
13. ALM Allgemeines Lineares Modell.PDF (2.3 MB)
- 13a. ALM Allgemeines Lineares Modell II.PDF (2.7 MB)
- 13b. Bootstrap bei Allgemeinem Linearem Modell III.PDF
14. Ereignisanalyse: Sterbetafel-Methode, Kaplan-Meier-Schätzer, Cox-Regression.PDF (1,5 MB)
15. Faktorenanalyse.PDF (1,6 MB)
16. Konfirmatorische Faktorenanalyse.PDF (0,3 MB)
17. Clusteranalyse.PDF (3 MB)
18. Pisa 2012 Almo-Daten und Analyse-Programme.PDF (17 KB)
19. Guttman- und Mokken-Skalierung.PFD (0.8 MB)
20. Latent Structure Analysis.PDF (1 MB)
21. Statistische Algorithmen in C (80 KB)
22. Conjoint-Analyse (PDF 0,8 MB)
23. Ausreisser entdecken (PDF 170 KB)
24. Statistische Datenanalyse Teil I, Data Mining I
25. Statistische Datenanalyse Teil II, Data Mining II
26. Statistische Datenanalyse Teil III, Arbeiten mit Almo-Datenanalyse-System
27. Mehrfachantworten, Tabellierung von Fragen mit Mehrfachantworten (0.8 MB)
28. Metrische multidimensionale Skalierung (MDS) (0,4 MB)
29. Metrisches multidimensionales Unfolding (MDU) (0,6 MB)
30. Nicht-metrische multidimensionale Skalierung (MDS) (0,5 MB)
31. Pfadanalyse als wiederholte Regressionsanalyse (0,7 MB)
32. Datei-Operationen mit Almo (1,1 MB)
33. Wählerstromanalyse und Wahlhochrechnung (1,6 MB)
34. Soziometrie. Auswertung soziometrischer Daten (0,5 MB)
35. Konfidenzintervall und p-Wert beim Bootstrap-Verfahren (200 KB)

Inhaltsverzeichnis

Kurt Holm

P51.0 Einführung	4
P51.1 Überblick: Die Verfahren	5
P51.1.1 Einfache lineare Regression.....	6
P51.1.2 Multiple lineare Regression	9
P51.1.3 Quadratische Optimierung	13
P51.1.4 Weitere Modelle	18
P51.2 Wählerstromanalyse und Wahlhochrechnung	21
P51.3 Einige Probleme.....	21
P51.3.0 Nichtwähler und Ungültigwähler	21
P51.3.1 Zeitpunkt der Hochrechnung	22
P51.3.2 Uneinheitlicher Trend.....	22
P51.3.3 Auswahl der Stimmbezirke.....	23
P51.3.4 Prozentwerte oder absolute Stimmenzahlen	24
P51.3.5 Ungleiche Zahl der Wahlberechtigten bei den beiden Wahlen.....	24
P51.3.6 Splitterparteien	26
P51.3.7 Prognoseerfolg.....	26
P51.3.8 Mandatsverteilung	26
P51.3.9 Briefwahl und Wahlkarten.....	27
P51.4 Wählerstromanalyse und Hochrechnung mit ALMO.....	27
P51.4.1 Eingabe in das Almo-Programm zur Hochrechnung und Wählerstromanalyse	28
P51.4.1.1 Schritte der Hochrechnung	28
P51.4.2 Die Dateien der vergangenen und der aktuellen Wahl.....	29
P51.4.3 Gesamtzahl der Zuwanderer mit Prog51mb ermitteln	32
P51.4.4 Eingabe in das Hochrechnungsprogramm Prog51ml	37
P51.4.4.1 Erläuterungen zu den Eingabeboxen von Prog51ml.....	44
P51.4.5 Ausgabe der Ergebnisse aus Hochrechnung	75
P51.4.5.1 Übergangsmatrix aus quadratischer Optimierung und multipler Regression	83
P51.4.6 Hochrechnung mit oder ohne Zu- und Abwanderer.....	85
P51.4.7 Ausgabe der Ergebnisse aus Wählerstromanalyse	89
P51.4.7.1 Wählerstromanalyse bei vollständiger Auszählung	89
P51.4.7.2 Wählerstromanalyse bei unvollständiger Auszählung	94
P51.4.8 Hochrechnung/Wählerstromanalyse mit Gruppen von Wahlgemeinden	94
P51.4.8.1 Hochrechnung / Wählerstromanalyse für eine einzelne Gemeindegruppe	95
P51.4.8.2 Schluss von einer Gemeindegruppe auf das Gesamtergebnis.....	99
P51.4.8.3 "Gruppierte" Hochrechnung/Wählerstromanalyse.....	100
P51.4.9 Gewichtete Hochrechnung.....	108
P51.4.9.1 Berechnung der Gewichtszahlen.....	110

Johann Bacher

P51.5 Das Modell	115
P51.6 Modellprüfgrößen.....	122
P51.6.1 Eingabe in Programm	123
P51.6.2 Ergebnisse aus Programm.....	124
P51.7 Residuenanalyse.....	125
Literatur	128

P51.0 Einführung

Die Autoren

Das Programm zur Wahlhochrechnung und Wählerstromanalyse wurde von Johann Bacher, Helmut Gferer und Kurt Holm programmiert. Gferer hat das zentrale Unterprogramm zur quadratischen Optimierung geschrieben, wobei er den Algorithmus nach Wilkinson/Reinisch (1971) verwendet hat. Bacher hat große Programmteile der Wählerstromanalyse und die Unterprogramme für die Modell-Prüfgrößen und die Residuenanalyse programmiert. Holm hat das Hochrechnungsprogramm, die "gruppierte" und die "gewichtete" Hochrechnung und die Grafik programmiert sowie die Programm-Masken entwickelt. Alle Programmteile sind in C geschrieben.

Im vorliegenden Text wurden die Abschnitte P51.0 bis P51.4 von Holm verfasst und die Abschnitte P51.5 bis P51.7 von Bacher.

Die Programme

Im vorliegenden Text wird Programm **Prog51m1** ausführlich beschrieben. Es leistet Wahl-Hochrechnung und Wählerstromanalyse nach dem Kalkül der quadratischen Optimierung. In Almo sind dann weiter noch enthalten:

Prog28m2 zur Wählerstromanalyse. Es verwendet ebenfalls den Kalkül der quadratischen Optimierung und ist im Programmcode weitgehend gleich zu Prog51m1. Es kann jedoch nicht zur Hochrechnung eingesetzt werden. Die Ergebnisausgabe und die Grafik ist gegenwärtig noch nicht auf den Stand von Prog51m1 gebracht worden.

Prog53m3 und **Prog53m4** zur Hochrechnung. Die beiden Programme verwenden den Kalkül der einfachen Regressionsanalyse ohne und mit Konstante.

Prog53m2 zur Hochrechnung nach dem Kalkül der multiplen Regression.

Almo enthält dann noch Programme zur Mandatsverteilung. Diese und die Programme zur einfachen und multiplen Regression wurden von Holm geschrieben. Alle diese Programme werden gefunden durch Klick auf den Knopf "Verfahren/Wahlhochrechnung" am Oberrand des Almo-Fensters oder durch Klick auf den Knopf "alle Progs". Prog53m2 und Prog53m4 werden kurz in den Abschnitten P51.1.1 und P51.1.2 vorgestellt. Ansonsten wird im vorliegenden Text ausschließlich Prog51m1 (zur quadratischen Optimierung) behandelt.

Programm-Anwendung

Eine Wahlhochrechnung am Wahlabend durchzuführen ist nur einer großen Organisation mit behördlicher Unterstützung möglich. Für den Leser dieses

vorliegenden Textes wird dies nicht möglich sein. Er kann jedoch eine Hochrechnung *nach* der Wahl rechnen. Dies ist dann, wie wir noch zeigen werden, eine Wählerstromanalyse. Der Algorithmus der "quadratischen Optimierung", der in der Hochrechnung eingesetzt wird, liefert eine Übergangsmatrix von den Parteien der vergangenen Wahl zu den Parteien der aktuellen Wahl. Dieser Algorithmus ist dazu noch *konsistent*. D.h. sind alle Wahlgemeinden der Wahl ausgezählt, dann "prognostiziert" das Verfahren die tatsächlichen Wähleranteile der Parteien.

Die für die Wählerstromanalyse notwendigen Daten sind öffentlich zugänglich - tatsächlich jedoch nicht so schnell und komplikationslos, wie man sich das wünschen würde und auch nicht immer kostenfrei. Der Zweck unseres Programms und dieses Textes ist es also eher, dem Almo-Benutzer zu ermöglichen, mit Daten, die für ihn verfügbar sind, zu experimentieren und spezielle und vertiefte Wählerstromanalysen zu rechnen. Auch mit unseren Daten kann der Benutzer rechnen und experimentieren.

Unsere Daten

Wir verwenden die sehr alten aber vollständigen Daten der Landtagswahlen von 1997 in Oberösterreich und die Daten der vorausgegangenen Wahl 1991. Die Daten der Landtagswahl 1997 enthalten eine Variable, in der die Uhrzeiten angegeben sind, zu denen die Gemeinden ihre Wahlergebnisse an die Zentrale gemeldet haben. Der Benutzer, der mit unserem Prog51m1 experimentieren will, kann somit Hochrechnungen für beliebige Zeitpunkte rechnen. Da das tatsächliche Ergebnis bekannt ist, kann die Güte der "Prognose" zum jeweiligen Zeitpunkt beurteilt werden. In Almo verfügbar sind dann auch noch die Daten der Landtagswahlen 1985 zu 1991.

Der Begriff "Wahlgemeinde"

In den folgenden Ausführungen werden wir oft den Begriff "Wahlgemeinde" oder kurz "Gemeinde" verwenden. In unseren Beispiel-Daten ist das tatsächlich eine politische Gemeinde.

Allgemein gilt jedoch: Gemeint ist mit diesem Begriff die *Untersuchungseinheit* der Hochrechnung bzw. Wählerstromanalyse. In der Datenmatrix, die dem Kalkül unterworfen wird, ist dies eine Zeile. Eine "Wahlgemeinde" kann also auch ein "Wahlsprenkel" aus einer Großstadt sein.

P51.1 Überblick: Die Verfahren

Die einfachste Art und Weise eine Wahlhochrechnung durchzuführen, besteht darin, am Wahlabend zu beobachten, wie die Parteien A, B usw. prozentuell von einer Wahlgemeinde zur anderen gewinnen bzw. verlieren. Man beobachtet etwa, dass die Partei A in der Wahlgemeinde i, deren Wahlergebnisse um 18:15 Uhr eintreffen, 1,8% dazu gewinnen konnte. In der Gemeinde j (einige wenige Minuten später) gewinnt sie 1,9% dazu. So kann man etwa im Verlauf einer halben Stunde feststellen, dass die Partei A im Durchschnitt in den Gemeinden, deren Ergebnisse eingetroffen sind, 1,8% dazu gewinnt. Wenn man nun prognostiziert, dass die Partei A bei dieser Wahl 1,8% dazu gewinnen wird, dann wird man vermutlich nicht sehr weit vom tatsächlichen Wahlergebnis entfernt liegen.

P51.1.1 Einfache lineare Regression

Diese intuitive, etwas schlicht anmutende Methode kann man aufwerten, wenn man sie als "einfache lineare Regression mit Konstante" durchführt. Siehe dazu Bruckmann, einen der "Väter" der Wahlhochrechnung, der dieses Verfahren schon 1966 beschrieb (in Kapitel 3.3). Das Verfahren wurde von Holm programmiert. Es steht als Masken-Programm Prog53m4 und als Beispiel-Programm P53m4_Gr zur Verfügung. Man findet es durch Klick auf den Knopf "Verfahren/Wahlhochrechnung" am Oberrand des Almo-Fensters oder durch Klick auf den Knopf "alle Progs", ebenfalls am Oberrand des Almo-Fensters.

Die Partei X aus der *aktuellen* Wahl wird als abhängige Variable in eine Regressionsanalyse eingegeben. Als unabhängige Variable wird dieselbe Partei X aus der *vergangenen* Wahl verwendet. So wird für alle Parteien verfahren. Jede Partei der aktuellen Wahl muss ein Pendant aus der vergangenen Wahl besitzen. Tritt z.B. in der aktuellen Wahl eine neue Partei an, dann muss für sie - als ihr Pendant in der vergangenen Wahl - eine künstliche "Leer-Partei" (mit 0 Stimmen) eingesetzt werden. Auch im umgekehrten Fall, dass eine Partei der vergangenen Wahl bei der aktuellen nicht mehr antritt, sollte eine "Leer-Partei" für die aktuelle Wahl eingesetzt werden oder die Partei muss aus der Analyse herausgenommen werden.

Wir rechnen mit Beispiel-Programm **P53m4_Gr.Alm** eine Hochrechnung für die Landtagswahl 1997 auf Basis der Wahl von 1991. Wir verwenden die bereits oben beschriebenen Daten und rechnen bei vollständiger Auszählung aller Gemeinden. Die Prognose-Ergebnisse des Regressionsmodells können so mit den tatsächlichen Wahlergebnissen verglichen werden. Der Almo-Benutzer kann mit unseren Beispieldaten auch eine Hochrechnung für den Zeitpunkt 14.00 Uhr rechnen. Zu diesem Zeitpunkt waren erst 35 Gemeinden ausgezählt. In der Programm-Maske muss der Benutzer dazu die Eingabebox "Simulation einer Hochrechnung" öffnen und eintragen: `wenn Zeit97 kleiner 1400`

Der Almo-Benutzer findet das Programm P53m4_Gr durch Klick auf den Knopf "alle Progs" am Oberrand des Almo-Fensters. Das Programm entstand aus der Programm-Maske Prog53m4.Msk dadurch, dass verschiedene Kleinstparteien aus der aktuellen Wahl zu einer Partei "Sonstige97" zusammengefasst wurden. Alle diese Kleinstparteien waren bei der vergangenen Wahl nicht angetreten. Als Pendant zu "Sonstige97" wird eine künstliche "Leer-Partei" mit dem Namen "Sonstige91" bei der vergangenen Wahl eingesetzt.

Die Parteien, die bei der vergangenen Wahl 1991 und der aktuellen Wahl 1997 antraten, waren

Parteien 1997	Parteien 1991
OEVP97	OEVP91
SPOE97	SPOE91
FPOE97	FPOE91
GRUENE97	VGOE91 }
-	GAL91 } zusammengefasst zu: VGOE_GAL91
LIF97 }	-
NEUTRALE97 }	-
KPOE97 }	-
CWG97 } Sonstige97	-
FDP97 }	-
OENP97 }	-
Ungültigwähler97	Ungültigwähler91
Nichtwähler97	Nichtwähler91

VGOE91 und GAL91 schlossen sich 1997 zu GRUENE97 zusammen. Sie werden deswegen auch für 1991 zusammengefasst zu VGOE_GAL91. Die Kleinstparteien LIF97 bis OENP97 werden zu Sonstige97 zusammengefasst. Für 1991 wird als Pendant die "Leer-Partei" Sonstige91 gebildet.

Das Programm liefert folgende Matrix der Regressionskoeffizienten

	OEVP97	SPOE97	FPOE97	GRUENE97	Sonst97	Ungült97	Nichtw97
OEVP91	0.972	0	0	0	0	0	0
SPOE91	0	0.761	0	0	0	0	0
FPOE91	0	0	0.805	0	0	0	0
VGOE_GAL91	0	0	0	0.622	0	0	0
Sonst91	0	0	0	0	0	0	0
Ungült91	0	0	0	0	0	0.221	0
Nichtw91	0	0	0	0	0	0	0.615
Konstante	-0.021	0.020	0.048	0.017	0.031	0.018	0.083

Die Regressionskoeffizienten beginnen mit 0, können aber größer 1.0 werden. D.h. sie dürfen nicht als Übergangswahrscheinlichkeiten interpretiert werden. Die Konstanten können negativ sein, werden aber nicht größer als 1.0.

Die Regressionskoeffizienten stehen nur in der Diagonale. Matrixglieder außerhalb der Diagonale sind 0. Dadurch wird ersichtlich, dass das Wahlergebnis der Partei X97 aus der aktuellen Wahl durch ihr Pendant X91 aus der vergangenen Wahl direkt bestimmt wird. Aber nicht nur. Die Konstante weist darauf hin, dass auch die anderen Parteien das Wahlergebnis der Partei X97 beeinflussen.

Die Stimmensummen der Parteien bei der vergangenen Wahl 1991 waren folgende:

	Stimmen in vergangener Wahl	
OEVP91	357771	
SPOE91	248640	
FPOE91	140311	
VGOE_plus_GAL91	44853	
Sonstige91	0	("Leer-Partei")
Ungültigwähler91	16315	
Nichtwähler91	140355	

	948245	

Damit kann die für die aktuelle Wahl prognostizierte Stimmenszahl und daraus die Prozentverteilung errechnet werden. Für beispielsweise ÖVP97 entsteht

$$\text{ÖVP97} = r(\text{ÖVP}) * S(\text{ÖVP}) + k(\text{ÖVP}) * G$$

$r(\text{ÖVP})$ = Regressionskoeffizient für ÖVP (=0.972)

$S(\text{ÖVP})$ = Stimmen für ÖVP in vergangener Wahl (=357771)

$k(\text{ÖVP})$ = Konstanten-Koeffizient für ÖVP (= -0.021)

G = Gesamt-Stimmenszahl aller Gemeinden in der vergangener Wahl bzw. Wahlberechtigte in der vergangenen Wahl (=948245)

Wir rechnen mit 6 Kommastellen

$$\text{ÖVP97} = 0.971928 * 357771 - 0.020556 * 948245 = 328235.528268$$

Von der ÖVP91 und von den anderen Parteien aus 1991 sind 328235 Stimmen zur ÖVP97 abgeströmt - wobei nicht identifizierbar ist wieviele Stimmen von den

jeweils anderen Parteien aus 91 zur ÖVP97 geströmt sind. Das Wahlergebnis für eine Partei X97 in der aktuellen Wahl wird also hauptsächlich relativ zum Ergebnis dieser Partei X91 in der vergangenen Wahl prognostiziert. Wählerwanderungen zwischen den verschiedenen Parteien werden nicht erkannt. Dies ist eine sehr schlichte Hochrechnungs-Methode. Bei stabilen politischen Konstellationen kann sie jedoch durchaus erfolgreich sein. Wählerstromanalyse sind methodenbedingt nicht möglich.

Für alle Parteien entsteht entsprechend obiger Formel folgende Matrix

Multiplikation der Regressionskoeffizienten mit Gesamt-Stimmenzahl aus vergangener Wahl ergibt "Quasi-Wählerströme" (Abströme in Zahl der Stimmen)

	OEVP97	SPOE97	FPOE97	GRUENE97	Sonst97	Ungült97	Nichtw97	GESAMT
OEVP91	347727	0	0	0	0	0	0	347727
SPOE91	0	189149	0	0	0	0	0	189149
FPOE91	0	0	112931	0	0	0	0	112931
VGOEGAL91	0	0	0	27896	0	0	0	27896
Sonstig91	0	0	0	0	0	0	0	0
Ungült91	0	0	0	0	0	3599	0	3599
Nichtw91	0	0	0	0	0	0	86269	86269
Konstante	-19492	18500	45298	16421	29502	16902	79023	186153
GESAMT	328235	207648	158229	44317	29502	20500	165292	-

Die letzte Zeile "Gesamt" enthält die vom Modell prognostizierten Stimmen je Partei. Die letzte Spalte "Gesamt" ist identisch mit den Diagonalgliedern (ausgenommen die Konstante). Gewarnt sei vor folgender naheliegender Interpretation: Im Diagonalglied stehen die Stimmen, die von der Partei X91 an die Partei X97 abgeströmt sind und in der Konstanten die Stimmen, die zusammengenommen von den anderen Parteien zur Partei X97 abgeströmt sind. Das kann schon deswegen nicht stimmen weil die Konstanten-Stimmen negativ sein können (siehe ÖVP97). Abströme können nicht negativ sein. Die Überlegung einen negativen Abstrom als "Rückstrom" in der entgegengesetzten Richtung zu interpretieren würde bedeuten, dass von 1997 nach 1991 ein Zustrom erfolgt.

Das gesamte Prognose-Ergebnis - verglichen mit dem tatsächlichen Wahlergebnis ist folgendes

	Prognose für aktuelle Wahl		tatsächliches Wahlergebnis	
	Stimmen in aktueller Wahl	nur Parteien in Prozent	Stimmen in aktueller Wahl	nur Parteien in Prozent
OEVP97	328235	42.743	328095	42.688
SPOE97	207648	27.040	207781	27.034
FPOE97	158229	20.605	158555	20.630
Gruene97	44317	5.771	44435	5.781
Sonst97	29502	3.842	29714	3.866
Ungült97	20500	-	20622	-
Nichtw97	165292	-	165860	-
	953723	100.000	955062	100.000

Die prognostizierten Werte stimmen mit den tatsächlichen gut überein. Da wir jedoch mit vollständig ausgezählten Gemeinde-Ergebnissen gerechnet haben, wäre eine exakte Übereinstimmung zu erwarten gewesen. Das bedeutet, dass das Verfahren der "einfachen Regression mit Konstante" keine *konsistente* Schätzung liefern kann.

Dem Almo-Benutzer sei empfohlen mit unseren Beispieldaten eine Hochrechnung für den Zeitpunkt 14.00 Uhr zu rechnen. Zu diesem Zeitpunkt waren erst 35 Gemeinden ausgezählt. In der Programm-Maske muss der Benutzer dazu die Eingabebox "Simulation einer Hochrechnung" öffnen und eintragen:

wenn Zeit97 kleiner 14.00

Offenkundig ist, dass für eine neue Partei in der aktuellen Wahl kein Pendant aus der vergangenen Wahl vorhanden sein kann. Die neue Partei wird dadurch nur über die Konstante geschätzt. Das Verfahren der einfachen Regression mit Konstanter ist also für den Fall, dass bei der aktuellen Wahl neue Parteien antreten, eher nicht empfehlenswert. Auch werden Parteien, die bei der vergangenen Wahl angetreten sind, aber nicht mehr bei der aktuellen, nicht berücksichtigt. Es wird nicht untersucht, wohin deren Wähler abgeströmt sind. Die Methode der einfachen linearen Regression mit Konstante wird heute nicht mehr benutzt. Sie ist eher von historischen Interesse.

P51.1.2 Multiple lineare Regression

Nun ist es natürlich naheliegend, die einfache Regression zur multiplen Regression zu erweitern.

Dieses Verfahren liefert Prognoseergebnisse, die nicht viel schlechter sind als die der quadratischen Optimierung, die wir noch darstellen werden. Es ist sicher das zweit-beste Hochrechnungsverfahren. Es besitzt jedoch eine bedeutsame Einschränkung: Eine Wählerstromanalyse (in Zahl der Stimmen) ist nicht möglich.

In Almo ist das bei Bruckmann (1966) in Kapitel 2 beschriebene Verfahren als Programm-Maske Prog53m2 enthalten. Es wurde von Kurt Holm programmiert. Man findet das Programm durch Klick auf den Knopf "Verfahren/Wahlhochrechnung" am Oberrand des Almo-Fensters oder durch Klick auf den Knopf "alle Progs".

Jede Partei der aktuellen Wahl wird als abhängige Variable in eine Regressionsanalyse (ohne Konstante) eingeben. Unabhängige Variable sind dann alle Parteien der vergangenen Wahl. Als "Parteien" können selbstverständlich auch Nichtwähler usw. betrachtet werden.

Wir rechnen mit Beispiel-Programm **P53m2_St.Alm** eine Hochrechnung für die Landtagswahl 1997 auf Basis der Wahl von 1991. Wir verwenden die bereits oben beschriebenen Daten und rechnen bei *vollständiger Auszählung* aller Gemeinden. Die Prognose-Ergebnisse des Regressionsmodells können so mit den tatsächlichen Wahlergebnissen verglichen werden. Der Almo-Benutzer kann selbstverständlich auch eine Hochrechnung bei unvollständiger Auszählung rechnen, z.B. für den Zeitpunkt 14.00 Uhr. Zu diesem Zeitpunkt waren erst 35 Gemeinden ausgezählt. In der Programm-Maske muss er dazu die Eingabebox "Simulation einer Hochrechnung" öffnen und eintragen:

wenn Zeit97 kleiner 14.00

Der Almo-Benutzer findet das Programm P53m2_St.Alm durch Klick auf den Knopf "alle Progs" am Oberrand des Almo-Fensters. Das Programm entstand aus der Programm-Maske Prog53m2.Msk dadurch, dass verschiedene Kleinstparteien aus der aktuellen Wahl zu einer Partei "Sonst97" zusammengefasst wurden.

Das Programm liefert folgende Matrix der Regressionskoeffizienten

Matrix der Regressionskoeffizienten (bei vollständiger Auszählung)

OEVP97	SPOE97	FPOE97	GRUENE97	Sonst97	Ungueilt97	Nichtw97	GESAMT
--------	--------	--------	----------	---------	------------	----------	--------

OEVP91	0.93	0.03	0.05	0.02	0.02	0.02	0.03	1.10
SPOE91	-0.02	0.89	0.10	0.03	0.02	0.03	0.03	1.08
FPOE91	0.08	0.01	0.93	0.01	0.00	0.03	0.03	1.09
VGÖE91	0.49	0.32	-0.02	0.42	0.19	0.00	0.15	1.55
GAL91	0.02	-0.23	-0.44	1.06	0.39	-0.01	0.32	1.11
Ungue191	0.55	0.25	0.00	-0.18	-0.02	0.20	-0.23	0.56
Nichtw91	-0.21	-0.21	-0.03	-0.02	0.02	-0.01	1.09	0.65

Auf der Stirnseite der Matrix stehen die Parteien der vergangenen Wahl 1991 und die Pseudo-Parteien der Ungültigwähler und Nichtwähler. Im Kopf der Matrix stehen die Parteien und Pseudo-Parteien der aktuellen Wahl 1997. Unter Sonst97 wurden mehrere Kleinstparteien zusammengefasst.

Folgendes fällt auf:

- Die Regressionskoeffizienten können negativ sein, aber auch größer als 1,0. D.h. sie dürfen nicht als Übergangswahrscheinlichkeiten interpretiert werden. Das bedeutet aber auch, dass es nicht möglich ist Wählerströme in Stimmenzahlen zu ermitteln. Betrachten wir beispielsweise, das rechte obere Eck in der Übergangsmatrix (den Wert 0.93). Wir dürfen nicht interpretieren, die Wahrscheinlichkeit, dass Wähler der Partei X zur Partei Y überlaufen, sei 93%.
- Die Übergangswahrscheinlichkeiten einer Partei X der vergangenen Wahl zu den Parteien der aktuellen Wahl müssen zusammen 1.0 (bzw. 100%) ergeben. Die Spaltensumme in der Matrix ist jedoch nicht gleich 1.0. Siehe Spalte "Gesamt". Dieser Defekt kann jedoch durch zwei verschiedene Methoden behoben werden:

(1) Werden die Zu- und Abwanderer als Pseudoparteien miteinbezogen, dann wird die Spaltensumme gleich 1.0. Die Zu- bzw. Abwanderer werden je Gemeinde ermittelt. Die Pseudo-Partei der Zuwanderer und Abwanderer entsteht aus der Differenz der Wahlberechtigten in der aktueller Wahl minus der Wahlberechtigten in der vergangener Wahl. Siehe die genaue Formel in Abschnitt P51.4.3. Dort wird auch das Programm Prog51mb beschrieben, mit dem die Summe aller Zuwanderer ermittelt wird, die in die Programm-Maske Prog53m2 für die multiple Regression eingegeben werden muss.

(2) Auch gibt es die Möglichkeit die Regressionskoeffizienten so zu transformieren dass Spaltensummen von 1.0 entstehen. Bruckmann hat dafür eine Formel vorgeschlagen (1966, S. 21, 135, Gleichg 2.4.3) Im Almo-Programm Prog53m2 ist diese Transformation möglich. Wir werden (gegen Ende dieses Abschnitts) noch sehen, dass durch diese Transformation jedoch nicht viel zu gewinnen ist.

Die so transformierte Regressionskoeffizienten können aber immer noch außerhalb des Bereichs 0-1 liegen, so dass sie nicht als Übergangswahrscheinlichkeiten betrachtet werden dürfen. Dieser Defekt kann nur durch das im nächsten Abschnitt behandelte Modell der quadratischen Optimierung behoben werden.

Wir zeigen noch die Regressionskoeffizienten, die das Programm für Fall (1) und (2) ausgibt.

Regressionskoeffizienten mit Zu- und Abwanderer

	OEVP97	SPOE97	FPOE97	GRUENE97	Sonst97	Ungue1t97	Nichtw97	Abwand97	GESAMT
OEVP91	0.90	0.00	0.04	0.01	0.02	0.02	0.01	0.00	1.00
SPOE91	-0.04	0.87	0.09	0.02	0.02	0.03	0.02	0.00	1.00
FPOE91	0.06	-0.01	0.92	0.01	0.00	0.03	0.01	-0.02	1.00
VGÖE91	0.33	0.19	-0.10	0.37	0.16	-0.01	0.03	0.04	1.00
GAL91	0.00	-0.25	-0.45	1.05	0.39	-0.01	0.30	-0.04	1.00

Unguel91	0.64	0.33	0.04	-0.15	0.00	0.20	-0.16	0.09	1.00
Nichtw91	-0.14	-0.14	0.01	0.01	0.04	0.00	1.15	0.08	1.00
Zuwand91	0.29	0.25	0.15	0.10	0.06	0.03	0.22	-0.09	1.00

Regressionskoeffizienten nach Transformation auf Summe 1.0

	OEVP97	SPOE97	FPOE97	GRUENE97	Sonst97	Unguel97	Nichtw97	GESAMT
OEVP91	0.91	0.01	0.04	0.00	0.01	0.01	0.02	1.00
SPOE91	-0.03	0.88	0.08	0.02	0.01	0.02	0.02	1.00
FPOE91	0.06	-0.01	0.92	0.00	-0.01	0.02	0.02	1.00
VGOE91	0.41	0.24	-0.10	0.34	0.11	-0.08	0.07	1.00
GAL91	0.00	-0.24	-0.46	1.04	0.38	-0.02	0.30	1.00
Unguel91	0.61	0.31	0.06	-0.12	0.05	0.26	-0.16	1.00
Nichtw91	-0.16	-0.15	0.02	0.04	0.07	0.04	1.14	1.00

In beiden Matrizen sind die Summen der Regressionskoeffizienten gleich 1.0. Die Matrizen sind zwar sehr ähnlich, jedoch nicht identisch gleich.

Die Stimmensumme der Parteien bei der vergangenen Wahl war folgende:

	Stimmen bei vergangener Wahl
OEVP91	357771
SPOE91	248640
FPOE91	140311
VGOE91	20603
GAL91	24250
Unguel91	16315
Nichtw91	140355
Zuwanderer	31478

	979723

Wird obige Matrix der Regressionskoeffizienten mit Zu- und Abwanderer multipliziert mit den Stimmsummen der Parteien der vergangenen Wahl, dann entsteht die Matrix der "Quasi-Wählerströme" (Abströme) und das (prognostizierte) Wahlergebnis der aktuellen Wahl in Stimmen und Prozent.

Multiplikation der Regressionskoeffizienten mit Gesamt-Stimmenszahl aus vergangener Wahl ergibt "Quasi-Wählerströme" - Abströme in Zahl der Stimmen prognostiziert auf der Basis von 445 ausgezählten Gemeinden

letzte Spalte "Gesamt" enthaelt die tatsaechliche Gesamt-Stimmenszahl je Partei in der vergangenen Wahl (nur wenn Summe der Reg.koeff. in Zeile =1.0)

letzte Zeile "Gesamt" enthaelt die vom Modell prognostizierte Gesamt-Stimmenszahl je Partei in der aktuellen Wahl

Beachte: Es können nicht interpretierbare negative Wählerströme auftreten

	OEVP97	SPOE97	FPOE97	GRUENE97	Sonst97	Ungü97	Nichtw97	Abwander	GESAMT
OEVP91	322860	1254	13911	3439	5667	6549	4789	-697	357771
SPOE91	-9443	216972	21288	5345	4085	6841	4678	-1126	248640
FPOE91	7979	-1282	129463	833	-253	3920	2096	-2446	140311
VGOE91	6834	3870	-2149	7530	3245	-245	642	875	20603
GAL91	-67	-5989	-10969	25458	9470	-187	7387	-853	24250
Ungü91	10384	5305	723	-2416	42	3324	-2589	1543	16315
Nich91	-19466	-20164	1489	1215	5490	-402	160738	11455	140355
Zuwand	9014	7815	4799	3032	1969	821	6807	-2778	31478

GESAMT	328095	207781	158555	44435	29714	20622	184548	5973	-
Prozent	33.5%	21.2%	16.2%	4.5%	3.9%	2,1%	18.8%	0.6%	100.0%
nur Parteien									
Prozent	42.7%	27.0%	20.6%	5.8%	3.9%	-	-	-	100.0%

Wir betrachten die Spalte **FPOE97**. Der erste Wert 13911 oben in der Spalte bedeutet: Von der OEVP91 sind 13911 Wähler abgeströmt zur FPOE97. Der Wert 13911 entstand aus $0.03888 \cdot 357771$ (in der obigen Matrix der Regressionskoeffizienten wurden die Werte gerundet nur mit 2 Komastellen ausgegeben).

Die gesamte Spalte FPOE97 entstand aus folgender Multiplikation (wobei die Regressionskoeffizienten hier mit 5 Kommastellen geschrieben wurden).

Parteien von denen Abstrom erfolgt	Stimmen vergangene Wahl	* Regress. koeffiz. FPOE97	= Abströme zu FPOE97
OEVP91	357771	0.03888	13911
SPOE91	248640	0.08562	21288
FPOE91	140311	0.92269	129463
VGÖE91	20603	-0.10429	-2149
GAL91	24250	-0.45234	-10969
Ungült91	16315	0.04430	723
Nichtw91	140355	0.01061	1489
Zuwanderer	31478	0.15245	4799
			----- 158555

Was sofort auffällt sind die negativen Abströme von den Parteien der vergangenen Wahl. So sind von der GAL91 (einem Vorläufer der Grünen) -10969 Wähler zur FPOE97 abgeströmt. Abströme können nicht negativ sein. Die Überlegung einen negativen Abstrom als "Rückstrom" in der entgegengesetzten Richtung zu interpretieren würde bedeuten, dass von 1997 nach 1991 ein Zustrom erfolgt.

Das ist eine sehr bedeutsame Einschränkung des multiplen Regressions-Modells: Eine Wählerstromanalyse ist nicht möglich. Sie wird erst dadurch möglich, dass das Regressionsmodell durch die quadratische Optimierung erweitert wird.

Wir vergleichen nun die von P53m2_St prognostizierten Ergebnisse mit den tatsächlichen

	Prognose für aktuelle Wahl		tatsächliches Wahlergebnis	
	Stimmen in aktueller Wahl	nur Parteien in Prozent	Stimmen in aktueller Wahl	nur Parteien in Prozent
OEVP97	328095	42.688	328095	42.688
SPOE97	207781	27.034	207781	27.034
FPOE97	158555	20.630	158555	20.630
GRUENE97	44435	5.781	44435	5.781
Sonst97	29714	3.866	29714	3.866
Ungült97	20622	-	20622	-
Nichtw97	184548	-	184548	-
Abwander	5973	-	5973	-
	-----	-----	-----	-----
	979723	100.000	979723	100.000

Die vom Modell für die aktuelle Wahl reproduzierte Stimmenzahl und Prozentzahl je Partei und das tatsächliche Ergebnis stimmen überein. Was bei unserem Beispiel (bei dem alle Gemeinden ausgezählt sind) sichtbar wird, ist also, dass das Modell fähig ist, *konsistent* zu schätzen. Bei vollständiger Auszählung werden die

Stimmen und die prozentuellen Anteile der Parteien richtig prognostiziert. Das gilt jedoch nur, wenn Zu- und Abwanderer miteinbezogen werden.

Die Schätzungen, die entstehen wenn *ohne* Zu- und Abwanderer gerechnet wird, ist zwar sehr gut aber hinsichtlich der Stimmenzahl und der Prozentwerte nicht konsistent. Es entsteht auch keine konsistente Schätzung, wenn ohne Zu- und Abwanderer gerechnet wird, jedoch die Regressionssumme auf 1.0 transformiert wird. Die Transformierung bewirkt, dass bei vollständiger Auszählung, die prognostizierte Gesamt-Stimmenzahl aus allen Parteien und Pseudo-Parteien gleich der Gesamt-Stimmenzahl der vergangenen Wahl ist. Das prognostizierte Wahlergebnis wird sozusagen rückbezogen auf das vergangene.

P51.1.3 Quadratische Optimierung

Das Verfahren der quadratischen Optimierung ergänzt die Methode der multiplen Regression. Es liefert echte Übergangswahrscheinlichkeiten. Prog51m1 verwendet dieses Verfahren. Es wird im vorliegenden Dokument ausführlich beschrieben.

Betrachten wir wieder ein Beispiel. Wir rechnen, wie oben im Abschnitt über die multiple Regression, eine Hochrechnung für die Landtagswahl 1997 auf Basis der Wahl von 1991. Wir verwenden die bereits oben beschriebenen Daten und rechnen mit Programm **P51m1_St.Alm** eine Hochrechnung bei vollständiger Auszählung aller Gemeinden. Die Prognose-Ergebnisse des Regressionsmodells können so mit den tatsächlichen Wahlergebnissen verglichen werden. Der Almo-Benutzer kann selbstverständlich auch eine Hochrechnung bei unvollständiger Auszählung rechnen, z.B. für den Zeitpunkt 14.00 Uhr. Zu diesem Zeitpunkt waren erst 35 Gemeinden ausgezählt. In der Programm-Maske muss er dazu die Eingabebox "Simulation einer Hochrechnung" öffnen und eintragen:

wenn Zeit97 kleiner 14.00

Man findet das Programm P51m1_St.Alm durch Klick auf den Knopf "alle Progs" am Oberrand des Almo-Fensters. Das Programm entstand aus der Programm-Maske Prog51m1.Msk dadurch, dass verschiedene Kleinstparteien aus der aktuellen Wahl zu einer Partei "Sonst97" zusammengefasst wurden.

Das Programm liefert folgende Matrix der Übergangswahrscheinlichkeiten

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungült91	Nichtw97
OEVP97	87.85	0	0	0	0	31.91	0
SPOE97	1.41	78.18	0	0	0	17.70	0
FPOE97	4.39	7.31	85.74	0	0	1.36	0
GRUENE	1.54	3.15	2.78	42.51	68.76	3.75	0
Sonst97	1.76	3.21	2.96	25.73	10.50	16.22	0
Ungül97	2.18	2.24	1.40	0	0	29.07	0
Nichtw97	0.87	5.91	7.11	31.76	20.74	0	100.00
Gesamt	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Die Übergangswahrscheinlichkeiten werden in der Matrix als Prozentwerte angegeben.

Wenn man die Matrix der Regressionskoeffizienten aus der multiplen Regression mit der hier abgebildeten Übergangsmatrix vergleicht, dann muss man beachten,

dass bei letzterer die Zeilen und Spalten umgekehrt (transponiert) sind. Die aktuellen Parteien stehen in der Stirnseite der Matrix und die vergangenen Parteien im Kopf der Matrix.

Bei der vergangenen Wahl 1991 gab es zwei grüne Parteien, die VGÖE91 (eher bürgerlich) und die GAL91 (eher links). Für die aktuelle Wahl 1997 hatten sich die beiden zu der gemeinsamen Partei "Die Grünen" zusammengeschlossen. Interessant ist nun, dass von den Wählern der GAL91 69% zur neuen grünen Partei "geströmt" sind, von den Wählern der eher bürgerlichen VGÖE91 nur 43%.

Die Stimmensumme der Parteien bei der vergangenen Wahl waren folgende:

	Stimmen in vergangener Wahl
OEVP91	357771
SPOE91	248640
FPOE91	140311
VGÖE91	20603
GAL91	24250
Ungült91	16315
Nichtw91	140355

	948245

Wird die Übergangsmatrix multipliziert mit den Stimmsummen der Parteien der vergangenen Wahl, dann entsteht die Matrix der Wählerströme (Abströme) und das (prognostizierte) Wahlergebnis der aktuellen Wahl in Stimmen und Prozent.

Wählerstromanalyse (Abströme) in Stimmen und Prozent

	OEVP91	SPOE91	FPOE91	VGÖE91	GAL91	Ungült97	Nichtw97	Gesamt	Parteien in %
OEVP97	314296	0	0	0	0	5206	0	319501	42.688
SPOE97	5052	194399	0	0	0	2888	0	202339	27.034
FPOE97	15698	18176	120306	0	0	222	0	154402	20.630
GRUENE	5501	7824	3904	8758	16674	611	0	43271	5.781
Sonst97	6312	7981	4148	5301	2546	2647	0	28936	3.866
Ungült97	7797	5573	1970	0	0	4742	0	20082	-
Nichtw97	3115	14687	9983	6544	5030	0	140355	179714	-

Gesamt	357771	248640	140311	20603	24250	16315	140355	948245	100.000

Wie wurde multipliziert? Betrachten wir die Zeile FPOE97.

Parteien von denen Abstrom erfolgt	Stimmen vergangene Wahl	* Übergang whrschkt FPOE97	= Abströme zu FPOE97
OEVP91	357771	0.04388	15698
SPOE91	248640	0.07310	18176
FPOE91	140311	0.85742	120306
VGÖE91	20603	0	0
GAL91	24250	0	0
Ungült91	16315	0.01358	222
Nichtw91	140355	0	0

			154402

Die Übergangswahrscheinlichkeiten wurden in der Tabelle als Anteilswerte mit 5 Kommastellen angeschrieben.

Die Multiplikation kann auch matrixalgebraisch so ausgedrückt werden:

$$W = U * S$$

m = Zahl der Parteien und Pseudo-Parteien der aktuellen Wahl
(im Beispiel: 7)
n = Zahl der Parteien und Pseudo-Parteien der vergangenen Wahl
(im Beispiel: 7)
U = das ist die m*n Matrix der Übergangswahrscheinlichkeiten
(die Übergangswahrscheinlichkeiten in Form von Anteilswerten)
S = das ist eine n*n Diagonalmatrix, in deren Diagonale die
Stimmzahlen der Parteien der vergangenen Wahl stehen

Die letzte Zeile "Gesamt" in der Tabelle der Wählerströme enthält die tatsächliche Stimmzahl je Partei in der vergangenen Wahl. Die letzte Spalte "Gesamt" enthält die vom Modell prognostizierte Stimmzahl je Partei in der aktuellen Wahl.

Wird die Zeile "Gesamt" summiert, dann entsteht die Zahl der Wahlberechtigten in der vergangenen Wahl 1991. Dies sind: 948245. Wird die Spalte "Gesamt" summiert, dann entsteht dieselbe Zahl. Das bedeutet, die für 1997 prognostizierten Stimmen je Partei summieren sich zur Zahl der Wahlberechtigten der vergangenen Wahl und nicht zu der der aktuellen. Wir werden auf diesen Defekt noch eingehen.

Zuerst soll das tatsächliche Ergebnis mit dem vom Modell reproduzierten verglichen werden

	vom Modell reproduziert		tatsächliches Ergebnis	
	Stimmen	nur Parteien in Prozent	Stimmen	nur Parteien in Prozent
OEVP97	319501	42.688	328095	42.688
SPOE97	202339	27.034	207781	27.034
FPOE97	154402	20.630	158555	20.630
GRUENE97	43271	5.781	44435	5.781
Sonstig97	28936	3.866	29714	3.866
Ungueilt97	20082	-	20622	-
Nichtw97	179714	-	184548	-
	-----	-----	-----	-----
	948245	100.000	973750	100.000

Wir stellen fest:

1. Das vom Modell prognostizierte Prozent-Ergebnis ist exakt richtig. Das Modell der quadratischen Optimierung liefert eine konsistente Schätzung hinsichtlich der Prozentverteilung.
2. Die aufsummierte Zahl aller Stimmen der aktuellen Wahl (=948245) wird vom Modell auf die der Wahlberechtigten der vergangenen Wahl (=948245) gesetzt.
3. Daraus folgt, dass die für die aktuelle Wahl je Partei prognostizierten Stimmen relativ zueinander richtig sind, aber nicht absolut stimmen.
4. Die obige Matrix der Wählerströme zeigt die Abströme von den Parteien der vergangenen Wahl zu denen der aktuellen Wahl - *unter der Voraussetzung, dass die Zahl der Wahlberechtigten der aktuellen Wahl gleich derjenigen der vergangenen Wahl ist.*

Wir haben einen Einfluss ausgeblendet, der für die unter 2 bis 4 beschriebenen Defekte verantwortlich ist: Die Zahl der Wahlberechtigten ist in der Regel zwischen vergangener und aktueller Wahl ungleich. Dieses Problem wird dadurch gelöst, dass man zwei Pseudo-Parteien einführt: Die Zu- und Abwanderer je Wahlgemeinde - wobei diese sehr einfach die Differenz zwischen den

Wahlberechtigten je Gemeinde zu den beiden Wahlzeitpunkten bilden. Ist die Differenz

Wahlberechtigte in aktueller Wahl minus Wahlberechtigte in vergangener Wahl

positiv, dann ist die betreffende Gemeinde eine "Zuwanderer"-Gemeinde. Die Abwandererzahl ist dann 0. Ist die Differenz negativ, dann ist die Gemeinde eine "Abwanderer"-Gemeinde. Die Abwandererzahl wird absolut (also positiv) gesetzt. Die Zuwandererzahl ist dann 0. Siehe die ausführliche Darstellung in Abschnitt P51.4.3 und P51.4.6.

Die Zuwanderer werden als eine Pseudo-Partei aus der vergangenen Wahl betrachtet und die Abwanderer als eine Pseudo-Partei der aktuellen Wahl. Für unser Programm P51m1_St wird dann noch die Summe aller Zuwanderer aus allen Gemeinden gebraucht. Sie ist gewissermaßen die Gesamtzahl der "Stimmen", die die Pseudo-Partei der Zuwanderer in der vergangenen Wahl erhalten hat. Sie ist 31478. Sie kann durch die Programm-Maske Prog51mb ermittelt werden.

Wir rechnen nun eine Analyse mit Zu- und Abwanderern. Der Benutzer muss im Programm P51m1_St in der Eingabebox für Zu- und Abwanderer nur diese beide Begriffe einsetzen.

Es entsteht folgende Matrix der Übergangswahrscheinlichkeiten:

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungült91	Nichtw91	Zuwander
OEVP97	88.54	0	0	0	0	0	0	35.99
SPOE97	0	78.38	0	0	0	15.52	0	32.96
FPOE97	4.40	6.18	87.01	0	0	6.25	0	13.75
GRUENE	1.28	3.01	2.25	39.34	71.18	12.32	0	5.90
Sonst97	1.91	3.22	2.45	23.25	10.23	24.10	0	0.75
Ungül97	2.25	2.02	0.87	0	0	37.39	0	0.70
Nichtw97	1.62	5.94	7.42	27.05	18.59	0	100.00	9.96
Abwander	0	1.25	0	10.36	0	4.42	0	0
Gesamt	100.00	100.00	100.00	100.00	00.00	100.00	100.00	100.00

Wählerströme - Abströme in Zahl der Stimmen

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungül97	Nichtw97	Zuwander	Gesamt
OEVP97	316767	0	0	0	0	0	0	11328	328095
SPOE97	0	194874	0	0	0	2532	0	10374	207781
FPOE97	15755	15365	122086	0	0	1020	0	4330	158555
GRUENE	4575	7472	3156	8105	17261	2010	0	1856	44435
Sonst97	6823	8010	3442	4790	2481	3932	0	236	29714
Ungül97	8050	5034	1218	0	0	6100	0	219	20622
Nichtw97	5801	14767	10409	5573	4507	0	140355	3135	184548
Abwander	0	3118	0	2135	0	721	0	0	5973
Gesamt	357771	248640	140311	20603	24250	16315	140355	31478	-

Die letzte Zeile "Gesamt" enthält wieder die tatsächlichen Gesamt-Stimmenzahlen je Partei in der vergangenen Wahl. Die Zeilensumme beträgt jetzt 979723. Das ist die Wahlberechtigten-Zahl in der aktuellen Wahl 1997. Wird der letzte Wert in der

"Gesamt"-Zeile, die 31478 Zuwanderer herausgenommen, dann entsteht 948245. Das ist die Wahlberechtigtenzahl der vergangener Wahl 1991.

Die letzte Spalte "Gesamt" enthält die vom Modell prognostizierte Gesamt-Stimmenzahl je Partei in der aktuellen Wahl 1997. Wird die "Gesamt"-Spalte summiert, so entsteht mit 979723 die Wahlberechtigtenzahl der aktuellen Wahl 1997.

Wir vergleichen jetzt noch die prognostizierten mit den tatsächlichen Ergebnissen

	vom Modell reproduziert		tatsächliches Ergebnis	
	Stimmen	nur Parteien in Prozent	Stimmen	nur Parteien in Prozent
OEVP97	328095	42.688	328095	42.688
SPOE97	207781	27.034	207781	27.034
FPOE97	158555	20.630	158555	20.630
GRUENE97	44435	5.781	44435	5.781
Sonstige97	29714	3.866	29714	3.866
Ungült97	20622	-	20622	-
Nichtw97	184548	-	184548	-
Abwanderer	5973	-	5973	-
	-----	-----	-----	-----
	979723	100.000	979723	100.000

Die Stimmen und die Prozente stimmen exakt überein. Bei vollständiger Auszählung kann das Modell die Ergebnisse exakt reproduzieren. Das Modell liefert eine *konsistente* Schätzung.

Die quadratische Optimierung besitzt folgende Eigenschaften und Besonderheiten:

- a. Alle Übergangswahrscheinlichkeiten liegen zwischen 0 % und 100 %.
- b. Die Übergangswahrscheinlichkeit in einer Spalte addieren sich zu 100 %.
- c. Was in unserem Beispiel sichtbar wurde, ist folgende Besonderheit: Das Verfahren ist *konsistent*. Sind alle Wahlgemeinden der aktuellen Wahl ausgezählt, dann "prognostiziert" das Verfahren der quadratischen Optimierung die tatsächliche Prozentverteilung.
- d. Werden auch noch die "Zu- und Abwanderer", wie oben ausgeführt, in die Analyse mit einbezogen, dann stimmen bei vollständiger Auszählung nicht nur die prognostizierten Prozentwerte sondern auch die prognostizierten absoluten Stimmen mit den tatsächlichen überein.

Die quadratische Optimierung ist mathematisch etwas aufwendiger als die multiple Regression. Sie kann begriffen werden als ein Kleinste-Quadrate-Kalkül mit folgenden 3 *Nebenbedingungen*:

$$(1) 0 \leq p_{ij} \leq 1$$

$$(2) \sum_i p_{ij} = 1$$

$$(3) \sum_j p_{ij} \cdot n_j = r_i \cdot n$$

Die obigen Sätze a und b und die Formeln (1) und (2) bilden das "Grundmodell" der Hochrechnung und Wählerstromanalyse durch quadratische Optimierung. Es liefert noch keine *konsistente* Schätzung. Bei vollständiger Auszählung aller Gemeinden wird das Endergebnis nicht exakt reproduziert - wenn auch mit vollkommen ausreichender Genauigkeit.

Die wichtige Erweiterung aus Satz c wird durch Formel 3 realisiert. Erst durch die 3. Nebenbedingung liefert das Modell eine *konsistente* Schätzung. Bei vollständiger Auszählung wird das Endergebnis exakt reproduziert.

Siehe auch die ausführliche Modell-Darstellung in Abschnitt P51.5.

i = kennzeichnet eine Partei der aktuellen Wahl

j = der vergangenen Wahl

Als "Partei" können selbstverständlich auch die Nichtwähler, Ungültigwähler, Zu- und Abwanderer usw. in die Analyse miteinbezogen werden.

p_{ij} = Übergangswahrscheinlichkeit der Partei i (bei der vergangenen Wahl) zur Partei j (bei der aktuellen Wahl)

n_j = Stimmenzahl der Partei j bei der vergangenen Wahl

r_i = Stimmenanteil der Partei i bei der aktuellen Wahl

n = Gesamtzahl der Stimmen aus den schon ausgezählten Wahlgemeinden bei der vergangenen Wahl

Die Nebenbedingung 1 erzeugt Satz a.

2 erzeugt b

3 erzeugt c

P51.1.4 Weitere Modelle

Neben den hier beschriebenen Regressions-Modellen und dem Modell der quadratische Optimierung wird in der Literatur zur Wahlhochrechnung immer wieder hingewiesen auf

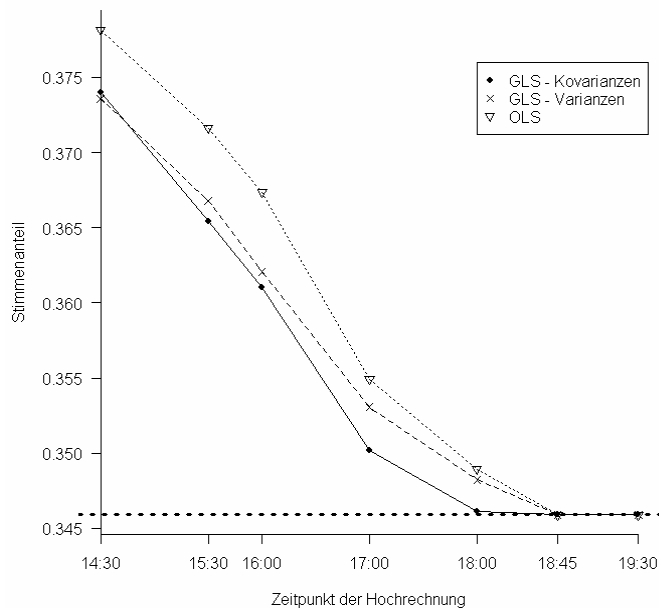
1. das Multinomial-Dirichlet Modell von Rosen, Jiang, King und Tanner (2001)
2. und das Multinomial-Log-Normal Modell von Greiner und Quinn(2009)
3. Verwandt damit ist das Multinomial-Modell, das Ledl (2007) beschreibt.

Einen guten Überblick über die beiden erstgenannten Modelle findet man in den Arbeiten von Klima(2016) und Kopecki(2016), die beide im Internet herunter geladen werden können.

Die Frage, welches Modell die Wahlergebnisse am besten prognostiziert, ist nicht

zu beantworten. Ledl verwendet verschiedene Wahldaten und vergleicht sein Multinomial-Modell mit dem "restringierten OLS-Modell". Das restringierte OLS-Modell ist weitgehend identisch mit dem hier vorgetragenen Modell der quadratischen Optimierung. Bei den vielen Vergleichen sieht er einen leichten aber nicht durchgängigen Vorteil für sein Modell. Hier sind 2 Beispiele aus den vielen Vergleichen, die er anstellt.

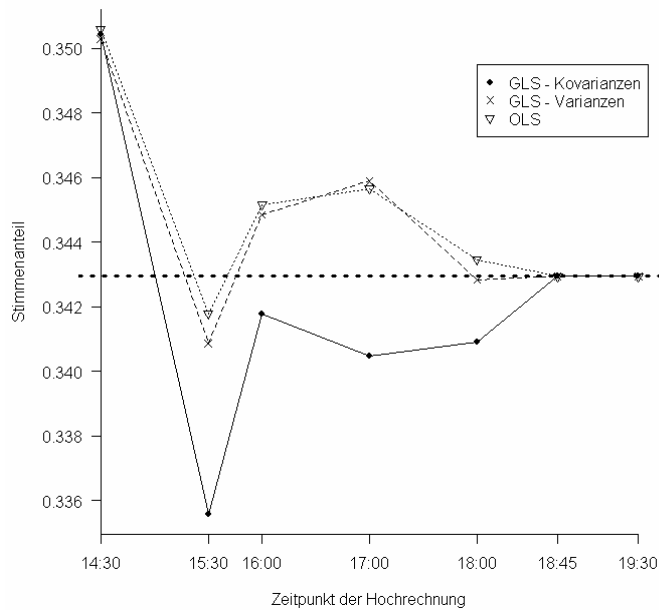
Wahlhochrechnung für die ÖVP in Oberösterreich EU-Wahl 2004 zu Ö.-Wahl 2002



entnommen aus Ledl (2007)

Mit "GLS-Kovarianzen" wird das Multinomial-Modell bezeichnet. "GLS-Varianzen" ist eine Variante dieses Modells. "OLS" entspricht dem Modell der quadratischen Optimierung. Beispiel: Bei der Hochrechnung um 17.00 Uhr prognostiziert das Multinomial-Modell für die ÖVP 35,0% und das OLS-Modell 35,5%. Das tatsächliche Wahlergebnis war 34,5%

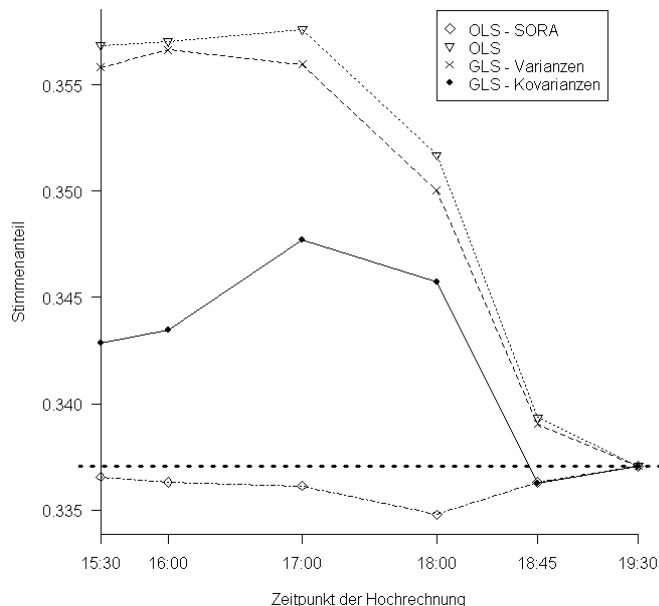
Wahlhochrechnung für die SPÖ in Oberösterreich EU-Wahl 2004 zu Ö.-Wahl 2002



entnommen aus Ledl (2007)

Die Prognose unterscheiden sich meistens nur im Bereich von Zehntel Prozenten. Ledl bezieht dann noch das Model des SORA-Instituts (das für das ORF-Fernsehen hochrechnet) in die Vergleiche mit ein. Dieses erweist sich in einem Fall als deutlich überlegen, wie in folgender Grafik ersichtlich wird

Wahlhochrechnung für die SPÖ bei Österreich-Wahl 2002 zu 1999



entnommen aus Ledl (2007)

Es ist ein "unrestringiertes OLS-Modell", das wohl weitgehend mit dem in Abschnitt P51.1.2 kurz beschriebenen multiplen Regressions-Modell identisch ist, jedoch eine Besonderheit aufweist: Es ist ein "gruppiertes Hochrechnungsmodell". Die Wahlgemeinden Österreichs werden in 5 Gruppen unterteilt - wobei die Gruppen

so gebildet werden, dass die Gemeinden einer Gruppe in ihren Übergangswahrscheinlichkeiten möglichst ähnlich sind - die 5 Gruppen sich in ihrer Übergangsmatrix jedoch möglichst deutlich unterscheiden. (Das SORA-Modell war jedoch auch dadurch überlegen, dass ihm Wahlergebnisse aus einigen Wiener Sprengeln sehr früh bekannt waren).

In Abschnitt P51.4.8 werden wir zeigen, wie mit unserem Programm Prog51m1 die "gruppierte Hochrechnung/Wählerstromanalyse" nach dem Modell der quadratischen Optimierung gerechnet werden kann.

Es darf wohl vermutet werden, dass durch die Weiterentwicklung der Modelle die Prognoseleistung nur noch minimal verbessert werden kann - dass es hingegen die Gruppierung (bzw. Clusterung) ist, die die Prognose entscheidend verbessern wird. Wichtig ist es dann herauszufinden, welche Merkmale bzw. Merkmalskombinationen für das Gruppieren relevant sind (und als Daten auch verfügbar sind).

P51.2 Wählerstromanalyse und Wahlhochrechnung

Von einer "Wählerstromanalyse" wird üblicherweise erwartet, dass sie zeigt woher die Stimmen, die die Partei X bei der aktuellen Wahl erhalten hat, gekommen sind. Sie soll die Frage beantworten: Von welchen Parteien und Pseudo-Parteien der vergangenen Wahl hat die Partei X ihre Stimmen gewonnen? Pseudo-Parteien sind z.B. Nichtwähler, Ungültigwähler, Erstwähler, Zuwanderer usw.

Unser Programm Prog51m1 berechnet für die jeweils ausgezählten Wahlgemeinden eine Matrix der Übergangswahrscheinlichkeiten von den Parteien der vergangenen Wahl zu denen der aktuellen Wahl. Es *prognostiziert* damit das endgültige Wahlergebnis. Dabei gilt, dass die Prognose umso besser wird je höher der Auszählungsgrad ist. Wenn schließlich alle Wahlgemeinden ausgezählt sind, dann liefert der im Prog51m1 verwendete Kalkül exakt das richtige Wahlergebnis. Wir werden das noch ausführlicher darstellen. Bei vollständiger Auszählung ist dann die Gesamtzahl der Stimmen, die eine Partei X erhalten hat, bekannt. Mit Hilfe der Matrix der Übergangswahrscheinlichkeiten kann dann leicht berechnet werden, von welchen Parteien und Pseudo-Parteien der vergangenen Wahl die Stimmen gewonnen wurden.

Aus der Hochrechnung wird dadurch eine Wählerstromanalyse. Anders formuliert: Die Wählerstromanalyse ist eine Hochrechnung bei vollständiger Auszählung.

Im Prinzip kann zu jedem Zeitpunkt, d.h. bei jedem Auszählungsgrad eine Wählerstromanalyse *prognostiziert* werden. Dies wird jedoch selten gemacht, da während der Hochrechnungs-Phase nur das vermutliche prozentuelle Wahlergebnis interessiert.

P51.3 Einige Probleme

Wir wollen nicht den Eindruck erwecken, es gebe keine Probleme. Natürlich gibt es die. Wir wollen im folgenden einige betrachten.

P51.3.0 Nichtwähler und Ungültigwähler

Was geschieht mit den Nichtwählern? Das ist ein leicht lösbares Problem. Sie

werden wie eine Partei betrachtet. In der Übergangsmatrix steht somit in einer der Spalten die "Partei" der Nichtwähler aus der vergangenen Wahl und in einer der Zeilen die "Partei" der Nichtwähler der aktuellen Wahl. Dies ermöglicht auch Wählerströme von und zu der "Partei" der Nichtwähler zu erfassen. Nicht selten sind derartige Ergebnisse besonders interessant. So kann sich etwa ergeben, dass ein hoher Prozentsatz der Nichtwähler aus der vergangenen Wahl zur neuen Partei C geströmt ist. Man wird dann interpretieren, dass C eine typische Protestwähler-Partei sei.

Die Zahl der Ungültigwähler wird in aller Regel sehr klein sein, so dass man sie aus der Analyse ausschließen kann. Es spricht jedoch nichts dagegen, sie auch, wie die Nichtwähler als eigene „Partei“ einzubeziehen – insbesondere, wenn man vermutet, dass unter ihnen ein größerer Teil aus Protest eine ungültige Stimme abgibt (die sogenannten "Weißwähler").

P51.3.1 Zeitpunkt der Hochrechnung

Wir wollen annehmen, dass bei einem Auszählungsgrad von 20 Prozent eine Hochrechnung durchgeführt wird. Wenn wir weiterhin unterstellen, dass die dabei ausgezählten Wahlgemeinden eine *Zufallsstichprobe* aus der Gesamtheit aller Wahlgemeinden bilden, dann können wir das Endergebnis mit einem bestimmten berechenbaren Vertrauensintervall prognostizieren.

Die Realität ist aber eine andere. Üblicherweise liegen die Auszählungsergebnisse von Städten, insbesondere von Großstädten sehr viel später vor als die aus ländlichen Gemeinden. Und die Erfahrung lehrt, dass die Übergangswahrscheinlichkeits-Matrizen von Städten und ländlichen Gemeinden (mehr oder weniger) differieren. Besonders dramatisch kann sich das bei der österreichischen Parlamentswahl auswirken. In Wien lebt beinahe 1/4 der österreichischen Bevölkerung. Und Wien ist "anders" (obwohl Wien glaubt, es wäre Österreich). Und die Wiener Wahlergebnisse liegen immer sehr viel später vor. Solange Wien oder zumindest Teile von Wien nicht ausgezählt sind, ist eine Hochrechnungsprognose noch unsicher.

Allgemein gilt also: Es muss anerkannt werden, dass eine Hochrechnung-Prognose zu einem frühen Zeitpunkt auf einer verzerrten Stichprobe von Wahlgemeinden aus der Gesamtheit aller Wahlgemeinden beruhen kann.

Eng verwandt mit diesem Problem ist das Problem des *uneinheitlichen Trends*.

P51.3.2 Uneinheitlicher Trend

Alle dargestellten Wahlhochrechnungsverfahren setzen voraus, dass die Wählerbewegungen in allen Gemeinden annähernd prozentuell gleich groß sind. Wenn in der kleinen 1000-Personen-Gemeinde die Partei A 1,8% gewinnt, dann wird unterstellt, dass in der Großstadt mit 500 000 Einwohnern der Gewinn in ungefähr derselben Höhe liegt. Dies ist offenkundig häufig nicht der Fall. Es kann beachtliche Unterschiede in den Wählerbewegungen von Stadt zu Land und zwischen den Regionen bzw. Bundesländern geben. Dieses Problem versucht man dadurch zu lösen, dass man in einem 1. Schritt mehrere getrennte Wählerstromanalysen/Hochrechnungen für möglichst homogene Gemeindegruppen durchführt und in einem 2. Schritt die prognostizierten Ergebnisse gewissermaßen zusammenzählt. So kann man etwa in Österreich für jedes der 9 Bundesländer eine getrennte Wahlhochrechnung rechnen. In einem 2. Schritt werden dann die 9 Ergebnisse zusammengezählt. Wir werden in Abschnitt P51.4.8 eine "gruppierte" Hochrechnung vorführen, bei der die 445 oberösterreichischen Wahlgemeinden in 4 Gruppen nach der Gemeindegröße zusammengefasst werden.

Clusteranalyse von Wahlgemeinden.

Wir unterstellen, dass für die Wahlgemeinden "Strukturmerkmale" bekannt sind, die mit dem Wählerverhalten der Gemeinden korrelieren. Gemeint sind Merkmale wie Geschlechterproportion, Alterspyramide, Berufsgruppen-Verteilung, Industrie-, Handels-, Landwirtschafts-Anteil, usw. Hier bietet es sich nun an, durch eine Clusteranalyse (mit diesen Strukturmerkmalen als Klassifikationsvariable) homogene Gemeindegruppen zu bilden, für die jeweils eine selbständige Hochrechnung bzw. Wählerstromanalyse gerechnet wird, und deren jeweilige Ergebnisse dann "zusammengezählt" werden. Siehe dazu Abschnitt P51.4.4.1, Eingabebox 12. Als Klassifikationsvariable könnten auch unmittelbar die stärksten Parteien (inklusive der "Partei" der Nichtwähler) der vergangenen Wahl mit ihren Prozentwerten eingesetzt werden. In Österreich verwendet ein Institut die SPÖ mit ihren Prozentwerten als einzige Partei zur Bildung von 5 Clustern.

Es ist sicherlich sehr arbeitsaufwändig, eine Hochrechnung zu organisieren, die das beschriebene Verfahren verwendet.

Der Trend kann im extremen Fall gegenläufig sein. Beispiel: Im Osten des Wahlgebiets verläuft er gegenläufig zum Westen. Der Normalfall dürfte aber sein, dass der Trend zwar gleichgerichtet aber deutlich verschieden stark ist. Das ist vor allem von Stadt zu Land so. Eine Möglichkeit dieses „Stadt-Land-Problem“ pragmatisch zu lösen besteht darin, dass man die Gemeinden unterschiedlich gewichtet. Betrachten wir ein Beispiel: Zu einem frühen Zeitpunkt der Hochrechnung liegen üblicherweise die Ergebnisse aus den Städten nur sehr unvollständig vor. Nun liegt der Gedanke nahe, die wenigen Städte, deren Ergebnisse vorliegen, so zu gewichten, dass das Verhältnis von Einwohnerzahl der Städte zu den ländlichen Gemeinden stimmt.

Gemeinden mit über etwa 20 000 Wahlberechtigten (also Städte) werden z.B. mit dem Faktor 1.25 gewichtet. Wenn beispielsweise 10 Städte vorhanden sind, dann gehen diese in die Hochrechnung so ein wie wenn sie 12.5 Städte wären. Oder anders formuliert: Zu den 10 vorhandenen Städten werden 2.5 artifizielle Städte erzeugt, deren Partei-Stimmenzahl dem Durchschnitt der 10 vorhandenen entspricht. Die Gewichtung gilt für alle "Parteien der aktuellen Wahl" und alle "Parteien der vergangenen Wahl", also auch für Nicht- und Ungültigwähler und für Zu- und Abwanderer. Siehe die ausführliche Darstellung in Abschnitt P51.4.9.

P51.3.3 Auswahl der Stimmbezirke

Wenn die Zahl der Stimmbezirke sehr groß ist und es (z.B. aus organisatorischen Gründen) nicht möglich ist, alle in eine Hochrechnung einzuschließen, dann kann man sich mit einer Stichprobe von Stimmbezirken begnügen. Die prognostizierte Stimmenverteilung wird allerdings dann weniger genau sein. Bei vollständiger Auszählung der Stichprobe sind dann die Ergebnisse um den Stichprobenfehler gegenüber dem tatsächlichen Wahlergebnis verzerrt.

Betrachten wir ein Beispiel: Für die Hochrechnung bei der deutschen Bundestagswahl müssen rund 8000 Stimmbezirke ausgewertet werden. Das die Hochrechnung durchführende Institut entschließt sich jedoch eine Stichprobe von rund 400 Stimmbezirken auszuwählen.

Aus der amtlichen Statistik kennt man die Verteilung der Gemeinden auf Größenklassen. Nun kann es sein, dass in der Stichprobe beispielsweise Stimmbezirke aus der Größenklasse 10000 bis 20000 Einwohner etwas unterrepräsentiert sind. Hier bietet es sich an, in die Datei der vergangenen Wahl eine zusätzliche Variable "Gewichtsfaktor" aufzunehmen. Diese Variable enthält für

jeden Stimmbezirk eine solche Gewichtungszahl, dass die Größenklassen-Verteilung exakt der tatsächlichen entspricht.

Man kann auch für weitere Merkmale (z.B. für die "Katholiken-Protestanten-Verteilung) Gewichtungen einführen - wobei es sich empfiehlt die Gewichtungszahlen aus verschiedenen Merkmalen zu multiplizieren und nur eine Variable "Gewichtsfaktor" zu bilden.

P51.3.4 Prozentwerte oder absolute Stimmenzahlen

Was soll der „input“ in eine Hochrechnung sein? Sollen die prozentuellen Ergebnisse der Parteien je Gemeinde oder die absoluten Stimmenzahlen als Eingabewerte verwendet werden? Dieses Problem steht in Verbindung mit dem bereits dargestellten „Stadt-Land-Problem“. Werden die Prozentergebnisse als Eingabewerte verwendet, dann wird eine Großstadt so behandelt wie ein kleines Dorf. Wird mit absoluten Stimmenzahlen gerechnet, dann erhalten Städte ein überwältigendes Gewicht. Der Logik der Hochrechnung, dass die Stimmenverschiebungen in allen Wahlgemeinden *prozentuell* ungefähr gleich groß sind, würde die Eingabe von Prozentwerten entsprechen. Einen Kompromiss, der sich in der Praxis gut bewährt hat, schlägt Bruckmann (1966, S. 20) vor. Als Eingabedaten werden verwendet

$$\frac{S_{ik}}{\sqrt{n_k}}$$

S_{ik} = Stimmenzahl der Partei i in der Gemeinde k

n_k = Wahlberechtigte in der Gemeinde k

Dieser Kompromiss hat auch eine statistische Begründung: die vorhandene „Heteroskedastizität“ wird verringert. In unseren Programmen Prog51m1 und Prog28m2 zu Hochrechnung und Wählerstromanalyse haben wir diese Transformation etwas modifiziert. Im Nenner steht - auch für die Parteien der vergangenen Wahl - die Wurzel aus den Wahlberechtigten der Gemeinde k aus der *aktuellen* Wahl. Dadurch wird erreicht, dass die oft ungleichen Wahlberechtigtenzahlen der Gemeinden aus beiden Wahlen ungefähr aneinander angepasst werden - wobei die Stimmenrelationen zwischen den Parteien erhalten bleiben.

P51.3.5 Ungleiche Zahl der Wahlberechtigten bei den beiden Wahlen

Wenn die Gesamtzahl der abgegebenen Stimmen in den Gemeinden zur aktuellen und zur vergangenen Wahl verschieden groß ist, dann kann man die prozentuellen Gewinne bzw. Verluste der Parteien nicht richtig berechnen. Dieses Problem kann man, wie wir dies schon gezeigt haben, dadurch lösen, dass man die Nichtwähler als „Partei“ in das Modell mit einbezieht.

Wenn jedoch die Zahl der Wahlberechtigten bei den beiden Wahlen verschieden ist, dann ist das Problem dadurch noch nicht gelöst. Es ist durchaus möglich, dass es durch Zu- bzw. Abwanderung und durch ein Ungleichgewicht bei der Zahl der Erstwähler im Verhältnis zur Zahl der Verstorbenen in den Wahlgemeinden zu ungleichen Zahlen der Wahlberechtigten bei der vergangenen und der aktuellen Wahl kommt. In diesem Fall besteht die Lösung des Problems darin, die Erstwähler je Gemeinde wie eine Partei als unabhängige Variable und die Gestorbenen je

Gemeinde als abhängige Variable in das Modell aufzunehmen. Das ermöglicht dann Aussagen darüber, zu welchen Parteien die Erstwähler "geströmt" sind und skurrilerweise auch, in welchem Maße die Wähler der Parteien der vergangenen Wahl in die Gruppe der Gestorbenen "geströmt" sind. Damit erkennen wir, welche Parteien vorzugsweise von alten und umgekehrt von jungen Menschen gewählt werden. "Alte Parteien" sind jene, deren Wähler aus der vergangenen Wahl überproportional in die Gruppe der Gestorbenen geströmt sind. Und "junge Parteien" sind jene, deren Wähler unterproportional zu den Gestorbenen geströmt sind.

Da es mit einem gewissen Aufwand verknüpft ist, die Zahlen der Erstwähler und der Gestorbenen je Wahlgemeinde zu erheben und, in der Vorbereitung der Hochrechnung, in den Computer einzugeben, ist es eher selten, dass diese beiden Gruppen berücksichtigt werden. Wir haben eine derartige Hochrechnung (als Wählerstromanalyse) bei der Landtagswahl 1991 (im Vergleich zu 1985) gerechnet. Siehe dazu das Programm P51_LW91_85. Man findet das Programm durch Klick auf den Knopf "alle Progs" am Oberrand des Almo-Fensters.

Die ungleiche Zahl der Wahlberechtigten zu den beiden Wahlen hat jedoch auch noch andere Ursachen, beispielsweise den ungleichen Zuzug und Abzug von Personen zu und von den Gemeinden.

Durch die Einführung von zwei "neuen" Variablen kann man erreichen, dass die Wahlpopulationen zu beiden Zeitpunkten identisch sind. Diese beiden Variablen sind "Zuwanderungsgewinn" und "Abwanderungsverlust". Sie werden wie folgt gebildet:

Zuwanderungsgewinn:

Diese Variable ist gleich 0, wenn die Wahlpopulation in einer Gemeinde abgenommen hat, sonst ist sie gleich der Differenz aus den Wahlberechtigten der aktuellen Wahl und der vergangenen Wahl.

Abwanderungsverlust:

Diese Variable ist - im Unterschied zur Variablen Zuwanderungsgewinn - gleich 0, wenn die Wahlpopulation in einer Gemeinde zugenommen hat, sonst ist sie gleich der Differenz aus den Wahlberechtigten der aktuellen Wahl und der vergangenen Wahl.

Die Variable Zuwanderungsgewinn wird als eine Partei der vergangenen Wahl (also als *unabhängige Variable*) in das zu schätzende Modell aufgenommen, die Variable Abwanderungsverlust als eine Partei der aktuellen Wahl (also als *abhängige Variable*). Die Frage ist, ob das Konsequenzen hat für Übergangswahrscheinlichkeiten zwischen den echten Parteien.

Wer sind die Zu- und Abwanderer?

Wir haben seither Zu- und Abwanderer als rechnerische Größen begriffen, die aus der Differenz der beiden Wahlpopulationen gebildet werden. Aber wer sind die Zuwanderer und Abwanderer? Dies sind z. B.

- a. Personen, die zwischen den beiden Wahlen verstorben sind
- b. Personen, die zwischen den beiden Wahlen das Wahlalter erreicht haben (Jungwähler)
- c. sonstige Personen, die das Wahlrecht erlangt haben (hauptsächlich durch Einbürgerung)
- d. Personen, die in die Gemeinde zugezogen bzw. weggezogen sind

Zwei Fragen stellen sich nun:

- Sind die Übergangswahrscheinlichkeiten zwischen den echten Parteien andere, wenn mit oder ohne Zu- und Abwanderung gerechnet wird.
- Sind die Übergangswahrscheinlichkeiten zwischen den echten Parteien abhängig davon, aus welchen konkreten Personen die Zu- und Abwanderer zusammengesetzt sind.

Macht es beispielsweise einen Unterschied, ob die Zu- und Abwanderer primär aus den Jungwählern/Verstorbenen bestehen oder aus dem Zustrom in die Ballungsräume aus den umliegenden Kleingemeinden.

Wir haben oben in Abschnitt P51.1.3 ausgeführt, dass das Verfahren der quadratischen Optimierung *konsistent* ist. Bei vollständiger Auszählung der Wahlgemeinden (also für den Fall der Wählerstromanalyse) reproduziert das Verfahren exakt die richtigen Prozentwerte für die echten Parteien. Und dies gilt unabhängig davon, ob die Zu- und Abwanderer in den Kalkül einbezogen sind oder nicht.

Das erlaubt dann den folgenden Schluss: Für die Übergangsmatrix der echten Parteien ist es belanglos, aus welchen Personengruppen die Zu- und Abwanderer zusammengesetzt sind.

Das gilt aber nicht für die unvollständige Auszählung, also für den Fall der Hochrechnung. Wir können aber behaupten, dass die Übergangsmatrix der echten Parteien nur geringfügig und mit zunehmender Auszählung immer weniger davon beeinflusst wird, ob Zu- und Abwanderer als Variable sich überhaupt im Modell befinden und wenn ja, aus welchen Personengruppen sie sich zusammensetzen. Siehe dazu auch die ausführliche Darstellung in Abschnitt P51.4.6 "Möglichkeit 2".

P51.3.6 Splitterparteien

Bei nationalen Wahlen, aber auch bei Landtagswahlen treten oft „Splitterparteien“ auf, die nur minimale Stimmzahlen für sich gewinnen können. Hier hat es sich bewährt, diese Parteien in einer Partei, mit dem Namen „Sonstige“, zusammenzufassen. Dies kann man sowohl bei den Kleinparteien der vergangenen als auch der aktuellen Wahl tun. Werden Parteien zusammengefasst, dann ändern sich auch minimal die Ergebnisse für jene Parteien, die von der Zusammenfassung nicht betroffen sind.

P51.3.7 Prognoseerfolg

Den Erfolg einer Hochrechnung kann man natürlich erst nach der Auszählung aller Stimmen beurteilen. Eine gewisse Überprüfung ist aber auch im Verlauf der Hochrechnung möglich. Ein Beispiel: Bei einer Landtagswahl, bei der 35 von 445 Gemeinden ausgezählt sind, korrelieren die Stimmenergebnisse dieser 35 Gemeinden für die Partei i mit dem vom Hochrechnungs-Modell für diese 35 Gemeinden prognostizierten Stimmenergebnissen mit $r^2=0.9997$. Entsprechend sind die r^2 -Werte für alle Parteien.

Anders ausgedrückt: Die vom Hochrechnungsverfahren errechnete Übergangsmatrix erzeugt für die 35 Gemeinden ein Stimmenergebnis, das mit dem tatsächlichen hervorragend übereinstimmt. Im Prinzip wird mit dieser Berechnung überprüft, ob die Grundannahme der Hochrechnung, dass in allen Gemeinden die prozentuellen Stimmenverschiebungen ungefähr gleich sind, zutrifft.

P51.3.8 Mandatsverteilung

Bei einer Hochrechnung sollte auch die Mandatsverteilung prognostiziert werden.

Die Berechnung der Mandate, die die Parteien erzielen ist relativ kompliziert. Im Almo-Hochrechnungsprogramm kann der Benutzer in einem speziellen Eingabefeld eine „1“ oder eine „2“ eingeben. Damit werden dann Mandats-Berechnungen ausgelöst, die nur teilweise den gesetzlichen Regelungen in Deutschland und Österreich entsprechen.

Gibt der Benutzer "1" ein, dann wird die Mandatsverteilung nach dem System von **Hare/Niemeyer** ermittelt. Dabei gelten folgende Voraussetzungen: (1) Jeder Wähler hat nur 1 Stimme, (2) alle Mandate sind Listenmandate, (3) es werden nur Parteien berücksichtigt, die auf oder über x% (üblicherweise 5%) liegen oder Direktmandate (vermutlich) erhalten, (4) wenn eine Partei mehr als 50% besitzt aber nicht die Mandatsmehrheit erhalten würde, dann wird ihr ein Restmandat zugewiesen. Diese Art der Mandatsermittlung gilt für das 1. Zuteilungsverfahren der deutschen Bundestagswahl und für einige Bundesländer (nach dem Wahlgesetz bis 2011).

Wenn der Benutzer "2" eingibt, dann wird das Verfahren nach **d'Hondt** durchgeführt. Dabei gilt wie bei Hare/Niemeyer obiger Satz (1) bis (3). Satz (4) gilt beim d'Hondt'schen Verfahren nicht. In Almo ist das mehrstufige Verfahren, das bei der deutschen Bundestagswahl angewendet wird, in separaten Programmen enthalten. Siehe dazu die Programm-Masken Prog52m1 und Prog52m2. In beiden Programmen wird die Mandatsverteilung nach dem bis 2011 geltenden Bundeswahlgesetz berechnet. Die Reform der Sitzverteilung von 2011, 2013 und später ist im Programm nicht berücksichtigt. Das Programm muss dafür noch angepasst werden. Das Daten-Beispiel, das in beiden Programmen gerechnet wird, verwendet "erfundene" Daten. Die Original-Daten der deutschen Bundestagswahl stehen uns nicht zur Verfügung. Zu dem Zeitpunkt, als wir das rechnen wollten, waren sie nur zu einem exorbitant hohen Preis erhältlich.

P51.3.9 Briefwahl und Wahlkarten

Die rechtlichen Regelungen für die Briefwahl und die Wahl mit Wahlkarte sind von Land zu Land und auch von einer Wahl zur nächsten so verschieden, dass wir es hier nicht wagen eine generelle Behandlungsweise für die Hochrechnung bzw. Wählerstromanalyse vorzuschlagen. Zu überlegen ist, ob für die Briefwähler eine eigene Wählerstromanalyse oder sogar eine eigene Hochrechnung durchgeführt werden kann.

P51.4 Wählerstromanalyse und Hochrechnung mit ALMO

Das Almo-Wahlhochrechnungssystem umfasst folgende Programme:

1. Wählerstromanalyse und Hochrechnung durch quadratische Optimierung mit **Prog51m1**

Im Almo-Statistiksystem sind neben Prog51m1 noch folgende weitere "Spezialprogramme" enthalten. Diese sind nicht hier im vorliegenden Dokument sondern im älteren Handbuch "P51 Wahl-Hochrechnung und Wählerstromanalyse" beschrieben.

2. Eingabe der Wahlergebnisse der ausgezählten Wahlgemeinden in eine Datei
3. Wahlhochrechnung durch multiple Regression (Prog53m2)
4. Wahlhochrechnung durch einfache Regression mit Konstante (Prog53m4)

5. Wahlhochrechnung durch einfache Regression ohne Konstante (Prog53m3)
6. Einbeziehung von Exitpoll-Daten in Wahlhochrechnung (Prog51m4). Dies ist ein "Experimentalprogramm".
7. Dreistufige Mandatsverteilung für die deutsche Bundestagswahl bei vollständiger Auszählung (Prog52m2). Entspricht dem bis 2011 geltenden Bundeswahlgesetz. Die Reform der Sitzverteilung von 2011 und 2013 und später ist im Programm nicht berücksichtigt
8. Hochrechnung der (dreistufigen) Mandatsverteilung für die deutsche Bundestagswahl bei unvollständiger Auszählung (Prog52m1). Entspricht dem bis 2011 geltenden Bundeswahlgesetz. Die Reform der Sitzverteilung von 2011 und 2013 und später ist im Programm nicht berücksichtigt

Bei den in 2 bis 5 genannten Programmen wird das absolute und das prozentuelle Stimmenergebnis prognostiziert sowie die Mandatsverteilung (wahlweise) nach dem Verfahren von Hare/Niemeyer oder d'Hondt berechnet.

P51.4.1 Eingabe in das Almo-Programm zur Hochrechnung und Wählerstromanalyse

Wir wollten für die Programm-Eingabe keine künstlich erzeugten Daten verwenden. Aktuelle Daten standen uns nicht zur Verfügung - jedoch die sehr alten aber vollständigen Daten der Landtagswahlen von 1997 in Oberösterreich und die Daten der vorausgegangenen Wahl 1991. Wir haben damit die Möglichkeit eine Hochrechnung im nachhinein und eine Wählerstromanalyse zu rechnen.

Von der Landtagswahl 1997 sind uns sogar die Uhrzeiten bekannt, zu denen die Gemeinden ihre Wahlergebnisse an die Zentrale gemeldet haben. Wir können also die Hochrechnung auch für beliebige Zeitpunkte durchführen und damit das Wahlergebnis "prognostizieren". Da wir das tatsächliche Ergebnis kennen, vermögen wir die Güte der "Prognose" zum jeweiligen Zeitpunkt zu beurteilen.

P51.4.1.1 Schritte der Hochrechnung

Die Schritte, die bei einer Hochrechnung bzw. Wählerstromanalyse mit Almo zurückgelegt werden müssen, sind diese:

1. Die Daten der vergangenen Wahl mussten in den älteren Programm-Versionen in eine Direktzugriffs-Datei geschrieben werden. Das geschah mit Prog00mk. In der aktuellen Version entfällt dieser Schritt. Normalerweise liegen die Daten im freien Format vor. Almo konvertiert sie jetzt programm-intern in das direkte Format. Zu den Almo-Dateiformaten siehe das Almo-Dokument 32 "Datei-Operationen mit Almo", Abschnitt 1.
2. Sollen Zu- und Abwanderer miteinbezogen werden, dann sollte die Gesamtzahl der Zuwanderer aus *allen* Wahlgemeinden des Wahlgebiets ermittelt werden. Dies geschieht mit nachfolgend beschriebenen **Prog51mb**. Siehe Abschnitt P51.4.3. Almo bietet folgende 3 Möglichkeiten, mit Zu- und Abwanderer umzugehen. Sie werden in Abschnitt P51.4.6 nochmals ausführlich erläutert.

- (1) Zu- und Abwanderer miteinbeziehen und die Gesamtzahl der Zuwanderer aus der gesamten Wahlpopulation dem Hochrechnungsprogramm Prog51m1 mitteilen
- (2) Zu- und Abwanderer miteinbeziehen. Deren Gesamtzahl aber nur aus der Zahl der ausgezählten Gemeinden ermitteln, d.h. aus den Gemeinden, die zum Zeitpunkt der Hochrechnung schon ausgezählt sind.
- (3) Zu- und Abwanderer ausschließen.

Nur bei der 1. Möglichkeit muss die Gesamtzahl der Zuwanderer bekannt sein. Diese ist allerdings die optimale Möglichkeit. Prog51mb muss als nur für diesen Fall gerechnet werden. Zu beachten ist, dass die Gesamtzahl der Zuwanderer für *alle* Wahlgemeinden des *gesamten* Wahlgebiets vom Benutzer in die Programm-Maske Prog51m1 eingetragen werden muss.

3. Die jeweils in der Zentrale eintreffenden Wahlergebnisse der einzelnen Gemeinden müssen in die "Datei der aktuellen Wahl" eingeschrieben werden. Siehe das Beispiel im nachfolgenden Abschnitt. Am einfachsten ist so vorzugehen:

Trifft in der Zentrale das Ergebnis aus einer Wahlgemeinde ein, dann wird die schon vorhandene Datei der aktuellen Wahl geladen (d.h. in ein Fenster gestellt) und das soeben eingetroffene Ergebnis als letzte Zeile dazu geschrieben.

Almo bietet aber auch alternativ zwei spezielle Dateneingabe-Programme und eine einfache Eingabemethode an. Der Benutzer erreicht sie durch Klick auf das Menü "Almo/Dateneingabe". Der Benutzer lese zuerst die Hilfe und die Manuals. Wurden die Daten mit einem fremden Programm geschrieben, dann ist entscheidend, dass die Daten sich in einem der drei von Almo lesbaren Formaten befinden - vorzugsweise im Format FREI. Der Benutzer klicke im Prog51m1 auf den Hilfefknopf in der Eingabebox für die Datei der aktuellen Wahl.

Für unsere Beispieldaten der Landtagswahl müssen die Wahlergebnisse nicht mehr in eine Datei geschrieben werden. Die gesamte Datei der aktuellen Wahl liegt ja bereits vor.

4. Das eigentliche Hochrechnungs- und Wählerstromanalyse-Programm ist dann **Prog51m1**. Es wird in Abschnitt P51.4.4 ausführlich beschrieben.

Will der Almo-Benutzer nur mit unseren Daten aus der Landtagswahl rechnen, dann benötigt er die in Punkt 2 und 3 genannten Programme nicht. In der Program-Maske Prog51m1 sind bereits die Dateien eingesetzt, ebenso die Gesamtzahl der Zuwanderer. Auch die gesamte Datei der aktuellen Wahl liegt bereits vor.

P51.4.2 Die Dateien der vergangenen und der aktuellen Wahl

Die Datei der aktuellen Wahl

Wir haben zwei Dateien. Die eine Datei ist die der *aktuellen* Wahl. Sie enthält die Gemeinden, die ihre Wahlergebnisse schon an die Zentrale gemeldet haben. Im Verlauf der Hochrechnung wächst sie um eine Gemeinde nach der anderen, also um eine Datenzeile nach der anderen. In unseren Beispieldaten verfügen wir schon über alle Gemeinden (zusätzlich mit der Uhrzeit der Meldung an die Zentrale). Betrachten wir einen Ausschnitt aus der Datei der aktuellen Wahl:

```

40101 136473 92046 2366 89680 27351 30249 19902 7409 3093 485 455 207 226 303 2336
40201 28829 21214 561 20653 5173 8084 4771 1598 668 157 131 71 0 0 1851
40301 38449 28111 654 27457 7957 8513 7445 2094 978 169 95 113 93 0 2032
.      .      .      .      .      .      .      .      .      .      .      .      .      .      .
41824 1779 1593 26 1567 622 482 334 72 27 16 4 6 4 0 1642

```

Die Daten befinden sich im Format "frei". Die Zahlenwerte sind durch ein Blank getrennt. In der 1. Spalte steht die Gemeinde-Nummer. So ist z.B. 40101 die Gemeindenummer von Linz. Darauf folgen die Variablen Wahlberecht97, AbgegebenStimm97 usw. Ab der 6. Spalte beginnen die Parteien mit OEVP97, dann SPOE97 usw. In der Programm-Maske Prog51m1 entsprechen die in die Eingabebox "Variablennamen" eingetragenen Namen der Reihenfolge der Spalten in der Datei. Dies sind die Namen:

Namen der Parteien aus der aktuellen Wahl

↔	Name 1 =GemeindeNr97;	Parteien aus der aktuellen Wahl
↔	Name 2 =Wahlberecht97;in der Reihenfolge, in der sie
↔	Name 3 =AbgegebenStimm97;	in der Datei hintereinander stehen
↔	Name 4 =UnguelStimm97;	
↔	Name 5 =GuelStimm97;	
↔	Name 6 =OEVP97;	
↔	Name 7 =SPOE97;	
↔	Name 8 =FPOE97;	
↔	Name 9 =GRUENE97;	
↔	Name 10 =LIF97;	
↔	Name 11 =NEUTR97;	
↔	Name 12 =KPOE97;	
↔	Name 13 =CWG97;	
↔	Name 14 =FDP97;	
↔	Name 15 =OENP97;	
↔	Name 16 =Zeit97;	

Die Datei der vergangenen Wahl

Die andere Datei ist die der *vergangenen* Wahl von 1991. Sie liegt vollständig vor, d.h. sie umfasst alle Gemeinden mit ihren Daten und Wahlergebnissen. Betrachten wir ebenfalls einen Ausschnitt aus dieser Datei:

```

40101 139812 103617 1979 101638 33494 40312 18993 4830 4009
40201 29454 22977 533 22444 6504 9592 4727 452 1169
40301 38005 29453 538 28915 9434 11165 6333 705 1278
.      .      .      .      .      .      .      .      .      .
41824 1469 1353 10 1343 610 455 238 21 19

```

Die Daten wurden ebenfalls im freien Format geschrieben. Die 1. Spalte ist wieder die Gemeindenummer. Die Wahlergebnisse beginnen in der 6. Spalte mit OEVP91. Die Reihenfolge der Variablen kann in Programm-Maske Prog51m1 aus der Eingabebox für die Variablennamen der vergangenen Wahl entnommen werden. Dies sind die Namen:

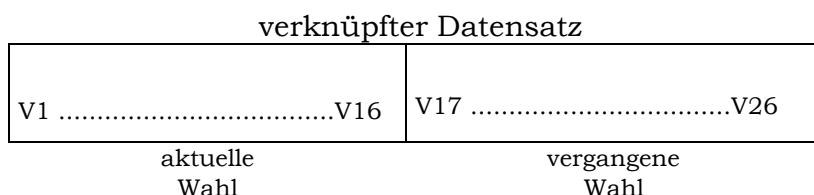
Namen der Parteien aus der vergangenen Wahl

↔	Name 17 =GemeindeNr91;	Parteien aus der vergangenen Wahlin der Reihenfolge, in der sie in der Datei hintereinander stehen Die Nummerierung muss an die letzte Nummer aus der aktuellen Wahl anschliessen
↔	Name 18 =Wahlberecht91;	
↔	Name 19 =AbgegebenStimm91;	
↔	Name 20 =UnguelStimm91;	
↔	Name 21 =GuelStimm91;	
↔	Name 22 =OEVP91;	
↔	Name 23 =SPOE91;	
↔	Name 24 =FPOE91;	
↔	Name 25 =VGOE91;	
↔	Name 26 =GAL91;	

Wir werden anschließend noch erläutern, warum wir den Variablen der vergangenen Wahl Variablennummern ab V17 zuweisen.

Wird eine Hochrechnung gerechnet, z.B. nachdem 35 Gemeinden mit ihren Ergebnissen vorliegen - die Datei der aktuellen Wahl also erst 35 Datensätze umfasst - dann liest Almo den ersten Datensatz und sucht in der (vollständigen) Datei der vergangenen Wahl nach derselben Gemeinde und verknüpft die beiden Datensätze zu einem gemeinsamen. Danach liest Almo den 2. Datensatz aus der Datei der aktuellen Wahl ein und sucht in der Datei der vergangenen Wahl nach derselben Gemeinde und so weiter für die vorhanden 35 Datensätze. Zum Schluss sind dann 35 verknüpfte Datensätze vorhanden.

Der "verknüpfte" Datensatz sieht dann folgendermaßen aus:



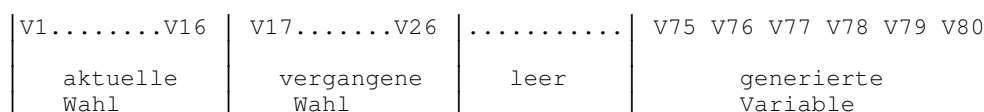
Von Variable 1 bis 16 reichen die Variablen der aktuellen Wahl. V1 ist die Gemeindenummer, V7 die Stimmenzahl für die SPÖ usw.

Von V17 bis V26 reichen die Variablen der vergangenen Wahl. V17 ist hier die Gemeindenummer, V23 die SPÖ usw.

V1 ist gleich V17.

Die Verknüpfung der beiden Datensätze zu einem gemeinsamen nimmt Almo im Verlauf der Hochrechnung in Prog51m1 selbst vor.

Dem verknüpften Datensatz werden nun noch die *generierten* Variablen der Nichtwähler (V75,76), der "sonstigen Parteien" (V77,78) und der Zu- und Abwanderer (V79,80) angehängt. Der gesamte Datensatz ist dann folgender:



V75 V76 =Nichtwähler aus aktueller, vergangener Wahl
V77 V78 =Sonstige aus aktueller, vergangener Wahl
V79 V80 =Zu-, Abwanderer

P51.4.3 Gesamtzahl der Zuwanderer mit Prog51mb ermitteln

Hochrechnung bzw. Wählerstromanalyse kann muss aber nicht Zu- und Abwanderer als Pseudo-Partei einschließen. Siehe dazu P51.1.3, P51.4.6 und P51.5. Werden sie mit eingeschlossen, dann wird, wenn alle Gemeinden ausgezählt sind, die Stimmenzahl, die den Parteien zuströmt, exakt vom Modell reproduziert.

In unserem Hochrechnungsmodell sind die beiden Variablen "Zuwanderer" und "Abwanderer" formal so definiert:

```
Zuwanderer in Gemeinde i =
    Wahlberechtigte in Gemeinde i bei aktueller Wahl
    minus
    Wahlberechtigte in Gemeinde i bei vergangener Wahl

Abwanderer in Gemeinde i =
    Wahlberechtigte in Gemeinde i bei vergangener Wahl
    minus
    Wahlberechtigte in Gemeinde i bei aktueller Wahl

Zuwanderer = Summe aller Zuwanderer aus allen Gemeinden
Abwanderer = Summe aller Abwanderer aus allen Gemeinden
```

Ergibt sich in der Gemeinde i für Zu- bzw. Abwanderer eine negative Zahl, so wird diese auf 0 gesetzt. Eine Gemeinde kann entweder nur Zuwanderung oder nur Abwanderung besitzen.

Zu- bzw. Abwanderer sind also nicht nur Personen die physisch einer Gemeinde zu- gewandert bzw. aus ihr abgewandert sind. Zu ihnen zählen auch Erstwähler und Verstorbene. Sie sind rein rechnerische Größen, die sich aus der Differenz der Wahlberechtigten der beiden Wahlen ergeben. Siehe auch Abschnitt P51.3.5.

Für die Wahl-Hochrechnung wird die Summe *aller* Zuwanderer gebraucht. Die Gesamtzahl der Abwanderer muss vom Benutzer nicht eingegeben werden.

Sollen Zu- und Abwanderer miteinbezogen werden, dann muss dem Programm die Gesamtzahl der Zuwanderer aus *allen* Wahlgemeinden in der Eingabebox 10 im Hochrechnungsprogramm Prog51m1 mitgeteilt werden. Der Grund dafür ist folgender: Im Kalkül der Hochrechnung/Wählerstromanalyse werden die Abwanderer als abhängige Variable, d.h. als Pseudo-Partei der aktuellen Wahl behandelt und die Zuwanderer als unabhängige Variable, d.h. als Pseudo-Partei der vergangenen Wahl. Für die Parteien und die Pseudo-Parteien der vergangenen Wahl müssen die Stimmenzahlen bekannt sein. Almo errechnet aus den zum jeweiligen Zeitpunkt ausgezählten Wahlgemeinden die Matrix der Übergangswahrscheinlichkeiten. Diese wird mit den Stimmenzahlen aus der vergangenen Wahl multipliziert. Damit entsteht dann das für die aktuelle Wahl prognostizierte Ergebnis.

Die Gesamtzahl der Zuwanderer aus allen Gemeinden kann nur errechnet werden, wenn die Zahl der Wahlberechtigten für jede einzelne Gemeinden aus vergangener und aktueller Wahl (schon vor der Hochrechnung) bekannt ist. Das ist normalerweise auch der Fall.

Mit **Prog51mb** kann dann die Gesamtzahl der Zuwanderer ermittelt werden. Wie dieses Programm in Almo gefunden und gestartet wird, ist im nachfolgenden Abschnitt P51.4.4 im Unterabschnitt "Kurzanleitung" beschrieben.

Prog51mb.Msk
"Zuwanderer" für Wahl-Hochrechnung ermitteln

Gebraucht wird

1. Eine aktuelle Datei der Wahlgemeinden. Sie muss enthalten
 - die Zahl der Wahlberechtigten je Wahlgemeinde für die aktuelle Wahl
 - eine Gemeindenummer, die als Verbindungsvariable zur vergangenen Wahl eingesetzt werden kann
2. Eine Datei der Wahlgemeinden zum Zeitpunkt der vergangenen Wahl. Sie muss enthalten
 - die Zahl der Wahlberechtigten je Wahlgemeinde für die vergangene Wahl.
 - eine Gemeindenummer, die als Verbindungsvariable zur aktuellenn Wahl eingesetzt werden kann


Sinnvollerweise wird man die Datei der vergangen Wahl verwenden. Im Beispiel ist dies die Datei "LWahl91.fre".

Siehe die ausführliche Erläuterung in Abschnitt P51.4.2 aus Almo-Dokument 33 "Wählerstromanalyse und Wahlhochechnung"

Programm-Bedienung ---> Hilfe



Speicher fuer x Variable Hilfe


Vereinbare Variable= 80

 Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert



Variablennamen Hilfe


Namen der Variablen aus der aktuellen Wahl

 Name 1 =GemeindeNr97;
 Name 2 =Wahlberecht97;


 erzeuge weitere Namensfelder für Variable der aktuellen Wahl



Namen der Variablen aus der vergangenen Wahl



 Name 3 =GemeindeNr91;
 Name 4 =Wahlberecht91;




 erzeuge weitere Namensfelder für Variable der vergangenen Wahl

aktuelle Datei der Wahlgemeinden



 ".\Testdat\ZahlWahlberecht97.fre"

 frei Format der Daten  Hilfe


  V1:2 der Datensatz enthält diese Variablen
Bei Format DIREKT schreiben Sie: alle_v



  V1  Hilfe



Verbindungsvariable aus der aktuellen Wahl zur vorausgegangenen Wahl
z.B. die Gemeindefnummer aus aktueller Wahl
Schreiben Sie V.. (nicht den Variablennamen)




 Wenn Dateiformat erweitert-FREI oder FIX  Hilfe

Datei der Wahlgemeinden aus vergangenen Wahl



 ".\Testdat\LWahl91.fre"


 frei Format der Daten  Hilfe

  V3:12 der Datensatz enthält diese Variablen


  V3  Hilfe


Verbindungsvariable (z.B. Gemeindefnummer) aus der vergangenen Wahl (schreiben Sie V..)


 Wenn Dateiformat erweitert-FREI oder FIX  Hilfe

 Umkodierungen und Kein-Wert-Angaben

Zuwanderer aus Wahlberechtigten ermitteln

 Wahlberecht97 Wahlberechtigte in aktueller Wahl

 Wahlberecht91 Wahlberechtigte in vergangener Wahl

 Option: Ein- und Ausschliessen von Wahlgemeinden

Programmende

Normalerweise wird man schon bevor gewählt wird, eine Datei finden können, in der neben anderen Merkmalen auch die Zahl der Wahlberechtigten der Gemeinden des Wahlgebiets enthalten ist. Wir benötigen für unser Programm nur diese Zahl und dazu die Gemeindefnummer.

In unserem Beispiel haben wir unterstellt, dass wir für die aktuelle Wahl 1997 über eine Datei verfügen, die nur die beiden Variablen

V1 GemeindeNr97


V2 Wahlbrecht97



für sämtliche Wahlgemeinden des Wahlgebiets enthält. Ein Ausschnitt aus dieser Datei ist etwa folgender



V1 GemeindeNr97	V2 Wahlbrecht97
40101	136473
40201	28829
40301	38449
40401	3587
.	.
.	.
.	.




In der Programm-Maske wird diese Datei über folgende Eingabebox definiert

aktuelle Datei der Wahlgemeinden

 **\".\\Testdat\\ZahlWahlberecht97.fre\"**

 **frei** Format der Daten  **Hilfe**


  **V1:2** der Datensatz enthält diese Variablen
Bei Format DIREKT schreiben Sie: alle_V



  **V1**  **Hilfe**



Verbindungsvariable aus der aktuellen Wahl zur vorausgegangenen Wahl
z.B. die **Gemeindenummer aus aktueller Wahl**
Schreiben Sie V.. (nicht den Variablennamen)




Als Gemeindedatei der *vergangenen* Wahl 1991 haben wir die Datei "LWahl91.fre" verwendet, die auch im Hochrechnungsprogramm Prog51m1 eingesetzt wird.

Datei der Wahlgemeinden aus vergangenen Wahl

 **\".\\Testdat\\LWahl91.fre\"**

 **frei** Format der Daten  **Hilfe**

  **V3:12** der Datensatz enthält diese Variablen

  **V3**  **Hilfe**

Verbindungsvariable (z.B. Gemeindenummer) aus der vergangenen Wahl (schreiben Sie V..)

Diese Datei umfasst die Variablen

- V3 GemeindeNr91
- V4 Wahlberecht91
- V5 AbgegebenStimm91
- V6 UngültStimm91
- V7 GültStimm91
- V8 OEP91
- V9 SPOE91
- V10 FPOE91
- V11 VGOE91
- V12 GAL91

Etwas verwirrend ist, dass die Variablen mit den Nummern V3 bis V12 eingelesen

werden. In Programm Prog51m1 hingegen werden Sie mit V17 bis V26 eingelesen. Der Grund dafür ist folgender:

Die beiden Gemeindedateien für die aktuelle und die vergangene Wahl werden von Almo zu einer gemeinsamen Datei verknüpft. Das funktioniert so: Almo liest den ersten Datensatz aus der Gemeindedatei für die aktuelle Wahl ein. Danach durchsucht es die Gemeindedatei der vergangenen Wahl nach derselben Gemeindenummer. Ist diese Gemeinde gefunden, dann wird ihr Datensatz *hinter* den aus der aktuellen Wahl gestellt. Dadurch entsteht eine neuer gemeinsamer Datensatz. Auf diese Weise wird für alle Datensätze der aktuellen Gemeindedatei der zugehörige Datensatz aus der Gemeindedatei der vergangenen Wahl gefunden und angefügt.

Bei dieser Verknüpfung der beiden Dateien wird nach der letzten Nummer der aktuellen Wahl kontinuierlich weiter nummeriert. Wichtig ist zu erkennen, dass nach der Verknüpfung die Variablen `GemeindeNr91` und `Wahlberecht91` aus der vergangenen Wahl jetzt die Variablennummern V3 und V4 besitzen. Das ist wichtig für die Zuweisung von Namen in der Eingabebox "Variablenamen". Für die Programm-Maske wird nur der Name für die Wahlberechtigten gebraucht. Für die Gemeindenummer und die weiteren Variablen muss keine Name vergeben werden (kann jedoch).

Ein Ausschnitt aus der Datei der vergangenen Wahl ist folgender

V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
GemeindeNr91	Wahlberecht91								
40101	139812	103617	1979	101638	33494	40312	18993	4830	4009
40201	29454	22977	533	22444	6504	9592	4727	452	1169
40301	38005	29453	538	28915	9434	11165	6333	705	1278
40401	3529	3078	61	3017	1200	966	753	24	74
.
:	:	:	:	:	:	:	:	:	:
.

Nach der Verknüpfung mit der Gemeindedatei der aktuellen Wahl entstehen folgende Datensätze:

aktuelle Wahl		vergangene Wahl									
V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
40101	136473	40101	139812	103617	1979	101638	33494	40312	18993	4830	4009
40201	28829	40201	29454	22977	533	22444	6504	9592	4727	452	1169
40301	38449	40301	38005	29453	538	28915	9434	11165	6333	705	1278
40401	3587	40401	3529	3078	61	3017	1200	966	753	24	74
.
.

Die Gemeindenummer in V1 ist gleich der in V3.

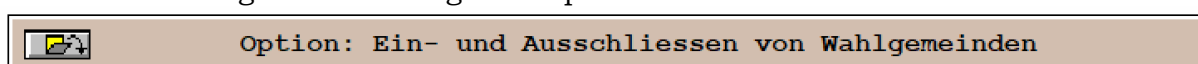
Nachdem für Gemeinde i der verknüpfte Datensatz erzeugt ist, kann die Differenz gebildet werden.

$$\text{Zuwanderer in Gemeinde } i = V2 - V4 = \text{Wahlberecht97} - \text{Wahlberecht91}$$

$$\text{Abwanderer in Gemeinde } i = V4 - V2 = \text{Wahlberecht91} - \text{Wahlberecht97}$$

Ist die Differenz negativ, dann wird sie auf 0 gesetzt. Die Gesamtzahl der Zuwanderer ist dann die Summe der Zuwanderer aus allen Gemeinden.

Die vorletzte Eingabebox ist folgende Optionsbox



Die Optionsbox muss nur geöffnet werden, wenn eine *gruppierte Hochrechnung / Wählerstromanalyse* gerechnet werden soll. In Abschnitt P51.4.4.1 Eingabebox 12 wird diese Variante ausführlich beschrieben.

Almo gibt folgendes Ergebnis aus dem Prog51mb aus:

Ergebnisse aus Almo

```
-----  
Zahl der Wahlgemeinden = 445  
Gesamtzahl der Wahlberechtigten der vergangenen Wahl = 948245  
Gesamtzahl der Wahlberechtigten der aktuellen Wahl = 973750  
  
Gesamtzahl der Zuwanderer = 31478      Abwanderer = 5973
```

Die Abwandererzahl wird in der Programm-Maske Prog51m1 für Hochrechnung bzw. Wählerstromanalyse nicht gebraucht.

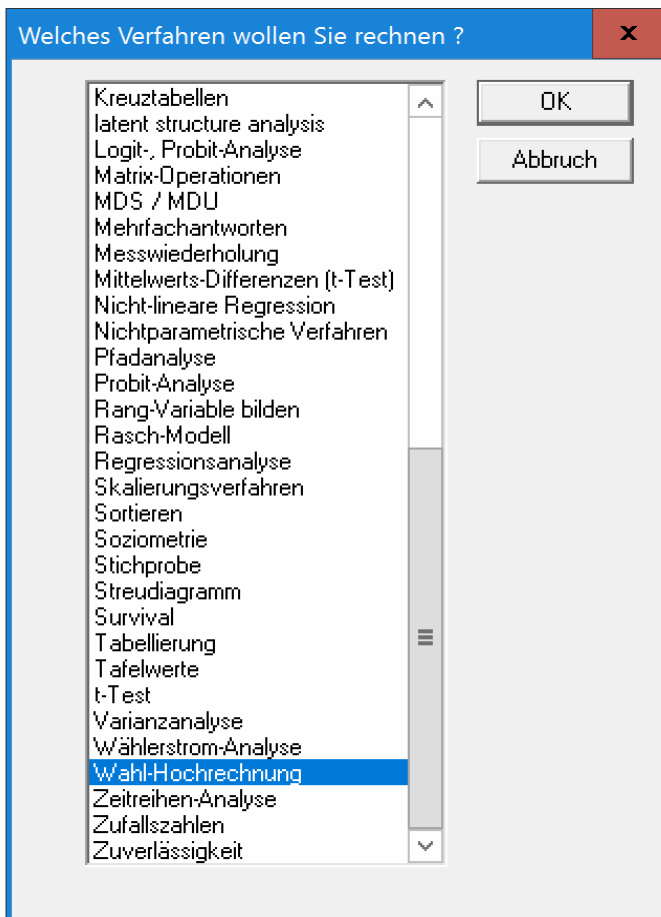
P51.4.4 Eingabe in das Hochrechnungsprogramm Prog51m1

Wir zeigen im folgenden, wie das Maskenprogramm zur Hochrechnung Und Wählerstromanalyse durch "quadratische Optimierung" auszufüllen ist. Die anderen Maskenprogramme für die Hochrechnung durch "multiple Regression" und durch "einfache Regression" mit und ohne Konstante sind weitgehend gleich. Wir werden im vorliegenden Dokument nicht darauf eingehen.

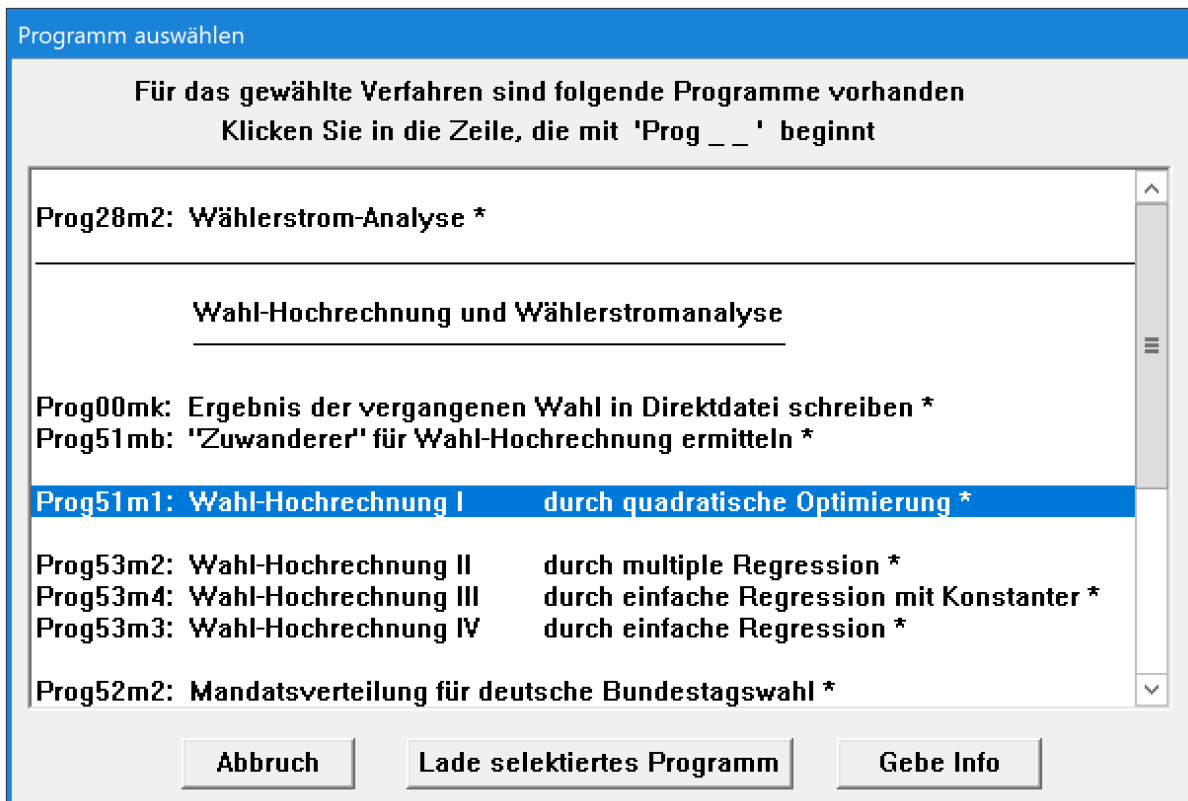
Wie man mit ALMO rechnet ist im Handbuch Teil 1 – Bedienungsanleitung Abschnitt 4 ausführlich dargestellt. Wir geben im folgenden eine Kurzanleitung.

Kurzanleitung:

- a. Starten Sie Almo. Es erscheint das noch leere Almo-Fenster. In der Knopfleiste am Oberrand des Almofensters
- b. klicken Sie auf den Knopf "Verfahren". Almo präsentiert Ihnen dann eine Übersicht über die vorhandenen Verfahren.



Klicken Sie auf "Wahl-Hochrechnung" oder "Wählerstromanalyse".
Dann erscheint eine Übersicht über die vorhandenen Maskenprogramme



c. Klicken Sie nun auf 'Prog51m1.Msk: Wahl-Hochrechnung I' und dann am unteren Rand 'Lade Maskenprogramm' (oder gleich Doppelklick auf Prog51m1.Msk). Es erscheint eine Eingabemaske, die auf den folgenden Seiten abgebildet ist.

d. Nachdem Sie diese Eingabemaske entsprechend Ihren Analyseabsichten ausgefüllt haben, klicken Sie in der Knopfleiste auf den Knopf "Rechne".

e. Almo präsentiert Ihnen eine Dialogbox, in der Ihnen für das Abspeichern des Almo-Maskenprogramms ein Name vorgeschlagen wird. Klicken Sie OK.

f. Dann wird Ihnen eine weitere Dialogbox präsentiert, in der Ihnen für das Abspeichern des Ergebnisses ein Dateiname vorgeschlagen wird. Klicken Sie OK.

g. Almo bringt jetzt ein neues Fenster (mit dem Titel "Almo rechnet") auf den Bildschirm, in dem Sie die Rechenfortschritte Ihres Programms verfolgen können.

h. Wenn als Status 'fertig' gemeldet wird, dann klicken Sie in diesem Fenster auf den Knopf LADE ERGEBNIS.

i. Almo lädt das Ergebnis in ein Fenster des Editor, wo Sie es sich anschauen können.

j. Wenn Sie durch die "Ergebnis-Datei" scrollen oder blättern, dann werden Sie einen großen Knopf "Grafik" finden. Klicken Sie auf diesen.

k. Almo öffnet seinen Grafik-Editor und stellt die Tabelle aus der Ergebnis-Datei als Diagramm dar.

1. Mit Klick auf den Knopf "Zurück" gelangen Sie wieder in die Ergebnis-Datei.

Dies ist nun die Programm-Maske zur Hochrechnung bzw. Wählerstromanalyse. Als Beispiel wurde die Landtagswahl in Oberösterreich 1997 verwendet. Die Daten sind also sehr alt. Wie bereits oben in Abschnitt P51.4.1 ausgedrückt, liegen uns die Daten zu dieser Wahl und der vorausgegangen Wahl von 1991 vollständig vor. Wir verfügen sogar über Strukturmerkmale der Wahlgemeinden und z.B. auch über die Uhrzeiten, zu denen die Gemeinden ihre Wahlergebnisse an die Zentrale gemeldet haben. Wir können also die Hochrechnung auch für beliebige Zeitpunkte durchführen und damit das Wahlergebnis für verschiedene Zeitpunkte des Wahl-Nachmittags und -Abends "prognostizieren".

Prog51m1.Msk
Wahl-Hochrechnung
Wählerstromanalyse
durch quadratische Optimierung

Prognose der Ergebnisse der aktuellen Wahl
auf Basis der Ergebnisse der vergangenen Wahl

Doppelklick
|

weitere Beispiel-Programme:

Landtagswahl	-->	[".\Almo_Bsp\P51_LW91_85.Alm"]
Almo-Dokument 33, P51.5	-->	[".\Almo_Bsp\P51Bspl.Alm"]
"gruppierte" Hochrechnung	->	[".\Almo_Msk\Prog51m1_Grp1.Msk"]

siehe
Almo-Dokument 33 "Wählerstromanalyse und Wahlhochrechnung"

Programm-Bedienung ---> Hilfe

Speicher fuer x Variable Hilfe

Vereinbare Variable= 100

Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

Datei der Variablennamen

zeige = Namensdatei in Output zeigen
 leer = nicht zeigen

Namen der Variablen aus der aktuellen Wahl

in der Reihenfolge, in der sie in der Datei hintereinander stehen

Leere alle Eingabefelder dieser Sub-Box

<input type="button" value="↔"/>	Name 1 =GemeindeNr97;
<input type="button" value="↔"/>	Name 2 =Wahlberecht97;
<input type="button" value="↔"/>	Name 3 =AbgegebenStimm97;
<input type="button" value="↔"/>	Name 4 =UnguelstStimm97;
<input type="button" value="↔"/>	Name 5 =GuelstStimm97;
<input type="button" value="↔"/>	Name 6 =OEVP97;
<input type="button" value="↔"/>	Name 7 =SPOE97;
<input type="button" value="↔"/>	Name 8 =FPOE97;
<input type="button" value="↔"/>	Name 9 =GRUENE97;
<input type="button" value="↔"/>	Name 10 =LIF97;
<input type="button" value="↔"/>	Name 11 =NEUTR97;
<input type="button" value="↔"/>	Name 12 =KPOE97;
<input type="button" value="↔"/>	Name 13 =CWG97;
<input type="button" value="↔"/>	Name 14 =FDP97;
<input type="button" value="↔"/>	Name 15 =OENP97;
<input type="button" value="↔"/>	Name 16 =Zeit97;

erzeuge weitere Namensfelder für Variable der aktuellen Wahl

Namen der Variablen aus der vergangenen Wahl

in der Reihenfolge, in der sie in der Datei hintereinander stehen
 Nummerierung muss an letzte Nummer aus aktueller Wahl anschliessen

<input type="button" value="↔"/>	Name 17 =GemeindeNr91;
<input type="button" value="↔"/>	Name 18 =Wahlberecht91;
<input type="button" value="↔"/>	Name 19 =AbgegebenStimm91;
<input type="button" value="↔"/>	Name 20 =UnguelstStimm91;
<input type="button" value="↔"/>	Name 21 =GuelstStimm91;
<input type="button" value="↔"/>	Name 22 =OEVP91;
<input type="button" value="↔"/>	Name 23 =SPOE91;
<input type="button" value="↔"/>	Name 24 =FPOE91;
<input type="button" value="↔"/>	Name 25 =VGOE91;
<input type="button" value="↔"/>	Name 26 =GAL91;

erzeuge weitere Namensfelder für Variable der vergangenen Wahl

Generierte Variable **Hilfe**

die nicht in den Dateien der beiden Wahlen enthalten sind

↔ **Name 75 =Nichtw97;**
 ↔ **Name 76 =Nichtw91;**



↔ **Name 77=Sonstige97;**
 ↔ **Name 78=Sonstige91;**

↔ **Name 79=Zuwanderer;** **Hilfe**
 ↔ **Name 80=Abwanderer;**


[...] erzeuge zusätzliche Namensfelder

Variablenamen in Datei speichern **Hilfe**

Eingabefeld leer = nicht speichern

↔  


Datei der aktuellen Wahl

 **\".\\Testdat\\LWahl97.fre\"**

frei Format der Daten **Hilfe**

↔ **v1:16** der Datensatz enthält diese Variablen
 Bei Format DIREKT schreiben Sie: alle_V

↔ **v1** **Hilfe**
 Verbindungsvariable aus der aktuellen Wahl
 zur vorausgegangenen Wahl
 z.B. die Gemeindenummer. Schreiben Sie V..

 **Wenn Dateiformat erweitert-FREI oder FIX** **Hilfe**

Parteien der aktuellen Wahl **Hilfe**

"echte" Parteien

↔ **OEVP97, SPOE97, FPOE97, GRUENE97, LIF97**
 ↔ **NEUTR97, KPOE97, CWG97, FDP97, OENP97**
 ↔


Pseudo-Parteien



↔ **UnguelstStimm97** Ungültig-Wähler
 ↔ **Nichtw97** Nichtwähler



Gesamtzahl der Wahlberechtigten bei der aktuellen Wahl




↔ **973750** wenn Gesamtzahl nicht bekannt ist
 dann Eingabefeld leer machen



Datei der vergangenen Wahl


 ".\Testdat\LWahl91.fre"



 frei Format der Daten 



  V17:26 der Datensatz enthält diese Variablen



  V17  Verbindungsvariable (z.B. Gemeindenummer)
aus der vergangenen Wahl (schreiben Sie V..)

 Wenn Dateiformat erweitert-FREI oder FIX 



Parteien der vergangenen Wahl 
"echte" Parteien



  OEVP91, SPOE91, FPOE91, VGOE91, GAL91


  |




  |




Pseudo-Parteien




  UnguelStimm91 Ungültig-Wähler




  Nichtw91 Nichtwähler

Zu- und Abwanderer miteinbeziehen 



   Zuwanderer Zuwanderer



   31478 Gesamtzahl der Zuwanderer in allen Gemeinden



   Abwanderer Abwanderer



   Zahl wird nicht benötigt


Nichtwähler, Ungültigwähler und "Sonstige" bilden
Umkodierungen und Variablenmanipulationen

  Nichtw91 = Wahlberecht91 - AbgegebenStimm91 ;

  Nichtw97 = Wahlberecht97 - AbgegebenStimm97 ;

  |

  |

 [...] erzeuge zusätzliche Eingabefelder

Option: Gruppen von Wahlgemeinden bilden

Option: Simulation einer Hochrechnung

Option: Zeige ausgewählte Wahlgemeinden

Gewichtung

Modell-Optionen

Option: Modellanpassung und Residuenanalyse

Wählerstromanalyse

↑↓ 1
 Wählerstromanalyse für ausgezählte Gemeinden
 0 = keine Wählerstromanalyse
 1 = Wählerstromanalyse durchführen

Grafik-Optionen für Wählerstrom-Analyse

Option: Mandatsverteilung ermitteln

Option: Einige Ergebnisse speichern

Option: Diagramm der Veränderung

Grafik-Optionen

Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix

Almo-Programmtext in Ergebnisliste zeigen

← ↓ zeige **Hilfe**
 zeige = Almo-Programmtext zeigen
 Eingabefeld leer = nicht zeigen

P51.4.4.1 Erläuterungen zu den Eingabeboxen von Prog51m1

Eingabebox 1: Speicher

Speicher fuer x Variable **Hilfe**

Vereinbare Variable= **80**

Die Zahl der vereinbarten Variablen muss mindestens so groß sein wie die höchste

im Maskenprogramm vorkommende Variablennummer. In unserem Fall ist dies 80. Das ist die Variablennummer der "Abwanderer". Der Benutzer kann auch eine höhere Zahl einsetzen.

Eingabebox 2: Namen der Parteien aus der aktuellen Wahl

zeige = Namensdatei in Output zeigen
 leer = nicht zeigen

in der Reihenfolge, in der sie in der Datei hintereinander stehen

Leere alle Eingabefelder dieser Sub-Box

<input type="button" value="↔"/>	Name 1 =GemeindeNr97;
<input type="button" value="↔"/>	Name 2 =Wahlberecht97;
<input type="button" value="↔"/>	Name 3 =AbgegebenStimm97;
<input type="button" value="↔"/>	Name 4 =UnguelstStimm97;
<input type="button" value="↔"/>	Name 5 =GuelstStimm97;
<input type="button" value="↔"/>	Name 6 =OEVP97;
<input type="button" value="↔"/>	Name 7 =SPOE97;
<input type="button" value="↔"/>	Name 8 =FPOE97;
<input type="button" value="↔"/>	Name 9 =GRUENE97;
<input type="button" value="↔"/>	Name 10 =LIF97;
<input type="button" value="↔"/>	Name 11 =NEUTR97;
<input type="button" value="↔"/>	Name 12 =KPOE97;
<input type="button" value="↔"/>	Name 13 =CWG97;
<input type="button" value="↔"/>	Name 14 =FDP97;
<input type="button" value="↔"/>	Name 15 =OENP97;
<input type="button" value="↔"/>	Name 16 =Zeit97;

erzeuge weitere Namensfelder für Variable der aktuellen Wahl

Zuerst erhalten die Variablen der aktuellen Wahl je einen Namen und zwar in der Reihenfolge in der die Variablen in der Datei aufeinanderfolgen. In unserem Fall ist die 1. Variable die Gemeindenummer, die 2. ist die Zahl der Wahlberechtigten, die 6. Variable ist die Stimmenzahl der ÖVP usw. bis zur 16. Variablen. Dies ist die Uhrzeit der Meldung des Wahlergebnisses an die Rechenzentrale. An alle Variablennamen haben wir die Zahl "97" angehängt um auszudrücken, dass es sich um die Wahl 1997 handelt.

Die Variablennamen aus der aktuellen plus der vergangenen Wahl können auch in eine separate Datei geschrieben werden. Man erspart es sich dann, die Namen in die Eingabefelder zu schreiben. Die Art und Weise wie Namen geschrieben werden ist wenig kompliziert. Klicken Sie auf die beiden Hilfeknöpfe in der Eingabebox. Almo erklärt Ihnen dann ausführlich, wie Variablennamen geschrieben werden.

Wenn Sie auf den Knopf mit dem doppelköpfigen Pfeil klicken,



dann wird das Eingabefeld gelöscht. Durch nochmaligem Klick wird der zuvor gelöschte Text wieder zurückgeholt.

Am Ende der Eingabebox ist der Knopf "Erzeuge weitere Namensfelder ...".



Wenn sie mehr als 16 Variable in der Datei der aktuellen Wahl haben, dann klicken Sie auf diesen Knopf. Es wird dann ein zusätzliches leeres Eingabefeld erzeugt, in das Sie den Variablennamen schreiben können.

Eingabebox 3: Namen der Parteien der vergangenen Wahl.

Namen der Variablen aus der vergangenen Wahl Hilfe

in der Reihenfolge, in der sie in der Datei hintereinander stehen
Numerierung muss an letzte Nummer aus aktueller Wahl anschliessen

↔	Name 17 =GemeindeNr91;
↔	Name 18 =Wahlberecht91;
↔	Name 19 =AbgegebenStimm91;
↔	Name 20 =UnguelStimm91;
↔	Name 21 =GuelStimm91;
↔	Name 22 =OEV91;
↔	Name 23 =SPOE91;
↔	Name 24 =FPOE91;
↔	Name 25 =VGOE91;
↔	Name 26 =GAL91;

... erzeuge weitere Namensfelder für Variable der vergangenen Wahl

Hier gilt die Erläuterung von Eingabebox 2 entsprechend. Beachten Sie aber, dass die Variablen-Numerierung hier mit 17 weiterläuft. Siehe dazu die Ausführung über "verknüpfte" Datensätze in P51.4.2.

Eingabebox 4: Namen für generierte Variable

Generierte Variable Hilfe

die nicht in den Dateien der beiden Wahlen enthalten sind

↔	Name 75 =Nichtw97;
↔	Name 76 =Nichtw91;
↔	Name 77=Sonstige97;
↔	Name 78=Sonstige91;
↔	Name 79=Zuwanderer;
↔	Name 80=Abwanderer;

... erzeuge zusätzliche Namensfelder Hilfe

Hier werden Namen für Variable geschrieben, die im Programm erst erzeugt werden, also nicht in der Datei der aktuellen Wahl und nicht in der Datei der vergangenen Wahl enthalten sind. In unserem Falle werden weiter unten im Programm in Eingabebox 11 die Nichtwähler der vergangenen und der aktuellen Wahl berechnet. Wir haben deswegen hier in Eingabebox 4 die Namen "Nichtw97" und "Nichtw91" eingeführt, denen wir die Variablennummern 75 und 76 zugewiesen haben. Diese Nummern sind an und für sich beliebig. Sie dürfen aber

selbstverständlich nicht im Bereich von 1 bis 26 liegen, da diese für die Variablen der Dateien der vergangenen und der aktuellen Wahl verwendet werden. Die Variablennummern müssen aber selbstverständlich innerhalb der in Eingabebox 1 vereinbarten Variablenzahl liegen.

Wir haben auch folgende zwei Namen eingeführt: "Sonstige97" und "Sonstige91", obwohl wir im nachfolgenden Programm keinen Gebrauch von ihnen machen. Es bestünde jedoch die Möglichkeit in Eingabebox 11 die sehr kleinen Parteien aus der Wahl 97 zu einer Partei "Sonstige97" zusammenzufassen z.B. durch die Anweisung





```
Sonstige97 = KPOE97 + CWG97 + FDP97 + OENP97;
```



Wir haben dann noch 2 weitere Namen eingeführt: "Zuwanderer" und "Abwanderer". Wir werden später, wenn wir Eingabebox 7 erläutern darauf zurückkommen.



Werden mehrere Parteien zu einer Partei zusammengefasst, dann ändern sich auch die Ergebnisse jener Parteien, die von der Zusammenfassung nicht betroffen sind. Die Übergangswahrscheinlichkeiten werden verändert. Diese Änderung sind in der Regel so minimal, dass sie vernachlässigt werden können.

Eingabebox 5 und 6: Datei der aktuellen Wahl

Datei der aktuellen Wahl

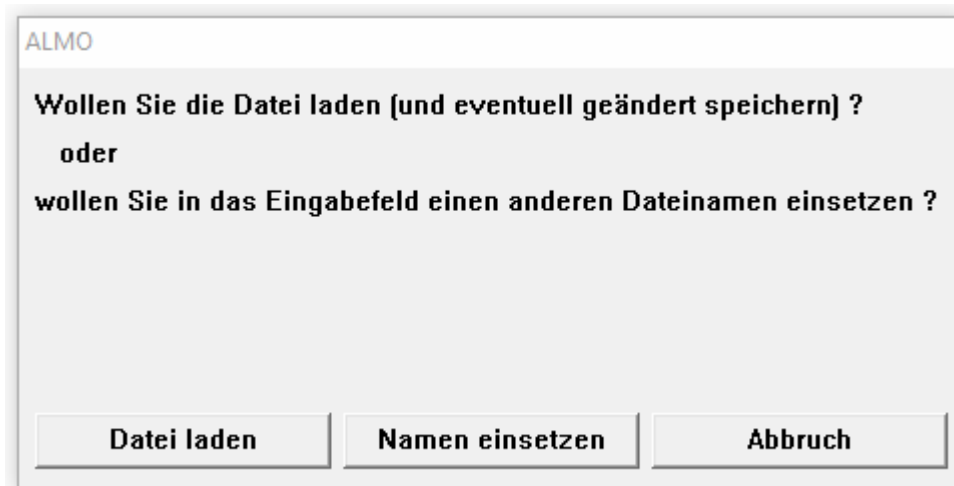
	".\Testdat\LWahl97.fre"	
	frei	Format der Daten  Hilfe
	V1:16	der Datensatz enthält diese Variablen Bei Format DIREKT schreiben Sie: alle_v

	V1	 Hilfe Verbindungsvariable aus der aktuellen Wahl zur vorausgegangenen Wahl z.B. die Gemeindenummer. Schreiben Sie V..
---	-----------	---

	Wenn Dateiformat erweitert-FREI oder FIX	 Hilfe
---	---	--

Eingabefeld 1: Name der Datei

Der Benutzer sollte sich die Datei der aktuellen Wahl anschauen. Nach Klick auf das Öffne-Symbol vor dem Eingabefeld mit dem Dateinamen präsentiert Almo folgende Dialogbox:



Dann Klick auf den Knopf "Datei laden". Almo zeigt dann die Datei in einem eigenen Fenster:

```

40101 136473 92046 2366 89680 27351 30249 19902 7409 3093 485 455 207 226 303 2336
40201 28829 21214 561 20653 5173 8084 4771 1598 668 157 131 71 0 0 1851
40301 38449 28111 654 27457 7957 8513 7445 2094 978 169 95 113 93 0 2032
40401 3587 2987 100 2887 1085 710 860 136 37 32 9 14 0 4 1644
40402 1686 1416 31 1385 827 241 235 40 7 11 3 13 0 8 1613
. . . . .
. . . . .

```

Die Daten sind durch ein oder mehrere Blanks getrennt. Eine Zeile enthält die Wahlergebnisse einer Gemeinde. Der 1. Wert ist die Gemeindenummer, der 2. die Zahl der Wahlberechtigten usw. wie in der Eingabebox für die Variablennamen definiert. Bei einer echten Hochrechnung würden die Ergebnisse einer Gemeinde nach der anderen in der Zentrale eintreffen. Die Datei würde so um eine Zeile nach der anderen wachsen. Almo rechnet mit den *gespeicherten* Daten. Wenn der Benutzer im Daten-Fenster Werte ändert, dann muss er speichern, wenn Almo die Änderung berücksichtigen soll. Wir empfehlen, das Datenfenster, nachdem man es angeschaut hat, wieder zu löschen.

Eingabefeld 2: Format der Daten

Almo unterscheidet 3 Formate, in denen die Daten vorliegen können. Wird in obiger Eingabebox auf den Hilfefknopf geklickt, dann erhält man eine ausführliche Erläuterung. Siehe auch das Almo-Dokument 32 "Datei-Operationen mit Almo", Abschnitt 1. Im folgenden Text werden nur sehr kurz die 3 Almo-Formate erklärt.

1. Format FREI

Beispiel: 1 2 23.5 Beamter 23 0 4

Die Werte eines Datensatzes sind durch Blank und/oder Komma oder das Tabulatorzeichen voneinander getrennt. Der Punkt dient als Dezimalpunkt. Das kann geändert werden.

Format FREI mit *Tabulator* als Trennzeichen.

Die Daten, die eingelesen werden sollen, wurden ursprünglich aus einem anderen System bzw. Programm mit dem Tabulator als Trennzeichen ausgegeben. Bei Excel haben sie dann üblicherweise die Namens-Erweiterung ".txt", bei SPSS und vielen anderen Statistiksystemen die Erweiterung ".dat". In einem solchen Fall muss in

der Eingabebox im Eingabefeld "Format" geschrieben werden

-2 und nicht **frei**

2. Format FIX

Die Werte stehen ohne Trennzeichen direkt hintereinander.

Beispiel:1223.5Beamter2304.....

Das fixe Format muss noch durch die FELD-Anweisung ergänzt werden. Das Format ist sehr fehleranfällig. Wir raten ab, es zu benutzen.

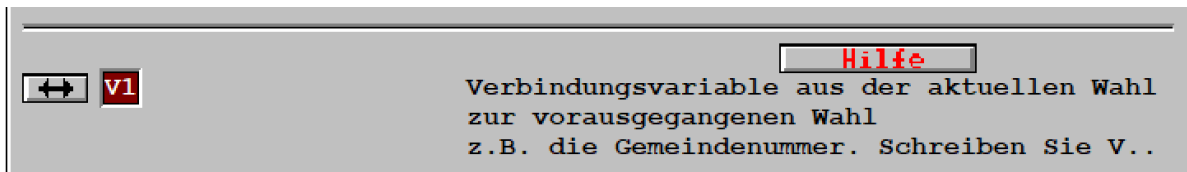
3. Format DIREKT

Dies ist ein Almo-spezifisches binäres Format. Es ist sehr schnell, erlaubt den DIREKT-Zugriff auf ausgewählte Datensätze (sogar Variable) und ermöglicht das platzsparende Speichern der Daten mit 1 bis 8 Byte (je nach Größe des Variablenwertes).

Eingabefeld 3: Die Daten werden diesen Variablen zugewiesen

Beim Format FREI und FIX werden die Werte des Datensatzes sukzessive eingelesen, so wie sie auf dem Speichermedium hintereinander stehen. Welche Variablen-Nummern dem jeweiligen Wert zugefügt werden, ist durch den Benutzer frei wählbar. Werden Variablennamen verwendet, so müssen deren Nummern natürlich den Variablennummern entsprechen. Wir verwenden für die Variablen der aktuellen Wahl die Nummern V1 bis 16 und für die der vergangenen Wahl die anschließenden Nummern V17 bis 26.

Eingabefeld 4: "Verbindungsvariable".



In P51.4.2 haben wir ausgeführt, dass die Daten aus der vergangenen und der aktuellen Wahl gemeindeweise verknüpft werden. Das geht jedoch nur, wenn in beiden Dateien eine Verbindungsvariable vorhanden ist. Normalerweise ist dies die Gemeindenummer. In unserem Falle ist dies V1. Schreiben Sie hier nicht den Namen sondern die Nummer mit dem "V" davor.

Eingabebox 7: Parteien der aktuellen Wahl

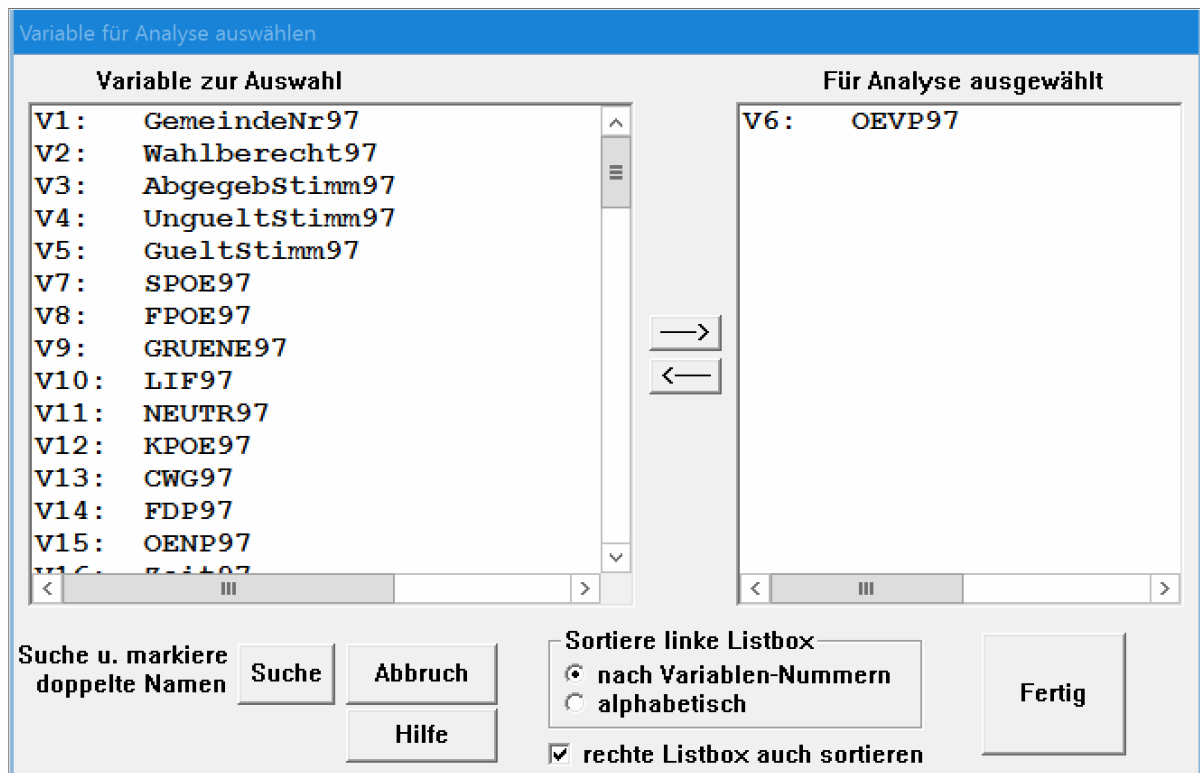
Parteien der aktuellen Wahl		Hilfe
"echte" Parteien		
↔	☐☐	OEVP97, SPOE97, FPOE97, GRUENE97, LIF97
↔	☐☐	NEUTR97, KPOE97, CWG97, FDP97, OENP97
↔	☐☐	
Pseudo-Parteien		
↔	UnguelstStimm97	Ungültig-Wähler
↔	Nichtw97	Nichtwähler
Gesamtzahl der Wahlberechtigten bei der aktuellen Wahl		
↔	973750	wenn Gesamtzahl nicht bekannt ist dann Eingabefeld leer machen

Sie haben hier 3 Eingabefelder, in die Sie die Parteienamen eintragen können. Sie können es sich aber auch bequem machen und nur die Variablen-Nummern eintragen; in unserem Falle wäre das V6:V15 (der Doppelpunkt heißt "bis"). Wenn Sie anstelle der sehr kleinen Parteien NEUTR97, KPOE97, CWG97, FDP97, OENP97 die Partei "Sonstige97" gebildet haben, dann müssen Sie im 2. Eingabefeld anstelle der dort angegebenen Parteien den Namen "Sonstige97" einsetzen.

Es gibt auch noch eine weitere Möglichkeit, die Parteien einzugeben. Vor dem Eingabefeld befindet sich ein 2. Knopf



Wenn Sie auf diesen klicken, dann öffnet Almo die Variablenauswahl-Box. Sie sehen folgendes



Klicken Sie auf eine Variable in der Listbox "Variable zur Auswahl". Dann klicken Sie auf den Pfeilknopf. Die Variable wird dann in die Listbox "Für Analyse ausgewählt" transportiert. Der "Transport" kann auch in der umgekehrten Richtung erfolgen.

Mit Doppelklick befördern Sie sofort einen Variablennamen von der einen Listbox in die andere.

Weitere Informationen zur Bedienung der Variablenauswahl-Box erhalten Sie, wenn Sie in der Eingabebox auf den Hilfe-Knopf klicken

"Echte" Parteien und Pseudo-Parteien

In der ersten Sub-Box werden die echten Parteien angegeben. Werden echte Parteien zusammengefasst, z.B. zur Partei "Sonstige", dann ist auch diese eine echte Partei und muss in die erste Subbox eingetragen werden. Also errechnet die Prozentverteilung für die echten Parteien - ohne die Pseudo-Parteien. In der zweiten Subbox werden die Pseudo-Parteien, wie Ungültigwähler, "Weißwähler", Nichtwähler usw. eingetragen.

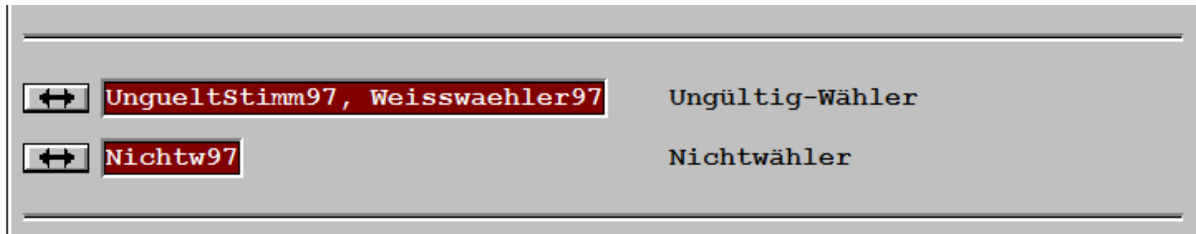
Die Nicht- und Ungültigwähler (und auch die "Abwanderer" aus Eingabebox 10) werden im Kalkül der Hochrechnung/Wählerstromanalyse mit den echten Parteien zur Gruppe der *abhängigen Variablen* zusammengefasst.

Im letzten Eingabefeld

wird die Gesamtzahl der Wahlberechtigten eingegeben. Also kann dann zum jeweiligen Hochrechnungszeitpunkt den "Auszahlungsgrad" mitteilen. Bleibt das Eingabefeld leer, dann entfällt diese Mitteilung.

Ungültig-Wähler und Weißwähler

Im mittleren Abschnitt der Eingabebox 7 werden die Ungültig-Wähler und die Nichtwähler eingetragen. Bei manchen Wahlen wird bei der Auszählung zwischen ungültig ausgefüllten Stimmzetteln und "weißen" Stimmzetteln unterschieden. Weißwähler sind Personen, die keine der auf dem Stimmzettel angebotenen Parteien präferieren und in gewisser Weise mit leerem Stimmzettel gegen die etablierten Parteien protestieren. In den mittleren Abschnitt der Eingabebox wird eingetragen:



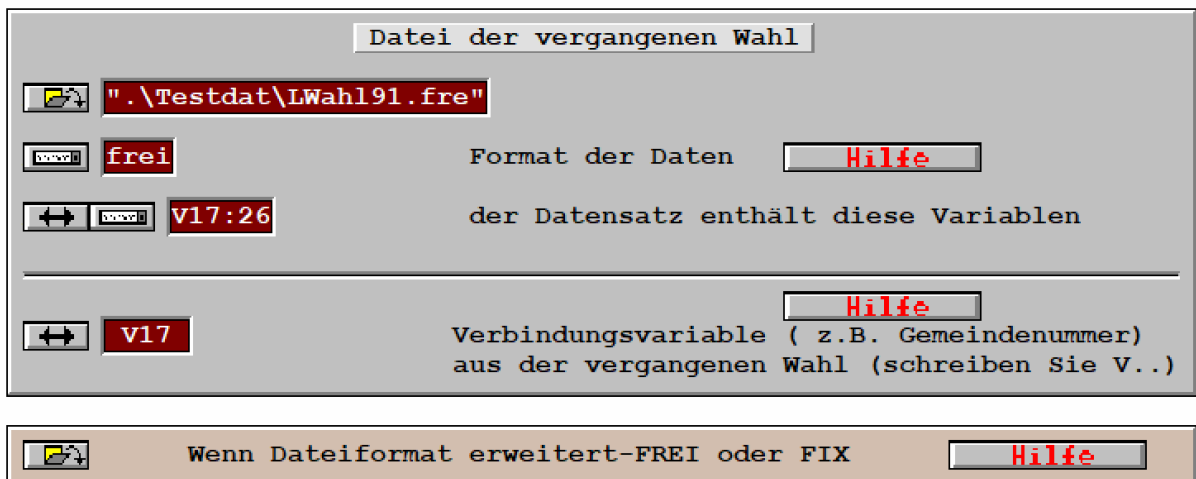
The screenshot shows a dialog box with two rows of input fields. The first row has a double-headed arrow icon, a text box containing "UnguelStimm97, Weisswaehler97", and the label "Ungültig-Wähler". The second row has a double-headed arrow icon, a text box containing "Nichtw97", and the label "Nichtwähler".

Voraussetzung ist natürlich, dass die Weißwähler in der Datei der aktuellen Wahl als Variable Nr. xx vorhanden sind. In der Eingabebox 2 für die Variablennamen werden sie dann so definiert

```
Name xx = Weisswaehler;
```

Wenn in der Datei der vergangenen Wahl die Weißwähler als eine Variable vorhanden sind, dann können auch diese in die Hochrechnung bzw. Wählerstromanalyse mit einbezogen werden.

Eingabebox 8: Datei der vergangenen Wahl



The screenshot shows a dialog box titled "Datei der vergangenen Wahl". It contains several input fields and labels: a file path ".\Testdat\LWahl91.fre", a format field "frei" with a "Hilfe" button, a variable list field "V17:26" with a "Hilfe" button, and a connection variable field "V17" with a "Hilfe" button. Below the dialog box is a separate bar with a file icon, the text "Wenn Dateiformat erweitert-FREI oder FIX", and a "Hilfe" button.

Den Variablen der Datei der vergangenen Wahl werden beim Einlesen die Variablennummern 17 bis 26 zugewiesen. In der Eingabebox muss nun noch angegeben werden, welche Nummer die Verbindungsvariable besitzt. Dies ist die Gemeindenummer. Schreiben Sie aber hier nicht den Namen sondern die Nummer, also V17. Also benötigt diese Verbindungsvariable, um während der Hochrechnung die beiden Dateien der aktuellen und der vergangenen Wahl gemeindeweise zu verknüpfen.

Eingabebox 9: Parteien der vergangenen Wahl

Parteien der vergangenen Wahl		Hilfe
"echte" Parteien		
↔	☐☐	OEVP91, SPOE91, FPÖ91, VGOE91, GAL91
↔	☐☐	
↔	☐☐	
Pseudo-Parteien		
↔	UnguelstStimm91	Ungültig-Wähler
↔	Nichtw91	Nichtwähler

"Echte" Parteien und Pseudo-Parteien

Die Eingabebox "Parteien der vergangenen Wahl" besteht aus 2 Abschnitten. Im oberen Abschnitt werden die echten Parteien angegeben, im unteren Abschnitt die Pseudo-Parteien der Ungültig- und Nicht-Wähler (und eventuell noch weitere, wie z.B. die "Weiß-Wähler", die absichtlich einen leeren Stimmzettel abgeben - siehe oben bei Eingabebox 7.

Die in den Eingabefeldern angegebenen Parteien plus die Nichtwähler plus die Ungültigwähler sollten die Gesamtzahl der Wahlberechtigten ergeben. Ohne die Nichtwähler müsste die Zahl der abgegebenen Stimmen entstehen.

Die Nicht- und Ungültigwähler (und auch die "Zuwanderer") werden im Kalkül der Hochrechnung/Wählerstromanalyse mit den echten Parteien zur Gruppe der *unabhängigen Variablen* zusammengefasst.

Die Eingabebox 9 ist ansonsten entsprechend Eingabebox 7 auszufüllen.

Eingabebox 10: Zu- und Abwanderer einbeziehen

Zu- und Abwanderer miteinbeziehen		Hilfe
↔	↓	Zuwanderer
	↔	31478
		Zuwanderer Gesamtzahl der Zuwanderer in allen Gemeinden
↔	↓	Abwanderer
		Abwanderer Zahl wird nicht benötigt

Der Benutzer muss entscheiden, ob er Zuwanderer und Abwanderer in die Analyse mit einbeziehen will. Er muss *beide* ein- oder ausschließen.

Sie können mit Klick auf den nach unten weisenden Pfeil den Namen "Abwanderer" und "Zuwanderer" einsetzen.

Wenn Sie hier "Zuwanderer" bzw. "Abwanderer" angeben, dann signalisieren Sie Also, dass Sie den Zuwanderungsgewinn bzw. den Abwanderungsverlust in der Wahlpopulation der Wahlgemeinden berücksichtigen wollen. Dadurch wird erreicht, dass die Gesamtzahl der Stimmen zu beiden Wahlzeitpunkten gleich groß

wird. Für unsere Beispieldaten ergaben sich bei *vollständiger Auszählung* diese Ergebnisse:

	Wahl 1991 Stimmen		Wahl 1997 Stimmen
-----		-----	
OEVP91	357771	OEVP97	328095
SPOE91	248640	SPOE97	207781
FPOE91	140311	FPOE97	158555
VGOE91	20603	GRUENE97	44435
GAL91	24250	LIF97	16021
UnguelStimm91	16315	NEUTR97	6247
Nichtw91	140355	KPOE97	2156
Zuwanderer	31478	CWG97	2957
	-----	FDP97	1454
zusammen	979723	OENP97	879
		UnguelStimm97	20622
		Nichtw97	184548
		Abwanderer	5973

		zusammen	979723

Die Gesamt-Stimmenzahl der vergangenen Wahl wird an die der aktuellen Wahl angepasst. Bei beiden Wahlen wurden je 979723 "Stimmen" vergeben. Dadurch wird es möglich bei der Wählerstromanalyse genau zu reproduzieren, wieviele Stimmen die Parteien erhielten und anzugeben wieviel Stimmen den jeweiligen Parteien zugeströmt sind (bzw. von ihnen abgeströmt sind).

In unserem Hochrechnungsmodell sind die beiden Variablen "Zuwanderer" und "Abwanderer" formal so definiert:

```
Zuwanderer in Gemeinde i =
    Wahlberechtigte in Gemeinde i bei aktueller Wahl
    minus
    Wahlberechtigte in Gemeinde i bei vergangener Wahl

Abwanderer in Gemeinde i =
    Wahlberechtigte in Gemeinde i bei vergangener Wahl
    minus
    Wahlberechtigte in Gemeinde i bei aktueller Wahl

Zuwanderer = Summe aller Zuwanderer aus allen Gemeinden
Abwanderer = Summe aller Abwanderer aus allen Gemeinden
```

Ergibt sich in der Gemeinde i für Zu- bzw. Abwanderer eine negative Zahl, so wird diese auf 0 gesetzt. Eine Gemeinde kann entweder nur Zuwanderung oder nur Abwanderung besitzen.

Zu- bzw. Abwanderer sind also nicht nur Personen die physisch einer Gemeinde zu- gewandert bzw. aus ihr abgewandert sind. Zu ihnen zählen auch Erstwähler und Verstorbene. Sie sind rein rechnerische Größen, die sich aus der Differenz der Wahlberechtigten der beiden Wahlen ergeben. Siehe auch Abschnitt P51.3.5.

Für die Wahl-Hochrechnung wird die Summe *aller* Zuwanderer gebraucht. Die Gesamtzahl der Abwanderer muss vom Benutzer nicht eingegeben werden. Also ermittelt sie selbst. Für die Wählerstromanalyse nach vollständiger Auszählung sind die folgenden Ausführungen, wie noch gezeigt werden wird, irrelevant.

Es werden folgende drei Möglichkeiten angeboten:

Möglichkeit 1: Hochrechnung mit Zu- und Abwanderer und Gesamtzahl der

Zuwanderer

Dies sollte der Normalfall sein. Mit Hilfe von Programm Prog51mb kann die Zuwandererzahl errechnet werden - vorausgesetzt die Zahl der Wahlberechtigten aus der vergangenen und der aktuellen Wahl ist bekannt und eine entsprechende (amtliche) Statistik steht dem Benutzer schon vor dem Wahlzeitpunkt zur Verfügung.

Möglichkeit 2: Hochrechnung ohne Zu- und Abwanderer

Die Eingabefelder in der Eingabebox werden leer gemacht. Die Übergangsmatrix, die das Almo-Programm ausgibt, ist gegenüber der aus der 1. Möglichkeit verschieden. Damit ändert sich auch das prognostizierte Hochrechnungsergebnis in absoluten Stimmenzahlen und in Prozentwerten. Die Unterschiede sind jedoch fast immer gering.

Möglichkeit 3: Hochrechnung mit Zu- und Abwanderer ohne Gesamtzahl der Zuwanderer

Zu- und Abwanderer werden mit eingeschlossen. Die Gesamtzahl der Zuwanderer kann jedoch nicht mitgeteilt werden, weil keine Statistik der Wahlberechtigten (insbesondere) der kommenden aktuellen Wahl verfügbar ist - oder einfach nur, weil es zu aufwendig ist, diese Zahl zu ermitteln. Das entsprechende Eingabefeld für die Gesamtzahl in der Eingabebox bleibt leer. Almo ermittelt in diesem Fall die Zuwandererzahl für die *ausgezählten* Wahlgemeinden und setzt diese als Ersatzwert ein.

Das hat folgend Auswirkung: Die Matrix der Übergangswahrscheinlichkeiten (die ja nur aus den ausgezählten Gemeinden errechnet wird) ist also dieselbe, wie bei Möglichkeit 1. Die prognostizierte prozentuelle Verteilung und die absoluten Stimmenzahlen sind jedoch anders als bei Möglichkeit 1.

Bei vollständiger Auszählung, wenn *alle* Wahlgemeinden ausgezählt sind, ist der Ersatzwert gleich dem tatsächlichen. Bei vollständiger Auszählung wird die Hochrechnung zur Wählerstromanalyse. Für diese ist es also irrelevant, ob die Zuwandererzahl vom Benutzer in das Eingabefeld eingesetzt wurde oder nicht. Das gilt nicht, wenn mit einer Stichprobe von Wahlgemeinden aus der aktuellen Wahl (also nicht *allen* Wahlgemeinden) hochgerechnet wurde.

Eine 4. Möglichkeit (mit der der Benutzer experimentieren kann) sei hier nur angedeutet: Es wird mit Zu- und Abwanderer gerechnet. Als Gesamtzahl wird 0 eingegeben.

Siehe zu dieser Thematik auch die ausführliche Darstellung in Abschnitt P51.4.6.

Wird eine Residuenanalyse gerechnet (Abschnitt 51.7), dann sollten Sie die Zu- und Abwanderer herauszunehmen.

Eingabebox 11: Nichtwähler etc. bilden. Umkodierungen und Variablenmanipulationen

Nichtwähler, Ungültigwähler und "Sonstige" bilden
Umkodierungen und Variablenmanipulationen

↔ Nichtw91 = Wahlberecht91 - AbgegebenStimm91 ;
↔ Nichtw97 = Wahlberecht97 - AbgegebenStimm97 ;

↔
↔

erzeuge zusätzliche Eingabefelder

Im 1. Eingabefeld müssen die Nichtwähler der vergangenen Wahl im zweiten die der aktuellen Wahl eingetragen werden.

Die Nichtwähler werden als Differenz der Wahlberechtigten minus der abgegebenen Stimmen ermittelt. Sollten die abgegebenen Stimmen nicht als Variable in der Datei enthalten sein, dann müssten alle Parteistimmen plus die ungültigen Stimmen addiert werden. In unserem Beispiel würde die Anweisung für die vergangene Wahl dann so lauten:

```
Nichtw91= Wahlberecht91- (OVP91+SPOE91+FPOE91+VGOE91+GAL91 +UnguelStimm91);
```

und für die aktuelle Wahl

```
Nichtw97= Wahlberecht97- (OEVP97+SPOE97+FPOE97+GRUENE97+LIF97+NEUTR97  
+KPOE97+CWG97+FDP97+OENP97 +UnguelStimm97);
```

Vergessen Sie nicht das Semikolon als Abschluss der Anweisung. Diese lange Anweisung muss in das Eingabefeld geschrieben werden. Es expandiert dabei nach rechts

↔ Nichtw97=Wahlberecht97- (OEVP97+SPOE97+FPOE97+GRUENE97+LIF97+NEUTR97+KPOE97+CWG97+FDP97+OENP97 +UnguelStimm97);

Sinnvoll ist es in diesem Fall die Variablennamen durch die Variablennummern zu ersetzen. Beide können auch gemischt werden. Die Anweisung lautet dann

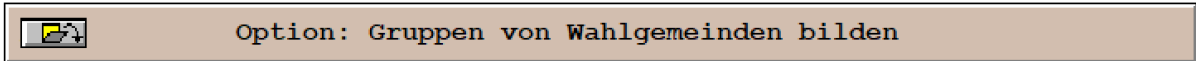
```
Nichtw97= Wahlberecht97- Summe (V6:15,4);
```

Im unteren Abschnitt der Eingabebox können beliebige Variablenmanipulationen eingesetzt werden. Insbesondere kann hier die "Sonstigen-Partei" gebildet werden. Wenn Sie z.B. die kleinen Parteien zusammenfassen wollen und eine Partei "Sonstige97" bilden wollen, dann müssen Sie hier eintragen

```
Sonstige97 = NEUTR97 + KPOE97 + FDP97 + OENP97;
```

Vergessen Sie nicht das Semikolon als Abschluss der Anweisung. Eine Anweisung kann über mehrere Eingabefelder gehen. Sie können aber auch ein Eingabefeld über den rechten Rand der Eingabebox hinaus ausdehnen.

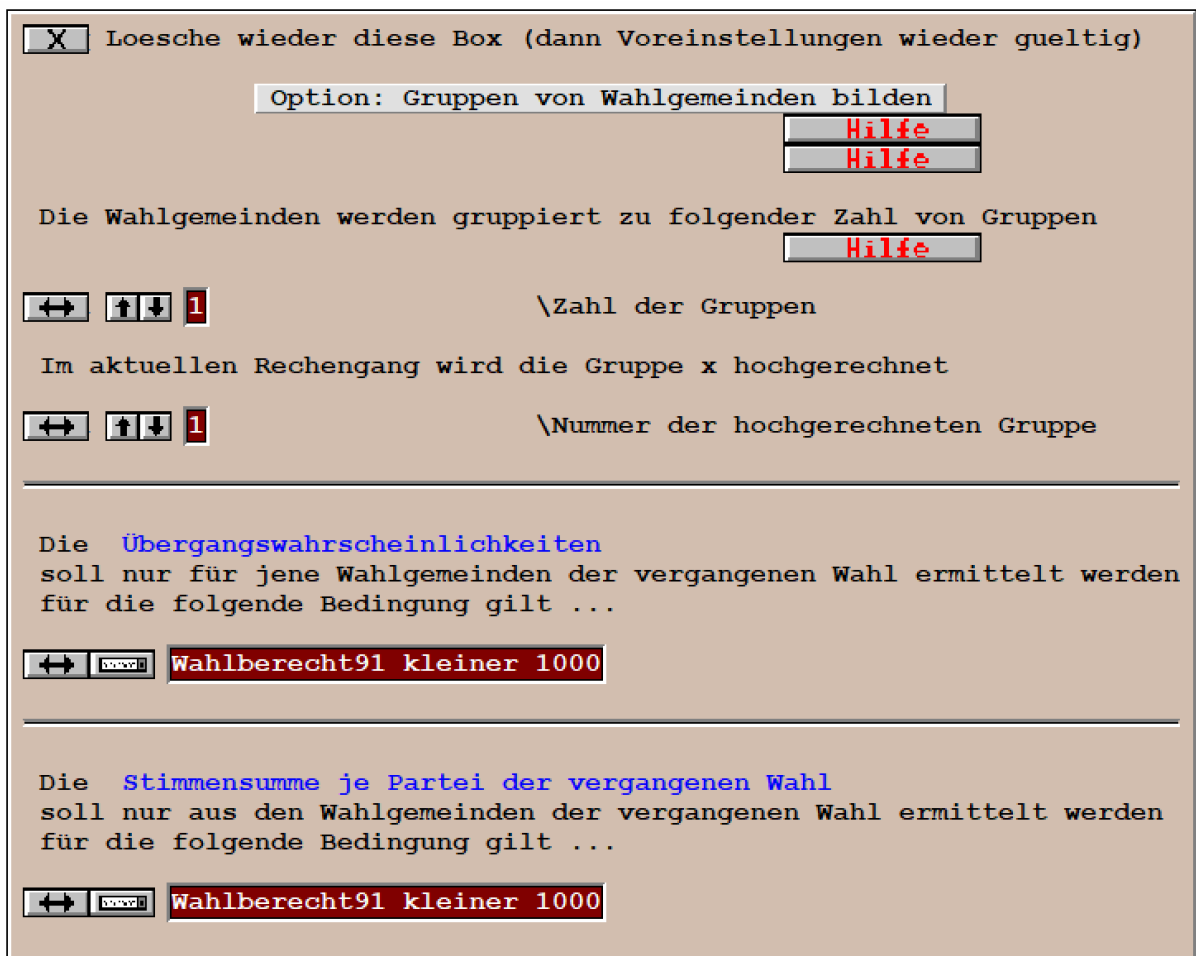
Eingabebox 12: Hochrechnung/Wählerstromanalyse mit Gruppen von Wahlgemeinden



Dem Benutzer werden 3 Möglichkeiten angeboten:

1. "Gruppierte" Hochrechnung. Die Wahlgemeinden werden zu mehreren (möglichst homogenen) Gruppen zusammengefasst. "Homogen" heißt, dass die Gemeinden einer Gruppe sehr ähnliche Übergangswahrscheinlichkeiten aufweisen. Für jede Gruppe wird eine eigene Hochrechnung bzw. Wählerstromanalyse gerechnet. Die Ergebnisse der einzelnen Gruppen werden zur Prognose für das gesamte Wahlgebiet summiert.
2. Hochrechnung für eine einzelne Gruppe von Wahlgemeinden (z.B. Hochrechnung nur für ein Bundesland oder z.B. nur für den ländlichen Raum des Wahlgebiets)
3. Hochrechnung einer einzelnen Gemeindegruppe, deren Übergangsmatrix verwendet wird um das Ergebnis für das gesamte Wahlgebiet zu prognostizieren. Eine eher ungewöhnliche Vorgehensweise.

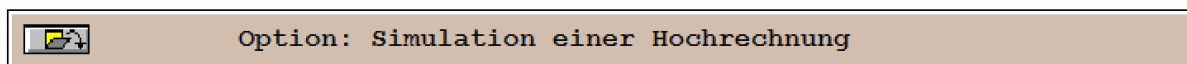
Wird die Optionsbox geöffnet, dann sieht man folgendes



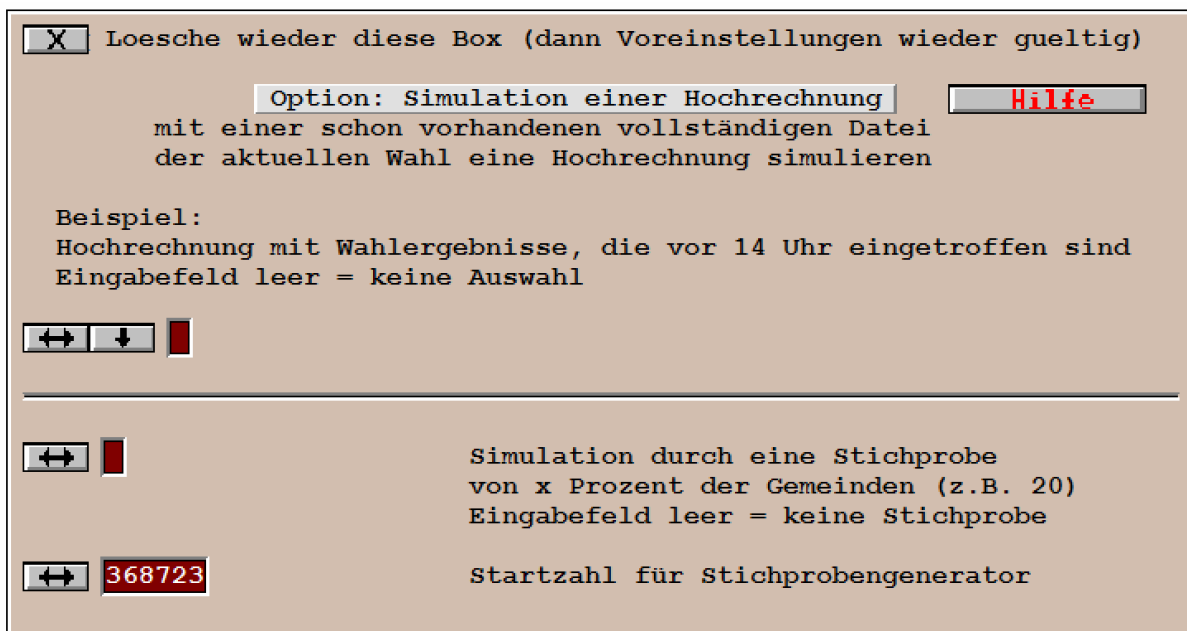
Die abgebildete Optionsbox ist für die 2. Möglichkeit, d.h. die Hochrechnung bzw. Wählerstromanalyse für eine einzelne Gemeindegruppe ausgefüllt. Es soll beispielsweise eine Wählerstromanalyse für den ländlichen Raum (definiert als Gemeinden mit bis 1000 Wahlberechtigten) gerechnet werden.

Wir werden diese Optionsbox und die 3 Möglichkeiten nicht hier sondern im später folgenden Abschnitt P51.4.8 "Hochrechnung/Wählerstromanalyse mit Gruppen von Gemeinden" ausführlich darstellen.

Eingabebox 13: Simulation einer Hochrechnung



Wird die Optionsbox geöffnet, dann sieht man folgendes:

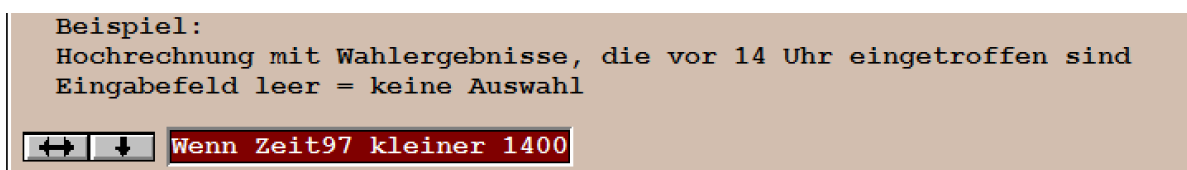


Sie können im Nachhinein eine Hochrechnung simulieren. Voraussetzung dafür ist, dass Sie die vollständigen Dateien zweier aufeinander folgender Wahlen besitzen.

1. Eingabefeld:

In unseren Beispieldaten der Landtagswahl in Oberösterreich 1997 ist in Variable16 der aktuellen Wahl die Zeit enthalten, zu der das Wahlergebnis in der Rechenzentrale eingetroffen ist. Bei kleinen ländlichen Gemeinden ist das nicht selten schon gegen 14 Uhr.

Wenn in der Datei der aktuellen Wahl eine Variable vorhanden ist, die die Uhrzeit enthält, dann können Sie ausprobieren, welches Ergebnis eine Hochrechnung z.B. um 18.00, um 18.30 usw. erbracht hätte. Klicken Sie auf den Knopf mit dem nach unten gerichteten Pfeil und füllen Sie dann das Eingabefeld so aus

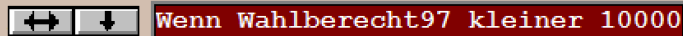


Beachten Sie, dass Almo keine Uhrzeitangabe mit Doppelpunkt, also z.B. "18:30", kennt. In der Datei muss die Variable als Ganzzahl oder als Dezimalzahl

geschrieben sein, also z.B. als "1830" oder als "18.30". In unseren Daten ist sie als Ganzzahl enthalten.

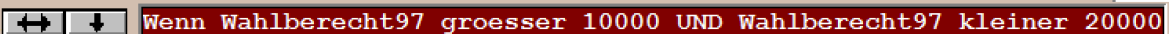
Ein weiteres Beispiel:

Da am Wahlabend üblicherweise zuerst die Ergebnisse aus kleineren Gemeinden eintreffen, kann man das in folgender Weise simulieren:



↔ ↓ Wenn Wahlberecht97 kleiner 10000

Generell gilt: In das Eingabefeld kann eine einfache "Wenn"-Anweisung in Almo-Syntax geschrieben werden, dabei ist auch eine "UND" bzw. "ODER"-Verknüpfung zulässig. Beispiel: Es soll eine Hochrechnung aus Gemeinden zwischen 10000 und 20000 Wahlberechtigten gerechnet werden. Die "Wenn"-Anweisung lautet dann



↔ ↓ Wenn Wahlberecht97 groesser 10000 UND Wahlberecht97 kleiner 20000

Die Wenn-Anweisung ist ausführlich beschrieben im Almo-Dokument Nr. 0 "Arbeiten mit Almo", Abschnitt P0.7.1

Die komplexeren Anweisungen der Form "Wenn..Dann..Sonst..EndeWenn" sind hier nicht zulässig.

2. Eingabefeld: Stichprobe

Wenn Sie hier beispielsweise "20" angeben, dann wird aus den vorhandenen Wahlgemeinden eine Stichprobe von 20% für die Hochrechnung verwendet.

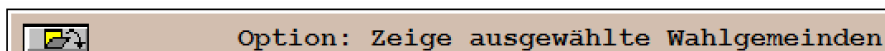
3. Eingabefeld: Startzahl für Zufallszahlen

Wenn Sie das Programm ein 2. Mal rechnen, dann wird genau dieselbe Stichprobe von Wahlgemeinden gezogen - es sei denn Sie ändern die Startzahl des Almo-internen Zufallsgenerators, z.B. auf 368725.

Weitere Simulationsmöglichkeiten

Da am Wahlabend üblicherweise zuerst die Ergebnisse aus kleineren Gemeinden eintreffen, kann man das in folgender Weise simulieren: Die Optionsbox "Ein- und Ausschliessen von Wahlgemeinden" wird geöffnet. Siehe dazu oben Eingabebox 12.

Eingabebox 14: Ausgewählte Wahlgemeinden zeigen



Option: Zeige ausgewählte Wahlgemeinden

Wird die Optionsbox geöffnet, dann sieht man folgendes:

Mann klickt in die eingeblendete Zeile "Wenn...". Also setzt diesen Text in das Eingabefeld ein. Man füllt dann etwa in folgender Weise aus:

```
Wenn Wahlberecht97 groesser 20000 Dann Gewicht0 = 1.25; Ende_Wenn
```

Möglich wäre es auch die Wahlberechtigtenzahl aus der vergangenen Wahl einzusetzen, um die zu gewichteten Gemeinden zu selektieren. Wir werden in Abschnitt P51.4.9 darauf eingehen. Die Anweisung würde dann lauten:

```
Wenn Wahlberecht91 groesser 20000 Dann Gewicht0 = 1.25; Ende_Wenn
```

Also benutzt die interne Systemvariable Gewicht0 um die Gewichtung vorzunehmen. Ausgezählte Gemeinden mit über 20 000 Wahlberechtigten (also Städte) werden dann mit dem Faktor 1.25 gewichtet. Wenn beispielsweise 10 ausgezählte Städte vorhanden sind, dann gehen diese in die Hochrechnung so ein wie wenn sie 12.5 Städte wären. Oder anders formuliert: Zu den 10 vorhandenen Städten werden 2.5 artifizielle Städte erzeugt, deren Partei-Stimmenzahlen dem Durchschnitt der 10 vorhandenen entspricht.

Im späteren Abschnitt P51.4.9 wird die "gewichtete" Hochrechnung ausführlich dargestellt.

Gewichtung durch eine Variable

Ein weiteres Beispiel für eine alternative Gewichtung ist folgendes: Für die Hochrechnung bei der deutschen Bundestagswahl müssen nicht notwendigerweise die rund 8000 Wahlgemeinden in die Hochrechnung einbezogen werden. Man kann auch eine Stichprobe von rund 400 Wahlgemeinden auswählen. Aus der amtlichen Statistik kennt man die Verteilung der Gemeinden auf Größenklassen. Nun kann es sein, dass in der Stichprobe beispielsweise Wahlgemeinden aus der Größenklasse 10000 bis 20000 Einwohner etwas unterrepräsentiert sind. Hier bietet es sich an, in die Datei der vergangenen Wahl eine zusätzliche Variable "Gewichtsfaktor" einzuschreiben. Sie wird in der Datei der vergangenen Wahl als letzte Variable nach hinten gestellt und erhält die Variablennummer 27. In der Eingabebox der Variablennamen wird hinzugefügt

```
Name 27 = Gewichtsfaktor;
```

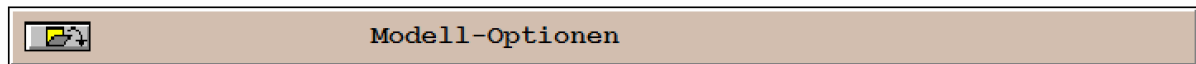
Diese Variable enthält für jede Wahlgemeinde eine solche Gewichtszahl, dass die Größenklassen-Verteilung exakt der tatsächlichen entspricht. Man kann auch für weitere Merkmale (z.B. für "katholische Gemeinde - evangelische Gemeinde) Gewichte einführen - wobei die Gewichtszahlen aus verschiedenen Merkmalen miteinander zu multiplizieren sind und somit nur die eine Variable "Gewichtsfaktor" bilden.

In der Eingabebox "Gewichtung" wird man im hier beschriebenen Falle in das Eingabefeld schreiben

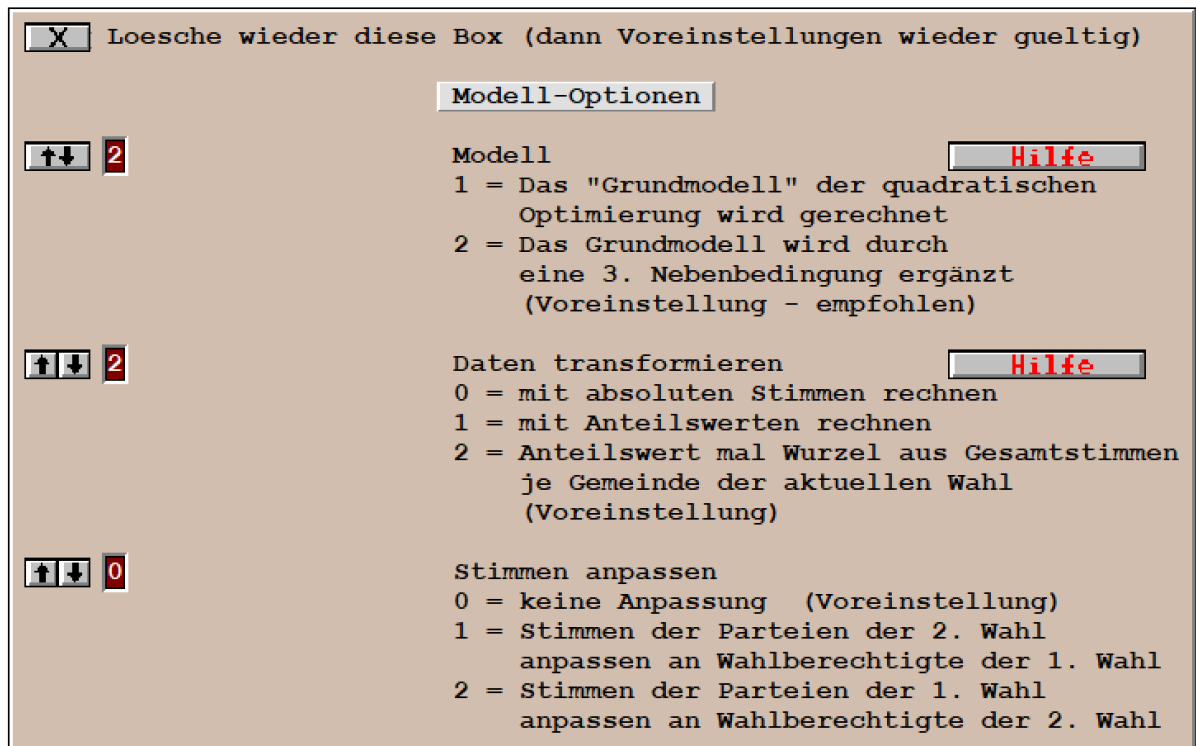
```
Gewicht0 = Gewichtsfaktor;
```

Siehe dazu im Almo-Dokument Nr. 0 "Arbeiten mit Almo", Abschnitt P0.7.2.1 und P0.7.2.2.

Eingabebox 16: Modell-Optionen



Wird die Optionsbox geöffnet, dann sieht man folgendes:



Eingabefeld 1: Modell

Wird "1" eingesetzt dann wird das "Grundmodell der quadratischen Optimierung" gerechnet. Der Kalkül erfüllt dann die in Abschnitt P51.1.3 beschriebenen Bedingungen (1) und (2).

Wird "2" eingesetzt, dann wird in das Grundmodell der Hochrechnung durch quadratische Optimierung die zusätzliche Bedingung (3) aus P51.1.3 eingeführt. Sie bewirkt folgendes: Wenn alle Gemeinden ausgezählt sind, dann stimmen die vom Modell prognostizierten Prozentwerte für die Parteien mit den tatsächlichen exakt überein. Siehe dazu die ausführliche Darstellung in P51.5. Werden auch noch Zu- und Abwanderer in die Analyse eingeführt, dann stimmen nicht nur die Prozentwerte sondern auch die absoluten Stimmen überein. Aber auch wenn die Auszählung noch unvollständig ist (wie das nun einmal bei der Hochrechnung der Fall ist), dann kommen die prognostizierten Prozentwerte den tatsächlichen sehr viel näher.

Werden bei Modell=1 die beiden Pseudo-Parteien der "Abwanderer" und der "Zuwanderer" in die Analyse miteinbezogen, dann stimmt bei vollständiger Auszählung die Summe aller prognostizierten Stimmen aus allen Gemeinden mit der tatsächlichen Stimmensumme der aktuellen Wahl überein. Da aber die prognostizierte Prozentverteilung nicht exakt ist, kann auch die prognostizierte Stimmzahl je Partei nicht exakt mit der tatsächlichen übereinstimmen.

Folgende Konstellationen können unterschieden werden:

Konstellation 1: Modell: 2 Zu- und Abwanderer: einbezogen

Folge bei vollständiger Auszählung:

- Die prognostizierte Prozentverteilung der Parteien und Pseudo-Parteien stimmt exakt mit der tatsächlichen überein
- Die prognostizierten Stimmzahlen der Parteien und Pseudo-Parteien stimmen exakt mit den tatsächlichen überein

Konstellation 2: Modell: 2 Zu- und Abwanderer: nicht einbezogen

Folge bei vollständiger Auszählung:

- a. Die prognostizierte Prozentverteilung der Parteien und Pseudo-Parteien stimmt exakt mit der tatsächlichen überein
- b. Die prognostizierten Stimmen für die Parteien und Pseudo-Parteien sind so berechnet, dass ihre Summe (aus allen Gemeinden) gleich der Gesamt-Stimmensumme aus der vergangenen Wahl ist.
- c. Die Relation der prognostizierten Stimmen der Parteien und Pseudo-Parteien stimmt exakt mit der tatsächlichen überein

Konstellation 3: Modell: 1 Zu- und Abwanderer: einbezogen

Folge bei vollständiger Auszählung:

- a. Die prognostizierte Prozentverteilung der Parteien und Pseudo-Parteien stimmt nicht exakt mit der tatsächlichen überein
- b. Die prognostizierten Stimmen für die Parteien und Pseudo-Parteien sind so berechnet, dass ihre Summe (aus allen Gemeinden) gleich der Gesamt-Stimmensumme aus der aktuellen Wahl ist.
- c. Die Relation der prognostizierten Stimmen der Parteien und Pseudo-Parteien stimmt nicht exakt mit der tatsächlichen überein

Konstellation 4: Modell: 1 Zu- und Abwanderer: nicht einbezogen

Folge bei vollständiger Auszählung:

- a. Die prognostizierte Prozentverteilung der Parteien und Pseudo-Parteien stimmt nicht exakt mit der tatsächlichen überein
- b. Die prognostizierten Stimmen für die Parteien und Pseudo-Parteien sind so berechnet, dass ihre Summe (aus allen Gemeinden) gleich der Gesamt-Stimmensumme aus der vergangenen Wahl ist.
- c. Die Relation der prognostizierten Stimmen der Parteien und Pseudo-Parteien stimmt nicht exakt mit der tatsächlichen überein

In der Optionsbox "Modelloptionen" kann der Benutzer auch entscheiden mit welcher der folgenden drei Datentransformationen gerechnet werden soll. Siehe die nachfolgende Beschreibung zu Eingabefeld 2. Hier sei nur kurz erläutert:

0 = keine Transformation. Es wird mit den Stimmzahlen der Parteien gerechnet.
Auswirkung: Große Wahlgemeinden beeinflussen die Übergangsmatrix stark.

1 = mit Anteilswert rechnen
Auswirkung: Bei der Berechnung der Übergangsmatrix bleibt die Größe der Wahlgemeinden unberücksichtigt.

2 = rechnen mit Anteilswert mal Wurzel aus Gesamtstimmzahl

Die oben beschriebenen Konstellationen gelten für Transformation 0 oder 2 nicht jedoch für Transformation 1. Bei dieser kann auch nicht mit Modell=2 eine Prozentverteilung prognostiziert werden, die bei vollständiger Auszählung mit der tatsächlichen exakt übereinstimmt.

Durch Transformation 0 oder 2 werden selbstverständlich unterschiedliche Übergangsmatrizen und dadurch auch unterschiedliche Wählerströme (in Stimmen) erzeugt. Bei Modell=2 prognostizieren jedoch trotzdem beide die oben beschriebene Konstellation 1 bzw. 2.

Eingabefeld 2: Daten transformieren

Die Daten können transformiert werden.

Wir verwenden folgende Notation:

$p(ik)$ = Anteilswert der Partei k in Wahlgemeinde i
 $v(ik)$ = Zahl der Stimmen der Partei k in Wahlgemeinde i - bei vergangener Wahl
 $a(ik)$ = Zahl der Stimmen der Partei k in Wahlgemeinde i - bei aktueller Wahl
 $V(i)$ = Gesamtstimmzahl in Wahlgemeinde i - bei vergangener Wahl
 $A(i)$ = Gesamtstimmzahl in Wahlgemeinde i - bei aktueller Wahl

Mit "Partei" sind alle echten Parteien und auch die "Quasi-Parteien" (z.B. Nichtwähler, Ungültigwähler, Zu- und Abwanderer etc.) gemeint

In das Eingabefeld kann eingegeben werden:

0 = keine Transformation. Es wird mit den Stimmzahlen gerechnet.
Auswirkung: Große Wahlgemeinden beeinflussen die Übergangsmatrix stark.

1 = Je Wahlgemeinde i wird für jede Partei k ihr Anteilswert $p(ik)$ gerechnet

transformierter Wert bei vergangener Wahl $p(ik) = v(ik) / V(i)$
transformierter Wert bei aktueller Wahl $p(ik) = a(ik) / A(i)$

Auswirkung: Bei der Berechnung der Übergangsmatrix bleibt die Größe der Wahlgemeinden unberücksichtigt.

2 = Je Wahlgemeinde i wird der Anteilswert $p(ik)$ einer Partei k multipliziert mit der Wurzel aus der Gesamtstimmzahl in Gemeinde i aus der aktuellen Wahl $A(i)$

transformierter Wert
bei vergangener Wahl = $\text{Wurzel}[A(i)] * p(ik)$
= $\text{Wurzel}[A(i)] * v(ik) / V(i)$

transformierter Wert
bei aktueller Wahl = $\text{Wurzel}[A(i)] * p(ik)$
= $\text{Wurzel}[A(i)] * a(ik) / A(i)$
= $a(ik) / \text{Wurzel}[A(i)]$

Zu beachten ist, dass auch bei der vergangenen Wahl mit der Wurzel der Gesamtstimmzahl A aus der aktuellen Wahl multipliziert wird. Damit wird erreicht, dass die Gesamtstimmzahlen der Gemeinden aus beiden Wahlen ungefähr aneinander angepasst wird - wobei die Stimmenrelationen erhalten bleiben. Der Einfluss der Größe der Wahlgemeinde auf die Übergangsmatrix liegt zwischen den Transformationen 0 und 1. Transformation 2 geht zurück auf einen Vorschlag von Bruckmann (G. Bruckmann, 1966, S.20), der hier modifiziert wurde.

Die eindeutig besten Ergebnisse erzielen Sie, wenn Sie "2" eintragen.

Eingabefeld 3: Stimmen anpassen

Wenn Sie "1" oder "2" einsetzen, dann wird die Zahl der Wahlberechtigten aus vergangener und aktueller Wahl rechnerisch gleichgemacht. Wir raten eher ab. Setzen Sie "0" ein.

Durch die Anpassung soll die Zahl der Wahlberechtigten in den beiden Wahlen gleich gemacht werden. Genauer formuliert: Die Stimmensumme der Parteien und Pseudo-Parteien, die im Programm als unabhängige Variable eingesetzt werden, soll gleich derjenigen der abhängigen Variablen werden. Das wird dadurch erreicht, dass die Stimmzahl der Parteien und Pseudo-Parteien in der Relation der Wahlberechtigten der 1. zur 2. Wahl erhöht bzw. verringert wird. Die Stimmen-Anpassung liefert auch bei voller Auszählung kein exakt richtiges Ergebnis. Das Verfahren ist eher nicht zu empfehlen.

Betrachten wir ein Beispiel:

Für die Wahlgemeinde xx wurden folgende Ergebnisse für die vergangene und die aktuelle Wahl gefunden

	Ergebnis 1. Wahl		

Gemeinde xx	Partei_A1	Partei_B1	Wahlberecht1
	40	80	160
	Ergebnis 2. Wahl		

Gemeinde xx	Partei_A2	Partei_B2	Partei_C2 Wahlberecht2

Wahlberecht1 = Stimmensumme aller Parteien plus Pseudo-Parteien (aller unabhängigen Variablen) aus vergangener Wahl in Gemeinde xx
 Wahlberecht2 = Stimmensumme aller Parteien plus Pseudo-Parteien (aller abhängigen Variablen) aus aktueller Wahl in Gemeinde xx

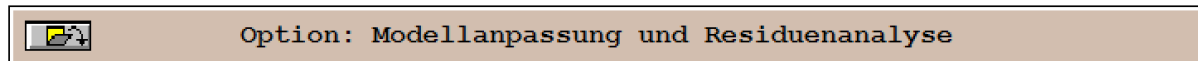
Die Stimmen der Partei_A1 aus der vergangenen Wahl soll angepasst werden an die Stimmenzahl der Wahlberechtigte der 2. Wahl

Partei_A1_angepasst = Partei_A1 * Wahlberecht2 / Wahlberecht1 = 40*180/160 = 45

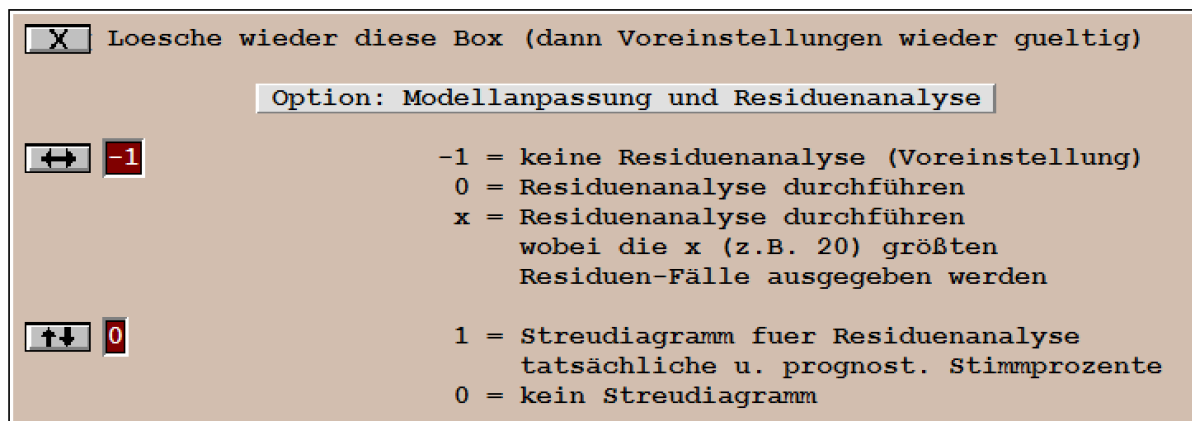
Stimmensumme der Partei_A2 aus aktueller Wahl anpassen an Wahlberechtigte der 1. Wahl

Partei_A2_angepasst = Partei_A2 *Wahlberecht1 / Wahlberecht2 = 50*160/180 = 44.4444

Eingabebox 17: Modellanpassung und Residuenanalyse



Wird die Optionsbox geöffnet, dann sieht man folgendes:



In P51.6 und P51.7 ist die Residuenanalyse ausführlich beschreiben.

Während einer laufenden Hochrechnung wird man keine Residuenanalyse durchführen. Es wäre aber zulässig. Setzen Sie also in das Eingabefeld "-1".

Wenn Sie jedoch mit unserem Hochrechnungsprogramm experimentieren, also simulierte Hochrechnungen durchführen, dann spricht nichts dagegen, eine Residuenanalyse zu rechnen. Setzen Sie in das 1.Eingabefeld "0", dann wird eine Residuenanalyse gerechnet. Setzen Sie eine Zahl größer 0, z.B. "5" dann wird eine Residuenanalyse gerechnet, wobei Ihnen die 5 Gemeinden ausgegeben werden, bei denen die Differenz zwischen prognostiziertem Wahlergebnis und tatsächlichem Ergebnis am größten ist.

Almo liefert einen sehr umfangreichen Output. Wir haben diesen Output mit unseren Beispieldaten dadurch erzeugt, dass wir in der Eingabebox "Simulation" gesetzt haben "Wenn Zeit97 kleiner 1400". Wir zeigen hier nur einen kurzen Auszug:

Variable

R-Quadrat

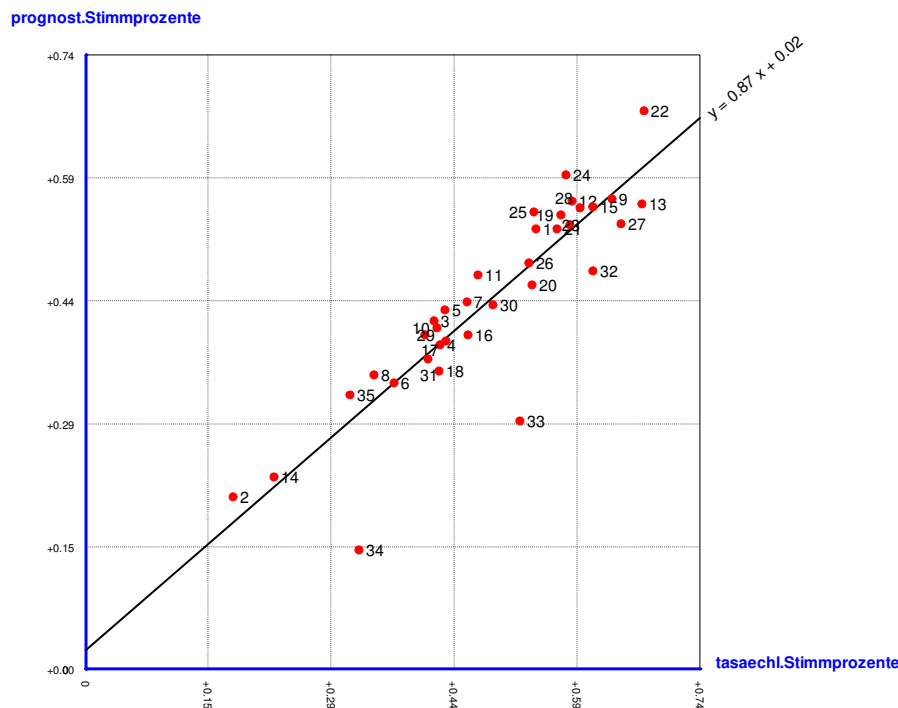
V6	OEVP97	0.99971919
V7	SPOE97	0.99986028
V8	FPOE97	0.99992946
V9	GRUENE97	0.99901297
V10	LIF97	0.99336317
V11	NEUTR97	0.99323621
V12	KPOE97	0.96084554
V13	CWG97	0.99707572
V14	FDP97	0.98501360
V15	OENP97	0.94680534
V4	UnguelStimm	0.99821540
V75	Nichtw97	0.99974308
V80	Abwanderer	0.91369561

Nachdem 35 von 445 Gemeinden ausgezählt sind, korrelieren die Stimmergebnisse dieser 35 Gemeinden für die ÖVP mit dem vom Hochrechnungs-Modell für diese 35 Gemeinden prognostizierten Stimmergebnissen mit $r^2=0.99971919$.

Anders ausgedrückt: Die vom Hochrechnungsverfahren errechnete Übergangsmatrix erzeugt für die 35 Gemeinden ein Stimmenergebnis, das mit dem tatsächlichen hervorragend übereinstimmt.

Wird im 2. Eingabefeld "1" eingesetzt, dann erzeugt Almo für jede Partei ein Streudiagramm. Wir zeigen hier das für die ÖVP. Jeder Punkt bildet eine Wahlgemeinde ab - mit dem tatsächlichen Prozentergebnis (auf der x-Achse) und dem vom Modell prognostizierten (auf der y-Achse).

Streudiagramm



Wenn der Benutzer wissen möchte, welche Wahlgemeinde sich hinter dem jeweiligen roten Punkt verbirgt, dann muss er die Eingabebox 14 "Ausgewählte Wahlgemeinden zeigen" öffnen und dort das Wort "zeige" einschreiben. Almo liefert dann folgende Ausgabe (gekürzt)

Gemeinden, die in Auswahl aufgenommen wurden:

- 1 40403
- 2 40404

```

3      40408
.      .
.      .
33     41801
34     41802
35     41809

```

Der im Diagramm weit links stehende Punkt 2 entspricht also der Wahlgemeinde mit dem Code 40404.

Almo kann die Streudiagramme nur für die ersten 8 Parteien ausgeben - entsprechend der Folge, mit der die Parteien in der Eingabebox 7 "Parteien der aktuellen Wahl" hintereinander stehen. Sind mehr als 8 Parteien vorhanden und möchte der Benutzer alle Streudiagramme sehen, dann muss er eine zweite Analyse mit umgestellter Reihenfolge rechnen.

Durch obiges Streudiagramm wird das Prinzip grafisch veranschaulicht, dem das Almo-Hochrechnungsprogramm folgt: Die Übergangsmatrix wird so gewählt, dass die (quadrierten) Differenzen zwischen den prognostizierten und den tatsächlichen Stimmanteilen der Parteien sich zu einem Minimum summieren. Dies ist das wohlbekannte Prinzip der Kleinste-Quadrate-Schätzung.

Eingabebox 18: Wählerstromanalyse

Wählerstromanalyse

Wählerstromanalyse für ausgezählte Gemeinden
0 = keine Wählerstromanalyse
1 = Wählerstromanalyse durchführen

Hochrechnung und Wählerstromanalyse werden in Prog51m1 nicht unterschieden. Demzufolge kann der Benutzer bei jedem Auszählungsgrad eine Wählerstromanalyse anfordern.

Wir rechnen eine Wählerstromanalyse bei unvollständiger Auszählung mit 35 Wahl-Gemeinden. Das wesentliche Ergebnis aus der Wählerstromanalyse ist die Übergangsmatrix der prozentuellen und der absoluten Abströme in Zahl der Stimmen. Wird in der Eingabebox eine 0 eingesetzt, dann wird nur nachfolgende Übergangsmatrix ausgegeben.

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Abstroeme errechnet aus 35 ausgezaehlten Gemeinden (die Summe der Werte in einer Spalte ist 100 %)

	OEVP91	SPOE91	FPOE91	VGÖE91	GAL91	Ungue1t91	Nichtw91	Zuwander
OEVP97	84.52	0	0	0	0	78.58	0	27.35
SPOE97	0	75.96	0	95.80	0	20.53	0	20.23
FPOE97	4.69	2.43	85.42	0	0	0	0	46.84
GRUENE97	2.12	5.27	0.00	0	75.13	0	0	5.57
LIF97	0.40	2.50	3.21	4.20	0	0	0	0
NEUTR97	1.38	0	0.16	0	0	0	0	0
KPOE97	0.44	0	0	0	0	0	0	0
CWG97	0.81	0	0	0	0	0.89	0	0
FDP97	0.15	0	0	0	0	0	0	0
OENP97	0.29	0	0	0	0	0	0	0
Ungue1tS	2.91	0	4.89	0	0	0	0	0
Nichtw97	1.45	13.84	4.72	0	24.87	0	100.00	0
Abwander	0.86	0	1.59	0	0	0	0	0
Gesamt	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Wird in der Eingabebox eine 1 eingesetzt, dann wird zusätzlich noch die nachfolgende Matrix der Abströme in Zahl der Stimmen ausgegeben.

prognostizierte Waehlerstroeme - Abstroeme in Zahl der Stimmen

die Gesamt-Stimmenzahl in der letzten Zeile ergibt die tatsaechliche
Stimmenzahl in der 1. Wahl
die Gesamt-Stimmenzahl in der letzten Spalte ergibt die prognostizierte
Stimmenzahl in der 2. Wahl

prognostiziert auf der Basis von 35 ausgezaehlten Gemeinden

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguel91	Nichtw91	Zuwander	Gesamt
OEVP97	302406	0	0	0	0	12820	0	8611	323837
SPOE97	0	188879	0	19738	0	3350	0	6369	218336
FPOE97	16770	6035	119856	0	0	0	0	14745	157405
GRUENE97	7572	13094	7	0	18220	0	0	1753	40646
LIF97	1420	6220	4505	865	0	0	0	0	13010
NEUTR97	4924	0	222	0	0	0	0	0	5146
KPOE97	1562	0	0	0	0	0	0	0	1562
CWG97	2890	0	0	0	0	145	0	0	3035
FDP97	542	0	0	0	0	0	0	0	542
OENP97	1020	0	0	0	0	0	0	0	1020
Unguel97	10417	0	6865	0	0	0	0	0	17282
Nichtw97	5177	34412	6624	0	6030	0	140355	0	192598
Abwander	3072	0	2233	0	0	0	0	0	5306
Gesamt	357771	248640	140311	20603	24250	16315	140355	31478	979725

Sind noch nicht alle Gemeinden ausgezählt, dann haben wir mit obiger Tabelle das mögliche Ergebnis einer Wählerstromanalyse *prognostiziert*.

Betrachten wir die OEVP91. Aus der 1. Übergangsmatrix ergibt sich, dass 84.52% ihrer Wähler, das sind (wie aus der 2. Übergangsmatrix ersichtlich wird) 302406, wieder zur OEVP97 gegangen. Zur SPOE97 ist niemand übergelaufen, aber 4.69%, das sind 16770 Wähler, sind zur FPOE97 abgewandert usw.

Almo macht diese Wählerwanderungen auch grafisch sichtbar, wenn der Benutzer die nachfolgende Optionsbox öffnet. Wir werden das in Abschnitt P51.4.7.1 zeigen.

Wurde in der Eingabebox 18 "Wählerstromanalyse" eine 1 eingesetzt, dann können auch die *Zuströme* zu den Parteien der aktuellen Wahl sichtbar gemacht werden. Die Optionsbox 16 "Modell-Optionen" muss dabei geschlossen bleiben. Almo rechnet dann mit Modell=2.

Wir zeigen hier die Übergangsmatrix der prozentuellen Zuströme

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Zuströme
errechnet aus 35 ausgezaehlten Gemeinden
(die Summe der Werte in einer Zeile ist 100 %)

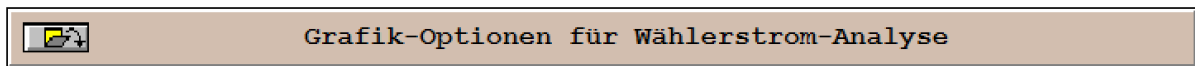
	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguel91	Nichtw91	Zuwander	Gesamt
	V22	V23	V24	V25	V26	V20	V76	V79	V151
OEVP97	93.16	0	0	0	0	4.61	0	2.23	100.00
SPOE97	0	86.63	0	8.23	0	2.17	0	2.97	100.00
FPOE97	9.29	2.76	81.08	0	0	0	0	6.87	100.00
GRUENE97	18.71	26.75	0.02	0	50.88	0	0	3.64	100.00
LIF97	11.17	40.44	43.28	5.12	0	0	0	0	100.00
NEUTR97	94.79	0	5.21	0	0	0	0	0	100.00
KPOE97	100.00	0	0	0	0	0	0	0	100.00
CWG97	94.46	0	0	0	0	5.54	0	0	100.00
FDP97	100.00	0	0	0	0	0	0	0	100.00
OENP97	100.00	0	0	0	0	0	0	0	100.00
Unguel97	55.40	0	44.60	0	0	0	0	0	100.00
Nichtw97	2.90	15.96	4.54	0	3.82	0	72.77	0	100.00

Abwanderer	52.97	0	47.03	0	0	0	0	0	100.00
------------	-------	---	-------	---	---	---	---	---	--------

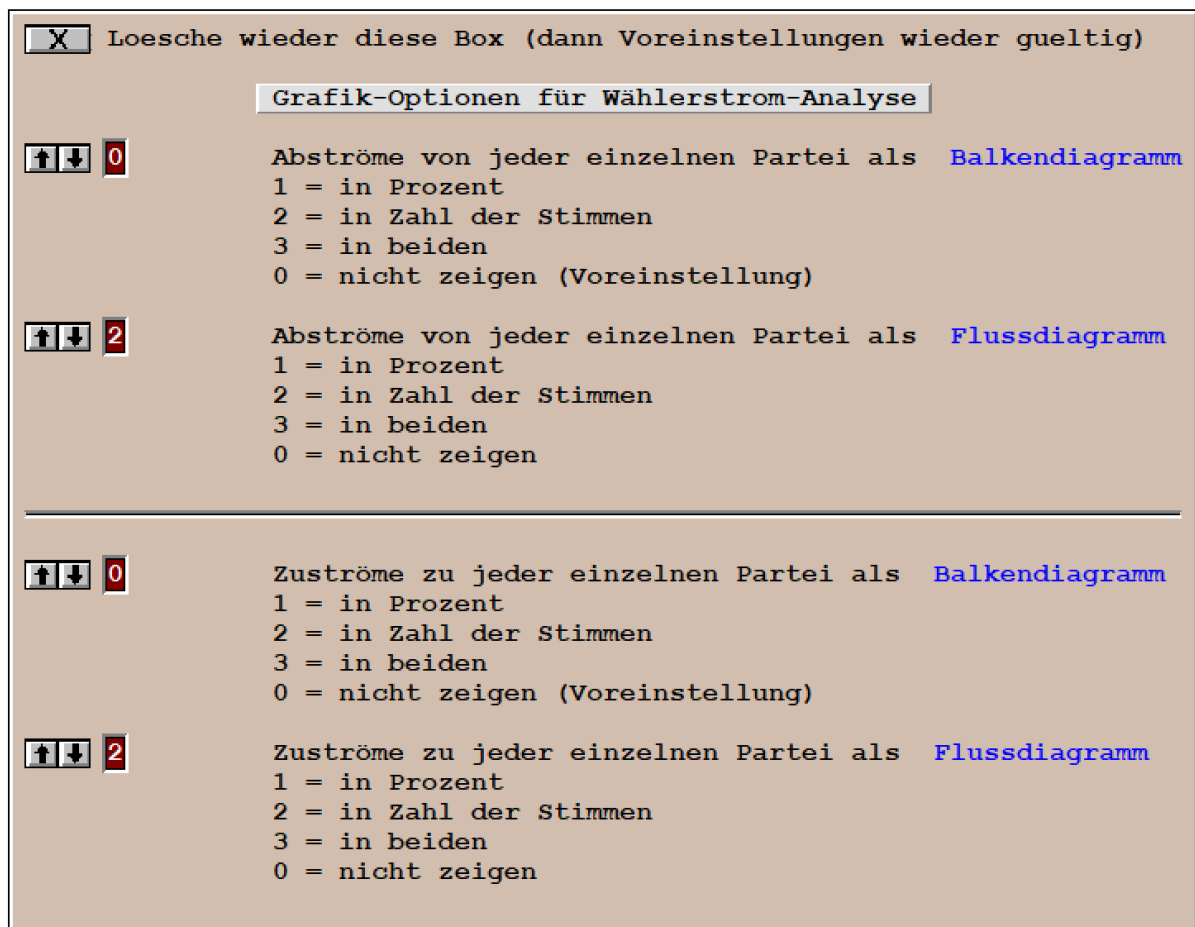
Die Übergangsmatrix der Zuströme ergibt sich sehr einfach dadurch, dass die Abströme-Matrix in Zahl der Stimmen in jeder Zeile von links nach rechts prozentuiert wird.

Die Zuströme-Matrix in Zahl der Stimmen ist bei Modell 2 identisch mit der Abströme-Matrix in Stimmen.

Eingabebox 19: Grafik-Optionen für die Wählerstromanalyse

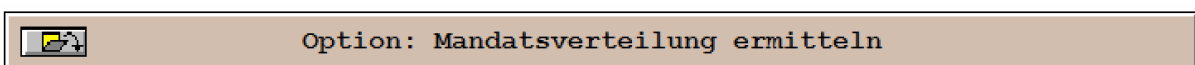


Wird die Optionsbox geöffnet, dann sieht man folgendes:



Die Optionsbox wird in Abschnitt P51.4.7 ausführlich erläutert. Dort werden auch die jeweils entstehenden Grafiken abgebildet

Eingabebox 20: Mandatsverteilung



Wird die Optionsbox geöffnet, dann sieht man folgendes:

X
Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)

Option: Mandatsverteilung ermitteln

Hilfe

↑ ↓
1

Mandatsverteilung ermitteln

0 = keine Mandatsverteilung (Voreinstellung)

1 = Mandatsverteilung nach Hare/Niemeyer

2 = Mandatsverteilung nach d'Hondt

↔
5.0

Sperrprozent für Mandatsgewinn

↔
0

Zahl der zu vergebenden Mandate

↔
LIF97

Parteien, die unter Sperrprozent bleiben
aber Direktmandate erhalten

Wenn Sie "0" eingeben, dann wird keine Mandatsverteilung ermittelt.

Geben Sie "1" ein, dann wird die Mandatsverteilung nach dem System von **Hare/Niemeyer** ermittelt. Dabei gelten folgende Voraussetzungen: (1) Jeder Wähler hat nur 1 Stimme, (2) alle Mandate sind Listenmandate, (3) es werden nur Parteien berücksichtigt, die auf oder über x% liegen oder Direktmandate (vermutlich) erhalten, (4) wenn eine Partei mehr als 50% besitzt aber nicht die Mandatsmehrheit erhalten würde, dann wird ihr ein Restmandat zugewiesen. Diese Art der Mandatsermittlung gilt für das 1. Zuteilungsverfahren der deutschen Bundestagswahl und für einige Bundesländer.

Wenn Sie "2" eingeben, dann wird das Verfahren nach **d'Hondt** durchgeführt. Dabei gilt wie bei Hare/Niemeyer obiger Satz (1) bis (3). Satz (4) gilt beim d'Hondt'schen Verfahren nicht.

Da das Verfahren der Mandatsermittlung von Land zu Land verschieden ist, muss Algo jeweils angepasst werden. In Algo ist gegenwärtig vorhanden, wie beschrieben, das Hare/Niemeyer-Verfahren, das d'Hondt'sche Verfahren und das mehrstufige Verfahren, das bei der deutschen Bundestagswahl angewendet wird (nach dem Wahlgesetz bis 2011). Die Reform der Sitzverteilung von 2011, 2013 und später ist im Programm nicht berücksichtigt. Siehe oben Abschnitt P51.3.8. Siehe auch hierzu die Programm-Maske Prog52m1.Msk.

Im 2. Eingabefeld geben Sie die "Sperrprozent" an, d.h. Sie geben an wieviel Prozent der Stimmen eine Partei erhalten muss, damit Sie in die Mandatsverteilung miteinbezogen wird.

Im 3. Eingabefeld ist dann noch anzugeben, wieviele Mandate insgesamt zu vergeben sind und im 4. Eingabefeld welche Parteien, die zwar unter x% bleiben aber Direktmandate erhalten. Da man dies zum Zeitpunkt der Hochrechnung noch nicht genau weiß, ist man hier auf Vermutungen angewiesen. Wir geben 200 als Zahl der zu vergebenden Mandate ein und nennen LIF97 als Partei, die vermutlich Direktmandate erhält. Wenn Sie weitere Parteien in das Eingabefeld schreiben wollen, dann müssen diese durch Beistrich getrennt werden. Wir erhalten folgenden Output.

Prognostizierte Mandate fuer Parteien aus 2. Wahl
(prognostiziert auf der Basis von 35 ausgezählten Gemeinden)

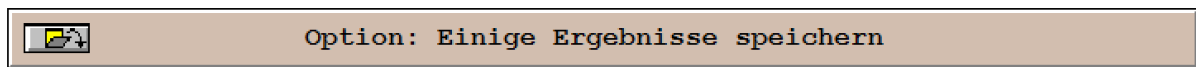
ganzzahlige Restmandate Mandatszähl

Mandate

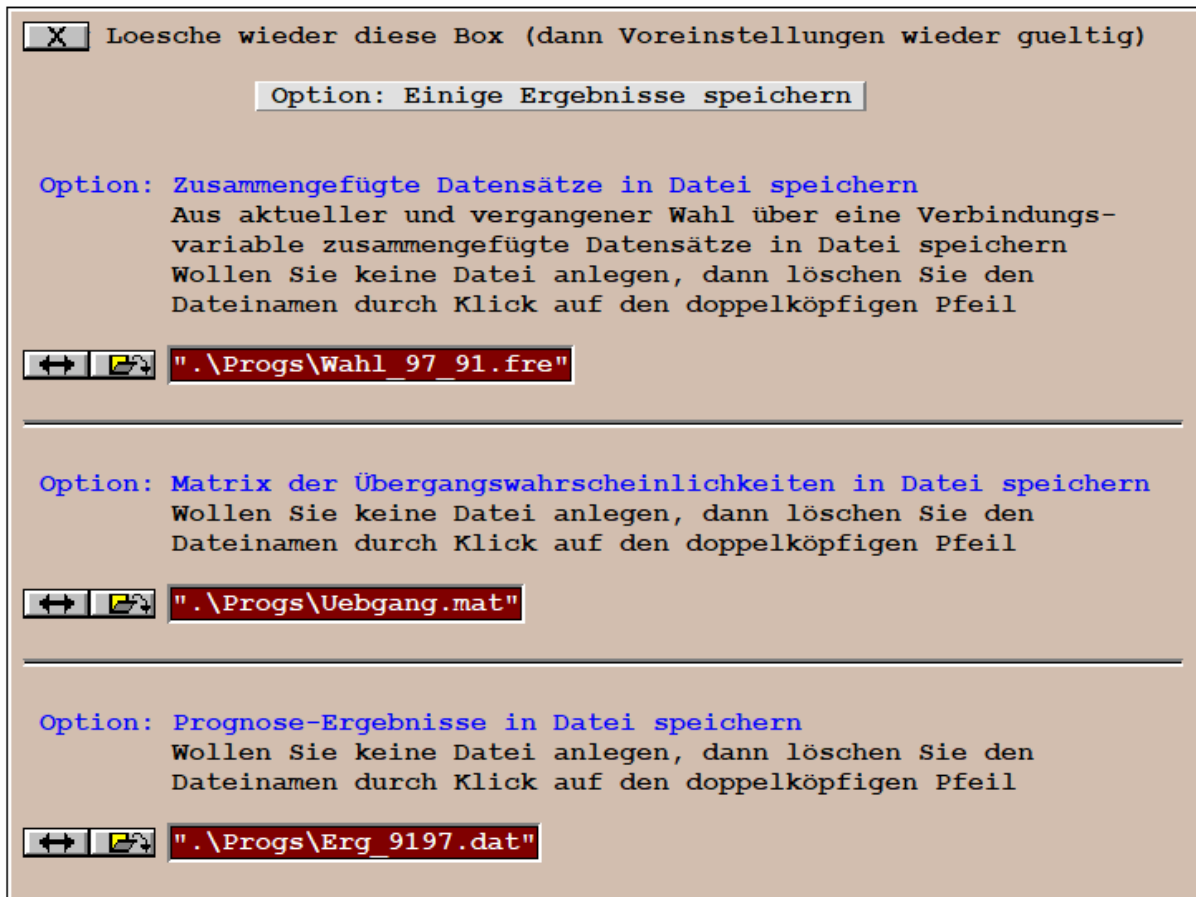
V6	OEVP97	85	1	86
V7	SPOE97	57	1	58
V8	FPOE97	41	1	42
V9	GRUENE97	10	1	11
V10	LIF97	3	-	3
V11	NEUTR97	-	-	-
V12	KPOE97	-	-	-
V13	CWG97	-	-	-
V14	FDP97	-	-	-
V15	OENP97	-	-	-
		----	----	----
zusammen		196	4	200

***** V10 LIF97 liegt unter 5% erhaelt aber mindestens 3 Direktmandate

Eingabebox 21: Einige Ergebnisse in Datei speichern



Wird die Optionsbox geöffnet, dann sieht man folgendes:



1. Eingabefeld: Zusammengeführte Datensätze in eine Datei speichern. Die Wahldaten werden von Prog51m1 in folgender Weise eingelesen: Zuerst wird aus der Datei der aktuellen Wahl der Datensatz einer Wahlgemeinde mit der Gemeindenummer *i* gelesen. In unserem Beispiel besteht dieser Datensatz aus den Variablen V1 bis V16. Dann wird in der Datei der vergangenen Wahl die Wahlgemeinde *i* gesucht und nachdem sie gefunden wurde, wird der betreffende Datensatz eingelesen und an den Datensatz aus der aktuellen Wahl angefügt - in unserem Beispiel als die Variablen V17 bis V26. Dieser gemeinsame Datensatz wird

dann in die angegebene Datei (im Beispiel: "Wahl_97_91.fre") gespeichert. So entsteht eine neue Datei, die je Wahlgemeinde die Variablen der aktuellen und der vergangenen Wahl enthält. Die neue Datei kann in Prog28m2 eingelesen werden. Dieses Programm verlangt als Dateneingabe *eine gemeinsame* Datei aus der aktuellen und der vergangenen Wahl.

BEACHTTE: Die Eingabe ist im Prinzip nur sinnvoll, nachdem alle Wahlgemeinden ausgezählt sind. Wird während der Hochrechnung, in unserem Beispiel etwa nach 35 von 445 Gemeinden, eine zusammengefügte Datei ausgegeben, dann umfasst diese natürlich nur diese 35 Gemeinden.

2. Eingabefeld: Matrix der Übergangswahrscheinlichkeiten in Datei speichern. Dies wird man tun, wenn man anschließend an die Hochrechnung mit Prog52m1 die Mandatsverteilung der österreichischen Nationalratswahl oder der deutschen Bundestagswahl ermitteln möchte. Wird das Eingabefeld leer gemacht, dann wird nicht gespeichert. BEACHTTE: Prog52m1 ist nicht auf dem letzten Stand. Die Reform der Sitzverteilung für den österreichischen Nationalrat und für den deutschen Bundestag von 2011, 2013 und später ist im Programm nicht berücksichtigt.

3. Eingabefeld: Prognose-Ergebnisse in Datei speichern

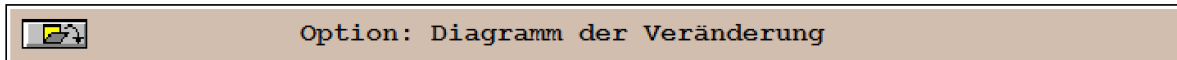
Die Prognose-Ergebnisse aus den aufeinander folgenden Hochrechnungen werden zusammen mit der Uhrzeit in eine Datei gespeichert. Man erhält auf diese Weise ein Protokoll. Also erzeugt beispielsweise folgende Datei:

```
Typ eR = einfache Regression
      kR = einfache Regression mit Konstanter
      mR = multiple Regression
      qO = quadratische Optimierung

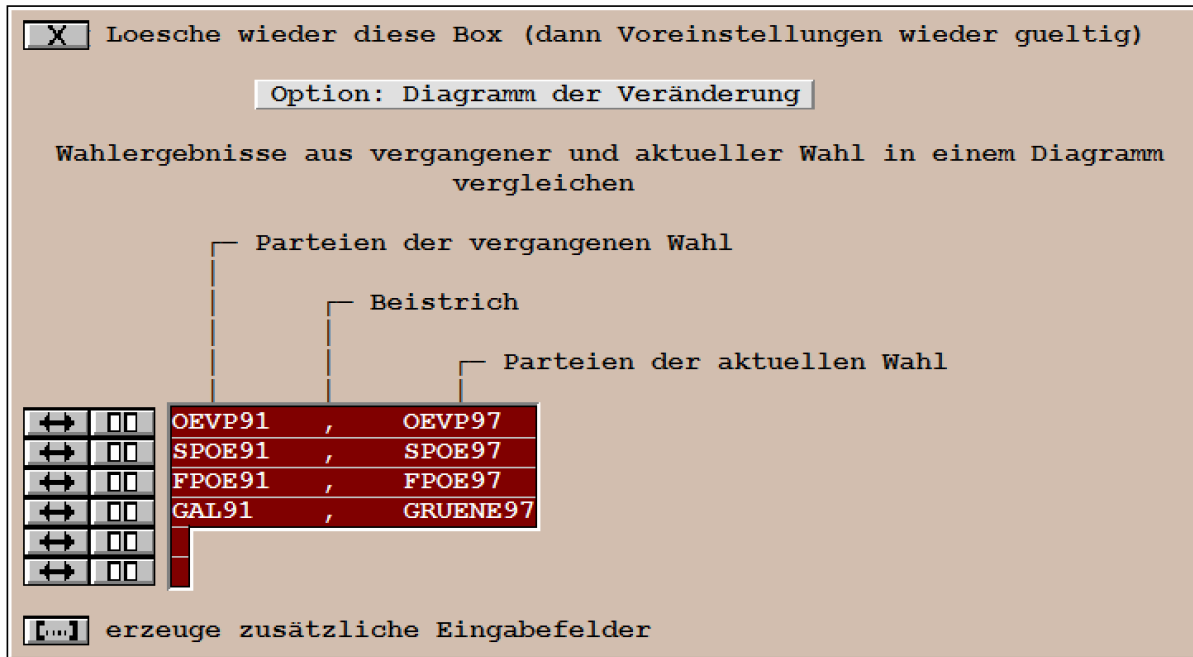
Zahl  = Zahl der ausgezählten Stimmbezirke
Zeit  = Uhrzeit der Hochrechnung
```

Typ	Zahl	Zeit	OEVP97	SPOE97	FPOE97	GRUENE	LIF97	NEUTR9	KPOE97	CWG97	FDP97	OENP97
qO	35	14:00	42.36	28.56	20.59	5.32	1.70	0.67	0.20	0.40	0.07	0.13
mR	35	14:02	40.27	29.80	21.27	5.43	1.98	0.60	0.17	0.29	0.11	0.07
qO	90	15:00	43.43	27.99	20.36	5.17	1.68	0.67	0.16	0.36	0.07	0.12

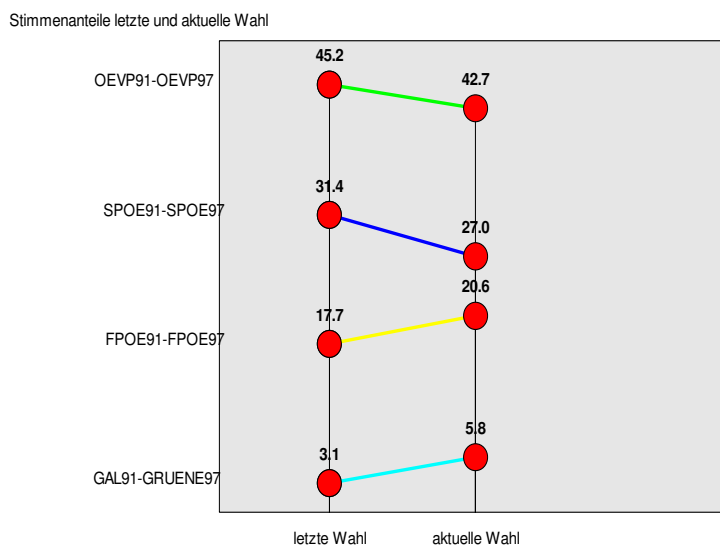
Eingabebox 22: Diagramm der Veränderung erstellen



Wird die Optionsbox geöffnet, dann sieht man folgendes

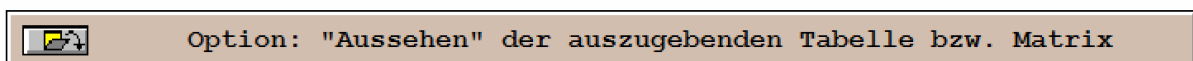


Es kann ein Liniendiagramm erstellt werden, in dem die Wahlergebnisse der vergangenen und der aktuellen Wahl verglichen werden. Beispiel:



Der Benutzer kann dieses Diagramm im Grafik-Editor in vielfältiger Weise verändern. Sie können etwa die Bezeichnungen verschieben. Oder: Klicken Sie beispielsweise auf den Knopf "Diverse Positionen" auf der linken Seite des Grafikeditors und selektieren Sie dann aus der Auswahl möglicher Positionen die erste. Sie erhalten eine 3D-Grafik. Mit Klick auf den Knopf "Einsetzen" wird die so veränderte Grafik in das Ergebnisfenster übergeben

Eingabebox 23: "Aussehen" der auszugebenden Tabellen und Matrizen



Für unser Beispiel ist es empfehlenswert die Optionsbox zu öffnen und so auszufüllen:

X Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)
 Leere alle Eingabefelder

Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix

maximale Breite der Tabelle bzw. Matrix
 Breite einer Tabellen- bzw. Matrixspalte
 Länge des Variablennamens vorne in Tabelle/Matrix
 Länge des Ausprägungsnamens vorne in Tabelle/Matrix

Kommastellen für Werte in Tabelle bzw. Matrix

Die maximale Breite der Tabellenausgabe wurde auf 120 Zeichen gesetzt, damit die Tabellen nicht umgebrochen wird. Voreingestellt sind ohnehin 120 Zeichen.

Mit diesen Eingabewerten erzeugt Almo folgende Matrix, die hier nur teilweise gezeigt wird.

		OEVP91	SPOE91	FPOE91
		V22	V23	V24
OEVP97	V6	88.53	0	0
SPOE97	V7	0	78.64	0
FPOE97	V8	4.40	6.37	87.25
GRUENE97	V9	1.29	3.02	2.03
.
.

```

|---A-----|B|---C----| ---C----|---C----|---C----
      9         8         8         8         8
  
```

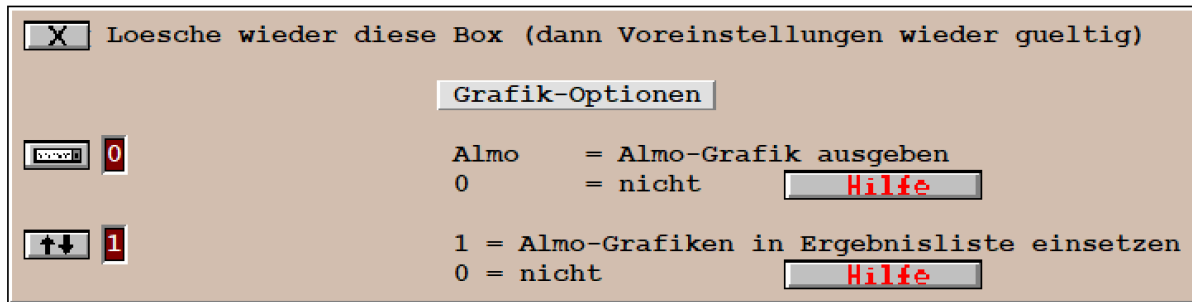
- A = 9 Länge des Variablennamens vorne in Matrix
- B = 0 Länge des Ausprägungsnamens vorne in Matrix
Da keine Ausprägungsnamen vorhanden sind wird B von Almo intern auf 0 gesetzt - unabhängig davon welchen Wert der Benutzer einsetzt
- C = 8 Breite einer Matrixspalte

Ist der Variablen- oder Ausprägungsnamen länger als der Platz, der ihm zur Verfügung gestellt wird, dann wird er entsprechend abgeschnitten.

Eingabebox 24: Grafik-Optionen

Grafik-Optionen

Wird die Optionsbox geöffnet, dann sieht man folgendes



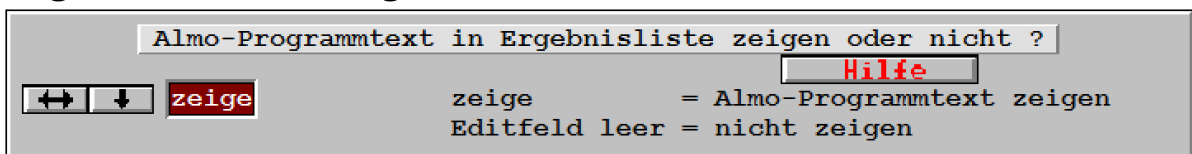
1. Eingabefeld: Klicken Sie auf den Knopf vor dem Eingabefeld und selektieren Sie dann entweder "Almo" oder "0".

0 = Es werden keine Grafiken ausgegeben. Dies gilt prinzipiell - auch für den Fall, dass in einer Eingabeboxen der Programm-Maske eine Grafik angefordert wurde

Almo = Grafiken werden ausgegeben.

2. Eingabefeld: Wird "1" eingegeben, dann werden die Grafiken an der entsprechenden Stelle der Ergebnisliste abgebildet. Bei "0" wird an dieser Stelle nur ein Grafik-Knopf eingesetzt. Durch Klick auf diesen wird die zugehörige Grafik im Grafik-Editor gezeigt.

Eingabebox 25: Almo-Programmtext



Wenn Sie das Eingabefeld leer machen, dann wird der Programmtext, in den Almo das Maskenprogramm übersetzt, nicht gezeigt. Wir raten davon eher ab, da eventuelle Fehlermeldungen von Almo, dann nicht einer Textstelle zugeordnet werden können.

P51.4.5 Ausgabe der Ergebnisse aus Hochrechnung

Wie oben bereits erwähnt, können wir mit den uns zur Verfügung stehenden Daten eine Hochrechnung im Nachhinein rechnen. In der Datei für die Landtagswahl in Oberösterreich 1997 befindet sich eine Variable, in der die Uhrzeit vermerkt ist, zu der die jeweilige Gemeinde ihre Wahlergebnisse an die Zentrale gemeldet hat. Da kleine, ländliche Gemeinden manchmal schon lange vor dem offiziellen Wahlschluss ihre Wahllokale schließen, besteht die Möglichkeit z.B. schon um 14:00 Uhr eine Hochrechnung zu erstellen. (Zu diesem Zeitpunkt waren erst 35 Gemeinden ausgezählt)

Die nachfolgende Ausgabe bezieht sich auf eine derartige Hochrechnung. Wir haben die Eingabebox "Option Simulation einer Hochrechnung" geöffnet und so ausgefüllt:

X Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)

mit einer schon vorhandenen vollständigen Datei
der aktuellen Wahl eine Hochrechnung simulieren

Beispiel:
Hochrechnung mit Wahlergebnisse, die vor 14 Uhr eingetroffen sind
Eingabefeld leer = keine Auswahl

Wenn Zeit97 kleiner 1400

Simulation durch eine Stichprobe
von x Prozent der Gemeinden (z.B. 20)
Eingabefeld leer = keine Stichprobe

Startzahl für Stichprobengenerator

Außerdem rechnen wir mit Zu- und Abwanderern als Quasi-Partei. Die Eingabebox ist folgende:

Zuwanderer
 Gesamtzahl der Zuwanderer in allen Gemeinden

Abwanderer
Zahl wird nicht benötigt

Wir unterstellen hier, dass wir die Gesamtzahl der Zuwanderer zuvor von allen 445 Gemeinden ermitteln konnten (etwa mit Prog51mb, siehe Abschnitt P51.4.3).

Beachte: Es muss die Gesamtzahl der Zuwanderer aus allen Zuwanderungsgemeinden in das Eingabefeld eingesetzt werden - auch wenn nur, wie in unserem Beispiel, mit 35 ausgezählten Gemeinden gerechnet wird.

Wir verzichten zunächst auf eine Wählerstromanalyse - obwohl eine solche auch bei unvollständiger Auszählung möglich wäre. Dadurch wird die Menge an Ergebnissen geringer und überschaubarer. In die Eingabebox muss 0 eingetragen werden

Wählerstromanalyse für ausgezählte Gemeinden
0 = keine Wählerstromanalyse
1 = Wählerstromanalyse durchführen

Damit die Tabellen und Matrizen in ganzer Breite abgebildet werden können (also nicht umgebrochen werden müssen) haben wir die Optionsbox "Aussehen der Tabellen und Matrizen" geöffnet und die Ausgabebreite auf 110 Zeichen gesetzt

X Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)
 ↔ Leere alle Eingabefelder

Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix

↔ maximale Breite der Tabelle bzw. Matrix
 ↔ Breite einer Tabellen- bzw. Matrixspalte
 ↔ Länge des Variablennamens vorne in Tabelle/Matrix
 ↔ Länge des Ausprägungsnamens vorne in Tabelle/Matrix
 ↔ Kommastellen für Werte in Tabelle bzw. Matrix

Die Ergebnisse

Almo liefert folgende Ergebnisse, die wir hier verkürzt wiedergeben:

Es werden 35 Wahlgemeinden von 455 verarbeitet
 Von 979723 Wahlberechtigten sind 29901 ausgezählt
 das entspricht einem Auszählungsgrad von 3.05 %

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Abstroeme
 errechnet aus 35 ausgezählten Gemeinden
 (die Summe der Werte in einer Spalte ist 100 %)

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguel91	Nichtw91	Zuwander
OEVP97	84.5	0	0	0	0	78.6	0	27.4
SPOE97	0	76.0	0	95.8	0	20.5	0	20.2
FPOE97	4.7	2.4	85.4	0	0	0	0	46.8
GRUENE97	2.1	5.3	0.0	0	75.1	0	0	5.6
LIF97	0.4	2.5	3.2	4.2	0	0	0	0
NEUTR97	1.4	0	0.2	0	0	0	0	0
KPOE97	0.4	0	0	0	0	0	0	0
CWG97	0.8	0	0	0	0	0.9	0	0
FDP97	0.2	0	0	0	0	0	0	0
OENP97	0.3	0	0	0	0	0	0	0
Unguel97	2.9	0	4.9	0	0	0	0	0
Nichtw97	1.4	13.8	4.7	0	24.9	0	100.0	0
Abwander	0.9	0	1.6	0	0	0	0	0
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Diese Matrix der Übergangswahrscheinlichkeiten wurde nach dem Kalkül der quadratischen Optimierung mit Modell 2 aus 35 ausgezählten Gemeinden ermittelt. Sie ist das Kernstück des Hochrechnungsverfahrens. Wir nennen sie auch "Matrix der Abstromwahrscheinlichkeiten".

Die Tabelle ist spaltenweise zu interpretieren. Wir betrachten die Spalte OEVP91. Von den Wählern der 35 ausgezählten Gemeinden, die in der vergangenen Wahl 1991 ÖVP gewählt haben, sind 84,5% in der aktuellen Wahl treu geblieben. D.h. 15,5% haben eine andere Partei gewählt. Zur SPÖ hat niemand gewechselt. Zur FPÖ sind 4,7% gewechselt usw. In den Gemeinden (von den 35 ausgezählten), die zwischenzeitlich einen Abwanderungsverlust erlitten haben, hat die ÖVP 0,9% ihrer Wähler von 1991 durch die Abwanderung verloren.

Ungewöhnlich ist, dass von den ÖVP91-Wählern keiner in der aktuellen Wahl zur SPÖ gewechselt ist. Der Kalkül liefert hier eine exakte 0 als Übergangswahrscheinlichkeit. Wir werden im folgenden Abschnitt P51.4.5.1 auf dieses Problem eingehen.

Von wem hat nun die ÖVP97 Stimmen bekommen? Dazu muss die Matrixzeile OEVP97 inspiziert werden. Lediglich in der Spalte "Unguel91" und "Zuwanderer" ist in der ÖVP-Zeile eine Übergangswahrscheinlichkeit größer 0 enthalten. 78,6%

der Ungültigwähler von 1991 haben jetzt, 1997, die ÖVP gewählt. In den Gemeinden (von den 35 ausgezählten) die zwischenzeitlich einen Zuwanderungsgewinn erzielten, haben 27,4% der Zugewanderten jetzt die ÖVP gewählt.

Die Frage

"Wer hat jetzt, in der aktuellen Wahl, seine Stimme von der Partei X91 zur Partei Y97 gegeben?"

ist besser zu beantworten, wenn die Matrix der "Zustromwahrscheinlichkeiten" betrachtet wird. Um diese zu bekommen muss in der Eingabebox "Wählerstromanalyse" eine 1 eingesetzt werden.

Wählerstromanalyse

1

Wählerstromanalyse für ausgezählte Gemeinden
0 = keine Wählerstromanalyse
1 = Wählerstromanalyse durchführen

Almo liefert dann zusätzlich folgende Matrix der "Zustromwahrscheinlichkeiten":

Wählerstroeme - Zustroeme in Prozent - Prognose fuer das Endergebnis
 prognostiziert auf Basis von 35 ausgezaehlten Gemeinden
 (die Summe der Werte in einer Zeile ist 100 %)

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguel97	Nichtw97	Zuwander	Gesamt
OEVP97	93.4	0	0	0	0	4.0	0	2.7	100.0
SPOE97	0	86.5	0	9.0	0	1.5	0	2.9	100.0
FPOE97	10.7	3.8	76.1	0	0	0	0	9.4	100.0
GRUENE97	18.6	32.2	0.0	0	44.8	0	0	4.3	100.0
LIF97	10.9	47.8	34.6	6.7	0	0	0	0	100.0
NEUTR97	95.7	0	4.3	0	0	0	0	0	100.0
KPOE97	100.0	0	0	0	0	0	0	0	100.0
CWG97	95.2	0	0	0	0	4.8	0	0	100.0
FDP97	100.0	0	0	0	0	0	0	0	100.0
OENP97	100.0	0	0	0	0	0	0	0	100.0
Unguel97	60.3	0	39.7	0	0	0	0	0	100.0
Nichtw97	2.7	17.9	3.4	0	3.1	0	72.9	0	100.0
Abwander	57.9	0	42.1	0	0	0	0	0	100.0

Die Stimmen, die die ÖVP in der aktuellen Wahl in den 35 ausgezählten Gemeinden erhielt stammen zu 93,4% von den ÖVP-Wählern der vergangenen Wahl 1991, zu 4% von den Wählern die 1991 ungültig gewählt haben und zu 2,7% von den Personen die in den Gemeinden mit einem Zuwanderungsgewinn zugewandert sind.

Zahl der tatsächlichen Stimmen der Parteien aus allen Gemeinden aus vergangener Wahl

		absolut	in Prozent	nur Parteien in Prozent
V22	OEVP91	357771	36.518	45.197
V23	SPOE91	248640	25.379	31.411
V24	FPOE91	140311	14.321	17.726
V25	VGOE91	20603	2.103	2.603
V26	GAL91	24250	2.475	3.064
V20	UnguelStimm91	16315	1.665	-
V76	Nichtw91	140355	14.326	-
V79	Zuwanderer	31478	3.213	-
		-----	-----	-----
	zusammen	979723	100.000	100.000
	Wahlbeteiligung		82.461	

Diese Tabelle enthält die Wahlergebnisse aus der vergangenen Wahl für alle Gemeinden. Sie

wird mit obiger Übergangsmatrix (der Abströme) multipliziert (siehe oben Abschnitt P51.1.3)
Dadurch entsteht nachfolgende Tabelle.

prognostizierte Wählerströme: Abströme in Zahl der Stimmen
Letzte Zeile: Tatsächliche Stimmzahl in der vergangenen Wahl
Letzte Spalte: Prognostizierte Stimmzahl in der aktuellen Wahl
prognostiziert auf der Basis von 35 ausgezählten Gemeinden

	OEVP91	SPOE91	FPOE91	VGÖE91	GAL91	Ungült91	Nichtw91	Zuwand	Summe 2. Wahl
OEVP97	302406	0	0	0	0	12820	0	8611	323837
SPOE97	0	188879	0	19738	0	3350	0	6369	218336
FPOE97	16770	6035	119856	0	0	0	0	14745	157405
GRÜNE97	7572	13094	7	0	18220	0	0	1753	40646
LIF97	1420	6220	4505	865	0	0	0	0	13010
NEUTR97	4924	0	222	0	0	0	0	0	5146
KPOE97	1562	0	0	0	0	0	0	0	1562
CWG97	2890	0	0	0	0	145	0	0	3035
FDP97	542	0	0	0	0	0	0	0	542
OENP97	1020	0	0	0	0	0	0	0	1020
Ungült97	10417	0	6865	0	0	0	0	0	17282
Nichtw97	5177	34412	6624	0	6030	0	140355	0	192598
Abwander	3072	0	2233	0	0	0	0	0	5306
Summe 1. Wahl	357771	248640	140311	20603	24250	16315	140355	31478	979723

Aus dieser Tabelle können dann die prognostizierten Stimmzahlen der Parteien aus der letzten Spalte "Summe 2.Wahl" entnommen werden und die Prozentwerte, die sie gewonnen haben, errechnet werden. Damit entsteht die folgende Tabelle der prognostizierten Wahlergebnisse:

Endergebnis - Prognose
=====

Zahl der prognostizierten Stimmen der Parteien aus aktueller Wahl
prognostiziert auf der Basis von 35 ausgezählten Gemeinden

		absolut	in Prozent	nur Parteien in Prozent
V6	OEVP97	323837	33.054	42.357
V7	SPOE97	218336	22.285	28.558
V8	FPOE97	157405	16.066	20.588
V9	GRÜNE97	40646	4.149	5.316
V10	LIF97	13010	1.328	1.702
V11	NEUTR97	5146	0.525	0.673
V12	KPOE97	1562	0.159	0.204
V13	CWG97	3035	0.310	0.397
V14	FDP97	542	0.055	0.071
V15	OENP97	1020	0.104	0.133
V4	UngültStimm97	17282	1.764	-
V75	Nichtw97	192598	19.658	-
V80	Abwanderer	5306	0.542	-
		-----	-----	-----
	zusammen	979723	100.000	100.000
	Wahlbeteiligung (prognostiziert)		79.800	

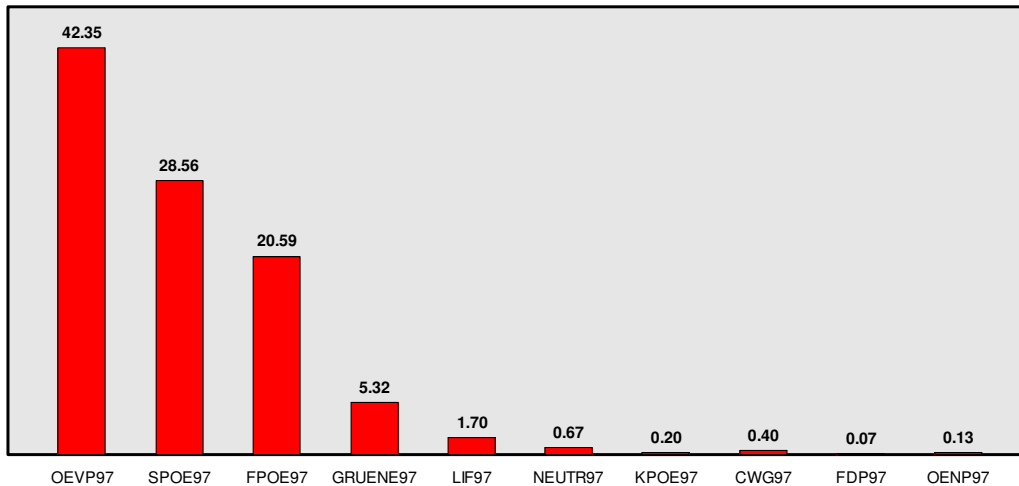
In obiger Tabelle der Wählerströme fällt auf, dass sich die obere und untere Randsumme auf denselben Wert von 979723 summieren. Die Tabelle ist also "in sich stimmig". Diese "Stimmigkeit" ist der Einführung der Zu- und Abwanderer als Quasi-Parteien zu verdanken. Siehe dazu die ausführliche Darstellung des Modells der Quadratischen Optimierung in Abschnitt P51.5.

979723 minus der Zahl der Abwanderer müsste die Zahl 973750 der Wahlberechtigten 1997 ergeben. Das ist nicht ganz exakt der Fall, da die Zahl der Abwanderer vom Modell geschätzt wurde und nach nur 35 ausgezählten Gemeinden noch nicht genau stimmen kann. Es entsteht

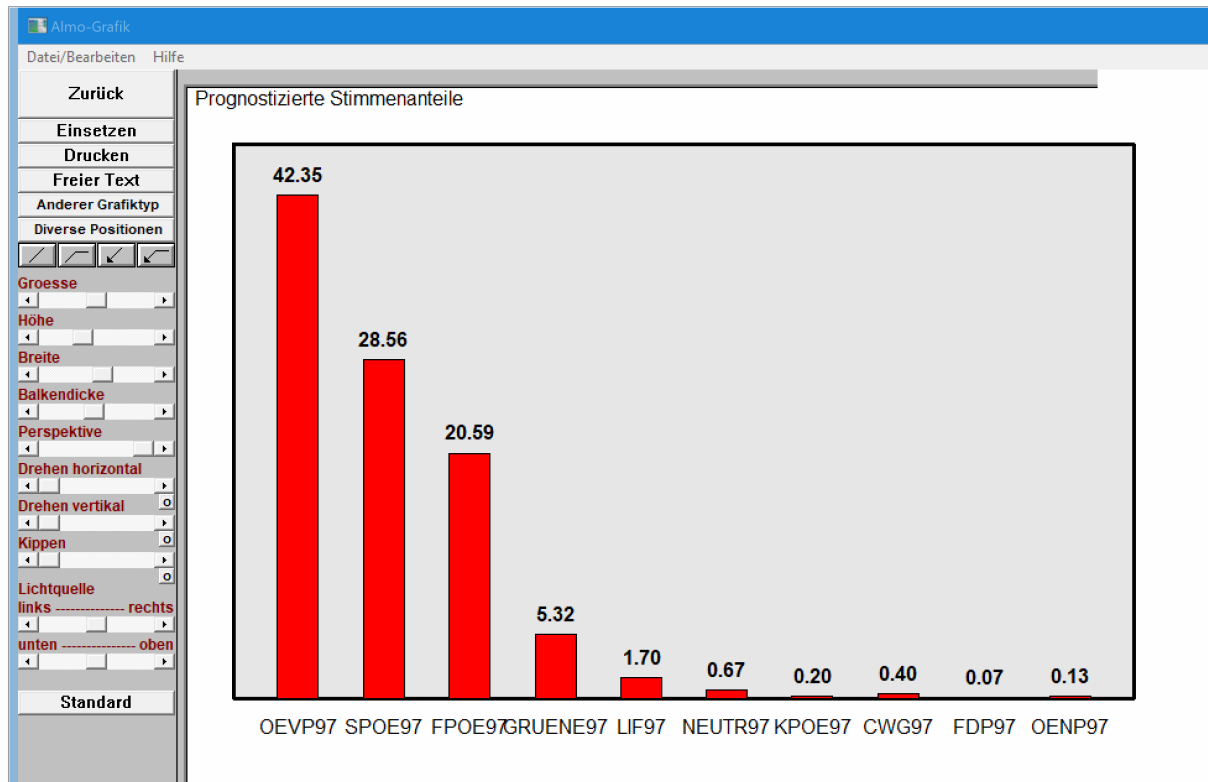
974417. Bei einer Hochrechnung bzw. Wählerstromanalyse nach vollständiger Auszählung wird aber exakt die Zahl der Wahlberechtigten 1997 errechnet.

Almo gibt die Prozentergebnisse noch als Grafik aus

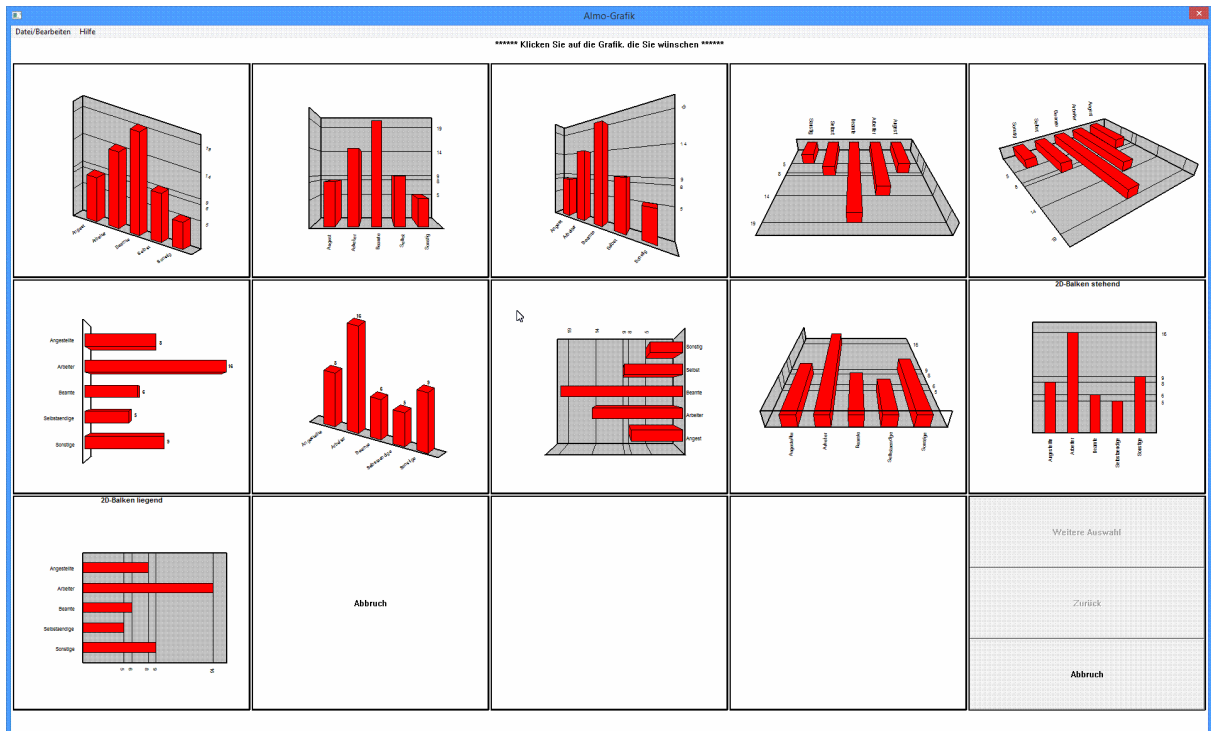
Prognostizierte Stimmenanteile



Die Grafik kann vom Benutzer in folgender Weise verändert werden: Im Ausgabefenster befindet sich oberhalb der Grafik ein großer Grafik-Knopf. Durch Klick auf diesen gelangt man in den Grafik-Editor. Siehe dazu das Almo-Dokument "AnleitungGrafik.pdf" im Almo-Wurzelverzeichnis. Man sieht folgendes. Wir zeigen nur einen Ausschnitt

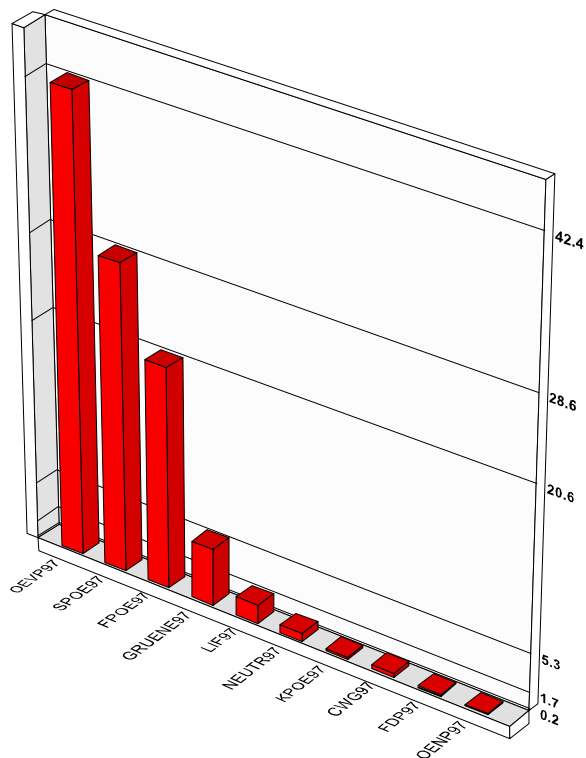


In der linken Leiste klickt man auf "Diverse Positionen". Almo zeigt dann eine Auswahl von alternativen Positionen, in denen die Grafik dargestellt werden kann.



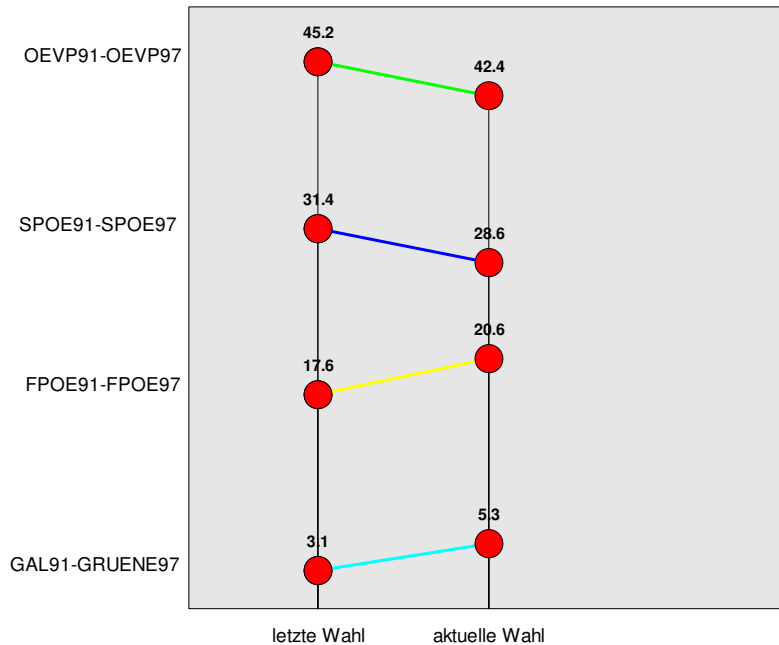
Man klickt in der 1. Reihe auf das 1. Bild. Die Grafik wird dann in dieser Form abgebildet. In der Werkzeug-Leiste auf der rechten Seite des Grafik-Editors klickt man dann noch auf "an Wand (ungl.)" und "3D-Schrift". So entsteht diese Grafik

Prognostizierte Stimmenanteile



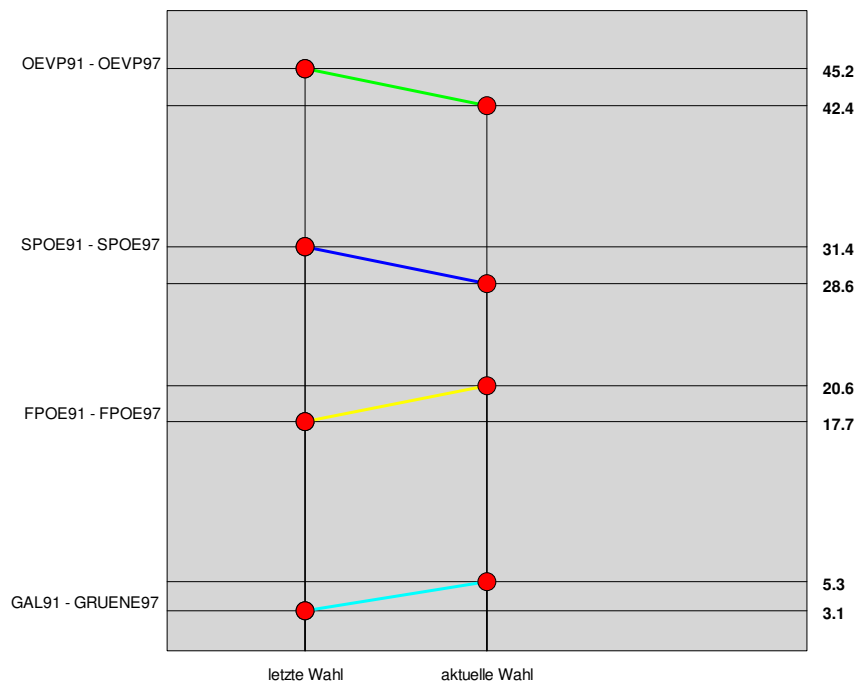
Wurde die Optionsbox "Diagramm der Veränderung" geöffnet und so ausgefüllt wie oben in der Erläuterung zu Eingabebox 22 gezeigt, dann zeichnet Almo noch folgendes Diagramm

Stimmenanteile letzte und aktuelle Wahl



Der Benutzer kann auch diese Grafik in vielfältiger Weise im Almo-Grafik-Editor verändern. In der rechten Werkzeugleiste setzen wir "Masszahlen wohin?" auf "an Wand (ungleich)"

Stimmenanteile letzte und aktuelle Wahl



Da uns das tatsächliche Endergebnis der Landtagswahl von 1997 bekannt ist,

können wir nun vergleichen wie gut unsere Prognose aus den 35 ausgezählten Gemeinden ist:

	prozentuelles Wahlergebnis	
	prognostiziert aus 35 Gemeinden	tatsächlich aus 445 Gemeinden
OEVP97	42.357	42.688
SPOE97	28.558	27.034
FPOE97	20.588	20.630
GRUENE97	5.316	5.781
LIF97	1.702	2.084
NEUTR97	0.673	0.813
KPOE97	0.204	0.281
CWG97	0.397	0.385
FDP97	0.071	0.189
OENP97	0.133	0.114
UnguelStimm97	-	-
Nichtw97	-	-
Abwanderer	-	-
	-----	-----
zusammen	100.000	100.000

Die Differenzen sind gering. Allerdings war diese Landtagswahl eine "unaufgeregte" Wahl, so dass es kein großes Risiko war, schon nach 35 ländlichen Gemeinden das Endergebnis zu prognostizieren.

Wird die Optionsbox "Mandatsverteilung" geöffnet und ausgefüllt, dann prognostiziert Almo noch die Mandate. (Wir haben unterstellt, dass 1997 insgesamt 56 Mandate zu vergeben waren).

Prognostizierte Mandate fuer Parteien aus aktueller Wahl
prognostiziert auf der Basis von 35 ausgezaehlten Gemeinden
gerechnet nach dem Verfahren von d'Hondt

	Mandate

V6 OEVP97	25
V7 SPOE97	16
V8 FPOE97	12
V9 GRUENE97	3
V10 LIF97	-
V11 NEUTR97	-
V12 KPOE97	-
V13 CWG97	-
V14 FDP97	-
V15 OENP97	-

zusammen	56

Die Mandate wurden hier nach dem d'Hondtschen Verfahren ermittelt. Das war für die oberösterreichischen Landtagswahlen damals korrekt.

P51.4.5.1 Übergangsmatrix aus quadratischer Optimierung und multipler Regression

Wir wollen die Übergangsmatrix aus der quadratischen Optimierung vergleichen mit derjenigen der multiplen Regression. Siehe dazu oben die Abschnitte P51.1.2 und P51.1.3. Wir rechnen mit Prog53m2 eine Analyse mit dem Verfahren der multiplen linearen Regression und mit Prog51m1 eine Analyse mit dem Verfahren der quadratischen Optimierung. Prog53m2 findet man nach Klick auf den Knopf "Verfahren" am Oberrand des Almo-Fensters, dann Wahlhochrechnung oder durch

Klick auf den Knopf "alle Progs".

Von der quadratischen Optimierung wurde in Abschnitt P51.1.3 gesagt, dass sie eine Regression mit 2 bzw. 3 Nebenbedingungen ist. Es interessiert nun die Frage "Wie verschieden sind die Koeffizienten, die den Übergang von den Parteien der vergangenen Wahl zu denen der aktuellen Wahl ausdrücken?"

Wir rechnen beide Verfahren mit vollständiger Auszählung und Zu- und Abwanderer. Aus der multiplen Regression erhalten wir folgende Übergangsmatrix. (Um vergleichen zu können haben wir die Matrix transponiert)

Matrix 1: Regressionskoeffizienten aus Prog53m2 (multiple Regression)

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungueelt91	Nichtw91	Zuwander
OEVP97	0.9024	-0.0380	0.0569	0.3317	-0.0028	0.6365	-0.1387	0.2864
SPOE97	0.0035	0.8726	-0.0091	0.1879	-0.2470	0.3251	-0.1437	0.2483
FPOE97	0.0389	0.0856	0.9227	-0.1043	-0.4523	0.0443	0.0106	0.1525
GRUENE97	0.0096	0.0215	0.0059	0.3655	1.0498	-0.1481	0.0087	0.0963
LIF97	-0.0028	0.0063	0.0010	0.1470	0.2365	-0.0741	0.0428	0.0559
NEUTR97	0.0087	0.0118	0.0023	-0.0140	0.0634	0.0438	-0.0154	0.0020
KPOE97	0.0001	0.0030	-0.0035	0.0022	0.0215	0.0065	0.0077	0.0029
CWG97	0.0074	-0.0014	-0.0017	-0.0063	0.0365	0.0450	-0.0041	-0.0011
FDP97	0.0029	-0.0000	-0.0095	0.0242	0.0334	0.0039	0.0016	0.0053
OENP97	-0.0005	-0.0033	0.0096	0.0043	-0.0008	-0.0226	0.0064	-0.0025
Ungueelt97	0.0183	0.0275	0.0279	-0.0119	-0.0077	0.2037	-0.0029	0.0261
Nichtw97	0.0134	0.0188	0.0149	0.0312	0.3046	-0.1587	1.1452	0.2162
Abwander	-0.0019	-0.0045	-0.0174	0.0425	-0.0352	0.0946	0.0816	-0.0883
Gesamt	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Die Übergangsmatrix besteht aus Regressionskoeffizienten. Die Summe der Regressionskoeffizienten in einer Spalte ist (nach dem Kalkül von Bruckmann und wenn Zu- und Abwanderer miteinbezogen werden) zwar 1.0, einige Regressionskoeffizienten sind jedoch negativ oder auch größer 1.0. Der Regressionskoeffizient in der Matrixzelle GAL91-GRUENE97 ist größer 1.0. Die Regressionskoeffizienten können also nicht als Übergangswahrscheinlichkeiten verstanden werden.

Matrix 2: Übergangswahrscheinlichkeiten aus Prog51m1 (quadratische Optimierung)

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungueelt91	Nichtw91	Zuwande
OEVP97	0.8853	0	0	0	0	0	0	0.3604
SPOE97	0	0.7864	0	0	0	0.1141	0	0.3303
FPOE97	0.0440	0.0637	0.8725	0	0	0.0122	0	0.1380
GRUENE97	0.0129	0.0302	0.0203	0.3920	0.7637	0.0569	0	0.0611
LIF97	0.0005	0.0232	0.0264	0.2264	0.0063	0.0940	0	0.0008
NEUTR97	0.0101	0.0003	0	0	0	0.1569	0	0
KPOE97	0.0008	0.0006	0	0	0	0.1049	0	0
CWG97	0.0061	0	0	0	0	0.0466	0	0
FDP97	0.0025	0	0	0	0	0.0352	0	0
OENP97	0	0	0	0	0	0.0533	0	0
Ungueelt97	0.0225	0.0222	0.0111	0	0	0.3236	0	0.0074
Nichtw97	0.0152	0.0588	0.0698	0.2694	0.2300	0	1.0000	0.1019
Abwander	0	0.0146	0	0.1123	0	0.0023	0	0
Gesamt	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

In der Matrix verblüffen die Wahrscheinlichkeiten von exakt 0. Betrachten wir die Spalte SPOE91. Kein Wähler der SPÖ aus der vergangenen Wahl ist in der aktuellen Wahl zur ÖVP gewechselt. Der entsprechende Regressionskoeffizient in Matrix 1 ist negativ. Auch die Regressionskoeffizienten der SPÖ91 gegenüber CWG97, FDP97 und OENP97 sind negativ. Die entsprechenden Übergangswahrscheinlichkeiten aus

der quadratischen Optimierung sind 0. Das gilt fast immer, aber nicht prinzipiell. Der Regressionskoeffizient der SPOE91 gegen Abwanderer ist negativ, die entsprechenden Übergangswahrscheinlichkeiten ist jedoch größer 0.

Regressionskoeffizienten und Übergangswahrscheinlichkeiten variieren weitgehend "parallel". Sie sind aber nicht durch eine lineare Transformation verbunden. Dieses Thema wird ausführlich von Ledl (2007, S.47-53) diskutiert.

P51.4.6 Hochrechnung mit oder ohne Zu- und Abwanderer

Wir rechnen nun eine Hochrechnung ohne Zu- und Abwanderer. Wir verwenden wieder die Daten der bis 14.00 Uhr ausgezählten 35 Gemeinden

	Hochrechnung mit Zu- und Abwanderern	Hochrechnung ohne Zu- und Abwanderern
OEVP97	42.357	42.433
SPOE97	28.558	28.684
FPOE97	20.588	20.522
GRUENE97	5.316	5.263
LIF97	1.702	1.632
NEUTR97	0.673	0.652
KPOE97	0.204	0.206
CWG97	0.397	0.403
FDP97	0.071	0.071
OENP97	0.133	0.134
UngeltStimm	-	-
Nichtw97	-	-
Abwanderer	-	X
	-----	-----
	100.000	100.000

Die Prozentverteilung unterscheiden sich so minimal, dass man sich fragen muss, ob die Mühe lohnt, die Zuwandererzahl zu ermitteln. In unserem Beispiel erklärt sich dieser minimale Unterschied jedoch dadurch, dass die 35 ausgezählten ländlichen Gemeinden eine sehr ähnliche Sozialstruktur besitzen.

In Prog51m1 (und auch in Prog53m2 für die multiple lineare Regression) werden folgende drei Möglichkeiten angeboten:

Möglichkeit 1: Hochrechnung mit Zu- und Abwanderer

Dazu muss die Gesamtzahl der Zuwanderer aus der *gesamten* Wahlpopulation dem Programm mitgeteilt werden. Diese Zahl kann durch Prog51mb (siehe Abschnitt P51.4.3) ermittelt werden. Voraussetzung dafür ist jedoch, dass eine entsprechende Statistik der Wahlberechtigten aus *allen* Wahlgemeinden des *gesamten* Wahlgebiets für die anstehende aktuelle Wahl und die vergangene Wahl verfügbar ist.

Möglichkeit 2: Hochrechnung ohne Zu- und Abwanderer

Die Eingabefelder in der entsprechenden Eingabebox 10 werden leer gemacht.

Bei unvollständiger Auszählung

Die Übergangsmatrizen aus Möglichkeit 1 und 2 sind verschieden. Damit ändert sich auch das prognostizierte Hochrechnungsergebnis in absoluten Stimmzahlen und in Prozentwerten. Die Unterschiede sind jedoch gering.

Bei vollständiger Auszählung:

Dies gilt auch bei vollständiger Auszählung (wenn alle Wahlgemeinden mit ihren

Ergebnissen vorliegen) - jedoch mit einer wesentlichen Ausnahme:

Die prozentuelle Verteilung auf die Parteien ist bei Möglichkeit 1 und 2 gleich und sie ist voll identisch mit der tatsächlichen.

Die Verteilung in absoluten Stimmen ist jedoch nur bei Möglichkeit 1 identisch mit der tatsächlichen.

Betrachten wir die Übergangsmatrizen für Analysen bei vollständiger Auszählung einmal mit und einmal ohne Zu- und Abwanderer

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Abstrome
mit Zu- und Abwanderern

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungueelt	Nichtw9	Zuwande
	V22	V23	V24	V25	V26	V20	V76	V79
OEVP97	88.53	0	0	0	0	0	0	36.04
SPOE97	0	78.64	0	0	0	11.41	0	33.03
FPOE97	4.40	6.37	87.25	0	0	1.22	0	13.80
.
.

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Abstrome
ohne Zu- und Abwanderern

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungueelt	Nichtw9
	V22	V23	V24	V25	V26	V20	V76
OEVP97	87.84	0	0	0	0	32.00	0
SPOE97	1.36	78.42	0	0	0	15.33	0
FPOE97	4.28	7.44	85.95	0	0	0	0
.
.

Die Übergangswahrscheinlichkeiten sind zwischen den beiden Matrizen verschieden. Die 1. Matrix bezieht sich jedoch auf eine Gesamtzahl von

$$948245 \text{ aus der vergangenen Wahl} + 31478 \text{ Zuwanderer} = 979723$$

und die der 2. Matrix nur auf die 948245 aus der vergangenen Wahl.

Die 2. Matrix ist sozusagen "heruntergerechnet" auf die kleineren Gesamtzahl. Im Prinzip sind die beiden Matrizen gleich.

Wenn bei vollständiger Auszählung das Endergebnis vorliegt, dann interessiert nur noch die Wählerstromanalyse. Man möchte die Zu- und Abströme der Stimmen zwischen den Parteien von der vergangenen zur aktuellen Wahl erkunden. Und hier macht es nun einen Unterschied ob man mit oder ohne Zu- und Abwanderer gerechnet hat. Werden Zu- und Abwanderer miteinbezogen, dann summieren sich die Wählerstimmen, die einer Partei von den anderen zuströmen zu der Zahl an Stimmen, die sie in der aktuellen Wahl tatsächlich erhalten hat. Betrachten wir ÖVP, SPÖ und FPÖ.

Vollständige Auszählung: Zu- u. Abströme in Stimmen mit Zu- und Abwanderer

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungueelt	Nicht97	Zuwande	Gesamt97	%
OEVP97	316751	0	0	0	0	0	0	11344	328095	42.688
SPOE97	0	195523	0	0	0	1861	0	10397	207781	27.034
FPOE97	15745	15846	122419	0	0	200	0	4345	158555	20.630
.
.

Gesamt91	357771	248640	140311	20603	24250	16315	140355	31478	979723
----------	--------	--------	--------	-------	-------	-------	--------	-------	--------

Die Zahlen in der letzten Spalte "Gesamt97" sind identisch mit den tatsächlich erhaltenen Stimmenzahlen in der aktuellen Wahl 1997. Werden die Stimmenzahlen in der Spalte "Gesamt97" addiert, dann ergibt sich die Zahl der Wahlberechtigten in 1997(=979723). Die Gesamtstimmen in der untersten Zeile "Gesamt91" sind die von den Parteien in der vergangenen Wahl 1991 gewonnen Stimmen. Wird diese unterste Zeile aufaddiert, dann entsteht (dadurch dass die Zuwanderer hinzukommen) ebenfalls die Zahl 979723. Die Tabelle ist also voll "in sich stimmig".

Betrachten wir nun die Ergebnisse, die aus der Analyse ohne Zu- und Abwanderer entstanden.

Vollständige Auszählung: Zu- u. Abströme in Stimmen **ohne** Zu- und Abwanderer

	OEVP91	SPOE91	FPOE91	VGÖE91	GAL91	Ungült	Nicht97	Gesamt97	%
OEVP97	314280	0	0	0	0	5221	0	319501	42.688
SPOE97	4859	194979	0	0	0	2500	0	202339	27.034
FPOE97	15301	18504	120598	0	0	0	0	154402	20.630
:	:	:	:	:	:	:	:	:	:
Gesamt91	357771	248640	140311	20603	24250	16315	140355	948245	

Die Stimmenzahlen für die aktuelle Wahl in der letzten Spalte "Gesamt97" sind nicht mehr identisch mit den tatsächlich erhaltenen. Sie summieren sich zur Zahl der in der vergangenen Wahl 1991 von allen Parteien zusammen (inklusive Ungültig- und Nichtwähler) gewonnenen Stimmen (=948245). Die Stimmenzahl für die aktuelle Wahl kann also nur für den unwahrscheinlichen Fall korrekt prognostiziert werden, wenn die Wahlbeteiligung in beiden Wahlen in allen Gemeinden dieselbe ist. Die Relation der Stimmen zwischen den Parteien wird jedoch korrekt prognostiziert und dadurch auch ihre Prozentanteile

Möglichkeit 3: Hochrechnung mit Zu- und Abwanderer ohne Gesamtzahl

Zu- und Abwanderer werden mit eingeschlossen. Die Gesamtzahl der Zuwanderer kann jedoch nicht mitgeteilt werden, weil keine Statistik der Wahlberechtigten der kommenden aktuellen Wahl verfügbar ist - oder einfach nur, weil es zu aufwendig ist, diese Zahl zu ermitteln. Das entsprechende Eingabefeld in Eingabebox 10 bleibt leer. Die Eingabebox 10 ist also folgende:

Zu- und Abwanderer miteinbeziehen Hilfe

↔ ↓	↔	Zuwanderer	Zuwanderer Gesamtzahl der Zuwanderer in allen Gemeinden
↔ ↓		Abwanderer	Abwanderer Zahl wird nicht benötigt

Almo ermittelt in diesem Fall die Zuwandererzahl für die *ausgezählten* Wahlgemeinden und setzt diese als Ersatzwert für die unbekannte Gesamtzahl ein. Das hat nachfolgend beschriebene Auswirkung.

Bei unvollständiger Auszählung.

Wir nehmen wieder an, dass nur die 35 bis 14.00 Uhr ausgezählten Gemeinden vorliegen. In die Optionsbox "Simulation einer Hochrechnung" wurde also

eingesetzt:



Zwar ist die Gesamtzahl aller Zuwanderer aus allen Gemeinden nicht bekannt, aber die Zahl der Zu- bzw. Abwanderer aus jeder der 35 ausgezählten Gemeinden ist ermittelbar. Die Matrix der Übergangswahrscheinlichkeiten (die ja nur aus den ausgezählten Gemeinden errechnet wird) ist also dieselbe, wie bei Möglichkeit 1. Betrachten wir die Übergangsmatrizen aus beiden Analysen für ÖVP, SPÖ und FPÖ

Tabelle A

Übergangsmatrix wenn Gesamtzahl der Zuwanderer mitgeteilt

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguelte	Nichtw91	Zuwand
OEVP97	84.5	0	0	0	0	78.6	0	27.4
SPOE97	0	76.0	0	95.8	0	20.5	0	20.2
FPOE97	4.7	2.4	85.4	0	0	0	0	46.8
.

Tabelle B

Übergangsmatrix wenn Gesamtzahl der Zuwanderer nicht mitgeteilt

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguelte	Nichtw91	Zuwand
OEVP97	84.5	0	0	0	0	78.6	0	27.4
SPOE97	0	76.0	0	95.8	0	20.5	0	20.2
FPOE97	4.7	2.4	85.4	0	0	0	0	46.8
.

Tabelle A und B sind identisch. Die Übergangswahrscheinlichkeiten werden mit den Stimmzahlen aus der vergangenen Wahl multipliziert. Daraus entsteht dann die Prognose der Stimmzahl (und der Prozente) für die noch laufende aktuelle Wahl.

Tabelle C

Zu- u. Abströme in Stimmzahl wenn Gesamtzahl der Zuwanderer mitgeteilt

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguelte	Nicht91	Zuwand	Gesamt97	%
OEVP97	302406	0	0	0	0	12820	0	8611	323837	42.357
SPOE97	0	188879	0	19738	0	3350	0	6369	218336	28.558
FPOE97	16770	6035	119856	0	0	0	0	14745	157405	20.588
.

Tabelle D

Zu- u. Abströme in Stimmzahl wenn Gesamtzahl der Zuwanderer nicht mitgeteilt

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguelte	Nicht91	Zuwand	Gesamt97	%
OEVP97	302406	0	0	0	0	12820	0	227	315454	42.984
SPOE97	0	188879	0	19738	0	3350	0	168	212134	28.905
FPOE97	16770	6035	119856	0	0	0	0	389	143050	19.492
.

Die Zuströme von den Parteien aus 91 zu ÖVP97, SPÖ97 und FPÖ97 sind in beiden Tabellen C und D gleich - mit Ausnahme der letzten Spalte, der Zuwanderer. Die Folge davon ist, dass die Gesamtstimmzahl in der letzten Spalte in Tabelle D anders ist als in Tabelle C. So erhält beispielsweise die ÖVP97 in Tabelle C 323837 Stimmen, in Tabelle D hingegen nur 315454.

Das wirkt sich natürlich auch verändernd auf die Prozentwerte aus (die wir hinter die Tabellen C und D geschrieben haben).

Bei vollständiger Auszählung

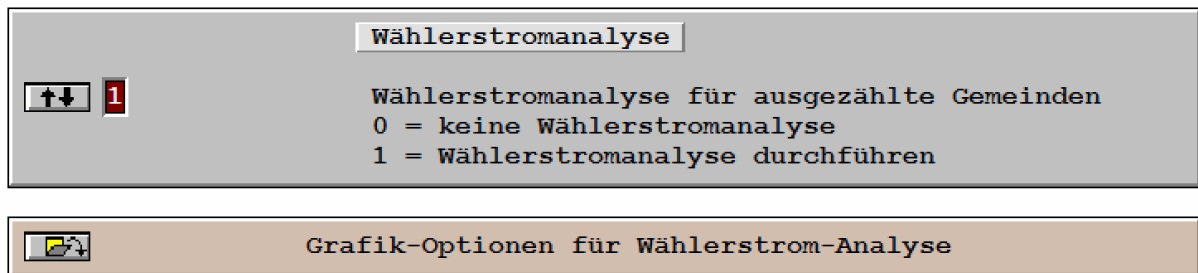
Sind *alle* Wahlgemeinden des *gesamten* Wahlgebiets ausgezählt, dann ist die Situation unproblematisch. Also ermittelt dann die korrekte Gesamtzahl der Zuwanderer. Das gilt nicht, wenn mit einer Stichprobe von Wahlgemeinden aus der aktuellen Wahl (also nicht *allen* Wahlgemeinden) hochgerechnet wurde.

Die Frage, die sich nun abschließend stellt, lautet: Wie soll man sich verhalten, wenn man die Gesamtzahl der Zuwanderer nicht ermitteln kann. Die Empfehlung lautet: Wenn der Auszählungsgrad für die Hochrechnung noch gering ist, ist es sicherer ohne Zu- und Abwanderer zu rechnen. Ist er jedoch hoch, dann mag die Prognose des Endergebnisses besser sein, wenn Zu- und Abwanderer eingeschlossen werden und Also die fehlende Zuwanderer-Gesamtzahl aus den ausgezählten Gemeinden als Ersatzwert einsetzt.

P51.4.7 Ausgabe der Ergebnisse aus Wählerstromanalyse

P51.4.7.1 Wählerstromanalyse bei vollständiger Auszählung

Bei vollständiger Auszählung sind Hochrechnung und Wählerstromanalyse identisch. In Prog51m1 wird in der Eingabebox 18 folgende Einstellung vorgenommen:



Wählerstromanalyse

Wählerstromanalyse für ausgezählte Gemeinden
0 = keine Wählerstromanalyse
1 = Wählerstromanalyse durchführen

Grafik-Optionen für Wählerstrom-Analyse

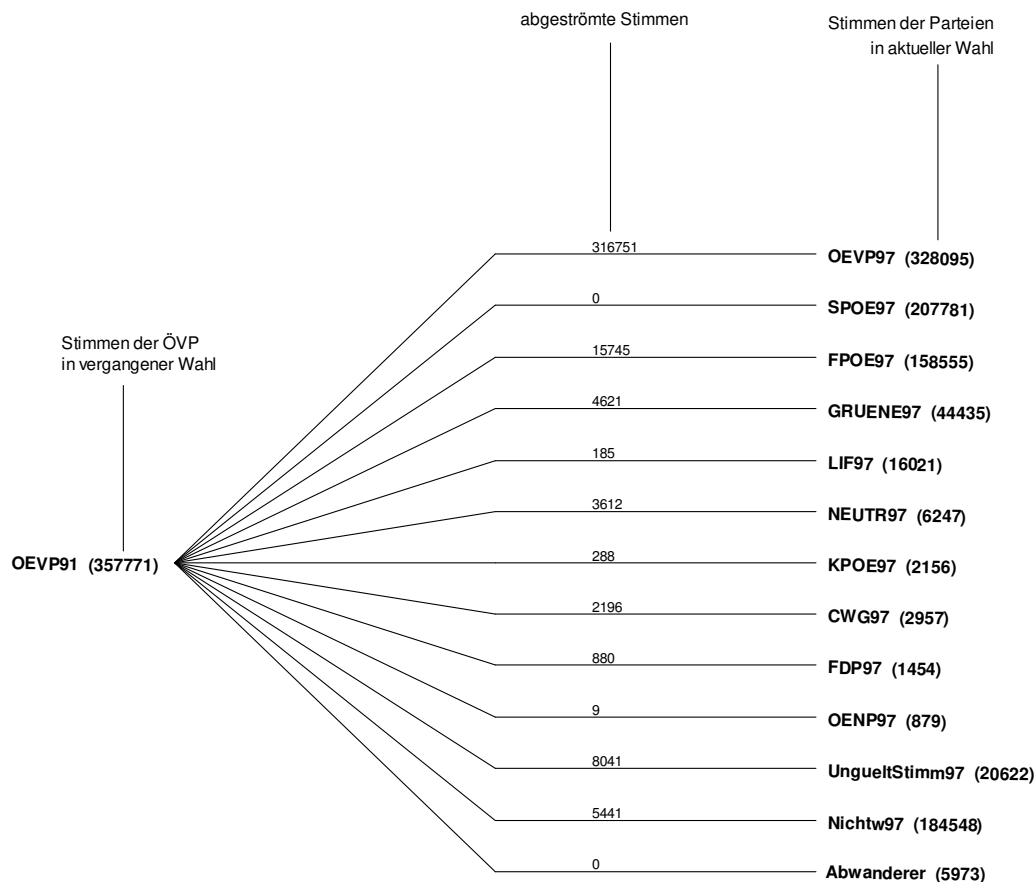
Die Ergebnisse, die Also ausgibt, sind im ersten Teil identisch mit denen einer Hochrechnung (ohne Wählerstromanalyse). Wird die Optionsbox "Grafik-Optionen für Wählerstromanalyse" *nicht* geöffnet, dann werden noch die weiter unten im Text abgebildeten zwei Matrizen der *Zuströme in Prozent* und *in Stimmen* ausgegeben, jedoch keine Grafik gezeichnet.

Wird die Optionsbox "Grafik-Optionen für Wählerstromanalyse" geöffnet, dann sieht man folgendes:

Gesamt	357771	248640	140311	20603	24250	16315	140355	31478	979723
--------	--------	--------	--------	-------	-------	-------	--------	-------	--------

Die Tabelle zeigt die Abströme bei vollständiger Auszählung aller 445 Gemeinden. Die in der Überschrift als "prognostiziert" bezeichnete Stimmenzahl ist also die tatsächliche.

Die 1. Spalte der Tabelle enthält die Abströme von der ÖVP 91 zu den Parteien aus 1997. Das Flussdiagramm dafür ist folgendes:



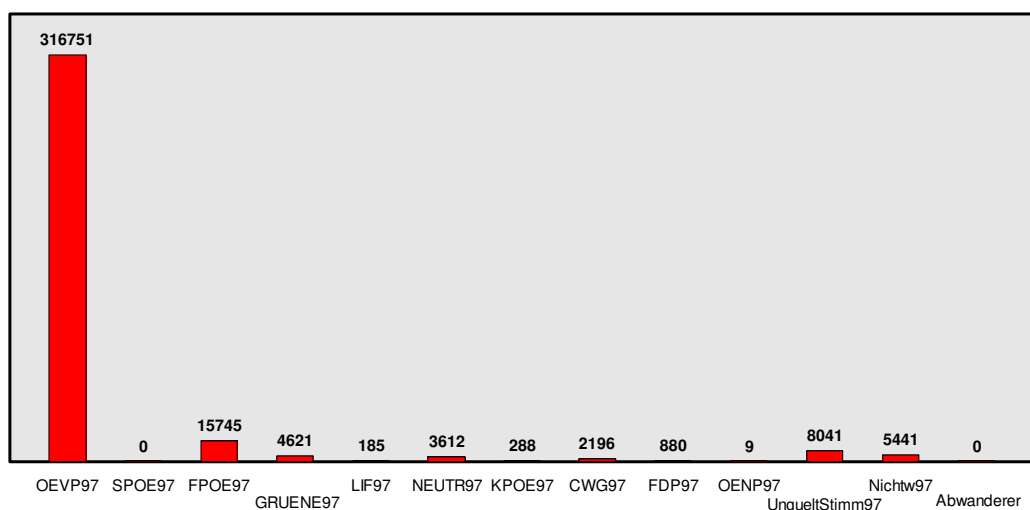
Die ÖVP hat in der vergangenen Wahl 1991 357771 Stimmen erhalten. Zu welchen Parteien der aktuellen Wahl 1997 sind diese Stimmen "abgeströmt" ?

1. An die FPÖ97 wurden 15745 Wähler verloren (die 1991 noch zur ÖVP gegangen waren)
2. An die SPÖ97 wurden keine Wähler verloren.
3. Der ÖVP sind im Wahljahr 1997 316751 Wähler von den 357771 aus 1991 treu geblieben usw.

Werden alle Abströme von der ÖVP91 summiert, dann entsteht die Stimmenzahl der ÖVP von 1991 (=357771). Das gilt aber nur, wenn in der Wählerstromanalyse die Zu- und Abwanderer miteinbezogen wurden und mit Modell 2 gerechnet wurde.

Sollen die Abströme als Balkendiagramm ausgegeben werden (was auch zusätzlich möglich ist), dann erhält man beispielsweise für die ÖVP91 folgende Grafik

Wahlerstroeme in Zahl der Stimmen
Abstroeme von OEVP91 zu ...



Hinweis: Im Grafik-Editor können Beschriftungen verschoben werden. Siehe dazu Almo-Dokument "AnleitungGrafik.pdf" im Almo-Wurzelverzeichnis.

P51.4.7.1.2 Tabelle und Diagramm der Zuströme

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Zustroeme
Wahlerstroeme - Zustroeme in Prozent - Prognose fuer das Endergebnis
prognostiziert auf Basis von 445 ausgezählten Gemeinden
(die Summe der Werte in einer Zeile ist 100 %)

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguelt91	Nichtw91	Zuwander	Gesamt
OEVP97	96.5	0	0	0	0	0	0	3.5	100.0
SPOE97	0	94.1	0	0	0	0.9	0	5.0	100.0
FPOE97	9.9	10.0	77.2	0	0	0.1	0	2.7	100.0
GRUENE97	10.4	16.9	6.4	18.2	41.7	2.1	0	4.3	100.0
LIF97	1.2	36.0	23.1	29.1	0.9	9.6	0	0.2	100.0
NEUTR97	57.8	1.2	0	0	0	41.0	0	0	100.0
KPOE97	13.4	7.3	0	0	0	79.4	0	0	100.0
CWG97	74.3	0	0	0	0	25.7	0	0	100.0
FDP97	60.5	0	0	0	0	39.5	0	0	100.0
OENP97	1.0	0	0	0	0	99.0	0	0	100.0
Unguelt97	39.0	26.7	7.5	0	0	25.6	0	1.1	100.0
Nichtw97	2.9	7.9	5.3	3.0	3.0	0	76.1	1.7	100.0
Abwander	0	60.6	0	38.7	0	0.6	0	0	100.0

Die Tabelle zeigt die Zuströme von Wählern zu den Parteien 1997 in Prozentzahlen. Betrachten wir die 3. Zeile. Von den 100%, die in der aktuellen Wahl 1997 die FPÖ gewählt haben, sind 9.9% ehemalige ÖVP-Wähler von 1991, 10% ehemalige SPÖ-Wähler von 1991 usw.

Wahlerstroeme - Zustroeme in Zahl der Stimmen - Prognose fuer das Endergebnis
prognostiziert fuer das Endergebnis auf der Basis von 445 ausgezählten Gemeinden
(Matrix ist identisch mit der fuer die Abstroeme)

letzte Zeile "Gesamt" enthaelt die tatsaechliche Gesamt-Stimmenzahl je Partei in der vergangenen Wahl

letzte Spalte "Gesamt" enthaelt die vom Modell prognostizierte Gesamt-Stimmenzahl je Partei in der aktuellen Wahl

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Unguelt91	Nichtw91	Zuwander	Gesamt
OEVP97	316751	0	0	0	0	0	0	11344	328095
SPOE97	0	195523	0	0	0	1861	0	10397	207781
FPOE97	15745	15846	122419	0	0	200	0	4345	158555
GRUENE97	4621	7518	2849	8076	18520	928	0	1924	44435
LIF97	185	5763	3698	4664	152	1533	0	27	16021
NEUTR97	3612	75	0	0	0	2560	0	0	6247
KPOE97	288	156	0	0	0	1711	0	0	2156

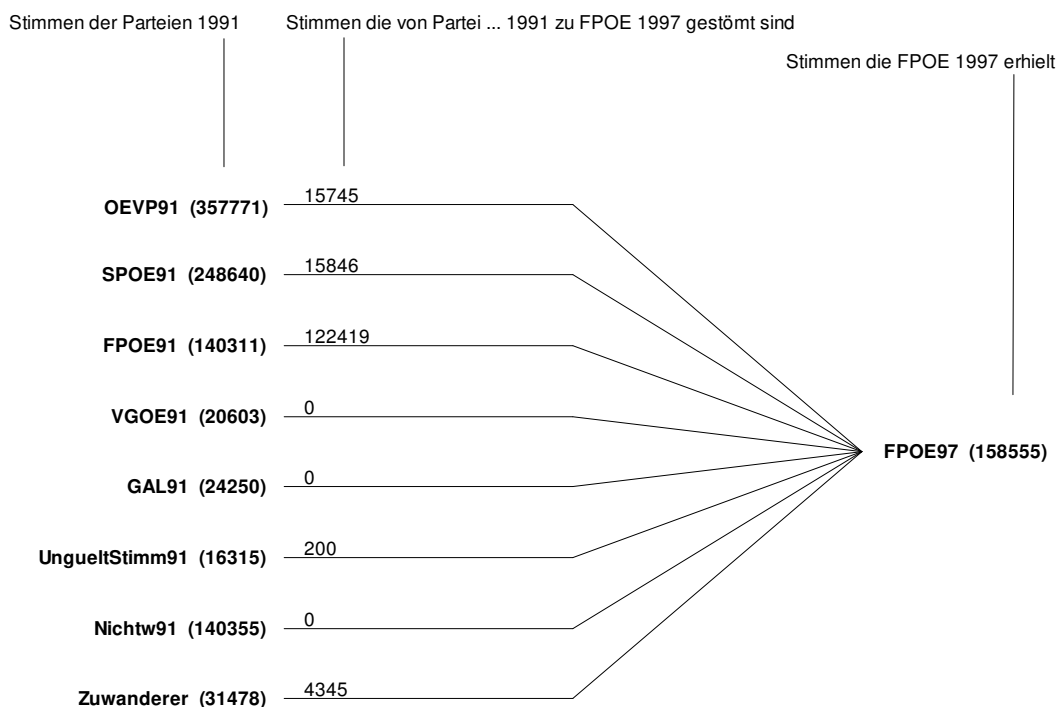
CWG97	2196	0	0	0	0	761	0	0	2957
FDP97	880	0	0	0	0	574	0	0	1454
OENP97	9	0	0	0	0	870	0	0	879
Ungueilt97	8041	5515	1551	0	0	5280	0	234	20622
Nichtw97	5441	14622	9794	5550	5578	0	140355	3207	184548
Abwander	0	3622	0	2314	0	37	0	0	5973
Gesamt	357771	248640	140311	20603	24250	16315	140355	31478	979723

Betrachten wir wieder die 3. Zeile: Die FPÖ hat in der aktuellen Wahl 158555 Stimmen gewonnen. Davon haben in der vergangenen Wahl noch 15745 die ÖVP gewählt. 15846 der FPÖ-Wähler haben 1991 noch SPÖ gewählt und 122419 Wähler sind der FPÖ treu geblieben.

Da Zu- und Abwanderer in die Analyse miteinbezogen wurden und die Zahl der Zuwanderer dem Programm mitgeteilt wurde, ist die Tabelle "stimmig". Die summierte Gesamt-Spalte und Gesamt-Zeile ergibt die Zahl der Wahlberechtigten von 979723 für die aktuelle Wahl.

Im nachfolgenden Flussdiagramm stehen links die Parteien der vergangenen Wahl mit den Stimmen, die sie damals erhielten. Ganz rechts steht die Partei X der aktuellen Wahl mit den Stimmen, die sie jetzt gewonnen hat. Auf den Linien stehen die zugeströmten Stimmen.

Wir betrachten nun als Beispiel das Flussdiagramme, das die *Zuströme* der Parteien von 1991 zur FPÖ 1997 darstellt.



Die Frage, die dieses Diagramm beantworten soll, lautet: Von wem hat die FPÖ 1997 ihre 158555 Stimmen bekommen? Die Antwort lautet:

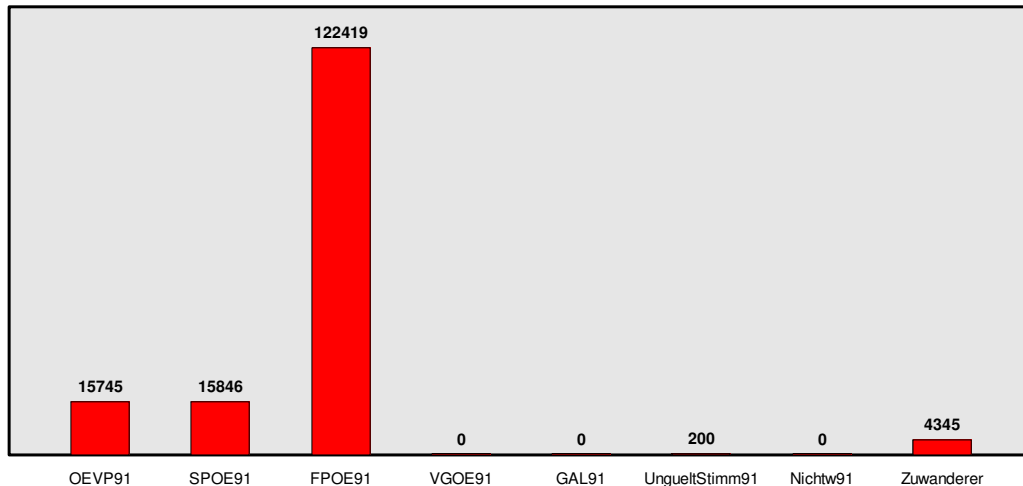
1. Von den 357771 Personen, die 1991 die ÖVP gewählt haben, sind 15745 bei der Wahl 1997 zur FPÖ gewechselt.

2. Von den 248640 Personen, die 1991 die SPÖ gewählt haben, sind 15846 bei der Wahl 1997 zur FPÖ gegangen. usw.

Werden alle zugeströmt Stimmen addiert, dann entsteht die Stimmenzahl der FPÖ von 1997 (=158555). Das gilt aber nur, wenn in der Wählerstromanalyse die Zu- und Abwanderer miteinbezogen wurden und mit Modell 2 gerechnet wurde.

In der Optionsbox wird auch angeboten, die Zuströme als Balkendiagramm grafisch auszugeben. Für FPÖ97 würde folgende Grafik entstehen:

Wählerstroeme in Zahl der Stimmen
Zustroeme zu FPOE97 von ...



P51.4.7.2 Wählerstromanalyse bei unvollständiger Auszählung

Prog51m1 erlaubt es, auch während der Hochrechnungsphase, wenn also die Auszählung noch unvollständig ist, eine Wählerstromanalyse anzufordern. Die Eingabeboxen für die Wählerstromanalyse und die Optionsbox für die Grafik-Optionen zur Wählerstromanalyse müssen entsprechend ausgefüllt werden.

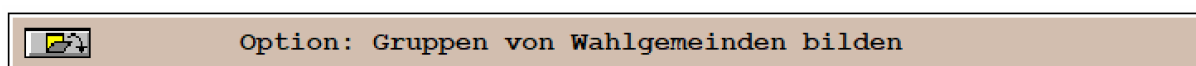
Die Ausgabe ist dieselbe wie bei vollständiger Auszählung. Man muss sich jedoch im Klaren darüber sein, dass eine Wählerstromanalyse *prognostiziert* wird, deren Ergebnisse von den letztendlichen Wählerströmen noch mehr oder weniger abweichen wird. Keinesfalls sollte man zu früh eine solche Analyse rechnen. Will der Benutzer mit unseren Beispieldaten rechnen, dann muss er in Eingabebox 13: Simulation einer Hochrechnung eingeben:

Wenn Zeit97 kleiner 1400

Die Ergebnisse sind dann schon nach 35 ausgezählten Gemeinden sehr gut.

P51.4.8 Hochrechnung/Wählerstromanalyse mit Gruppen von Wahlgemeinden

Im Maskenprogramm Prog51m1 ist folgende Eingabebox 12 enthalten:



Dem Benutzer werden 3 Möglichkeiten angeboten:

1. Hochrechnung bzw Wählerstromanalyse für eine einzelne Gruppe von Wahlgemeinden (z.B. Hochrechnung nur für ein Bundesland oder z.B. nur für den ländlichen Raum des Wahlgebiets)
2. Hochrechnung einer einzelnen Gemeindegruppe, deren Übergangsmatrix verwendet wird um das Ergebnis für das gesamte Wahlgebiet zu prognostizieren. Das ist eine eher ungewöhnliche Vorgehensweise.
3. "Gruppierte" Hochrechnung. Die Gemeinden werden zu mehreren (möglichst homogenen) Gruppen zusammengefasst. Für jede Gruppe wird eine eigene Analyse gerechnet. Die Ergebnisse der einzelnen Gruppen werden zur Prognose für das gesamte Wahlgebiet summiert.

Hinweis:

Wir werden zuerst die Möglichkeit 1 darstellen und dabei die wesentlichen Begriffe erläutern, die für das Hochrechnen mit Gruppen gebraucht werden. Der Leser lese also dieses erste Kapitel, auch wenn ihn primär die als 3. Möglichkeit genannte "gruppierte" Hochrechnung interessiert.

P51.4.8.1 Hochrechnung / Wählerstromanalyse für eine einzelne Gemeindegruppe

Es sollen nur Wahlgemeinden, die durch ein spezifisches Merkmal (oder Merkmalskombination) gekennzeichnet sind, in die Hochrechnung einbezogen werden. Beispiel: Es soll eine Hochrechnung oder Wählerstromanalyse für ländliche Gemeinden bis zu einer Größe von 1000 Wahlberechtigten gerechnet werden. Die Optionsbox wird geöffnet und in folgender Weise ausgefüllt

Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)

Option: Gruppen von Wahlgemeinden bilden

Die Wahlgemeinden werden gruppiert zu folgender Zahl von Gruppen

\Zahl der Gruppen

Im aktuellen Rechengang wird die Gruppe x hochgerechnet

\Nummer der hochgerechneten Gruppe

Die **Übergangswahrscheinlichkeiten** soll nur für jene Wahlgemeinden der vergangenen Wahl ermittelt werden für die folgende Bedingung gilt ...

Die **Stimmensumme je Partei der vergangenen Wahl** soll nur aus den Wahlgemeinden der vergangenen Wahl ermittelt werden für die folgende Bedingung gilt ...

In der 1. Sektion der Optionsbox wird mitgeteilt, dass nur eine Gruppe gebildet wird.

In der 2. Sektion wird angegeben aus welchen Gemeinden die Übergangsmatrix errechnet werden soll. In unserem Beispiel wird eingetragen

Wahlberecht91 kleiner 1000

Damit formuliert der Benutzer die Bedingung nach der die Wahlgemeinden selektiert werden sollen. In dieser Bedingung sollten nur Variable auftreten, die aus der Datei der vergangenen Wahl stammen. D. h. die Gemeinden, die für die Gruppe ausgewählt werden, sollten über Variable aus der Datei der vergangenen Wahl definiert werden. Wir werden anschließend im Unter-Abschnitt "Die Selektions-variable" nochmals ausführlich darauf eingehen.

Almo selektiert dann aus den (zum Hochrechnungs-Zeitpunkt bereits) ausgezählten Gemeinden der aktuellen Wahl jene, die zu den ausgewählten Gemeinden der vergangenen Wahl passen, d.h. diejenigen welche die gleiche Gemeindenummer haben. Bereits ausgezählte Gemeinden der aktuellen Wahl, die größer als 1000 Wahlberechtigte sind, werden also nicht berücksichtigt.

In der 3. Sektion wird angegeben, für welche Gemeinden der vergangenen Wahl die Stimmensummen je Partei ermittelt werden soll. Hier *muss* wie in der 2. Sektion eingetragen werden

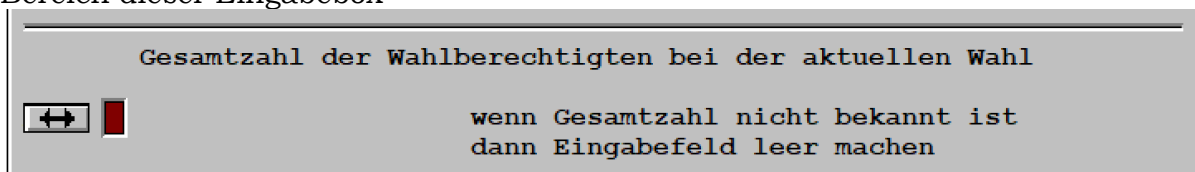
Wahlberecht91 kleiner 1000

Für die so ausgewählten Gemeinden wird die Übergangsmatrix dann mit den Stimmensummen aus der vergangenen Wahl multipliziert. Damit ergeben sich die für die aktuelle Wahl prognostizierten Stimmen bzw. Prozente je Partei und die entsprechende Wählerstromanalyse. Die Prognose gilt nur für die Gruppe von Gemeinden kleiner 1000.

Die Auswahl-Bedingung muss durch eine Variable definiert sein, die sich auf die *vergangene* Wahl bezieht, die somit in der Datei der vergangenen Wahl enthalten ist. Da Stimmensumme und Übergangsmatrix für dieselbe Gemeindegruppe berechnet werden müssen, folgt daraus, dass die Auswahl-Bedingung für die Übergangsmatrix (in der 2. Sektion) identisch sein muss. Wir werden anschließend im Unter-Abschnitt "Die Selektionsvariable" auf dieses Thema zurückkommen.

Für diese Hochrechnung/Wählerstromanalyse einer einzelnen Gruppe muss in Prog51m1 in folgenden Eingabeboxen noch eingetragen bzw. geändert werden:

a. In der Eingabebox 7 "Parteien der aktuellen Wahl" wird das Eingabefeld für die Gesamtzahl der Wahlberechtigten leer gelassen. Wir zeigen nur den unteren Bereich dieser Eingabebox



Man kann aber auch diese Gesamtzahl mit der entsprechend ausgefüllten Programm-Maske Prog51mb berechnen und sich ausgeben lassen. "Prog51mb_Zuwand_Grp1" ist bereits für unsere Beispieldaten der Landtagswahl

Auch "UND" sowie "ODER" sind möglich. Beispiel:

```
V5 gleich 7 UND V10 groesser V20  
V6 kleiner V5 ODER V12 nichtgleich 2
```

Bei "UND" müssen beide Bedingungen erfüllt sein bei "ODER" genügt es wenn mindestens eine der beiden erfüllt ist. Siehe die ausführliche Darstellung im Almo-Dokument Nr.0 "Arbeiten mit Almo", Abschnitt P 0.7.1.

Die Selektionsvariable

Der Benutzer definiert in der 2. und 3. Sektion der Eingabebox die Bedingung nach der Gemeinden für die Hochrechnung selektiert werden. Die Frage ist nun, welche Variablen dürfen in dieser Bedingung vorkommen?

Selektionsvariable für die Übergangsmatrix.

In Abschnitt P51.4.2 haben wir gezeigt, wie Almo die beiden Dateien der aktuellen und der vergangenen Wahl verknüpft. Hier nochmals kurz wiederholt: Für die gerade eingeleseene Gemeinde aus der Datei der aktuellen Wahl sucht Almo in der Datei der vergangenen Wahl nach der dazu passenden Gemeinde (mit der gleichen Gemeindenummer). Der Benutzer hat dafür in der Eingabebox die Gemeindenummer als Verbindungsvariable angegeben. Ist sie gefunden, dann werden die beiden Datensätze zu einem gemeinsamen verknüpft.

Der Benutzer kann nun in der 2. Sektion eine Variable aus diesem gemeinsamen Datensatz verwenden um festzulegen, welche spezifischen Gemeinden ausgewählt werden sollen. Er schreibt in unserem Datenbeispiel in das Eingabefeld die Bedingung `Wahlberecht91 kleiner 1000`. Er verwendet also eine Variable aus der *vergangenen* Wahl. Danach werden für die Analyse nur jene gemeinsamen Datensätze verwendet, die diese Bedingung erfüllen. Im Prinzip möglich wäre es auch eine Variable aus der *aktuellen* Wahl einzusetzen, etwa zu schreiben `Wahlberecht97 kleiner 1000`. Dabei können jedoch Probleme auftreten, auf die wir gleich anschließend eingehen werden.

Zulässig ist auch eine Kombination von Merkmalen. Beispiel: Man will eine Wählerstromanalyse rechnen für kleine überwiegend landwirtschaftliche Gemeinden. Wir unterstellen, dass in der Datei der vergangenen Wahl eine Variable `Landwirtschaft91` vorhanden ist, die den Prozentanteil der landwirtschaftlich Beschäftigten enthält. Die Bedingung würde dann lauten

```
Wahlberecht91 kleiner 1000 und Landwirtschaft91 grösser 50
```

Selektionsvariable für die Stimmensumme.

Die Stimmensumme ist selbstverständlich nur aus der vergangenen Wahl bekannt. Es muss also eine Variable der vergangenen Wahl als Selektionsvariable verwendet werden. Der Benutzer schreibt in der 3. Sektion der Eingabebox `Wahlberecht91 kleiner 1000`. Wurde in der 2. Sektion für die Übergangsmatrix `Wahlberecht97` geschrieben, dann muss bedacht werden, dass sich die Gemeindegröße von 91 auf 97 verändert haben kann. In unserem Hochrechnungsprogramm würde die Übergangsmatrix für Gemeinden errechnet werden, die 97 kleiner 1000 waren. Diese würde dann multipliziert werden mit den Stimmensummen je Partei für die Gemeinden, die 91 kleiner 1000 waren. Letztendlich muss der Benutzer entscheiden, ob er Wahlprognosen bzw. Wählerstromanalysen akzeptieren will, bei denen Übergangsmatrix und Stimmensumme nicht aus der exakt gleichen Wählerpopulation stammen.

Übergangsmatrix und Stimmensumme

Wichtig ist folgendes zu erkennen: Der Eintrag in der 3. Sektion bestimmt für welche Gemeinde-Gruppe X eine Wahlprognose hochgerechnet bzw. eine

Wählerstromanalyse erstellt werden soll - wobei eine Übergangsmatrix eingesetzt wird, die für die Gemeindegruppe Y ermittelt wird. Offensichtlich macht es nur Sinn wenn X gleich Y ist, d.h. wenn Übergangsmatrix und Stimmensummen für dieselbe Gemeindegruppe errechnet wurden. Möglicherweise ist aber z.B eine Analyse sinnvoll, bei der die Übergangsmatrix gebildet wird nach der Bedingung

Wahlberecht91 kleiner 1000

und die Stimmensumme nach der Bedingung

OEVP91/AbgegebenStimm91 groesser 0.5

Die Übergangsmatrix wird hier für eine Gruppe von sehr kleinen Wahlgemeinden ermittelt und die Stimmensummen für eine Gruppe von Gemeinden die ÖVP-Hochburgen sind (ÖVP-Anteil grösser 50%). Es wird also eine Wahlprognose hochgerechnet bzw. eine Wählerstromanalyse erstellt für ÖVP-Hochburgen - wobei eine Übergangsmatrix verwendet wird, die aus sehr kleinen Gemeinden errechnet wurde.

Wir überlassen es dem Benutzer darüber nachzudenken, ob bei derartigen Analysen spezielle Erkenntnisse gewonnen werden können und ob er sich Konstellationen vorstellen kann, bei der solche Analysen Sinn machen. Also erlaubt jedenfalls derartige Analysen.

P51.4.8.2 Schluss von einer Gemeindegruppe auf das Gesamtergebnis

Wir erzeugen einen Sonderfall der Hochrechnung für eine einzelne Gruppe, wenn wir die Eingabebox 12 in folgender Weise ausfüllen:

X Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)

Option: Gruppen von Wahlgemeinden bilden

Die Wahlgemeinden werden gruppiert zu folgender Zahl von Gruppen

\Zahl der Gruppen

Im aktuellen Rechengang wird die Gruppe x hochgerechnet

\Nummer der hochgerechneten Gruppe

Die **Übergangswahrscheinlichkeiten** soll nur für jene Wahlgemeinden der vergangenen Wahl ermittelt werden für die folgende Bedingung gilt ...

Die **Stimmensumme je Partei der vergangenen Wahl** soll nur aus den Wahlgemeinden der vergangenen Wahl ermittelt werden für die folgende Bedingung gilt ...

Das letzte Eingabefeld bleibt leer. Das bedeutet: Also berechnet in diesem Fall zwar die Übergangsmatrix nur für die selektierten Wahlgemeinden. Multipliziert diese aber dann mit der gesamten Stimmensumme je Partei der vergangenen Wahl. Es wird also auf Basis einer Gruppe das Wahlergebnis für das gesamte Wahlgebiet prognostiziert. Das kann nur in einem ganz speziellen Fall sinnvoll sein. Programm-Maske Prog51m1 ist ansonsten auszufüllen, wie oben für die Analyse einer einzelnen Gruppe beschrieben - mit einer Ausnahme: Sollen Zu- und Abwanderer miteinbezogen werden, dann muss die Zahl der Zuwanderer für das gesamte Wahlgebiet und nicht nur für die einzelne Gruppe eingesetzt werden.

P51.4.8.3 "Gruppierte" Hochrechnung/Wählerstromanalyse

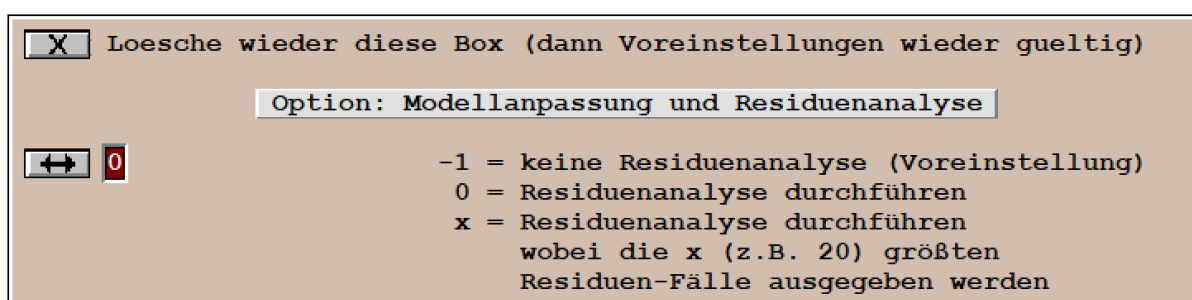
Gemeinden werden zu Gruppen zusammengefasst, z.B. in 4 Gruppen nach der Gemeindegröße. Die Gemeinden können auch durch Merkmalskombinationen gruppiert werden. Die einzelne Gruppe sollte maximal homogen sein. Die Gruppen im Vergleich zueinander sollten maximal heterogen sein. In anderen Worten: Die Gemeinden einer Gruppe sollten in ihren Übergangswahrscheinlichkeiten (von Partei X zu Y) maximal ähnlich sein. Die Übergangsmatrizen der Gruppen sollten sich jedoch deutlich unterscheiden.

Für jede einzelne Gruppe wird eine eigenständige Hochrechnung gerechnet. Für das gesamte Wahlgebiet entsteht die Hochrechnung bzw. Wählerstromanalyse dann dadurch, dass die für jede Gruppen prognostizierten Stimmenzahlen je Partei summiert werden.

Die Hochrechnungen für alle Gruppen sollten zum gleichen Auszählungszeitpunkt erfolgen. Beispiel: Bis zum Zeitpunkt 18.10 Uhr ist eine bestimmten Anzahl von Wahlgemeinden ausgezählt. Diese Gemeinden werden, wie beschrieben, gruppiert, hochgerechnet und summiert.

Die Gruppierungsmerkmale können als Variable in der Datei der vergangenen Wahl bereits enthalten sein. Beispiel: Die Gemeindegröße (=Zahl der Wahlberechtigten). Die Gruppierung kann dann unmittelbar im Hochrechnungsprogramm Prog51m1 in Eingabebox 12 veranlasst werden. Oder: Die Gruppierung muss zuvor und unabhängig von Prog51m1 vorgenommen werden. Die Gruppenzugehörigkeit wird dann als zusätzliche Variable in die Datei der vergangenen Wahl übertragen. Wir werden später zeigen, wie im Anschluss an eine Clusteranalyse die Cluster-zugehörigkeit in die Datei der vergangenen Wahl eingetragen wird.

Als Indikator und Maß für die Homogenität einer Gruppe kann die *mittlere absolute prozentuelle Abweichungen der Stimmenanteile* verwendet werden. Je kleiner diese Abweichung ist umso homogener ist die Gruppe. Siehe dazu Abschnitt P51.6.2. Um die Abweichung zu ermitteln, muss die Optionsbox 17 "Modellanpassung und Residuenanalyse" geöffnet werden und in das 1. Eingabefeld eine "0" eingesetzt werden. Wir zeigen nur den oberen Bereich dieser Eingabebox



Im Bestreben möglichst homogene Gruppen zu bilden, kann man übertreiben. Man fasst z.B. die Wahlgemeinden der Größe nach zu 20 Gruppen zusammen. Einzelne Gruppen können dann so wenige Gemeinden beinhalten, dass die Übergangsmatrix nicht mehr berechnet werden kann. Das kann insbesondere beim Hochrechnen zu einem frühen Zeitpunkt des Wahltages geschehen. Man wird also zunächst ungruppiert hochrechnen und erst zu einem späteren Zeitpunkt gruppiert. Immer möglich ist eine gruppierte Wählerstromanalyse nachdem alle Gemeinden ausgezählt sind.

Wir werden im Folgenden ausführlich eine "gruppierte" Hochrechnungen mit 4 Gruppen vortragen. Die Programme dazu sind unter den Namen Prog51m1_Grp1 bis Prog51m1_Grp4 in Almo enthalten. Der Benutzer findet sie nach Klick auf den Knopf "alle Progs" am Oberrand des Almo-Fensters oder über "Verfahren/Wahlhochrechnung".

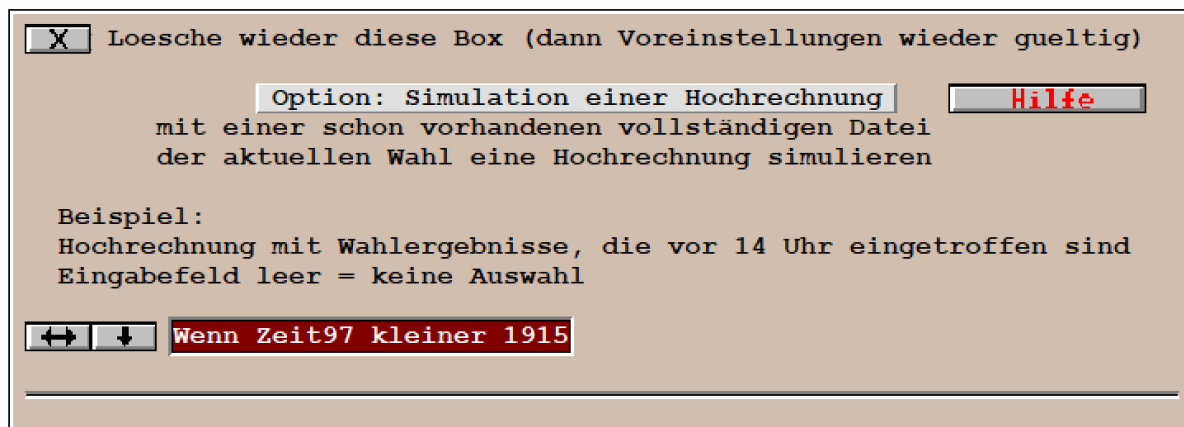
Wir gruppieren nach der Gemeindegröße (bzw. Zahl der Wahlberechtigten) in folgende 4 Gruppen

Gruppe 1:	bis 999	Wahlberechtigte
Gruppe 2:	1000-2999	Wahlberechtigte
Gruppe 3:	3000-7999	Wahlberechtigte
Gruppe 4:	8000 u. mehr	Wahlberechtigte

Unsere Beispieldaten der oberösterreichischen Landtagswahl eignen sich nicht gut für eine Gruppierung der Gemeinden in 4 Größen-Gruppen. Da Oberösterreich nur eine Großstadt und nur 3 Mittelstädte besitzt, musste Gruppe 4 (die "städtische Gruppe") schon bei 8000 Wahlberechtigten beginnen, um mindestens 11 Wahlgemeinden zu bekommen. Um es gleich vorwegzunehmen: Eine gruppierte Analyse für unsere Daten rentiert sich kaum. Für eine große nationale Wahl wird sie sich immer lohnen. So hat das SORA-Institut, das für das österreichische Fernsehen Hochrechnungen durchgeführt hat, die Gemeinden in 5 Gruppen zusammengefasst (allerdings nicht nach der Gemeindegröße). Siehe dazu die Ausführungen von Hofinger/ Ogris (2002) und Ledl (2007, S.46 ff.).

Der Benutzer muss für jede Gruppe ein eigenes Programm erstellen. Die Programm-Maske Prog51m1 wird 4 Mal ausgefüllt. Der Benutzer betrachte die bereits oben erwähnten 4 Programme Prog51m1_Grp1 bis Prog51m1_Grp4.

Wir simulieren in unserem Beispiel eine Hochrechnung um 19:15 Uhr am Wahlabend. Die Optionsbox 13 "Simulation einer Hochrechnung" wird folgendermaßen ausgefüllt.



In der Programm-Maske Prog51m1 müssen dann noch folgende Eingabeboxen geändert werden. Wir haben das bereits oben bei der Hochrechnung für eine einzelne Gemeindegruppe in den Punkten a und b gezeigt, so dass wir das hier nicht wiederholen werden

Die Eingabebox 12 wird für die Gruppe 1 so ausgefüllt (siehe Prog51m1_Grp1)

X **Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)**

Option: Gruppen von Wahlgemeinden bilden

Die Wahlgemeinden werden gruppiert zu folgender Zahl von Gruppen

\Zahl der Gruppen

Im aktuellen Rechengang wird die Gruppe x hochgerechnet

\Nummer der hochgerechneten Gruppe

Die **Übergangswahrscheinlichkeiten** soll nur für jene Wahlgemeinden der vergangenen Wahl ermittelt werden für die folgende Bedingung gilt ...

Die **Stimmensumme je Partei der vergangenen Wahl** soll nur aus den Wahlgemeinden der vergangenen Wahl ermittelt werden für die folgende Bedingung gilt ...

In der 1. Sektion der Box wird mitgeteilt, dass insgesamt 4 Gruppen gerechnet werden sollen und dass mit dem vorliegenden jetzt zu rechnenden Programm die Gruppe 1 gerechnet werden soll.

In den 4 Programmen wird als Auswahl-Bedingung eingesetzt

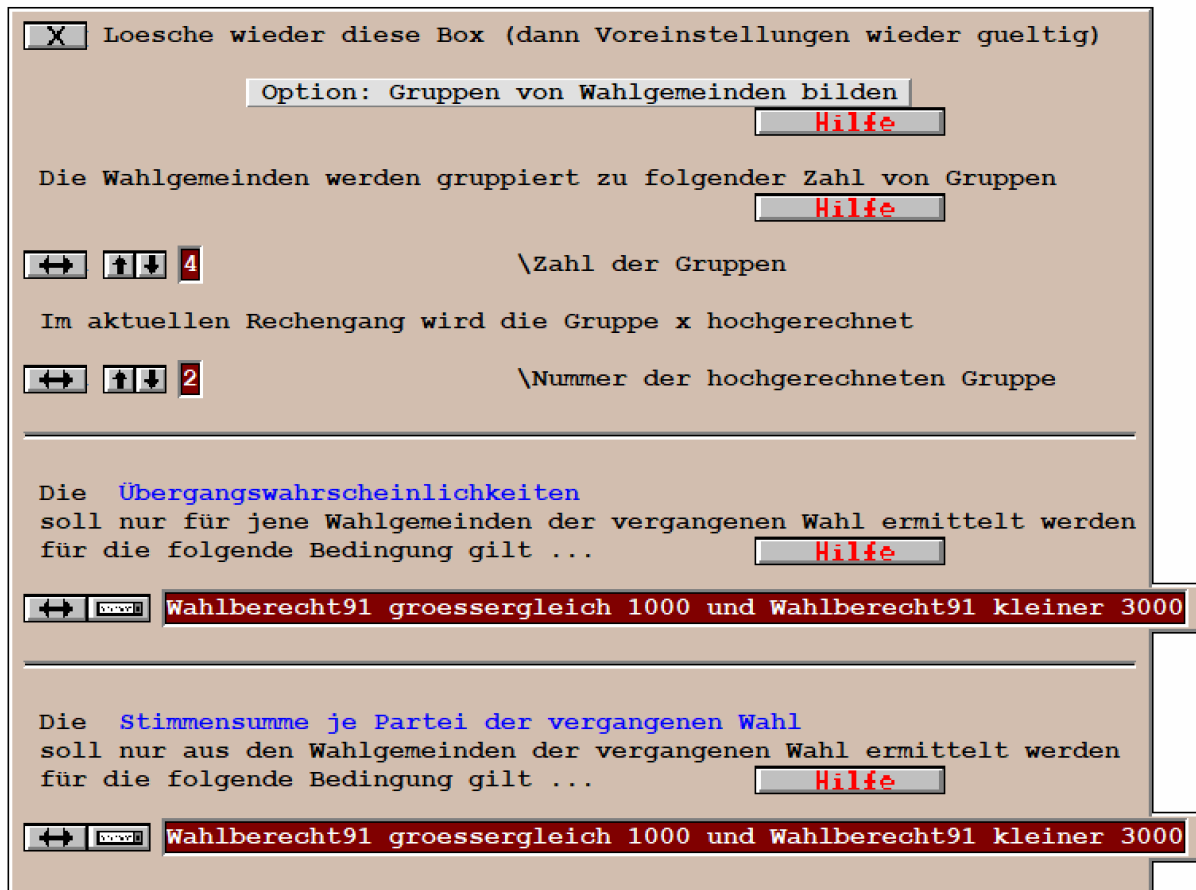
```

in Prog51m1_Grp1:   Wahlberechtigt91 kleiner           1000
in Prog51m1_Grp2:   Wahlberechtigt91 groessergleich  1000 und
                   Wahlberechtigt91 kleiner 30000
in Prog51m1_Grp3:   Wahlberechtigt91 groessergleich  3000 und
                   Wahlberechtigt91 kleiner 80000
in Prog51m1_Grp4:   Wahlberechtigt91 groessergleich  8000

```

Die Auswahl-Bedingung wird in der 2. und 3. Sektion der Eingabebox identisch eingesetzt.

Wir zeigen hier noch die Eingabebox für die 2. Gruppe (siehe Prog51m1_Grp2)



Der Benutzer rechnet zuerst Prog51m1_Grp1 für die 1. Gruppe. Also gibt das Ergebnis aus und erzeugt im Almo-Ordner "Progs" zwei Dateien ("a_gruppd" und "a_gruppstrom"), in welchen die Zwischenergebnisse für das schlussendliche Gesamtergebnis gespeichert werden.

Im ersten Teil der Ausgabe werden wir über die vergangene Wahl informiert (Ausgabe gekürzt)

Zahl der eingelesenen Datensätze: 175

```

OEVP91          58406
SPOE91          21194
FPOE91          16540
VGOE91           1558
GAL91           2089
UngueeltStimm   2105
Nichtw91        10100

```

Gruppe 1 umfasst 175 Gemeinden. Danach werden die Wahlergebnisse der vergangenen Wahl für die Gruppe 1 angegeben.

Wir zeigen jetzt noch die Matrix der Wählerströme für die Gruppe 1 und die Wahlprognose für Gruppe 1. Die Hochrechnung wurde um 19:15 Uhr gerechnet.

Für Gruppe 1 werden 173 Datensätze verarbeitet

Von den 175 Gemeinden der Gruppe 1 waren also um 19.15 Uhr 173 ausgezählt.

Wählerstroeme - Zu- und Abstroeme in Zahl der Stimmen - **Prognose fuer GRUPPE 1** prognostiziert auf der Basis von 173 ausgezaehlten Gemeinden

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungueelt91	Nichtw91	Zuwander	Gesamt
OEVP97	50714	0	0	557	646	1587	0	1836	55340

SPOE97	0	17824	0	165	0	507	0	567	19063
FPOE97	2527	553	14309	0	0	0	764	978	19131
GRUENE97	858	549	29	696	1443	0	0	488	4062
LIF97	118	494	183	140	0	0	68	159	1161
NEUTR97	464	220	33	0	0	0	38	9	764
KPOE97	134	57	0	0	0	0	0	0	191
CWG97	366	24	0	0	0	11	0	0	401
FDP97	131	0	0	0	0	0	0	0	131
OENP97	0	0	90	0	0	0	0	0	90
UnguelSt	1264	465	602	0	0	0	0	62	2394
Nichtw97	1831	647	1293	0	0	0	9230	1025	14026
Abwander	0	361	0	0	0	0	0	0	361
Gesamt	58406	21194	16540	1558	2089	2105	10100	5123	117115

Die Gesamt-Stimmenzahl in der letzten Zeile der Matrix ergibt die tatsächliche Stimmenzahl je Partei in der vergangenen Wahl für Gruppe 1.

Die Gesamt-Stimmenzahl in der letzten Spalte ergibt die prognostizierte Stimmenzahl je Partei in der aktuellen Wahl für Gruppe 1.

Wahlergebnis - Prognose Gruppe 1

Zahl der prognostizierten Stimmen der Parteien aus aktueller Wahl
prognostiziert auf der Basis von 173 ausgezählten Gemeinden

		absolut	in Prozent	nur Parteien in Prozent
V6	OEVP97	55340	47.253	55.156
V7	SPOE97	19063	16.277	19.000
V8	FPOE97	19131	16.335	19.068
V9	GRUENE97	4062	3.469	4.049
V10	LIF97	1161	0.991	1.157
V11	NEUTR97	764	0.652	0.761
V12	KPOE97	191	0.163	0.190
V13	CWG97	401	0.342	0.399
V14	FDP97	131	0.112	0.131
V15	OENP97	90	0.077	0.090
V4	UnguelStimm	2394	2.044	-
V75	Nichtw97	14026	11.976	-
V80	Abwanderer	361	0.308	-
		-----	-----	-----
zusammen		117115	100.000	100.000

Wahlbeteiligung
(prognostiziert) 87.716

Danach rechnet der Benutzer nacheinander die Programme für Gruppe 2 und 3. Also gibt jeweils die Ergebnisse aus. In den beiden Zwischendateien "a_gruppd" und "a_gruppst" werden die Gruppenergebnisse aufsummiert. Abschließend rechnet der Benutzer das Programm für die letzte Gruppe 4. Also erkennt dass dies die letzte Gruppe ist. Es gibt die Ergebnisse für diese Gruppe aus und zusätzlich noch das aufsummierte Gesamtergebnis für die gesamte Wählerpopulation. Dies ist folgendes:

===== Summiertes Endergebnis fuer die 4 Gruppen =====

Summiertes Endergebnis fuer die 4 Gruppen - Prognose fuer aktuelle Wahl

Zahl der prognostizierten Stimmen der Parteien aus aktueller Wahl
prognostiziert auf der Basis von 4 Gruppen von Wahlgemeinden
mit zusammen 432 ausgezählten Wahlgemeinden

		absolut	in Prozent	nur Parteien in Prozent
V6	OEVP97	328501	33.530	42.667
V7	SPOE97	210463	21.482	27.336
V8	FPOE97	157443	16.070	20.449

V9	GRUENE97	43579	4.448	5.660
V10	LIF97	16092	1.643	2.090
V11	NEUTR97	6660	0.680	0.865
V12	KPOE97	2131	0.217	0.277
V13	CWG97	3023	0.309	0.393
V14	FDP97	1188	0.121	0.154
V15	OENP97	843	0.086	0.110
V4	UngueiltStimm	20825	2.126	-
V75	Nichtw97	184937	18.876	-
V80	Abwanderer	4037	0.412	-
		-----	-----	-----
zusammen		979723	100.000	100.000
Wahlbeteiligung (prognostiziert)			80.712	

Wahlerstroeme: Zu- und Abstroeme in Zahl der Stimmen -Prognose fuer das Endergebnis prognostiziert auf der Basis von 4 Gruppen von Wahlgemeinden mit zusammen 432 ausgezählten Wahlgemeinden

	OEVP91	SPOE91	FPOE91	VGOE91	GAL91	Ungueilt91	Nichtw91	Zuwander	Gesamt97
OEVP97	312168	0	2045	2784	646	1587	0	9271	328501
SPOE97	0	202216	0	1255	0	507	0	6485	210463
FPOE97	9305	20917	120121	452	360	1403	764	4120	157443
GRUENE97	7021	7501	1270	5855	17166	840	1427	2500	43579
LIF97	4476	4174	1023	2211	195	258	2278	1478	16092
NEUTR97	3779	1578	156	0	0	1100	38	10	6660
KPOE97	584	1442	0	27	0	0	0	77	2131
CWG97	2265	80	0	0	0	678	0	0	3023
FDP97	852	156	0	49	0	0	0	132	1188
OENP97	446	76	293	0	28	0	0	0	843
UngueiltS	7503	3472	1503	86	0	7650	0	610	20825
Nichtw97	9371	6131	13899	7885	3134	1874	135848	6795	184937
Abwander	0	897	0	0	2721	419	0	0	4037
Gesamt 91	357771	248640	140311	20603	24250	16315	140355	31478	979723

===== Ende summiertes Ergebnis =====

Die Gesamt-Stimmenzahl in der letzten Zeile der Matrix ergibt die tatsächliche Stimmenzahl je Partei in der vergangenen Wahl für das gesamte Wahlgebiet. Die Gesamt-Stimmenzahl in der letzten Spalte ergibt die prognostizierte Stimmenzahl je Partei in der aktuellen Wahl für das gesamte Wahlgebiet.

Nun muss man sich fragen "hat sich der Aufwand für die gruppierte Hochrechnung rentiert?"

Wir vergleichen die prognostizierten Ergebnisse der gruppierten mit der normalen ungruppierten Hochrechnung und mit den tatsächlichen Endergebnissen

	tatsächlich %	gruppiert %	ungruppiert %
OEVP97	42.688	42.667	42.708
SPOE97	27.034	27.336	27.420
FPOE97	20.630	20.449	20.514
GRUENE97	5.781	5.660	5.653
LIF97	2.084	2.090	1.937
NEUTR97	0.813	0.865	0.863
KPOE97	0.281	0.277	0.251
CWG97	0.385	0.393	0.394
FDP97	0.189	0.154	0.173
OENP97	0.114	0.110	0.086
	-----	-----	-----
	100.000	100.000	100.000

Vergleicht man mit dem tatsächlichen Endergebnis der Wahl, so stellt man fest: Die Wahlergebnisse aus der gruppierten Hochrechnung sind für ÖVP, SPÖ und

Grüne minimal besser als die aus der normalen ungruppierten Hochrechnung und für die FPÖ minimal schlechter. Die Unterschiede sind jedoch so gering, dass man konstatieren muss, dass sich der Aufwand für die Gruppierung in unserem Falle nicht gelohnt hat.

Nun wollen wir noch überprüfen, ob die Gruppierung nach der Gemeindegröße einigermaßen homogene Gruppen geschaffen hat. Wir entnehmen den Homogenitätsindex (die *mittlere absolute prozentuelle Abweichung der Stimmenanteile*) den Ergebnislisten aus den 4 Gruppen

	ungruppiert	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4
OEVP97	5.833	5.603	5.089	5.040	4.840
SPOE97	12.457	15.786	9.689	5.797	7.470
FPOE97	11.838	13.831	10.986	8.329	3.942
GRUENE97	29.051	38.312	18.114	12.161	7.583

LIF97	66.404	67.066	37.605	23.784	12.639
NEUTR97	65.242	73.779	52.169	34.454	25.531
KPOE97	61.466	24.477	59.190	48.218	24.251
CWG97	70.662	51.936	89.454	77.910	59.901
FDP97	26.894	17.613	31.817	32.998	24.003
OENP97	13.908	13.915	13.853	17.434	18.625
UnguelStimm	35.668	42.640	26.749	14.790	8.992
Nichtw97	15.044	18.102	10.563	6.563	5.387
Abwanderer	11.929	11.109	8.459	6.799	16.084

zusammen	32.800	30.321	28.749	22.637	16.865

Wir betrachten nur die Parteien, die über der (in Österreich geltenden) 4%-Schwelle liegen. Für Gruppe 2, 3 und 4 ist in allen Fällen der Homogenitätsindex besser als für die (nicht-gruppierte) Gesamtgruppe aller Gemeinden. Dass die Homogenität in Gruppe 1 der Kleinstgemeinden unter 1000 Wahlberechtigten für SPÖ, FPÖ und Grüne schlechter ist als in der Gesamtgruppe aller Gemeinden überrascht zunächst. Möglicherweise muss unterschieden werden zwischen traditionellen Dörfern einerseits und andererseits Gemeinden im Umfeld der Städte mit hohem Pendleranteil. Die Übergangsmatrizen dieser zwei Untergruppen sind wahrscheinlich verschieden.

Gruppierung nach Wähleranteil einer Partei aus der vergangenen Wahl

Es kann sehr sinnvoll sein, nach dem Wahlergebnis für eine Partei der vergangenen Wahl zu gruppieren. So hat das SORA-Institut, das für das österreichische Fernsehen Hochrechnungen durchgeführt hat, die Gemeinden nur nach dem SPÖ-Prozentwert in 5 Gruppen gruppiert. Siehe dazu die Ausführungen von Hofinger/ Ogris (2002) und Ledl (2007, S.46 ff.). Diese Gruppierungsvariable wurde vermutlich verwendet, weil sie stellvertretend für eine Reihe nicht erhobener Hintergrundvariablen steht. Sie korreliert sicher mit der Variablen "Stadt-Land" bzw. mit der Gemeindegröße. Sie versucht aber wohl auch Gemeindegruppen zu bilden, die bezüglich der wirtschaftlichen Struktur verschieden aber in sich homogen sind.

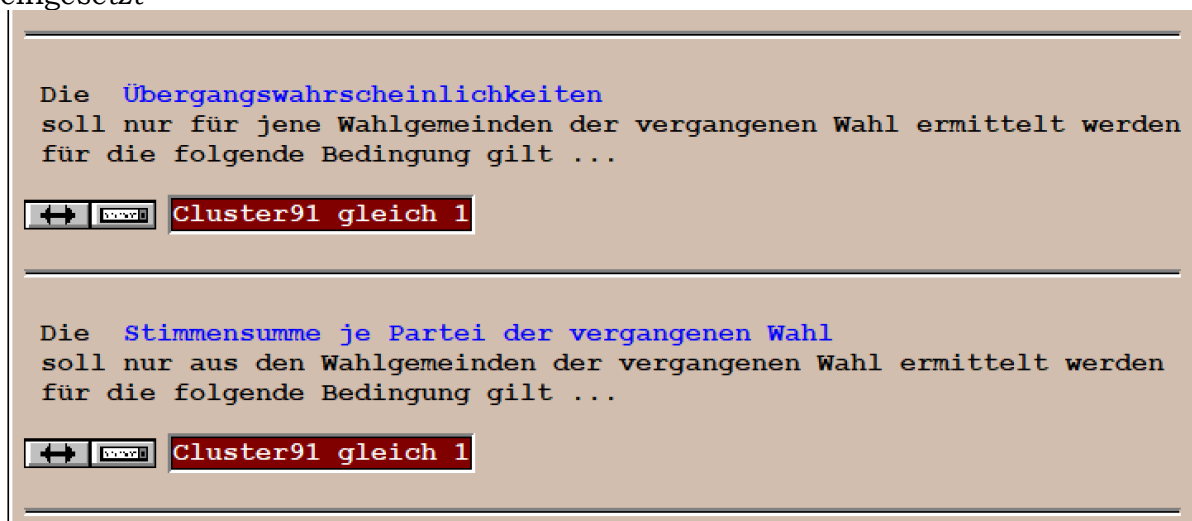
Gruppierung nach Bundesländern

In Österreich kann für die 9 Bundesländer je eine Hochrechnung gerechnet werden. Die je Bundesland prognostizierten Stimmen für die einzelnen Parteien werden summiert. So entsteht eine Hochrechnung für Gesamt-Österreich. Man wird so vorgehen, wenn man unterstellt, dass sich die Wähler der 9 Bundesländer so unterscheiden, dass 9 deutlich verschiedene Übergangsmatrizen entstehen. In

Deutschland hat es sich erwiesen, dass die Übergangsmatrizen der "neuen" Bundesländer sich deutlich von denen der westdeutschen unterscheiden. Eine Gruppierung nach Bundesländern wird die Prognose sicherlich verbessern.

"Geclusterte" Hochrechnung/Wählerstromanalyse

Der Gedanke ist naheliegend, die Gemeinden durch eine vorausgehende Clusteranalyse zu gruppieren. Für jedes Gemeinde-Cluster wird eine selbständige Hochrechnung gerechnet. Die prognostizierten Stimmen aus den Clustern werden je Partei summiert. So entsteht die Prognose für das gesamte Wahlgebiet. Die Clusterung der Gemeinden kann nicht im Rahmen des Hochrechnungsprogramms Prog51m1 geschehen. Das muss vorher geschehen. Man wird so vorgehen: Mit dem "k-means"-Clusterprogramm Prog37m1 oder Prog37m3 werden die Gemeinden der *vergangenen* Wahl zu homogenen Clustern zusammengefasst. Die Clusterzugehörigkeit der einzelnen Gemeinden wird als zusätzliche letzte Variable an den Datensatz angehängt. In den beiden genannten Clusteranalyse-Programmen muss dazu die Optionsbox "Clusterzugehörigkeit der Objekte in Datei speichern" geöffnet werden. Im Hochrechnungs-Programm Prog51m1 wird dann in die Optionsbox 12 eingesetzt



..... und die Hochrechnung gerechnet. Danach wird Prog51m1 für "Cluster91 gleich 2" gerechnet usw. Zum Schluss werden dann die pro Cluster prognostizierten Stimmen parteiweise addiert. Das Wort "Cluster91" ist der vom Benutzer eingegebene Namen für die angehängte zusätzliche Variable.

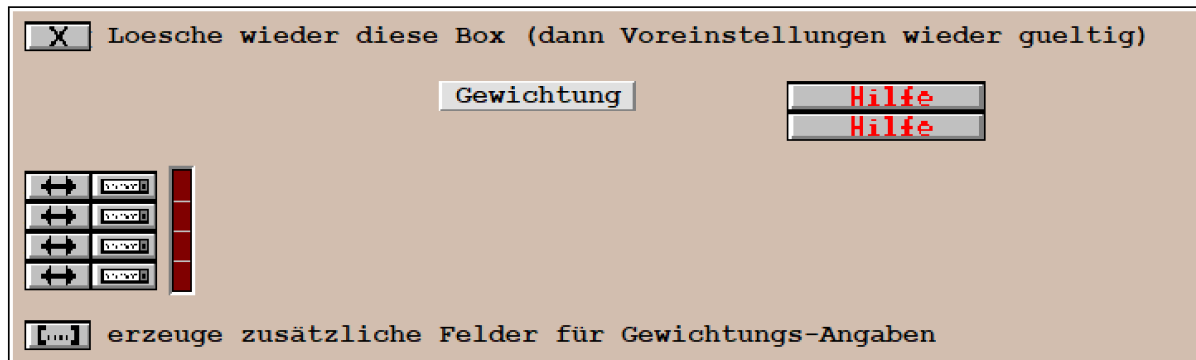
Die entscheidende Frage ist natürlich "welche Klassifikations-Merkmale sollen eingesetzt werden"? Prinzipiell wird die Antwort lauten "solche, die mit dem Wahlverhalten korrelieren - und auch pro Gemeinde als statistische Daten bekannt und zugänglich sind". Das können Strukturmerkmale sein, wie Geschlechterproportion, Alterspyramide, Arbeiter-Anteil, Zahl der Wahlberechtigten usw. Merkmale, in denen sich die Gemeinden wenig unterscheiden (wie vermutlich bei der Geschlechterproportion) sind wenig brauchbar. Naheliegend ist es auch sehr einfach nach den Anteilen der stärksten Parteien in der *vergangenen* Wahl zu clustern. In Österreich wird man hier die Parteien SPÖ, ÖVP, FPÖ und Nichtwähler einsetzen. Ob sich der Aufwand für eine "geclusterte Hochrechnung" rentiert, kann nur im nachhinein und im Vergleich mit der einfachen ungruppierten Hochrechnung festgestellt werden. Eine "geclusterte Wählerstromanalyse" kann später nach der Wahl in aller Ruhe gerechnet werden - wobei diese auch für verschiedene Zeitpunkte des Wahl-Nachmittags (z.B. 16.00 Uhr, 17.00 Uhr usw.) gerechnet wird.

P51.4.9 Gewichtete Hochrechnung

Prog51m1 enthält folgende Eingabebox 15



Wird die Optionsbox geöffnet, dann sieht man folgendes:



Wir haben bereits oben diese Eingabebox 15 "Gewichtung" an folgendem Beispiel kurz erläutert: Zu einem frühen Zeitpunkt der Hochrechnung liegen üblicherweise die Ergebnisse aus den Städten nur sehr unvollständig vor. Nun liegt der Gedanke nahe, die wenigen Städte, deren Ergebnisse vorliegen, so zu gewichten, dass das Verhältnis von Wählerzahl der Städte zu der Wählerzahl der übrigen Gemeinden stimmt.

Was wird gewichtet?

Gewichtet wird der "verknüpfte" Datensatz wie er am Ende von Abschnitt P51.4.2 beschrieben und durch ein kleines Schema illustriert wurde - genauer: gewichtet werden die im "verknüpften" Datensatz enthaltenen *Parteien und Pseudo-Parteien* der aktuellen Wahl und der vergangenen Wahl. Dazu gehören also auch die Nicht- und Ungültigwähler und die Zu- und Abwanderer. Hier nochmals wiederholt:

V1.....V16	V17.....V26	V75 V76 V77 V78 V79 V80
aktuelle Wahl	vergangene Wahl	leer	generierte Variable

Generierte Variable:
V75 V76 =Nichtwähler aus aktueller, vergangener Wahl
V77 V78 =Sonstige aus aktueller, vergangener Wahl
V79 V80 =Zu-, Abwanderer

Alle Gemeinden gewichten

Wird gewichtet, dann müssen korrekterweise alle Gemeinden gewichtet werden. Betrachten wir ein Beispiel: Die Wahlgemeinden werden der Größe nach in 4 Gruppen eingeteilt. Dazu verwenden wir die Datei der vergangenen Wahl

Gruppe 1: Dörfer, Wahlberechtigte bis 1000
Gruppe 2: Gemeinden mit 1000 bis 3000 Wahlberechtigten
Gruppe 3: Gemeinden mit 3000 bis 8000
Gruppe 4: Städte: ab 8000

Die Gruppenbildung

Betrachten wir ein Beispiel: Wir haben die Wahlgemeinden in 4 Gruppen nach der Gemeindegröße (genauer nach der Zahl der Wahlberechtigte) eingeteilt. Um 17.30 Uhr wird eine Hochrechnung gerechnet. Wenn wir unterstellen, dass die zu diesem Zeitpunkt ausgezählten Stimmen eine Zufallsstichprobe aus allen im Wahlgebiet vorhandenen Stimmen darstellen, dann muss der Auszählungsgrad um 17 Uhr 30 in den vier Gruppen annähernd gleich groß sein. Wir wissen aber aus Erfahrung, dass z. B. viele Städte ihre Wahlergebnisse erst sehr viel später an die Zentrale melden können. So ist um 17.30 Uhr der Auszählungsgrad in der Gruppe der Städte sehr viel kleiner als in den anderen Gruppen. Durch entsprechende Gewichtung der vorhandenen Stimmen aus den Städten muss das ausgeglichen werden.

Es gilt also das Prinzip: Wenn Gemeindegruppen sich in ihrem Auszählungsgrad zum Hochrechnungszeitpunkt unterscheiden, dann muss das durch entsprechende Gewichtung ausgeglichen werden. Der Zweck der gewichteten Hochrechnung ist es, die unterschiedlichen Auszählungsgrade der Gemeindegruppen auszugleichen.

In unserem Beispiel wird aber das vermutliche Wahlergebnis nur dann besser prognostiziert, wenn die Wähler in den Städten sich anders verhalten als die Wähler der anderen Gruppen, genauer formuliert, wenn die Gruppe der Städte eine andere Übergangsmatrix besitzt als die anderen Gruppen. Das kann aber bei der gewichteten Hochrechnung nur im nachhinein durch eine "Hochrechnung für eine einzelne Gemeindegruppe" (siehe Abschnitt P51.4.8.1) überprüft werden. Letzten Endes ist es der Hochrechner bzw. Wahlforscher, der auf Grund seiner Kenntnis der Wahlpopulation und der Erfahrungen aus vergangenen Wahlen, entscheidet, ob eine bestimmte Gruppenbildung sich rentiert. Es muss aber festgehalten werden, dass die Prognose nicht systematisch schlechter wird, wenn Gruppen gebildet werden, die zwar unterschiedliche Auszählungsgrade zum Hochrechnungszeitpunkt besitzen, aber gleiche Übergangsmatrizen besitzen. Die gewichtete Hochrechnung kann nicht schlechter werden als die ungewichtete.

Die Frage lautet nun: Welche Variable oder Variablenkombinationen sollen zur Bildung der Gruppen verwendet werden?

Es müssen Variable sein, die für den unterschiedlichen Auszählungsgrad zum Hochrechnungs-Zeitpunkt verantwortlich sind. In erster Linie ist dies wohl die Gemeindegröße, genauer die Zahl der Wahlberechtigten in der vergangenen Wahl, möglich wäre aber auch die Zahl der abgegebenen Stimmen in der vergangenen Wahl. Für unser Beispiel der Landtagswahl kommen demgemäß die Variablen `Wahlberecht91` oder `AbgegebenStimm91` in Frage.

Unterschied zur "gruppierten" Hochrechnung

Bei der in Abschnitt P51.4.8 behandelten "*gruppierten*" Hochrechnung ist die Logik der Gruppenbildung eine andere. Bei ihr sollten sich die Gruppen in den Übergangswahrscheinlichkeiten deutlich unterscheiden. Der Auszählungsgrad zum Hochrechnungszeitpunkt ist irrelevant.

Die *gruppierte* Hochrechnung wird eingesetzt, wenn Gruppen mit unterschiedlichen Übergangsmatrizen vorliegen, die *gewichtete* Hochrechnung, wenn die Gruppen zum Hochrechnungszeitpunkt unterschiedliche Auszählungsgrade besitzen.

In Bezug auf den Hochrechnungs-Kalkül ist der Unterschied folgender: Bei der *gruppierten* Hochrechnung werden je Gruppe eigene Matrizen der Übergangswahrscheinlichkeiten errechnet und in der Folge eigene Wahlergebnisse je Gruppe prognostiziert. Die Prognose für das gesamte Wahlgebiet entsteht dann aus der Addition der prognostizierten Gruppenergebnisse.

Bei der *gewichteten* Hochrechnung wird aus den unterschiedlich gewichteten Gruppenergebnissen *eine gemeinsame* Matrix der Übergangswahrscheinlichkeiten errechnet und daraus das Wahlergebnis für das gesamten Wahlgebiet abgeleitet.

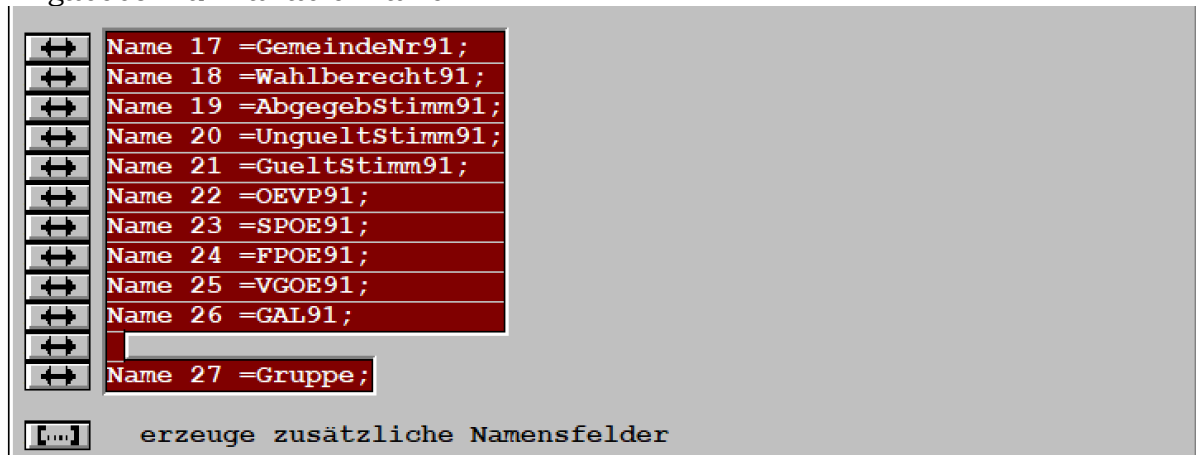
P51.4.9.1 Berechnung der Gewichtszahlen

Die gewichtete Hochrechnung soll im Folgenden an einem Rechenbeispiel vorgeführt werden. Wir müssen allerdings vorausschicken, dass unsere Daten der Landtagswahl sich für eine gewichtete Hochrechnung wenig eignen. So sind in unseren Daten nur 4 Städte mit mehr als 20 000 Einwohner vorhanden. Wir haben deswegen für unser Rechenbeispiel schon Gemeinden ab 8000 Wahlberechtigten als Städte klassifizieren müssen, um genügend Gemeinden der Größenordnung "Stadt" zu erhalten. Die Ergebnisse, die unser nachfolgendes Rechenbeispiel liefert, sind deswegen kaum besser als die einer normalen ungewichteten Hochrechnung.

1. Rechenschritt:

Im ersten Rechenschritt berechnen wir mit der Programm-Maske **Prog18m2_vergangen** die prozentuelle Verteilung der Wahlberechtigten auf die 4 Gemeindegruppen. Das Programm wird gefunden nach Klick auf den Knopf "alle Progs" am Oberrand des Almo-Fensters. Wir verwenden die Datei der vergangenen Wahl. Aus der Programm-Maske zeigen wir nur folgende zwei Eingabeboxen:

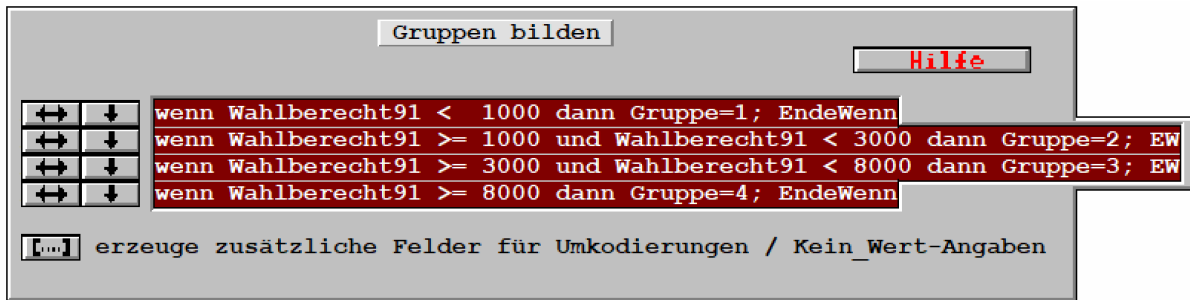
Eingabebox für Variablennamen



Die Gruppenzugehörigkeit der Gemeinden ist in der Datei der vergangenen Wahl als Variable nicht enthalten. Wir generieren dafür eine Variable. Dazu verwenden wir die freie (sonst nicht gebrauchte) Variable 27 und weisen ihr in obiger Eingabebox den Namen "Gruppe" zu.

Gebraucht werden eigentlich nur die beiden Namen `Wahlberecht91` und `Gruppe`. Die anderen Namen müssten nicht geschrieben werden.

Die Gemeindegruppen werden dann in "Prog18m2_vergangen" in folgender Eingabebox gebildet



Das Wort "EW" am Ende des 2. und 3. Eingabefelds ist eine zulässige Kurzform von "EndeWenn". Das Programm liefert folgende Ausgabe (gekürzt)

Ergebnisse für vergangene Wahl

Zahl der Wahlgemeinden je Gruppe

Gruppe 1	175
Gruppe 2	212
Gruppe 3	47
Gruppe 4	11

445

Wahlberechtigte je Gruppe

Gruppe 1	111992
Gruppe 2	346980
Gruppe 3	192459
Gruppe 4	296814

948245

Prozent je Gruppe

Gruppe 1	11.81
Gruppe 2	36.59
Gruppe 3	20.30
Gruppe 4	31.30

100.00

Diese Prozentzahlen müssen bekannt sein, bevor hochgerechnet wird. "Prog18m2_vergangen" muss also vor dem Hochrechnungsprogramm Prog51m1 gerechnet werden.

2. Rechenschritt:

Um 17.30 Uhr soll eine Hochrechnung gerechnet werden. Zuvor muss nun mit der Programm-Maske **Prog18m2_aktuell** die Prozentverteilung auf die 4 Gemeindegruppen zum Zeitpunkt 17.30 Uhr ermittelt werden. Diese Programm-Maske wird gefunden nach Klick auf den Knopf "alle Progs" am Oberrand des Almo-Fensters.

Ausgewertet wird jetzt die bis zum Zeitpunkt 17.30 Uhr vorhandene "verknüpfte" Datei, wie sie oben am Beginn des Abschnitts schematisch dargestellt ist. Die verknüpfte Datei besteht aus den *ausgezählten* Gemeinden aus der aktuellen Wahl und den über die Gemeindenummer als Verbindungsvariable angefügten gleichen Gemeinden aus der vergangenen Wahl. Deswegen müssen in die Eingabebox für die Namen auch die der aktuellen Wahl und der vergangenen Wahl eingetragen werden. Wieder muss eine Variable mit dem Namen "Gruppe" generiert werden. Dafür wird die freie Variable 17 verwendet.

Eingabeboxen für Variablennamen der aktuellen Wahl

Namen der Variablen aus der aktuellen Wahl Hilfe

in der Reihenfolge, in der sie in der Datei hintereinander stehen

Leere alle Eingabefelder dieser Sub-Box

<input type="checkbox"/>	Name 1 =GemeindeNr97;
<input type="checkbox"/>	Name 2 =Wahlberecht97;
<input type="checkbox"/>	Name 3 =AbgegebenStimm97;
<input type="checkbox"/>	Name 4 =UnguelteStimm97;
<input type="checkbox"/>	Name 5 =GuelteStimm97;
<input type="checkbox"/>	Name 6 =OEVP97;
<input type="checkbox"/>	Name 7 =SPOE97;
<input type="checkbox"/>	Name 8 =FPOE97;
<input type="checkbox"/>	Name 9 =GRUENE97;
<input type="checkbox"/>	Name 10 =LIF97;
<input type="checkbox"/>	Name 11 =NEUTR97;
<input type="checkbox"/>	Name 12 =KPOE97;
<input type="checkbox"/>	Name 13 =CWG97;
<input type="checkbox"/>	Name 14 =FDP97;
<input type="checkbox"/>	Name 15 =OENP97;
<input type="checkbox"/>	Name 16 =Zeit97;

erzeuge weitere Namensfelder für Variable der aktuellen Wahl

Eingabebox für Variablennamen der vergangenen Wahl

Namen der Variablen aus der vergangenen Wahl Hilfe

in der Reihenfolge, in der sie in der Datei hintereinander stehen
Numerisierung muss an letzte Nummer aus aktueller Wahl anschliessen

<input type="checkbox"/>	Name 17 =GemeindeNr91;
<input type="checkbox"/>	Name 18 =Wahlberecht91;
<input type="checkbox"/>	Name 19 =AbgegebenStimm91;
<input type="checkbox"/>	Name 20 =UnguelteStimm91;
<input type="checkbox"/>	Name 21 =GuelteStimm91;
<input type="checkbox"/>	Name 22 =OEVP91;
<input type="checkbox"/>	Name 23 =SPOE91;
<input type="checkbox"/>	Name 24 =FPOE91;
<input type="checkbox"/>	Name 25 =VGOE91;
<input type="checkbox"/>	Name 26 =GAL91;

erzeuge weitere Namensfelder für Variable der vergangenen Wahl

Eingabebox für Namen der generierten Variablen

Generierte Variable Hilfe

die nicht in den Dateien der beiden Wahlen enthalten sind

<input type="checkbox"/>	Name 27 =Gruppe;
--------------------------	------------------

erzeuge zusätzliche Namensfelder

Da wir in unserem Rechenbeispiel die Hochrechnung aus den vollständig vorhandenen Daten der aktuellen Wahl simulieren, müssen wir in "Prog18m2_aktuell" in der Eingabebox für "Ein- und Ausschliessen" eintragen:

Ein- und Ausschliessen von Untersuchungseinheiten

Untersuchungseinheiten in Analyse **Hilfe**
 nur **einschließen**
 wenn

Zeit97 kleiner 1730

Untersuchungseinheiten aus Analyse **Hilfe**
ausschließen
 wenn

Die Eingabebox in "Prog18m2_aktuell" für die Gruppenbildung ist folgende:

Gruppen bilden **Hilfe**

wenn Wahlberecht91 < 1000 dann Gruppe=1; EndeWenn
 wenn Wahlberecht91 >= 1000 und Wahlberecht91 < 3000 dann Gruppe=2; EW
 wenn Wahlberecht91 >= 3000 und Wahlberecht91 < 8000 dann Gruppe=3; EW
 wenn Wahlberecht91 >= 8000 dann Gruppe=4; EndeWenn

erzeuge zusätzliche Felder für Umkodierungen / Kein_Wert-Angaben

Die Eingabe ist dieselbe wie bei "Prog18m2_vergangen".

Das Programm liefert folgendes Ergebnis (gekürzt):

Ergebnisse für aktuelle Wahl. Auszählung: 17.30 Uhr

Ausgezaehlte Wahlgemeinden je Gruppe

Gruppe 1	167
Gruppe 2	158
Gruppe 3	19
Gruppe 4	3

	347

Wahlberechtigte je Gruppe

Gruppe 1	105679
Gruppe 2	250186
Gruppe 3	73980
Gruppe 4	29960

	459805

Prozent je Gruppe

Gruppe 1	22.98
Gruppe 2	54.41
Gruppe 3	16.09
Gruppe 4	6.52

	100.00

Von den bis 17.30 Uhr eingetroffenen Wahlergebnissen stammen aus Gruppe 1 (den Dörfern) 20.65% der ausgezählten Wahlberechtigten. Die Dörfer umfassen jedoch nur 11.81% aller Wahlberechtigten (wie aus den Ergebnissen von "Prog18m2_vergangen" ersichtlich wurde). Um 17.30 Uhr sind also die Dörfer überrepräsentiert. Das muss durch entsprechende Gewichtung ausgeglichen

werden. Die Gewichtungszahl dafür ist

$$g1 = 11.81/22.98 = 0.51$$

Für alle 4 Gruppen entstehen so folgende Gewichtungszahlen

Gruppe 1 $g1 = 11.81/22.98 = 0.51$
 Gruppe 2 $g2 = 36.59/54.41 = 0.67$
 Gruppe 3 $g3 = 20.30/16.09 = 1.26$
 Gruppe 4 $g4 = 31.30/6.52 = 4.80$

Um 17 Uhr 30 sind also besonders die Städte (Gruppe 4) unterrepräsentiert. Dies muss ausgeglichen werden. So werden die Stimmen für die Parteien in den um 17.30 Uhr ausgezählten 3 Städten mit 4.80 multipliziert.

3. Rechenschritt:

In der Eingabebox 15 "Gewichtung" des Hochrechnungsprogramms **Prog51m1** muss dann folgendes eingetragen werden. Dann wird Prog51m1 gerechnet.

Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)

Gewichtung Hilfe

Hilfe

Wenn Wahlberecht91 < 1000 dann Gewicht0=11.81/22.98; EndeWenn

Wenn Wahlberecht91 >=1000 und Wahlberecht91 <3000 dann Gewicht0=36.59/54.41; EW

Wenn Wahlberecht91 >=3000 und Wahlberecht91 <8000 dann Gewicht0=20.30/16.09; EW

Wenn Wahlberecht91 >=8000 dann Gewicht0=31.30/6.52; EndeWenn

erzeuge zusätzliche Felder für Gewichtungs-Angaben

Die so ausgefüllte Programm-Maske ist in Almo unter dem Namen **Prog51m1 gewichtet** enthalten. Sie wird gefunden durch Klick auf den Knopf "Verfahren" dann "Wahlhochrechnung" oder auf den Knopf "alle Progs" am Oberrand des Almo-Fensters. Das Programm ist identisch mit dem allgemeinen Hochrechnungsprogramm Prog51m1, mit der einer Ausnahme, dass die Optionsbox für die Gewichtung geöffnet wurde und, wie oben gezeigt, ausgefüllt wurde.

Das Programm liefert folgendes Ergebnis. Wir zeigen nur das prognostizierte Wahlergebnis und geben noch zum Vergleich das Ergebnis der normalen ungewichteten Hochrechnung und das tatsächliche Endergebnis an.

Zahl der prognostizierten Stimmen der Parteien aus aktueller Wahl in Prozent

		gewichtet nur Parteien in Prozent	tatsächlich nur Parteien in Prozent	ungewichtet nur Parteien in Prozent
V6	OEVP97	42.756	42.688	42.949
V7	SPOE97	26.879	27.034	27.263
V8	FPOE97	20.392	20.630	20.449
V9	GRUENE97	6.022	5.781	5.781
V10	LIF97	2.134	2.084	1.870
V11	NEUTR97	0.863	0.813	0.838
V12	KPOE97	0.234	0.281	0.202
V13	CWG97	0.402	0.385	0.403
V14	FDP97	0.132	0.189	0.146
V15	OENP97	0.186	0.114	0.099
V4	Ungült.Stimmen97	-	-	-
V75	Nichtwähl97	-	-	-
V80	Abwanderer	-	-	-
		-----	-----	-----
zusammen		100.000	100.000	100.000

Die gewichtete Analyse ist teilweise besser aber auch teilweise schlechter als die

ungewichtete Analyse. Auch ist die Differenz zum tatsächlichen Ergebnis vernachlässigbar gering. Die Gewichtung hat sich für unsere Daten nicht rentiert. Wir haben bereits darauf hingewiesen, dass sich unsere Daten der Landtagswahl nur wenig für eine gewichtete Hochrechnung eignen. So sind nur 4 Städte mit mehr als 20 000 Einwohner vorhanden. Wir haben deswegen in unserem Beispiel schon Gemeinden ab 8000 Wahlberechtigten als Städte klassifizieren müssen um eine ausreichende Anzahl von Städten zu erhalten. Gemeinden mit einer Einwohnerzahl um 8000 besitzen aber wohl nicht die Übergangsmatrix, die für echte Städte charakteristisch ist.

Drei Anmerkungen sind noch notwendig:

1. In unserem Beispiel werden die Gemeindegruppen gemäß den Wahlberechtigten der vergangenen Wahl gebildet. Wenn eine vollständige und aktuelle Datei der Wahlberechtigtenzahl aller Gemeinden verfügbar ist, dann kann in "Prog18m2_vergangen" diese Datei anstelle der Datei der vergangenen Wahl eingetragen werden. In den 3 Programmen "Prog18m2_vergangen", "Prog18m2_aktuell" und "Prog51m1_gewichtet" kann dann mit der aktuellen Wahlberechtigtenzahl `Wahlberecht97` gerechnet werden.

2. Die Gruppenbildung kann von einem Hochrechnungszeitpunkt zum nächsten geändert werden. Dann muss selbstverständlich auch "Prog18m2_vergangen" neu gerechnet werden - was nicht notwendig wäre, wenn die Gruppenbildung unverändert bleibt.

3. Die Gewichtung könnte weiter verfeinert werden, wenn die möglicherweise unterschiedlichen Wahlbeteiligungen in den Gemeindegruppen mit eingerechnet werden. Das würde dadurch erreicht, dass nicht die Zahl der Wahlberechtigten zur Gruppierung verwendet wird sondern die Zahl der abgegebenen Stimmen, also die Variable `V19 AbgegebenStimm91`

P51.5 Das Modell

Das "Grundmodell" der Wählerstromanalyse/Hochrechnung wird durch die beiden Bedingungen (1) und (2) aus Abschnitt P51.1.3 gebildet. Das heißt, es werden nur die dort beschriebenen Restriktionen bei der Schätzung berücksichtigt. Diese sind:

1. Jede Übergangswahrscheinlichkeit (Abstromwahrscheinlichkeit) soll größer/gleich 0 bzw. 0% und kleiner/gleich 1 bzw. 100% sein.
2. Die Summe der Übergangswahrscheinlichkeiten einer Partei zum Zeitpunkt t_1 zu den Parteien zum Zeitpunkt t_2 soll gleich 1 bzw. 100% sein.

Wenn wir mit $p(i,j)$ die Übergangswahrscheinlichkeit von der Partei j zum Zeitpunkt t_1 zur Partei i zum Zeitpunkt t_2 bezeichnen, so lassen sich diese beiden Restriktionen ausdrücken als:

$$1. \quad 0 \leq p(i, j) \leq 1 \quad \text{für alle } i, j$$

$$2. \quad \sum_i p(i, j) = 1 \quad \text{für alle } j$$

Wir nennen dies das *Grundmodell der Hochrechnung bzw. Wählerstromanalyse*.

Davon, dass die Ergebnisse diese beiden Restriktionen erfüllen, kann man sich an Hand der Ergebnisse überzeugen. Die in Abschnitt P51.4.5 mit der Überschrift "Matrix der Uebergangswahrscheinlichkeiten in Prozent" ausgegebene Matrix enthält die Übergangswahrscheinlichkeiten als Spaltenprozente. Alle Werte liegen in dem Intervall zwischen 0 und 100%. Die Spaltensummen sind 100%.

Für die weiteren Ausführungen soll ein einfacheres (fiktives) Beispiel verwendet werden: Zum ersten Zeitpunkt t_1 (=1. Wahl) sollen zwei Parteien vorliegen, die mit APARTEI1 und BPARTEI1 bezeichnet sind. Zum zweiten Zeitpunkt t_2 (=2. Wahl) kandidieren drei Parteien, die mit APARTEI2, BPARTEI2 und CPARTEI2 bezeichnet werden. CPARTEI2 ist also eine neue Partei.

Der Benutzer findet das Programm dazu nach Klick auf den Knopf "alle Progs" am Oberrand des Almo-Fensters (1) als Maskenprogramme mit den Namen **P51Bspl.Alm**. Dies ist das Maskenprogramm Prog51m1, ausgefüllt für das hier verwendete fiktive Beispiel. Im Programm wird mit "Modell=1" und ohne Zu- und Abwanderer gerechnet. Wir rechnen teilweise auch mit **Prog28m1** (das ebenfalls durch Klick auf den Knopf "alle Progs" geladen wird). Das ist die für unsere Beispieldaten ausgefüllte Programm-Maske Prog28m2 (die nur für die Wählerstromanalyse bei vollständiger Auszählung verwendet werden kann). Die beiden Programme liefern eine weitgehend gleiche Ergebnis-Ausgabe.

Im Beispiel wird eine Hochrechnung bei vollständiger Auszählung gerechnet - d.h. eine *Wählerstromanalyse*. Dadurch habe wir die Möglichkeit zu überprüfen, ob unser Modell das richtige Ergebnis reproduziert.

Das "Grundmodell" der Wählerstromanalyse/Hochrechnung liefert folgendes Ergebnis:

Unabhaengige Variable (Variable der 1. Stufe)
V1 APartei1
V2 BPartei1

Abhaengige Variable (Variable der 2. Stufe)
V3 APartei2
V4 BPartei2
V5 CPartei2

=====
Es werden 61 Datensaeetze verarbeitet
=====

Matrix der Uebergangswahrscheinlichkeiten in Prozent
(die Summe der Werte in einer Spalte ist 100 %)

		AParteil	BParteil
		V1	V2
APartei2	V3	88.4476	0
BPartei2	V4	0	96.7413
CPartei2	V5	11.5524	3.2587
		100.0000	100.0000

Die berechneten Übergangswahrscheinlichkeiten $p(i,j)$ erfüllen die beiden Restriktionen. Alle Werte sind kleiner/gleich 0% bzw. größer/gleich 100%. Die Spaltensummen sind 100%, wie man leicht nachrechnen kann. Für die Partei APARTEI1 der 1. Wahl ergibt sich folgende Spaltensumme: $88.4476 + 0 + 11.5524 = 100.00\%$. 88% der Wähler der Partei APARTEI1 der 1. Wahl haben bei der 2. Wahl

die Partei APARTEI2 gewählt, 12% die CPARTEI2.

Diese Wählerströme werden von ALMO auch in Absolutzahlen ausgegeben:

Wählerstroeme (Abstroeme)
(bezogen auf Stimmen der Parteien aus der 1. Wahl)

		AParteil BParteil		
		V1	V2	
APartei2	V3	2460	0	2460
BPartei2	V4	0	2413	2412
CPartei2	V5	321	81	402
		2781	2494	5275

"Bezogen auf Stimmen der Parteien aus der 1. Wahl" bedeutet, dass die Spaltensummen den Gesamtstimmen der Parteien der 1. Wahl entsprechen. Die Gesamtzahl der Stimmen ist daher die Gesamtstimmenzahl der 1. Wahl, siehe nachfolgende Ergebnisausgabe.

Zahl der Stimmen der Parteien aus 1. Wahl			
		absolut	in Prozent
V1	AParteil	2781	52.720
V2	BParteil	2494	47.280
zusammen		5275	100.000

Zahl der vom Modell reproduzierten Stimmen der Parteien aus 2. Wahl			
		absolut	in Prozent
V1	A_Partei_2	2460	46.630
V2	B_Partei_2	2413	45.739
V3	C_Partei_2	403	7.631
zusammen		5276	100.000

Anmerkung: Der Kalkül in Prog51ml reproduziert die Stimmenzahl je Partei als Dezimalzahl. Die Stimmenzahl kann aber nur ganzzahlig sein. Durch das notwendige Runden können dann tatsächliche und reproduzierte Stimmenzahl minimal differieren. Das ist in unserem Beispiel geschehen. Die Spaltensumme müsste 5275 sein.

Vergleichen wir nun diese Ergebnisse mit der tatsächlichen Stimmenverteilung aus der 2. Wahl.

tatsächliche Zahl der Stimmen der Parteien aus 2. Wahl			
		absolut	in Prozent
V1	A_Partei_2	2531	46.629
V2	B_Partei_2	2482	45.726
V3	C_Partei_2	415	7.646
zusammen		5428	100.000

Es zeigt sich folgendes:

1. Die vom Grundmodell reproduzierte Prozentverteilung stimmt mit der tatsächlichen (wenn auch nur sehr minimal) nicht überein.
2. Die vom Modell reproduzierte Stimmenzahl je Partei stimmt mit der tatsächlichen nicht überein.
3. Die vom Modell reproduzierte Stimmensumme entsprechen der tatsächlichen Stimmensumme aus der 1. Wahl und nicht der 2. Wahl

Aufgrund des Modells werden beispielsweise für die Partei APARTEI2 2460 Stimmen (=46.630%) vorausgesagt. Die Partei APARTEI2 hat tatsächlich aber 2531 Stimmen (46.629%) erhalten.

Eine exakte Reproduktion der Stimmverteilung der 2. Wahl kann durch folgende zwei Maßnahmen erreicht werden:

1. In das Modell der Wählerstromanalyse wird zusätzlich die Nebenbedingung aufgenommen, dass die relative Stimmverteilung der 2. Wahl durch das Modell reproduziert wird. Formal ausgedrückt bedeutet diese Nebenbedingung

$$\sum_j p(i, j) \cdot n(j, t_1) = r(i) \cdot n(t_1),$$

wobei

i = Partei aus der 2. Wahl

j = Partei aus der 1. Wahl

p(i, j) = Übergangswahrscheinlichkeit zu Partei i in der zweiten Wahl von Partei j aus der 1. Wahl

n(j, t₁) = Zahl der Stimmen der Partei j in der 1. Wahl

r(i) = Stimmenanteil der Partei i in der zweiten Wahl

n(t₁) = Gesamtzahl der Stimmen aus der 1. Wahl.

In unserem Beispiel ist n(j=1, t₁) gleich 2781 (=die Stimmenzahl der Partei APARTEI_1), n(t₁) ist gleich 5275 (=Gesamtzahl der Stimmen aus der 1. Wahl) und r(i=1) ist gleich 0.46626 bzw. 46.626% (=der Stimmenanteil der Partei APARTEI_2).

Diese zusätzliche Nebenbedingung wird beim Maskenprogramm Prog51m1 (und auch bei Prog28m1) dadurch eingeführt, dass in der Eingabebox „Modell-Optionen“ das 1. Eingabefeld auf „2“ gesetzt wird. In unserem Beispiel führt diese Anweisung dann zu folgendem Ergebnis:

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Abstrome
(die Summe der Werte in einer Spalte ist 100 %)

		A_Partei_1	B_Partei_1
		V1	V2
A_Partei	V3	88.4451	0
B_Partei	V4	0	96.7137
C_Partei	V5	11.5549	3.2863
		100.00	100.00

Die Übergangsmatrix ist etwas verändert.

Wahlerstroeme - Abstrome in Zahl der Stimmen

		A_Partei_1	B_Partei_1	
		V1	V2	
A_Partei	V3	2460	0	2460
B_Partei	V4	0	2412	2412
C_Partei	V5	321	82	403
		2781	2494	5275

Die Abströme in Stimmen scheinen unverändert zu sein. Tatsächlich sind jedoch Änderungen vorhanden, die allerdings nur sichtbar würden, wenn die im Kalkül errechneten Dezimalwerte für die Ausgabe nicht gerundet würden. Wir haben die Daten für unser Beispiel etwas unglücklich gewählt. Bei den echten Daten aus der Landtagswahl, die wir in die Programm-Maske Prog51m1 eingesetzt haben, sind die Abströme für Modell=1 und Modell=2 (ohne Zu- und Abwanderer) deutlich verschieden.

Zahl der vom Modell reproduzierten Stimmen der Parteien aus 2. Wahl		
	absolut	in Prozent
A_Partei_2	2460	46.629
B_Partei_2	2412	45.726
C_Partei_2	403	7.646
	-----	-----
zusammen	5275	100.000

Dies ist nun die wesentliche Verbesserung, die durch die 3. Nebenbedingung erzeugt wurde: Die prognostizierte (relative) Stimmenverteilung für die 2. Wahl stimmt mit der empirischen Stimmenverteilung überein. Der prognostizierte Stimmenanteil für die Partei APARTEI2 ist gleich dem tatsächlichen Stimmanteil von 46.629%, jener für die Partei BPARTEI2 gleich dem tatsächlichen Stimmanteil von 45.726% usw.

Die absoluten Stimmzahlen stimmen dagegen nicht überein, da zu beiden Zeitpunkten eine unterschiedlich große Wahlpopulation vorliegt. In der 1. Wahl betrug diese 5275, in der 2. Wahl 5428. Nur wenn zu beiden Zeitpunkten die Zahl der Wähler gleich groß ist, führt die zusätzliche Nebenbedingung dazu, dass auch die Gesamtstimmen der Parteien der 2. Wahl reproduziert werden.

2. Durch die Einführung von zwei "neuen" Variablen kann man nun erreichen, dass die Wahlpopulation zu beiden Zeitpunkten identisch ist. Diese beiden Variablen sind "Zuwanderungsgewinn" und "Abwanderungsverlust" und werden wie folgt gebildet:

Zuwanderungsgewinn: Diese Variable ist gleich 0, wenn die Wahlpopulation in einer Gemeinde abgenommen hat, sonst ist sie gleich der Differenz aus den Wählern der 2. Wahl und der 1. Wahl. Bezeichnen wir mit $n(k, t_1)$ die Zahl der Wähler der ersten Wahl in der Gemeinde k und mit $n(k, t_2)$ jene der zweiten Wahl, so ist die Variable Zuwanderungsgewinn definiert als:

$$\text{Zuwanderungsgewinn} = \begin{cases} 0 & \text{wenn } n(k, t_2) \leq n(k, t_1) \\ n(k, t_2) - n(k, t_1) & \text{sonst} \end{cases}$$

Abwanderungsverlust: Diese Variable ist - im Unterschied zur Variablen Zuwanderungsgewinn - gleich 0, wenn die Wahlpopulation in einer Gemeinde zugenommen hat, sonst ist sie gleich der Differenz aus den Wählern der 1. Wahl und der 2. Wahl. Formal ausgedrückt ergibt sich folgende Definition:

$$\text{Abwanderungsverlust} = \begin{cases} 0 & \text{wenn } n(k, t_1) \leq n(k, t_2) \\ n(k, t_1) - n(k, t_2) & \text{sonst} \end{cases}$$

Die Berechnung dieser beiden Variablen erfolgt programm-intern. Wird Prog51m1 für eine Hochrechnung (bei unvollständiger Auszählung) verwendet, dann muss der Benutzer die Gesamtzahl der Zuwanderer in Eingabebox 10

eintragen. Wird Prog51m1 (bei vollständiger Auszählung) für eine Wählerstromanalyse oder wird Prog28m2 verwendet, dann entfällt diese Eingabe.

Die Variable `Abwanderer` wird als abhängige Variable in das Modell aufgenommen, die Variable `Zuwanderer` als unabhängige Variable. Ihre Interpretation hängt davon ab, ob in dem Modell andere demographische Veränderungen der Wahlpopulation, wie z.B. Jungwähler und Sterbefälle berücksichtigt werden. Ist dies nicht der Fall, kann nur von einem allgemeinen Zuwanderungsgewinn bzw. Abwanderungsverlust gesprochen werden. Ein Zuwanderungsgewinn kann beispielsweise bedeuten, dass die Jungwähler die Sterbefälle überwiegen oder dass zwar die Sterbefälle die Jungwähler überwiegen, die Zuwanderungen aber dieses Defizit und die Abwanderungen kompensieren.

In unserem Beispiel führt die Verwendung des Modells 2 und der beiden zusätzlichen Variablen `Zuwanderer` und `Abwanderer` zu folgenden Ergebnissen:

Ergebnisse aus ALMO

 Unabhaengige Variable (Variable der 1. Stufe)

V1 AParteil
 V2 BParteil
 V6 Zuwander

Abhaengige Variable (Variable der 2. Stufe)

V3 APartei2
 V4 BPartei2
 V5 CPartei2
 V7 Abwander

=====
 Es werden 61 Datensaeetze verarbeitet
 =====

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Abstroeme
 (die Summe der Werte in einer Spalte ist 100 %)

		AParteil	BParteil	Zuwandere		
		V1	V2	V6		
APartei2	V3	88.2830	0	44.8820		
BPartei2	V4	0	96.4753	44.9145		
CPartei2	V5	11.1417	3.5247	10.2035		
Abwander	V7	0.5753	0	0		

Waehlerstroeme - Abstroeme in Zahl der Stimmen

		AParteil	BParteil	Zuwander		
		V1	V2	V6		
APartei2	V3	2455	0	76	2531	
BPartei2	V4	0	2406	76	2482	
CPartei2	V5	310	88	17	415	
Abwander	V7	16	0	0	16	
		2781	2494	169	5444	

Matrix der Uebergangswahrscheinlichkeiten in Prozent - Zustroeme
 (die Summe der Werte in einer Zeile ist 100 %)

		AParteil	BParteil	Zuwander
		V1	V2	V6
APartei2	V3	97.00	0	3.00
BPartei2	V4	0	96.94	3.06
CPartei2	V5	74.66	21.18	4.16
Abwander	V7	100.00	0	0

Waehlerstroeme - Zustroeme in Zahl der Stimmen
 (die Summe der Werte in einer Zeile ergibt die Stimmenzahl in der 2. Wahl)
 (ist identisch mit Zahl der Abstroeme)

		AParteil	BParteil	Zuwander	
		V1	V2	V6	
APartei2	V3	2455	0	76	2531
BPartei2	V4	0	2406	76	2482
CPartei2	V5	310	88	17	415
Abwander	V7	16	0	0	16
		2781	2494	169	5444

Zahl der Stimmen der Parteien aus 1. Wahl

		absolut	in Prozent
V1	A_Partei	2781	51.084
V2	B_Partei	2494	45.812
V6	Zuwander	169	3.104
zusammen		5444	100.000

tatsaechliche Zahl der Stimmen der Parteien aus 2. Wahl

	absolut	in Prozent	nur Parteien in Prozent
A_Partei_2	2531	46.492	46.629
B_Partei_2	2482	45.591	45.726
C_Partei_2	415	7.623	7.646
Abwanderer	16	0.294	-
zusammen		5444	100.000

vom Modell reproduzierte Zahl der Stimmen der Parteien aus 2. Wahl

	absolut	in Prozent	nur Parteien in Prozent
A_Partei_2	2531	46.492	46.629
B_Partei_2	2482	45.591	45.726
C_Partei_2	415	7.623	7.646
Abwanderer	16	0.294	-
zusammen		5444	100.000

Wenn wir in der Tabelle der absoluten Wählerströme die Zeilen- und Spaltensumme berechnen, so ergeben sich die Gesamtstimmen der einzelnen Parteien zu beiden Zeitpunkten. Die Wahlpopulation zu beiden Zeitpunkten ist gleich. Sie hat sich durch die Hinzunahme der beiden Pseudo-Parteien der Ab- und Zuwanderer auf 5444 erhöht. Da durch dieses Vorgehen die Gesamtwählerstimmen der Parteien der 2. Wahl reproduziert werden, ist auch eine zeilenweise Prozentuierung der Tabelle der absoluten Wählerströme zulässig. Dadurch können folgende Fragestellungen beantwortet werden: Wie sehen die Rekrutierungsmuster der Parteien aus? Woher kommen die Wähler z.B. der Partei APARTEI2 usw. Die zeilenweise Prozentuierung wird in der Matrix der Zuströme vorgenommen. Diese Matrix ist zeilenweise zu lesen: Die Partei APARTEI_2 rekrutiert sich zu 97% aus Wählern der Partei APARTEI_1 und zu 3% aus Zuwanderungsgewinnen usw.

Beachten: Die bei Modell 2 berechneten Zustromprozentage sind nicht identisch mit den Ergebnissen einer Analyse, bei der die un- und abhängigen Variablen vertauscht und somit die Stimmenergebnisse der 1. Wahl durch jene der 2. Wahl "vorausgesagt" werden sollen.

P51.6 Modellprüfgrößen

Wir haben bisher die Ergebnisse der Wählerstromanalyse/Hochrechnung interpretiert, ohne über die Modellanpassung Bescheid zu wissen. Wie brauchbar ist das Modell der Wählerstromanalyse/Hochrechnung für prognostische Zwecke? Mit dieser Frage ist gemeint, wie gut die Ergebnisse der 2. Wahl aufgrund der Übergangswahrscheinlichkeiten und der Ergebnisse der 1. Wahl vorausgesagt werden können.

Die im Folgenden dargestellte Modellanpassung und Residuenanalyse wird üblicherweise zusammen mit einer Wählerstromanalyse gerechnet. Sie können jedoch bei Prog51m1 auch in der Phase der Hochrechnung bei unvollständiger Auszählung durchgeführt werden. Sie sagen dann darüber aus, wie gut die Modell-Ergebnisse an die Ergebnisse der *ausgezählten* Wahlgemeinden angepasst sind.

Bezeichnen wir mit $n(k,i,t_2)$ die tatsächliche Zahl der Wähler der Partei i in der Gemeinde k bei der 2. Wahl (=Zeitpunkt t_2) und mit $nn(k,i,t_2)$ die mit Hilfe unseres Modells prognostizierte Zahl der Wähler der Partei i in der Gemeinde k bei der 2. Wahl, so gibt die Größe

$$APA(k,i) = \frac{|n(k,i,t_2) - nn(k,i,t_2)|}{n(k,i,t_2)}$$

Auskunft über die absolute prozentuelle Abweichung der prognostizierten Stimmenanzahl für die Partei i in der Gemeinde k zum Zeitpunkt t_2 bezogen auf tatsächliche Stimmenanzahl der Partei i in der Gemeinde k zum Zeitpunkt t_2 . Umso größer der Wert, desto höher ist die absolute prozentuelle Abweichung. Ist $n(k,i,t_2)$ gleich 0, wird $APA(k,i)$ gleich der absoluten Abweichung der tatsächlichen und prognostizierten Stimmen gesetzt, um Divisionen mit 0 zu vermeiden. Dies führt dazu, dass die entsprechenden Einheiten in einer Residuenanalyse i.d.R. sehr große Abweichungen aufweisen, obwohl die absoluten Unterschiede nur gering sind. Die Bezeichnung "*absolute prozentuelle Abweichung*" darf daher *nicht* dahingehend mißverstanden werden, dass *absolute Differenzen* erfasst werden,

sondern dass der *Absolutbetrag von relativen (=prozentuellen) Abweichungen in die Berechnung eingeht*. Die prognostizierte Stimmzahl $nn(k,i,t_2)$ einer Partei i in der Gemeinde k wird auf Grund der Übergangswahrscheinlichkeiten und der Stimmenverteilung der 1. Wahl ($=t_1$) berechnet mit

$$nn(k,i,t_2) = \sum_j p(i,j) \cdot n(k,j,t_1)$$

Aus den absoluten prozentuellen Abweichungen $APA(k,i)$ für die Parteien i in den Gemeinden k können folgende Maßzahlen zur Beurteilung der Modellanpassung berechnet werden:

mittlere absolute prozentuelle Abweichungen für die Partei i

$$MAPA(i) = \sum_k APA(k,i) / K \quad \text{und}$$

mittlere absolute prozentuelle Abweichungen insgesamt

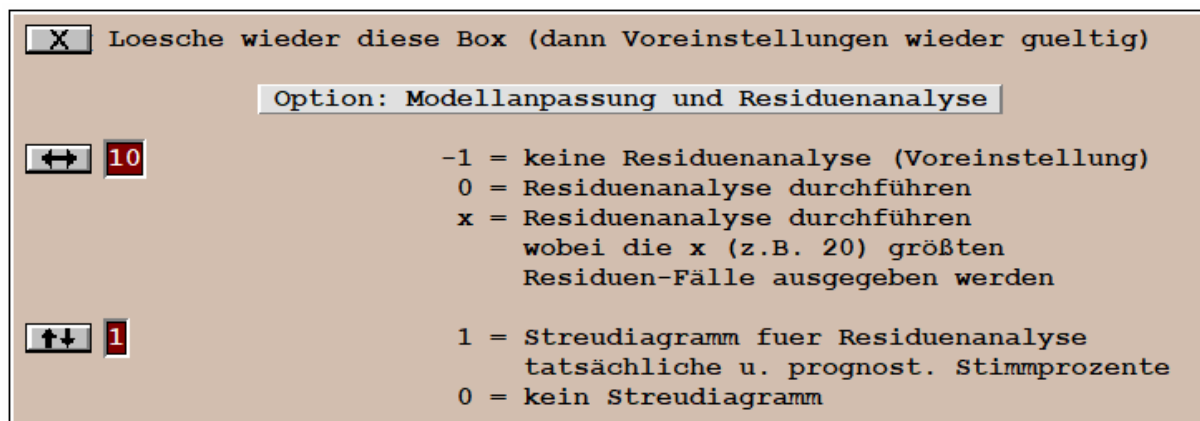
$$MAPA = \sum_i MAPA(i) / I,$$

wobei K die Zahl der Gemeinden und I die Zahl der Parteien der 2. Wahl sind.

Wir rechnen die Modellanpassung und anschließend im nächsten Abschnitt die Residuenanalyse wobei wir Modell 2 verwenden.

P51.6.1 Eingabe in Programm

Im Maskenprogramm Prog51m1 bzw. im Beispielpogramm P51Bsp1 muss die Optionsbox "Modellanpassung und Residuenanalyse" geöffnet wurde.



Im 1. Eingabefeld muss dann "0" oder eine Zahl größer 0 eingesetzt werden. Wir haben "10" geschrieben. Wird im 2. Eingabefeld "1" eingesetzt, dann wird ein Streudiagramm gezeichnet. Wir haben oben in Abschnitt P51.4.4.1 bei der Erläuterung zur Eingabebox 17 ein Streudiagramm abgebildet.

Im Maskenprogramm Prog28m1 (für die Wählerstromanalyse, ohne Hochrechnung) muss in der Eingabebox "Optionen" eingesetzt werden:

Optionen	
↑↓ 2	Modell 1 = Ermittle Abströme von Parteien der 1. Wahl zu den Parteien der 2. Wahl 2 = zusätzlich: Ermittle Zuströme zu Parteien der 2. Wahl aus den Parteien der 1. Wahl
↑↓ 2	Gewichtung 0 = mit absoluten Stimmen rechnen 1 = mit Prozentwerten rechnen 2 = Prozent*Wurzel aus Gesamtstimmen je Gemeinde aus 2. Wahl
↔ 61	Residuenanalyse 0 = keine Residuenanalyse n = Residuenanalyse mit n=Zahl der Fälle
↔ 10	x = die x größten Residuen-Fälle werden ausgegeben 0 = keine Ausgabe

Im 3. Eingabefeld (Residuenanalyse) muss die Zahl der Wahlgemeinden eingetragen werden, für die eine Wählerstromanalyse gerechnet wird. Werden also beispielsweise 450 Gemeinden untersucht, ist 450 zu schreiben. Auch die Angabe einer größeren Zahl ist zulässig. Eine kleinere Zahl führt dagegen zu einer Fehlermeldung, da ALMO in diesem Fall nicht den erforderlichen Speicher reserviert hat.

P51.6.2 Ergebnisse aus Programm

Pruefung der Modellanpassung fuer absolute Stimmen:

=====

mittlere absolute prozentuelle Abweichungen der absoluten Stimmen

V3	APartei2	3.248
V4	BPartei2	3.250
V5	CPartei2	20.532
V7	Abwander	33.666

zusammen		15.174

Die mittlere bzw. durchschnittliche Abweichung für die absoluten Stimmen der Partei APARTEI2 in einer Gemeinde beträgt 3.2%. Für die "kleinen" Parteien (CPARTEI2 und Abwanderer) ergeben sich allgemein größere absolute prozentuelle Abweichungen, obwohl die "absoluten" Differenzen nur gering sein können. Eine prozentuelle Abweichung von durchschnittlich 20.5% kann bei einer Partei mit einer geringen Stimmenzahl bedeuten, dass anstelle von 5 tatsächlichen Stimmen nur 4 oder 6 Stimmen prognostiziert werden.

Die mittleren absoluten prozentuellen Abweichungen werden von ALMO nicht nur für die absoluten Stimmen sondern auch für die Stimmenanteile berechnet.

mittlere absolute prozentuelle Abweichungen der Stimmenanteile

V3	APartei2	3.248
V4	BPartei2	3.250
V5	CPartei2	20.532
V7	Abwander	11.096
zusammen		9.53

Bezugsgrößen sind hier nicht die absoluten Stimmen der Parteien in den einzelnen Gemeinden, sondern deren Stimmenanteile. Bezogen auf den Gesamtstimmanteil von 46.5% der Partei APARTEI2 bedeuten die 3.3% Abweichungen, dass in einer Gemeinde beispielsweise anstelle von 46.5% ein Stimmenanteil von 48% oder von 45% prognostiziert wird. Die Abweichungen von 20.5% der "kleinen" Partei CPARTEI2 bedeuten, dass bezogen auf den Gesamtstimmanteil von 7.6% in einer Gemeinde beispielsweise anstelle von 7.6% ein Stimmenanteil von 9% oder von 6% vorausgesagt wird. Wie das Beispiel zeigt, sollten - wie bereits mehrfach erwähnt - prozentuelle Abweichungen von 20% oder mehr bei kleinen Parteien nicht überbewertet werden.

P51.7 Residuenanalyse

In unserem Beispiel sind die 10 größten prozentuellen Abweichungen der absoluten Stimmen:

Datensatz	Variable	absolute proz.Abw.	empirischer Wert	prognostizierter Wert	
49	V5	CPartei2	104.580	3.000	6.137
30	V7	Abwander	93.096	4.000	0.276
51	V7	Abwander	91.178	3.000	0.265
17	V7	Abwander	90.411	3.000	0.288
55	V7	Abwander	88.493	2.000	0.230
59	V5	CPartei2	77.574	4.000	7.103
17	V5	CPartei2	76.280	4.000	7.051
22	V7	Abwander	74.685	1.000	0.253
29	V7	Abwander	74.685	1.000	0.253
47	V7	Abwander	74.685	1.000	0.253

In dem Datensatz (=Wahlgemeinde) 49 hat die Partei CPARTEI2 3 Stimmen (=empirischer Wert) erhalten. Die prognostizierte Stimmenzahl ist gleich 6.137. Die absolute prozentuelle Abweichung ist daher gleich $|3.000 - 6.137| / 3.000 = 104.580$. Der prognostizierte Wert weicht um über 100% von dem empirischen Wert ab.

Normalerweise ist man wenig daran interessiert, die Ergebnisse für die Quasi-Partei der Abwanderer zu prognostizieren und die Güte dieser Prognose zu überprüfen. Man rechnet dann eine Analyse ohne Zu- und Abwanderer. Die 10 größten Abweichungen entstehen dann ausschließlich bei der neuen C_Partei_2. Hier wird ersichtlich, dass die Wahlergebnisse für neue Parteien schwieriger zu prognostizieren sind.

Datensatz	Variable	absolute proz.Abw.	empirischer Wert	prognostizierter Wert	
49	V3	C_Partei_2	96.787	3.000	5.904
17	V3	C_Partei_2	78.942	4.000	7.158
59	V3	C_Partei_2	74.808	4.000	6.992
27	V3	C_Partei_2	61.610	4.000	6.464
22	V3	C_Partei_2	59.967	4.000	6.399
6	V3	C_Partei_2	59.145	4.000	6.366

30	V3	C_Part_2	37.875	5.000	6.894
37	V3	C_Part_2	34.200	10.000	6.580
5	V3	C_Part_2	31.938	5.000	6.597
58	V3	C_Part_2	30.077	10.000	6.992

ALMO führt nach der Ausgabe der x-größten prozentuellen Abweichungen für diese eine Analyse durch, in der untersucht wird, ob sich die größten Abweichungen auf bestimmte Datensätze (Wahlgemeinden) verteilen. In unserem Beispiel (mit Abwanderer) ergibt sich folgendes Bild:

Verteilung der Abweichungen auf Datensätze (Wahlgemeinden)

Datensatz	n=
49	1
30	1
51	1
17	2
55	1
59	1
22	1
29	1
47	1

Verteilung der Abweichungen auf Variable

Variable	n=
V5 CPartei2	3
V7 Abwander	7

Nur bei einem Datensatz (=Datensatz 17) treten mehr als eine Abweichung auf. Von einer Konzentration von Abweichungen auf bestimmte Datensätze kann daher kaum gesprochen werden. Dennoch wird man die größten Abweichungen auf eventuelle Eingabefehler hin untersuchen. Im Unterschied zur Verteilung der größten Abweichungen auf die Datensätze zeigt sich eine Konzentration auf die beiden kleinen Parteien CPARTEI2 und Abwanderer.

Die x-größten Abweichungen werden auch für Stimmenanteile berechnet. Für diese zeigen wir noch die Modelldiagnose aus dem Regressionsansatz

Modelldiagnosen aus Regressionsansatz:

Variable		SQ(gesamt)	SQ(erkl.)	SQ(Fehler)
V1	A_Part_2	94400427.25	94400264.24	163.01
V2	B_Part_2	90213229.67	90213066.50	163.18
V3	C_Part_2	2492428.69	2492290.66	138.03
V9	Abwanderer	4853.61	4815.88	37.73
zusammen		187110939.21	187110223.85	501.95

Variable	R-Quadrat
V1 A_Part_2	0.99999827
V2 B_Part_2	0.99999819
V3 C_Part_2	0.99994462
V9 Abwanderer	0.99222612

zusammen

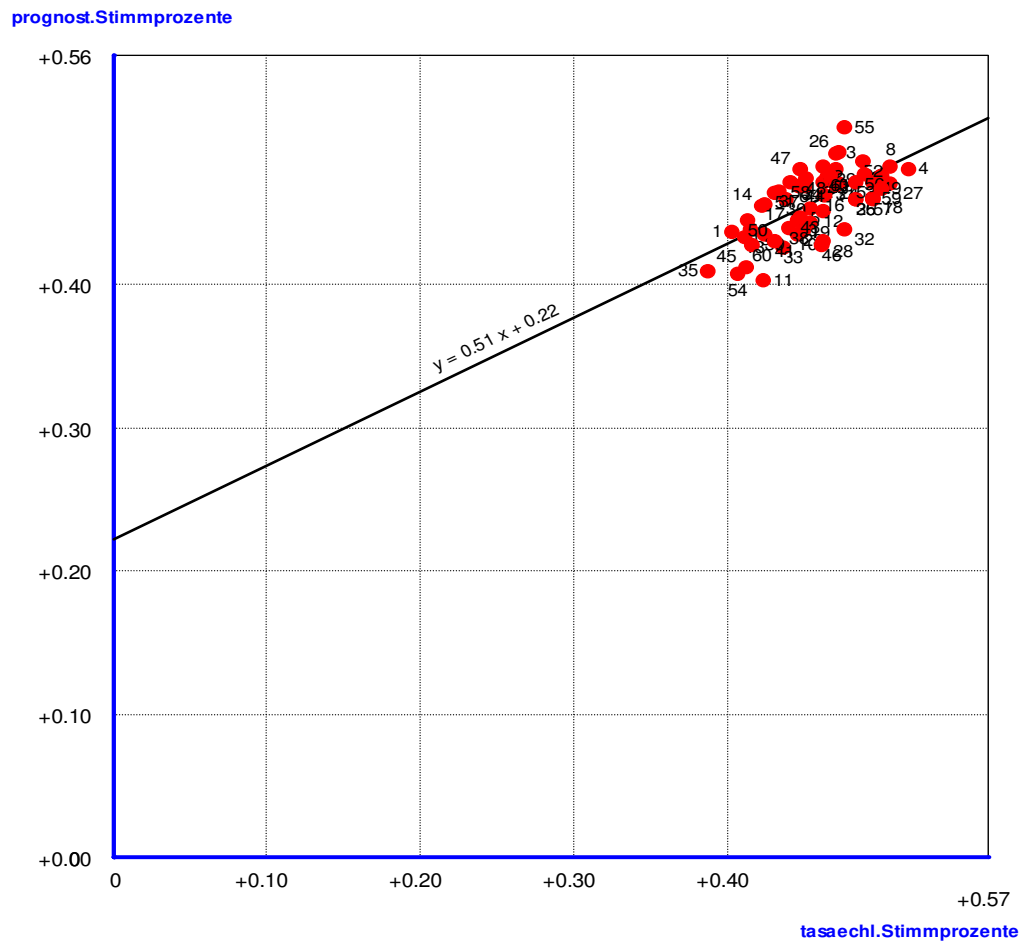
0.99999732

$SQ(erk\ l.)$ ist die Abweichungsquadratsumme, die das Modell von der $SQ(gesamt)$, also der gesamten Abweichungsquadratsumme erklärt. R -Quadrat bezeichnet die quadrierte Korrelation zwischen tatsächlichem und reproduziertem Stimmenanteil

Mit Prog51Bsp werden dann noch Streudiagramme für die Parteien gezeichnet. Wir zeigen hier das Streudiagramm für die B_Partei_2. Ein Punkt bildet eine Wahlgemeinde ab, wobei sein Wert auf der x-Achse den tatsächlichen Anteilswert und auf der y-Achse den vom Modell reproduzierten Anteilswert abbildet.

B_Partei_2 - Streudiagramm der tatsächlichen u. prognostizierten Stimmprozente

Streudiagramm



Literatur

G. Bruckmann: Schätzung von Wahlresultaten aus Teilergebnissen, Wien, Würzburg, 1966

Greiner, D. J. und Quinn, K. M. (2009). R*C Ecological Inference: Bounds, Correlations, Flexibility and Transparency of Assumptions. *Journal of the Royal Statistical Society A*, 172 (1), S.67-81.

Internet: <https://www.law.berkeley.edu/files/GQ-JRSSA.pdf>

C. Hofinger/ G. Ogris: Orakel der Neuzeit: Was leisten Wahlbörsen, Wählerstromanalysen und Wahltagshochrechnungen? *Österreichische Zeitschrift für Politikwissenschaft*, 2002, 31, 143-58.

Klima, Andre: Ökologische Inferenz und hybride Modelle, Schätzung der Wählerwanderung in Mehrparteiensystemen, 2016

Internet: https://edoc.ub.uni-muenchen.de/20327/1/Klima_Andre.pdf

Kopecki, Petra: Analyse der Wählerwanderung zwischen der Oberbürgermeisterwahl (2015) und Bundestagswahl (2013) in Mannheim anhand von Aggregat- und Individualdaten, 2016

Internet: https://epub.ub.uni-muenchen.de/29574/1/MA_Kopecki.pdf

Ledl, Thomas: Modellierung von Wechselwählerverhalten als Multinomial-experiment, 2007

Internet: <http://homepage.univie.ac.at/thomas.ledl/>

E. Neuwirth: Prognoserechnung am Beispiel der Wahlhochrechnung; In: P. Mertens (Hg.), *Prognoserechnung*, Heidelberg, physica-Verlag, 1994

E. Neuwirth: Schätzung von Wählerübergangswahrscheinlichkeiten; In: M. Holler, (Hg.), *Wahlanalyse*, tuduv-Verlag, München, 1984

Rosen, O., Jiang, W., King, G. und Tanner, M. A. (2001). Bayesian and Frequentist Inference for Ecological Inference: The R*C Case. *Statistica Neerlandica*, 55 (2), 2001, S.134-156.

Internet: <http://gking.harvard.edu/files/gking/files/em.pdf?m=1360038990>

J. H. Wilkinson/C. Reinisch: *Lineare Algebra*, Berlin-Heidelberg-New York, 1971