TECHNICAL REPORT

Prognoserechnung am Beispiel der Wahlhochrechnung

Erich Neuwirth

TR-93-8

May 1993

Institute of Statistics, Operations Research,
and Computer Methods
University of Vienna
Universitätsstraße 5/9, A-1010 Wien, Austria

Erschienen in: Peter Mertens, Prognoserechnung (physica Verlag, 1994)

Prognoserechnung am Beispiel der Wahlhochrechnung
Erich Neuwirth

Einleitung:

Prognoserechnung, also Vorhersage von numerischen Daten, ist eine sehr allgemeine statistische Methodik. Eines der in der Öffentlichkeit relativ bekannten Einsatzgebiete ist die Wahlhochrechnung am Wahltag während der laufenden Auszählung der Stimmen. Wir wollen uns im folgenden mit den mathematisch-statistischen Grundlagen des Verfahrens beschäftigen. Zusätzlich werden wir uns aber auch mit grundsätzlichen Fragen darüber, welchen Zweck Wahlhochrechnungen über den reinen Unterhaltungswert hinaus überhaupt erfüllen, befassen.

1. Mathematisch-statistische Modellgrundlagen

Wenn wir im folgenden von Wahlhochrechnung sprechen, dann wollen wir darunter immer ein Verfahren verstehen, daß man auch als „Schätzung von Wahlresultaten aus Teilergebnissen“ (siehe auch [2]und [4]) bezeichnen kann. Die Datengrundlage besteht also aus dem kompletten Wahlergebnis einer Vorwahl und Ergebnissen der neuen Wahl, allerdings nur für einen Teil des Wahlgebiets. Diese Art der Wahlprognose unterscheidet sich ganz wesentlich von Vorhersagen, die auf vor der Wahl durchgeführten Meinungsumfragen beruhen, weil die Datenquelle wesentlich zuverlässiger ist.
Sehr oft wählt man als Vorwahl einfach die letzte gleichartige Wahl (etwa bei Wahlen zum Bundestag die letztvergangene Bundestagswahl). In manchen Fällen ist das aber nicht so einfach möglich. Ein sehr einprägsamer derartiger Fall war die Volksabstimmung über das Kernkraftwerk Zwentendorf in Österreich im Jahr 1976.
Die Daten über beide Wahlgänge müssen in einer relativ feinen regionalen Gliederung vorliegen. Normalerweise stehen Wahlergebnisse mehr oder weniger automatisch für die einzelnen Gemeinden des Wahlgebiets zur Verfügung, daher liegen die Daten auch in ausreichender Feingliederung vor.
Die zentrale Annahme für das im folgenden beschriebene Hochrechnungsverfahren ist nun folgende: wir gehen davon aus, daß das Übergangsverhalten der Wähler im gesamten Wahlgebiet homogen ist. Vereinfacht gesprochen heißt das, daß diejenigen Wähler, die bei der Vorwahl eine bestimmte Partei gewählt haben, sich bei der neuen Wahl in allen Gemeinden des betrachteten Gebietes nach dem selben Schlüssel auf die bei der neuen Wahl kandidierenden Parteien verteilen. Diese grundlegende Modellannahme entzieht sich im wesentlichen einer empirischen Verifizierung. Allerdings zeigt die Erfahrung, daß die Qualität der auf diesem Modell beruhenden Prognosen sehr hoch ist. Sollte diese Modellannahme in ihrer reinen Form nicht erfüllt sein (was in der Regel der Fall ist), dann handelt es sich trotzdem um kein unlösbares Problem. In diesem Fall teilt man nämlich einfach das gesamte Wahlgebiet in mehrere Subregionen ein, für die man dann annehmen kann, daß die Homogenitätsannahme zumindest näherungsweise erfüllt ist. Darauf aufbauend führt man für jedes dieser Teilgebiete eine eigene Hochrechnung durch. Wie sehen nun die Modellannahmen in mathematischer Formulierung aus?

Wir gehen davon aus, daß das Hochrechnungsgebiet aus N Gemeinden besteht. In jeder dieser Gemeinden gibt es I Parteien bei der Vorwahl und J Parteien bei der neuen Wahl. Daher gibt es ein altes und ein neues Wahlergebnis, das wir wie folgt beschreiben:

$X_i^{(n)}$	Stimmen der Partei i in Gemeinde n bei der Vorwahl
$Y_j^{(n)}$	Stimmen der Partei j in Gemeinde n bei der neuen Wahl

Die Annahme der „homogenen Wählerströme im Wahlgebiet“ läßt sich mathematisch so ausdrücken:
In allen Gemeinden gelten folgende Gleichungen:

$$Y_j^{(n)} = \sum_{i=1}^{I} X_i^{(n)} p_{ij} \quad \text{für } j = 1..J, n = 1..N$$

Diese Gleichungen bedeuten ganz einfach, daß das Stimmenergebnis der neuen Wahl zustandekommt, indem in allen Gemeinden derselbe Prozentsatz jener Wähler, die bei der Vorwahl Partei i gewählt haben, bei der neuen Wahl Partei j wählen, und daß diese Tatsache für alle $i = 1..I$ und $j = 1..J$ gilt.
Grundlage dieses Modells ist, daß bei beiden Wahlgängen dieselben Wähler wahlberechtigt waren, was natürlich eine vereinfachende Annahme ist. Wir werden uns in einem späteren Abschnitt noch einmal näher mit diesem Problem beschäftigen, derzeit nehmen wir die vereinfachende Annahme einfach als gegeben an.
Wir können jetzt natürlich noch die Gesamtzahl der Wahlberechtigten in allen Gemeinden durch eine Gleichung ausdrücken:

$$S^{(n)} = \sum_{i=1}^{I} X_i^{(n)} = \sum_{j=1}^{J} Y_j^{(n)}$$

Ebenso können wir die Anteile der einzelnen Parteien als Gleichungen ausdrücken:

$$x_i^{(n)} = \frac{X_i^{(n)}}{S^{(n)}}$$

$$y_j^{(n)} = \frac{Y_j^{(n)}}{S^{(n)}}$$

Genauso wie für die Stimmen gelten unsere Modellgleichungen dann auch für die Anteile:

$$y_j^{(n)} = \sum_{i=1}^{I} x_i^{(n)} p_{ij} \quad \text{für } j = 1..J, n = 1..N$$

Die Hochrechnung hat nun folgende Aufgabe:
Während des Hochrechnungszeitraums liegen die Ergebnisse der Vorwahl schon komplett vor, die Ergebnisse der aktuellen Wahl jedoch nur teilweise. Daher wird mit statistischen Methoden versucht, aus den Daten aller Gemeinden, wo bereits beide Wahlergebnisse verfügbar sind, die Übergangskoeffizienten (auch als Wählerströmekoeffizienten bezeichnet) zu schätzen. Wenn wir diese Schätzung mit $\hat{p}_{ij}$ bezeichnen, dann können wir mittels der Gleichung

$$\hat{y}_j^{(n)} = \sum_{i=1}^{I} x_i^{(n)} \hat{p}_{ij}$$

eine Schätzung für das Ergebnis der aktuellen Wahl in Gemeinde n auch dann berechnen, wenn aus dieser Gemeinde noch kein Ergebnis der aktuellen Wahl vorliegt. Man braucht jetzt nur noch diese Schätzung für die noch nicht

ausgezählten Gemeinden und das Ergebnis der aktuellen Wahl in den bereits ausgezählten Gemeinden zu addieren und bekommt so bereits die Wahlhochrechnung. Sie ändert sich laufend, weil die Datengrundlage der Schätzung der $\hat{p}_{ij}$ jene Gemeinden sind, aus denen bereits Ergebnisse der aktuellen Wahl vorliegen, und diese Gemeindegruppe während der Hochrechnung natürlich ständig größer wird.
Die $\hat{p}_{ij}$ sind aber nicht nur als Hilfsmittel für die Wahlhochrechnung interessant. Sie liefern gleichzeitig eine „Wählerstromanalyse“. Diese Koeffizienten geben an, wie sich die Wähler der einzelnen Parteien bei einer Vergleichswahl bei der neuen Wahl auf die kandidierenden Gruppen verteilt haben, und genau das ist die Frage, die eine Wählerstromanalyse beantworten soll. Eine nach dieser Methode durchgeführte Wahlhochrechnung liefert als Zusatznutzen also gleich noch eine Wählerstromanalyse.

2. Statistische Schätzverfahren für die Modellparameter

In die Überlegungen des vorigen Abschnitts sind stochastische Komponenten noch überhaupt nicht eingeflossen. Die Modellgleichungen sind von den idealisierten Annahmen räumlich homogener Übergangswahrscheinlichkeiten ausgegangen. Diese Annahmen haben dann zu eigentlich deterministischen Gleichungen über den Zusammenhang zwischen Ergebnis der Vorwahl und der aktuellen Wahl geführt.
Die stochastische Komponente wird mit folgendem Modell beschrieben.
Wir gehen davon aus, daß wir das Verhalten der einzelnen Wahlberechtigten als unabhängige Wiederholungen einer Multinomialverteilung beschreiben können. Ein Wähler, der bei der Vorwahl Partei i gewählt hat, möge bei der aktuellen Wahl mit Wahrscheinlichkeit p_{ij} Partei j wählen.
Wir wollen den zu einem solchen Wähler gehörenden Zufallsvektor mit $\vec{X}_i$ bezeichnen. (Der Index i ist notwendig, weil die Verteilung des Zufallsvektor je nach Wahlverhalten bei der Vorwahl verschieden ist.) Aufgrund der Definition von $\vec{X}_i$ gilt, daß der Erwartungeswertvektor $E(\vec{X}_i)$ die Komponenten p_{ij} hat.
Wenn wir jetzt das Wahlergebnis in Gemeinde n (in Stimmen) als Zufallsvektor $\vec{Y}^{(n)}$ betrachten, dann läßt sich dieser Zufallsvektor folgendermaßen darstellen:

$$\vec{Y}^{(n)} = \sum_{i=1}^{I} \sum_{k=1}^{X_i} \vec{X}_{i,k}^{(n)}$$

Dabei sind die $\vec{X}_{i,k}$ für gleiche i und verschiedene k unabhängige Zufallsvektoren mit Multinomialverteilung (mit dazugehörendem gleichen Wahrscheinlichkeitsvektor p_{ij}). Da ein solcher Vektor nur das Übergangsverhalten eines einzelnen Wählers beschreibt, hat er in genau einer Komponente die Zahl 1 und und allen anderen Komponenten die Zahl 0 stehen. Daher gilt für den Erwartungswert der j-ten Komponente von $\vec{Y}^{(n)}$ (wegen der Summierung aus den $\vec{X}_{i,k}^{(n)}$)

$$E(\vec{Y}_j^{(n)}) = \sum_{i=1}^{I} X_i^{(n)} p_{ij}$$

Die Stimmenergebnisse in den einzelnen Gemeinden, also die $Y_j^{(n)}$ sind aber nichts anderes als Realisationen der Zufallsvariablen $\vec{Y}_j^{(n)}$. Somit genügen unsere

Zufallsvariablen, für die wir Realisationen kennen, einem linearen Modell. Daher liegt es nahe, zur Schätzung der p_{ij}, die die Parameter des linearen Modells sind, regressionsanalytische oder zumindest verwandte Verfahren zu verwenden. Es gibt (vor allem in der soziologischen und politologischen Literatur) auch immer wieder Arbeiten, die diese Gleichung zur Grundlage einer multiplen Regressionsanalyse machen und die Koeffizienten, die die Regressionsschätzung ergibt, einfach als Wählerströme interpretieren. In einer derartigen Regressionsanalyse hätte man als jeweils einzige abhängige Variable die Ergebnisse einer Partei in allen Gemeinden zu wählen, als unabhängige Variable die Ergebnisse aller kandidierenden Parteien bei der Vorwahl in allen Gemeinden.
Allerdings ist ein derartiges Vorgehen aus mehreren Gründen methodisch nicht zu rechtfertigen. Da es sich bei den p_{ij} um Wahrscheinlichkeiten handelt, wissen wir, daß für alle p_{ij} gilt

$$0 \le p_{ij} \le 1$$

Außerdem gehören alle p_{ij} mit festem i zu einer festen Multinomialverteilung, daher gilt

$$\sum_{j=1}^{J} p_{ij} = 1 \quad \text{für alle } i$$

Die Einhaltung der Summenbedingung ist im simplifizierenden Regressionsmodell sogar gewährleistet (sie folgt aus linearen Abhängigkeiten zwischen den in Betracht kommenden Variablen), die Nichtnegativitätsbedingung wird aber normalerweise nicht eingehalten. Solange die Regressionskoeffizienten nur intern für die Hochrechnung verwendet werden, ist diese Inkonsistenz zwar störend, ein derartiges Vorgehen kann aber immer noch zu brauchbaren Hochrechnungen führen. Sollen jedoch auch Wählerstromanalysen publiziert werden, dann sind die Resultate dieses Verfahrens unbrauchbar, weil es negative Wählerströme nicht geben kann.
Dieses Problem ist aber natürlich mit statistischen Hilfsmitteln zu lösen. Bei der Regressionsanalyse handelt es sich ganz einfach um den Maximum-Likelihood-Schätzer in einem linearen Modell ohne weitere Restriktionen über die Parameter. Wir können dieselbe Likelihood-Funktion untersuchen und bei der Optimierungsaufgabe zur Ermittlung des besten Schätzers einfach nur Parameterkombinationen zulassen, die die Konsistenzbedingungen (Nichtnegativität und Summenbedingung) erfüllen.
Es gibt aber noch eine zusätzliche Komplikation, die schon beim linearen Modell ohne Nebenbedingungen auftritt: Standardvoraussetzung einer gewöhnlichen multiplen Regressionsanalyse ist die Homoskedastizität. Die Varianzen aller in Modell auftretenden Zufallsvariablen müssen gleich sein. Wir müssen also die Varianzen unserer Zufallsvariablen (zunächst einmal bei festem j) ansehen.

$$\vec{Y}_j^{(n)} = \sum_{i=1}^{I} \sum_{k=1}^{X_i} \vec{X}_{i,k}^{(n)}$$

Die einzelnen j-ten Komponenten $\vec{X}_{i,k}^{(n)}$ sind binomialverteilte unabhängige Zufallsvariable, haben also die Varianz jeweils $p_{ij}(1-p_{ij})$. Die Gesamtvarianz von $\vec{Y}_j^{(n)}$ ist daher

$$\sigma^2(\vec{Y}_j^{(n)}) = \sum_{i=1}^{I} X_i^{(n)} p_{ij}(1-p_{ij}) = S^{(n)} \sum_{i=1}^{I} x_i^{(n)} p_{ij}(1-p_{ij})$$

$S^{(n)}$ ist die Zahl der Wahlberechtigten in der Gemeinde n. Diese Zahl unterliegt sehr starken Schwankungen. Wir können daher die Annahme gleicher Varianz für alle Zufallsvariablen nicht ohne weiteres als bestätigt ansehen. Allerdings sind die $S^{(n)}$ bekannt. Der restliche Term in unserem Varianzausdruck ist jedoch unbekannt, weil er die Parameter p_{ij} enthält, die wir erst schätzen wollen. Andrerseits kann man davon ausgehen, daß in einem politisch einigermaßen homogenen Teilgebiet unseres gesamten Untersuchungsgebiets die $x_i^{(n)}$ nicht allzu verschieden sind. Dann sind auch die Summen $\sum_{i=1}^{I} x_i^{(n)} p_{ij}(1-p_{ij})$ nicht allzu verschieden und wir machen keinen großen Fehler, wenn wir annehmen, daß die Varianzen annähernd folgende Gleichung erfüllen:

$$\sigma^2(\vec{Y}_j^{(n)}) = CS^{(n)}$$

Dabei ist C eine unbekannte Konstante.

Ein klassisches Resultat aus der Theroie der linearen Modelle besagt, daß bei ungleichen Varianzen, die bis auf einen multiplikativen Faktor bekannt sind, in Analogie zum varianzhomogenen linearen Modell ebenfalls ein optimaler Schätzer existiert. In unserem Fall ist dieser optimale Schätzer für die p_{ij} als Lösung des Minimierungsproblems

$$\sum_{n=1}^{N} \frac{1}{S^{(n)}} (Y_j^{(n)} - \sum_{i=1}^{k} X_i^{(n)} p_{ij})^2 = \min$$

zu finden. Im Falle des Problems ohne Nebenbedingungen gibt es für dieses Problem natürlich auch explizite Lösungsgleichungen. Wir beschäftigen uns aber gerade mit einem Fall, bei dem die p_{ij} zusätzliche Nebenbedingungen erfüllen müssen.

Im einfachen Fall ohne Nebenbedingungen könnte man dieses Optimierungsproblem getrennt für jedes j lösen und erhielte so Schätzwerte für die gesuchten Parameter. Wir wollen aber noch zusätzlich, daß folgende Gleichungen und Ungleichungen erfüllt sind:

$$0 \le p_{ij} \le 1$$

$$\sum_{j=1}^{J} p_{ij} = 1 \text{ für alle } i$$

Das Problem dabei ist, daß die Gleichungen Koeffizienten miteinander verknüpfen, die in verschiedenen Optimierungsproblemen vorkommen, (wir haben für jeden Wert von j ein anderes Optimierungsproblem zu lösen). Die Ungleichungen verknüpfen aber p_{ij} mit gleichem i und verschiedenem j. Wir können daher die Optimierungsprobleme nicht unabhängig voneinander lösen, so wie das im Falle ohne Nebenbedingungen möglich wäre. Als vereinfachender Ausweg bietet sich an, die zu minimierenden Zielfunktionen einfach zu addieren und dann diese neue Gesamtzielfunktion unter Einhaltung der Nebenbedingungen zu minimieren. Das zu lösende Optimierungsproblem lautet also:

$$\sum_{j=1}^{J} \sum_{n=1}^{N} \frac{1}{S^{(n)}} (Y_j^{(n)} - \sum_{i=1}^{k} X_i^{(n)} p_{ij})^2 = \min$$

wobei

$$0 \le p_{ij} \le 1$$

$$\sum_{j=1}^{J} p_{ij} = 1 \text{ für alle } i$$

Statistisch wäre diese Addition gerechtfertigt, wenn die entsprechenden Zufallsvariablen für verschiedene Werte von j unabhängig wären, weil dann die Likelihood-Funktionen miteinander multipliziert werden könnten und daher die negativen Logarithmen der Likelihood-Funktionen, und das sind gerade unsere zu minimierenden Funktionen, addiert werden könnten. Diese Annahme ist sicher nicht ganz richtig. Die korrekte analytische Berechnung des Modells unter Berücksichtigung dieser Tatsache ist im Prinzip auch möglich, würde aber den Aufwand zur Berechnung der Schätzer der p_{ij} enorm erhöhen. Da die Hochrechnung aber mit relativ einfacher Software durchgeführt werden soll und außerdem die rein pragmatisch betrachtete Qualität der Schätzer ein sehr wichtiges Kriterium ist, haben wir uns entschlossen, dieses stark vereinfachende Modell trotz bekannter Modellungenauigkeiten einzusetzen. In einem weiteren Abschnitt werden wir uns die Güte der Vorhersagen in einem realen Fall genauer ansehen.
Zusammenfassend können wir feststellen, daß die zuletzt angegebenen Gleichungen einen Kompromiß zwischen voller mathematischer Modellgenauigkeit und einfach zu rechnendem Analyseverfahren darstellen. Wir verwenden dieses Verfahren seit einigen Jahren bei gesamtösterreichischen Wahlgängen zur Wahlhochrechnung und erzielen damit eine ganz beachtliche Vorhersagegenauigkeit.
Die früher in Österreich eingesetzten Verfahren (für eine vollständige methodische Beschreibung siehe [2]) haben dieses Modell noch weiter vereinfacht und dabei grob gesagt den multivariaten Ansatz unseres Modells durch einen univariaten Ansatz ersetzt.
Gehen wir einmal davon aus, daß bei den beiden betrachteten Wahlgängen dieselben Parteien kandidieren. Wenn man dann die Gleichung

$$E(\vec{Y}_j^{(n)}) = \sum_{i=1}^{I} X_i^{(n)} p_{ij}$$

für eine Partei j betrachtet, die nur Stimmen verliert, dann sind ja für $i \neq j$ alle $p_{ij} = 0$. Daher reduziert sich diese Gleichung auf

$$E(\vec{Y}_j^{(n)}) = X_j^{(n)} p_{jj}$$

Für eine Partei, die nur gewinnt (und keine eigenen Wähler verliert), gilt

$$E(\vec{Y}_j^{(n)}) = X_j^{(n)} + \sum_{i \neq j} X_i^{(n)} p_{ij}$$

Wenn wir jetzt noch (das ist natürlich wieder eine zusätzliche Vereinfachung) annehmen, daß diese Partei von allen anderen Parteien einen jeweils gleichen Anteil von deren Wählern hinzugewinnt, daß wir also $p_{ij} = q_j$ annehmen können, dann gilt folgende Modellgleichung:

$$E(Y_j^{(n)}) = X_j^{(n)} + \sum_{i \neq j} X_i^{(n)} q_j = (1 - q_j) X_j^{(n)} + q_j S^{(n)}$$

Dieser Ansatz ist im wesentlichen wieder univariat. Pragmatisch wurde dann so vorgegangen, daß man die p_{jj} und die q_j in univariaten Modellen schätzte und dann für die einzelnen Parteien jeweils nach dem bei der Wahl zu beobachtenden Gewinnen oder Verlusten die eher passende Methode zur Vorhersage verwendete. Solange diese Methode bei gesamtösterreichischen Wahlen zur Wahlhochrechnung verwendet wurde, war aber immer wieder ein Problem zu beobachten. Die Ergebnisse aus Wien liegen immer sehr spät vor, weil das gesamte Bundesland Wien erst zum letztmöglichen Zeitpunkt Wahlschluß hat. Die vorher bereits berechneten Wahlprognosen mußten also immer Wien aus den anderen Bundesländern hochrechnen. Dabei war dann fast immer festzustellen, daß sich die Prognose quantitativ entscheidend änderte, sobald die ersten Wiener Ergebnisse in der Hochrechnung berücksichtigt werden konnten.

3. Ein empirischer Fall: Bundespräsidentenwahl 1992 in Österreich

Das beschriebene Verfahren wurde auch bei der Bundespräsidentenwahl 1992 in Österreich eingesetzt.Wir wollen uns den ersten Wahlgang etwas näher ansehen. Als Vergleichswahl verwenden wir die Nationalratswahl 1990. Die 4 Kandidaten dieser Wahl waren relativ eindeutig den 4 größeren Parteien der Nationalratswahl zuzuordnen. Deshalb war es auch möglich, zumindest versuchsweise die im vorigen Abschnitt beschriebenen vereinfachenden univariaten Modellvarianten zu verwenden.

Eine weitere methodische Bemerkung ist noch notwendig. Sämtliche beschriebenen Verfahren beruhen darauf, daß man die p_{ij} im Prognosegebiet als konstant voraussetzen kann. Diese Annahme ist aber für das gesamte Bundesgebiet Österreich nicht gerechtfertigt. Daher wird zunächst einmal für jedes Bundesland getrennt eine Hochrechnung durchgeführt. Diese Hochrechnungen werden dann zu einer bundesweiten Hochrechnung kombiniert. Dabei tritt aber ein Problem auf: Es gibt Bundesländer, für die ziemlich lange überhaupt keine Ergebnisse vorliegen. Daher sind in diesen Bundesländern auch keine Schätzungen für die p_{ij} verfügbar.

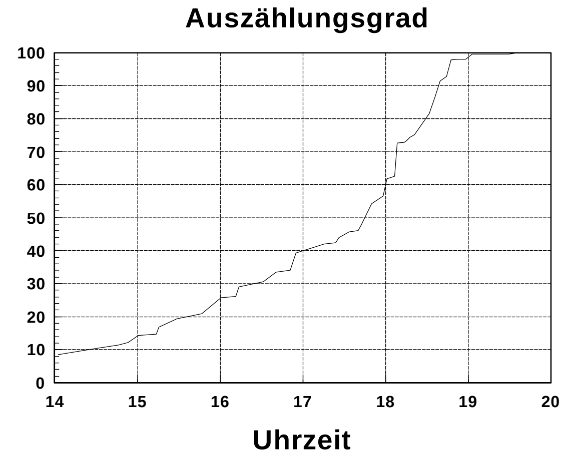
Der pragmatische Ausweg aus diesem Problem besteht darin, in diesem Fall ein gewichtetes Mittel der p_{ij} aus den anderen Bundesländern zu verwenden.

Gewichtungsfaktor ist dabei vor allem der Auszählungsgrad in diesen Bundesländern. Zusätzlich werden die Gewichte aber noch mit Konstanten auf- oder abgewertet, wobei diese Konstanten anderen Bundesländern mit ähnlicher politischer Konstellation und in geographischer Nähe zum betrachteten Bundesland ein höheres Gewicht einräumen als den anderen Bundesländern.

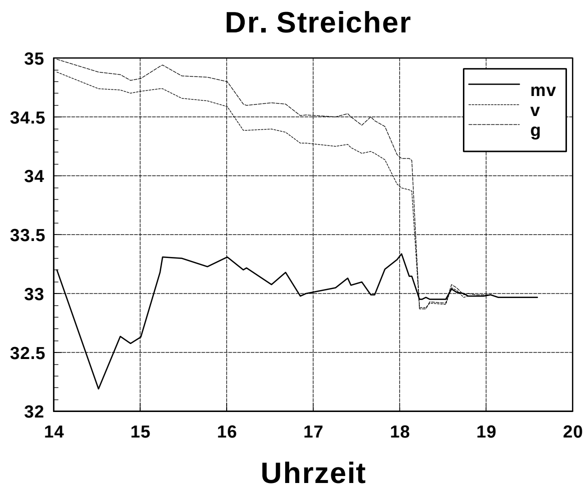
Vergleicht man die vereinfachenden univariaten Ansätze und den vollen multivariaten Ansatz unter diesem Gesichtspunkt, dann ist die entscheidende Frage, ob der multivariate Ansatz bei diesem Kombinationsverfahren bessere Vorhersagewerte in Bundesländern, aus denen noch keine Einzelergebnisse vorliegen, liefert.

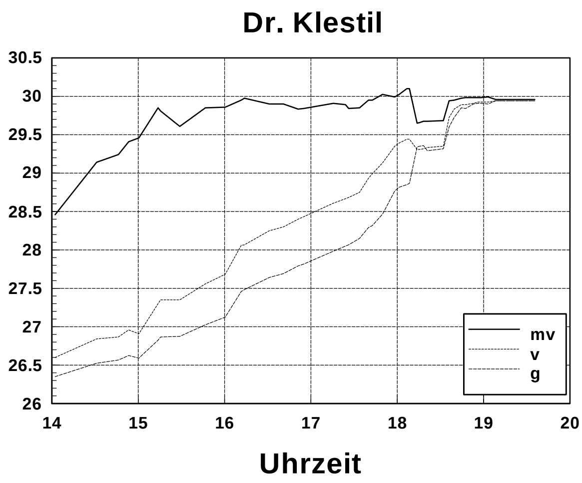
Sehen wir uns das also im Falle der Bundespräsidentenwahl an.

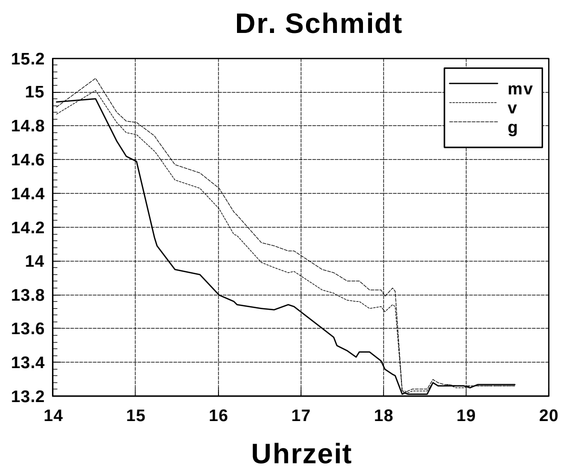
Der Auszählungsgrad in Abhängigkeit von der Uhrzeit wird in folgender Graphik dargestellt

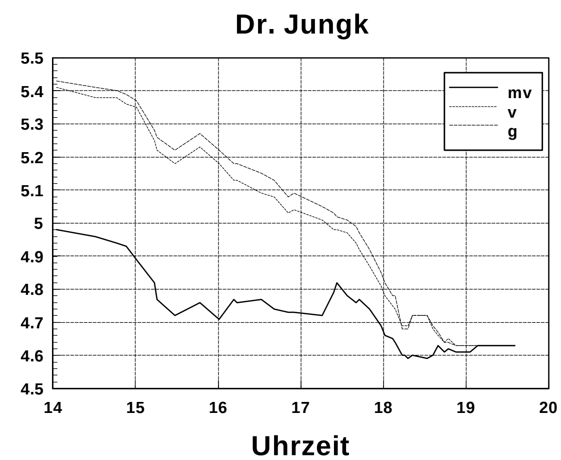


In den folgenden Graphiken wollen wir die Prognosen nach den drei verschiedenen Methoden für die einzelnen Kandidaten darstellen. Die ausgewiesenen Prozentsätze sind immer Anteile an Wahlberechtigten. Da die Prognosemethodik wie im vorigen Abschnitt dargestellt auch die Gruppe der Nichtwähler im Modell berücksichtigt, gibt diese Darstellungsform ein besseres Bild des tatsächlichen Verhaltens des Prognosemodells als die Anteile an gültigen Stimmen. In der Legende bedeutet mv immer die multivariate Schätzung, g die Schätzung, die bei gewinnenden Parteien eher paßt, und v die Schätzung, die bei verlierenden Parteien eher paßt.



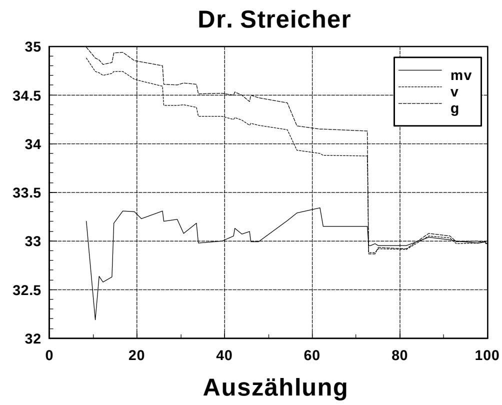


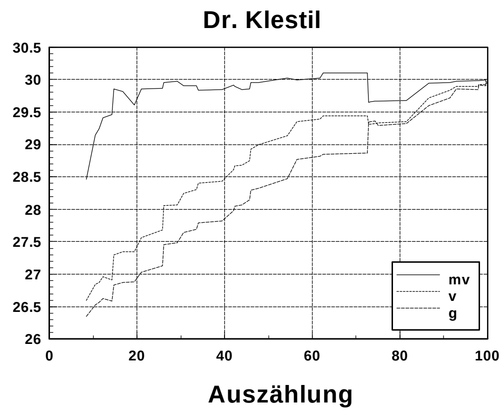


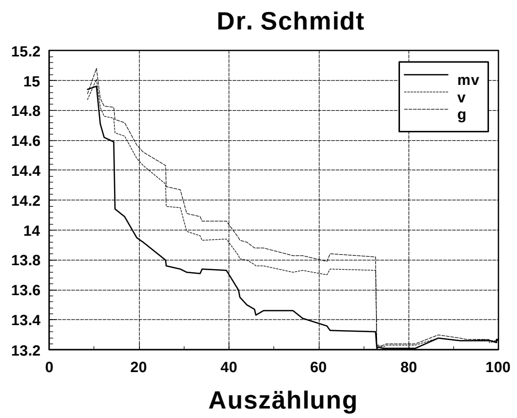


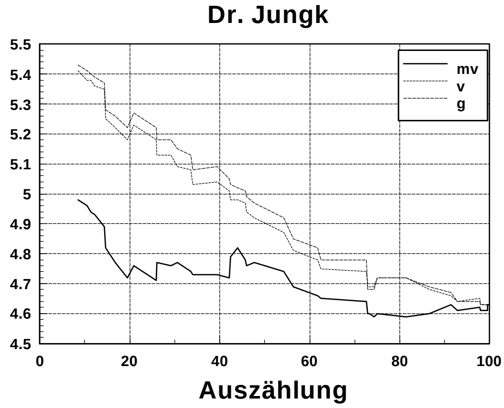
Diese Graphiken zeigen deutlich die Überlegenheit der multivariaten Schätzung gegenüber den univariaten Schätzungen. Besonders auffällig ist die „Knickstelle“ knapp nach 18 Uhr. Dabei handelt es sich um den im vorigen Abschnitt bereits erwähnten „Wien-Knick“. Zu diesem Zeitpunkt stehen üblicherweise die ersten Teilergebnisse aus Wien für die Hochrechnung zur Verfügung. Da die Mehrheitsverhältnisse in Wien generell von „Restösterreich“ verschieden sind, kann man speziell bei univariaten Modellen nicht erwarten, daß sich Trends aus einem eher simplistischen Modell einfach übertragen lassen. Dieser schon bei früheren Hochrechnungen immer wieder beobachtbare „Wien-Effekt“ war für den

Autor dieses Beitrags auch der Grund, den multivariaten Ansatz der Hochrechnung zur Erhöhung der Vorhersagegenauigkeit zu analysieren.
Aus methodischen Gründen ist es, zumindest post festum, auch ganz nützlich, dieselben Daten noch einmal graphisch darzustellen, diesmal aber mit dem Auszählungsgrad entlang der x-Achse. Bei diesen Graphiken werden Bruchstellen, die durch massiertes Eintreffen von Daten zu bestimmten Zeiten auftreten, etwas gedehnt und dadurch geglättet. Man erhält so also vom methodischen Standpunkt aus interessantere Kurven über den Verlauf der Prognosen.









Auch bei dieser Darstellungsform bleiben (besonders bei Dr. Klestil und Dr. Schmidt) bei den univariaten Prognosen die Bruchstellen deutlich sichtbar.
Beide Sätze von Graphiken liefern deutliche Hinweise, daß der multivariate Ansatz, also die möglichst vollständige statistische Erfassung aller möglichen Wählerströme, nicht nur das theoretisch-methodisch einleuchtendere Modell ist, sondern daß die Vorhersagegüte der Hochrechnung merkbar (und zwar sogar für Laien merkbar) zunimmt.

4. Allgemeine Überlegungen

Eine im Zusammenhang mit der Wahlhochrechnung immer wieder gestellte Frage lautet: Lohnt es überhaupt den ganzen Aufwand, ein Ergebnis, das man 2 bis 3 Stunden später sowieso genau kennt, schon vorherzusagen? Die erste Antwort lautet natürlich: Eigentlich lohnt es sich nicht unbedingt.
Prinzipiell werden Hochrechnungen trotzdem durchgeführt, weil Radio und Fernsehen dankbar sind, vor dem Vorliegen des endgültigen Wahlergebnisses berichtenswerte Nachrichten geliefert zu bekommen. Das sollte aber für einen Statistiker nicht als Begründung ausreichen, den ganzen aufwendigen Apparat „in Betrieb zu nehmen“.
Es gibt aber noch einen weiteren Grund, der aus der Sicht des Statistikers für die Durchführung einer Wahlhochrechnung spricht. Die wissenschaftliche Statistik kämpft immer wieder mit einem Legitimationsproblem. In den meisten Fällen sind Analysen und Vorhersagen nicht unmittelbar empirisch verifizierbar. Das ist bei der Wahlhochrechnung anders. In dieser speziellen Situation kann der Statistiker Vorhersagen machen, die für jedermann wenige Stunden später nachprüfbar werden. Wenn diese Vorhersagen relativ treffsicher sind (und sich viele Wahlrechnungen hindurch immer wieder bewähren), dann besteht doch zumindest die Hoffnung, daß die Öffentlichkeit der statistischen Methodik eine gewisse Glaubwürdigkeit zubilligt. Der übliche „Test“ für ein wissenschaftliches Modell oder genereller für eine angewandte Wissenschaft besteht darin, in bestimmten Situationen Vorhersagen zu machen und diese Vorhersagen dann mit empirischen Beobachtungen zu vergleichen. Die Statistik hat nur ganz selten die Chance, sich dieser „Nagelprobe“ im Rahmen einer größeren Öffentlichkeit zu stellen. Die Wahlhochrechnung ist eine der ganz wenigen Chancen, diese Nagelprobe bei sehr hohem Sichtbarkeitsgrad vorzunehmen.
Außerdem beruhen Wählerstromanalysen mit Wahlergebnissen (also nicht mit Meinungsumfragedaten) ebenfalls auf derselben Methodik. Solche Analysen werden in der Öffentlichkeit aber immer wieder in Frage gestellt. Man könne

Wählerströme nicht errechnen, weil die Stimmen „kein Mascherl haben“, heißt es dann sinngemäß. Als Antwort auf diese negative Einschätzung der Möglichkeiten der Wahlanalyse ist dann ein Hinweis darauf, daß die Wahlhochrechnung auf denselben Prinzipien wie die Wählerstromanalyse beruht, und bei methodisch einwandfreier Durchführung einigermaßen zuverlässig funktioniert, unter Umständen doch vertrauensbildend. Nach wie vor besteht in der Öffentlichkeit das Problem, daß man zwischen Analysen aufgrund von Umfragen und Analysen aufgrund von Wahlergebnisse nicht genau genug unterscheidet. Da besteht dann die Gefahr, daß die immer schlechter werdende Vorhersagegenauigkeit der Umfrageanalysen die Glaubwürdigkeit aller Wahlanalysen negativ beeinflußt. Treffsichere Wahlhochrechnungen, bei deren Präsentation auch darauf hingewiesen wird, daß sie sich methodisch von Prognosen aus Umfragedaten deutlich unterscheiden, können auch in solchen Fällen ein Beitrag dazu sein, die Glaubwürdigkeit von statistischen Prognosemethoden zu erhöhen. Der Autor des vorliegenden Beitrags muß in diesem Fall allerdings anmerken, daß es ziemlich schwierig ist, Journalisten in der doch sehr gespannten Situation einer laufenden Wahlhochrechnung dazu zu bringen, in kurzen, aber korrekten Formulierungen die Datenbasis und die verwendete Methodik der Wahlhochrechnung einigermaßen klar darzustellen und insbesondere auf den Unterschied zwischen auf ausgezählten Ergebnissen beruhenden Wahlhochrechnung und auf Umfragen basierenden Prognosen hinzuweisen.

Literatur:

[1] Brown P. und Payne, C.,: Election Night Forecasting. *Journal of the Royal Statistical Society, A,* 1975, 138, S. 463-498.
[2] Bruckmann G.: Schätzung von Wahlresultaten aus Teilergebnissen, Wien 1966.
[3] McCarthy C. und Ryan, T.M.: Estimates of Voter Transition Probabilities from the British General Elections of 1974, *Journal of the Royal Statistical Society, A,* 1977, 140, S. 78-85.
[4] Neuwirth E.: Schätzung von Wählerübergangswahrscheinlichkeiten. *In: Wahlanalyse (Hrsg. M. Holler),* tuduv-Verlag 1984.