Department of Political Science, University of Aarhus Tel.: 89 42 12 71
Universitetsparken, DK-8000 Århus C Fax.: 86 13 98 39
Søren Risbjerg Thomsen E-mail:srt@ps.au.dk

February 2006

Models for Multi-Level Voting Behaviour

Exercises, Session 8 – Ecological inference using regression analysis

The purpose of ecological inference is to estimate individual-level covariation from ecological covariation. Individual-level covariation does not describe the behavior of the single individual voter, but it describes covariation across individuals. Examples are tables for voter transitions between parties from one election to the next and tables for party choice within different social classes, in short: class voting. In this session we will show how estimates of individual covariation can be derived from aggregate covaration using the classical regression approach.

When estimating 2×2 tables such as Table 8.1 we consider dichotomous individual-level voting behavior such as voting for one of the socialist parties versus for one of the non socialist parties. Another example could be belonging to the working class versus belonging to the middle class. When we consider voter transitions in 2×2 tables p is the proportion (a value between 0 and 1) voting socialist at the second election among those who also voted socialist at the first election. q is the proportion voting socialist at the second elections among non socialist at the first election. X_i is the percentage (a value between 0 and 100) voting socialist at the first elections in geographical district no. i and Y_i is the percentage voting socialist at the second election in the same district.

Table 8.1 Transitions rates between socialist and non socialist parties

	2001		
1998	Socialist	Non socialist	Total
Socialist	p	1-p	1
Non soc.	q	1-q	1

When we consider class-voting in 2×2 tables p is the proportion voting socialist in a single election among working-class voters, while q is the proportion voting socialist among middle-class voters. X_i is the percentage belonging to the working class in geographical unit no. g and Y_g is the percentage voting socialist in the unit at a single election.

Data preparation

The do-file for this session is DKdis03.do. For data preparation we use the commands for constructing party percentages and percentages in socio-economic groups from DKdis01_Problems.do. Further to create shares of dichotomies we use the commands in Table 8.2.

Table 8.2 Commands for creating shares of dichotomies

```
* Voting dichotomies

* Pct voting for socialist parties
egen validvot98 = rsum(dpoe98 - xdp98)
gen psoc98 = (dpoe98+dpf98+dpa98)/validvot98*100
egen validvot01 = rsum(dpoe01 - xdp01)
gen psoc01 = (dpoe01+dpf01+dpa01)/validvot01*100

* Pct non socialist voters
gen pnso98 = 100 - psoc98
gen pnso01 = 100 - psoc01

* Class dichotomies

* Pct working class voters 1999
gen pwrk99 = (oclwc99+ocwrk99+ocupl99)/(votep99-ocret99)*100

* pct middle class voters 1999
gen pmid99 = 100 - pwrk99
```

The percent of socialist votes is computed in shares of valid votes, so we are excluding spoilt votes and abstainers from the analysis. Notice that we are using the rsum() function, which adds all variables from dpoe98 to xdp98 by setting missing values equal to zero. If we instead were using the sum() function, valid votes would be set to missing in cases where votes for independent candidates (xdp98) were missing. When computing share of voters in the working class (including lower white collar and unemployed) we exclude retired persons. We use class data from 1999 both for the analysis of voting in 1998 and 2001 assuming that class proportions only change slowly.

Estimating 2×2 tables

The Goodman model for estimating p (the proportion voting socialist at the second election among those who voted socialist at the first election) and q (the proportion voting socialist at the second election among those who voted non socialist at the first election) is

$$Y_g = pX_g + q(100 - X_g) + e_g \ ; \quad g = 1, 2, \ . \ . \ . \ . \ , N \ , \tag{8.1}$$

and solving for X_g gives

$$Y_g = 100q + (p - q)X_g + e_g \ ; \quad g = 1, 2, \ . \ . \ . \ . \ , N \ . \tag{8.2}$$

i.e., by a regression analysis across units with Y as dependent variable and X as independent variable the intercept is

$$\text{intercept} = 100q \tag{8.3}$$

and the slope is

$$\text{slope} = p - q \,. \quad (8.4)$$

Hence, the individual-level voting behavior q and p are estimated

$$q = \text{intercept}/100 \quad (8.5)$$

$$p = \text{slope} + q \,. \quad (8.6)$$

Table 8.3 shows the commands for estimating the coefficients in equation (8.2) and drawing the relevant scatterplot.

Table 8.3 Commands for estimating dichotomous voter transitions and drawing relevant scattergram

```
* Ecological inference by regression
************************************

* 2 x 2 tables

* Voter transitions

* Simple regression analysis weighted by unit size
regress psoc01 psoc98 [aweight=validvot01 ]

* Get coefficients
matrix coefs = e(b)
gen a = coefs[1,2]
gen b = coefs[1,1]
gen pct = (_n-1)/(_N-1)*100
gen tendency = a + b*pct

* Draw symmetric scattergram
twoway (scatter psoc01 psoc98, sort) (line tendency pct, sort clpat(solid)),/*
*/ ytitle(Socialist parties 2001: Pct, margin(medsmall)) yscale(range(0 100))/*
*/ xtitle(Socialist parties 1998: Pct, margin(medsmall)) xscale(range(0 100))/*
*/ xlabel(0(10) 100, grid) legend(off) ysize(4) xsize(5)/*
*/ ylabel(0(10) 100, grid) legend(off)/*
*/graphregion(fcolor(white) lcolor(black))
drop a b pct tendency
```

The simple regression analysis with percent share of socialist in 2001 as dependent variable and percent share of socialist in 1998 as independent variable is weighted by unit size. Table 8.4 shows the results. According to equation (8.5) q (the proportion voting socialist at the second election among those who voted non socialist at the first election) is equal to -0.0057 i.e., a negative proportion of about 0.6 percent. Since this is not logical possible it is customary to set this proportion equal to 0. Then, according to equation (8.6) p (the proportion voting socialist at the second election among those who voted socialist at the first election) is estimated to be 0.832 + 0 = 0.832, or 83.2 percent.

[Mention other regression approaches, e.g. Gary king – Demonstrate the King ei program?]

Table 8.4 Results of simple regression analysis

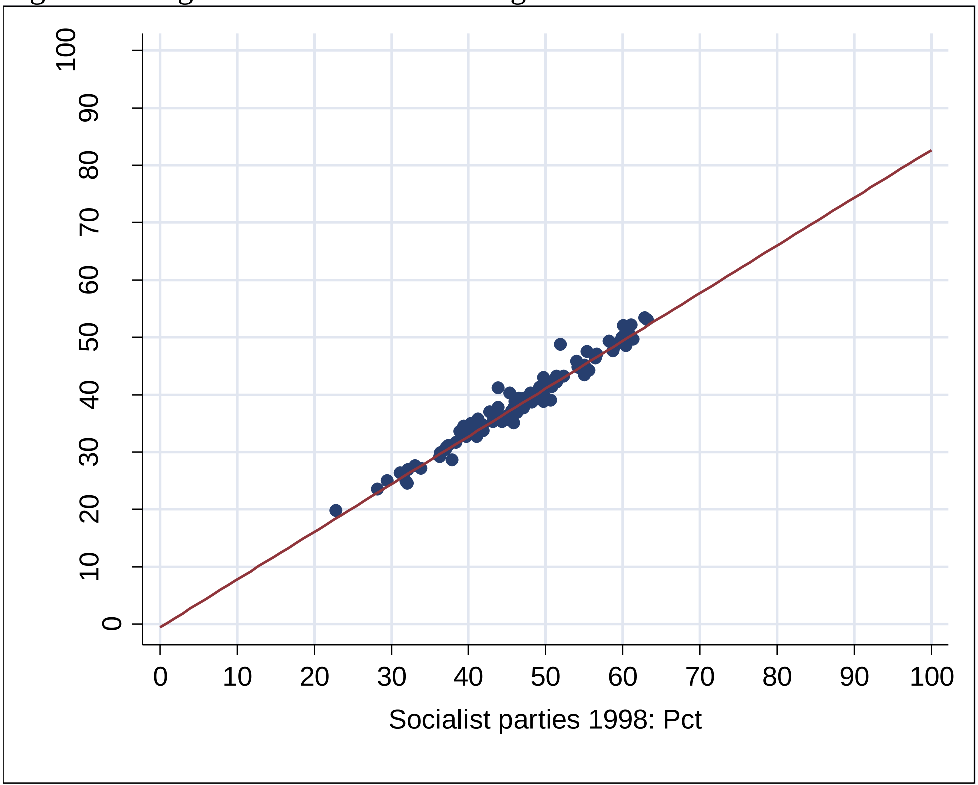
```
      Source |       SS       df       MS              Number of obs =     103
-------------+------------------------------           F(  1,   101) = 3009.55
       Model |  4272.0065     1   4272.0065            Prob > F      =  0.0000
    Residual | 143.367805   101  1.41948322            R-squared     =  0.9675
-------------+------------------------------           Adj R-squared =  0.9672
       Total |  4415.3743   102  43.2879834            Root MSE      =  1.1914

------------------------------------------------------------------------------
      psoc01 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      psoc98 |   .8322944   .0151714    54.86   0.000     .8021984    .8623904
       _cons |  -.5714065   .7102109    -0.80   0.423    -1.980274    .8374609
------------------------------------------------------------------------------
```

When drawing the relevant graph we get the estimates from the memory as usual, but also create a variable called pct by

```
gen pct = (_n-1)/(_N-1)*100
```

where _n is the unit number (from 1 to 103) and _N is the total number of units (103). The variable varies from 0 to 100 and is useful for drawing the straight line in Figure 8.1.

Figure 8.1 Regression line for estimating voter transitions from 1998 to 2001



One can see from the graph that in a hypothetical district with 100 percent voting socialistic in 1998, about 83 pct. would vote socialistic again, and in a hypothetical district where 0 percent voted socialistic in 1998 about 0 percent would vote socialistic in 2001. Thus the estimates of p and q are based on extrapolation outside the range of the independent variable.

The estimates of voter transitions by ecological regression are shown in Table 8.5. The results look somewhat extreme. Thus, it is not likely that none of the former non socialist should shift to the socialists. Unfortunately, with the regression method one can often find outright wrong estimates less than o or greater that 100 percent.

Table 8.5 Regression estimates of voter transitions

1998\2001	Soc.	Non soc.	Total
Soc.	83	17	100
Non soc.	0	100	100
Total	38	62	100

It is also possible to estimate *p* and *q* directly from equation (8.1) by separate multiple regression analysis for each of the two outcomes in 2001 (socialist or nonsocialis) by considering *X* as the one independent variable and (100-*X)* as the other independent variable in both regressions. One should think that this would be impossible because of the perfect collinearity between *X* and *100-X*, but it is actually possible if on omit the intercept in the regression analysis (no constant) (c.f. Achen & Shively 1995, p. 33, footnote 3). Table 8.6 shows the commands for the direct estimation of the transition proportions, and Table 8.7 show the results.

Table 8.6 Commands for direct estimation of transition proportions

```
* Direct estimates of transition proportions

* Voter transitions to soc 2001
regress psoc01 psoc98 pnso98 [aweight=validvot01], noconstant

* Voter transitions to non soc 2001
regress pnso01 psoc98 pnso98 [aweight=validvot01], noconstant
```

Table 8.7 Direct estimates of transition proportions

```
      Source |       SS       df       MS              Number of obs =     103
-------------+------------------------------           F(  2,   101) =53493.47
       Model |   151866.18     2  75933.0898           Prob > F      =  0.0000
    Residual |  143.367805   101  1.41948322           R-squared     =  0.9991
-------------+------------------------------           Adj R-squared =  0.9990
       Total |  152009.547   103  1475.82085           Root MSE      =  1.1914

------------------------------------------------------------------------------
      psoc01 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      psoc98 |   .8265804   .0082509   100.18   0.000     .8102127     .842948
      pnso98 |  -.0057141   .0071021    -0.80   0.423    -.0198027    .0083746
------------------------------------------------------------------------------


      Source |       SS       df       MS              Number of obs =     103
-------------+------------------------------           F(  2,   101) =       .
       Model |  402066.208     2  201033.104           Prob > F      =  0.0000
    Residual |   143.36777   101  1.41948287           R-squared     =  0.9996
-------------+------------------------------           Adj R-squared =  0.9996
       Total |  402209.576   103  3904.94734           Root MSE      =  1.1914

------------------------------------------------------------------------------
      pnso01 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      psoc98 |   .1734196   .0082509    21.02   0.000      .157052    .1897873
      pnso98 |   1.005714   .0071021   141.61   0.000     .9916254    1.019803
------------------------------------------------------------------------------
```

In Table 8.7 one can see directly from the first regression analysis that about 83 pct. (0.8265) of those who voted socialist in 1998 (soc98) stayed with the socialist in 2001 (soc01) and about 0 percent (-0.0057) went from non socialist in 1998 (nso98) to soc01. From the second regression analysis one can see that about 17 pct. went from soc98 to nso01 and that about 100 pct. Of those who voted nso98 stayed with nso01. Be aware that R^2 now is conspicuously high since it has a different interpretation as the explained share of variation in Y^2 instead of the usual explained share of variation in $(Y-\bar{Y})^2$.

By exactly the same technique we estimate the proportions voting either socialist or non socialist in 1998 within either the working class or the middle class. The line graph is shown in Figure 8.2 and the direct estimates in Table 8.9.

Table 8.8 Commands for estimation of class voting

```
* Class voting

* Simple regression analysis weighted by unit size
regress psoc98 pwrk99 [aweight=validvot98]

* Get coefficients
matrix coefs = e(b)
gen a = coefs[1,2]
gen b = coefs[1,1]
gen pct = (_n-1)/(_N-1)*100
gen tendency = a + b*pct

* Draw symmetric scattergram
twoway (scatter psoc98 pwrk99, sort) (line tendency pct, sort clpat(solid)),/*
*/ ytitle(Socialist parties 1998: Pct, margin(medsmall)) yscale(range(0 100))/*
*/ xtitle(Working class: Pct, margin(medsmall)) xscale(range(0 100))/*
*/ xlabel(0(10) 100, grid) legend(off) ysize(4) xsize(5)/*
*/ ylabel(0(10) 100, grid) legend(off)/*
*/graphregion(fcolor(white) lcolor(black))
drop a b tendency

* Class and socialist vote 1998
regress psoc98 pwrk99 pmid99 [aweight=validvot98 ], noconstant

* Class and non socialist vote 1998
regress pnso98 pwrk99 pmid99 [aweight=validvot98 ], noconstant
```

Figure 8.2 Regression line for estimating class voting in 1998

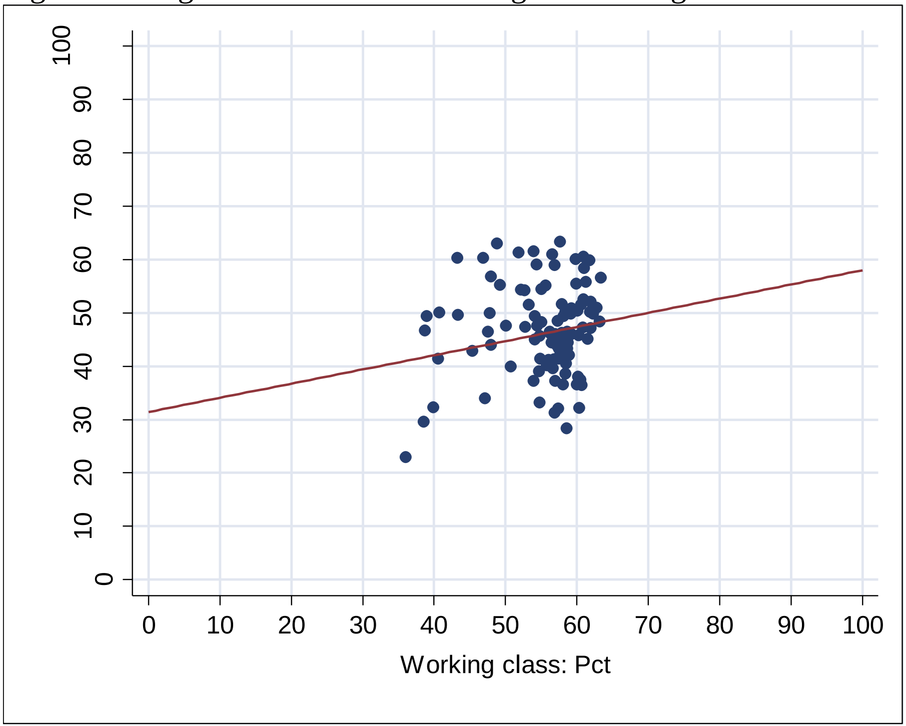


Table 8.9 Direct estimates of class voting

```
      Source |       SS       df       MS              Number of obs =     103
-------------+------------------------------           F(  2,   101) = 1862.30
       Model |  219915.236     2  109957.618           Prob > F      =  0.0000
    Residual |   5963.4386   101  59.0439465           R-squared     =  0.9736
-------------+------------------------------           Adj R-squared =  0.9731
       Total |  225878.674   103  2192.99684           Root MSE      =   7.684

------------------------------------------------------------------------------
      psoc98 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      pwrk99 |   .5797731   .0591402     9.80   0.000     .4624548    .6970914
      pmid99 |   .3142573   .0737674     4.26   0.000     .1679226    .4605919
------------------------------------------------------------------------------


      Source |       SS       df       MS              Number of obs =     103
-------------+------------------------------           F(  2,   101) = 2528.33
       Model |  298565.147     2  149282.574           Prob > F      =  0.0000
    Residual |  5963.43841   101  59.0439446           R-squared     =  0.9804
-------------+------------------------------           Adj R-squared =  0.9800
       Total |  304528.586   103  2956.58821           Root MSE      =   7.684

------------------------------------------------------------------------------
      pnso98 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      pwrk99 |   .4202269   .0591402     7.11   0.000     .3029086    .5375452
      pmid99 |   .6857427   .0737674     9.30   0.000     .5394081    .8320774
------------------------------------------------------------------------------
```

The class data is from 1999, but since the geographical distribution of class only changes slowly this is of little importance. The estimates in Table 8.9 show that 58 percent of the working class voted for the socialist parties while only 31 percent of the middle class did so. Correspondingly 42 percent of the working class voted for the non socialists and 68 percent of the middle class did so. These estimates are probably somewhat unstable because of the relatively weak linear relation in Figure 8.2 (R^2 of the simple regression is only 0.04).

Validity of the estimates

The validity of the estimates of the individual voting behavior using the classical regression approach is questionable. The individual level estimates of the 2 x 2 tables above from the Danish election surveys are somewhat different.[1] Thus the percent of stable socialist voters from 1998 to 2001 was 89 pct in the surveys, while the ecological estimate were 83 pct; and the percent of stable non-socialist voters in the survey was 91 pct, while the ecological estimate was 100 pct. In the 1998 survey the proportion voting socialist were 55 pct in the working class and 39 pct in the middle class, while the ecological estimates were 58 and 31. These ecological estimates are not terrible wrong but tend to show a stronger association in the 2 x 2 table than is the case with the survey results. The class estimates are probably also somewhat unreliable because of the weak linear association in Figure 8.2.

A major problem with the regression technique is that p and q are assumed to be constant across units, which is not necessarily the case. This is especially evident when we acquire inadmissible values below zero or above one.

Estimating m×n tables

For m×n tables the voter transitions p_{jk} is the proportion choosing party no. j at the second election out of those who chose party no. k at the first election. X_{ji} is percentage support for party no. j at the first elections in unit no. i and Y_{ji} is the percentage voting for the same party in the same unit at the second election.

For m×n tables for class-voting p_{jk} is proportion choosing party no. j at a single election out of those belonging to class no. k . X_{ji} is percentage belonging to class no. j Y_{ji} is the percentage voting for party no. j in the same unit.

Table 8.10 shows the notation for voter transitions between more than two parties from one election to the next.

Table 8.10 Notation for voter transitions in a m×n table

Party no. at first election	Party no. at second election 1	2	. . .	k	. . .	n	Total
1	p_{11}	p_{12}	. . .	p_{1k}	. . .	p_{1n}	1
2	p_{21}	p_{22}	. . .	p_{2k}	. . .	p_{2n}	1
.	. . .	. . .	. . .	. . .	. . .	. . .	.
j	p_{j1}	p_{j2}	. . .	p_{jk}	. . .	p_{jn}	1
.	. . .	. . .	. . .	. . .	. . .	. . .	.
m	p_{m1}	p_{m2}	. . .	p_{mk}	. . .	p_{mn}	1

p_{jk} is the proportion choosing party no. k at the second election out of those who chose party no. j at the first election. The Goodman model for estimating all transition rates in column no. k of Table 8.10 is

[1] The estimates with survey data is made with the do-files Master98b.do and Master01b.do by weighting the data so that the party shares in the surveys are the same as in the election results. The results are shown in detail in Surveys.xls.

$$Y_{ki} = p_{1k}X_{1i} + p_{2k}X_{2i} + \ldots + p_{jk}X_{ji} + \ldots + p_{mk}X_{mi} + e_{ki}; \quad i = 1, 2, \ldots, N \tag{8.7}$$

According to this model (8.7) we can estimate the transition rates $p_{1k}, p_{2k}, \ldots, p_{mk}$ by multiple regression analysis with percentage support Y_k at the second election for party no. k as dependent variable and the percentage support $X_1, X_2, \ldots, X_m$ at the first election for each of the parties as independents variables. Even though the independent variables are linear dependent by

$$X_1 + X_2 + \ldots + X_m = 100 , \tag{8.8}$$

multicollinearity is avoided by regression through the origin i.e., no intercept. Again, one should be aware that R^2 for each k indicates the explained share of variation in Y_k^2 instead of the usual explained share of variation in $(Y_k - \bar{Y}_k)^2$. Again, it is most practical to estimate all columns in Table 8.10, even though the last one can be computed from the others since each row add up to one. There is no guarantee that this method avoids inadmissible values below zero or above 1.

Exactly the same method can also be used to estimate a table with occupational groups as rows and party choice as columns. Table 8.11 shows the commands estimating voter transitions between all parties) (except the very small Democratic Reform Party, the independent candidates and spoilt votes) from 1998 to 2001 and for estimating vote for all parties within all occupational groups.

Table 8.11 Commands for estimating m x n tables, voter transitions and class vote

```
* m x n tables

* Multi-party voter transitions
foreach v of varlist dpoe01 dpf01 dpa01 dpb01 dpd01 dpq01 dpc01 /*
  */ dpv01 dpo01 dpz01 absdp01 {
regress p`v' pdpoe98 pdpf98 pdpa98-pdpz98 pabsdp98 [aweight=votdp01 ], no-
constant
  }

* Multi-party class voting
foreach v of varlist dpoe01 dpf01 dpa01 dpb01 dpd01 dpq01 dpc01 /*
  */ dpv01 dpo01 dpz01 absdp01 {
regress p`v' pocfar99-pocret99 [aweight=votdp01 ], noconstant
  }
```

The output is somewhat voluminous since one has to make a multiple regression analysis for each column of the table. As examples we show as well ecological estimates of voter transitions to the social democrats (party a) in Table 8.12, and the class voting for the same party in Table 8.13.

Table 8.12 Voter transitions to the Social Democrats

pdpa01	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pdpoe98	.1702075	.1535079	1.11	0.270	-.1346724	.4750873
pdpf98	-.0639724	.1168673	-0.55	0.585	-.2960808	.1681361
pdpa98	.8469291	.0247744	34.19	0.000	.7977249	.8961333
pdpb98	.1273828	.1519552	0.84	0.404	-.1744133	.4291789
pdpd98	.0097727	.2178115	0.04	0.964	-.4228197	.4423651
pdpq98	.1103663	.1073247	1.03	0.306	-.1027898	.3235224
pdpc98	-.0040661	.0501007	-0.08	0.935	-.1035703	.0954382
pdpv98	.0263883	.0300951	0.88	0.383	-.0333832	.0861597
pdpo98	-.0044897	.0924878	-0.05	0.961	-.1881785	.1791991
pdpz98	.0157998	.0652722	0.24	0.809	-.1138365	.1454361
pabsdp98	-.1455527	.0668825	-2.18	0.032	-.2783871	-.0127183

Table 8.13 Class voting for the Social Democrats

pdpa01	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pocfar99	-.0957321	.3048029	-0.31	0.754	-.7008423	.5093781
pocslf99	-2.417848	.6354048	-3.81	0.000	-3.679286	-1.15641
pochwc99	.3633342	.1642196	2.21	0.029	.0373172	.6893512
poclwc99	.2418965	.1781257	1.36	0.178	-.1117277	.5955206
pocwrk99	.5693752	.1078236	5.28	0.000	.3553183	.7834321
pocupl99	1.776762	.4711709	3.77	0.000	.8413698	2.712155
pocstd99	-.1626851	.1047115	-1.55	0.124	-.3705636	.0451934
pocret99	.2988578	.0544024	5.49	0.000	.1908553	.4068602

The results are not entirely meaningless since many former Social Democrats from 1998 seem to vote for the party again in 2001. Similarly the high proportion of workers voting for the party in 2001 is likely. But the large negative share of former abstainers voting for the Social democrats in 2001 (-15 pct in Table 8.12) and the 177 pct unemployed voting for the party (Table 8.13) are inadmissible. Compare with survey results in Survey.xls.

Repairing the regression approach

Because of the frequent occurrence of inadmissible estimates with the Goodman model, researches have tried to modify it in many ways without altering the fundamental asymmetric approach with dependent and independent manifest variables (c.f. Achen and Shively, 1995, chapter 5-6).

The first step is to acknowledge that the transition rates are not constant across units. For the 2×2 table the transition rates *p* and *q* in the (Table 8.1) can have different values in different units. Thus equation (8.1) should be rewritten

$$Y_i \;=\; p_i X_i + q_i(100 - X_i) + e_i \;\; ; \quad i = 1, 2, \;.\;.\;.\;.\;, N\,, \tag{8.9}$$

However, if p_i and q_i are allowed to vary freely this “accounting equation” is trivial because it is always true. So, one must impose restrictions on p_i and q_i to estimate individual level voting behavior.

For example one could argue, that p and q could vary across units by linear dependence on some third variable Z (Crewe and Paine, 1976). Hence

$$\begin{aligned} p_i &= a_1 + b_1 Z_i \\ q_i &= a_2 + b_2 Z_i \end{aligned} \qquad i = 1, 2, \; . \; . \; . \; . \; , N \tag{8.10}$$

Substituting (8.10) into (8.9) yields

$$Y_i = 100a_1 + (a_1 - a_2)X_i + 100b_2 Z_i + (b_1 - b_2)Z_i X_i + e_i \;; \quad i = 1, 2, \; . \; . \; . \; . \; , N \tag{8.11}$$

This model can be estimated with Y as dependent variable and X, Z and the interaction term ZX as independent variables, and we can then compute the four parameters a_1 , a_2 , b_1 , and b_2 . Further, average values of p and q in the transition table can be estimated by inserting the average value of Z in (8.10).

The trouble with a model as (8.10) is, that one can always experiment with different Z-variables until a reasonable result is achieved. However, this does not guarantee that the same Z-variable will work just as well in another election. Further, it does not guarantee admissible estimates.

Another approach builds on the precondition, that the transition rates should be logical possible. A simple version, “the method of bounds” were suggested by Duncan and Davis (1953; cited from Achen and Shively, 1995, pp. 191-192). It does not give a precise estimate of the transition rates, but the upper and lower limit of each rate within each district. For example, within a district where 90 pct. are working class and 60 pct. are Democrats, it is impossible that less than 30 percent of all voters are working-class Republicans. And it is impossible that more than 40 percent belong to the same group, so the proportion of working class voters voting for Republicans should vary between 33 and 44 percent. However, if the marginal proportions are more balanced the boundaries are wider.

A combination of the regression method and the method of bounds is the so-called “constrained regression” (Lewin et al., 1972; Hoschka and Schunk, 1975) that estimates constant transition rates under the side-condition that they should be within the logical possible range. However, this method tends to give estimates too close to the borders of the possible range.

A recent approach by Gary King (1997) hypothesizes that the two-dimensional distribution (p, q) across units should follow a special distribution within the possible range, the truncated bivariate normal distribution. However, estimates by the King method seem not to be very different from those obtained with the Goodman model, and the method is apparently not able to handle tables larger than 2×3.

A common problem with all the cited methods, except the Duncan and Davis method of bounds, is that they are asymmetric, by modeling dependent variables of voting behavior as conditioned on independent variables such as previous voting behavior or social class. As discussed in Session 7, this approach can often give biased results, since it disregards random disturbances in the independent variables. An alternative approach is to regard all manifest variables as conditioned on latent variables – the latent structure approach.

Problems

Problem 1
Estimate the 2 x 2 table for class voting in 2001

Problem 2
Estimate multi-party class voting for 1998.