Chapter 15
Verhulst and Poisson Distributions

The logistic curve (or Verhulst sigmoid curve) sometimes is used in spatial econometrics (see, e.g., Domencich and McFadden, 1975; Paelinck and Klaassen, 1979, pp. 68–72, 156–168). Two examples will be given hereafter, one for estimation in the binary case, the other for a dynamic specification. A related Poisson distribution problem is then treated; the latter distribution is less frequently used, because count data have to be available for econometric treatment.

15.1 Robust Estimation in the Binary Case: A Linear Logistic Estimator (LLE)

For a binary variable $z = 1$, let

$$d_{1i} \triangleq 1 - (1 + \exp(\mathbf{a'x_i} + b_i))^{-1}, \tag{15.1}$$

which, for a variable with subscript *i*,. is the natural distance between *1* and the logistic curve; equally, for $z = 0$, let for a variable with subscript *j*

$$d_{0j} \triangleq (1 + \exp(\mathbf{a'x_j} + b_j))^{-1} \tag{15.2}$$

which is that distance of the logistic curve from *0*.

Furthermore let

$$\mathbf{a'x_{1i}} + b_i = -\ln\left(d_{i1}{}^{-1} - 1\right) \triangleq \delta_{1i}, \tag{15.3}$$

and

$$-\mathbf{a'x_{0j}} - b_j = -\ln\left(d_{0j}{}^{-1} - 1\right) \triangleq \delta_{0j}, \tag{15.4}$$

D.A. Griffith, J.H.P. Paelinck, *Non-standard Spatial Statistics and Spatial Econometrics*, Advances in Geographic Information Science 1,
DOI 10.1007/978-3-642-16043-1_15,

In both cases, $\partial\delta/\partial d > 0$, and if $d_i = d_j$, then $\delta_i = \delta_j$.

Minimizing $\sum_i \delta_{1i} + \sum_j \delta_{0j}$, and normalizing the vector **δ**, **i** being the unit column vector, namely

$$\min \mathbf{i}'\delta - \lambda/2(\boldsymbol{\delta}'\boldsymbol{\delta} - c), \tag{15.5}$$

yields

$$\mathbf{i} = \lambda\boldsymbol{\delta} = \mathbf{X}^*\mathbf{a}, \tag{15.6}$$

with

$$\mathbf{X}^* = \begin{bmatrix} \mathbf{X_1} \\ -\mathbf{X_0} \end{bmatrix} \tag{15.7}$$

and ***a*** including the constant. Because $\lambda < 0$, switching the sign of $\mathbf{X_0}$ in Eq. (15.7), replaces **i** by

$$\mathbf{i}^* = \begin{bmatrix} -\mathbf{i} \\ \mathbf{i} \end{bmatrix}$$

A linear estimator of ***a*** is given by

$$\mathbf{a} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{i}, \tag{15.8}$$

with λ conveniently set equal to *–1*.

Therefore, for unit λ:

$$\mathbf{V(a)} = (\mathbf{X}'\mathbf{X})^{-1}, \tag{15.9}$$

and hence pseudo-*t* values can be computed as follows

$$t_k = a_k/\sqrt{x_{kk}}, \tag{15.10}$$

The method was applied to the following (unique) explanatory variable: 0*.4, 0.5, 0.6, 0.92, 0.95, 0.98*; Table 15.1 lists the results (pseudo-*t* values in parentheses).

The results are graphically presented in Fig. 15.1.

Table 15.1 Estimation results for model (15.8)

Parameters	Values
Slope	−2.2169 (−40.7040)
Constant	1.1158 (30.3286)
R^2	0.9988

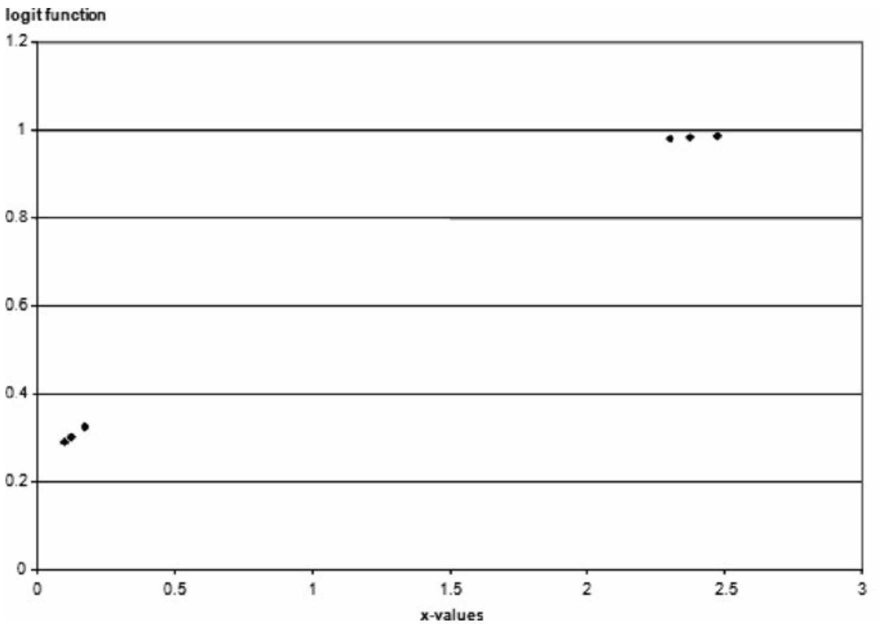

Fig. 15.1 The logistic resulting from Table 15.1

15.2 A Logistic Dynamic Share Model

Let $0 < a_{ij} < 1$ be the share of sector i in region j; $\sum_i a_{ij} = 1, \forall j$.

Let the model be specified as follows

$$a_{ijt} = \left[1 + \exp\left(\sum{}_{ij}\alpha_{ij,t-1} + \beta^{ij}\right)\right]^{-1} \tag{15.11}$$

a generalized logistic function. The superscripts refer to the subscripts of the left hand member.

From Eq. (15.11) one can derive

$$\ln(a_{ijt}^{-1} - 1) = \sum{}_{ij}\alpha_{ij}^{ij}a_{ij,,t-1} + \beta^{ij}, \tag{15.12}$$

In equilibrium, $a_{ijt} = a_{ij,t-1}$, $\forall i,j$. Thus,

$$\ln(a_{ijt}{}^{-1} - 1) = \sum_{ij}\alpha^{ij}{}_{ij}a_{ijt} + \beta^{\,ij} \tag{15.13}$$

$$= \alpha^{ij}{}_{ij}a_{ijt} + r^{ij}, \tag{15.14}$$

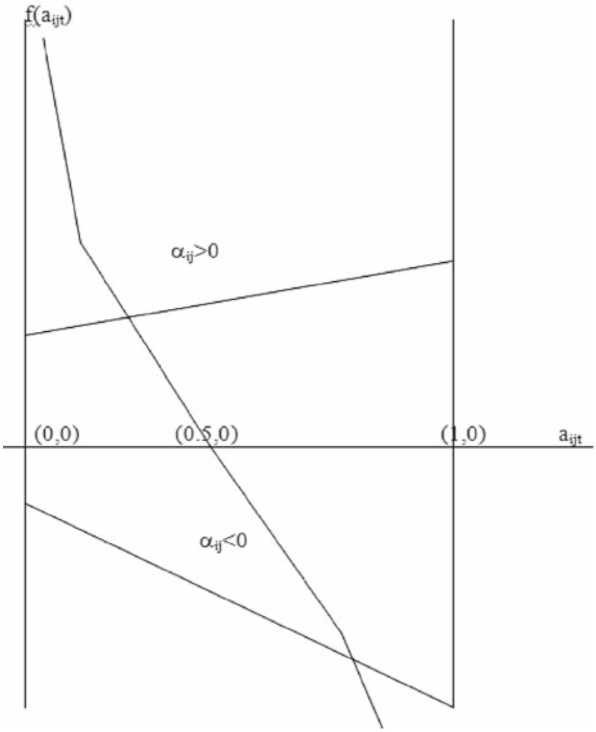

Fig. 15.2 Equilibrium solutions for (15.13) and (15.14)

where r denotes the remaining terms. There are two possibilities according to the sign of α_{ij}. Figure 15.2 shows how possible solutions look (recall that $0 < a_{ij} < 1$).

In terms of stability, the following points can be made:

(1) There exists a confiner defined as follows:

$$\max_{\mathbf{a}} \sum_{ij} \alpha^{ij}{}_{ij} a_{ij} + \beta^{ij} \tag{15.15}$$

$$\min_{\mathbf{a}} \sum_{ij} \alpha^{ij}{}_{ij} a_{ij} + \beta^{ij} \tag{15.16}$$

$$\text{s.t} \sum\nolimits_{i} a_{ij} = 1, \forall j \tag{15.17}$$

$$1 \geq a_{ij} \geq 0, \forall i, j \tag{15.18}$$

Table 15.2 Test numbers for model (15.11)

t	a_{11}	a_{21}	a_{12}	a_{22}
1	0.30	0.70	0.65	0.35
2	0.32	0.68	0.64	0.36
3	0.32	0.68	0.64	0.36
4	0.34	0.66	0.62	0.38
5	0.37	0.63	0.62	0.38
6	0.38	0.62	0.61	0.39
7	0.39	0.61	0.59	0.41
8	0.41	0.59	0.57	0.43
9	0.41	0.59	0.57	0.43
10	0.42	0.58	0.56	0.44

(2) One can linearize (Taylor-expansion around 0.5) the left-hand member of (15.12), yielding

$$\mathbf{a_t} \approx -0.25\,\mathbf{A}\mathbf{a_{t-1}} - 0.25(\mathbf{b} - 2\mathbf{i}). \tag{15.19}$$

Convergence depends on $\mathbf{-i < \lambda(A) < i}$, whereas divergence is constrained by the confiner. In the case of convergence, the attractor is:

$$\mathbf{a}^{\mathrm{o}} = -\mathbf{0}.25(\mathbf{I} + 0.25\mathbf{A})^{-1}(\mathbf{b} - \mathbf{2i}). \tag{15.20}$$

However, given the approximation, conditions (15.17) are not necessarily satisfied.

The following (fictional) numbers, reported in Table 15.2, have been used to test the model.

Estimation (see Table 15.3) was performed by minimizing the sum of squares between the observed a_{ijt}s and the SDLS endogenously generated ones (see Sect. 11.1.3). The resulting overall R^2 is 0.9989, and, moreover, conditions (15.17) are very closely satisfied in both regions, with erratic divergences not exceeding 2% [see comments about Eq. (15.20)].

Starting from the last observations, a 24-period simulation was performed. Table 15.4 shows the results, which again obey conditions (15.17) very closely. The simulation reveals the inherent dynamics of the model, which could hardly be deduced from the "observed" series; Figs. 15.3 and 15.4 portray this once more.

15.3 A Linear Poisson Distribution Estimator

The Poisson probability mass function is given by

$$\mathrm{p(n)} = \exp(-\mu)\mu^{\mathrm{n}}/\mathrm{n}!. \tag{15.21}$$

Table 15.3 Estimation results from Table 15.2

Features	α_{11}	α_{21}	α_{12}	α_{22}		
Parameters	–6.16267	11.78416	9.939417	–13.3512		
	4.125329	3.303637	4.513609	–8.34668		
	–6.61019	2.508025	–7.97125	3.446991		
	2.07069	–3.84603	–10.9558	6.180226		
	3.352769	–6.93088	2.237161	6.066311	Conditions (15.17)	
SDLS variables	0.475024	0.505434	0.443078	0.572783	0.980458	1.015861
	0.316985	0.680277	0.63812	0.360616	0.997262	0.998736
	0.324253	0.675817	0.640868	0.357918	1.000071	0.998786
	0.343588	0.656204	0.626998	0.373463	0.999792	1.000461
	0.36097	0.640596	0.616677	0.384327	1.001566	1.001004
	0.379807	0.620659	0.601062	0.401076	1.000466	1.002138
	0.394141	0.608216	0.591866	0.409607	1.002357	1.001473
	0.40885	0.590948	0.575795	0.425779	0.999798	1.001574
	0.414217	0.587613	0.571102	0.427662	1.00183	0.998764
	0.417111	0.57954	0.557563	0.439539	0.996651	0.997101
Conditions (15.12)	1.64E–09	–5.2E–10	1.2E–09	3.81E–10		
	1.27E–10	–1.1E–10	–1E–10	2.49E–11		
	1.1E–10	–1.5E–10	–8.6E–11	3.11E–11		
	1.37E–10	–6.8E–11	–1.5E–10	–1.3E–10		
	3.32E–13	–1E–10	–2.9E–10	–1.5E–10		
	–2.4E–10	1.02E–10	–6.4E–10	–4.1E–10		
	–1.9E–09	1.23E–10	–4.5E–10	5.47E–10		
	–1.2E–09	2.86E–10	–7.5E–10	1.23E–11		
	–3.3E–09	8.85E–11	2.75E–10	2.02E–09		
SDLS minus observed a_{ijt}	0.003015	0.000277	0.00188	0.000616		
	0.004253	0.004183	0.000868	0.002082		
	0.003588	0.003796	0.006998	0.006537		
	0.00903	0.010596	0.003323	0.004327		
	0.000193	0.000659	0.008938	0.011076		
	0.004141	0.001784	0.001866	0.000393		
	0.00115	0.000948	0.005795	0.004221		
	0.004217	0.002387	0.001102	0.002338		
	0.002889	0.00046	0.002437	0.000461		

Its mean μ can be written as a function of various factors, x_k; assume a linear function. Then

$$\ln[p(n_i)] = -\mu_i + n_i \ln \mu_{iI} - \ln n_i! \tag{15.22}$$

The first-order maximum likelihood conditions (the second-order ones also are satisfied, as Eq. (15.22) is concave in the parameters) are, for some parameter *a*

$$\sum_i x_{ik} = \sum_i n_i x_{ik}/\mu_i \tag{15.23}$$

Table 15.4 Simulation results

t	a_{11}	a_{21}	a_{12}	a_{22}	Conditions (15.17)	
1	0.42	0.58	0.65	0.35	1	1
2	0.602225	0.445547	0.496556	0.49466	1.047771	0.991216
3	0.671098	0.357977	0.399227	0.585599	1.029075	0.984826
4	0.663613	0.346834	0.367438	0.615801	1.010447	0.983239
5	0.611098	0.384416	0.383735	0.600675	0.995515	0.98441
6	0.537536	0.447841	0.428239	0.559126	0.985378	0.987366
7	0.457233	0.522775	0.487681	0.503906	0.980008	0.991587
8	0.382394	0.596658	0.551243	0.445049	0.979051	0.996292
9	0.325752	0.655956	0.606947	0.393426	0.981708	1.000373
10	0.295642	0.690855	0.644244	0.358716	0.986496	1.00296
11	0.291214	0.700626	0.659679	0.344255	0.99184	1.003934
12	0.304921	0.691703	0.657195	0.346493	0.996625	1.003688
13	0.328039	0.672193	0.643538	0.359153	1.000231	1.002691
14	0.353605	0.648826	0.624817	0.37652	1.002431	1.001337
15	0.376836	0.626439	0.60558	0.394357	1.003274	0.999937
16	0.394782	0.608214	0.588933	0.409789	1.002995	0.998723
17	0.406029	0.595901	0.576753	0.421086	1.001931	0.997839
18	0.410465	0.589988	0.569829	0.427519	1.000452	0.997348
19	0.409008	0.589897	0.568006	0.42923	0.998905	0.997236
20	0.403273	0.59429	0.570392	0.427041	0.997563	0.997433
21	0.395192	0.601406	0.575628	0.422208	0.996599	0.997836
22	0.386659	0.609425	0.582185	0.416144	0.996083	0.998328
23	0.379219	0.616776	0.588645	0.410161	0.995995	0.998806
24	0.373874	0.622369	0.593931	0.40526	0.996243	0.999191
25	0.371017	0.625685	0.597426	0.402014	0.996702	0.99944

If, on average, $\mu_i = 1$, condition (15.23) also is satisfied on average. Consequently

$$\mu_i / n_i = 1, \forall_i \tag{15.24}$$

from which an OLS estimator can be derived as

$$\mathbf{a} = (\mathbf{X'X})^{-1}\mathbf{X'i} \tag{15.25}$$

where for the constant term elements n_i^{-1} appear in **X,** because each element of the usual unit column vector has to be divided by the counts observed.

The model has been applied to the data reported in Table 15.5.

Table 15.6 presents the results (pseudo-t values in parentheses) for a one-variable (x_i) with the n_i^{-1} terms as required.

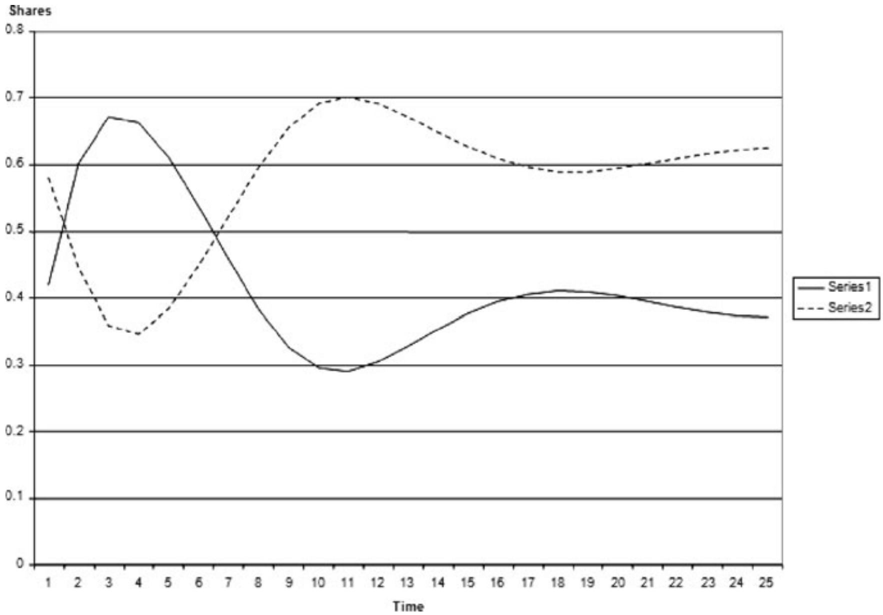

Fig. 15.3 First region simulatioins results graphed

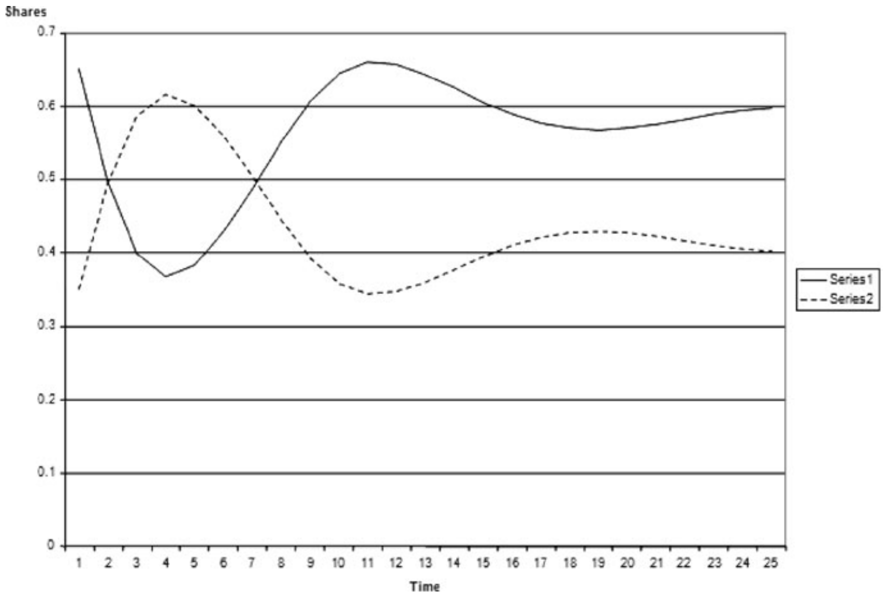

Fig. 15.4 Second region, simulation results graphed

Table 15.5 Data to apply Eq. (15.25)

x_i	n_i
3	1
7	2
8	4
11	5
13	7
17	9

Table 15.6 Estimation results using data from Table 15.5

Parameters	Values
Slope	0.4796 (5.7466)
Constant	–0.4436 (–1.2584)

15.4 Conclusion

Again very robust and simple estimators have been developed for the Verhulst and Poisson curve parameters. Although the processes might be complex, they are readily calibrated.

The examples have shown that the obtained estimation results are readily usable for consistent simulation, which moreover reveals properties that the original series do not show at once. This demonstrates the utility of longer term extrapolations, as the function—in this case, the Verhulst function—does not lead to analyses close to that of classical dynamics (see Chap. 11).