

## Capitolo 8

### MODELLI DI REGRESSIONE LOGISTICA

#### PER LE VARIABILI CONTESTUALI

Si faccia riferimento:

- all'Appendice A per il catalogo completo delle variabili, per una loro sommaria descrizione e per un elenco delle modalità da esse assumibili;
- all'Appendice B per i codici dei programmi SAS relativi alla creazione dei dataset [in particolare al programma *Dati00.sas* per la creazione del dataset di partenza e le varie riclassificazioni effettuate];
- all'Appendice C per i codici dei programmi SAS relativi ai modelli elaborati; in nota a ciascun titolo di paragrafo viene riportato il nome del programma SAS di riferimento.

## 8.1 L'ALFABETIZZAZIONE<sup>1</sup>

In questo capitolo andremo infine a studiare alcuni degli elementi contestuali che si sono rivelati significativi fattori discriminanti nella determinazione del rischio nei modelli precedenti. Una volta riconosciuti infatti *i fattori caratterizzanti i gruppi precursori della transizione*, lo scopo (come già introdotto nel Capitolo 4) diventa quello di isolare le cause prime delle trasformazioni in atto, di capire quale contesto sia stato il loro naturale terreno di sviluppo ed in quali altre situazioni sia invece possibile intervenire per favorirne la maturazione.

### LA SCELTA DELLE VARIABILI ESPLICATIVE E LA CREAZIONE DEL DATASET

Sono state inizialmente selezionate come candidate esplicative le seguenti variabili:

v013	Età attuale - Classi quinquennali
v024	Regione di residenza
v025	Ambiente di residenza: urbano o rurale
v026	Ambiente di residenza: tipo di centro abitato
v103	Ambiente di residenza d'infanzia: tipo di centro abitato
v106	Maggior grado di istruzione iniziato
v131	Etnia

Si noti come siano state introdotte nuove variabili contestuali tra le esplicative per riuscire a comprendere e a localizzare il meglio possibile le trasformazioni in atto, il ruolo della cultura, delle tradizioni, dell'ambiente d'origine e dell'ambiente di residenza attuale.

### L'ELABORAZIONE DEL MODELLO

Esaminando le stime che si ottengono dall'elaborazione del modello prima di introdurre un processo di selezione automatica, appare evidente la fortissima e ovvia dipendenza dell'alfabetizzazione dal grado d'istruzione.

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	wald Chi-Square	Pr > ChiSq
v106	Higher	1	21.8991	924.9	0.0006	0.9811
v106	Primary	1	5.0746	0.1352	1408.0739	<.0001
v106	Secondary	1	22.0237	281.2	0.0061	0.9376

---

<sup>1</sup> Vedi *Mod18.sas* [Appendice C].

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
V106 Higher	vs No education	>999.999	<0.001	>999.999
V106 Primary	vs No education	159.908	122.675	208.440
V106 Secondary	vs No education	>999.999	<0.001	>999.999

Il significato delle due variabili è troppo simile; il grado d'istruzione discrimina già in modo netto il possesso o meno dell'attributo che il modello si pone di studiare: *v106* viene perciò forzatamente esclusa dall'insieme delle esplicative.

Introdotti un processo di selezione stepwise e le variabili d'interazione, si ottengono i risultati seguenti.

Model Information

Data Set	DATI97.DATI18	
Response Variable	v108c1	Alfabetizzazione - Nuova classificazione
Number of Response Levels	2	
Number of Observations	8542	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Response Profile

Ordered Value	v108c1	Total Frequency
1	Reads	2151
2	Cannot read	6391

Probability modeled is v108c1='Reads'.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2339.1629	45	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
6.8126	5	0.2350

Summary of Stepwise Selection

Step	Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	V103		4	1	1714.6230	.	<.0001
2	V026		3	2	232.5850	.	<.0001
3	V131		11	3	224.9015	.	<.0001
4	V013		6	4	179.4794	.	<.0001
5	V024		3	5	36.4972	.	<.0001
6	V013*V024		18	6	39.6140	.	0.0024

Step	Variable Label
1	Childhood place of residence
2	De facto place of residence
3	Ethnicity
4	Age 5-year groups
5	Region
6	

Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
V013	6	50.5084	<.0001
V024	3	13.4447	0.0038
V026	3	162.2783	<.0001
V103	4	387.1515	<.0001
V131	11	197.1487	<.0001
V013*v024	18	39.2115	0.0027

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.7665	0.1294	35.0966	<.0001
V013 20-24	1	0.0825	0.1655	0.2487	0.6180
V013 25-29	1	-0.2770	0.1688	2.6928	0.1008
V013 30-34	1	0.1769	0.1864	0.9010	0.3425
V013 35-39	1	-0.3800	0.1907	3.9714	0.0463
V013 40-44	1	-0.8973	0.2193	16.7409	<.0001
V013 45-49	1	-1.3631	0.2842	22.9997	<.0001
V024 Central	1	-0.1937	0.1624	1.4236	0.2328
V024 North East	1	0.0647	0.1961	0.1089	0.7415
V024 South	1	0.4497	0.2001	5.0501	0.0246
V026 Countryside	1	-1.1703	0.1328	77.7157	<.0001
V026 Small city	1	-0.3545	0.1446	6.0130	0.0142
V026 Town	1	-0.2101	0.1443	2.1183	0.1456
V103 Abroad	1	-1.2949	0.2171	35.5734	<.0001
V103 City	1	0.00285	0.1234	0.0005	0.9816
V103 Countryside	1	-1.5651	0.1064	216.3093	<.0001
V103 Town	1	-0.2102	0.1194	3.0969	0.0784
V131 Balant	1	0.4440	0.3091	2.0632	0.1509
V131 Bambara	1	0.2697	0.2099	1.6499	0.1990
V131 Diola	1	0.9460	0.1464	41.7768	<.0001
V131 Mancagne	1	1.9000	0.5128	13.7310	0.0002
V131 Mandingue/Soce/M	1	-0.4158	0.1535	7.3366	0.0068
V131 Manjaak	1	0.3580	0.3199	1.2524	0.2631
V131 Not Senegalese	1	-0.3513	0.2475	2.0147	0.1558
V131 Other	1	0.6900	0.2000	11.8978	0.0006
V131 Poular	1	-0.2913	0.0866	11.3094	0.0008
V131 Sarakole/Soninke	1	0.3643	0.2026	3.2322	0.0722
V131 Serer	1	0.7065	0.0883	64.0195	<.0001
V013*v024 20-24 Central	1	-0.1998	0.2136	0.8755	0.3494
V013*v024 20-24 North East	1	-0.3706	0.2723	1.8525	0.1735
V013*v024 20-24 South	1	-0.2043	0.2581	0.6267	0.4286
V013*v024 25-29 Central	1	-0.3600	0.2300	2.4492	0.1176
V013*v024 25-29 North East	1	-0.1852	0.2903	0.4071	0.5235
V013*v024 25-29 South	1	-0.3883	0.2693	2.0793	0.1493
V013*v024 30-34 Central	1	-1.0494	0.2516	17.3960	<.0001
V013*v024 30-34 North East	1	-0.8113	0.3159	6.5971	0.0102
V013*v024 30-34 South	1	-1.0390	0.2949	12.4181	0.0004
V013*v024 35-39 Central	1	-0.5810	0.2576	5.0868	0.0241
V013*v024 35-39 North East	1	-0.2774	0.3211	0.7468	0.3875
V013*v024 35-39 South	1	-1.0945	0.3408	10.3126	0.0013
V013*v024 40-44 Central	1	0.2562	0.2915	0.7723	0.3795
V013*v024 40-44 North East	1	-0.0171	0.3965	0.0019	0.9656
V013*v024 40-44 South	1	-0.6025	0.3664	2.7035	0.1001
V013*v024 45-49 Central	1	-0.0318	0.3888	0.0067	0.9347
V013*v024 45-49 North East	1	-0.3535	0.5826	0.3683	0.5439
V013*v024 45-49 South	1	-0.1987	0.4498	0.1951	0.6587

### Odds Ratio Estimates

Effect		Point Estimate	95% wald Confidence Limits	
V026 Countryside	vs Capital, large c	0.310	0.239	0.402
V026 Small city	vs Capital, large c	0.702	0.528	0.931
V026 Town	vs Capital, large c	0.811	0.611	1.076
V103 Abroad	vs Capital, large c	0.274	0.179	0.419
V103 City	vs Capital, large c	1.003	0.787	1.277
V103 Countryside	vs Capital, large c	0.209	0.170	0.258
V103 Town	vs Capital, large c	0.810	0.641	1.024
V131 Balant	vs Wolof/Lebou	1.559	0.851	2.858
V131 Bambara	vs Wolof/Lebou	1.310	0.868	1.976
V131 Diola	vs Wolof/Lebou	2.575	1.933	3.431
V131 Mancagne	vs Wolof/Lebou	6.686	2.447	18.265
V131 Mandingue/Soce/M	vs Wolof/Lebou	0.660	0.488	0.891
V131 Manjaak	vs Wolof/Lebou	1.430	0.764	2.677
V131 Not Senegalese	vs Wolof/Lebou	0.704	0.433	1.143
V131 Other	vs Wolof/Lebou	1.994	1.347	2.951
V131 Poular	vs Wolof/Lebou	0.747	0.631	0.886
V131 Sarakole/Soninke	vs Wolof/Lebou	1.440	0.968	2.141
V131 Serer	vs Wolof/Lebou	2.027	1.705	2.410

### Association of Predicted Probabilities and Observed Responses

Percent Concordant	82.3
Percent Discordant	17.2
Percent Tied	0.6

### COMMENTI

- ✎ Il risultato più importante è forse il concreto e progressivo aumento della popolazione femminile alfabetizzata nelle nuove generazioni: come già sottolineato più volte, questo risultato concretizza l'influenza che tale fattore esercita realmente sull'avvio del calo della fecondità, altrimenti attestata ad un livello potenziale<sup>2</sup>.
- ✎ *v025*, la variabile che in tutti i modelli precedenti rappresentava l'effetto dell'urbanizzazione, in tal caso non risulta significativa. Non è un risultato sorprendente: viene di fatto sostituita da *v026*, indicante il tipo di centro abitato, a cui *v025* risulta ovviamente fortemente associata.
- ✎ Risulta discriminante anche l'ambiente di residenza d'infanzia, considerato separatamente da quello attuale e introdotto con la variabile *v103*. La formazione culturale e sociale di un individuo, d'altronde, nasce e matura inizialmente in tale contesto; è ovvio che le attitudini comportamentali ne vengano segnate anche nel caso di successivi trasferimenti.

---

<sup>2</sup> Vedi paragrafo 4.2.

✎ I risultati relativi all'etnia e alla regione possono apparire anomali se pensiamo alle caratteristiche delle singole popolazioni, alla loro religione, alle loro attività produttive, alle loro tradizioni<sup>3</sup>. Dobbiamo però tenere in considerazione il fatto che l'effetto esercitato da un fattore differenziale fondamentale come l'urbanizzazione, che potrebbe rappresentare buona parte delle principali differenze tra popolazioni di diversa etnia, viene stimato separatamente.

---

<sup>3</sup> Vedi paragrafo 2.3.

## 8.2 IL GRADO D'ISTRUZIONE<sup>4</sup>

### LA SCELTA DELLE VARIABILI ESPLICATIVE E LA CREAZIONE DEL DATASET

Vengono prese in considerazione tre diverse dicotomizzazioni di *v106*, associate rispettivamente ad un grado d'istruzione raggiunto (iniziato, non necessariamente portato a termine) primario, secondario e “alto”.

Sono state selezionate come candidate esplicative le seguenti variabili:

v013	Età attuale - Classi quinquennali
v024	Regione di residenza
v025	Ambiente di residenza: urbano o rurale
v026	Ambiente di residenza: tipo di centro abitato
v103	Ambiente di residenza d'infanzia: tipo di centro abitato
v131	Etnia

### L'ELABORAZIONE DEL MODELLO

Consideriamo innanzitutto i risultati ottenuti per il modello relativo al raggiungimento del grado d'istruzione primario (iniziato, non necessariamente portato a termine).

```
PROC LOGISTIC data=dati97.dati19 descending;
class v013(ref='15-19') v024(ref='West') v025(ref='Urban')
      v026(ref='Capital, large city') v103(ref='Capital, large city')
      v131(ref='Wolof/Lebou')
/ param=ref;
model v106_1 = v013 v024 v025 v026 v103
              v131
              v013*v024 v103*v025
/ selection=stepwise sle=0.05 sls=0.01;
Run;
```

#### Model Information

Data Set	DATI97.DATI19	
Response Variable	v106_1	Iniziato grado di istruzione primario
Number of Response Levels	2	
Number of Observations	8553	
Model	binary logit	
Optimization Technique	Fisher's scoring	

#### Response Profile

Ordered Value	v106_1	Total Frequency
1	Sì	2558
2	No	5995

Probability modeled is v106\_1='Sì'.

NOTE: 40 observations were deleted due to missing values for the response or explanatory variables.

---

<sup>4</sup> Vedi *Mod19.sas* [Appendice C].



Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2954.9506	45	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
6.0894	5	0.2976

Summary of Stepwise Selection

Step	Entered	Effect Removed	DF	Number In	Score Chi-Square	wald Chi-Square	Pr > ChiSq
1	V103		4	1	1955.9187	.	<.0001
2	V131		11	2	324.9851	.	<.0001
3	V013		6	3	334.1854	.	<.0001
4	V026		3	4	301.1775	.	<.0001
5	V024		3	5	59.2721	.	<.0001
6	V013*V024		18	6	49.0063	.	0.0001

Step	Variable Label
1	Childhood place of residence
2	Ethnicity
3	Age 5-year groups
4	De facto place of residence
5	Region
6	

Type III Analysis of Effects

Effect	DF	wald Chi-Square	Pr > ChiSq
V013	6	87.6006	<.0001
V024	3	30.3838	<.0001
V026	3	219.9480	<.0001
V103	4	459.2183	<.0001
V131	11	271.8855	<.0001
V013*v024	18	48.4534	0.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	wald Chi-Square	Pr > ChiSq
Intercept	1	1.4024	0.1412	98.6017	<.0001
V013 20-24	1	-0.0864	0.1757	0.2418	0.6229
V013 25-29	1	-0.5806	0.1772	10.7312	0.0011
V013 30-34	1	-0.1488	0.1946	0.5846	0.4445
V013 35-39	1	-0.8241	0.1985	17.2434	<.0001
V013 40-44	1	-1.3895	0.2254	38.0068	<.0001
V013 45-49	1	-1.9055	0.2870	44.0779	<.0001
V024 Central	1	-0.2414	0.1635	2.1796	0.1399
V024 North East	1	0.2308	0.1946	1.4066	0.2356
V024 South	1	0.6986	0.2044	11.6821	0.0006
V026 Countryside	1	-1.2739	0.1342	90.0457	<.0001
V026 Small city	1	-0.3131	0.1477	4.4959	0.0340
V026 Town	1	-0.1773	0.1467	1.4592	0.2271
V103 Abroad	1	-1.3147	0.2054	40.9538	<.0001
V103 City	1	-0.1252	0.1275	0.9648	0.3260
V103 Countryside	1	-1.7597	0.1078	266.5365	<.0001
V103 Town	1	-0.3285	0.1225	7.1930	0.0073
V131 Balant	1	0.6300	0.2984	4.4575	0.0347
V131 Bambara	1	0.2896	0.2136	1.8374	0.1753
V131 Diola	1	1.2446	0.1492	69.6034	<.0001
V131 Mancagne	1	2.5472	0.6162	17.0861	<.0001
V131 Mandingue/Soce/M	1	-0.1542	0.1465	1.1076	0.2926
V131 Manjaak	1	0.6238	0.3269	3.6419	0.0563
V131 Not Senegalese	1	-0.5932	0.2401	6.1033	0.0135

V131	Other		1	0.6760	0.2049	10.8835	0.0010
V131	Poular		1	-0.3255	0.0859	14.3666	0.0002
V131	Sarakole/Soninke		1	0.3589	0.1981	3.2829	0.0700
V131	Serer		1	0.8591	0.0876	96.1014	<.0001
V013*v024	20-24	Central	1	-0.3281	0.2195	2.2354	0.1349
V013*v024	20-24	North East	1	-0.2777	0.2683	1.0714	0.3006
V013*v024	20-24	South	1	-0.2750	0.2664	1.0656	0.3019
V013*v024	25-29	Central	1	-0.3387	0.2328	2.1169	0.1457
V013*v024	25-29	North East	1	-0.2195	0.2867	0.5862	0.4439
V013*v024	25-29	South	1	-0.5580	0.2740	4.1485	0.0417
V013*v024	30-34	Central	1	-0.9369	0.2520	13.8186	0.0002
V013*v024	30-34	North East	1	-0.8757	0.3128	7.8388	0.0051
V013*v024	30-34	South	1	-1.3290	0.2993	19.7125	<.0001
V013*v024	35-39	Central	1	-0.4826	0.2602	3.4397	0.0636
V013*v024	35-39	North East	1	-0.1948	0.3167	0.3784	0.5385
V013*v024	35-39	South	1	-1.1692	0.3332	12.3144	0.0004
V013*v024	40-44	Central	1	0.4831	0.2917	2.7427	0.0977
V013*v024	40-44	North East	1	0.2549	0.3788	0.4530	0.5009
V013*v024	40-44	South	1	-0.5728	0.3565	2.5813	0.1081
V013*v024	45-49	Central	1	0.1634	0.3847	0.1804	0.6710
V013*v024	45-49	North East	1	-0.4789	0.5876	0.6641	0.4151
V013*v024	45-49	South	1	-0.3877	0.4472	0.7516	0.3860

#### Odds Ratio Estimates

Effect		Point Estimate	95% wald Confidence Limits	
V026 Countryside	vs Capital, large c	0.280	0.215	0.364
V026 Small city	vs Capital, large c	0.731	0.547	0.977
V026 Town	vs Capital, large c	0.838	0.628	1.117
V103 Abroad	vs Capital, large c	0.269	0.180	0.402
V103 City	vs Capital, large c	0.882	0.687	1.133
V103 Countryside	vs Capital, large c	0.172	0.139	0.213
V103 Town	vs Capital, large c	0.720	0.566	0.915
V131 Balant	vs Wolof/Lebou	1.878	1.046	3.370
V131 Bambara	vs Wolof/Lebou	1.336	0.879	2.030
V131 Diola	vs Wolof/Lebou	3.472	2.591	4.651
V131 Mancagne	vs Wolof/Lebou	12.771	3.817	42.733
V131 Mandingue/Soce/M	vs Wolof/Lebou	0.857	0.643	1.142
V131 Manjaak	vs Wolof/Lebou	1.866	0.983	3.541
V131 Not Senegalese	vs Wolof/Lebou	0.553	0.345	0.885
V131 Other	vs Wolof/Lebou	1.966	1.316	2.938
V131 Poular	vs Wolof/Lebou	0.722	0.610	0.855
V131 Sarakole/Soninke	vs Wolof/Lebou	1.432	0.971	2.111
V131 Serer	vs Wolof/Lebou	2.361	1.988	2.804

#### Association of Predicted Probabilities and Observed Responses

Percent Concordant	84.1
Percent Discordant	15.4
Percent Tied	0.5

Seguono i risultati per il modello relativo al raggiungimento del grado d'istruzione secondario. In questo caso nessuna delle due variabili d'interazione risulta significativa.

#### Model Information

Data Set	DATI97.DATI19	
Response Variable	v106_2	Iniziato grado di istruzione secondario
Number of Response Levels	2	
Number of Observations	8553	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Response Profile

Ordered Value	v106_2	Total Frequency
1	Si	864
2	No	7689

Probability modeled is v106\_2='si'.

NOTE: 40 observations were deleted due to missing values for the response or explanatory variables.

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1315.8289	24	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
49.5939	26	0.0035

Summary of Stepwise Selection

Step	Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	V103		4	1	1104.1337	.	<.0001
2	V026		3	2	155.3636	.	<.0001
3	V131		11	3	101.4157	.	<.0001
4	V013		6	4	31.6162	.	<.0001

Step	Variable Label
1	Childhood place of residence
2	De facto place of residence
3	Ethnicity
4	Age 5-year groups

Type III Analysis of Effects

Effect	DF	wald Chi-Square	Pr > ChiSq
V013	6	31.0086	<.0001
V026	3	165.1438	<.0001
V103	4	207.3937	<.0001
V131	11	97.9409	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	wald Chi-Square	Pr > ChiSq
Intercept	1	-0.5953	0.1118	28.3654	<.0001
V013 20-24	1	0.0973	0.1144	0.7239	0.3949
V013 25-29	1	-0.1206	0.1237	0.9512	0.3294
V013 30-34	1	-0.0408	0.1324	0.0950	0.7579
V013 35-39	1	-0.4000	0.1454	7.5636	0.0060
V013 40-44	1	-0.5266	0.1734	9.2283	0.0024
V013 45-49	1	-0.7663	0.2237	11.7306	0.0006
V026 Countryside	1	-1.7836	0.1404	161.3655	<.0001
V026 Small city	1	-0.5852	0.1382	17.9387	<.0001
V026 Town	1	-0.6616	0.1384	22.8480	<.0001
V103 Abroad	1	-0.9034	0.2839	10.1283	0.0015
V103 City	1	-0.0180	0.1400	0.0165	0.8977
V103 Countryside	1	-1.7843	0.1442	153.1051	<.0001
V103 Town	1	-0.1982	0.1409	1.9794	0.1595

V131	Balant	1	0.7865	0.4204	3.5002	0.0614
V131	Bambara	1	0.2904	0.2635	1.2144	0.2705
V131	Diola	1	0.7428	0.1555	22.8065	<.0001
V131	Mancagne	1	1.7125	0.4844	12.5002	0.0004
V131	Mandingue/Soce/M	1	-0.2468	0.2113	1.3643	0.2428
V131	Manjaak	1	0.6768	0.3657	3.4254	0.0642
V131	Not Senegalese	1	0.4592	0.3034	2.2912	0.1301
V131	Other	1	0.9371	0.2243	17.4505	<.0001
V131	Poular	1	0.00630	0.1136	0.0031	0.9558
V131	Sarakole/Soninke	1	0.5482	0.2629	4.3495	0.0370
V131	Serer	1	0.8289	0.1181	49.2957	<.0001

Odds Ratio Estimates

Effect		Point Estimate	95% wald Confidence Limits	
V013 20-24 vs 15-19		1.102	0.881	1.379
V013 25-29 vs 15-19		0.886	0.696	1.130
V013 30-34 vs 15-19		0.960	0.741	1.244
V013 35-39 vs 15-19		0.670	0.504	0.891
V013 40-44 vs 15-19		0.591	0.420	0.830
V013 45-49 vs 15-19		0.465	0.300	0.721
V026 Countryside	vs Capital, large c	0.168	0.128	0.221
V026 Small city	vs Capital, large c	0.557	0.425	0.730
V026 Town	vs Capital, large c	0.516	0.393	0.677
V103 Abroad	vs Capital, large c	0.405	0.232	0.707
V103 City	vs Capital, large c	0.982	0.747	1.292
V103 Countryside	vs Capital, large c	0.168	0.127	0.223
V103 Town	vs Capital, large c	0.820	0.622	1.081
V131 Balant	vs Wolof/Lebou	2.196	0.963	5.005
V131 Bambara	vs Wolof/Lebou	1.337	0.798	2.241
V131 Diola	vs Wolof/Lebou	2.102	1.550	2.851
V131 Mancagne	vs Wolof/Lebou	5.543	2.145	14.323
V131 Mandingue/Soce/M	vs Wolof/Lebou	0.781	0.516	1.182
V131 Manjaak	vs Wolof/Lebou	1.968	0.961	4.029
V131 Not Senegalese	vs Wolof/Lebou	1.583	0.873	2.869
V131 Other	vs Wolof/Lebou	2.553	1.644	3.962
V131 Poular	vs Wolof/Lebou	1.006	0.805	1.257
V131 Sarakole/Soninke	vs Wolof/Lebou	1.730	1.034	2.896
V131 Serer	vs Wolof/Lebou	2.291	1.818	2.887

Association of Predicted Probabilities and Observed Responses

Percent Concordant	84.4
Percent Discordant	14.8
Percent Tied	0.8

Consideriamo infine il modello relativo al raggiungimento del grado d'istruzione cosiddetto "alto".

Dalle stime dei parametri e degli odds ratio relativi a *v013*, capiamo che probabilmente c'è qualche problema in relazione alla classe di riferimento di tale variabile.

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	wald Chi-Square	Pr > ChiSq
V013 20-24	1	5.3029	20.6763	0.0658	0.7976
V013 25-29	1	2.8335	20.6928	0.0188	0.8911
V013 30-34	1	3.6007	20.6749	0.0303	0.8617
V013 35-39	1	5.0200	20.6701	0.0590	0.8081
V013 40-44	1	4.9064	20.6721	0.0563	0.8124
V013 45-49	1	5.9358	20.6705	0.0825	0.7740

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits
V013 20-24 vs 15-19	200.913	<0.001 >999.999
V013 25-29 vs 15-19	17.005	<0.001 >999.999
V013 30-34 vs 15-19	36.622	<0.001 >999.999
V013 35-39 vs 15-19	151.414	<0.001 >999.999
V013 40-44 vs 15-19	135.155	<0.001 >999.999
V013 45-49 vs 15-19	378.350	<0.001 >999.999

Le donne con meno di 20 anni vengono escluse dall'analisi per evidente scarsità di donne in età '15-19' che abbiano raggiunto un grado d'istruzione "alto"; si impone di conseguenza anche uno slittamento della classe di riferimento di v013 a '20-24'.

Model Information

Data Set	DATI97.DATI19
Response Variable	v106_3 Iniziato grado di istruzione alto
Number of Response Levels	2
Number of Observations	6607
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	v106_3	Total Frequency
1	Sì	70
2	No	6537

Probability modeled is v106\_3='Sì'.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	203.9725	7	<.0001

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
39.8292	39	0.4330

Summary of Stepwise Selection

Step	Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	V026		3	1	281.7526	.	<.0001
2	V103		4	2	20.9918	.	0.0003

Step	Variable Label
1	De facto place of residence
2	Childhood place of residence

Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
V026	3	48.8447	<.0001
V103	4	17.1359	0.0018

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.5102	0.1694	219.5164	<.0001
V026 Countryside	1	-3.2749	0.6563	24.9034	<.0001
V026 Small city	1	-2.6260	0.5198	25.5181	<.0001
V026 Town	1	-1.3776	0.4236	10.5757	0.0011
V103 Abroad	1	0.4759	0.5098	0.8713	0.3506
V103 City	1	0.3777	0.3514	1.1553	0.2824
V103 Countryside	1	-2.2712	0.6636	11.7156	0.0006
V103 Town	1	-0.4190	0.4357	0.9249	0.3362

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits
V026 Countryside vs Capital, large c	0.038	0.010 0.137
V026 Small city vs Capital, large c	0.072	0.026 0.200
V026 Town vs Capital, large c	0.252	0.110 0.578
V103 Abroad vs Capital, large c	1.609	0.593 4.372
V103 City vs Capital, large c	1.459	0.733 2.905
V103 Countryside vs Capital, large c	0.103	0.028 0.379
V103 Town vs Capital, large c	0.658	0.280 1.545

Association of Predicted Probabilities and Observed Responses

Percent Concordant	87.7
Percent Discordant	5.9
Percent Tied	6.5

**COMMENTI**

- ✎ I risultati dei primi due modelli considerati in questo paragrafo sono abbastanza conformi a quelli ottenuti nel caso dell'alfabetizzazione; le considerazioni a riguardo si muovono quindi nella stessa direzione.
- ✎ L'unica differenza significativa riguarda il modello sul raggiungimento di un grado d'istruzione "alto"; in tal caso, infatti, l'età attuale della donna non risulta significativa e le variabili relative all'ambiente di residenza, sia d'infanzia che attuale, sembrano sufficienti a spiegare la dipendenza analizzata dal modello. L'interpretazione di tale risultato potrebbe risiedere nel fatto che un grado d'istruzione "alto" rimanga prerogativa delle grandi città ma che, anche in tali ambienti, per ora non vi siano grandi progressi nella sua diffusione.

## 8.3 L'URBANIZZAZIONE<sup>5</sup>

### LA SCELTA DELLE VARIABILI ESPLICATIVE E LA CREAZIONE DEL DATASET

L'ultimo modello preso in considerazione è quello relativo all'urbanizzazione del territorio.

Sono state selezionate come candidate esplicative le seguenti variabili:

v013	Età attuale - Classi quinquennali
v024	Regione di residenza
v103	Ambiente di residenza d'infanzia: tipo di centro abitato
v106	Maggior gradi di istruzione iniziato
v108cl	Alfabetizzazione - Nuova classificazione
v131	Etnia
s503	Evento unione (Sperimentato=1)
v502cl	Stato attuale di unione - Dicotomica
poligam	Poligamia
figlio1	Evento 1° figlio (Sperimentato=1)

### L'ELABORAZIONE DEL MODELLO

Con questo modello vorremmo cercare di spiegare e capire le trasformazioni in corso a livello di urbanizzazione; l'ambiente di residenza d'infanzia potrebbe però già raccogliere gran parte di questi cambiamenti, motivo per cui *v103* viene esclusa dal modello.

Allo stesso tempo vengono escluse *v106* e *v108cl*, visto che con questi fattori potrebbe facilmente sussistere una dipendenza inversa a quella considerata.

		<u>Model Information</u>	
Data Set		DATI97.DATI20	
Response Variable		v025urb	Ambiente di residenza urbano
Number of Response Levels		2	
Number of Observations		8587	
Model		binary logit	
Optimization Technique		Fisher's scoring	

<u>Response Profile</u>		
Ordered Value	v025urb	Total Frequency
1	Sì	3053
2	NO	5534

Probability modeled is v025urb='Sì'.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.

<u>Testing Global Null Hypothesis: BETA=0</u>			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2130.8733	23	<.0001

<sup>5</sup> Vedi *Mod20.sas* [Appendice C].

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
25.4480	19	0.1463

Summary of Stepwise Selection

Step	Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	V024		3	1	1392.7346	.	<.0001
2	v502c1		1	2	482.6698	.	<.0001
3	V131		11	3	166.0348	.	<.0001
4	V013		6	4	84.9166	.	<.0001
5	poligam		1	5	21.1747	.	<.0001
6	S503		1	6	20.1133	.	<.0001
7	figlio1		1	7	5.7169	.	0.0168
8		figlio1	1	6	.	5.7045	0.0169

Step	Variable Label
1	Region
2	Stato di unione attuale - Dicotomica
3	Ethnicity
4	Age 5-year groups
5	Poligamia
6	Evento unione
7	Evento 1° figlio
8	Evento 1° figlio

Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
V013	6	113.4844	<.0001
V024	3	964.8716	<.0001
V131	11	155.7700	<.0001
S503	1	19.9472	<.0001
v502c1	1	64.2001	<.0001
poligam	1	24.2528	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.5444	0.0797	375.4191	<.0001
V013 20-24	1	0.5381	0.0879	37.4342	<.0001
V013 25-29	1	0.7082	0.1017	48.4641	<.0001
V013 30-34	1	0.8994	0.1079	69.4573	<.0001
V013 35-39	1	1.0666	0.1107	92.7766	<.0001
V013 40-44	1	1.0049	0.1218	68.0497	<.0001
V013 45-49	1	1.0017	0.1357	54.4698	<.0001
V024 Central	1	-2.0755	0.0685	919.2192	<.0001
V024 North East	1	-1.7082	0.0864	390.8374	<.0001
V024 South	1	-1.4880	0.1015	215.0630	<.0001
V131 Balant	1	-0.4209	0.2720	2.3942	0.1218
V131 Bambara	1	0.5349	0.1892	7.9911	0.0047
V131 Diola	1	-0.1353	0.1319	1.0529	0.3048
V131 Mancagne	1	1.7234	0.5244	10.8010	0.0010
V131 Mandingue/Soce/M	1	-0.1727	0.1227	1.9813	0.1592
V131 Manjaak	1	0.4942	0.2926	2.8518	0.0913
V131 Not Senegalese	1	0.1019	0.1684	0.3664	0.5450
V131 Other	1	0.7108	0.1873	14.4057	0.0001
V131 Poular	1	-0.3402	0.0754	20.3379	<.0001
V131 Sarakole/Soninke	1	-0.1118	0.1860	0.3613	0.5478
V131 Serer	1	-0.7609	0.0828	84.3890	<.0001
S503 In unione (ora o	1	-0.6005	0.1345	19.9472	<.0001
v502c1 Attualmente in u	1	-0.9480	0.1183	64.2001	<.0001
poligam Sì	1	-0.3275	0.0665	24.2528	<.0001



Odds Ratio Estimates

Effect		Point Estimate	95% wald Confidence Limits	
V013	20-24 vs 15-19	1.713	1.442	2.035
V013	25-29 vs 15-19	2.030	1.663	2.478
V013	30-34 vs 15-19	2.458	1.989	3.037
V013	35-39 vs 15-19	2.905	2.339	3.610
V013	40-44 vs 15-19	2.732	2.151	3.468
V013	45-49 vs 15-19	2.723	2.087	3.553
V024	Central vs West	0.125	0.110	0.144
V024	North East vs West	0.181	0.153	0.215
V024	South vs West	0.226	0.185	0.276
V131	Balant vs wolof/Lebou	0.656	0.385	1.119
V131	Bambara vs wolof/Lebou	1.707	1.178	2.474
V131	Diola vs wolof/Lebou	0.873	0.674	1.131
V131	Mancagne vs wolof/Lebou	5.603	2.005	15.660
V131	Mandingue/Soce/M vs wolof/Lebou	0.841	0.662	1.070
V131	Manjaak vs wolof/Lebou	1.639	0.924	2.909
V131	Not Senegalese vs wolof/Lebou	1.107	0.796	1.540
V131	Other vs wolof/Lebou	2.036	1.410	2.938
V131	Poular vs wolof/Lebou	0.712	0.614	0.825
V131	Sarakole/Soninke vs wolof/Lebou	0.894	0.621	1.288
V131	Serer vs wolof/Lebou	0.467	0.397	0.550
S503	In unione (ora o vs Mai in unione	0.549	0.421	0.714
v502c1	Attualmente in u vs Attualmente non	0.388	0.307	0.489
poligam	Sì vs No	0.721	0.633	0.821

Association of Predicted Probabilities and Observed Responses

Percent Concordant	77.8
Percent Discordant	21.4
Percent Tied	0.8

**COMMENTI**

- ☒ Tutte le regioni apportano, come ci aspettavamo, un contributo negativo al rischio di base associato alla regione costiera dell'ovest, notoriamente la più moderna ed urbanizzata.
- ☒ I risultati relativi all'etnia sono ragionevolmente conformi alle origini, all'identità e alle peculiarità delle singole popolazioni.
- ☒ Per quanto riguarda invece le altre variabili, i risultati di questo modello sono anomali e difficilmente interpretabili; la causa di ciò potrebbe essere imputabile al fatto che alcune dipendenze siano inverse a quelle considerate e quindi vadano a falsare gli effetti attribuiti agli altri fattori.

