

Modelli statistici e strumenti grafici per opzioni politiche e variabili di contesto

Nathalie Sitzia

Dipartimento di Scienze Statistiche, Università di Napoli Federico II
E-mail: nathaliesitzia@iol.it

Summary: In this paper we discuss some statistical models for political preferences and we propose a graphical representation for the analysis of structural and ideological discrepancy or homogeneity among several political beliefs. At this aim, the opinion poll of a sample of university students is studied to get an empirical check of the appropriateness of the proposal. The results show the coherent interaction between graphic representation and modelling analysis in social and political researches.

Keywords: Generalized Linear Models, Ranks Jittering, Political Preferences

1. Introduzione

Il potenziale rapporto esistente tra scelte politiche e caratteristiche individuali è stato da sempre considerato oggetto di studio per il notevole interesse che suscita nel dibattito politico-sociale. In tale ambito, si registrano una serie di luoghi comuni che insistono su una corrispondenza, piuttosto rigida, fra il gruppo sociale di appartenenza e quello politico di riferimento. Negli ultimi anni, la rigidità di tale schematizzazione è venuta meno, come dimostrano gli eventi elettorali in Italia e nel resto del mondo. Emerge dunque, ancor più forte, l'esigenza di comprendere se e come oggi le opinioni politiche di un individuo si evolvano e siano connesse con il suo background culturale, sociale e familiare.

La letteratura statistica ha in genere privilegiato le tematiche connesse ai cosiddetti “mutamenti” elettorali (Brown and Payne, 1986) e alla scelta fra candidati alternativi (Keener and Waldman, 1985), anche con attenzione a questioni specifiche, come per esempio il problema della collocazione dei candidati sulla scheda elettorale (Upton and Brook, 1974).

In tale contesto, il presente lavoro è motivato dall’esigenza di mostrare come gli strumenti dell’analisi statistica possano essere efficacemente proposti ed utilizzati per la verifica delle interrelazioni tra opzione politica, variabili individuali e di contesto. A tal fine, si applicano differenti modelli probabilistici e si propone una rappresentazione grafica delle preferenze in modo da giungere ad una valutazione comparata delle informazioni ottenute da una specifica indagine svolta su giovani universitari.

Nel paragrafo 2 sono discusse la pianificazione, i contenuti e le modalità dell’indagine. Quindi, sono sviluppati alcuni modelli probabilistici e vengono esplorati degli strumenti grafici nell’ambito della classe dei modelli lineari generalizzati. Più precisamente, nel paragrafo 3 si esamina un modello di regressione logistica rivolto a comprendere quali caratteristiche individuali e di contesto spieghino la scelta politica dell’individuo. Nel paragrafo 4, si valuta –mediante un modello log-lineare- il grado di impegno sociale degli studenti in base alla scelta politica espressa. Infine, mediante una adeguata rappresentazione grafica differenziata rispetto alle scelte ideologiche, si è cercato di collegare l’analisi delle priorità avvertite verso tematiche notevoli ad una ipotetica base ideologica dell’individuo. Alcune considerazioni finali concludono il lavoro.

2. Obiettivi e pianificazione dell’indagine

Poche settimane prima delle elezioni politico-amministrative 2001, è stato somministrato agli studenti delle Facoltà di Scienze Politiche e di Sociologia dell’Università degli Studi di Napoli Federico II, un questionario diretto a verificare le connessioni tra una serie di tematiche attuali e le tradizionali opzioni politiche (Sinistra, Centro,

Destra) mediante una serie di domande ove era necessario esprimere graduatorie di preferenza.

Il questionario si articola in due gruppi di quesiti: nel primo, si cerca di pervenire alla misura di variabili individuali (sesso, età, ecc.) e di contesto (ambiente familiare, socio-culturale, ecc.); nel secondo gruppo, si cerca di individuare la base ideologica dello studente attraverso la quale egli esprime la sua aderenza ad una determinata posizione politica. Quest'ultimo aspetto si ricollega implicitamente attraverso le priorità che egli fornisce alle problematiche legate al mondo del lavoro (disoccupazione, flessibilità e mobilità), alla sicurezza e alla giustizia, ecc. ed esplicitamente grazie alla graduatoria che il rispondente formula rispetto alle grandi opzioni politiche correnti.

La somministrazione del questionario è avvenuta nella stessa giornata nelle due Facoltà per evitare distorsioni dovute ad esternalità occasionali (ovvero problematiche di attualità o di cronaca proposte da mass-media in giorni diversi). Inoltre, le modalità di somministrazione e raccolta hanno rigorosamente garantito l'anonimato e la *privacy*.

Il tasso di risposta ha superato ampiamente il 90% per entrambe le Facoltà. L'indagine si basa su 241 questionari completi, 86 provenienti da studenti di Sociologia e 155 da Scienze Politiche. In particolare, il collettivo esaminato presenta le seguenti caratteristiche salienti schematizzate nella Tabella 1.

Tabella 1. Caratteristiche del collettivo esaminato

Facoltà	Sesso	Numero studenti	Età media	Età mediana
Sociologia	Maschi	20	25.43	24.25
	Femmine	66	23.13	23.08
Scienze Politiche	Maschi	63	21.21	19.58
	Femmine	92	20.13	19.42
Totale		241	21.67	20.25

Si osservi che, ovviamente, i rispondenti non rappresentano un campione casuale degli studenti delle due Facoltà ma sono rappresentativi di quella popolazione di studenti maggiormente

sensibili alle problematiche sociali e politiche (date le Facoltà cui sono iscritti) e che partecipa con una certa regolarità ai corsi.

Nei paragrafi seguenti, vengono anzitutto discussi i risultati ottenuti mediante l'uso dei modelli logit e log-lineare, rispettivamente, mentre la successiva rappresentazione grafica viene proposta quale ulteriore conferma delle precedenti risultanze, in modo da collegare le priorità espresse verso alcune tematiche correnti con le personali opzioni politiche.

3. Il modello di regressione logistico

La prima analisi modellistica effettuata ha riguardato la probabilità di essere di Sinistra $\Pr(S=1)$, di Centro $\Pr(C=1)$, di Destra $\Pr(D=1)$ e, a tal fine, il modello prescelto per studiare la dipendenza delle variabili risposta dalle variabili esplicative individuali è stato il modello lineare generalizzato (*GLM*) di tipo logistico.

La componente casuale di tali modelli è costituita, rispettivamente, da una variabile casuale (v.c.) Binomiale nella quale il parametro π rappresenta, di volta in volta, la probabilità di aderire all'opzione ideologica considerata.

Considerando, genericamente, la v.c. risposta $\Pr(Y=1)$, le osservazioni sulle p variabili esplicative sono contenute nella matrice \mathbf{X} le cui dimensioni sono $(n \times (p+1))$, per la presenza del vettore unitario nella prima colonna. Secondo la notazione tradizionale, x_{ij} indica il valore assunto dalla variabile X_j , $j=1,2,\dots,p$ nell'unità statistica i -esima, $i=1,2,\dots,n$. I $(p+1)$ parametri sono contenuti nel vettore $\mathbf{b} = (\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_p)'$, mentre \mathbf{x} rappresenta una generica riga della matrice \mathbf{X} .

La *funzione legame* che lega la variabile risposta al predittore lineare \mathbf{xb} è la funzione logit (McCullagh and Nelder, 1989):

$$\log \left(\frac{p}{1-p} \right) = \mathbf{b}_0 + \mathbf{b}_1 X_1 + \dots + \mathbf{b}_p X_p$$

per la quale si ha:

$$\mathbf{p} = \Pr(Y = 1) = \frac{e^{x\mathbf{b}}}{1 + e^{x\mathbf{b}}} = \frac{1}{1 + e^{-x\mathbf{b}}};$$

espressioni simili verranno utilizzate per $\Pr(S=1)$, $\Pr(C=1)$ e $\Pr(D=1)$.

Il coefficiente \mathbf{b}_0 esprime il livello base *degli odds* del verificarsi dell'evento a prescindere dai valori e dalle modalità delle variabili esplicative. I coefficienti $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p$, invece, sono i logaritmi degli *odds ratio* associati alle relative variabili esplicative X_1, X_2, \dots, X_p .

Di ogni modello sono stati stimati i coefficienti \mathbf{b} massimizzando la log-verosimiglianza :

$$\log L(\mathbf{b}, y) \propto \sum_{i=1}^n \sum_{j=1}^p y_i x_{ij} \mathbf{b}_j - \sum_{i=1}^n N_i \log \left(1 + \exp \left(\sum_{j=1}^p x_{ij} \mathbf{b}_j \right) \right)$$

dove N_i rappresenta il numero di soggetti accomunati da uguali valori delle variabili esplicative.

Inoltre, per verificare la bontà di adattamento dei modelli è stata calcolata la devianza:

$$dev = 2 \sum_i \sum_j \left(f_{ij} \times \log \left(\frac{f_{ij}}{\hat{f}_{ij}} \right) \right),$$

dove con f_{ij} e con \hat{f}_{ij} sono indicate, rispettivamente, le frequenze osservate nel campione e quelle teoriche ottenute sotto l'ipotesi di validità del modello stimato. Quindi, dal modello *massimale* sono stati rimossi quei termini che risultavano meno significativi e si sono confrontate le devianze, giungendo poi ad un modello *minimale* adeguato (Dobson, 1990).

Specificamente, le variabili individuali e di contesto prescelte per l'analisi sono: la Facoltà a cui lo studente è iscritto (FAC, dove Sociologia = 0 e Scienze Politiche = 1); il Sesso (SEX); l'Età (ETA); il numero di componenti del nucleo familiare (NFA); il Diploma secondario di provenienza (DIP); la professione del capo famiglia (PCF); l'aver o meno una madre lavoratrice (MAL); l'essere o meno iscritto ad associazioni politico-sindacali (POS), religiose (REL), di volontariato (VOL), culturali (CUL) o sportive (SPO); l'aver o meno

almeno un parente iscritto ad associazioni politiche (PAP); l'aver o meno votato in precedenza (VOT); le preferenze sportive rispetto a sport di squadra o individuali (PSP); il possesso o meno di un lettore DVD (DVD); il navigare o meno quotidianamente in Internet (INT).

Per procedere all'analisi relativa alla probabilità di essere di Sinistra, di Centro e di Destra sono stati stimati sequenzialmente (mediante il *software* S-Plus) i modelli riportati nelle Tabelle 2-4.

Tabella 2. Modelli logit per studenti di Sinistra

Modelli	Devianza	Gradi di libertà
$\Pr(S = 1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_1FAC + \mathbf{b}_2SEX + \mathbf{b}_3ETA + \mathbf{b}_4NFA + \mathbf{b}_5DIP + \mathbf{b}_6PCF + \mathbf{b}_7MAL + \mathbf{b}_8REL + \mathbf{b}_9VOL + \mathbf{b}_{10}POS + \mathbf{b}_{11}CUL + \mathbf{b}_{12}SPO + \mathbf{b}_{13}PAP + \mathbf{b}_{14}VOT + \mathbf{b}_{15}PSP + \mathbf{b}_{16}DVD + \mathbf{b}_{17}INT)\}^{-1}$	291.45	222
$\Pr(S = 1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_1FAC + \mathbf{b}_4NFA + \mathbf{b}_7MAL + \mathbf{b}_{11}CUL + \mathbf{b}_{14}VOT + \mathbf{b}_{16}DVD + \mathbf{b}_{17}INT)\}^{-1}$	294.44	232
$\Pr(S = 1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_1FAC + \mathbf{b}_7MAL + \mathbf{b}_{14}VOT + \mathbf{b}_{17}INT)\}^{-1}$	301.66	236

Tabella 3. Modelli logit per studenti di Centro

Modelli	Devianza	Gradi di libertà
$\Pr(C = 1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_1FAC + \mathbf{b}_2SEX + \mathbf{b}_3ETA + \mathbf{b}_4NFA + \mathbf{b}_5DIP + \mathbf{b}_6PCF + \mathbf{b}_7MAL + \mathbf{b}_8REL + \mathbf{b}_9VOL + \mathbf{b}_{10}POS + \mathbf{b}_{11}CUL + \mathbf{b}_{12}SPO + \mathbf{b}_{13}PAP + \mathbf{b}_{14}VOT + \mathbf{b}_{15}PSP + \mathbf{b}_{16}DVD + \mathbf{b}_{17}INT)\}^{-1}$	209.21	222
$\Pr(C = 1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_7MAL + \mathbf{b}_8REL + \mathbf{b}_9VOL + \mathbf{b}_{10}POS + \mathbf{b}_{11}CUL + \mathbf{b}_{14}VOT + \mathbf{b}_{16}DVD)\}^{-1}$	214.71	232
$\Pr(C = 1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_7MAL + \mathbf{b}_9VOL + \mathbf{b}_{16}DVD)\}^{-1}$	224.19	236

Tabella 4. Modelli logit per studenti di Destra

Modelli	Devianza	Gradi di libertà
$\Pr(D=1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_1FAC + \mathbf{b}_2SEX + \mathbf{b}_3ETA + \mathbf{b}_4NFA + \mathbf{b}_5DIP + \mathbf{b}_6PCF + \mathbf{b}_7MAL + \mathbf{b}_8REL + \mathbf{b}_9VOL + \mathbf{b}_{10}POS + \mathbf{b}_{11}CUL + \mathbf{b}_{12}SPO + \mathbf{b}_{13}PAP + \mathbf{b}_{14}VOT + \mathbf{b}_{15}PSP + \mathbf{b}_{16}DVD + \mathbf{b}_{17}INT)\}^{-1}$	187.949	222
$\Pr(D=1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_1FAC + \mathbf{b}_4NFA + \mathbf{b}_7MAL + \mathbf{b}_9VOL + \mathbf{b}_{10}POS + \mathbf{b}_{13}PAP + \mathbf{b}_{14}VOT + \mathbf{b}_{17}INT)\}^{-1}$	190.997	232
$\Pr(D=1) = \{1 + \exp(\mathbf{b}_0 + \mathbf{b}_1FAC + \mathbf{b}_4NFA + \mathbf{b}_{13}PAP+)\}^{-1}$	200.424	237

Al fine di individuare il modello “ottimale”, ci si è basati sul test asintotico che confronta le devianze tra due modelli ($LR = dev(M_1) - dev(M_2)$) con il valore critico della v.c. $\mathbf{C}_{(g)}^2$, essendo i gradi di libertà g espressi dalla differenza tra il numero dei parametri dei due modelli.

Alla fine di un’ articolata selezione si è optato per l’adozione dei seguenti modelli logit le cui stime sono riportate nelle Tabelle 5-7.

Tabella 5. Modello logit per studenti di Sinistra

Variabili	$\hat{\mathbf{b}}_i$	$t_{\mathbf{b}_i}$
Costante	-0.698	-1.206
Facoltà	-0.735	-2.386
Madre lavoratrice	0.472	1.645
Votato	1.199	2.276
Internet	0.691	1.654

Tabella 6. Modello logit per studenti di Centro

Variabili	$\hat{\mathbf{b}}_i$	$t_{\mathbf{b}_i}$
Costante	-1.285	-5.737
Madre lavoratrice	-1.250	-3.252
Associazioni volontariato	0.961	2.433
Letto DVD	0.988	2.093

Tabella 7. Modello logit per studenti di Destra

Variabili	\hat{b}_i	t_{b_i}
Costante	0.238	0.247
Facoltà	1.583	3.390
Componenti nucleo familiare	-0.622	-2.729
Parente politicizzato	-1.218	-2.522

Da tali risultati si possono derivare alcune interessanti considerazioni che, ora brevemente, riassumiamo. In primo luogo, emerge che se si passa dalla Facoltà di Sociologia a quella di Scienze Politiche, *ceteris paribus*, gli *odds* di avere uno studente di Sinistra diminuiscono del 52% ($\exp(-0.73) = 0.48$). Al contrario, gli *odds* aumentano del 60% ($\exp(0.47) = 1.60$) nel caso in cui la madre lavora. Invece, se lo studente ha già votato, la probabilità di essere di Sinistra cresce più del doppio ($\exp(1.20) = 3.32$, cioè del 232%) e, nel caso in cui navighi in Internet quotidianamente, tale probabilità cresce quasi del doppio ($\exp(0.69) = 1.99$).

Gli studenti che hanno una madre lavoratrice hanno una minor probabilità di essere di Centro ($\exp(-1.25) = 0,28$ quindi del 72%) rispetto a coloro che hanno una madre casalinga. Al contrario, tale probabilità è superiore del 161% e del 166% ($\exp(0.96) = 2.61$; $\exp(0.99) = 2.66$), rispettivamente, nei casi di iscritti ad associazioni di volontariato, rispetto a coloro che non lo sono, e di quando la famiglia possieda un lettore DVD, rispetto a chi non possiede tali strumenti, tuttora relativamente costosi.

Analizzando, infine, i coefficienti relativi alla probabilità di essere di Destra emerge che gli *odds* aumentano di $\exp(1.58) = 4.85$, cioè del 385%, nel passare da Sociologia a Scienze Politiche, mentre diminuiscono di $\exp(-0.62) = 0.53$, cioè del 47% per ogni incremento unitario del numero di componenti del nucleo familiare. Le probabilità di essere di Destra per chi ha almeno un componente familiare iscritto ad un'associazione politica è inferiore di $\exp(-1.22) = 0.29$, cioè del 71%, rispetto a coloro che non ne hanno.

Sulla base di tali risultanze, il profilo che emerge relativamente al particolare collettivo esaminato, è che:

i) gli studenti iscritti alla Facoltà di Sociologia che hanno una madre lavoratrice, navigano in Internet e hanno già usufruito del diritto-dovere al voto (e che hanno di riflesso un'età relativamente più elevata rispetto ai loro colleghi), sono di Sinistra con più elevata probabilità;

ii) gli studenti che hanno una madre casalinga, sono iscritti ad associazioni di volontariato e posseggono strumenti tecnologici come i lettori DVD, sono di Centro con più elevata probabilità;

iii) gli studenti iscritti alla Facoltà di Scienze Politiche, che hanno famiglie poco numerose e parenti non iscritti ad associazioni politiche sono di Destra con probabilità più elevata.

4. Il modello log-lineare

Per valutare il grado di impegno sociale degli studenti esaminati si è utilizzato un modello per dati di conteggio, noto in letteratura come *modello log-lineare* (Bishop, Fienberg and Holland, 1995). Infatti, in tale caso, la variabile risposta è il numero (=conteggio) delle associazioni cui lo studente è iscritto mentre le variabili esplicative considerate sono quelle del precedente paragrafo, cui si aggiunge la caratteristica essere di Sinistra (S), di Centro (C) oppure di Destra (D).

Quindi, tenuto conto delle modalità di somministrazione e di raccolta del questionario, è lecito ipotizzare che la variabile risposta (numero di associazioni alle quali lo studente è iscritto) si distribuisca come una v.c. di Poisson $Y \sim Po(\mathbf{I})$. D'altra parte, poiché tale variabile deve assumere solo valori non negativi (qualunque siano i valori delle variabili esplicative), si sceglie di costruire un modello per il logaritmo del valore medio della v.c. Poisson:

$$\log(\mathbf{I}_i) = \sum_{j=1}^p x_{ij} \mathbf{b}_j .$$

La funzione di log-verosimiglianza è:

$$\log L(\mathbf{b}, \mathbf{y}) \propto \sum_{i=1}^n \left[\sum_{j=1}^p y_i x_{ij} \mathbf{b}_j - \exp \left(\sum_{j=1}^p x_{ij} \mathbf{b}_j \right) \right].$$

Come per l'analisi precedente, attraverso la stima dei parametri e il confronto fra le devianze, si passa dal modello massimale ad un modello minimale adeguato. La successiva Tabella 8, che riassume alcuni confronti fra tali modelli, evidenzia altresì come alla fine sia opportuno limitarsi a considerare come variabili esplicative l'Età, l'uso regolare di Internet e l'adesione ideologica al Centro.

Quindi, la successiva Tabella 9 mostra le stime dei coefficienti per un modello per il logaritmo del valore medio:

$$\log(I_i) = \mathbf{b}_0 + \mathbf{b}_3 \text{ETA} + \mathbf{b}_{13} \text{INT} + \mathbf{b}_{14} \text{C} .$$

Tabella 8. Confronto fra modelli log-lineari

Modelli	Devianza	Gradi di libertà
$\log(I_i) = \mathbf{b}_0 + \mathbf{b}_1 \text{FAC} + \mathbf{b}_2 \text{SEX} + \mathbf{b}_3 \text{ETA} + \mathbf{b}_4 \text{NFA} + \mathbf{b}_5 \text{DIP} + \mathbf{b}_6 \text{PCF} + \mathbf{b}_7 \text{MAL} + \mathbf{b}_8 \text{PAP} + \mathbf{b}_9 \text{VOT} + \mathbf{b}_{10} \text{PSP} + \mathbf{b}_{11} \text{DVD} + \mathbf{b}_{12} \text{INT} + \mathbf{b}_{13} \text{S} + \mathbf{b}_{14} \text{C} + \mathbf{b}_{15} \text{D}$	305.523	223
$\log(I_i) = \mathbf{b}_0 + \mathbf{b}_1 \text{FAC} + \mathbf{b}_3 \text{ETA} + \mathbf{b}_6 \text{PCF} + \mathbf{b}_{10} \text{PSP} + \mathbf{b}_{12} \text{INT} + \mathbf{b}_{14} \text{C}$	309.528	233
$\log(I_i) = \mathbf{b}_0 + \mathbf{b}_1 \text{FAC} + \mathbf{b}_3 \text{ETA} + \mathbf{b}_{12} \text{INT} + \mathbf{b}_{14} \text{C}$	315.638	237

Tabella 9. Stima dei coefficienti del modello log-lineare

Variabili	$\hat{\mathbf{b}}_i$	$t_{\mathbf{b}_i}$
Costante	-0.660	-2.102
Età	0.030	2.308
Internet	0.391	2.572
Centro	0.272	1.957

In sostanza, si può spiegare il numero medio di associazioni cui lo studente è iscritto tramite la seguente funzione:

$$\hat{I} = \exp\{-0.66 + 0.030 \times \text{Età} + 0.39 \times \text{Internet} + 0.27 \times \text{Centro}\} .$$

Ora, poiché le variabili *Internet* e *Centro* sono entrambe dicotomiche (SI=1; NO=0) è possibile rappresentare tale funzione -al variare della variabile continua *Età*- per ciascuno dei quattro possibili profili che si ottengono combinando le modalità delle due variabili dicotomiche. La

Figura 1 e la Tabella 10 evidenziano, rispettivamente, i profili che si ottengono (in termini di valore medio atteso) ed il grafico delle corrispondenti funzioni al variare dell'età.

Tabella 10. Numero di associazioni atteso

Internet (SI=1; NO=0)	Centro (NO=0)	Modello stimato	Età	Valore medio previsto del numero di associazioni
0	0	$0.517*(1.03)^{Età}$	18	0.886
			20	0.942
			22	1.000
			24	1.061
			26	1.127
1	0	$0.764*(1.03)^{Età}$	18	1.309
			20	1.390
			22	1.476
			24	1.568
			26	1.665
0	1	$0.677*(1.03)^{Età}$	18	1.161
			20	1.233
			22	1.309
			24	1.391
			26	1.476
1	1	$1.000*(1.03)^{Età}$	18	1.716
			20	1.822
			22	1.934
			24	2.054
			26	2.186

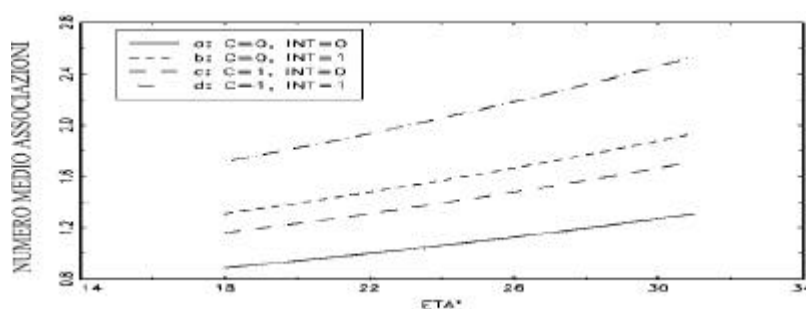


Figura 1. Numero medio di associazioni in funzione dell'età

Le evidenze che emergono dalla Tabella precedente e, soprattutto, dalla sottostante Figura 1, possono essere così schematizzate:

i) a parità di età, il numero di associazioni cui lo studente partecipa aumenta se è di Centro;

ii) tale partecipazione, *ceteris paribus*, aumenta ancor di più se lo studente naviga quotidianamente su Internet.

In breve, da tali analisi sembra che gli studenti universitari più impegnati socialmente, culturalmente e politicamente siano anche gli studenti più adulti, spesso di posizione politica di Centro e, soprattutto, che utilizzano regolarmente le moderne tecnologie informatiche per la comunicazione.

5. Analisi delle priorità in rapporto alla scelta politica

L'ultima analisi che esamineremo riguarda il confronto delle priorità espresse dagli studenti rispetto alle seguenti tematiche: *salute, sicurezza, lavoro, ambiente, giustizia, immigrazione* assegnando ad essi i ranghi 1,2,...,6 in ragione dell'ordine di importanza attribuita. Grazie ai dati a disposizione, si è dedotta una graduatoria di attenzione rispetto alle suddette problematiche, attraverso il confronto tra la media dei ranghi.

Per analizzare le differenze oppure le omogeneità comportamentali di diversi gruppi (studenti di Scienze Politiche e di Sociologia, di Sinistra, di Centro e di Destra, maschi e femmine) si è fatto ricorso anzitutto ad analisi tabellari di tipo standard e, poi, ad una proposta di grafico polare *jittered* che si è dimostrata capace di evidenziare con particolare efficacia i vari aspetti del fenomeno in studio.

Nello specifico, rispetto alle convenzioni abituali (Marden, 1995; Taplin, 1997), abbiamo preferito utilizzare i ranghi "complementari", cioè quelli definiti in modo che la preferenza massima risulti pari a 6. Tale modifica induce una migliore leggibilità delle scelte dello studente perché conferisce alle tabelle ed ai grafici seguenti, un messaggio della rilevanza della risposta fornita dallo studente direttamente connessa all'importanza delle scelte effettuate.

Tabella 11. Ranghi medi per Opinione politica

Tematiche	Sinistra	Centro	Destra
Salute	5,22	5,61	5,69
Lavoro	4,54	4,20	4,43
Giustizia	3,72	3,79	3,64
Sicurezza	3,05	3,12	3,24
Ambiente	2,89	2,77	2,71
Immigrazione	1,58	1,49	1,28

Tabella 12. Ranghi medi per Facoltà di appartenenza

Tematiche	Sociologia	Scienze Politiche
Salute	5.337	5.406
Lavoro	4.616	4.361
Giustizia	3.593	3.801
Sicurezza	3.093	3.096
Ambiente	2.845	
Immigrazione	1.490	

Tabella 13. Ranghi medi rispetto al Sesso

Tematiche	Maschi	Femmine
Salute	5.144	5.506
Lavoro	4.578	4.386
Giustizia	3.819	3.677
Sicurezza	3.060	3.113
Ambiente	3.000	2.746
Immigrazione	1.397	1.569

Tabella 14. Ranghi medi per Impegno politico di un familiare

Tematiche	Famiglia con impegno politico	Famiglia senza impegno politico
Salute	5.463	5.156
Lavoro	4.361	4.703
Giustizia	3.740	3.687
Sicurezza	3.084	3.125
Ambiente	2.887	2.687
Immigrazione	1.463	1.640

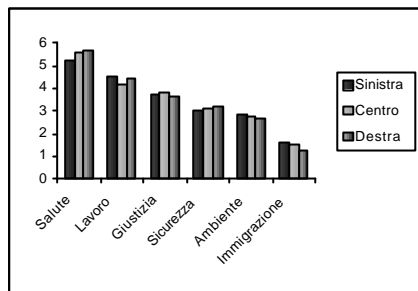


Figura 2. Graduatoria rispetto ai ranghi medi per Opinione politica

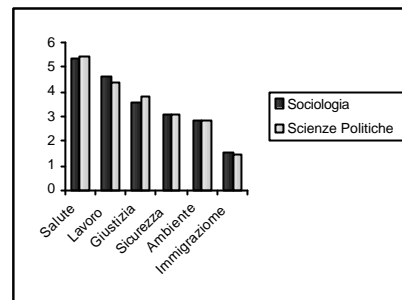


Figura 3. Graduatoria rispetto ai ranghi medi per Facoltà

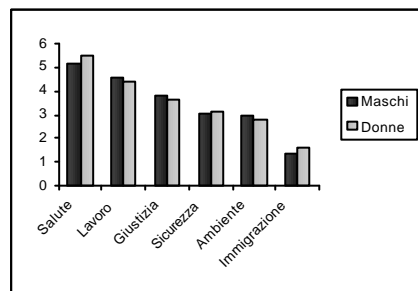


Figura 4. Graduatoria rispetto ai ranghi medi per Sesso

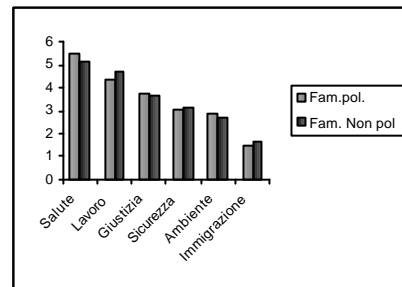


Figura 5. Graduatoria rispetto ai ranghi medi per Impegno politico di un familiare

Ebbene, contrariamente rispetto a quanto viene descritto da alcuni stereotipi tradizionali, le graduatorie espresse dai tre gruppi politici prima individuati sono pressoché identiche. Al primo posto si trova, sempre, la *salute* che precede nell'ordine: il *lavoro*, la *giustizia*, la *sicurezza*, l'*ambiente* e l'*immigrazione*. In effetti, sia la Tabella 11, nella quale sono riportati i ranghi medi, che la corrispondente Figura 2, mostrano che, all'interno di una gerarchia identica di priorità, le differenze espresse fra i tre gruppi sono oggettivamente assai modeste.

Inoltre, l'omogeneità tra le priorità tematiche non muta al cambiare delle variabili di riferimento. Infatti, non si riscontrano sostanziali differenze nelle graduatorie dei gruppi distinti rispetto alla Facoltà di appartenenza (Tabella 12 e Figura 3), al Sesso (Tabella 13 e Figura 4) e alla più o meno marcata Politicizzazione familiare (Tabella 14 e Figura 5).

Da tali analisi sembra che si possano formulare le seguenti considerazioni:

i) è confermato un certo “appiattimento” ideologico-culturale delle nuove generazioni, spesso denunciato da sociologi e politologi;

ii) le problematiche più diffuse dai mass-media sono state interiorizzate, evidenziando un certo contrasto tra quelle “personali” (la propria *salute*, la ricerca del proprio *lavoro*, la *giustizia* e la *sicurezza*) e quelle riguardanti la “collettività” (*l'ambiente* e *l'immigrazione*);

iii) la *giustizia* e la *sicurezza* sembrano viaggiare di pari passo, sino a far sorgere il fondato sospetto che gli studenti non abbiano distinto tra amministrazione e funzionamento della giustizia da un lato e prevenzione ed efficienza nella tutela delle persone e della proprietà dall'altro, come invece avviene nel dibattito politico. In buona sostanza, si sospetta che i rispondenti abbiano collegato le due problematiche al comune denominatore della criminalità diffusa.

Se questi dati fossero segnali di un più generale livellamento occorrerebbe riflettere seriamente, tenuto conto che l'indagine è stata effettuata all'interno di due Facoltà i cui studenti, in linea di principio, dovrebbero essere maggiormente attenti alle problematiche politiche e a quelle che da esse derivano.

Allo scopo di rendere ancora più evidenti tali considerazioni, abbiamo proposto per i ranghi una particolare rappresentazione grafica, ottenuta mediante una riformulazione originale del cosiddetto *jittering*. Infatti, questa tipologia di grafici consente di analizzare il *pattern* generale della graduatoria individuati anche, se presenti, eventuali casi anomali di comportamento.

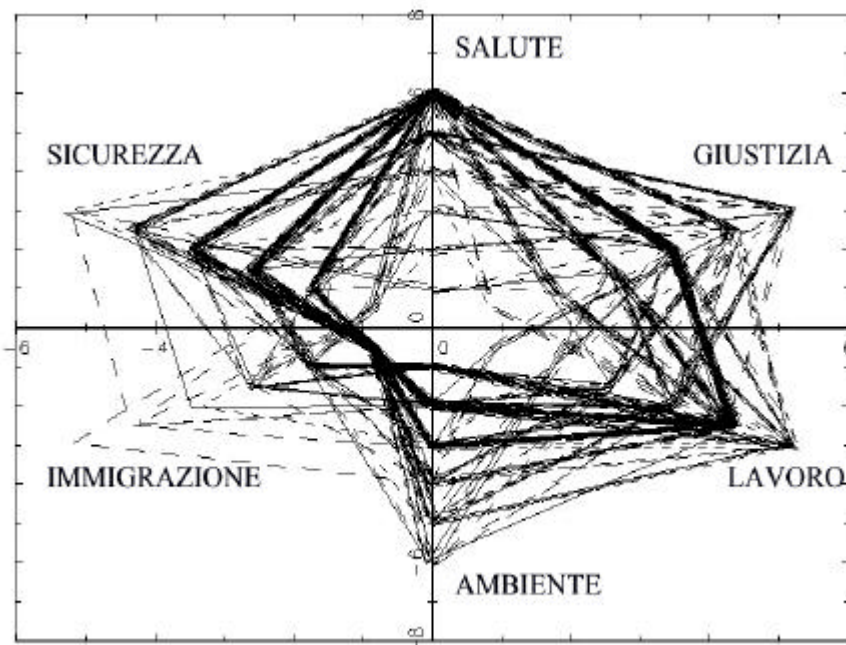


Figura 6. Priorità assegnate alle 6 tematiche

In sostanza, in un diagramma polare diviso in 6 settori (quante sono le possibili modalità da ordinare), si sono rappresentati i ranghi *jittered* r_{ij}^* , ove $r_{ij}^* = r_{ij} + \mathbf{e}_{ij}$, essendo r_{ij} il rango scelto dall' i -esimo soggetto per la j -esima modalità ($j=1,2,\dots,6$), ed \mathbf{e}_{ij} una v.c. Normale di valore medio nullo e con varianza sufficientemente piccola affinché la rappresentazione complessiva non risulti alterata.

In tal modo, si realizzano molteplici obiettivi:

- i) ciascun individuo viene descritto da un pattern che manifesta ed evidenzia la sua particolare configurazione;
- ii) i grafici non si sovrappongono e manifestano con chiarezza eventuali addensamenti che indicherebbero una tipologia identica;
- iii) si individua visivamente un pattern prevalente all'interno dei gruppi e nel loro confronto.

Attraverso tali rappresentazioni sono state analizzate le priorità espresse da tutti gli intervistati (Figura 6) e quelle riferite a singoli

sottogruppi: Sesso, Facoltà, Impegno politico familiare (Figura 7), Collocazione ideologica (Figura 8).

Si osservi che, in ciascuno dei grafici seguenti, il primo settore (quello in alto al centro) corrisponde alla tematica *salute*, cui seguono, nell'ordine e in senso orario, *giustizia*, *lavoro*, *ambiente*, *immigrazione* e *sicurezza*.

Osservando le grandi concentrazioni di linee si evince con facilità che la graduatoria emersa dalle precedenti analisi trova conferma in tale grafico. Trovano altresì conferma le considerazioni rispetto alla sospetta confusione tra *giustizia* e *sicurezza*. Le linee si concentrano, infatti, per la prima nei ranghi intermedi 2, 3 e 4, e per la seconda in quelli 2, 3, 4 e 5, equamente. Si attestano agli ultimi posti della graduatoria *ambiente* e *immigrazione*: per quest'ultima modalità l'addensamento in corrispondenza del rango 1 è piuttosto vistosa.

Dai grafici della Figura 7 emergono una serie di considerazioni che qui riassumiamo brevemente:

i) parte del sottogruppo riferito alle donne si discosta dalla graduatoria generale dando precedenza alle problematiche riguardanti *giustizia*, *sicurezza* e, anche se in misura ridotta, *immigrazione*, mentre i maschi danno maggiore rilevanza a quella riguardante il *lavoro*;

ii) gli studenti e le studentesse che maggiormente si discostano dalla graduatoria di base appartengono alla Facoltà di Scienze Politiche;

iii) le difformità rispetto al pattern principale si riscontrano per lo più nella popolazione di studenti che non proviene da una famiglia politicizzata.

Maggiore interesse assume, infine, il confronto tra i grafici relativi ai sottogruppi di Sinistra, Centro e Destra (Figura 8), dai quali si possono trarre alcune considerazioni riassuntive:

i) il sottogruppo di Sinistra si manifesta con il maggior numero di anomalie. A differenza di quanto viene generalmente dichiarato, con eccessiva semplificazione, molti di tali studenti ritengono rilevanti e prioritarie tematiche quali *sicurezza*, *giustizia* e *immigrazione*. Per contro, si conferma un certo interesse per l'*ambiente* da parte dei giovani di Sinistra, avendo alcuni studenti espresso una priorità per tale tematica;

ii) gli studenti aderenti alle opzioni di Centro e di Destra confermano sostanzialmente la graduatoria principale. Pochi studenti danno precedenza alle problematiche riguardanti *immigrazione* o *ambiente*. La *sicurezza*, tematica ritenuta generalmente di “destra”, non riscuote il primo posto nella graduatoria per nessuno studente di tali sottogruppi.

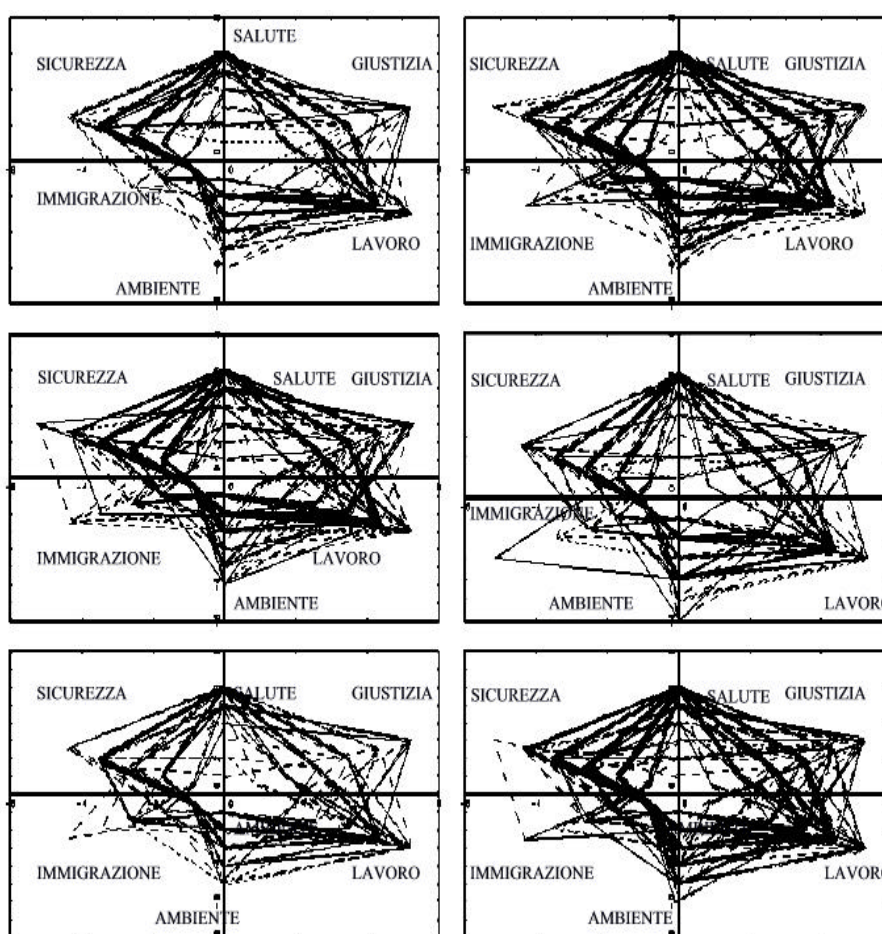


Figura 7. Priorità delle tematiche distinte (dall'alto in basso) per Sesso, Facoltà e Impegno politico di un familiare.

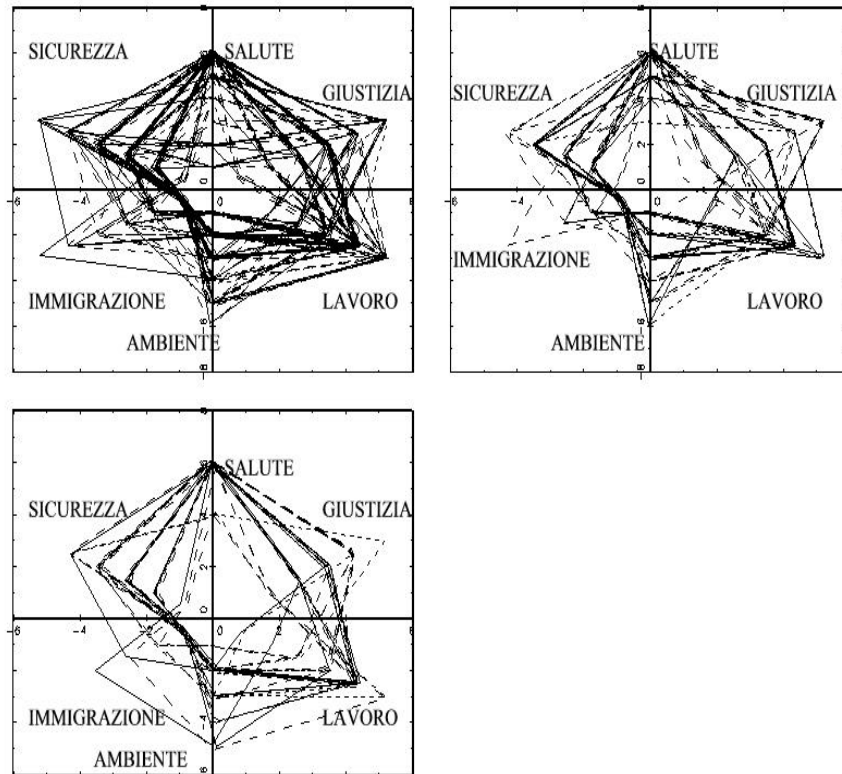


Figura 8. Priorità delle 6 tematiche distinte per ideologia politica

6. Considerazioni finali

In questo lavoro sono stati esaminati modelli probabilistici e proposti strumenti grafici per analizzare le differenze e le omogeneità in un particolare campione di studenti universitari, con riferimento alle principali posizioni politiche.

Attraverso un'indagine condotta su studenti delle Facoltà di Scienze Politiche e di Sociologia dell'Università degli Studi di Napoli Federico II sono state individuate alcune variabili individuali e di contesto che appaiono statisticamente connesse alla scelta politica e

all'impegno sociale. In tale ambito, è emersa con chiara evidenza una sostanziale omogeneità degli atteggiamenti e delle priorità espresse dagli studenti, a prescindere dalle differenti scelte politiche.

L'approccio ha mostrato, su uno specifico caso di studio, la coerente interazione tra strumenti grafici di natura statistica e analisi modellistica. In particolare, la proposta di visualizzare i ranghi *jittered* mediante una grafica di tipo polare, ha consentito di trasformare la singola permutazione delle preferenze in un *pattern* specifico con il quale confrontare le differenti similarità.

Infine, rispetto all'approccio tradizionale derivante dai modelli di regressione, i risultati qui discussi hanno evidenziato come i modelli lineari generalizzati producano risultati di agevole e coerente interpretazione in presenza di variabili qualitative, che sono usualmente diffuse nelle analisi socio-politiche.

Ringraziamenti: Il lavoro ha usufruito di contributi finanziari nell'ambito di progetti di ricerca afferenti al Dipartimento di Scienze Statistiche dell'Università degli Studi di Napoli Federico II. Si ringraziano gli anonimi referees che hanno contribuito in modo sostanziale a migliorare la presentazione dei risultati.

Riferimenti Bibliografici

Bishop Y.M.M, Fienberg S.E., Holland P.W. (1975) *Discrete Multivariate Analysis: Theory and Practice*, MIT Press, Cambridge, MA.

Brown P.J. and Payne C.D. (1986) Aggregate Data, Ecological Regression, and Voting Transition, *Journal of the American Statistical Association*, 81, 452-460.

Dobson A.J. (1990) *An Introduction to Generalized Linear Models*, Chapman & Hall, London.

Keener R.W. and Waldman D.M. (1985) Maximum Likelihood Regression of Rank Censored Data, *Journal of the American Statistical Association*, 80, 385-392.

Marden J.I. (1995) *Analyzing and Modeling Rank Data*, Chapman & Hall, London.

McCullagh P. and Nelder J.A. (1989) *Generalized Linear Models* (2nd edition), Chapman & Hall, London.

Upton G.J.G. and Brook D. (1974) The Importance of Positional Voting Bias in British Elections, *Journal of Political Studies*, 22, 178-190.

Taplin R.H. (1997) The Statistical Analysis of Preference Data, *Applied Statistics*, 46, 493-512.