

# Un percorso: materiali per gli studenti

## 1 Motivazioni alla normale

### 1.1 I sondaggi

Per cercare di stabilire la posizione di una popolazione numerosa riguardo ad una determinata questione, sono molte le istituzioni che effettuano sondaggi<sup>1</sup> nei più svariati ambiti.

In particolare, i sondaggi che hanno valenza politica ed elettorale e che sono diffusi al pubblico, devono<sup>2</sup> essere pubblicati sul sito

[www.sondaggipoliticoelettorali.it](http://www.sondaggipoliticoelettorali.it) a cura del Dipartimento per l'Informatica e l'Editoria.

Vediamone uno: *Gli italiani ed il Referendum del 17 Aprile*

SONDAGGIO		
Dati Sondaggio	Domande	Conclusioni
Titolo del sondaggio		Soggetto che ha realizzato il sondaggio
Gli italiani ed il Referendum del 17 aprile		DEMOPOLIS - Istituto di Ricerche
Soggetto committente		Soggetto acquirente
Otto e Mezzo (LA7)		LA7 Srl
Data o periodo in cui è stato realizzato il sondaggio - Da		Data o periodo in cui è stato realizzato il sondaggio - A
22/03/2016		24/03/2016
Mezzo(i) di comunicazione di massa sul quale(i) è stato pubblicato o diffuso il sondaggio		Data di pubblicazione o diffusione
Programma Otto e Mezzo, diffuso anche da l'Espresso online e dai quotidiani Il Tirreno, Messaggero Veneto, Il Centro, La Nuova Sardegna, ecc.		30/03/2016
Popolazione di riferimento		Estensione territoriale del sondaggio
Popolazione italiana maggiorenne		Nazionale
Metodo di campionamento, inclusa l'indicazione se trattasi di campionamento probabilistico o non probabilistico, del panel e l'eventuale ponderazione		Consistenza numerica del campione di intervistati, numero dei non rispondenti e delle sostituzioni effettuate
Campione: probabilistico statisticamente rappresentativo dell'universo di riferimento, stratificato per genere, età, ampiezza demografica ed area geografica di residenza		Consistenza numerica del campione: 1.000 intervistati. Rispondenti: interviste complete effettuate 1.000 (33,1%); rifiuti/sostituzioni 2.024 (66,9%); totale contatti 3.024 (100%)
Rappresentatività del campione, inclusa l'indicazione del margine d'errore		Metodo raccolta delle informazioni
Rappresentatività dei risultati: popolazione italiana maggiorenne; margine massimo di errore 2,8%		Metodo di raccolta delle informazioni: cawi-cati per la somministrazione del questionario strutturato di rilevazione

SONDAGGIO - Domande e Risposte		
Dati Sondaggio	Domande	Conclusioni
N° Domanda	Domanda	
1	Lei sarebbe propenso a vietare alla scadenza il rinnovo delle attuali concessioni per l'estrazione in mare entro 12 miglia dalle coste?	
2	Per quali ragioni lei è favorevole al divieto di rinnovo delle concessioni?	
3	Per quali ragioni lei è contrario al divieto di rinnovo delle concessioni?	
4	Lei sa che il 17 aprile si terrà il Referendum sulle trivellazioni?	
Pagina 1 di 1 (4 elemento)		
Area Scheda Domanda e Risposta		
Testo Domanda		
Lei sarebbe propenso a vietare alla scadenza il rinnovo delle attuali concessioni per l'estrazione in mare entro 12 miglia dalle coste?		
Testo Risposta		
Sì: 74%		
No: 26%		
Dati ripercentralizzati in assenza del non sa (15%)		

Esaminiamo i termini per noi più significativi:

<sup>1</sup> I sondaggi vengono spesso fatti online mediante software quali cawi-cati, come nell'esempio riportato in figura (è indicato nella sezione "Metodo raccolta delle informazioni").

<sup>2</sup> Ciò è previsto dalla legge n. 28 del 22 Febbraio 2000, sulla par condicio.

- *Campione.* È importante avere informazioni sull'orientamento di voto di una popolazione o sull'opinione in merito ad una data questione. Un modo per ottenerlo è mediante le elezioni o i referendum; ma non si può coinvolgere l'intera popolazione tutte le volte. Pertanto, si considera un opportuno sottoinsieme della popolazione, detto appunto campione, e si registrano le preferenze espresse da esso. Le frequenze relative ottenute dal campione sono una stima delle preferenze dei singoli individui dell'intera popolazione.
- *Margine d'errore.* Se il sondaggio ha un margine massimo d'errore del 2,8%, significa che la probabilità dell'evento in questione (per esempio "l'individuo vota SI") si discosterà al massimo della quantità 0,028 dalla stima ottenuta sul campione...con probabilità "grande"<sup>3</sup>.
- *Consistenza numerica del campione.* Il numero di cittadini che costituiscono il campione è strettamente legato al margine di errore del sondaggio. Una questione significativa al riguardo è determinare la dimensione del campione affinché la stima fatta su esso sia "buona". Questa però è una questione articolata di cui non ci occupiamo.

## La questione

Una popolazione costituita da **10.000** individui è chiamata a votare tra due candidati, diciamo *A* e *B*. Mediante un sondaggio effettuato su un campione della popolazione, si è stimato che la probabilità che l'**individuo** sia favorevole ad *A* è del **40%**.

Sulla base di ciò, si vuole stimare il numero **F** di individui che voteranno *A*. Precisamente, quanto vale

$$P(3900 \leq F \leq 4100) ?$$

*Osservazioni:*

- In sostanza, a partire dalla percentuale di individui del campione che sono a favore di *A*, si vuole stimare la probabilità che il generico individuo della popolazione sia a favore di *A*.
- Gli estremi di variabilità stabiliti per *F* nel problema rappresentano rispettivamente il 39% e il 41% della popolazione. Pertanto la richiesta si può così riformulare: qual è la probabilità che il numero *F* sia compreso tra  $(40 - 1)\%$  e  $(40 + 1)\%$  della popolazione?

<sup>3</sup>Precisamente con probabilità che di solito è del 95% (in fisica, invece, si usano margini molto più bassi!). Secondo un approccio frequentista, su molti campionamenti ci si aspetta che circa il 95% delle volte, la probabilità *p* sulla popolazione cada nell'intervallo  $[\bar{p} - 0,028; \bar{p} + 0,028]$ , dove  $\bar{p}$  è la stima sul campione.