TEORIA E TECNICA DELL'INDAGINE STATISTICA E DEL CAMPIONAMENTO (MATR.DISPARI)

CAMPIONAMENTO CASUALE SEMPLICE INTERVALLI DI CONFIDENZA E NUMEROSITÀ OTTIMALE

MANUELA SCIONI
Dipartimento di Scienze Statistiche
manuela.scioni@unipd.it

una volta che ho la varianza di stima per la varianza campionaria posso utilizzarla per int conf che a sua volta mi servira per i calcoli della numerosità campionaria
conterra il vero valore del parametro
nella pratica conosciamo un solo campione cioe n ma grazie al lim centrale posso calcolare cosi tende ad una normale standardizzata

INTERVALLO DI CONFIDENZA PER LA MEDIA

Formulazione del teorema del limite centrale (valida per N, n e (N-n) grandi):

$$\frac{\bar{y}-\bar{Y}}{se(\bar{y})} = \frac{\bar{y}-\bar{Y}}{\frac{s}{\sqrt{n}}\sqrt{1-f}} \rightarrow N(0,1)$$

grazie alla appx di una normale anche da pop finite

Intervallo di confidenza

$$\bar{y} - z\alpha/_2 \frac{s}{\sqrt{n}}\sqrt{1-f} \leq \bar{Y} \leq \bar{y} + z\alpha/_2 \frac{s}{\sqrt{n}}\sqrt{1-f}$$

Se il parametro è un totale

$$N\bar{y} - z\alpha/_2 N\frac{s}{\sqrt{n}}\sqrt{1-f} \leq T \leq N\bar{y} + z\alpha/_2 N\frac{s}{\sqrt{n}}\sqrt{1-f}$$

Se il parametro è una proporzione

$$\hat{p} - z\alpha/_2 \sqrt{\frac{(1-f)\hat{p}(1-\hat{p})}{n-1}} \leq P \leq \hat{p} + z\alpha/_2 \sqrt{\frac{(1-f)\hat{p}(1-\hat{p})}{n-1}}$$

DETERMINAZIONE DELLA NUMEROSITÀ OTTIMALE

Fissato un livello di precisione (es: l'ampiezza dell'intervallo di confidenza), quanto deve essere n affinché tale livello sia garantito?

$$2\mathrm{D} = \text{ampiezza prefissata dell'intervallo di confidenza}$$

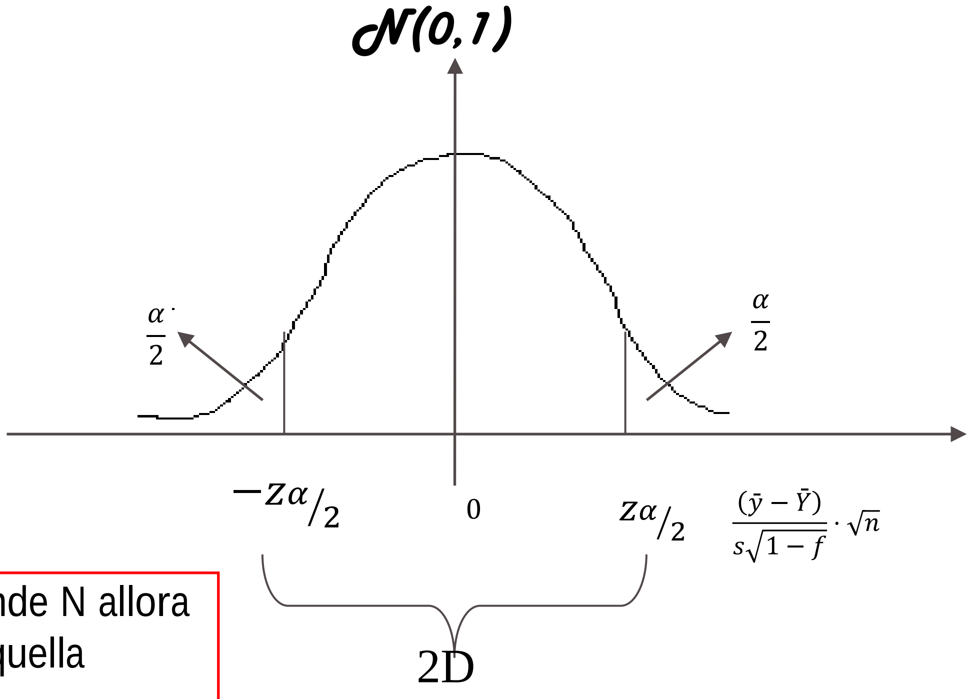
$$2D = 2z_{\alpha/2}\frac{S}{\sqrt{n}}\sqrt{1-f}$$

non lo consco e nemmeno la stima della variabiel di interesse xkr sto andando a stimare la numerosita
allora istat utilizza una stima di un indagine precedente es variabilita reddito , se dicotomica considero caso di max variazione,oppure indagine pilota

$$\frac{N \cdot D^2}{z^2_{\alpha/2} S^2} = \frac{1}{n}(N-n)$$

$$n\left(\frac{N \cdot D^2}{z^2_{\alpha/2} S^2} + 1\right) = N \quad \rightarrow \quad n = \frac{N}{\frac{N \cdot D^2}{z^2_{\alpha/2} S^2} + 1} = \frac{z^2_{\alpha/2} S^2}{D^2 + \frac{z^2_{\alpha/2} S^2}{N}}$$

se molto grande N allora tende a uno quella frazione



NUMEROSITÀ OTTIMALE: OSSERVAZIONI

come trovarer S^2?

- Per stimare n occorre stimare la varianza di Y, S^2. Come?
  - Attraverso indagini pilota.
  - Da informazioni ottenute da studi precedenti.
  - Da considerazioni sulla natura della variabile e della popolazione.
- Se Y è dicotomica, la situazione di massima variabilità si ha per p=0,5, da cui S^2=0,25 (vedere file aggiuntivo: Determinazione per n per proporzione).
- Se l'indagine è multiscopo, ad ogni variabile corrisponde un valore ottimale di n. Il valore finale di n sarà o il valore ottimale massimo o la media ponderata dei valori ottimali di n, con pesi proporzionali all'importanza della variabile:

di solito abb piu v. di interesse non solo una

$$n = \sum_g W_g n_g , \qquad \sum W_g = 1$$

NUMEROSITÀ CAMPIONARIA: BEN OLTRE IL LIVELLO DI PRECISIONE ATTESO…

La numerosità ottima di un campione è quella che permette di ottenere gli obiettivi dell'indagine al minimo costo, e sarà il più piccolo numero in base al quale le stime raggiungono il livello di attendibilità atteso dal ricercatore

I fattori che influiscono sulla determinazione della dimensione ottimale del campione possono essere:

- La varianza del fenomeno osservato
- L'errore campionario che si ritiene ammissibile
- Il costo (costo fisso + costo per osservazione)
- La presenza di molte variabili
- L'esigenza di stime significative per sub-popolazioni (domini di studio)