STIMA DEI PARAMETRI

STIMA PUNTUALE DEI PARAMETRI

Per STIMA PUNTUALE DEI PARAMETRI s'intende l'insieme dei metodi inferenziali che permettono di attribuire un valore ad un parametro della popolazione, utilizzando i dati di un campione casuale osservato ($x_1, x_2, \ldots, x_n$) ed elaborandoli.

Sia ($x_1, x_2, \ldots, x_n$) il campione casuale osservato e sia $g(\bullet)$ una funzione matematica a n variabili. Applicando $g(\bullet)$ al campione casuale osservato si ottiene un valore:

$$g(\bullet) = g(x_1, x_2, \ldots, x_n)$$

Si definisce STIMA il risultato ***g*** delle elaborazioni dei dati di un campione casuale osservato, al fine di stimare il parametro di interesse $\boldsymbol{\gamma}$ della popolazione.

Si definisce STIMATORE del parametro $\boldsymbol{\gamma}$ della popolazione la *statistica campionaria* $G = g(X_1, X_2, \ldots, X_n)$ utilizzata per stimare $\boldsymbol{\gamma}$.

Proprietà degli STIMATORI

A) CORRETTEZZA o NON DISTORSIONE

Uno stimatore G di definisce CORRETTO o NON DISTORTO se la sua media coincide con il parametro da stimare:

$$M(G) = \gamma$$

Se $M(G) \neq \gamma$, allora G si definisce stimatore *non corretto* o *distorto*: la quantità $D = M(G) - \gamma$ si dice ***distorsione***.

B) EFFICIENZA

Il concetto di *efficienza* di uno stimatore si basa sulla sua variabilità: uno stimatore sarà tanto più efficiente, ovvero preciso, quanto più piccola risulta la sua Varianza, e quindi il suo S.q.m.

Uno stimatore è tanto più efficiente quanto più piccolo è il suo errore medio $\sigma_G = \sqrt{Var(G)}$.

EFFICIENZA ASINTOTTICA

Se la varianza dello stimatore tende a zero per n (dimensione campionaria) che tende all'infinito:

$$\lim_{n \to +\infty} Var(G) = 0$$

allora G si definisce stimatore asintoticamente efficiente.

C) CONSISTENZA

Si definisce G stimatore CONSISTENTE del parametro γ se e solo se vale la seguente relazione:

$$\lim_{n \to +\infty} P(|G - \gamma| < \varepsilon) = 1$$

Teorema: Se uno stimatore G è corretto e asintoticamente efficiente, allora G è stimatore consistente di γ.

STIMA DEI PARAMETRI

STIMA PER INTERVALLO

La tecnica inferenziale della STIMA PER INTERVALLO ha lo scopo di fissare un intervallo entro il quale si trovi, con la probabilità assegnata (: ***1-α***), il valore vero del parametro ***γ*** incognito della popolazione.
Se questa probabilità ***1-α*** è vicina ad *uno*, allora si ottiene un intervallo che contiene quasi certamente il parametro. Invece la probabilità di estrarre un campione il cui intervallo non contenga il parametro è ***α*** e quindi è piccola: ***α*** esprime la misura del rischio di errore, ovvero la probabilità di sbagliare.

INTERVALLO DI CONFIDENZA o *INTERVALLO FIDUCIARIO*

$$1-\alpha = P\{\gamma_1 < \gamma < \gamma_2\}$$

La quantità ***(1-α)*** è il Coefficiente di confidenza.

Gli estremi γ_1 e γ_2 sono i Limiti di confidenza.

INTERVALLO DI CONFIDENZA per la Media $\boldsymbol{\mu}$ *della popolazione*

$$1-\alpha = P\{\mu_1 < \mu < \mu_2\}$$

<u>Se</u> la Media campionaria m si distribuisce NORMALMENTE, allora

$$1-\alpha = P\left\{-u_{\frac{\alpha}{2}} < u < +u_{\frac{\alpha}{2}}\right\}$$

Dove $u = u_m = \dfrac{m-\mu}{\frac{\sigma}{\sqrt{n}}}$ (nel caso di campione casuale CON ripetizione), quindi sostituendo si ottiene:

$$1-\alpha = P\left\{-u_{\frac{\alpha}{2}} < \frac{m-\mu}{\frac{\sigma}{\sqrt{n}}} < +u_{\frac{\alpha}{2}}\right\}$$

Da cui, effettuando alcuni passaggi algebrici ai fini di *isolare* **μ**, si ottiene la formula finale dell' *INTERVALLO DI CONFIDENZA* per la Media **μ** della popolazione:

$$1-\alpha = P\left\{m - u_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < m + u_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right\}$$

Invece, nel caso di campione casuale SENZA ripetizione e popolazione di dimensione N finita, nella formula della *Var(**m**)* bisogna tener conto del fattore di correzione, quindi:

$$u = u_m = \frac{m-\mu}{\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}}$$

Da sostituire nella formula:

$$1-\alpha = P\{\mu_1 < \mu < \mu_2\} = P\left\{-u_{\frac{\alpha}{2}} < u < +u_{\frac{\alpha}{2}}\right\} = P\left\{-u_{\frac{\alpha}{2}} < \frac{m-\mu}{\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}} < +u_{\frac{\alpha}{2}}\right\}$$

Da cui, effettuando alcuni passaggi algebrici ai fini di *isolare* **μ**, si ottiene la formula finale dell' *INTERVALLO DI CONFIDENZA* per la Media **μ** della popolazione:

$$1-\alpha = P\left\{m - u_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} < \mu < m + u_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}\right\}$$

<u>Se</u> la Media campionaria m <u>standardizzata</u> si distribuisce come una v.c. ***t*** *di Student,* allora

$$1-\alpha = P\left\{-t_{\frac{\alpha}{2}} < t < +t_{\frac{\alpha}{2}}\right\}$$

Dove $t = t_m = \frac{m-\mu}{\frac{s}{\sqrt{n}}}$, quindi sostituendo si ottiene:

$$1-\alpha = P\left\{-t_{\frac{\alpha}{2}} < \frac{m-\mu}{\frac{s}{\sqrt{n}}} < +t_{\frac{\alpha}{2}}\right\}$$

Da cui, effettuando alcuni passaggi algebrici ai fini di *isolare* **μ**, si ottiene la formula finale dell' *INTERVALLO DI CONFIDENZA* per la Media **μ** della popolazione:

$$1-\alpha = P\left\{m - t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} < \mu < m + t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}\right\}$$

Esercizio Da una popolazione con varianza $\sigma^2=49$ è stato estratto con reinserimento un campione casuale di n=100 unità; sia *50* il valore della *media campionaria*.
Determinare l'intervallo di confidenza al livello del 92% per la media della popolazione.

SOLUZIONI

Poiché n=100>30 allora la *media campionaria* si distribuisce normalmente.
Dato un livello di significatività pari al *92%=1-α*, segue *α=8%* e *α/2=4%*; quindi sulle tavole della v.c. Normale standardizzata bisogna individuare i valori

$$\pm u_{\frac{\alpha}{2}} = \pm u_{4\%} = \pm 1{,}75$$

$$1-\alpha = P\left\{m - u_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < m + u_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right\}$$

$$1-\alpha = P\left\{50 - 1{,}75 \cdot \frac{7}{\sqrt{100}} < \mu < 50 + 1{,}75 \cdot \frac{7}{\sqrt{100}}\right\}$$

$$0{,}92 = P\{48{,}775 < \mu < 51{,}225\}$$

Esercizio Da un campione casuale di 35 impiegati (: n=35) risulta che la retribuzione mensile (in migliaia di euro) presenta *media campionaria* pari a *2* e *varianza campionaria corretta* pari a *0,25*.
Determinare l'intervallo di confidenza al livello del 96% per la media della popolazione.

SOLUZIONI

Poiché n=35>30 allora la *media campionaria* si distribuisce normalmente.

$$1-\alpha = P\left\{m - u_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} < \mu < m + u_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}\right\}$$

$$1-\alpha = P\left\{2 - 2{,}05 \cdot \frac{0{,}5}{\sqrt{35}} < \mu < m + u_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}\right\}$$

$$0{,}96 = P\{1{,}82674 < \mu < 2{,}17326\}$$

INTERVALLO DI CONFIDENZA per la Varianza $\boldsymbol{\sigma^2}$ *della popolazione, nel caso di popolazione Normale*

$$1-\alpha = P\left\{\sigma_1^2 < \sigma^2 < \sigma_2^2\right\}$$

Poiché, se la popolazione d'origine è Normale la Varianza campionaria $\boldsymbol{S^2}$ si distribuisce come una *v.c. chi-quadrato*, a meno di una coefficiente di proporzionalità, allora si può scrivere:

$$1-\alpha = P\left\{\chi^2_{\left(1-\frac{\alpha}{2}\right)} < \chi^2 < \chi^2_{\frac{\alpha}{2}}\right\}$$

Dove $\chi^2 = \dfrac{S^2 \cdot \nu}{\sigma^2}$, quindi sostituendo si ottiene:

$$1-\alpha = P\left\{\chi^2_{\left(1-\frac{\alpha}{2}\right)} < \frac{S^2 \cdot \nu}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}\right\}$$

Da cui, effettuando alcuni passaggi algebrici ai fini di *isolare* $\boldsymbol{\sigma^2}$, si ottiene la formula finale dell' *INTERVALLO DI CONFIDENZA* per la Varianza $\boldsymbol{\sigma^2}$ della popolazione:

$$1-\alpha = P\left\{\frac{S^2 \cdot \nu}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{S^2 \cdot \nu}{\chi^2_{\left(1-\frac{\alpha}{2}\right)}}\right\}$$

Esercizio Da una popolazione distribuita normalmente viene estratto un campione casuale di 10 unità (: n=10); la varianza calcolata sul campione risulta pari a 600.
Determinare l'intervallo di confidenza al livello del 95% per la varianza della popolazione.

SOLUZIONI

$$1-\alpha = P\left\{\frac{S^2 \cdot \nu}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{S^2 \cdot \nu}{\chi^2_{\left(1-\frac{\alpha}{2}\right)}}\right\}$$

Dove *95%=1-α*, da cui segue *α=5%* e *α/2=2,5%*; quindi sulle tavole della v.c. "chi-quadrato" bisogna individuare i seguenti valori, in corrispondenza a ***ν****=n-1=10-1=****9***:

$$\chi^2_{\frac{\alpha}{2}} = \chi^2_{2,5\%} = 19{,}02$$

$$\chi^2_{\left(1-\frac{\alpha}{2}\right)} = \chi^2_{97,5\%} = 2{,}70$$

La varianza del campione è pari a 600, quindi bisogna effettuare la *correzione* per ottenere ***S²***:

$$S^2 = 600 \cdot \frac{n}{n-1} = 600 \cdot \frac{10}{9} = 666{,}67$$

Da cui, sostituendo:

$$1-\alpha = 0{,}95 = P\left\{\frac{666{,}67 \cdot 9}{19{,}02} < \sigma^2 < \frac{666{,}67 \cdot 9}{2{,}70}\right\} =$$
$$= P\{315{,}45 < \sigma^2 < 2222{,}22\}$$