LEZIONE 9 18/05/09

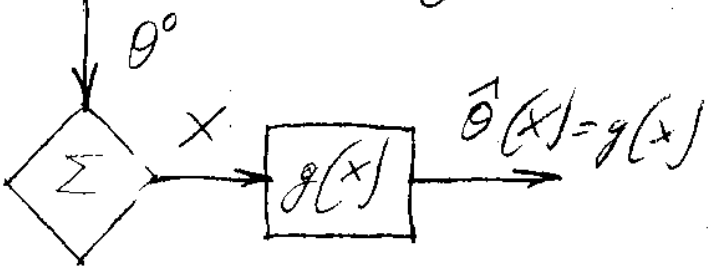
(Inizio della 2a parte del corso)

IL PROBLEMA DELLA STIMA

Si passa ora dal calcolo della probabilità alla statistica allo scopo di elaborare i dati.

Si tratta di raccogliere i dati e a partire da essi arrivare a determinare le quantità incognite di interesse.

La legge dei grandi numeri rappresenta un importante risultato acquisito nella prima parte del corso. Essa afferma che sotto determinate ipotesi, continuando a raccogliere dati e a calcolare le relative medie campionarie, queste convergono al valore teorico.



SCHEMA DEL PROBLEMA DELLA STIMA

Schematizzando il problema della stima, θ^0 è il parametro incognito, Σ è la sorgente casuale dei dati, x è il vettore dei dati,

$g(x)$ è lo stimatore, in pratica un algoritmo

$\hat{\theta}(x)$ è la stima, il risultato dell'elaborazione

I dati raccolti (x) dipendono dai parametri incogniti (θ^0) in modo generalmente sofisticato. Inoltre contengono sempre una componente casuale, quanto meno sono affetti da errori di misura.

x è un vettore di dati casuali e $g(x)$ è una funzione di v.c. Anche $\hat{\theta}(x)$, definita come $g(x)$ è una v.c.

Lavorando con gli stimatori, la base su cui si opera è casuale ed i risultati sono, a loro volta, casuali.

Esempio: consideriamo una v.c. gaussiana con media m e varianza σ^2 incognite.

In questo caso θ^0 è un vettore: $\theta^0 = \begin{bmatrix} m \\ \sigma^2 \end{bmatrix}$

150 Ipotizziamo che la sorgente dei dati Σ consista nell'operazione di osservazione dei dati

$$X_1, X_2, \dots, X_n \quad i.i.d.$$

~~con~~ gaussiane con media m e varianza σ^2.

Immaginiamo anche che sia possibile replicare n volte in maniera indipendente l'esperimento associato alla variabile gaussiana.

In queste ipotesi risulta:

$$X = [X_1 \ X_2 \ \dots \ X_n]^T$$

con X v.c. vettoriale con d.d.p.

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^2 \sigma^{2n}}} e^{-\frac{1}{2}\sum_{i=1}^{n} \frac{(x_i - m)^2}{\sigma^2}} \quad \text{(pag. 141)}$$

oppure

$$f_X(\theta) = \frac{1}{\sqrt{(2\pi)^2 \sigma^{2n}}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} (x_i - m)^2}$$

La $f_X(x)$ dipende da θ, essendo

$$\sigma^2 = \theta_2^0 \quad e \quad m = \theta_1^0$$

I dati dipendono dalle incognite che rappresentano i parametri incogniti della d.d.p.

Per ciò che attiene allo stimatore, si ha:

$$\hat{\theta} = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = \begin{bmatrix} g_1(X_1, X_2, \dots, X_n) \\ g_2(X_1, X_2, \dots, X_n) \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} X_i \\ \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \bar{X})^2 \end{bmatrix}$$

Le componenti tipiche dello stimatore sono la media e la varianza campionaria.

-) <u>Stima dei momenti di una v.c. (momenti campionari)</u>

Si consideri un esperimento che fornisce $X_1, X_2, \dots, X_n$, i.i.d.

Il problema da affrontare consiste nello stimare i momenti di X_i (media, valore quadratico medio, varianze, ...).

Introduciamo una classe di possibili stimatori con le definizioni seguenti

-) <u>Momento campionario di ordine k</u>

$$M_k := \frac{1}{n} \sum_{i=1}^{n} X_i^k$$

M_k è una v.c. poiché dipende dalle X_i che sono delle v.c.

Si noti che è indicato con la lettera maiuscola

-) <u>Momenti centrali campionari di ordine k</u>

$$S_k := \frac{1}{n} \sum_{i=1}^{n} (X_i - M_1)^k$$

Si noti che i momenti campionari di ordine k sono gli stimatori dei momenti di ordine k.
I momenti centrali campionari di ordine k sono gli stimatori dei momenti centrali di ordine k.

M_k è uno stimatore di $E[X_i^k]$

S_k è uno stimatore di $E[(X_i - m)^k]$ con $m = E[X_i]$

Analizziamo le proprietà di questi stimatori.

-) <u>Proprietà della media campionaria M_1</u>

$$M_1 = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X}$$

Ricordiamo che gli stimatori sono delle v.c.

-) $E[M_1] = m = E[X_i]$ (pag. 137)

-) $Var[M_1] = \frac{\sigma^2}{n}$ con $\sigma^2 = Var[X_i]$ (pag. 137)

-) d.d.p.

152 Se le v.c. $X_i \sim N(m, \sigma^2)$, allora M_1 è una v.c. gaussiana.
Infatti essa è una combinazione lineare di v.c. gaussiane

-) Proviamo a standardizzare definendo la v.c. Z

$$Z := \frac{M_1 - m}{\sqrt{\frac{\sigma^2}{n}}} = \frac{M_1 - m}{\frac{\sigma}{\sqrt{n}}}$$

Al denominatore appare la deviazione standard.
La trasformazione di standardizzazione è lineare, perciò Z è una v.c. gaussiana e standard.
Esprimiamo M_1 in funzione di Z

$$M_1 = m + \frac{\sigma}{\sqrt{n}} Z$$

m rappresenta il valore vero da stimare.
$\left(\frac{\sigma}{\sqrt{n}} Z\right)$ è un termine di errore che contiene Z che è una v.c. standard gaussiana con dispersione nota, ma contiene anche una parte che tende a zero, se n diventa grande.
Quindi l'errore tende a zero se n tende ad ∞.
Sotto queste considerazioni la media campionaria può essere vista come un trasduttore virtuale la cui qualità e precisione sono legate al numero di dati raccolti.

-) Supponiamo che X_i non siano gaussiani.
In questo caso il teorema centrale del limite garantisce che per $n \to \infty$, M_1 tende in distribuzione ad una v.c. gaussiana.

-) Proprietà dei momenti di ordine 2

$$M_2 := \frac{1}{n} \sum_{i=1}^{n} X_i^2 \quad \text{(stimatore di } E[X_i^2])$$

-) $E[M_2] = E[X_i^2]$

Dimostrazione:

$$E[M_2] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i^2] =$$

$$= \frac{1}{n}\, n\, E[X_i^2] = E[X_i^2]$$

Ricordiamo che il valore atteso della somma è la somma dei valori attesi e che questi sono tutti uguali.

•) <u>Varianza campionaria noto il valor medio</u>

Talvolta capita di conoscere il valor medio

$$m = E[X_i]$$

allora nell'espressione della varianza campionaria si utilizza il valore vero anziché la media campionaria.

$$S_m^2 := \frac{1}{n}\sum_{i=1}^{n} (X_i - m)^2$$

L'eventualità capita di rado.

<u>Proprietà della varianza campionaria noto il valor medio</u>

-) Valor atteso

$$E[S_m^2] = E\left[\frac{1}{n}\sum_{i=1}^{n} (X_i - m)^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[(X_i - m)^2] =$$

$$= \frac{1}{n}\sum_{i=1}^{n} \sigma^2 = \frac{1}{\cancel{n}}\cancel{n}\sigma^2 = \sigma^2$$

Ricordiamo che con σ^2 si indica la $Var[X_i]$

$$\sigma^2 = Var[X_i] \quad \text{con } X_i \text{ i.i.d.}$$

Tutti gli X_i hanno la stessa media e la stessa varianza.

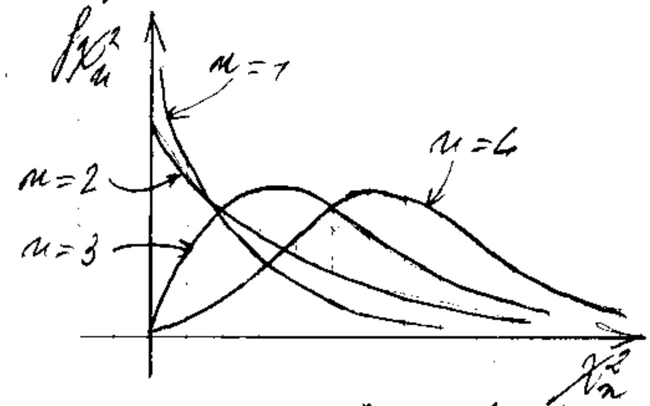
-) <u>La v.c. χ^2</u> ("chi quadrato")

Siano Z_i, con $i = 1, 2, \dots, n$, delle v.c. gaussiane standard i.i.d. In questa ipotesi

$$\chi^2_n = \sum_{i=1}^{n} Z_i^2$$

è detta "chi quadrato" ad n gradi di libertà.

<u>Proprietà</u>



•) $E[\chi^2_n] = n$

•) $Var[\chi^2_n] = 2n$

•) per $n \to \infty$ si nota:

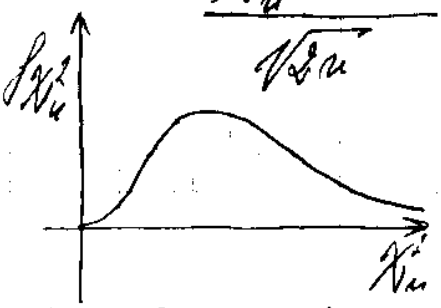
$$\frac{\chi^2_n - n}{\sqrt{2n}} \to Z$$

, con Z = gaussiana standard.

Ciò vale in virtù del teorema centrale del limite.

Ricordiamo che, con Z si indica una gaussiana standard.

La d.d.p. della χ^2_n per $n=1$ presenta un asintoto verticale mentre per $n=2$ ricorda una esponenziale, e per $n=3$ una lognormale!



ANDAMENTO TIPICO DI χ^2_n PER N GRANDE

•) <u>Varianza campionaria, noto il valor medio</u> (ripresa da pag. 153)

Se X_i sono delle Gaussiane, si può notare che

$$\frac{n S_m^2}{\sigma^2}$$

è una χ^2_n.

Dimostrazione:

$$\frac{n S_m^2}{\sigma^2} = \frac{\cancel{n}}{\sigma^2} \frac{1}{\cancel{n}} \sum_{i=1}^{n} (X_i - \mu)^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 =$$

$$= \sum_{i=1}^{n} Z_i^2$$

, con Z_i gaussiane standard (vedi pag. 7.)

Dunque,

$$\frac{n S_m^2}{\sigma^2} = \sum_{i=1}^{n} Z_i^2 = \chi^2_n$$

•) Varianza campionaria

È definita:

$$S_2 := \frac{1}{n} \sum_{i=1}^{n} (X_i - M_1)^2$$

Si noti che

$$S_2 \neq S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - M_1)^2$$

Proprietà di S_2

–) $E[S_2] = \frac{n-1}{n} \sigma^2$

Per dimostrarlo serve la proprietà

–) $S_2 = M_2 - M_1^2$, con M_2 = momento campionario di ordine 2 e

M_1 = media campionaria

La proprietà è simile a quella illustrata a pag. 84 e riferita alle v.c.

Dimostrazione della proprietà precedente

$$E[S_2] = E[M_2 - M_1^2] = E[M_2] - E[M_1^2] =$$
$$= m_2 - (Var[M_1] + E[M_1]^2)$$

ricordando che

$m_k = E[X_i^k]$, m_k = momento teorico di ordine k (pag. 151)

e $Var[M_1] = E[M_1^2] - E[M_1]^2$,

da cui

$$E[M_1^2] = Var[M_1] + E[M_1]^2$$

Riprendendo la dimostrazione

$$E[S_2] = m_2 - \left(\frac{\sigma^2}{n} + m^2\right) = \quad \text{(pag. 151)}$$
$$= m_2 - m^2 - \frac{\sigma^2}{n} = (m_2 - m^2) - \frac{\sigma^2}{n} =$$
$$= \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2$$

sempre ricordando che la differenza tra il valore quadratico medio e la media al quadrato è la varianza

$$\sigma^2 = m_2 - m_1^2$$

Gli stimatori visti finora avevano come baricentro proprio le quantità che dovevano stimare, mentre il baricentro di S_2 è un po' più piccolo di σ^2.

<u>Teorema</u> (di Fisher)

Se le variabili X_i sono i.i.d. e gaussiane, S_2 è una v.c.

$$\chi^2_{n-1}$$

(ad $n-1$ gradi di libertà) ed è indipendente dalla media campionaria, M_1, $\forall n$.

• Ronald Aylmer Fisher (Londra, 1890 – Adelaide 1962), statistico e matematico inglese

Ricordiamo che in generale la varianza campionaria dipende dalla media campionaria, come mostrato a pag. 155.

Nota: se le variabili X_i sono i.i.d. ma non gaussiane per $n \to \infty$, S_2 diventa gaussiana e indipendente da M_1 (risultato asintotico, per $n \to \infty$)

Il teorema precedente vale $\forall n$.

PROPRIETÀ DEGLI STIMATORI

Cerchiamo dei criteri di giudizio per valutare gli stimatori.

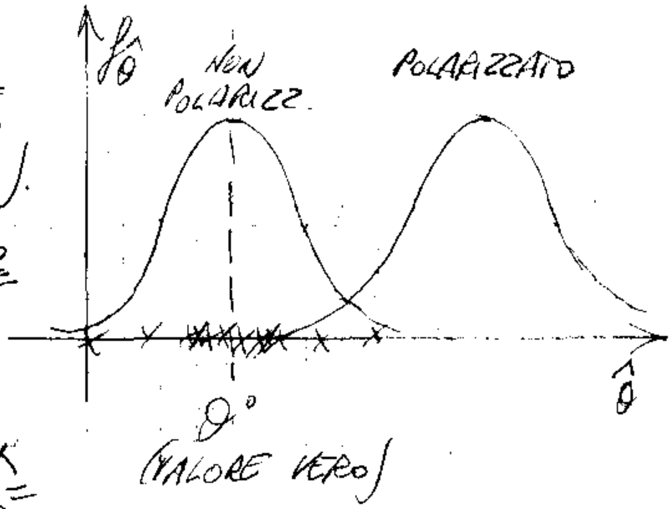
•) Proprietà di non polarizzazione

Lo stimatore $\hat{\vartheta}$ è detto non polarizzato (o corretto o non deviato o, in inglese unbiased) se:

$$E[\hat{\vartheta}] = \vartheta^o, \quad \text{con } \vartheta^o = \text{valore vero}$$

Gli stimatori sono v.c. cui può essere associata una d.d.p.

Si tratta di capire quale sia il baricentro della d.d.p. dello stimatore.



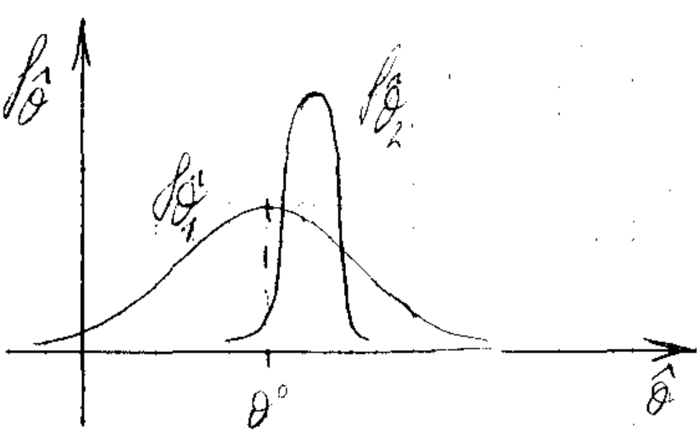
Per uno stimatore non polarizzato il baricentro della d.d.p. coincide con il valore vero. (schema di sinistra).

I valori forniti dallo stimatore contengono errori per difetto e per eccesso e la loro distribuzione è fornita dalla d.d.p.

I dati si distribuiscono attorno al valor ~~medio~~ vero.

Lo stimatore rappresentato dalla funzione a destra è tutto spostato e gli errori sono quasi tutti per eccesso.

La non polarizzazione è, in generale, una proprietà desiderabile.



Ci sono poi aspetti più importanti:

nello schema a sinistra $\hat{\vartheta}_2$ è polarizzato, ma migliore di $\hat{\vartheta}_1$, non polarizzato.

Pur non essendo polarizzato $\hat{\vartheta}_1$ può dare valori molto lontani dal valore vero ϑ^o.

•) Esempi di stimatori

-) Media campionaria: essendo $E[M_1] = \mu$, essa risulta non polarizzata

-) <u>Momento del secondo ordine</u>: non è polarizzato

$$E[M_2] = m_2$$

-) <u>Varianza campionaria</u>, <u>noto il valor medio</u>: non polarizzata

$$E[S_m^2] = \sigma^2$$

-) <u>S_2</u>: è polarizzato, essendo

$$E[S_2] = \frac{n-1}{n}\sigma^2$$

Asintoticamente diventa non polarizzato, poiché per $n \to \infty$

$$E[S_2] \to \sigma^2$$

Data la polarizzazione di S_2 è preferito il seguente stimatore

•) <u>Stimatore non polarizzato della varianza</u>

<u>Varianza campionaria</u> (<u>corretta nel senso di non polarizzata</u>).

$$S^2 := \frac{n}{n-1} S_2 = \frac{\cancel{n}}{n-1}\,\frac{1}{\cancel{n}}\sum_{i=1}^{n}(X_i - M_1)^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M_1)^2$$

$$E[S^2] = \frac{n}{n-1}E[S_2] = \frac{n}{n-1}\,\frac{n-1}{n}\sigma^2 = \sigma^2$$

Nota: non sempre è possibile, apportando uno stimatore polarizzato, eliminarne la polarizzazione, anche se ciò è vero in generale.

•) <u>Proprietà della consistenza</u>

$\hat\theta$ è consistente se, per $n \to \infty$, converge in probabilità a θ^0.

Come esempio, dimostriamo che i momenti campionari sono tutti consistenti.

Occorre dimostrare che M_k converge in probabilità a

$$m_k = E\left[X_i^k\right]$$

Definendo

$$Y_i := X_i^k$$

si nota che M_k è la media campionaria delle Y_i.
Infatti:

$$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Inoltre è:

$$E\left[X_i^k\right] = E\left[Y_i\right]$$

Applichiamo la legge dei Grandi Numeri alle Y_i, che assicura che la media campionaria delle Y_i converge alla media teorica.

La media campionaria di Y_i è rappresentata da M_k.
La media teorica degli Y_i sono i valori attesi degli X_i^k.

Ricordiamo che la legge dei Grandi Numeri garantisce la convergenza in media quadratica e, quindi, in probabilità.

Osservazione: apparentemente la legge dei Grandi Numeri riguarda solo la media campionaria, ma in realtà tutti i momenti campionari si possono vedere come delle medie campionarie.

Conseguenza: tutti i momenti campionari sono stimatori consistenti.

•) Teorema: anche i momenti ~~camp~~ centrali campionari sono consistenti.

La dimostrazione non viene fornita.

Nota: la consistenza è una proprietà desiderabile.

•) Proprietà: asintotica normalità

$\hat{\theta}$ si dice asintoticamente normale se, per $n \to \infty$ converge in distribuzione ad una v.c. Gaussiana.

Ad esempio, i momenti campionari sono tutti asintoticamente normali, come immediata conseguenza del teorema centrale del limite, poiché i momenti centrali sono medie di v.c. i.i.d.

Infatti qualsiasi momento centrale implica la somma di termini del tipo X_i^k.

Ridenominando gli X_i^2 con Y_i, si si accorge che si stanno eseguendo delle medie.

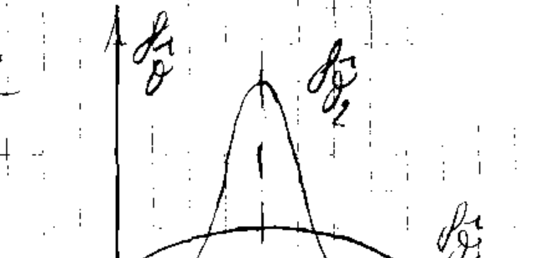
Applicando il teorema centrale del limite, per n grande, si ottiene una v.c. gaussiana.

La stessa conclusione può essere tratta per i momenti centrali campionari.

•) <u>Confronto tra stimatori</u> (con un numero finito di dati)

–) <u>Stimatori non polarizzati</u>

Cerchiamo dei criteri per valutare la bontà di uno stimatore non polarizzato quando si ha un numero finito di dati.



Nell'esempio mostrato a destra lo stimatore $\hat\theta_2$ è migliore avendo varianza minore e di conseguenza valori più concentrati attorno al valor vero.

Il criterio da seguire è la <u>varianza</u>: minore è la varianza, migliore è lo stimatore.

Nella ricerca dello stimatore migliore c'è un limite alla varianza che si può ottenere.

– Teorema: <u>disuguaglianza di Cramér-Rao</u>

Il teorema risale al 1945.

Sia $\hat\theta$ uno stimatore non polarizzato.

Allora, sotto condizioni di regolarità, vale la seguente disuguaglianza

• CARL HARALD CRAMÉR (Stoccolma 1893–1985) matematico e statistico svedese

• Calyampudi Radhakrishna Rao (n. Hadagali, India nel 1920) statistico indiano

$$\mathrm{Var}[\hat\vartheta] \geq \frac{1}{s}$$

con s = quantità d'informazione di Fisher essendo

$$s = -E\left[\frac{\partial^2}{\partial\vartheta^2} \ln f_X(\vartheta; X)\Big|_{\vartheta=\vartheta^\circ}\right]$$

-) Estensione al caso vettoriale (ϑ° e $\hat\vartheta$ vettori)

$$\mathrm{Var}[\hat\vartheta] \geq s^{-1}$$ con s = matrice di informazione di Fisher

$\mathrm{Var}[\hat\vartheta]$ = matrice (quadrata) varianze di $\hat\vartheta$.

-) Definizione: uno stimatore $\hat\vartheta_u$ non polarizzato si dice a minima varianza se la

$$\mathrm{Var}[\hat\vartheta_u] \leq \mathrm{Var}[\hat\vartheta],$$

con $\hat\vartheta$ è un qualsiasi stimatore non polarizzato.

Nota: se $\hat\vartheta$ raggiunge il limite di Cramér-Rao allora è uno stimatore a minima varianza.

Esempio (sulla media campionaria)

Calcolare la quantità di informazione di Fisher per il problema della stima di m in base ad

$$X_i \sim N(m, \sigma^2) \ \text{i.i.d.}$$

Risulta:

$$s = \frac{n}{\sigma^2}$$

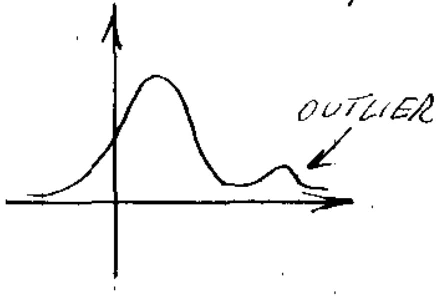
Ne consegue:

$$\mathrm{Var}[\hat\vartheta] \leq \frac{\sigma^2}{n}$$

Conseguenza: per v.c. $X_i \sim N(m, \sigma^2)$ i.i.d. la media campionaria m, che non è polarizzata e la cui varianza è σ^2/n è uno stimatore a minima varianza.

La varianza σ^2/n permette di raggiungere il limite di Cramér-Rao.

Osservazione: non è sempre vero che la media campionaria è la scelta migliore, per colpa del problema degli outlier.



Nella raccolta dei dati capita di avere valori anomali, errati o aberranti che fanno spostare di molto le stime.

Si può cercare di eliminare i dati non corretti, ammesso di poterli riconoscere.

Un altro metodo consiste nell'usare stimatori robusti che non risentono in modo pesante di eventuali outlier.

Usando, ad esempio, la mediana campionaria anziché la media campionaria si considera il dato centrale (dei dati ordinati in senso crescente).

Anche se la mediana si sposta, per colpa di qualche dato anomalo, essa si sposta poco perché, al centro, i dati sono molto fitti.

-) Stimatori non polarizzati.

Non ha senso cercare di ottenere solo varianza piccola.

Infatti considerando $\hat{\vartheta} = \nu$, cioè uno stimatore che restituisce sempre lo stesso valore, a prescindere dai dati sperimentali, si ha

$$Var[\hat{\vartheta}] = 0,$$

ma lo stimatore è pessimo.

Il vero scopo consiste nell'avere un errore

$$\hat{\vartheta} - \vartheta^\circ$$

"piccolo", cercando di rendere l'errore quadratico medio

$$E\left[(\hat{\vartheta} - \vartheta^\circ)^2\right]$$

il più piccolo possibile.

In inglese, l'errore quadratico medio si indica con MSE (Mean Square Error).

Proprietà di MSE

$$MSE = E\left[(\hat{\vartheta} - \vartheta^\circ)^2\right] = E\left[\left\{(\hat{\vartheta} - E[\hat{\vartheta}]) + (E[\hat{\vartheta}] - \vartheta^\circ)\right\}^2\right] =$$

$$= E\left[\left(\hat{\theta} - E[\hat{\theta}]\right)^2\right] + 2E\left[\left(\hat{\theta} - E[\hat{\theta}]\right)\left(E[\hat{\theta}] - \vartheta^\circ\right)\right] + E\left[\left(E[\hat{\theta}] - \vartheta^\circ\right)^2\right] = \quad (*)$$

$$= Var[\hat{\vartheta}] + 2\left(E[\hat{\vartheta}] - \vartheta^\circ\right) E\left[\hat{\vartheta} - E[\hat{\vartheta}]\right] + \left(E[\hat{\vartheta}] - \vartheta^\circ\right)^2 = \quad (**)$$

$$= Var[\hat{\vartheta}] + \left(E[\hat{\vartheta}] - \vartheta^\circ\right)^2$$

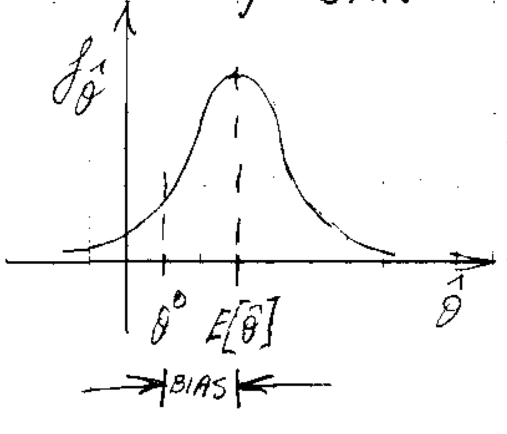
(*) Si osservi che:

- $E\left[\left(\hat{\theta} - E[\hat{\theta}]\right)^2\right]$ è la $Var[\hat{\theta}]$
- $E[\hat{\theta}] - \vartheta^\circ$ è una quantità deterministica ed esce dall'espressione del valor atteso.

(**) $E\left[\hat{\vartheta} - E[\hat{\vartheta}]\right] = E[\hat{\vartheta}] - E[\hat{\vartheta}] = 0$

Definiamo:

-) $MSE = E\left[\left(\hat{\theta} - \vartheta^\circ\right)^2\right]$
-) $BIAS = E[\hat{\vartheta}] - \vartheta^\circ$, differenza tra il baricentro dello stimatore e il valore vero della stima



Si ottiene:

$$MSE = Var[\hat{\vartheta}] + (BIAS)^2$$

Per un generico stimatore polarizzato o non polarizzato è importante avere un piccolo MSE, che risulta la somma di due contributi: la varianza dello stimatore e il BIAS elevato al quadrato.

Se lo stimatore non è polarizzato ($BIAS = 0$) occorre considerare la $Var[\hat{\theta}]$.

Se invece c'è il BIAS occorre cercare un compromesso tra BIAS e varianza.

Non è da scartare a priori uno stimatore polarizzato.

INTERVALLI DI CONFIDENZA

La stima di una serie di dati fornisce sempre una v.c., sia essa la media o la mediana campionaria.

Fornendo delle stime la correttezza impone di indicare il grado di affidabilità delle stime affinché si possa valutare l'attendibilità dei risultati forniti al cliente.

Esempio: stima della media m, usando la media campionaria M_1.

Una volta fornito M_1. Occorre dare un'idea dell'affidabilità della stima.

Per una data probabilità γ, si tratta di trovare un δ tale che

$$P(|M_1 - m| \le \delta) = \gamma$$

γ di solito è assunto pari a 0,95 (valore di affidabilità).

In altre parole, si va cercando un δ che garantisca che nel 95% dei casi l'errore sia $\le \delta$.

In tal caso, si userà il seguente intervallo di confidenza

$$I_\gamma = [M_1 - \delta, M_1 + \delta]$$

Nei sondaggi elettorali, δ è anche detto "forchetta".

Esempio: sia $\hat{\theta}$ uno stimatore gaussiano non polarizzato con varianza

$$\sigma^2_{\hat{\theta}}$$

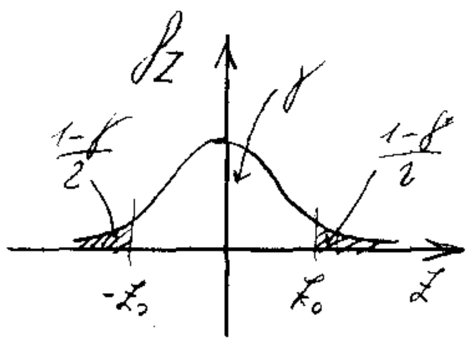
conosciuta. In altre parole

$$\hat{\theta} \sim N(\theta^0, \sigma^2_{\hat{\theta}})$$

Ci poniamo come obiettivo di determinare I_γ, per fissato.

Standardizziamo, ottenendo una gaussiana standard con

$$\frac{\hat{\theta} - \theta^0}{\sigma_{\hat{\theta}}}$$



Si cerca nella tabella della gaussiana standard il valore z_0, tale che

$$P(|Z| \le z_0) = \gamma$$

Ad esempio, se $\gamma = 0{,}95$, si ottiene

$$z_0 = 1{,}96$$

Inoltre,

$$P(|Z| \le z_0) = \gamma \Longrightarrow P\left(\left|\frac{\hat{\theta}-\theta^\circ}{\sigma_{\hat{\theta}}}\right| \le z_0\right) = \gamma \Longrightarrow$$

$$\Longrightarrow P\left(|\hat{\theta}-\theta^\circ| \le z_0 \sigma_{\hat{\theta}}\right) = \gamma$$

$z_0\sigma_{\hat{\theta}}$ è la quantità δ che si voleva determinare e rispetto alla quale il modulo dell'errore risulta minore.

Di conseguenza,

$$I_\gamma = [\hat{\theta}-\delta, \hat{\theta}+\delta] = [\hat{\theta} - z_0\sigma_{\hat{\theta}}, \hat{\theta} + z_0\sigma_{\hat{\theta}}]$$

Esempio di applicazione al caso della media campionaria.

Si considerino v.c. X_i i.i.d. con

$$\sigma^2 = Var[X_i]$$

supposta nota.

Ipotizziamo che n sia "grande" nel senso che sufficiente da poter applicare il teorema centrale del limite.

Determiniamo I_γ per la media campionaria, essendo

$$m = E[X_i]$$

Sfruttiamo il risultato precedente

$$\theta^\circ = m$$

$$\hat{\theta} = M_1$$

$$\sigma^2_{\hat{\theta}} = Var[\hat{\theta}] = Var[M_1] = \frac{\sigma^2}{n}$$

Risulta:

$$\delta = z_0 \sigma_{\bar\theta} = z_0 \frac{\sigma}{\sqrt{n}}$$

Ciò implica che

$$I_f = \left[M_1 - \frac{z_0 \sigma}{\sqrt{n}} , M_1 + \frac{z_0 \sigma}{\sqrt{n}} \right]$$

La quantità $\frac{\sigma}{\sqrt{n}}$ è spesso chiamata <u>standard error</u> (S.E.)

$$S.E. = \frac{\sigma}{\sqrt{n}}$$

Lo standard error è la deviazione standard per la media campionaria.

Nella stima della media campionaria si ha a che fare con due deviazioni standard: quella dei dati raccolti per l'analisi e quella della media campionaria, essendo questa una V.C.

σ è la deviazione standard dei dati

$\frac{\sigma}{\sqrt{n}}$ è la deviazione standard della media campionaria.