

Stimatori Inferenza I

Maura Mezzetti

Stimatori

- **Stima puntuale**: stima di un parametro sconosciuto della popolazione attraverso una statistica calcolata dall'osservazione campionaria
- **Intervallo di confidenza**: calcolo di un intervallo di valori a partire dalle osservazioni campionarie che includa il parametro incognito della popolazione con un grado di fiducia attribuito

Definizioni

- **Stimatori**: una variabile casuale funzione del campione $(X_1, X_2, \dots, X_n)$ la cui realizzazione è finalizzata al parametro θ
- **Stima**: valore numerico che lo stimatore assume in corrispondenza di un campione osservato

Stimatori

A partire dal campione definiamo lo *stimatore*

$$T = t(X_1, \dots, X_n)$$

T è una statistica campionaria. Essa viene utilizzata per stimare il parametro θ incognito della popolazione.

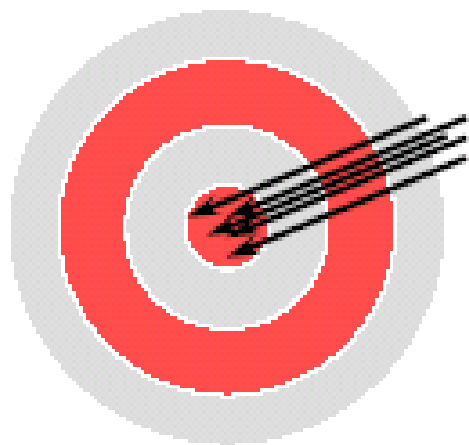
Si dice dunque che T è uno stimatore di θ . T è una v.c. e come tale è caratterizzata da una **distribuzione campionaria**.

In generale, definiamo stimatore di θ una v.c. funzione delle variabili casuali $X_1, X_2, \dots, X_n$, che generano i valori osservati del campione:

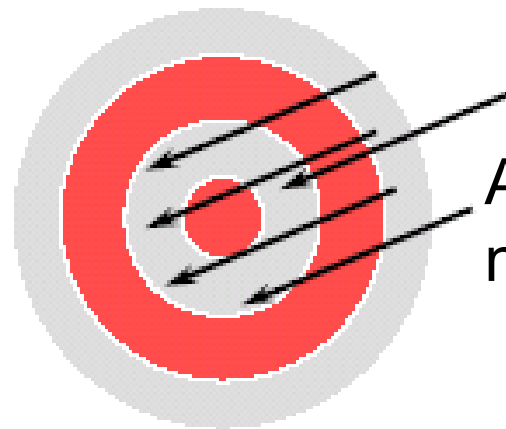
$$T = t(X_1, \dots, X_n)$$

Il valore assunto dallo stimatore T in un particolare campione estratto dalla popolazione $\{x_1, \dots, x_n\}$, fornisce la **stima puntuale** $t = t(x_1, \dots, x_n)$; t è un valore osservato, ovvero una realizzazione della v.c. T .

Accuratezza e Precisione



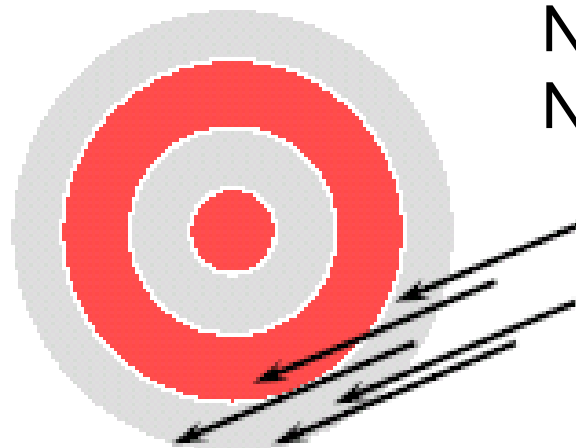
Accurato
and Preciso



Accurato
non Preciso



Preciso
non Accurato



Non Accurato
Non Preciso

Stimatore Corretto

T è uno stimatore **corretto** di θ se $E(T) = \theta$
si dice distorto altrimenti.

La distorsione (bias) è definita come

$$B(T) = E(T) - \theta$$

Esercizio

Sia $X \sim f(x; \theta)$, $E(X) = 2\theta$, (X_1, X_2, X_3) un campione casuale semplice, si consideri i seguenti stimatori di θ

$$T_1 = \frac{X_1 + X_2 - X_3}{2}$$

$$T_2 = \frac{X_1}{2}$$

$$T_3 = \bar{X}$$

Quali dei precedenti stimatori è corretto per θ ?

$$E(T_1) = E\left(\frac{X_1 + X_2 - X_3}{2}\right) = \frac{E(X_1) + E(X_2) - E(X_3)}{2} =$$

$$= \frac{2\theta + 2\theta - 2\theta}{2} = \theta$$

$$E(T_2) = E\left(\frac{X_1}{2}\right) = \frac{2\theta}{2} = \theta$$

$$E(T_3) = \left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{E(X_1) + E(X_2) + E(X_3)}{3} =$$

$$= \frac{2\theta + 2\theta + 2\theta}{3} = 2\theta$$

T_1 e T_2 stimatori corretti

T_3 DISTORTO con distorsione

$D(T_3) = E(T_3) - \theta = \theta$: **distorsione!**

Errore Quadratico Medio

di uno stimatore è definito come

$$EQM_T(\theta) = E(T - \theta)^2 \quad \theta \in \Theta$$

Formula ridotta:

$$EQM_T(\theta) = Var(T) + [E(T) - \theta]^2$$

$$\begin{aligned}
\text{MSE}(T) &= E[(T - \theta)^2] \\
&= E\{[T - E(T) + E(T) - \theta]^2\} \\
&= E\{[T - E(T)]^2\} + E\{[E(T) - \theta]^2\} \\
&\quad + 2E\{[T - E(T)][E(T) - \theta]\} \\
&= V(T) + \{[E(T) - \theta]^2\} \\
&\quad + 2[E(T) - \theta]E\{[T - E(T)]\} \\
&= V(T) + [B(T)]^2
\end{aligned}$$

Efficienza

T1 è più efficiente di T2 se e solo se

$$EQM(T1) \leq EQM(T2) \quad \forall \theta \in \Theta$$

Proprietà asintotiche: La Consistenza

Uno stimatore $T_n = t(X_1, X_2, \dots, X_n)$ definito su un campione casuale di ampiezza n è uno **stimatore consistente** del parametro θ se per

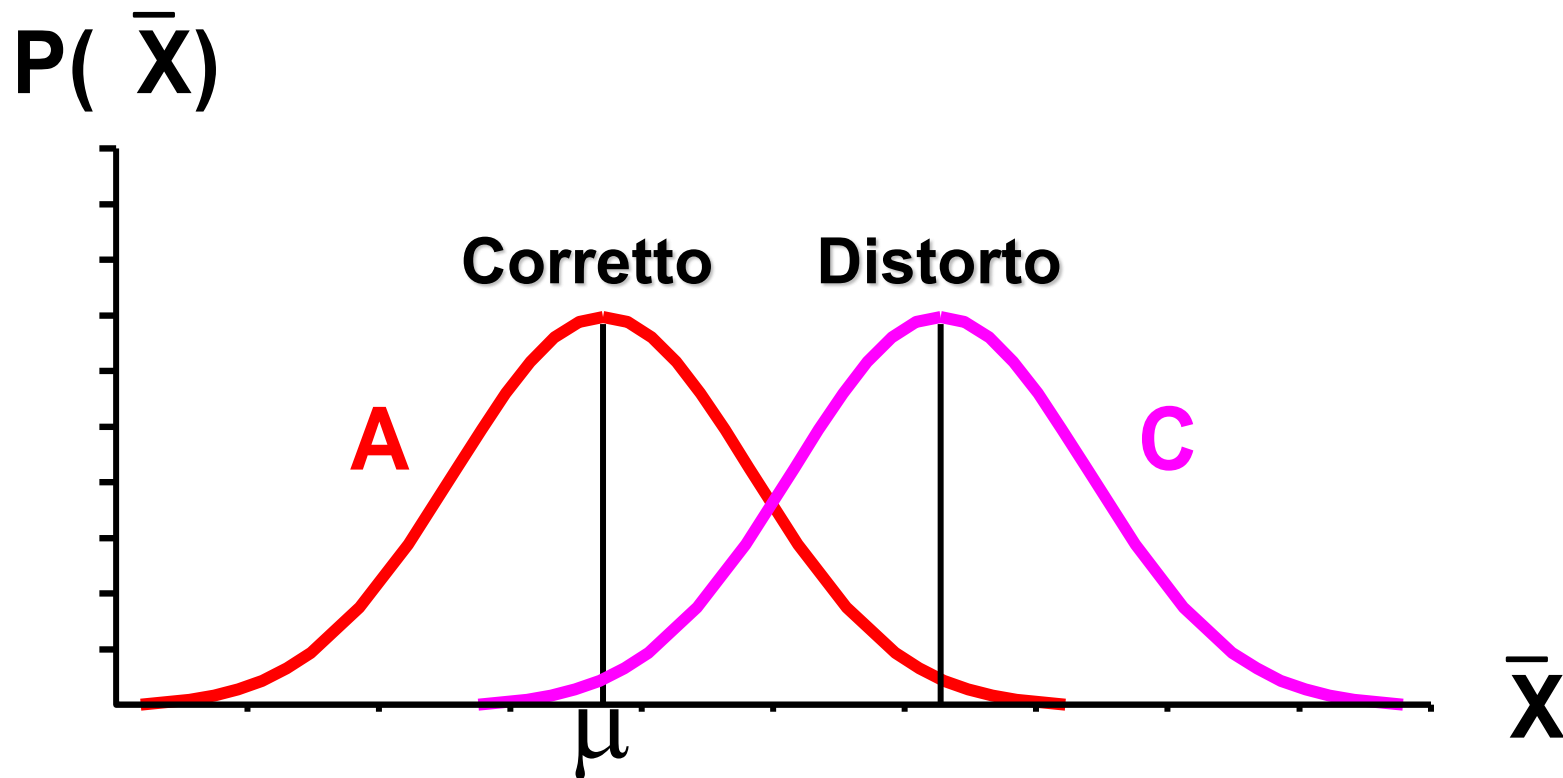
$$\lim_{n \rightarrow \infty} EQM_{T_n}(\theta) = 0 \quad \forall \theta \in \Theta$$

O equivalentemente se

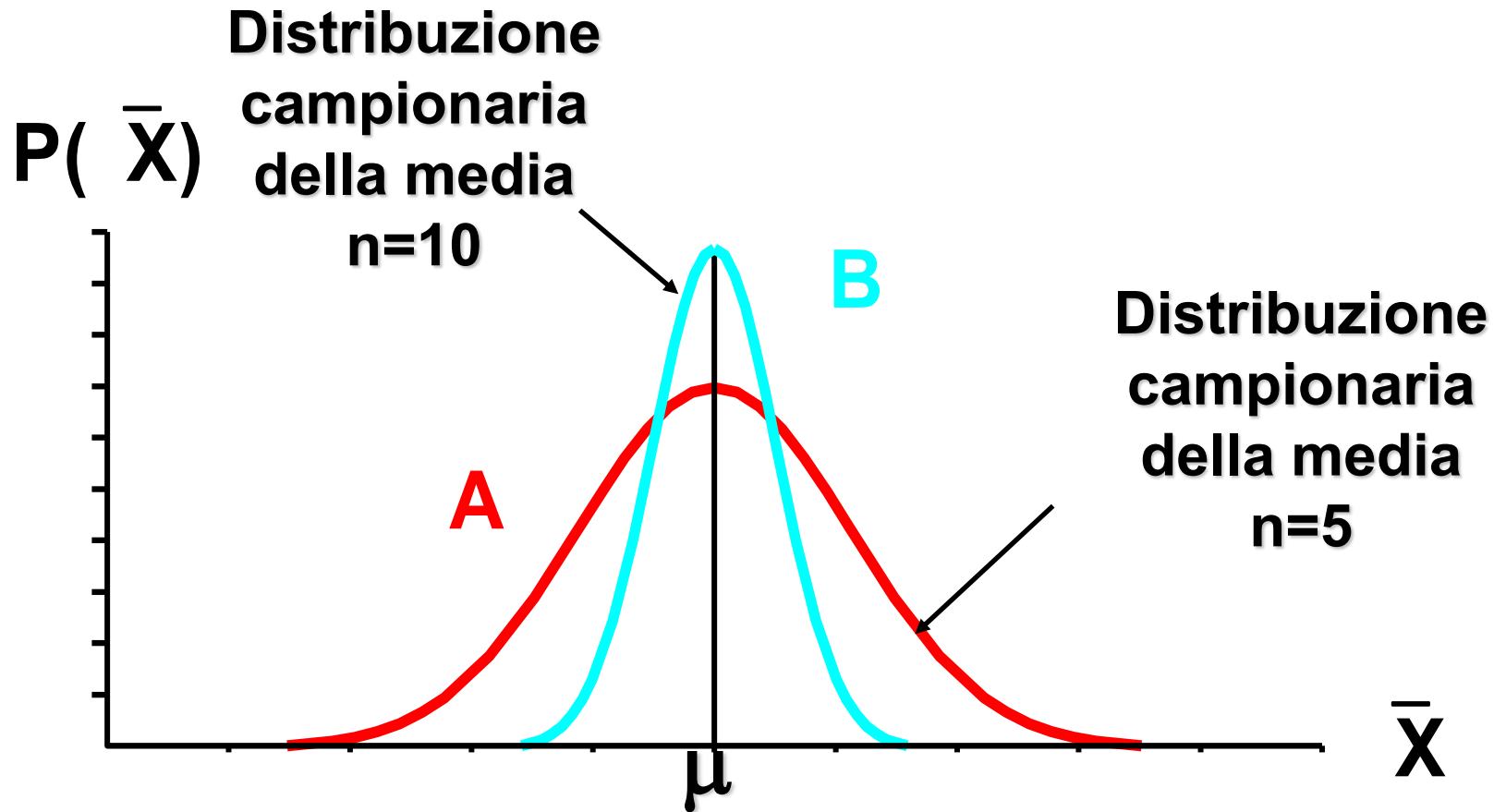
$$\lim_{n \rightarrow \infty} Bias(T_n) = 0 \quad \forall \theta \in \Theta$$

$$\lim_{n \rightarrow \infty} Var(T_n) = 0 \quad \forall \theta \in \Theta$$

Correttezza



Efficienza



La media campionaria è uno stimatore consistente per la media della popolazione

$$E(X)=\mu \quad \text{Var}(X)=\sigma^2$$

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{\text{Var}\left(\sum_{i=1}^n X_i\right)}{n^2} = \frac{\sum_{i=1}^n \text{Var}(X_i)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$
$$\lim_{n \rightarrow \infty} \text{Var}(\bar{X}_n) = 0$$

Poiché la media campionaria è uno stimatore corretto per la media della popolazione, la precedente condizione verifica che la media campionaria è uno stimatore consistente per la media della popolazione

La proporzione campionaria è stimatore
consistente della proporzione della
popolazione p

$X \sim \text{Bernoulli}(p)$

$E(X)=p \quad \text{Var}(X)=p(1-p)$

$$\text{Var}(\hat{p}) = \text{Var}\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{\text{Var}\left(\sum_{i=1}^n X_i\right)}{n^2} = \frac{\sum_{i=1}^n \text{Var}(X_i)}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

$$\lim_{n \rightarrow \infty} \text{Var}(\hat{p}_n) = 0$$

Esercizio

Sia X una v.a. con valore atteso μ e varianza σ^2 e sia (X_1, X_2, X_3, X_4) un campione casuale semplice distribuito come X . Il seguente stimatore è uno stimatore corretto per μ ? Valutarne inoltre l'errore quadratico medio.

$$T(X_1, X_2, X_3, X_4) = \frac{1}{2}X_1 - X_2 + X_3 + \frac{2}{3}X_4$$

Esercizio

Sia X una v.a. con valore atteso μ e varianza σ^2 e sia (X_1, X_2, X_3, X_4) un campione casuale semplice distribuito come X . Il seguente stimatore è uno stimatore corretto per μ ? Valutarne inoltre l'errore quadratico medio.

$$T(X_1, X_2, X_3, X_4) = \frac{1}{2}X_1 - X_2 + X_3 + \frac{2}{3}X_4$$

$$E(T) = \frac{1}{2}\mu - \mu + \mu + \frac{2}{3}\mu = \frac{7}{6}\mu \quad = \text{STIMATORE DISTORTO PER } \mu \Rightarrow \quad B(T) = \frac{7}{6}\mu - \mu = \frac{1}{6}\mu$$

$$EQM(T) = V(T) + B(T)^2 = \frac{97}{36}\sigma^2 + \frac{1}{36}\mu^2$$

Esercizio

Siano X_1 e X_2 due variabili aleatorie indipendenti estratte da una popolazione con $E(X) = \mu$ e $V(X) = \sigma^2$. Dimostrare che lo stimatore T

$$T = \frac{(X_1 - X_2)^2}{2}$$

è uno stimatore NON distorto della varianza della popolazione.

Esercizio

Occorre verificare che $E(T) = \sigma^2$

Sapendo che $\sigma^2 = Var(X) = E(X^2) - E(X)^2$

$$\rightarrow E(X^2) = \sigma^2 + \mu^2$$

Inoltre, poiché le v.a. sono indipendenti,

Da cui, $Y = X_1 - X_2$ $E(Y) = 0$ $Var(Y) = 2\sigma^2$

$$E(Y^2) = Var(Y) + E(Y)^2 = 2\sigma^2$$

Pertanto T è uno stimatore NON distorto della varianza della popolazione

Tuttavia, μ è incognito e viene stimato mediante $\bar{X}$. Si consideri dunque lo stimatore:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

Si dimostra che $\hat{\sigma}^2$ è distorto, ma asintoticamente corretto, ed in particolare:

$$E(\hat{\sigma}^2) = \frac{n-1}{n} \sigma^2 = \sigma^2 - \frac{1}{n} \sigma^2,$$

$$B(\hat{\sigma}^2) = -\frac{1}{n} \sigma^2 \rightarrow_{n \rightarrow \infty} 0$$

Dimostrazione

$$E(\hat{\sigma}^2) = \frac{1}{n} E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right]$$

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

$$= E \left\{ \sum_{i=1}^n [(X_i - \mu) - (\bar{X} - \mu)]^2 \right\}$$

$$= E \left\{ \sum_{i=1}^n [(X_i - \mu)^2 + (\bar{X} - \mu)^2 - 2(X_i - \mu)(\bar{X} - \mu)] \right\}$$

$$= E \left\{ \left[\sum_{i=1}^n (X_i - \mu)^2 \right] + n(\bar{X} - \mu)^2 - 2n(\bar{X} - \mu)^2 \right\}$$

$$= E \left\{ \left[\sum_{i=1}^n (X_i - \mu)^2 \right] - n(\bar{X} - \mu)^2 \right\}$$

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

$$= (n-1)\sigma^2$$

Pertanto,

$$\mathbb{E}(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$$

Stimatore corretto della varianza (varianza campionaria corretta):

$$S^2 = \frac{n}{n-1}\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Stima: campione osservato $\{x_1, \dots, x_n\}$.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Esempio

Un'impresa desidera stimare il tempo medio in giorni lavorativi che occorre per evadere un ordine. Per questo ha estratto un campione casuale di 12 ordini

$$\{x_1, \dots, x_{12}\} = \{20, 10, 15, 8, 10, 12, 9, 12, 14, 8, 13, 16\}$$

da cui risulta $\bar{x} = 12.25$ e $s^2 = 12.93$.

La stima della varianza della popolazione ci consente di completare la specificazione della distribuzione campionaria di $\bar{X}$, che dipende da σ^2 . Si ha infatti $V(\bar{X}) = \sigma^2/n$.

La stima della varianza dello stimatore $\bar{X}$ risulta dunque:

$$\hat{V}(\bar{X}) = \frac{s^2}{n}$$

e fornisce una misura (inversa) della precisione dello stimatore della media.

Esercizio

Sia X una variabile aleatoria tale che $E(X)=\theta$ e varianza 1, Considerare i due stimatori di θ ,

$$T_1 = \frac{X_1 + 2X_2 - X_3}{2}$$

$$T_2 = \frac{X_1 + X_2 - X_3}{3}$$

quale dei due stimatori è più efficiente?

$$E(T_1) = \frac{E(X_1) + 2E(X_2) - E(X_3)}{2} = \theta$$

$$E(T_2) = \frac{E(X_1) + E(X_2) - E(X_3)}{3} = \frac{\theta}{3}$$

$$Var(T_1) = \frac{Var(X_1) + 4Var(X_2) + Var(X_3)}{4} = \frac{3}{2}$$

$$Var(T_2) = \frac{Var(X_1) + Var(X_2) + Var(X_3)}{9} = \frac{1}{3}$$

$$MSE(T_1) = \frac{3}{2} \quad MSE(T_2) = \frac{1}{3} + \frac{4}{9} \theta^2$$

Example

If $X_1, \dots, X_3$ are i.i.d. as X , with expected value $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$, consider the following estimators of μ

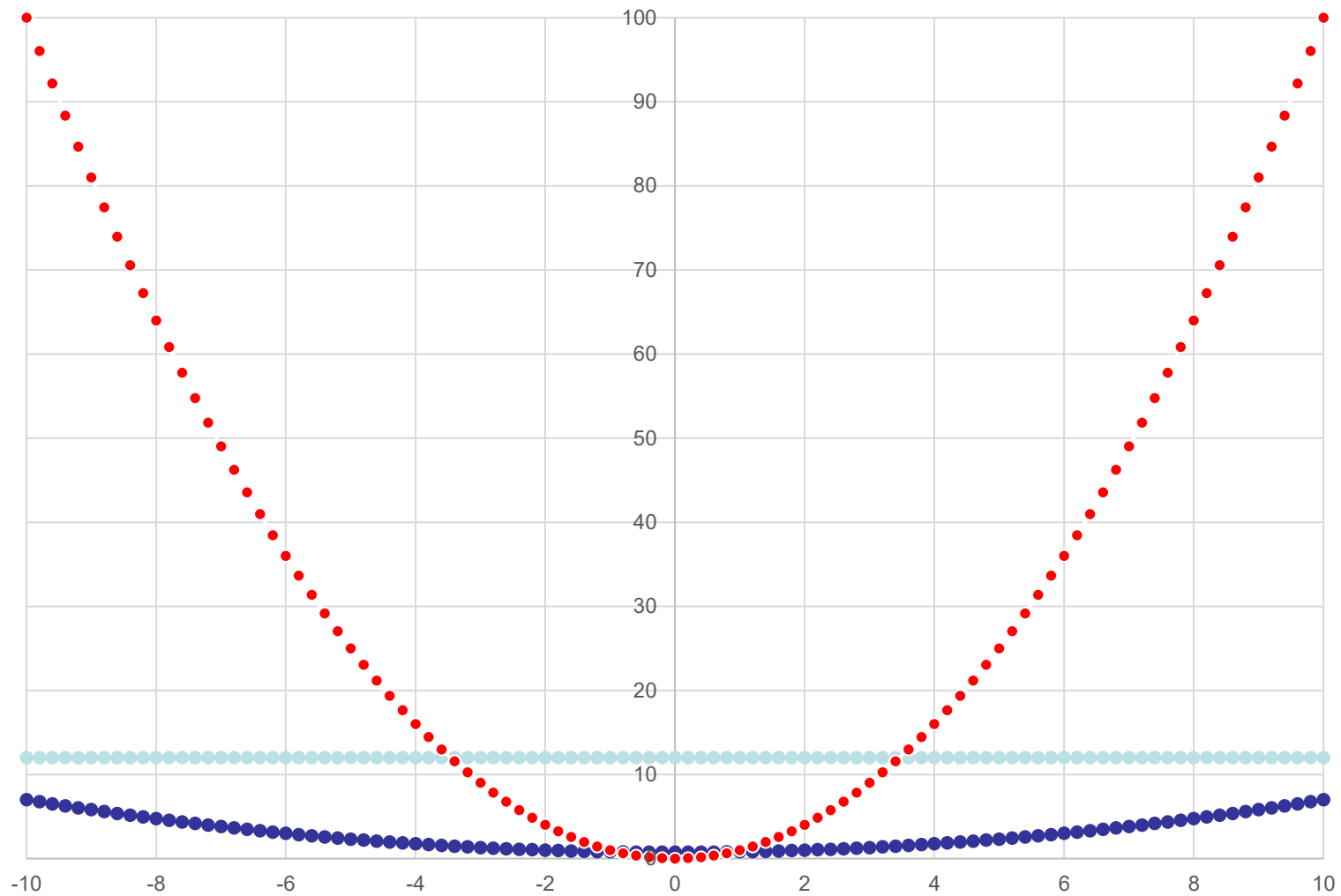
$$\begin{aligned}T_1 &= X_1 + X_2 - X_3 \\T_2 &= \frac{X_1 + X_2 + X_3}{4} \\T_3 &= 0\end{aligned}$$

Which estimator would you prefer?

$$\begin{aligned}E(T_1) &= E(X_1 + X_2 - X_3) = \mu & \text{Var}(T_1) &= \text{Var}(X_1 + X_2 - X_3) = 3\sigma^2 \\E(T_2) &= \frac{E(X_1 + X_2 + X_3)}{4} = \frac{3}{4}\mu & \text{Var}(T_2) &= \frac{\text{Var}(X_1 + X_2 + X_3)}{16} = \frac{3}{16}\sigma^2 \\E(T_3) &= 0 & \text{Var}(T_3) &= 0\end{aligned}$$

T_1 is the only unbiased estimator, T_3 is biased but it is the one with the lowest variance, let's compare the MSE.

$$\begin{aligned}MSE_{\mu}(T_1) &= 3 \times \sigma^2 \\MSE_{\mu}(T_2) &= \frac{3}{16}\sigma^2 + \frac{\mu^2}{16} \\MSE_{\mu}(T_3) &= \mu^2\end{aligned}$$



—●— MSE1 —●— MSE2 —●— MSE3

Esercizio

Siano (X_1, X_2, X_3, X_4) due variabili aleatorie indipendenti estratte da una popolazione con $E(X) = \mu$ e $V(X) = \sigma^2$. Dimostrare che lo stimatore T

$$T = \frac{(X_1 + X_2)^2}{2} - 2X_3X_4$$

è uno stimatore NON distorto della varianza della popolazione.

Esercizio

Siano X_1 e X_2 due variabili aleatorie indipendenti estratte da una popolazione con $E(X) = \mu$ e $V(X) = \sigma^2$.

Lo stimatore T

$$T = \frac{(X_1 + \bar{X}_n)}{2}$$

è uno stimatore distorto della media della popolazione? E' uno stimatore consistente?

Metodo dei momenti

Siano $(X_1, X_2, \dots, X_n)$ n variabili aleatorie indipendenti estratte da una popolazione X con probabilità $p(x; \theta)$ o densità $f(x; \theta)$.

La stima con il metodo dei momenti si ottiene risolvendo il sistema di equazioni ottenuto eguagliando i primi d momenti della variabile X ai primi d momenti campionari

Metodo dei momenti: Gaussiana

Siano $(X_1, X_2, \dots, X_n)$ n variabili aleatorie indipendenti estratte da una popolazione X Gaussiana con $E(X) = \mu$ e $V(X) = \sigma^2$

$$\hat{\mu} = \bar{y}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$$

Metodo dei momenti: Binomiale

Siano $(X_1, X_2, \dots, X_n)$ n variabili aleatorie indipendenti estratte da una popolazione X Bernoulliana con $E(X) = \pi$

$$\hat{\pi} = \bar{x}$$

Metodo dei momenti: Uniforme

Siano $(X_1, X_2, \dots, X_n)$ n variabili aleatorie indipendenti estratte da una popolazione X Uniforme(0, θ)

$$\hat{\theta} = 2\bar{x}$$

Esempio

Lancio del dado FINO a quando si osserva il numero 6:
X "numero di tentativi prima di osservare il numero 6"

Quali possibili valori può assumere X?

Con quale probabilità $X=0$? $X=1$? $X=2$?

$$P(X = x) = \left(\frac{5}{6}\right)^x \frac{1}{6} \quad x = 0, 1, 2, \dots$$

Distribuzione Geometrica

$$P(X = x) = p(1 - p)^x \quad x = 0, 1, 2, \dots$$

Metodo dei momenti: Geometrica

Siano $(X_1, X_2, \dots, X_n)$ n variabili aleatorie indipendenti estratte da una popolazione X Geometrica di parametro p

$$E(X) = \frac{1-p}{p}$$

$$\hat{p} = \frac{1}{1 + \bar{x}}$$

Metodo dei massima verosimiglianza

Il metodo della massima verosimiglianza fornisce una metodologia assai generale per la costruzione di stimatori.

Esso richiede la specificazione della distribuzione di probabilità o di densità della popolazione.

Metodo dei massima verosimiglianza: Esempio

Immaginiamo di avere 100 biscotti nel mio zaino. I biscotti sono di due tipi: biscotti al cioccolato o biscotti della fortuna. Inoltre sappiamo che il numero di biscotti della fortuna sono 10 oppure 90.



Estraiamo a caso un biscotto dallo zaino ed è un biscotto della fortuna.

Metodo dei massima verosimiglianza: Esempio

Estraiamo a caso un biscotto dallo zaino ed è un biscotto della fortuna.

Quale delle due risposte è più verosimile:



10 biscotti della fortuna e 90 biscotti di cioccolato

90 biscotti della fortuna e 10 biscotti di cioccolato

Metodo dei massima verosimiglianza

Siano $(X_1, X_2, \dots, X_n)$ n variabili aleatorie indipendenti estratte da una popolazione X con probabilità $p(x; \theta)$ o densità $f(x; \theta)$.

Definizione: la funzione di verosimiglianza rappresenta la probabilità di osservare un dato campione al variare del valore del parametro θ .

$$L(\theta) = P(X = x_1; \theta)P(X = x_2; \theta) \times \dots \times P(X = x_n; \theta)$$

Oppure

$$L(\theta) = f(x_1; \theta)f(x_2; \theta) \times \dots \times f(x_n; \theta)$$

Metodo dei massima verosimiglianza

Definizione: la funzione di verosimiglianza rappresenta la probabilità di osservare un dato campione al variare del valore del parametro θ .

$$L(\theta) = P(X = x_1; \theta)P(X = x_2; \theta) \times \cdots \times P(X = x_n; \theta)$$

Oppure

$$L(\theta) = f(x_1; \theta)f(x_2; \theta) \times \cdots \times f(x_n; \theta)$$

La **stima di massima verosimiglianza** di θ il valore $\hat{\theta}$ tale che il valore della verosimiglianza è massimo

$$\hat{\theta} = \arg \max L(\theta)$$

Esempio: popolazione Bernoulliana di parametro π

Siamo interessati a stimare π (es. frequenza relativa dei clienti soddisfatti).

Viene estratto un campione casuale di $n = 20$ unità.

$$X_i \sim \text{IID Bernoulli}(\pi), P(X_i = x_i) = \pi^{x_i}(1 - \pi)^{1-x_i}, x_i = 0, 1$$

Si supponga che il campione osservato presenti $\sum_{i=1}^n x_i = 15$ successi (valori $x_i = 1$) e 5 insuccessi (valori $x_i = 0$).

$$L(\pi) = \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n \pi^{x_i}(1 - \pi)^{1-x_i} = \pi^{\sum_i x_i} (1 - \pi)^{n - \sum_i x_i}$$

Pertanto, nel nostro caso,

$$L(\pi) = \pi^{15}(1 - \pi)^5, \quad \log L(\pi) = 15 \log \pi + 5 \log(1 - \pi)$$

La *stima* di massima verosimiglianza del parametro π è:

$$\hat{\pi} = \max_{\pi} L(\pi) = \max_{\pi} \left[\pi^{\sum_i x_i} (1 - \pi)^{n - \sum_i x_i} \right], \quad \pi \in [0, 1]$$

ovvero

$$\hat{\pi} = \max_{\pi} [\log L(\pi)] = \max_{\pi} \left\{ \sum_i x_i \log(\pi) + \left(n - \sum_i x_i \right) \log(1 - \pi) \right\}.$$

La soluzione può essere ottenuta differenziando e ponendo pari a zero la derivata prima.

$$\frac{d}{d\pi} \log L(\pi) = \frac{\sum_i x_i}{\pi} - \frac{n - \sum_i x_i}{1 - \pi} = 0$$

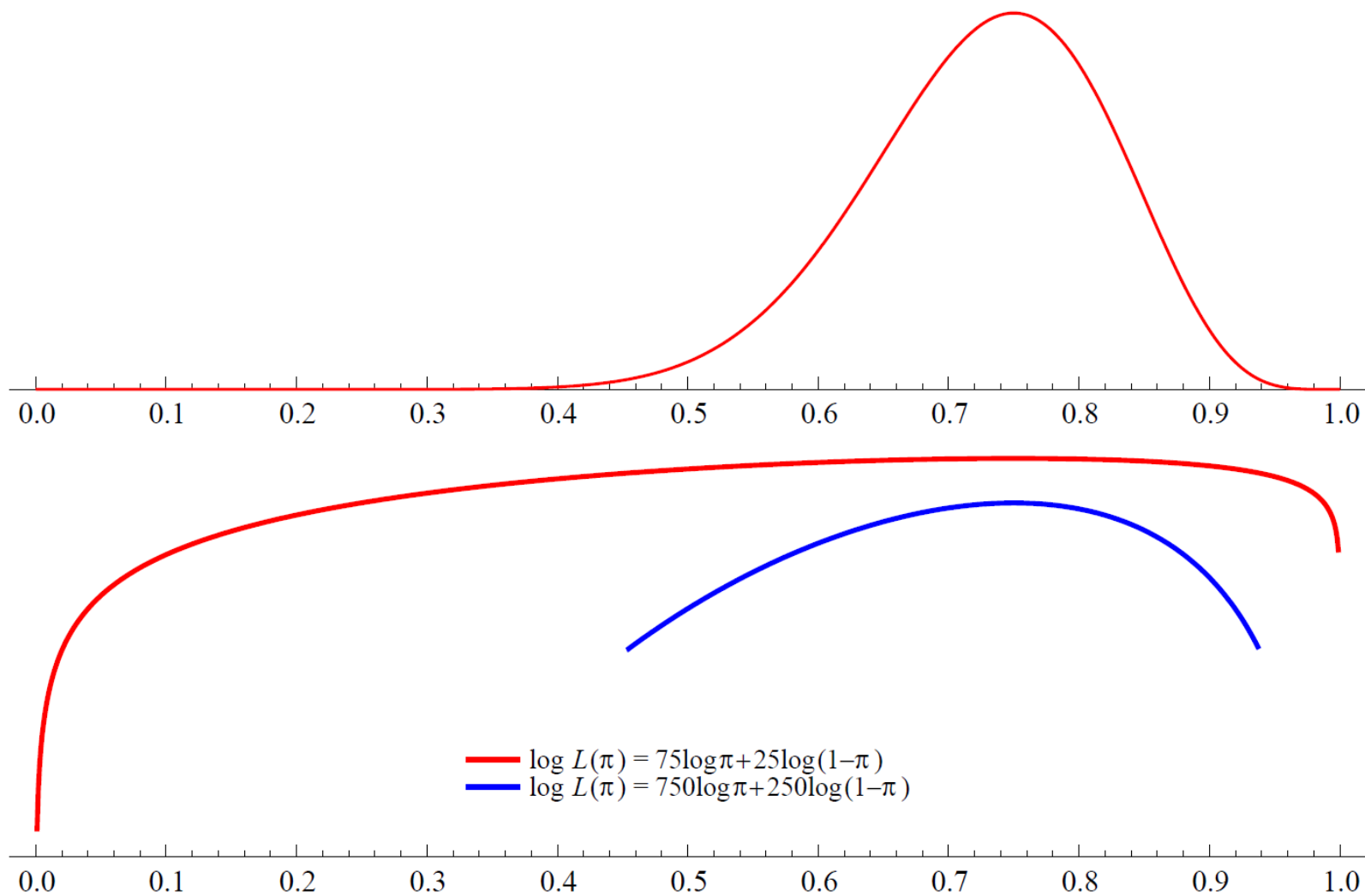
da cui segue

$$\hat{\pi} = \bar{X}$$

Questo dimostra che $\bar{X} = \sum X_i/n$ è lo *stimatore* di massima verosimiglianza di π .

Nell'esempio precedente $\hat{\pi} = 15/20 = 0.75$.

$$L(\pi) = \pi^{15}(1-\pi)^5$$



Metodo di Massima Verosimiglianza

σ^2 noto

$$f(y_i | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y_i - \mu)^2}$$

$$f(y_1, \dots, y_i, \dots, y_n | \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y_i - \mu)^2}$$

$$f(y_1, \dots, y_i, \dots, y_n | \mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2}$$

Metodo di Massima Verosimiglianza

σ^2 noto

$$L(\mu, \sigma^2 | y_i) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2}$$

$$l(\mu, \sigma^2) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2$$

$$\frac{dl(\mu, \sigma^2)}{d\mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \mu)$$

Metodo di Massima Verosimiglianza

σ^2 noto

$$0 = \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \mu)$$

$$\hat{\mu} = \bar{y}$$

$$\frac{d^2 l(\mu, \sigma^2)}{d\mu^2} = -\frac{n}{\sigma^2} < 0$$

Metodo di Massima Verosimiglianza Geometrica

$$p(x_i; \pi) = \pi(1 - \pi)^{x_i}$$

$$p(x_1, \dots, x_i, \dots, x_n | \pi) = \prod_{i=1}^n \pi(1 - \pi)^{x_i}$$

$$L(\pi) = \pi^n (1 - \pi)^{\sum x_i}$$

$$\log L(\pi) = n \log(\pi) + \sum x_i \log(1 - \pi)$$

$$\frac{d \log L(\pi)}{d\pi} = \frac{n}{\pi} - \frac{\sum x_i}{1 - \pi}$$

$$\frac{d^2 \log L(\pi)}{d\pi^2} = -\frac{n}{\pi^2} - \frac{\sum x_i}{(1 - \pi)^2} < 0$$

$$\frac{d \log L(\pi)}{d\pi} = 0$$

$$\hat{\pi} = \frac{1}{1 + \bar{x}}$$

Proprietà degli stimatori di massima verosimiglianza

Sotto opportune condizioni di regolarità, gli stimatori di massima verosimiglianza

- ▶ sono invarianti rispetto a trasformazioni (ad es. la stima di MV di σ è semplicemente $\sqrt{\hat{\sigma}^2}$);
- ▶ sono consistenti e asintoticamente corretti;
- ▶ la distribuzione asintotica è normale;
- ▶ sono asintoticamente efficienti.