UNIVERSITÀ
DEGLI STUDI DI TRIESTE

Dipartimento di scienze economiche,
aziendali, matematiche e statistiche
“Bruno de Finetti”

Statistica

Inferenza: alcuni approfondimenti

Francesco Pauli

A.A. 2016/2017

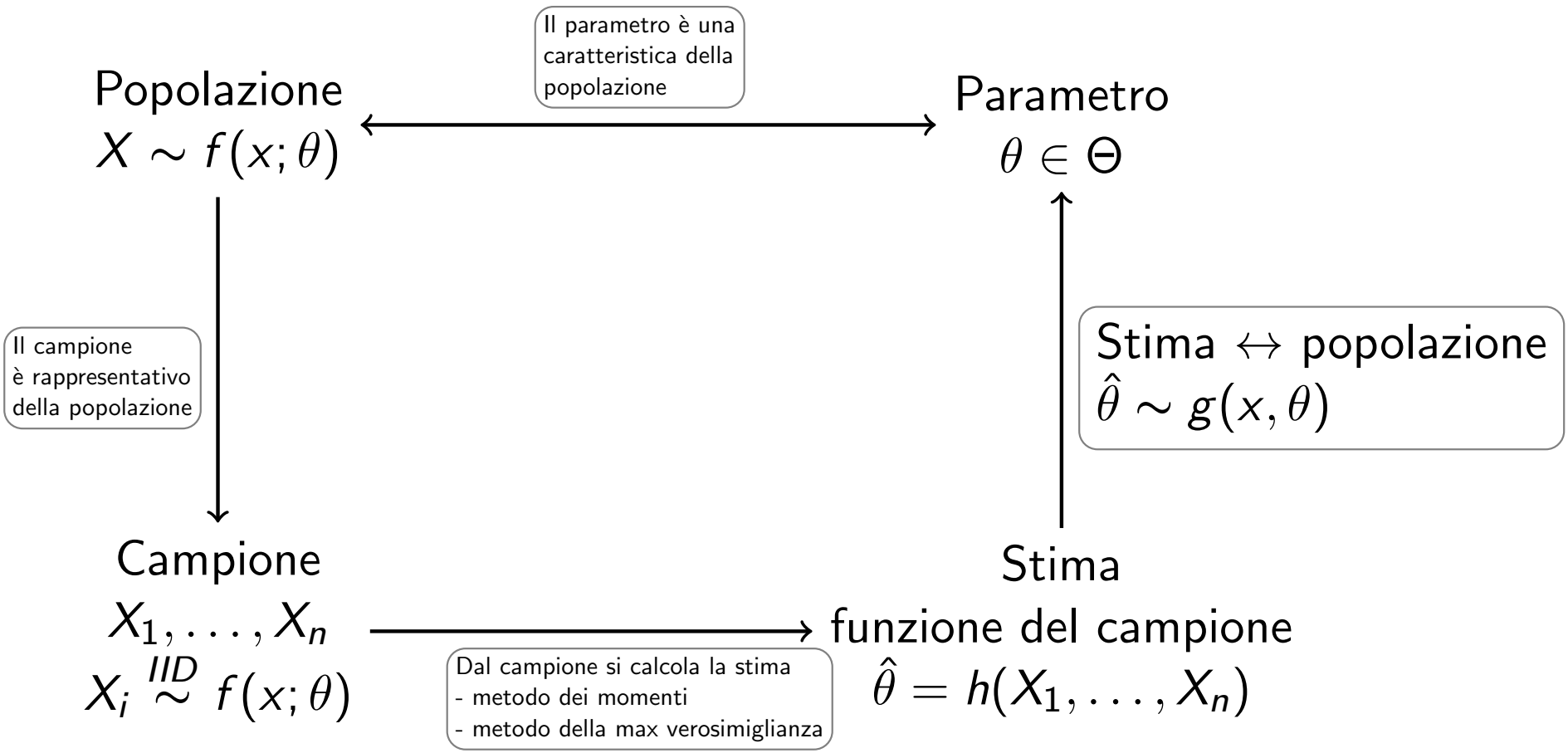
Indice

Distribuzione campionaria e precisione di uno stimatore

Ottenere uno stimatore

Proprietà di uno stimatore

Popolazione-campione-parametro-stima



Campione

Assumiamo di avere un campione di osservazioni indipendenti e aventi tutte la distribuzione della popolazione (di interesse).
In simboli, abbiamo un campione $X_1, \ldots, X_n$ indipendente e identicamente distribuito secondo $f(x;\theta)$ dove f è la distribuzione nella popolazione e θ è il parametro (o i parametri).

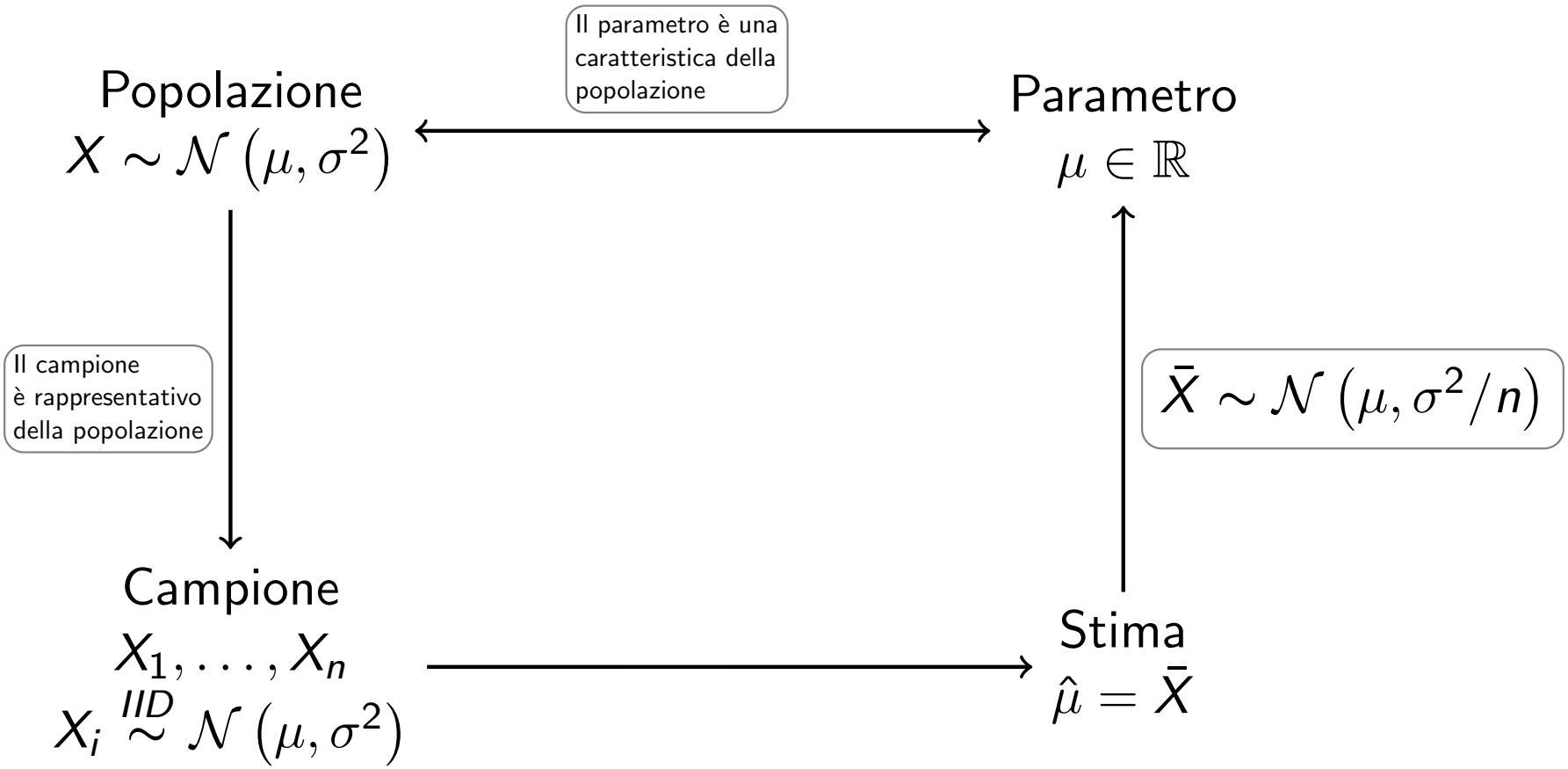
- Nell'esempio dei bianchi/neri osserviamo $X_1, \ldots, X_n$ indipendenti dove X_i vale 1 se l'i-esima osservazione è nera e 0 altrimenti e $P(X_i = 1) = \pi$.
- Nell'esempio del controllo di processo osserviamo $X_1, \ldots, X_n$ indipendenti dove $X_i \sim \mathcal{N}(\mu, \sigma^2)$.

Si dice **spazio campionario** l'insieme dei possibili campioni

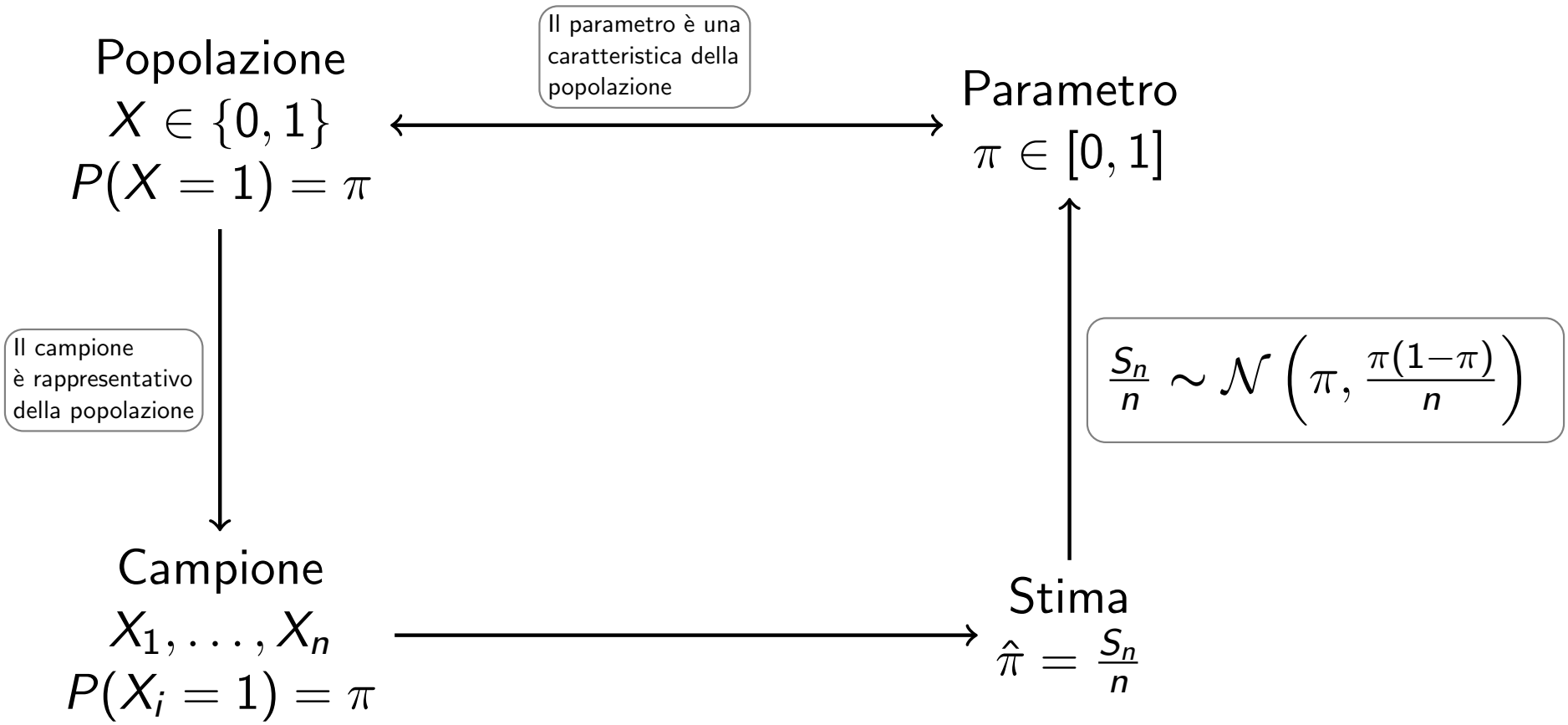
- Nell'esempio dei bianchi/neri lo spazio campionario è l'insieme delle sequenze di 0 e 1 di lunghezza n.
- Nell'esempio del controllo di processo lo spazio campionario è $\mathbb{R}^n$.

Le ipotesi fatte consentono di ottenere facilmente la distribuzione di probabilità del campione.

Popolazione-campione-parametro-stima: normale



Popolazione-campione-parametro-stima: proporzione



Distribuzione di probabilità del campione

Se ciascuna X_i ha distribuzione $f(x_i;\theta)$ ed essendo le X_i indipendenti, la distribuzione congiunta è

$$f(x_1,\ldots,x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

Nel caso dei bianchi/neri, tenendo presente che

$$X_i \in \{0,1\},\ P(X_i = 1) = \pi$$

si ha

$$P(X_1 = x_1 \cap \ldots \cap X_n = x_n) = \pi^{\sum_{i=1}^{n} x_i}(1-\pi)^{n-\sum_{i=1}^{n} x_i}$$

(cfr i ragionamenti fatti nel definire la distribuzione binomiale).

Distribuzione di probabilità del campione

Se ciascuna X_i ha distribuzione $f(x_i;\theta)$ ed essendo le X_i indipendenti, la distribuzione congiunta è

$$f(x_1,\ldots,x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

Nel caso della normale si ha

$$f(x_i;\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

e quindi

$$\begin{aligned} f(x_1,\ldots,x_n;\theta) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \end{aligned}$$

Un po' di definizioni

Una qualunque funzione del campione si chiama **statistica**.

- Proporzione campionaria e media campionaria nei due esempi precedenti sono statistiche.

Una statistica usata per stimare un parametro è detta **stimatore**, se il parametro è θ si indica spesso lo stimatore con $\hat{\theta}$.

- La proporzione campionaria nell'esempio dei b/n è uno stimatore per la proporzione di neri nella popolazione.
- La media campionaria nell'esempio industriale è uno stimatore per la media della popolazione.

La distribuzione di probabilità della statistica è detta **distribuzione campionaria**.

- La proporzione campionaria ha, approssimativamente, distribuzione campionaria normale.
- La media campionaria di un campione normale ha distribuzione campionaria normale.

Indice

Distribuzione campionaria e precisione di uno stimatore

Ottenere uno stimatore

Proprietà di uno stimatore

Ricavare uno stimatore: metodo dei momenti

Negli esempi precedenti (bianchi/neri e controllo di processo) abbiamo proposto degli stimatori in modo naturale

- la proporzione campionaria $\hat{\pi}$ come stimatore dell'analoga proporzione nella popolazione π;
- la media campionaria $\hat{\mu} = \bar{X}$ come stimatore della media della popolazione μ.

Questo modo di procedere "per analogia" è di carattere generale e prevede di eguagliare un momento campionario (media o varianza tipicamente) al corrispondente momento della popolazione.

Momento è un concetto più generale, per quello che abbiamo visto momenti sono la media, la media dei quadrati, la varianza.

Metodo dei momenti: uniforme

Nel caso della distribuzione uniforme (cfr esempio sulla statistica in guerra, anche se lì è discreta), se si assume che la popolazione X sia distribuita secondo un'uniforme tra 0 e θ, sicché

$$E(X) = \theta/2$$

si eguaglia allora la media teorica a quella campionaria

$$\bar{X} = E(X)$$

e uno stimatore per θ viene da

$$\bar{X} = E(X) = \frac{\theta}{2} \quad \Rightarrow \hat{\theta} = 2\bar{X}$$

Metodo dei momenti: esponenziale

Nel caso di una popolazione X con distribuzione esponenziale di parametro λ si ha

$$E(X) = 1/\lambda,$$

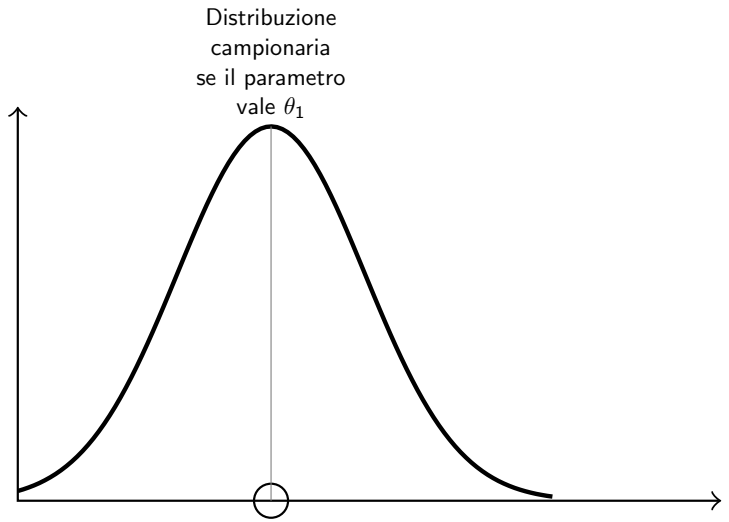
sichhé uno stimatore per λ viene da

$$\bar{X} = E(X) = \frac{1}{\lambda} \quad \Rightarrow \hat{\lambda} = \frac{1}{\bar{X}}$$

Nel caso della distribuzione di Poisson ...

Metodo della verosimiglianza

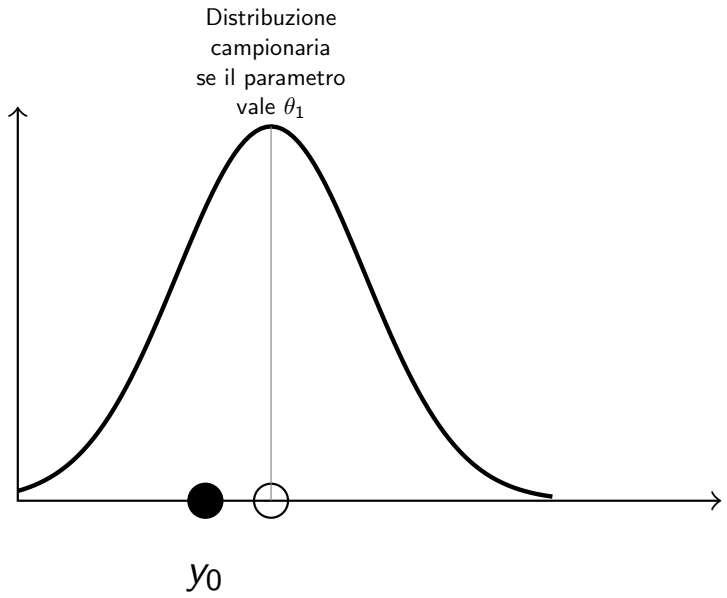
Partiamo della distribuzione campionaria per un dato valore del parametro, θ_1.



Metodo della verosimiglianza

Partiamo della distribuzione campionaria per un dato valore del parametro, θ_1.

Consideriamo il campione, y_0.

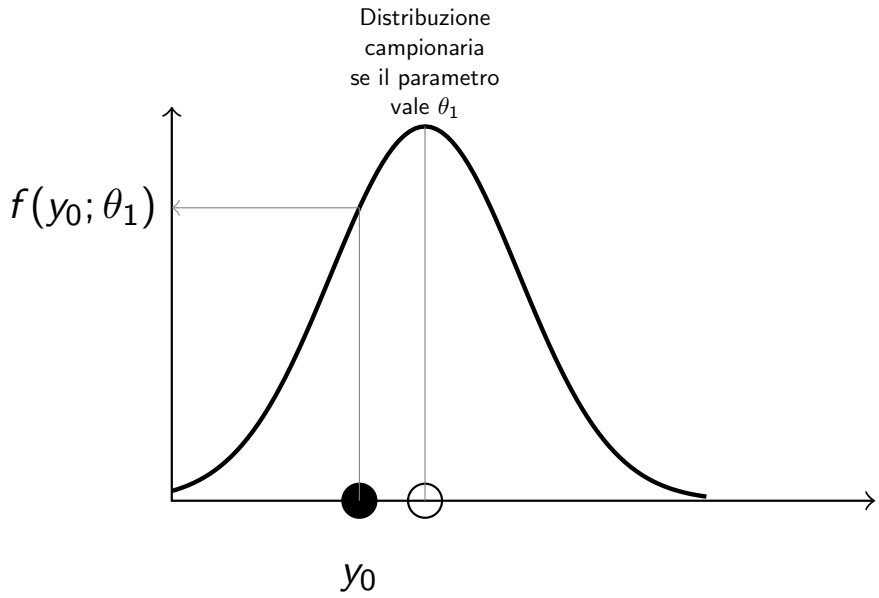


Metodo della verosimiglianza

Partiamo della distribuzione campionaria per un dato valore del parametro, θ_1.

Consideriamo il campione, y_0.

Possiamo calcolare la probabilità (o densità) del campione corrispondente al valore θ_1 del parametro.

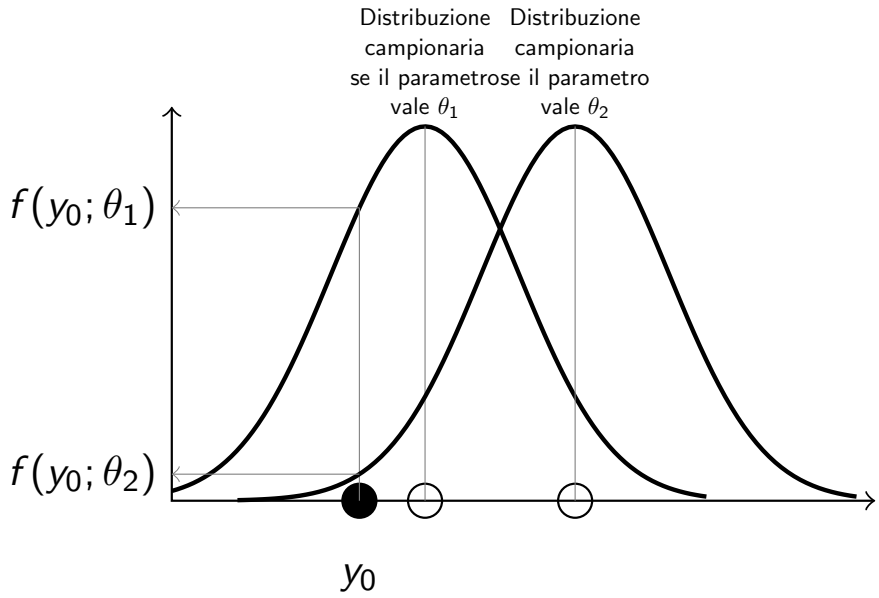


Metodo della verosimiglianza

Partiamo della distribuzione campionaria per un dato valore del parametro, θ_1.

Consideriamo il campione, y_0.

Possiamo calcolare la probabilità (o densità) del campione corrispondente al valore θ_1 del parametro.



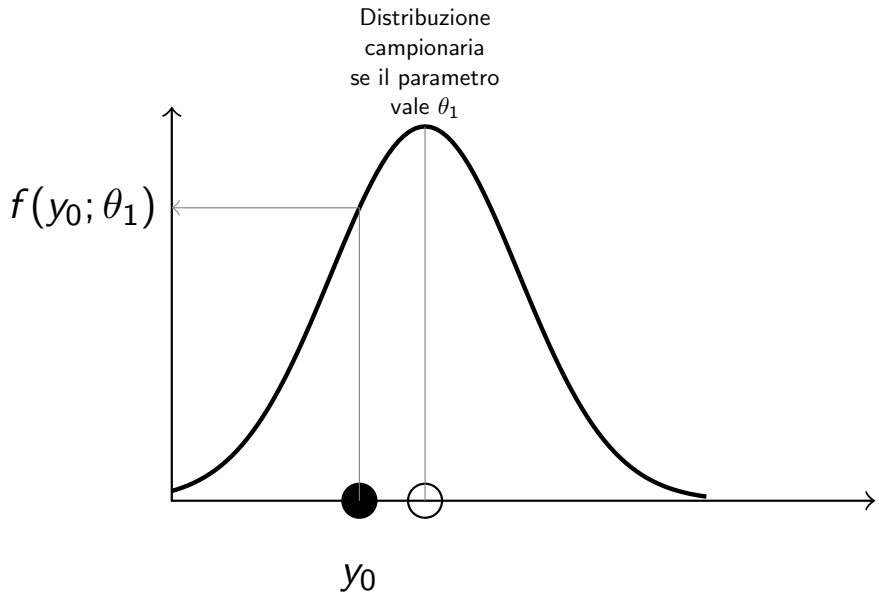
La stessa cosa può essere fatta per altri valori del parametro, qui θ_2.

Metodo della verosimiglianza

Partiamo della distribuzione campionaria per un dato valore del parametro, θ_1.

Consideriamo il campione, y_0.

Possiamo calcolare la probabilità (o densità) del campione corrispondente al valore θ_1 del parametro.



La stessa cosa può essere fatta per altri valori del parametro, qui θ_2.

Il valore θ_1 rende il campione più probabile che non θ_2, si dice che è più verosimile.

Con questo criterio si ottiene uno stimatore scegliendo **il valore di θ che rende più probabile il campione, detto stimatore di massima verosimiglianza.**

Massima verosimiglianza: esempio binomiale

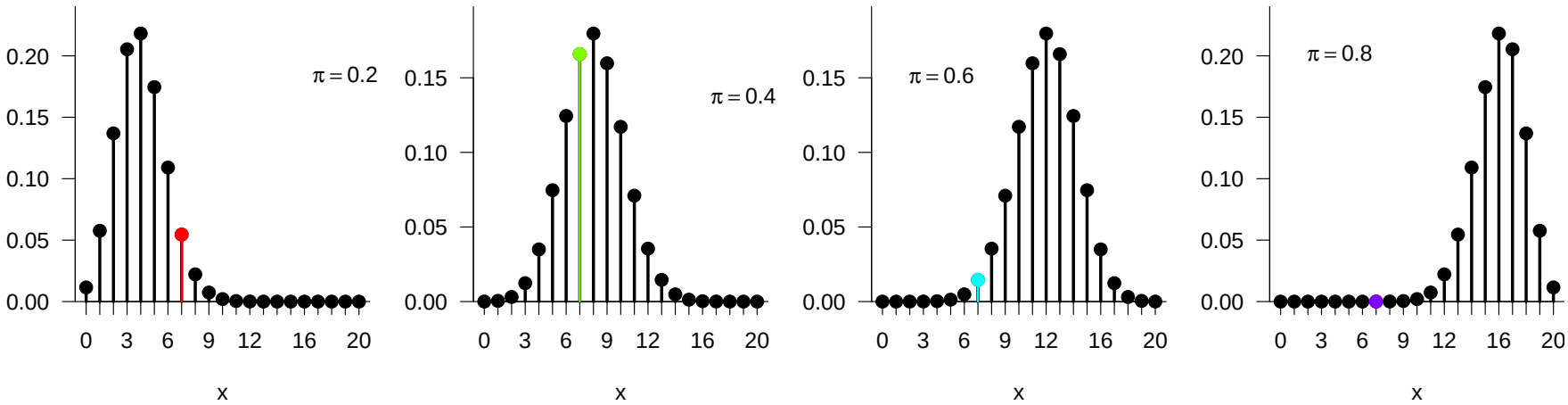
Consideriamo un campione da una Binomiale di dimensione n e parametro θ. La probabilità di un particolare campione x_0 è

$$P(X = x_0) = \binom{n}{x} \pi^{x_0} (1 - \pi)^{n - x_0}$$

Massima verosimiglianza: esempio binomiale

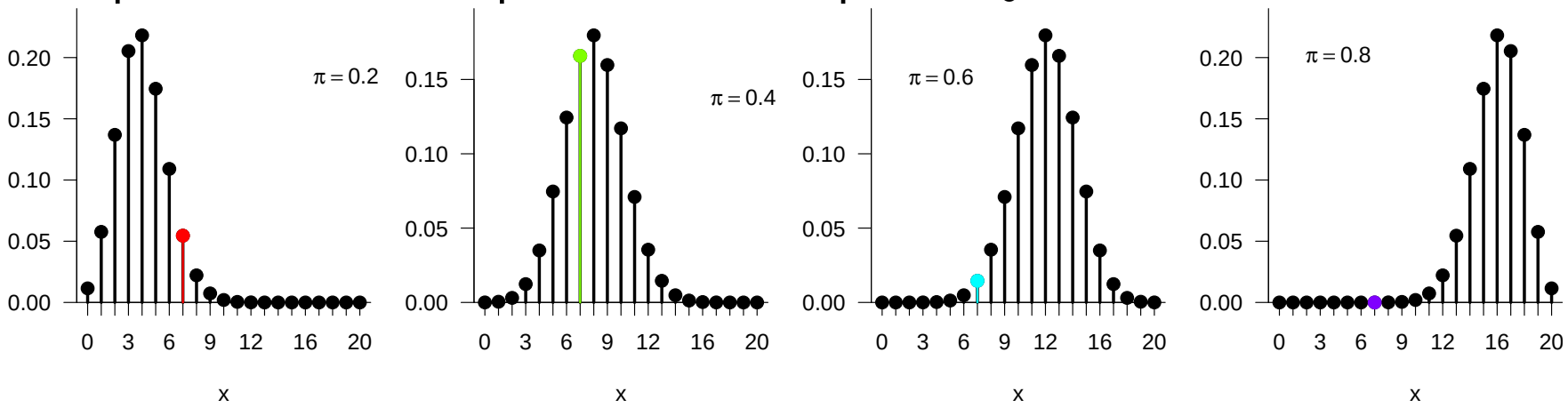
Consideriamo un campione da una Binomiale di dimensione n e parametro θ. La probabilità di un particolare campione x_0 è

$$P(X = x_0) = \binom{n}{x} \pi^{x_0} (1 - \pi)^{n - x_0}$$



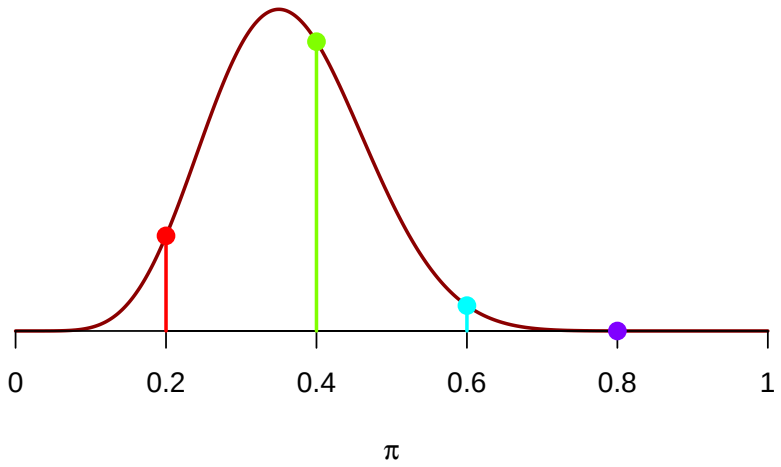
Massima verosimiglianza: esempio binomiale

Consideriamo un campione da una Binomiale di dimensione n e parametro θ. La probabilità di un particolare campione x_0 è



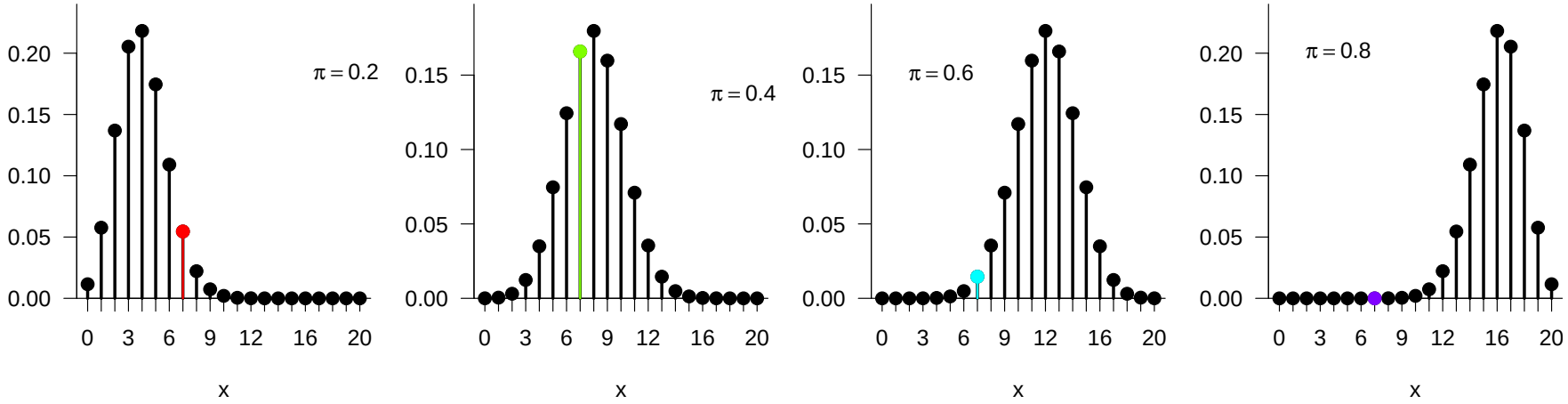
La funzione che descrive l'andamento di questa probabilità al variare di π è

$$L(\pi) = \binom{n}{x_0} \pi^{x_0} (1-\pi)^{n-x_0}$$



Massima verosimiglianza: esempio binomiale

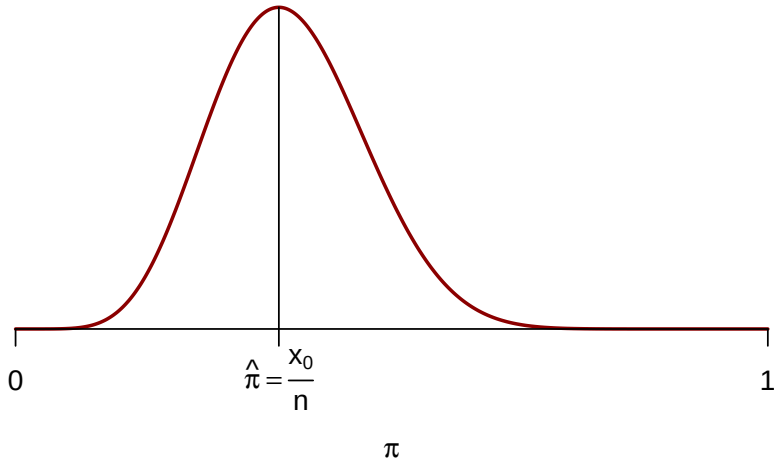
Consideriamo un campione da una Binomiale di dimensione n e parametro θ. La probabilità di un particolare campione x_0 è



La funzione che descrive l'andamento di questa probabilità al variare di π è

$$L(\pi) = \binom{n}{x_0} \pi^{x_0} (1-\pi)^{n-x_0}$$

Il massimo si ha in $\hat{\pi} = \frac{x_0}{n}$



Massimo della verosimiglianza binomiale

Dobbiamo ricavare il massimo di $L(\pi) = \binom{n}{x_0}\pi^{x_0}(1-\pi)^{n-x_0}$ ma questo equivale a trovare il massimo di

$$\ln L(\pi) = \ln \binom{n}{x_0} + x_0 \ln \pi + (n - x_0)\ln(1-\pi)$$

per la quale si ha la derivata

$$(\ln L(\pi))' = \frac{x_0}{\pi} - \frac{n - x_0}{1-\pi}$$

Cerchiamo i punti stazionari eguagliando la derivata a 0

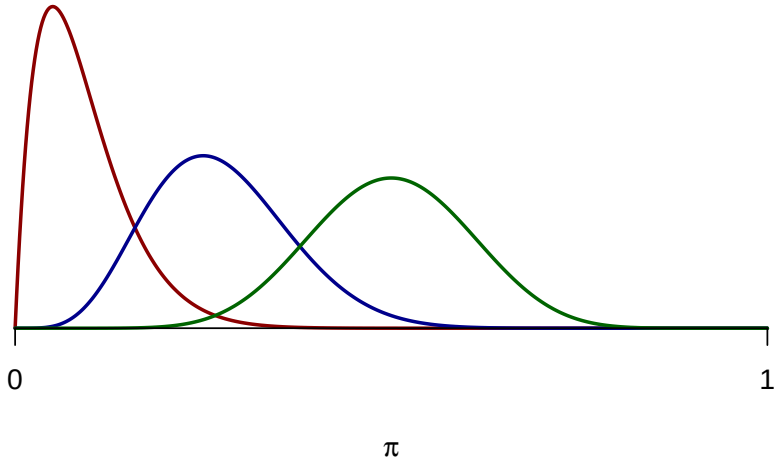
$$\frac{x_0}{\pi} - \frac{n - x_0}{1-\pi} = 0 \quad \Leftrightarrow \quad \hat{\pi} = \frac{x_0}{n}$$

che è un massimo in quanto la funzione ha la convessità verso il basso su tutto il dominio

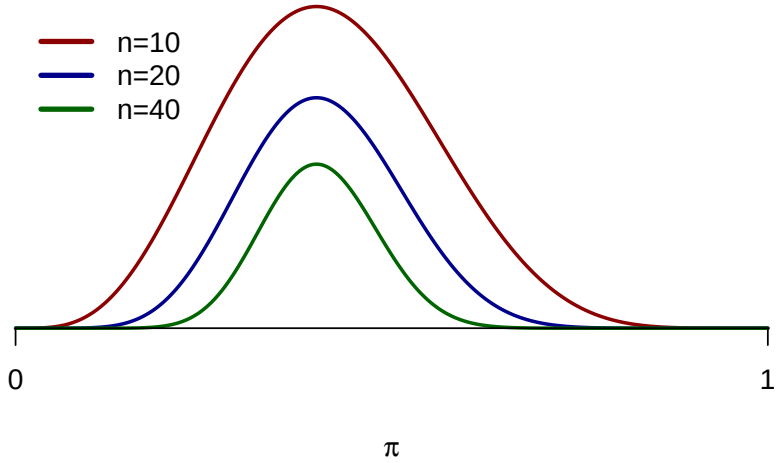
$$(\ln L(\pi))'' = -\frac{x_0}{\pi^2} - \frac{n - x_0}{(1-\pi)^2} < 0$$

Verosimiglianza binomiale

Funzione $L(\pi)$ per campioni di numerosità 20 con diversa proporzione campionaria.



Funzione $L(\pi)$ per campioni di diversa numerosità con la medesima proporzione campionaria.



Massimo della verosimiglianza: caso dell'esponenziale

Si ha un campione $X_1, \ldots, X_n$ da un'esponenziale di parametro λ,

$$f(x_i; \lambda) = \lambda e^{-\lambda x_i}$$

complessivamente, essendo le osservazioni indipendenti, la funzione di densità è il prodotto delle densità

$$f(x_1, \ldots, x_n; \lambda) = \prod_{i=1}^{n} f(x_i; \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

Per ottenere lo stimatore di massima verosimiglianza consideriamo la funzione

$$L(\lambda) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

il cui massimo è in

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$$

Massimo della verosimiglianza esponenziale

Per ottenere il massimo di

$$L(\lambda) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

si considera il trasformato logaritmico

$$\log(L(\lambda)) = n \log \lambda - \lambda \sum_{i=1}^n x_i$$

la cui derivata

$$(\log(L(\lambda)))' = \frac{n}{\lambda} - \sum_{i=1}^n x_i$$

si annulla per $\hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i}$, dove

$$(\log(L(\lambda)))'' = -\frac{n}{\lambda^2} < 0$$

e quindi si ha un massimo.

Indice

Distribuzione campionaria e precisione di uno stimatore

Ottenere uno stimatore

Proprietà di uno stimatore

Proprietà di uno stimatore

Per valutare la bontà di uno stimatore si guarda a come è fatta la sua distribuzione campionaria.

In particolare, si considerano alcune caratteristiche della distribuzione in relazione al parametro.

Il principio sottostante le proprietà che vengono illustrate è che vorremmo che fosse probabile che le stime che si ottengono cadano vicino al vero valore del parametro.

Proprietà di uno stimatore: correttezza

Definizione: correttezza di uno stimatore

Uno stimatore $\hat{\theta}$ per il parametro θ si dice corretto (o non distorto) se e solo se, qualunque sia θ,

$$E(\hat{\theta}) = \theta$$

Se uno stimatore è corretto, mediamente le stime hanno lo stesso valore del parametro.

Un esempio di stimatore distorto è x_n (il massimo) nell'esempio della statistica in guerra.

Proprietà di uno stimatore: correttezza

Definizione: correttezza di uno stimatore

Uno stimatore $\hat{\theta}$ per il parametro θ si dice corretto (o non distorto) se e solo se, qualunque sia θ,

$$E(\hat{\theta}) = \theta$$

- Lo stimatore proporzione campionaria per la proporzione nella popolazione è $\hat{\pi} = \frac{X}{n}$ dove $X \sim \text{Binomiale}(n, \pi)$, quindi

$$E(\hat{\pi}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}n\pi = \pi$$

- Lo stimatore media campionaria per la media di una popolazione è anch'esso corretto.

Proprietà di uno stimatore: consistenza

Definizione: consistenza di uno stimatore

Uno stimatore $\hat{\theta}_n$ per il parametro θ si dice consistente se e solo se, qualunque sia θ,

$$\hat{\theta}_n \underset{p}{\to} \theta$$

La proprietà di consistenza significa che se il campione diventa molto grande diventa arbitrariamente vicina a 1 la probabilità che lo stimatore cada vicino al parametro.

Si dimostra che una condizione sufficiente per la consistenza di $\hat{\theta}_n$ è che

$$E(\hat{\theta}_n) \to \theta \quad \text{e} \quad V(\hat{\theta}_n) \to 0$$

al crescere di n la variabilità tende a 0 e la distribuzione campionaria si concentra su θ

Proprietà di uno stimatore: consistenza

Definizione: consistenza di uno stimatore

Uno stimatore $\hat{\theta}_n$ per il parametro θ si dice consistente se e solo se, qualunque sia θ,

$$\hat{\theta}_n \underset{p}{\rightarrow} \theta$$

- La proporzione campionaria per la proporzione nella popolazione e la media campionaria per la media di una popolazione sono consistenti in virtù della legge dei grandi numeri.
- Si può provare la consistenza degli stimatori sopra esaminando media e varianza.

Confronto di stimatori: efficienza

Una quantificazione dell'errore medio che commetto se uso uno stimatore è

Definizione: errore quadratico medio

Si dice errore quadratico medio (EQM) dello stimatore $\hat{\theta}$ per il parametro θ la quantità

$$MSE(\hat{\theta}) = E((\hat{\theta} - \theta)^2)$$

Consideriamo due stimatori $\hat{\theta}_1$ e $\hat{\theta}_2$ per uno stesso parametro θ, è sensato confrontarli rispetto ai rispettivi EQM

Definizione: efficienza

Dati gli stimatori $\hat{\theta}_1$ e $\hat{\theta}_2$ per il parametro θ si dice che $\hat{\theta}_1$ è più efficiente di $\hat{\theta}_2$ se e solo se, qualunque sia θ,

$$MSE(\hat{\theta}_1) \leq MSE(\hat{\theta}_2)$$