

La media campionaria

MEDIA CAMPIONARIA

Date n v.a. X_1, \dots, X_n indipendenti e identicamente distribuite (in breve i.i.d.), la v.a.

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i,$$

è detta media campionaria.

In altre parole, se le X_1, \dots, X_n sono il risultato di n misurazioni (indipendenti) di una stessa quantità (aleatoria), la media campionaria è la media aritmetica di queste misurazioni.

La media campionaria

È naturale chiedersi alcune cose su \bar{X}_n :

- posso dire che legge ha?
- c'entra qualcosa con $\mathbb{E}(X_1)$?

Alla prima domanda rispondiamo che in generale non è semplice conoscere la legge di \bar{X}_n , ma se n è grande e ci si accontenta di un'approssimazione, il Teorema del Limite Centrale dà una risposta.

Alla seconda domanda risponderà la legge dei grandi numeri.

$$\text{Su } \mathbb{E}(X_1) \longleftrightarrow \bar{X}_n$$

Notiamo fin d'ora: se i valori che ognuna delle v.a. può assumere sono, ad esempio, v_1, v_2, v_3, v_4 e $f_{X_1}(v_1) = p_1, f_{X_1}(v_2) = p_2, f_{X_1}(v_3) = p_3, f_{X_1}(v_4) = p_4$, allora

$$\begin{aligned}\bar{X}_n &= \frac{1}{n} (v_1 \cdot \text{num. di volte che } v_1 \text{ è presente nelle osservazioni} \\ &\quad + v_2 \cdot \text{num. di volte che } v_2 \text{ è presente nelle osservazioni} \\ &\quad + v_3 \cdot \text{num. di volte che } v_3 \text{ è presente nelle osservazioni} \\ &\quad + v_4 \cdot \text{num. di volte che } v_4 \text{ è presente nelle osservazioni}) \\ &= v_1 \cdot \text{freq.rel. di } v_1 \\ &\quad + v_2 \cdot \text{freq.rel. di } v_2 \\ &\quad + v_3 \cdot \text{freq.rel. di } v_3 \\ &\quad + v_4 \cdot \text{freq.rel. di } v_4) \\ &\approx v_1 \cdot p_1 + v_2 \cdot p_2 + v_3 \cdot p_3 + v_4 \cdot p_4 = \mathbb{E}(X_1).\end{aligned}$$

Quindi $\bar{X}_n \approx \mathbb{E}(X_1)$ se pensiamo che le frequenze relative \approx probabilità.

$$\text{Su } \mathbb{E}(X_1) \longleftrightarrow \bar{X}_n$$

È dunque piuttosto naturale dire che,
SE frequenze relative \approx probabilità ALLORA ANCHE
 $\bar{X}_n \approx \mathbb{E}(X_1)$.

L'idea di fondo è che “frequenze relative $\not\approx$ probabilità” sia un fatto improbabile se n è grande.

L'affermazione precisa è la legge dei grandi numeri che vedremo a breve.

Il teorema del limite centrale

TLC

Sia $\{X_i\}_{i \geq 1}$ una successione di v.a. i.i.d. tutte con valore atteso $\mathbb{E}(X_i) = \mu$ e varianza finita $\text{Var}(X_i) = \sigma^2 > 0$. Sia

$$S_n^* = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}.$$

Allora per ogni $t \in \mathbb{R}$ vale

$$\lim_{n \rightarrow +\infty} \mathbb{P}(S_n^* \leq t) = \Phi(t).$$

$$S_n^* = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}.$$

è la standardizzata della media campionaria,

infatti

$$\begin{aligned}\mathbb{E}(\bar{X}_n) &= \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) \\ &= \frac{1}{n} \sum_{i=1}^n \mu = \mu.\end{aligned}$$

$$\begin{aligned}\text{Var}(\bar{X}_n) &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}.\end{aligned}$$

$$\lim_{n \rightarrow +\infty} \mathbb{P}(S_n^* \leq t) = \Phi(t) = \mathbb{P}(\mathcal{N}(0, 1) \leq t).$$

ci dice la legge di S_n^* si approssima, per n grande, con quella di una $\mathcal{N}(0, 1)$.

In altre parole, **se vogliamo calcolare probabilità relative a S_n^* possiamo utilizzare quelle relative a $\mathcal{N}(0, 1)$ come approssimazione.**

Approssimazione normale

qui con \approx indichiamo l'approssimazione delle probabilità di cui abbiamo appena discusso:

$$S_n^* \approx \mathcal{N}(0, 1)$$

$$\bar{X}_n \approx \mathcal{N}(\mu, \sigma^2/n)$$

$$X_1 + \cdots + X_n \approx \mathcal{N}(n\mu, n\sigma^2)$$

Pro

Questa approssimazione vale qualsiasi sia la legge delle X_1, \dots, X_n .

Contro

È un'approssimazione e vale solo se n è sufficientemente grande, ma quanto grande dipende dalla legge delle X_1, \dots, X_n (dunque in generale non si conosce!).

Approssimazione normale

$$S_n^* = \mathcal{N}(0, 1)$$

$$\bar{X}_n = \mathcal{N}(\mu, \sigma^2/n)$$

$$X_1 + \cdots + X_n = \mathcal{N}(n\mu, n\sigma^2)$$

Identità

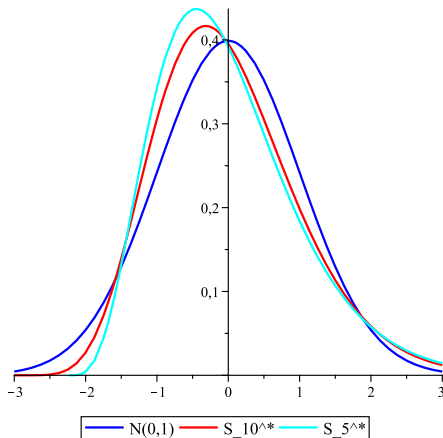
L'approssimazione diviene un'uguaglianza nel caso di variabili normali $\mathcal{N}(\mu, \sigma^2)$ i.i.d.

Quando vale l'approssimazione

- Se la legge delle X_1, \dots, X_n non è troppo asimmetrica, *a livello empirico* si è stabilito che $n \geq 30$ va bene.
- Se invece la legge di ciascuna X_i è $\mathcal{B}(p)$, l'approssimazione normale si considera valida se sia np che $n(1 - p)$ sono ≥ 5 .

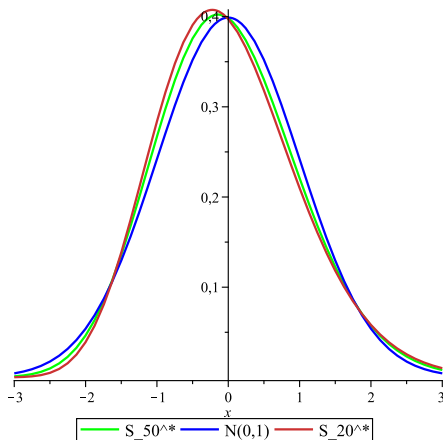
Il TLC visto con le densità

Prendiamo come esempio una successione di v.a. X_i di tipo esponenziale (non è importante qui sapere come sono definite) e confrontiamo il grafico della densità della standardizzata di \bar{X}_n con quello di $\mathcal{N}(0, 1)$:



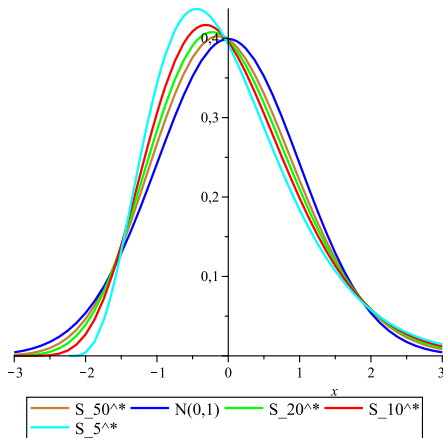
Il TLC visto con le densità

In realtà il teorema confronta le aree e non le curve, ma poiché siamo abituati a vedere la campana della densità $\mathcal{N}(0, 1)$ guardiamo questi grafici:



Il TLC visto con le densità

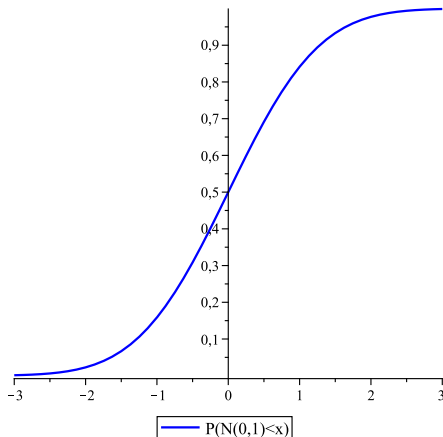
Vediamo infine tutte le densità: quella di $\mathcal{N}(0, 1)$ e quelle di S_5^* , S_{10}^* , S_{20}^* , S_{50}^* .



Il TLC visto con la funzione Φ

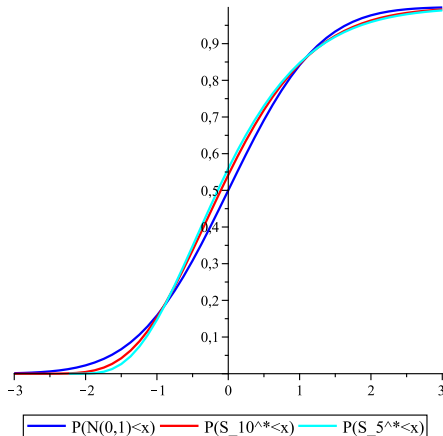
Invece della densità, guardiamo ora il grafico della funzione

$$x \mapsto \mathbb{P}(\mathcal{N}(0, 1) \leq x) = \Phi(x)$$



Grafici di approssimazione

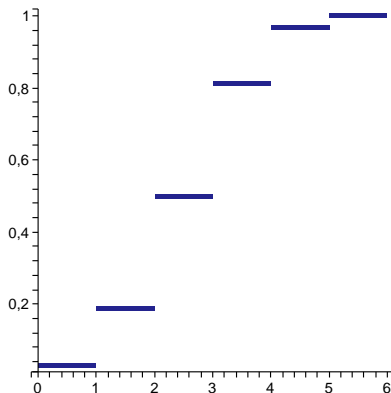
Confrontiamo con gli analoghi grafici per le medie campionarie standardizzate di prima (S_5^* e S_{10}^*):



Il TLC per somma di Bernoulli

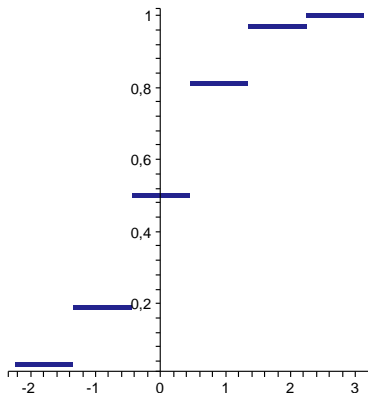
Consideriamo ora X_1, X_2, X_3, X_4, X_5 indipendenti e ciascuna con legge $\mathcal{B}(0.5)$. Sia $S_5 = X_1 + X_2 + X_3 + X_4 + X_5$ e guardiamo il grafico della funzione

$$x \mapsto \mathbb{P}(S_5 \leq x) = \mathbb{P}(B(5, 0.5) \leq x)$$



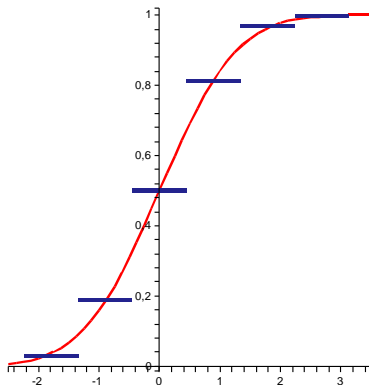
Il TLC per somma di Bernoulli

Standardizziamo S_5 e guardiamo il grafico di $\mathbb{P}(S_5^* \leq x)$



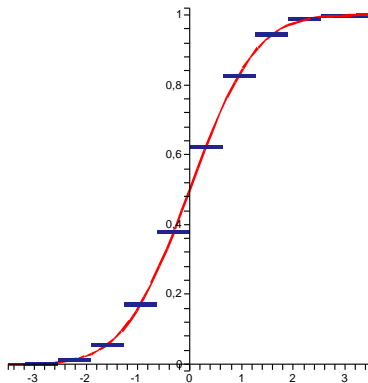
Il TLC per somma di Bernoulli

Confrontiamo il grafico di $\mathbb{P}(S_5^* \leq x)$ e quello di $\mathbb{P}(\mathcal{N}(0, 1) \leq x)$:



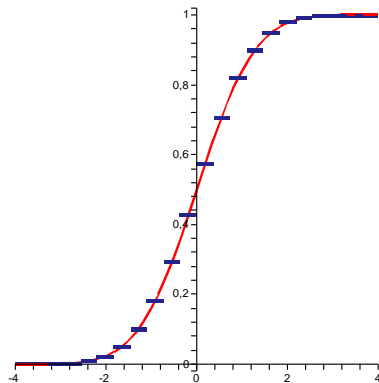
Il TLC per somma di Bernoulli

Confrontiamo il grafico di $\mathbb{P}(S_{10}^* \leq x)$ e quello di $\mathbb{P}(\mathcal{N}(0, 1) \leq x)$:



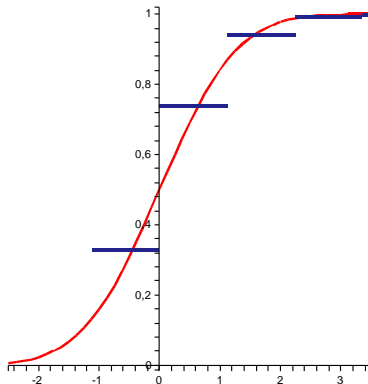
Il TLC per somma di Bernoulli

Confrontiamo il grafico di $\mathbb{P}(\mathbf{S}_{30}^* \leq x)$ e quello di $\mathbb{P}(\mathcal{N}(0, 1) \leq x)$:



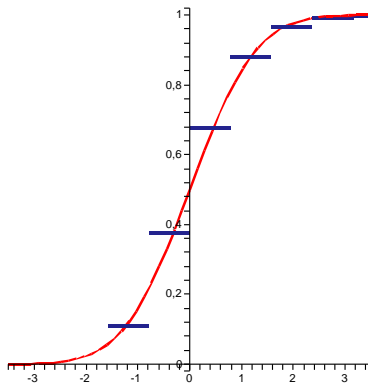
Il TLC per somma di Bernoulli

Prendiamo ora delle $\mathcal{B}(0.2)$: confrontiamo il grafico di $\mathbb{P}(S_5^* \leq x)$ e quello di $\mathbb{P}(\mathcal{N}(0, 1) \leq x)$:



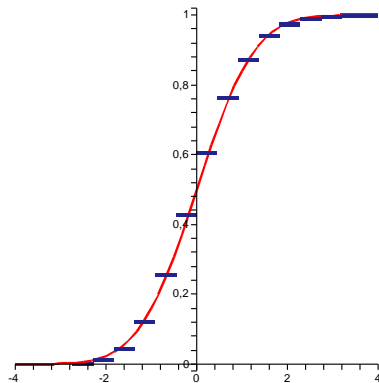
Il TLC per somma di Bernoulli

Confrontiamo il grafico di $\mathbb{P}(S_{10}^* \leq x)$ e quello di $\mathbb{P}(\mathcal{N}(0, 1) \leq x)$:



Il TLC per somma di Bernoulli

Confrontiamo il grafico di $\mathbb{P}(\mathbf{S}_{30}^* \leq x)$ e quello di $\mathbb{P}(\mathcal{N}(0, 1) \leq x)$:



La binomiale

Una delle approssimazioni normali viste è

$$X_1 + \cdots + X_n \approx \mathcal{N}(n\mu, n\sigma^2)$$

poiché una $\mathcal{B}(n, p)$ può essere vista come somma di n $\mathcal{B}(p)$ indipendenti ne ricaviamo

$$\mathcal{B}(n, p) \approx \mathcal{N}(np, np(1 - p)).$$

La moneta

Se lancio 1000 volte una moneta equilibrata, qual è la probabilità che escano meno di 490 teste? e più di 530? E che la differenza con 500 superi 5?

$$\mathbb{P}(\mathcal{B}(1000, 0.5) < 490) \approx \mathbb{P}(\mathcal{N}(500, 250) < 490)$$

$$\mathbb{P}(\mathcal{B}(1000, 0.5) > 530) \approx \mathbb{P}(\mathcal{N}(500, 250) > 530)$$

$$\mathbb{P}(|\mathcal{B}(1000, 0.5) - 500| > 5) \approx \mathbb{P}(|\mathcal{N}(500, 250) - 500| > 5)$$

Calcoliamo

$$\begin{aligned}\mathbb{P}(\mathcal{B}(1000, 0.5) < 490) &\approx \mathbb{P}(\mathcal{N}(500, 250) < 490) \\ &= \mathbb{P}\left(\mathcal{N}(0, 1) < \frac{490 - 500}{\sqrt{250}}\right) = \Phi(-0.63) \\ &= 1 - 0.73565 = 0.26435.\end{aligned}$$

$$\begin{aligned}\mathbb{P}(\mathcal{B}(1000, 0.5) > 530) &\approx \mathbb{P}(\mathcal{N}(500, 250)) > 530) \\ &= \mathbb{P}\left(\mathcal{N}(0, 1) > \frac{530 - 500}{\sqrt{250}}\right) = 1 - \Phi(1.90) \\ &= 1 - 0.97128 = 0.02872.\end{aligned}$$

$$\begin{aligned}\mathbb{P}(|\mathcal{B}(1000, 0.5) - 500| > 5) &\approx \mathbb{P}(|\mathcal{N}(500, 250) - 500| > 5) \\ &= \mathbb{P}\left(|\mathcal{N}(0, 1)| > \frac{5}{\sqrt{250}}\right) = 2(1 - \Phi(0.32)) \\ &= 2 \cdot (1 - 0.62552) = 0.74896.\end{aligned}$$

10000 lanci

Se lancio 10000 volte una moneta equilibrata, qual è la probabilità che escano meno di 4900 teste? e più di 5300? E che la differenza con 5000 superi 50?

$$\mathbb{P}(\mathcal{B}(10000, 0.5) < 4900) \approx \mathbb{P}(\mathcal{N}(5000, 2500) < 4900)$$

$$\mathbb{P}(\mathcal{B}(10000, 0.5) > 5300) \approx \mathbb{P}(\mathcal{N}(5000, 2500) > 5300)$$

$$\begin{aligned}\mathbb{P}(|\mathcal{B}(10000, 0.5) - 5000| > 50) \\ \approx \mathbb{P}(|\mathcal{N}(5000, 2500) - 5000| > 50)\end{aligned}$$

Calcoliamo

$$\begin{aligned}\mathbb{P}(\mathcal{B}(10000, 0.5) < 4900) &\approx \mathbb{P}(\mathcal{N}(5000, 2500) < 4900) \\ &= \mathbb{P}\left(\mathcal{N}(0, 1) < \frac{4900 - 5000}{\sqrt{2500}}\right) = \Phi(-2) \\ &= 1 - 0.97725 = 0.02275.\end{aligned}$$

$$\begin{aligned}\mathbb{P}(\mathcal{B}(1000, 0.5) > 5300) &\approx \mathbb{P}(\mathcal{N}(5000, 2500)) > 5300) \\ &= \mathbb{P}\left(\mathcal{N}(0, 1) > \frac{5300 - 5000}{\sqrt{2500}}\right) = 1 - \Phi(6) \approx 0\end{aligned}$$

$$\begin{aligned}\mathbb{P}(|\mathcal{B}(10000, 0.5) - 5000| > 5) \\ &\approx \mathbb{P}(|\mathcal{N}(5000, 2500) - 5000| > 50) \\ &= \mathbb{P}\left(|\mathcal{N}(0, 1)| > \frac{50}{\sqrt{2500}}\right) = 2(1 - \Phi(1)) \\ &= 2 \cdot (1 - 0.84134) = 0.31732.\end{aligned}$$

Media e varianza cosa dicono?

Ricordiamo: se X è una v.a.

$\mathbb{E}(X)$ è un indice di posizione, $\text{Var}(X)$ un indice di dispersione.

Vogliamo ora **quantificare il significato di “varianza = misura della dispersione”**: la disuguaglianza di Chebychev è quello che ci serve.



Pafnuty Lvovich Chebychev (1821-1894)

Disuguaglianza di Chebychev

Disuguaglianza di Chebychev

Sia X una v.a. con $\mathbb{E}(X)=\mu$ e $\text{Var}(X)=\sigma^2$. Sia ε un numero reale positivo prefissato. Vale la seguente disuguaglianza:

$$\mathbb{P}(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

o equivalentemente

$$\mathbb{P}(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}.$$

Se $\varepsilon = \delta\sigma$ ($\delta > 0$) allora si ha

$$\mathbb{P}(|X - \mu| \geq \delta\sigma) \leq \frac{1}{\delta^2}$$

$$\mathbb{P}(|X - \mu| < \delta\sigma) \geq 1 - \frac{1}{\delta^2}.$$

La prima versione

“Sciogliendo” il modulo in

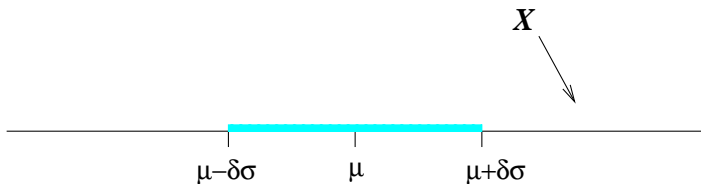
$$\mathbb{P}(|X - \mu| \geq \delta\sigma) \leq \frac{1}{\delta^2}$$

e il fatto che $(X \geq \mu + \delta\sigma)$ e $(X \leq \mu - \delta\sigma)$ sono eventi incompatibili otteniamo che

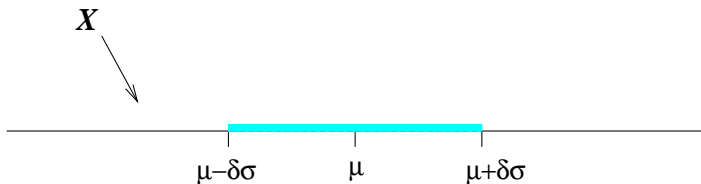
$$\mathbb{P}(X \geq \mu + \delta\sigma) + \mathbb{P}(X \leq \mu - \delta\sigma) \leq \frac{1}{\delta^2}$$

La prima versione

Significa che la probabilità che la v.a. X assuma valori



oppure



è minore di $\frac{1}{\delta^2}$.

La seconda versione

“Sciogliendo” il modulo in

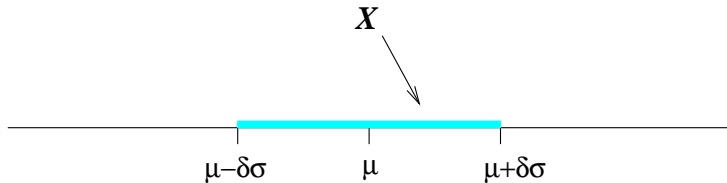
$$\mathbb{P}(|X - \mu| < \delta\sigma) \geq 1 - \frac{1}{\delta^2}$$

otteniamo che

$$\mathbb{P}(\mu - \delta\sigma < X < \mu + \delta\sigma) \geq 1 - \frac{1}{\delta^2}$$

La seconda versione

Significa che la probabilità che la v.a. X assuma valori



è maggiore o uguale a $1 - \frac{1}{\delta^2}$.

Chebychev con alcuni δ

Riscriviamo la seconda versione di Chebychev con $\delta = 2, 3, 5, 10$:

$$\mathbb{P}(\mu - 2\sigma < X < \mu + 2\sigma) \geq 1 - \frac{1}{4} = 0.75$$

$$\mathbb{P}(\mu - 3\sigma < X < \mu + 3\sigma) \geq 1 - \frac{1}{9} = 0.88$$

$$\mathbb{P}(\mu - 5\sigma < X < \mu + 5\sigma) \geq 1 - \frac{1}{25} = 0.96$$

$$\mathbb{P}(\mu - 10\sigma < X < \mu + 10\sigma) \geq 1 - \frac{1}{100} = 0.99$$

Significato

Ad esempio con $\delta = 5$ abbiamo che con una probabilità **almeno** del 96% X assume valori nell'intervallo $[\mu - 5\sigma, \mu + 5\sigma]$; con una probabilità **al massimo** del 4% X assume valori fuori da quell'intervallo.

Cosa mi dice Chebychev

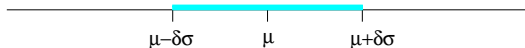
Non so prevedere esattamente il valore di X (perché è una v.a.), MA ho un intervallo di valori in cui è molto probabile che si trovi il valore che X assumerà.

Varianza=dispersione

$$\mathbb{P}(\mu - \delta\sigma < X < \mu + \delta\sigma) \geq 1 - \frac{1}{\delta^2}$$

Come influisce la varianza σ^2

La probabilità (almeno $1 - \frac{1}{\delta^2}$) è fissata se scelgo δ , ma quanto è largo l'intervallo dipende da σ^2 .

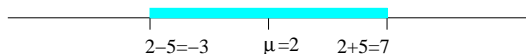


Esempi con varianze diverse

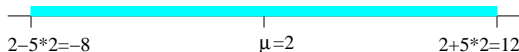
Sia X v.a. con $\mathbb{E}(X) = 2$ e $\text{Var}(X) = 1$ e sia Y v.a. con $\mathbb{E}(Y) = 2$ e $\text{Var}(Y) = 4$.

Cerchiamo un intervallo in cui X assuma valori con probabilità ≥ 0.96 e idem per Y .

Per X :



Per Y :



Differenze

Il centro dell'intervallo è lo stesso, ma **fissata la probabilità** l'intervallo è più ampio per la v.a. con varianza maggiore.

Varianza come misura dell'incertezza

Una varianza maggiore mi dà maggiore incertezza sull'esito dell'esperimento "qual è il valore assunto da X ". L'idea è resa quantitativa da Chebychev.

La legge dei grandi numeri

Legge dei grandi numeri Legge **forte** dei grandi numeri

Siano X_1, X_2, \dots, X_n v.a. indipendenti e identicamente distribuite. Sia $\mathbb{E}(X_i) = \mu$ e $\text{Var}(X_i) = \sigma^2$ per ogni i . Allora per ogni $\varepsilon > 0$ vale

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mu| > \varepsilon) = 0. \mathbb{P}\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mu\right) = 1.$$

Ricordiamo

Che tutti i valori attesi e tutte le varianze siano uguali è implicito nella richiesta **identicamente distribuite**. Quello che specifichiamo nelle ipotesi è che chiamiamo μ e σ^2 rispettivamente il valore atteso e la varianza.

\bar{X}_n è la media campionaria cioè $\frac{1}{n} \sum_{i=1}^n X_i$.

Osserviamo che il teorema vale anche se le v.a. non ammettono varianza, cioè anche se $\mathbb{E}((X_1 - \mathbb{E}(X_1))^2) = +\infty$.

Dimostrazione

Applichiamo la disuguaglianza di Chebychev alla v.a. \bar{X}_n :

$$\mathbb{P}\left(|\bar{X}_n - \mathbb{E}(\bar{X}_n)| \geq \delta \sqrt{\text{Var}(\bar{X}_n)}\right) \leq \frac{1}{\delta^2} \quad \text{per ogni } \delta > 0.$$

Ci serve il calcolo di $\mathbb{E}(\bar{X}_n)$ e di $\text{Var}(\bar{X}_n)$, si fa utilizzando le proprietà di valore atteso e varianza:

$$\begin{aligned}\mathbb{E}(\bar{X}_n) &= \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \mathbb{E}\left(\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \mu.\end{aligned}$$

$$\begin{aligned}\text{Var}(\bar{X}_n) &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}.\end{aligned}$$

Dimostrazione

Applichiamo la disuguaglianza di Chebychev alla v.a. \bar{X}_n :

$$\mathbb{P} \left(|\bar{X}_n - \mathbb{E}(\bar{X}_n)| \geq \delta \sqrt{\text{Var}(\bar{X}_n)} \right) \leq \frac{1}{\delta^2} \quad \text{per ogni } \delta > 0.$$

Ora abbiamo $\mathbb{E}(\bar{X}_n) = \mu$ e $\text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$.

$$\mathbb{P} \left(|\bar{X}_n - \mu| \geq \delta \frac{\sigma}{\sqrt{n}} \right) \leq \frac{1}{\delta^2} \quad \text{per ogni } \delta > 0.$$

Scegliamo $\delta = \varepsilon \sqrt{n} / \sigma$ (in modo che $\delta \frac{\sigma}{\sqrt{n}} = \varepsilon$):

$$\mathbb{P} \left(|\bar{X}_n - \mu| \geq \varepsilon \right) \leq \frac{\sigma^2}{\varepsilon^2 n} \quad \text{per ogni } \varepsilon > 0$$

Per $n \rightarrow \infty$ si ottiene la tesi.

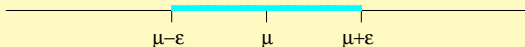
Cosa significa

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mu| > \varepsilon) = 0?$$

Penso di aver scelto ε (ad esempio $\varepsilon = 0.01$). $|\bar{X}_n - \mu|$ è la distanza fra la media campionaria \bar{X}_n e μ (che è il valore atteso di ciascuna X_i).

⇒ La probabilità che \bar{X}_n cada **fuori** dall'intervallo colorato in figura è trascurabile (≈ 0) se n è abbastanza grande.

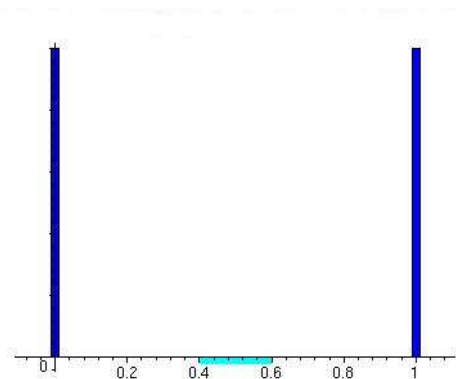
⇒ La probabilità che \bar{X}_n cada **dentro** l'intervallo colorato in figura è ≈ 1 se n è abbastanza grande.



La LGN vista con dei grafici

Prendiamo una variabile $X_1 \sim \mathcal{B}(0.5)$ e $\varepsilon = 0.1$. L'intervallo $(\mu - \varepsilon, \mu + \varepsilon)$ è $(0.4, 0.6)$ e in figura è rappresentato in azzurro (ricordiamo che $\mu = p = 0.5$).

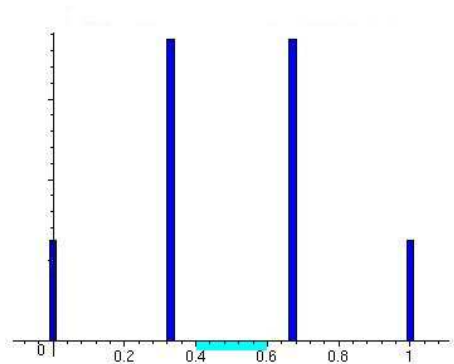
In blu la densità di $\bar{X}_1 = X_1$.



In questo caso $\mathbb{P}(|\bar{X}_1 - 0.5| > 0.1) = 1$.

La LGN vista con dei grafici

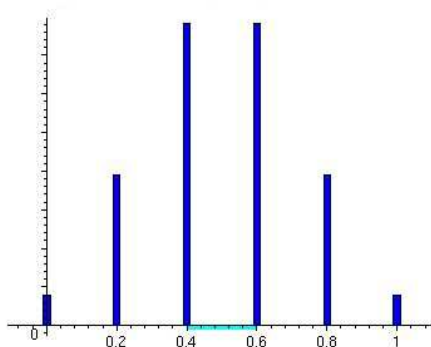
Prendiamo ora tre variabili indipendenti X_1, X_2, X_3 , tutte $B(0.5)$. In blu la densità di $\bar{X}_3 = (X_1 + X_2 + X_3)/3$.



In questo caso $\mathbb{P}(|\bar{X}_3 - 0.5| > 0.1) = 1$.

La LGN vista con dei grafici

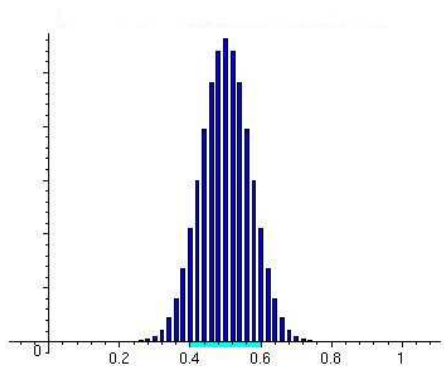
Prendiamo ora 5 variabili indipendenti X_1, X_2, X_3, X_4, X_5 , tutte $\mathcal{B}(0.5)$. In blu la densità di $\bar{X}_5 = (X_1 + \dots + X_5)/5$.



In questo caso $\mathbb{P}(|\bar{X}_5 - 0.5| > 0.1) = 0.375$.

La LGN vista con dei grafici

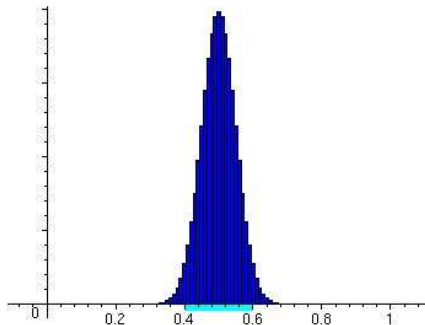
Prendiamo ora 50 variabili indipendenti X_1, \dots, X_{50} , tutte $B(0.5)$. In blu la densità di $\bar{X}_{50} = (X_1 + \dots + X_{50})/50$.



In questo caso $\mathbb{P}(|\bar{X}_{50} - 0.5| > 0.1) = 0.11832$.

La LGN vista con dei grafici

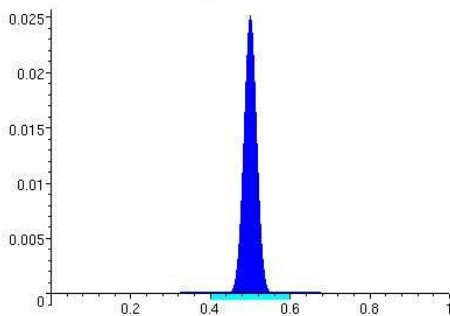
Prendiamo ora 100 variabili indipendenti X_1, \dots, X_{100} , tutte $B(0.5)$. In blu la densità di $\bar{X}_{100} = (X_1 + \dots + X_{100})/100$.



In questo caso $\mathbb{P}(|\bar{X}_{100} - 0.5| > 0.1) = 0.0352$.

La LGN vista con dei grafici

Prendiamo ora 1000 variabili indipendenti X_1, \dots, X_{1000} , tutte $\mathcal{B}(0.5)$. In blu la densità di $\bar{X}_{1000} = (X_1 + \dots + X_{1000})/1000$.



In questo caso $\mathbb{P}(|\bar{X}_{1000} - 0.5| > 0.1) = 2 \cdot 10^{-10}$.

Applicazione

Supponiamo di avere un “tipo” di v.a. (ovvero una legge) con valore atteso μ , e facciamo n osservazioni indipendenti di quel tipo. Significa che consideriamo n v.a. i.i.d. X_1, \dots, X_n .

La legge dei grandi numeri dice che il valore osservato per \bar{X}_n con grande probabilità è vicino a μ .

Applicazione

Se non conosco μ , posso stimarla con \bar{X}_n .

Se n è grande, la probabilità che i due valori siano “molto diversi” è quasi zero.

Media teorica \approx media osservata

In pratica confrontiamo μ , media teorica di ciascuna osservazione, con \bar{X}_n media osservata.

La legge dei grandi numeri rende precisa l'affermazione

$$\mu \approx \bar{X}_n.$$

Nella pratica μ non si conosce, ma si stima con la media di (molti) esperimenti.

La casalinga di Voghera

Ha ragione la nostra casalinga a dire che

frequenza con cui osservo A in n esperimenti $\xrightarrow{n \rightarrow \infty} \mathbb{P}(A)$?

Sì, ce lo dice la legge dei grandi numeri!

Infatti se A è un evento di cui non conosco la probabilità, posso fare n esperimenti e porre $X_i = 1$ se nell'esperimento i -esimo si è verificato A (=il caso ha pescato un caso da A), $X_i = 0$ altrimenti.

Allora le v.a. X_i sono $\mathcal{B}(p)$ dove $p = \mathbb{P}(A) = \mathbb{E}(X_1)$.

La casalinga di Voghera

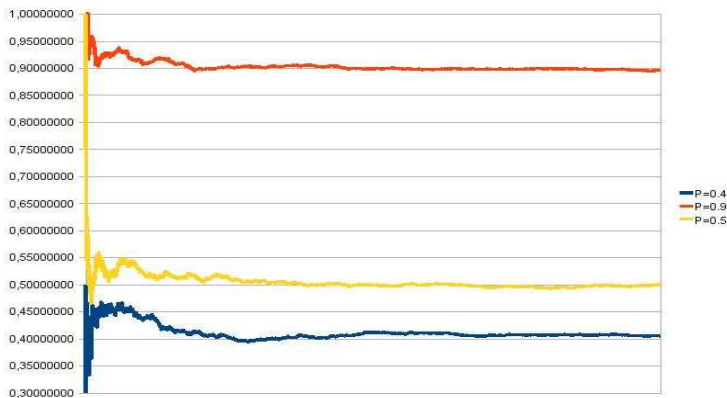
La legge dei grandi numeri dice che (nel senso dell'enunciato rigoroso...)

$$\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n} \xrightarrow{n \rightarrow \infty} p = \mathbb{E}(X_1).$$

Ma la somma delle X_i è il numero di volte che ho osservato A nei miei esperimenti, dunque \bar{X}_n è la frequenza relativa di A e per n grande risulta vicina a p , con grande probabilità.

Esperimento ripetuto

Nel seguente grafico sono rappresentate 3 successioni finite $\left\{ \frac{\sum_{i=1}^n X_i}{n} \right\}_{n=1, \dots, 4000}$ derivanti da esperimenti di Bernoulli con parametri 0.4, 0.5 e 0.9.



Un altro “uso comune” della LGN

Nel linguaggio comune si sente dire “per la legge dei grandi numeri” per intendere che anche una cosa improbabile, se si fa un numero grande di prove, capiterà.

Esempio. “Gioco una cinquina al lotto. So che la probabilità di vittoria è bassa, ma se gioco su tutte le ruote tutte le settimane, *per la legge dei grandi numeri*, prima o poi vincerò”.

Queste situazioni andrebbero trattate con il modello geometrico, ma anche la LGN ci dà una comprensione (anche se un po' empirica).

Un altro “uso comune” della LGN

Infatti sia A il mio evento “improbabile” (probabilità p piccola): per la LGN

$$\frac{\text{numero di volte in cui si verifica } A \text{ in } n \text{ esperimenti}}{n} \approx p.$$

Da cui

numero di volte in cui si verifica A in n esperimenti $\approx np$,

e se $n \geq \frac{1}{p}$ ne segue che $np \geq 1$ e il numero di volte in cui si verifica A in n esperimenti è almeno 1 (con grande probabilità).