1 Convergenza e approssimazione

Definizione 1
Sia X una v.a. e $\{X_n\} = \{X_1, X_2, \dots\}$ una successione di v.a. Si dice che $X_n \longrightarrow X$ quasi certamente ($X_n \to^{q.c.} X$) se $P(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$.

Definizione 2
Sia X una v.a. e $\{X_n\} = \{X_1, X_2, \dots\}$ una successione di v.a. Si dice che $X_n \longrightarrow X$ in probabilità ($X_n \to^{P} X$) se $\forall r > 0$, risulta $\lim_{n\to\infty} P(|X_n - X| > r) = 0$.
Ciò significa che per ogni $\epsilon > 0$ esiste un intero ν tale che, per $n > \nu$ risulta

$$P(|X_n - X| > r) < \epsilon.$$

Definizione 3
Sia X una v.a. e $\{X_n\} = \{X_1, X_2, \dots\}$ una successione di v.a. Si dice che $X_n \longrightarrow X$ in distribuzione ($X_n \to^{d} X$) se, per ogni punto x di continuità per F risulta $\lim_{n\to\infty} F_n(x) = F(x)$, dove F_n è la f.d.d. di X_n e F è la f.d.d. di X.

Osservazione Si può dimostrare che $X_n \to^{q.c.} X \;\Rightarrow\; X_n \to^{P} X$, ma non è vero il viceversa. Inoltre, se $X_n \to^{P} X$, allora $X_n \to^{d} X$.

Supponiamo di effettuare n lanci di una moneta per la quale si abbia che $P(\text{Testa}) = p \in (0, 1)$ in ogni lancio, e indichiamo con k il numero di Teste uscite. La quantità k/n rappresenta la proporzione di Teste ottenute negli n lanci. Se la moneta non è truccata (cioè $p = 1/2$) l'intuizione suggerisce che questa proporzione non si discosti troppo da $1/2$. Naturalmente è molto difficile che sia esattamente $k/n = 1/2$ (ciò vorrebbe dire che si sono ottenute esattamente tante Teste quante Croci), come pure è possibile che negli n lanci, per combinazione, si sia verificato un numero abnorme (molto grande o molto piccolo) di Teste, il che implicherebbe un valore della proporzione k/n distante da $1/2$. L'intuizione suggerisce ancora che, se n cresce, questo fenomeno dovrebbe tendere a sparire: se i primi lanci hanno dato una eccedenza di Teste, ciò dovrebbe poi essere compensato dai lanci successivi. In definitiva, se la moneta non è truccata, al crescere di n la proporzione di Teste uscite dovrebbe stabilizzarsi intorno al valore $1/2$; se poi la moneta è truccata e $p = P(\text{Testa}) \neq 1/2$, ci aspettiamo che la proporzione di Teste uscite si stabilizzi intorno al valore p. Questa situazione può essere modellizzata con una successione $\{X_1, X_2 \dots\}$ di v.a. indipendenti e Bernoulliane $B(1, p)$, dove supponiamo che $\{X_i = 1\}$ è l'evento "il lancio $i-$esimo ha dato Testa", $\{X_i = 0\}$ è l'evento "il lancio $i-$esimo ha dato Croce". Con questo modello il numero di Teste ottenute in n lanci è $X_1 + ... + X_n$ e la proporzione di Teste in n lanci (che abbiamo indicato prima con k/n) è

$$\bar{X}_n = \frac{X_1 + ... + X_n}{n}.$$

Dunque, se la moneta è equilibrata ci aspettiamo che, per n grande $\bar{X}_n$ assuma con piccola probabilità dei valori lontani da $1/2$; se la moneta è truccata ($p \neq 1/2$) per n grande $\bar{X}_n$ assumerà con piccola probabilità dei valori lontani da p. È un caso particolare di quanto afferma la *Legge dei Grandi Numeri*.

Theorem 1.1 (Legge dei grandi numeri) *Sia* $\{X_n\}$ *una successione di v.a. indipendenti ed equidistribuite, con media* μ *e varianza* σ^2. *Allora, detta* $S_n = \frac{X_1+\ldots+X_n}{n}$, *si ha:*

$$1)\ S_n \to^{q.c} \mu \text{ (legge forte dei grandi numeri)}$$

$$2)\ S_n \to^{P} \mu \text{ (legge debole dei grandi numeri).}$$

Dim. Proviamo solo 2). Si ha:

$$E(S_n) = \frac{1}{n}E(X_1 + X_2 + \cdots + X_n) = \frac{1}{n}[E(X_1) + E(X_2) + \cdots + E(X_n)] = \frac{n\mu}{n} = \mu,$$

$$Var(S_n) = \frac{1}{n^2}Var(X_1+X_2+\cdots+X_n) = \frac{1}{n^2}[Var(X_1)+Var(X_2)+\cdots+Var(X_n)] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n};$$

Per la disuguaglianza di Chebicev:

$$P(|S_n - \mu| > r) \leq \frac{Var(S_n)}{r^2} = \frac{\sigma^2}{nr^2} \to 0, \text{ per } n \to \infty.$$

Dunque, la legge (debole) dei grandi numeri (LGN) afferma che, per ogni $r > 0$:

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > r\right) \to 0, \text{ per } n \to \infty.$$

Esempio Si lancia una moneta e non si sa se essa sia equilibrata o meno. La LGN (debole) fornisce una stima per la probabilità di ottenere Testa in un singolo lancio. Infatti, se poniamo $X_i = 1$, se l'$i-$esimo lancio dà Testa, 0, altrimenti, allora le $X_i \sim B(1, p)$ e $X_1 + X_2 + \cdots + X_n$ conta il numero delle Teste uscite in n lanci e

$$S_n = (X_1 + X_2 + \cdots + X_n)/n \to^{P} E(X_i) = p.$$

In pratica, comunque, è possibile fare solo un numero finito di lanci, e quindi è utile calcolare l'errore che si commette stimando p con un n fissato. Si ha, per $r > 0$:

$$P\left(\left|S_n - p\right| > r\right) \leq \frac{Var(S_n)}{r^2} = \frac{1}{r^2}\ \frac{1}{n^2} \cdot np(1-p) = \frac{p(1-p)}{nr^2};$$

siccome $p(1-p) \leq 1/4,\ \forall p \in (0,1)$, otteniamo:

$$P\left(\left|S_n - p\right| > r\right) \leq \frac{1}{4}\frac{1}{nr^2}.$$

Ad esempio, se $n = 50$, la probabilità che p disti da S_n per più di $r = 0.1$ è $\leq \frac{1}{4\cdot 0.01} \cdot \frac{1}{50} = 0.5$; se $n = 1000,\ P\left(|S_n - p| > 0.1\right) \leq \frac{1}{4\cdot 0.01} \cdot \frac{1}{1000} = 0.025$.

Esempio Usare la disuguaglianza di Chebicev per mostrare che la probabilità che in n lanci di un dado non truccato il numero dei 6 usciti è compreso tra $\frac{1}{6}n - \sqrt{n}$ e $\frac{1}{6}n + \sqrt{n}$ è almeno $\frac{31}{36}$.
Si ha:

se X_i vale 1 se esce 6, e 0 altrimenti, allora $X = X_1 + X_2 + \cdots + X_n \sim B(n, \frac{1}{6})$; $E(X) = n/6$, $Var(X) = \frac{n}{6} \cdot \frac{5}{6} = \frac{5n}{36}$. Quindi:

$$P\left(\frac{1}{6}n - \sqrt{n} \le X \le \frac{1}{6}n + \sqrt{n}\right) = P\left(-\sqrt{n} \le X - \frac{1}{6}n \le \sqrt{n}\right)$$

$$= P(|X - E(X)| \le \sqrt{n}) = 1 - P(|X - E(X)| > \sqrt{n}),$$

che, per la disuguaglianza di Chebicev è

$$\ge 1 - \frac{1}{(\sqrt{n})^2} \cdot \frac{5n}{36} = 1 - \frac{5}{36} = \frac{31}{36}\ .$$

Pertanto, la probabilità cercata è $\ge \frac{31}{36}$.

Proposition 1.2 *Se $X_n \to^P X$, allora $aX_n + b \to^P aX + b$, $\forall a, b \in \mathbb{R}$.*

Dim. Se $r > 0$, si ha:

$$P(|aX_n + b - aX - b| > r) = P(|a||X_n - X| > r) = P(|X_n - X| > r') \to 0, \text{ per } n \to \infty,$$

dove abbiamo posto $r' = r/|a|$. Il fatto che $P(|X_n - X| > r') \to 0$ è conseguenza dell'ipotesi, cioè $X_n \to^P X$.

Osservazione Sia $\{X_n\}$ una successione di v.a. indipendenti ed ugualmente distribuite, con media μ e varianza σ^2 finite; la LGN permette di stimare la media μ, visto che $\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) \to^P \mu$.
Ci proponiamo ora di stimare la varianza σ^2.
Poniamo

$$\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2.$$

Sviluppando il quadrato, troviamo:

$$\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}_n X_i + \bar{X}_n^2\right) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n\,\frac{1}{n}\sum_{i=1}^{n} X_i + \bar{X}_n^2\frac{1}{n}\sum_{i=1}^{n} 1$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n \cdot \bar{X}_n + \bar{X}_n^2 \cdot 1 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n^2 + \bar{X}_n^2 = \left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - \bar{X}_n^2.$$

Per la LGN $\bar{X}_n \to^P \mu = E(X_1)$ e quindi $\bar{X}_n^2 \to^P \mu^2 = E^2(X_1)$; la LGN afferma anche che, se le v.a. $Y_i := X_i^2$ hanno varianza finita, ovvero se X_i ha momento di ordine 4 finito, allora:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i \to^P E(Y_1) = E(X_1^2).$$

Concludendo:

$$\bar{\sigma}_n^2 \to^P E(X_1^2) - E^2(X_1) = Var(X_1) = \sigma^2,$$

cioè, per n grande $\bar{\sigma}_n^2$ assume valori prossimi a σ^2 con grande probabilità.

In questi calcoli, in effetti abbiamo utilizzato, senza dimostrarle, alcune proprietà della convergenza in probabilità; ad esempio, se $X_n \to^P a$ e $Y_n \to^P b$, con $a, b \in \mathbb{R}$, allora $X_n + Y_n \to^P a + b$ e $X_n \cdot Y_n \to^P ab$.

Metodo Montecarlo
Sia $f : [0,1] \longrightarrow \mathbb{R}$ una funzione (deterministica) continua e limitata, e $\{X_n\}$ una successione di v.a. indipendenti ed uniformemente distribuite in $[0,1]$. Allora, $\{Y_n\} = \{f(X_n)\}$ è una successione di v.a. indipendenti con media finita

$$E(Y_n) = E[f(X_n)] = E[f(X_1)] = \int_0^1 f(x)dx$$

e varianza finita

$$Var(Y_n) = Var[f(X_n)] = Var[f(X_1)] = E(Y_1^2) - E^2(Y_1) = \int_0^1 f^2(x)dx - \left(\int_0^1 f(x)dx\right)^2,$$

essendo f continua e limitata. Allora, per la LGN applicata alla successione $\{Y_n\} = \{f(X_n)\}$, otteniamo che:

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \to^P \ E(Y_1) = E[f(X_1)] = \int_0^1 f(x)dx,$$

cioè

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \to^P \ \int_0^1 f(x)dx.$$

Ciò suggerisce un algoritmo numerico stocastico per calcolare l'integrale di una funzione $f(x)$.
1) Si generano $X_1, X_2, \ldots$ numeri aleatori (pseudo) uniformi in $[0,1]$ (tali generatori esistono, anche sulle calcolatrici tascabili).
2) Allora, la quantità

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) = \frac{1}{n}(f(X_1) + f(X_2) + \cdots + f(X_n))$$

è un'approssimazione, per n grande dell' integrale

$$\int_0^1 f(x)dx.$$

Naturalmente, lo stesso discorso si può ripetere con ovvie modifiche per stimare l'integrale di $f(x)$ su un intervallo $[a,b]$.

Riportiamo qui sotto uno pseudocodice per stimare, col metodo Montecarlo, il valore di $\int_0^1 x^2\ dx = 1/3$ (valore teorico). Utilizziamo la funzione $rand()$ che fornisce un numero pseudorandom con distribuzione uniforme in $(0,1)$, e l'algoritmo descritto sopra con $n = 2000$.

// Stima dell'integrale tra 0 e 1 di $f(x) = x^2$
// il valore teorico è 1/3,

$n = 2000;$
$s = 0;$
$for\ i = 1 : n$
$x = rand();$
$y = x * x;$
$s = s + y;$
end
$integrale = s/n$

2 Il Teorema limite centrale

Theorem 2.1 *Si $\{X_n\}$ una successione di v.a. indipendenti ed equidistribuite, con la stessa media $E(X_i) = \mu$ e la stessa varianza $Var(X_i) = \sigma^2$. Allora, per ogni $z \in \mathbb{R}$ si ha:*

$$P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq z\right) \longrightarrow \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx, \text{ per } n \to \infty.$$

Ciò significa che $S_n^* = \frac{X_1+X_2+\cdots+X_n-n\mu}{\sigma\sqrt{n}}$ è asintoticamente $\mathcal{N}(0,1)$, per $n \to \infty$; in altre parole, la f.d.d. di S_n^*, ovvero $P(S_n^* \leq z)$, si può approssimare, per n grande, con la f.d.d. $\Phi(z)$ di una v.a. $Z \sim \mathcal{N}(0,1)$, dove $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$.

Esempio 1 Una lampada ha un tempo di vita che segue una legge esponenziale di media $\mu = 10$ (giorni). Non appena la lampada smette di funzionare, viene sostituita con una nuova. Qual è la probabilità che 40 lampade siano sufficienti in un anno?

Sia X_i la durata della i-esima lampada; possiamo supporre che le v.a. X_i siano indipendenti, con distribuzione esponenziale di parametro $\lambda = 1/\mu = 1/10$ (l'unità di misura di tempo è il giorno); pertanto $\sigma^2 = Var(X_i) = 1/\lambda^2 = 100$. La probabilità cercata è:

$$P(X_1 + X_2 + \cdots + X_{40} \geq 365) = 1 - P(X_1 + X_2 + \cdots + X_{40} < 365);$$

Ma:

$$P(X_1 + X_2 + \cdots + X_{40} < 365) = P\left(\frac{X_1 + X_2 + \cdots + X_{40} - 40\mu}{\sigma\sqrt{40}} < \frac{365 - 40\mu}{\sigma\sqrt{40}}\right)$$

$$= P\left(\frac{X_1 + X_2 + \cdots + X_{40} - 40\mu}{10\sqrt{40}} < \frac{365 - 40 \cdot 10}{10\sqrt{40}}\right)$$

e, per il TLC, tale probabilità si può approssimare con

$$\Phi\left(\frac{365-400}{10\sqrt{40}}\right) = \Phi(-0.55) = 1-\Phi(0.55),$$

dove $\Phi(x)$ è la f.d.d. di una v.a. Gaussiana standard. Riprendendo il calcolo, una approssimazione della probabilità richiesta è:

$$P(X_1+X_2+\cdots+X_{40}\geq 365) \approx 1-[1-\Phi(0.55)] = \Phi(0.55) = 0.71\ .$$

Si osservi, comunque che il TLC vale per $n\to\infty$, mentre qui $n=40$ è finito! Sicuramente, questo tipo di approssimazione è tanto migliore, quanto n è grande.
In ogni caso, per il teorema di addizione di v.a. Gamma indipendenti, risulta $X_1+X_2+\cdots+X_{40}\sim\Gamma(40,1/10)$, per cui, detta F la f.d.d. di una v.a. $Gamma(40,1/10)$, si ha esattamente $P(X_1+X_2+\cdots+X_{40}\geq 365) = 1-F(365)$. In realtà, come abbiamo visto, siccome in questo caso $\alpha = m$ è intero (si ha $m=40$) si conosce la forma esplicita della f.d.d, che è:

$$F(x) = 1-e^{-\lambda x}\sum_{i=0}^{m-1}\frac{(\lambda x)^i}{i!},$$

per cui risulta infine, esattamente (senza approssimazione):

$$P(X_1+X_2+\cdots+X_{40}\geq 365) = 1-\left(1-e^{-365/10}\sum_{i=0}^{39}\frac{(365/10)^i}{i!}\right) = e^{-36.5}\sum_{i=0}^{39}\frac{(36.5)^i}{i!}.$$

Se si effettua il calcolo dell'ultima quantità qui sopra mediante un computer, si ottiene il valore 0.6975... che è il valore esatto della $P(X_1+X_2+\cdots+X_{40}\geq 365)$; come si vede, questo differisce di poco dal valore approssimato, 0.71 ottenuto con molta meno fatica (si consideri che per il calcolo esatto, occorre sommare 40 frazioni che hanno numeratori grandi e scomodi da calcolare, e denominatori altrettanto grandi; la somma ottenuta deve essere infine moltiplicata per $e^{-36.5}$).

Esempio 2 Si lanci 100 volte un dado truccato in modo tale che sia 1/3 la probabilità che esca 2. Usando il Teorema Limite Centrale, stimare la probabilità che esca al più 20 volte il numero 2
(i) senza utilizzare la correzione di continuità;
(ii) utilizzando la correzione di continuità.

Soluzione Siano X_i, $i=1,\ldots,100$ tali che $X_i=1$ se esce 2 all'i–esimo lancio, $X_i=0$ se non esce 2 all'i–esimo lancio. Si ha $E(X_i)=1/3$ e $Var(X_i)=\sigma^2=\frac{1}{3}\cdot\frac{2}{3}$. Indichiamo con $S_{100}=X_1+\cdots+X_{100}$. Allora $S_{100}\sim B(100,1/3)$ e quindi $E[S_{100}]=100\cdot 1/3=100/3$ e $Var[S_{100}]=100\cdot 1/3\cdot 2/3=200/9$.
(i) Sia $W\sim\mathcal{N}(0,1)$; utilizzando il Teorema limite centrale senza la correzione di continuità si ottiene:

$$P(X_1+X_2+\cdots+X_{100}\leq 20) = P\left(S_{100}\leq 20\right)$$
$$= P\left(\frac{X_1+X_2+\cdots+X_{100}-100\cdot 1/3}{\sigma\sqrt{100}}\leq\frac{20-100\cdot 1/3}{\sigma\sqrt{100}}\right)$$

$$= P\left(\frac{S_{100} - E(S_{100})}{\sqrt{Var(S_{100})}} \leq \frac{20 - E(S_{100})}{\sqrt{Var(S_{100})}}\right) =$$

$$P\left(\frac{S_{100} - 100/3}{10\sqrt{2}/3} \leq \frac{20 - 100/3}{10\sqrt{2}/3}\right) = P\left(\frac{S_{100} - 100/3}{10\sqrt{2}/3} \leq \frac{-40/3}{10\sqrt{2}/3}\right) \cong$$

$$P\left(W \leq \frac{-40/3}{10\sqrt{2}/3}\right) = \Phi(-2\sqrt{2}) =$$

$$= 1 - \Phi(2\sqrt{2}) = 1 - \Phi(2.83) = 1 - 0.99767 = 0.00233 \ .$$

(ii) Utilizziamo ora il Teorema limite centrale con la correzione di continuità; osserviamo che, siccome S_{100} assume soltanto valori interi, l'evento "$S_{100} \leq 20$ " coincide con l'evento "$S_{100} < 21$ ", o anche con l'evento "$S_{100} \leq 20 + \alpha$ ", con $\alpha \in (0, 1)$, e quindi $P(S_{100} \leq 20) = P(S_{100} \leq 20 + \alpha)$; per esempio, possiamo prendere $\alpha = 1/2$.
Il termine "approssimazione di continuità" si riferisce al fatto che, mentre S_{100} è una v.a. discreta a valori interi, la v.a. con la quale approssimiamo la sua distribuzione, grazie al TLC (cioé una v.a. Gaussiana standard), è invece una v.a. continua. Per cercare di diminuire la discrepanza che può esserci nell'approssimare la distribuzione di una v.a. discreta a valori interi con quella di una v.a. continua, si introduce appunto la cosiddetta "approssimazione di continuità".
Dunque, tenendo conto di quanto osservato sopra:

$$P\left(S_{100} \leq 20\right) = P\left(S_{100} \leq 20 + \frac{1}{2}\right) =$$

$$= P\left(\frac{S_{100} - E(S_{100})}{\sqrt{Var(S_{100})}} \leq 20.5 - E(S_{100})\right) =$$

$$= P\left(\frac{S_{100} - 100/3}{10\sqrt{2}/3} \leq \frac{20.5 - 100/3}{10\sqrt{2}/3}\right) =$$

$$= P\left(\frac{S_{100} - 100/3}{10\sqrt{2}/3} \leq \frac{-77/6}{10\sqrt{2}/3}\right) \cong P\left(W \leq \frac{-77/6}{10\sqrt{2}/3}\right) =$$

$$= \Phi\left(-\frac{77\sqrt{2}}{40}\right) = 1 - \Phi\left(\frac{77\sqrt{2}}{40}\right) =$$

$$= 1 - \Phi(2.72) = 1 - 0.99674 = 0.00326 \ .$$

Questo risultato è leggermente più preciso di quello ottenuto senza la correzione di continuità, cioé 0.00233 anche se la differenza è molto piccola.

3 Intervallo di confidenza (o fiducia) per la media incognita di una popolazione statistica

Sia $\{X_n\}$ una successione di v.a. indipendenti ed equidistribuite, con media μ e varianza σ^2, e sia $\bar{X}_n = (X_1 + X_2 + \cdots + X_n)/n$ la media aritmetica delle X_i. Allora, si ha:

$$P(|\bar{X}_n - \mu| > \alpha) = P(\{\bar{X}_n > \mu + \alpha\} \cup \{\bar{X}_n < \mu - \alpha\})$$

$$= P[X_1 + X_2 + \cdots + X_n > n(\mu + \alpha)] + P[X_1 + X_2 + \cdots + X_n < n(\mu - \alpha)]$$

che, per n grande, grazie all'approssimazione data dal TLC, vale circa

$$1 - \Phi\left(\frac{n(\mu+\alpha) - n\mu}{\sigma\sqrt{n}}\right) + \Phi\left(\frac{n(\mu-\alpha) - n\mu}{\sigma\sqrt{n}}\right) = 1 - \Phi\left(\frac{\alpha}{\sigma}\sqrt{n}\right) + \Phi\left(-\frac{\alpha}{\sigma}\sqrt{n}\right)$$
$$= \Phi\left(-\frac{\alpha}{\sigma}\sqrt{n}\right) + \Phi\left(-\frac{\alpha}{\sigma}\sqrt{n}\right) = 2\Phi\left(-\frac{\alpha}{\sigma}\sqrt{n}\right).$$

Quindi, se $\delta > 0$, otteniamo:

$$P(|\bar{X}_n - \mu| \leq \delta) \approx 1 - 2\Phi\left(-\frac{\delta}{\sigma}\sqrt{n}\right) = 1 - 2\left(1 - \Phi\left(\frac{\delta}{\sigma}\sqrt{n}\right)\right) = 2\Phi\left(\frac{\delta}{\sigma}\sqrt{n}\right) - 1.$$

Ora scegliamo $\delta = \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}$, dove $\phi_{1-\alpha/2}$ è il quantile di una Gaussiana standard di ordine $1 - \alpha/2$, cioè risulta $\Phi(\phi_{1-\alpha/2}) = 1 - \alpha/2$. Con questa scelta di δ, si ottiene quindi:

$$P\left(|\bar{X}_n - \mu| \leq \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}\right) \approx 2\Phi(\phi_{1-\alpha/2}) - 1 = 2(1 - \alpha/2) - 1 = 1 - \alpha.$$

Ma:

$$|\bar{X}_n - \mu| \leq \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2} \Leftrightarrow \mu \in I,$$

dove

$$I := \left[\bar{X}_n - \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2},\ \bar{X}_n + \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}\right], \tag{3.1}$$

quindi μ appartiene all'intervallo I con probabilità approssimativamente uguale a $1 - \alpha$. L'intervallo I si chiama *intervallo di confidenza di livello* $1-\alpha$ *per la media* μ. Allo scopo di individuare la posizione della media μ, l'intervallo I di confidenza ottenuto è tanto migliore, quanto minore è la sua ampiezza.

Esempio 1 Si esamina un campione di $n = 100$ componenti elettronici e si trova che la media campionaria del tempo di vita dei componenti è $\bar{x}_n = (x_1 + x_2 + \cdots + x_n)/n = 50$ (mesi). Ipotizzando che il tempo di vita di un componente del campione sia una v.a. X con varianza $\sigma^2 = 144$ e media μ (incognita), si trovi un intervallo di confidenza al livello $1 - \alpha = 0.95$ per la media incognita μ.

Soluzione. Per la (3.1) un intervallo I di confidenza a livello $1 - \alpha$ per la media incognita di una distribuzione avente varianza σ^2, è:

$$I = \left[\bar{x}_n - \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2},\ \bar{x}_n + \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}\right] \tag{$*$}$$

dove $\bar{x}_n$ è la media campionaria e ϕ_β è il quantile della Gaussiana standard, tale che $\Phi(\phi_\beta) = \beta$. Nel caso in esame, si ha $n = 100,\ \sigma = 12$, e la media campionaria è $\bar{x}_n = 50$; inoltre, da $1 - \alpha = 0.95$ segue $1 - \alpha/2 = 0.975$; dalla tavola dei valori di Φ si ricava $\Phi(1.96) = 0.975$, dunque $\phi_{1-\alpha/2} = 1.96$. Sostituendo in $(*)$, si ottiene che un intervallo di confidenza per la media μ di X, al livello 0.95 è:

$$I = \left[50 - \frac{12}{10} \cdot 1.96,\ 50 + \frac{12}{10} \cdot 1.96\right] = [47.648,\ 52.352]\ .$$

Ciò significa che con probabilità abbastanza alta (≥ 0.95) la media incognita μ si trova situata nell'intervallo $[47.648,\ 52.352]$.

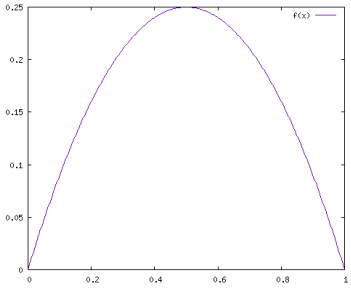

Figure 1: Grafico della funzione $\sigma^2(p) = p(1-p)$; il massimo si ottiene per $p = 1/2$ e vale $1/4$.

Esempio 2 (quando la varianza non è nota)
Un campione di 100 transistor viene estratto da una grossa fornitura e testato per rilevare eventuali imperfezioni. Si trova che 80 pezzi superano il controllo. Calcolare un intervallo di confidenza a livello $1-\alpha = 0.95$ per la percentuale p di transistor della fornitura accettabili.

Soluzione. Sia $X_i,\ i = 1, \ldots, n$ la v.a. che vale 1 se l'i–esimo transistor è privo di difetti, e 0 altrimenti. Le v.a. X_i sono indipendenti e Bernoulliane di parametro p incognito (quindi anche la media $\mu = p$ e la varianza $\sigma^2 = p(1-p)$ sono incognite); $X = X_1 + X_2 + \cdots + X_n$ rappresenta il numero totale di transistor accettabili (cioè privi di difetti). Il problema fornisce per la media campionaria di transistor accettabili il valore $\bar{x} = (x_1 + x_2 + \cdots + x_n)/n = 0.8$, ottenuta dagli n transistor testati. Per la solita formula (3.1), sappiamo che un intervallo I di confidenza a livello $1-\alpha$ per la media incognita di una distribuzione avente varianza σ^2, è:

$$I = \left[\bar{x} - \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}, \bar{x} + \frac{\sigma}{\sqrt{n}}\phi_{1-\alpha/2}\right] \qquad (*)$$

dove $\bar{x}$ è la media campionaria, e ϕ_β è il quantile della Gaussiana standard, tale che $\Phi(\phi_\beta) = \beta$. Nel caso in esame, si ha $n = 100,\ \bar{x} = 0.8$, ma σ è **incognita**. Da $1-\alpha = 0.95$ segue $1-\alpha/2 = 0.975$, e quindi dalla tavola dei valori di Φ si ricava $\phi_{1-\frac{\alpha}{2}} = 1.96$. Sostituendo in $(*)$, si ottiene l'intervallo:

$$I(\sigma) = \left[0.8 - \frac{\sigma}{10} \cdot 1.96,\ 0.8 + \frac{\sigma}{10} \cdot 1.96\right].$$

Poiché $\sigma = \sqrt{p(1-p)} \leq 1/2,\ \forall p \in [0,1]$, l'intervallo $I(\sigma)$ è certamente contenuto in

$$I = \left[0.8 - \frac{0.5}{10} \cdot 1.96,\ 0.8 + \frac{0.5}{10} \cdot 1.96\right] = [0.702,\ 0.898]\ ,$$

che è l'intervallo di confidenza cercato. Si noti che l'intervallo trovato ha un'ampiezza eventualmente maggiore di quella che si sarebbe trovata se σ fosse stata nota, avendo dovuto fare una maggiorazione.

In realtà, in casi come questo è possibile trovare una stima più stretta per l'intervallo di confidenza, utilizzando la distribuzione di Student (si veda il libro di testo, o l'esercizio 3.19 del libro di esercizi), che non tratteremo per esigenza di tempo.