Teoremi limite

Enrico Ferrero

27 febbraio 2007

LA DISEGUAGLIANZA DI CHEBYCHEV

Sia X una variabile aleatoria avente valore medio μ e varianza σ^2.

Sia poi K un arbitrario numero positivo. È possibile dimostrare che la probabilità di avere $|X-\mu| \geq K\sigma$ è sempre minore di $1/K^2$.

$$P[|X-\mu| \geq K\sigma] \leq \frac{1}{K^2}$$

Questa è relazione nota come la diseguaglianza di Chebychev.

Dimostriamo allora tale diseguaglianza limitandoci al caso delle distribuzioni continue.

Se $K < 1$ ciò è ovviamente verificato, essendo la probabilità $P \leq 1$.

Dobbiamo quindi rivolgere la nostra attenzione ai valori che può assumere la variabile aleatoria X per cui $|x - \mu| \geq K\sigma$ cioè per tutti quei valori di x fuori dall'intervallo $\mu \pm K\sigma$ tali da verificare la condizione

$$(x - \mu)^2 \geq K^2\sigma^2$$

Calcoliamo la varianza σ^2

$$\sigma^2 = \int_{-\infty}^{\mu-K\sigma} (x-\mu)^2 f(x)dx + \int_{\mu-K\sigma}^{\mu+K\sigma} (x-\mu)^2 f(x)dx + \int_{\mu+K\sigma}^{\infty} (x-\mu)^2 f(x)dx$$

Limitiamoci a considerare i valori di x tali che $(x-\mu)^2 \geq K^2\sigma^2$

$$\sigma^2 \geq K^2\sigma^2 \left[\int_{-\infty}^{\mu-K\sigma} f(x)dx + \int_{\mu+K\sigma}^{\infty} f(x)dx\right] + \int_{\mu-K\sigma}^{\mu+K\sigma} (x-\mu)^2 f(x)dx$$

Poichè

$$\int_{\mu-K\sigma}^{\mu+K\sigma} (x-\mu)^2 f(x)dx \geq 0$$

segue che

$$\sigma^2 \geq K^2\sigma^2 \left\{ \int_{-\infty}^{\mu-K\sigma} f(x)dx + \int_{\mu+K\sigma}^{\infty} f(x)dx \right\}$$

ovvero

$$P\left[|X-\mu| \geq K\sigma\right] \leq \frac{1}{K^2}$$

essendo per definizione

$$P\left[|X-\mu| \geq K\sigma\right] = \left\{ \int_{-\infty}^{\mu-K\sigma} f(x)dx + \int_{\mu+K\sigma}^{\infty} f(x)dx \right\}$$

La diseguaglianza di Chebychev fornisce un limite superiore alla probabilità che X scarti da μ, K volte σ.

In genere tale limite è molto più grande del valore effettivo di P che dipende dalla distribuzione.

Si noti che tale limitazione al valore della probabilità è indipendente dal tipo di distribuzione.

Ad esempio se poniamo K = 2

$$P[|X - \mu| \geq 2\sigma] \leq 0.25$$

(che implica $P[|X - \mu| < 2\sigma] \geq 0.75$)

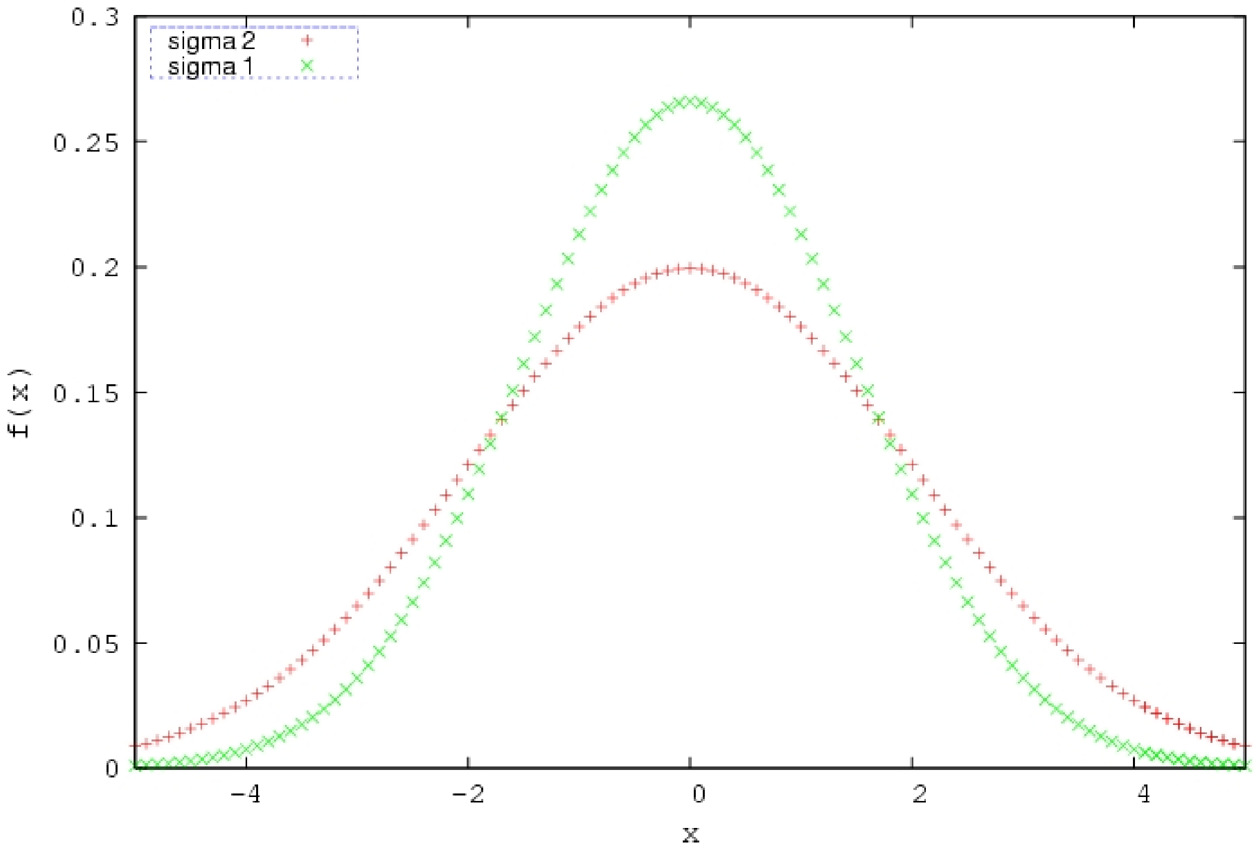
- Considerando il caso della distribuzione normale, ci rendiamo conto immediatamente che, a parità di deviazione standard σ, la probabilità che X sia fuori dall'intervallo $\mu \pm K\sigma$ è minore tanto più grande è K.

- X deve assumere valori sempre più lontani da μ e che quindi cadono sempre più sulle code della curva di Gauss.

- A parità di K, al diminuire della deviazione standard σ, la curva a campana di Gauss si restringe progressivamente, così che per avere un eguale valore di probabilità delle code si devono considerare intervalli di ampiezza diversa.

- In altre parole se considero due distribuzioni gaussiane aventi $\sigma_1 < \sigma_2$, allora X risulta molto più probabilmente concentrata attorno al valore medio μ nel caso della distribuzione con σ_1.

$$\sigma_1 < \sigma_2$$



LA LEGGE DEI GRANDI NUMERI

- Questa legge è connessa alla osservazione empirica (assunta dagli oggettivisti come definizione stessa di probabilità) che, detto N il numero di prove, la frequenza relativa di un evento tende alla probabilità quando N tende all'infinito.

- Questa osservazione, che vogliamo esprimere ora in termini matematici è anche nota come *legge empirica del caso*.

- In altre parole, si può dire che i fenomeni aleatori mostrano un certa regolarità esprimibile secondo una legge matematica:

  qualunque sia il fenomeno trattato ogni singolo risultato è casuale ed imprevedibile ma non influenza la distribuzione di probabilità dell'insieme di tutti i risultati, le deviazioni dal valor medio, inevitabili in ogni singolo esperimento, nel complesso di tutti gli esperimenti si compensano e la distribuzione di probabilità rimane stabile.

- Quando il numero di esperimenti è grande il valor medio cessa di essere casuale e può essere predetto con grande precisione.

Sia allora S_N una variabile aleatoria (ad esempio il numero di successi in N prove indipendenti); S_N esprime il possibile risultato di N tentativi a ciascuno dei quali associamo una variabile aleatoria X_i. Avremo allora

$$S_N = \sum_{i=1}^{N} X_i$$

X_i sono variabili aleatorie indipendenti aventi la stessa distribuzione di probabilità e quindi con lo stesso valor medio $E[X_i] = \mu$ e varianza $Var[X_i] = \sigma^2$.

Definiamo la frequenza relativa di successo in queste N prove nel modo seguente

$$\frac{S_N}{N} = \textit{frequenza relativa di successo}$$

ed il suo valore atteso è

$$E\left[\frac{S_N}{N}\right] = \frac{1}{N}\sum_{i=1}^{N} E[X_i] = \mu$$

e, poichè le X_i sono indipendenti

$$Var\left[\frac{S_N}{N}\right] = \frac{1}{N^2}\sum_{i=1}^{N} Var[X_i] = \frac{\sigma^2}{N}$$

Sia ora ϵ una quantità positiva arbitrariamente piccola; invocando la diseguaglianza di Chebychev, possiamo affermare che

$$P[|\frac{S_N}{N} - \mu| \geq \epsilon] \leq \frac{\sigma^2}{N\epsilon^2}$$

avendo posto

$$\left(\epsilon = K\frac{\sigma}{\sqrt{N}}\right)$$

In altri termini questo significa che la probabilità che $\frac{S_N}{N}$ si discosti da μ di ϵ, è tanto più piccola quanto più è alto il numero di prove N.

Quindi la variabile aleatoria $\frac{S_N}{N}$ converge in probabilità verso il suo valor medio quando N è sufficientemente grande

$$\lim_{N\to\infty} P[|\frac{S_N}{N} - \mu| \geq \epsilon] = 0$$

Nello stesso limite la frequenza relativa di successo tende al suo valore atteso

$$\lim_{N\to\infty} \frac{S_N}{N} = \mu$$

Per esempio consideriamo il lancio di una moneta, la frequenza relativa di successo è data dal rapporto tra il numero di volte che esce, per esempio, testa e il numero di lanci. Se il numero di lanci è molto grande questa tende al valore 1/2.

IL TEOREMA DEL LIMITE CENTRALE

- Il teorema del limite centrale asserisce che la distribuzione associata ad una variabile aleatoria S_N, somma di N variabili aleatorie indipendenti X_i, tende alla *distribuzione normale* al crescere di N, qualunque sia la distribuzione delle N variabili, purchè nessun termine della sommatoria sia predominante rispetto agli altri.

- Il teorema del limite centrale è l'estensione del *Teorema di De Moivre-Laplace* secondo il quale la distribuzione Binomiale può essere approssimata da una Gaussiana al crescere di N.

In particolare affinchè sia valido questo teorema, è condizione *sufficiente*, ma *non necessaria*, che le variabili X_i abbiano tutte la stessa distribuzione di probabilità con valore medio μ e varianza σ^2.

$$\lim_{N\to\infty} P[\mu + \sigma x_1 \leq S_N \leq \mu + \sigma x_2] = \int_{x_1}^{x_2} N(x)dx$$

dove N(x) e una distribuzione normale ridotta, cioè con $\mu = 0$ e $\sigma = 1$

Poichè per ipotesi $S_N = \sum_i X_i$ è definita da variabili X_i aventi tutte la stessa distribuzione (quindi stessi σ_i e μ_i) , possiamo scrivere

$$\mu = E[S_N] = \sum_{i=1}^{N} E[X_i] = NE[X_i] = N\mu_i$$

e, poichè le X_i sono indipendenti

$$\sigma^2 = \sum_{i=1}^{N} \sigma_i^2 = N\sigma_i^2$$

si dimostra che la variabile

$$S_N^* = \frac{S_N - \mu}{\sigma} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{X_i - \mu_i}{\sigma_i}$$

ha distribuzione normale ridotta e quindi la variabile aleatoria S_N ha distribuzione normale del tipo $n(x, \mu, \sigma)$

Nel caso di distribuzione Binomiale si ha:

$$\lim_{N\to\infty} P[np + \sqrt{npq}x_1 \le S_N \le np + \sqrt{npq}x_2] = \int_{x_1}^{x_2} N(x)dx$$

che rappresenta il *teorema di De Moivre-Laplace*

e per la distribuzione di Poisson

$$\lim_{N\to\infty} P[\mu + \sqrt{\mu}x_1 \le S_N \le \mu + \sqrt{\mu}x_2] = \int_{x_1}^{x_2} N(x)dx$$

Nei casi che si presentano in pratica N assume un valore finito, pertanto la funzione di distribuzione della variabile aleatoria oggetto di studio è solamente prossima a quella di Gauss:

$$f(x) \approx \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Questa approssimazione sarà tanto più buona quanto più N è grande.

Si noti che la variabile S_N rappresenta la media campionaria a meno della costante 1/N:

$$\bar{X} = \frac{1}{N} S_N = \frac{1}{N} \sum_i X_i$$

quindi ciò che si è detto per S_N è valido anche per la media, fatti salvi alcuni aggiustamenti di costanti moltiplicative:

la variabile $\frac{\bar{X}-\mu_i}{\sigma_i/\sqrt{N}}$ converge in distribuzione a una variabile normale ridotta avente valore atteso 0 e varianza 1.

Supponiamo allora che x_i sia il risultato di una misura della variabile aleatoria X. $E[X] = \mu$ rappresenta il valor vero della grandezza. Allora l'errore $\Delta X = X - \mu$ può essere considerato a sua volta come una variabile aleatoria determinata da tante cause indipendenti e definibile quindi come somma di processi indipendenti ΔX_i

$$\Delta X = \sum_i \Delta X_i$$

Questa e la ragione per cui sovente si assume che la distribuzione di probabilità degli errori casuali sia una Gaussiana a valor medio nullo.