Ricordiamo i concetti chiave per un uso avvertito della

LEGGE DEI GRANDI NUMERI

(ad esempio E. Ventsel, Teoria delle probabilità, cap XIII)

Richiamiamo innanzitutto la DISUGUAGLIANZA di CEBYSEV.

Date SPERANZA MATEMATICA $M[X] = m_x = \int_{-\infty}^{\infty} x f(x)\,dx$

e VARIANZA $D[X] = D_x = \int_{-\infty}^{\infty} (x-m_x)^2 f(x)\,dx$

e un qualsiasi $\alpha > 0$

LA PROBABILITA' che X devii dalla sua SPERANZA MATEMATICA non meno di α è LIMITATA da $\frac{D_x}{\alpha^2}$

$$\boxed{P(|X-m_x| \geq \alpha) \leq \frac{D_x}{\alpha^2}}$$

Infatti $P(|X-m_x|>\alpha) = \int_{|x-m_x|>\alpha} f(x)\,dx$

$$D_x = \int_{-\infty}^{\infty} (x-m_x)^2 f(x)\,dx = \int_{-\infty}^{\infty} |x-m_x|^2 f(x)\,dx \geq \int_{|x-m_x|>\alpha} |x-m_x|^2 f(x)\,dx$$

e dunque $D_x \left(\geq \int_{|x-m_x|>\alpha} |x-m_x|^2 f(x)\,dx \right) \geq \alpha^2 \int_{|x-m_x|>\alpha} f(x)\,dx = \alpha^2 P(|X-m_x|>\alpha)$

N.B. La disuguaglianza di Cebysev dà solo un LIMITE SUPERIORE!

ES. $P(|X-m_x| \geq 3\sigma_x) \leq \frac{D_x}{9\sigma_x^2} = \frac{1}{9}$

mentre PER LA GAUSSIANA questa probabilità è $\sim 3 \cdot 10^{-3}$!

Dalla disuguaglianza si ottiene il TEOREMA di CEBYSEV
(LEGGE dei GRANDI NUMERI)

Prima di tutto... Sia X una VARIABILE ALEATORIA con $M[X] = m_x$
e $D[X] = D_x$

Effettuiamo n PROVE e registriamo i valori $X_1, X_2, \dots, X_n$.

Consideriamo la MEDIA ARITMETICA $Y = \frac{1}{n} \sum_{i=1}^{n} X_i$

Essa é una FUNZIONE LINEARE delle VARIABILI INDIPENDENTI X_i (!)

e dunque la sua SPERANZA MATEMATICA vale

$$m_y = M[Y] \underset{(!)}{=} \left(\sum_i \int_{-\infty}^{\infty} x_i f(x_i)\, dx_i \right) \frac{1}{n} = \frac{1}{n} \sum_i M[X_i] = \frac{1}{n} n m_x = m_x$$

mentre la sua VARIANZA

$$D_y = D[Y] = \frac{1}{n^2} \sum_{i=1}^{n} D[X_i] = \frac{D_x}{n} \quad (*)$$

cioé
$$\boxed{\begin{aligned} m_y &= m_x \\ D_y &= \frac{D_x}{n} \end{aligned}}$$

(*) Attenzione! Provate a fare il conto! Troverete che ad esempio

$$D[X+Y] = D[X] + D[Y] + 2K_{xy}$$

↳ COVARIANZA $K_{xy} = \int_{-\infty}^{\infty} (x-m_x)(y-m_y) f(x,y)\, dx\, dy$

ma $K_{xy} = 0$ per VARIABILI INDIPENDENTI

Nota inoltre che (X_i indipendenti) $D\left[\sum_i a_i X_i\right] = \sum_i a_i^2 D[X_i]$

Vale allora che

PER n SUFFICIENTEMENTE GRANDE di PROVE INDIPENDENTI la MEDIA ARITMETICA DI UNA VARIABILE ALEATORIA CONVERGE IN PROBABILITA' ALLA SUA SPERANZA MATEMATICA

(N.B. X_n conv. in probabilità ad a $\rightsquigarrow P(|X_n - a| < \varepsilon) > 1-\delta$)

$$P\left(\left| \frac{1}{n} \sum_i X_i - m_x \right| < \varepsilon \right) > 1-\delta$$

Infatti $Y = \frac{1}{n} \sum_i X_i$ ha $m_y = m_x$ e $D_y = \frac{D_x}{n}$

e dunque (Cebysev) $P(|Y - m_y| \geq \varepsilon) \leq \frac{D_y}{\varepsilon^2} = \frac{D_x}{n\varepsilon^2}$

e dunque dati qualsiasi ε e δ posso trovare n : $\frac{D_x}{n\varepsilon^2} < \delta$

N.B. Teo di CEBYSEV GERERALIZZATO per $X_1\ X_2 \dots X_n$

di sper. mat. $m_{x_1}\ m_{x_2} \dots m_{x_n}$

e varianze $D_{x_1}\ D_{x_2} \dots D_{x_n}$, $\forall i\ D_{x_i} < L$:

la media aritmetica converge alla media aritmetica delle $\{m_{x_i}\}$ ...