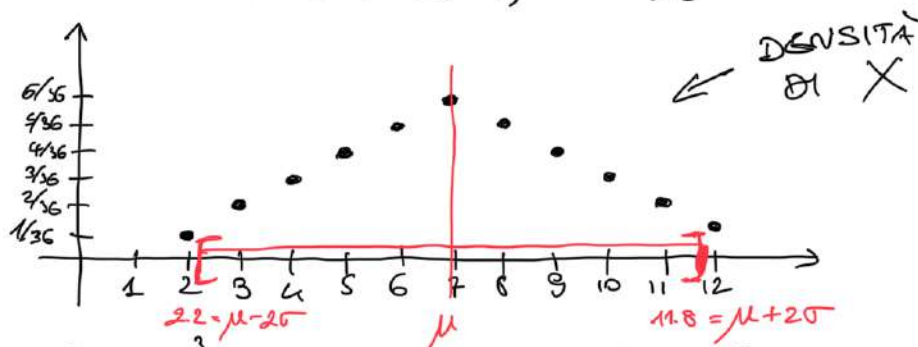


## ESEMPIO

CONSIDERIAMO LA VARIABILE ALEATORIA

$X =$  "LA SOMMA DEI PUNTEGGI DEL LANCIO DI 2 DADI"

$$\begin{aligned} E(X) &= 7 \\ \text{VAR}(X) &= \text{VAR}(X_1 + X_2) = \text{indipendenti} \\ &= \text{VAR}(X_1) + \text{VAR}(X_2) = \text{stesse legge} \\ &= 2 \text{VAR}(X_1) = 35/6 \end{aligned}$$



$$P(X \leq 3) = \sum_{k=2}^3 P(X=k) = \frac{1}{36} + \frac{2}{36} = \frac{3}{36}$$

$$P(7 \leq X \leq 9) = \sum_{k=7}^9 P(X=k) = \frac{6}{36} + \frac{5}{36} + \frac{4}{36} = \frac{15}{36}$$

VOGLIAMO CALCOLARE LA PROB PER UN INTERVALLO CENTRATO NELLA MEDIA, OVVERO  $7 =: \mu$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$$

$$\begin{aligned} \sigma &= \sqrt{\text{VAR}(X)} = \sqrt{35/6} \approx 2.4 \Rightarrow P\left(7 - 2\sqrt{\frac{35}{6}} \leq X \leq 7 + 2\sqrt{\frac{35}{6}}\right) = \\ &= P(2.2 \leq X \leq 11.8) = P(3 \leq X \leq 11) = \\ &= 1 - \underbrace{P(X=2)} - \underbrace{P(X=12)} = 1 - \frac{1}{36} - \frac{1}{36} = \frac{34}{36} \approx 0.94 = 94\% \end{aligned}$$

LA DISUGUAGLIANZA DI CHEBICHEV OFFRE UNA STIMA DI QUESTO INTERVALLO SENZA CONOSCERE LE PROBABILITA' MA SOLO IL VALORE ATTESO  $\mu$  E LA VARIANZA  $\sigma^2$

LA DISTANZA TRA  $X$  E LA MEDIA  $\mu$  È MINORE O UGUALE A  $2\sigma$

$$\mu - 2\sigma \leq X \leq \mu + 2\sigma \Leftrightarrow |X - \mu| \leq 2\sigma$$

$$X \leq \mu - 2\sigma \text{ oppure } X \geq \mu + 2\sigma \Leftrightarrow |X - \mu| \geq 2\sigma$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \geq 1 - \frac{1}{4} = \frac{3}{4} = 0.75 = 75\%$$

LA STIMA È GROSSOLANA

# TEOREMA (DISUGUAGLIANZA DI CHEBICHEV)

$$P(|X - \mu| \geq \delta \sigma) \leq \frac{1}{\delta^2} \quad \text{PER OGNI } \delta > 0$$

## DI MOSTRAZIONE

$$P\left(\left|\frac{X - \mu}{\sigma}\right| \geq \delta\right) \leq \frac{1}{\delta^2} \quad \leftarrow$$

STANDARDIZZATA DI  $X$  :  $E\left(\frac{X - \mu}{\sigma}\right) = 0$   $\text{Var}\left(\frac{X - \mu}{\sigma}\right) = 1$

$$Y := \frac{X - \mu}{\sigma} \quad E(Y) = 0 \quad \text{Var}(Y) = 1$$
$$\text{Var}(Y) = E(Y^2) - \underbrace{E(Y)^2}_{=0} = E(Y^2)$$

$$1 = E(Y^2) = \sum_k y_k^2 \cdot P(Y = y_k) \geq \text{(CONSIDERA SOLO ALCUNI VALORI DI } y_k)$$

$$\sum_{k: |y_k| \geq \delta} y_k^2 \cdot P(Y = y_k) \geq \left( \text{se } |y_k| \geq \delta \Rightarrow \underline{y_k^2 \geq \delta^2} \right)$$

$$\sum_{k: |y_k| \geq \delta} \delta^2 \cdot P(Y = y_k) = \delta^2 \sum_{k: |y_k| \geq \delta} P(Y = y_k)$$

$$= \delta^2 P(|Y| \geq \delta)$$

$$1 \geq \delta^2 P(|Y| \geq \delta) \Rightarrow P(|Y| \geq \delta) \leq \frac{1}{\delta^2} \quad \square$$

PERCHÉ LE FREQUENZE RELATIVE SI AVVICINANO ALLA PROBABILITÀ TEORICA ?

### ESEMPIO

$X_i$  = "ESCE TESTA ALL'  $i$ -ESIMO LANCIO DI UNA MONETA"

$X_i$  È UNA BERNOULLIANA  $E(X_i) = p = \frac{1}{2}$   
 $\underline{VAR(X_i)} = p \cdot (1-p) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} =: \sigma^2$

SE FACCIO  $m$  LANCI / CALCOLO LA FREQUENZA RELATIVA DELLE TESTE

$$\frac{X_1 + X_2 + \dots + X_m}{m} =: \bar{X}_m$$

MEDIA CAMPIONARIA

$$E(\bar{X}_m) = E\left(\frac{X_1 + X_2 + \dots + X_m}{m}\right) = \frac{1}{m} [E(X_1) + E(X_2) + \dots + E(X_m)] =$$
$$= \frac{1}{m} (p + p + \dots + p) = \frac{1}{m} \cdot m \cdot p = p \leftarrow \text{VALORE ATTESO DELLA MEDIA CAMPIONARIA}$$

$$VAR(\bar{X}_m) = VAR\left(\frac{X_1 + X_2 + \dots + X_m}{m}\right) = \frac{1}{m^2} [VAR(X_1) + VAR(X_2) + \dots + VAR(X_m)] =$$
$$= \frac{1}{m^2} [\sigma^2 + \sigma^2 + \dots + \sigma^2] = \frac{1}{m^2} \cdot m \cdot \sigma^2 = \frac{\sigma^2}{m}$$

$$E(\bar{X}_m) = \frac{1}{2} \quad VAR(\bar{X}_m) = \frac{1}{4m}$$

$$VAR(aX+b) = a^2 \cdot VAR(X)$$

SE  $m$  È GRANDE  $\Rightarrow VAR(\bar{X}_m) \rightarrow 0$ , OVVERO I VALORI CHE OTTENGO PER  $\bar{X}_m$  SONO SEMPRE PIÙ VICINI AL SUO VALORE ATTESO, CIOÈ AD  $\frac{1}{2}$

$$P\left(\underset{\uparrow}{\bar{X}_m} \text{ È VICINA AD } \underset{\uparrow}{\frac{1}{2}}\right) \rightarrow \underline{1} \quad \text{PER } m \rightarrow \infty$$

FREQ. RELATIVA  
" "  
MEDIA CAMPIONARIA

PROB. TEORICA  
" "  
VALORE ATTESO DELLA BERNOULLIANA

$$P\left(\bar{X}_m \text{ È LONTANA DA } \frac{1}{2}\right) \rightarrow 0 \quad \text{PER } m \rightarrow \infty$$

$$P\left(|\bar{X}_m - \mu| > \varepsilon\right) \rightarrow 0 \quad \text{PER } m \rightarrow \infty$$

NON È ALTRO CHE LA LEGGE DEI GRANDI NUMERI

## TEOREMA DI BERNOULLI (LEGGE DEI GRANDI NUMERI)

$$P(|\bar{X}_n - \mu| > \varepsilon) \rightarrow 0 \quad \begin{array}{l} \text{PER } n \rightarrow \infty \\ \text{PER OGNI } \varepsilon > 0. \end{array}$$

### DIMOSTRAZIONE

DISUGUAGLIANZA DI CHEBYSHEV:  $P(|X - \mu_X| \geq \delta \sigma_X) \leq \frac{1}{\delta^2}$  PER OGNI  $\delta > 0$

$$\Rightarrow P(|\bar{X}_n - \mu| \geq \delta \frac{\sigma}{\sqrt{n}}) \leq \frac{1}{\delta^2} \quad \text{PER OGNI } \delta > 0$$

$$(E(\bar{X}_n) = \mu, \text{ Var}(\bar{X}_n) = \frac{\sigma^2}{n} \Rightarrow \sigma_{\bar{X}_n} = \frac{\sigma}{\sqrt{n}})$$

DEVIAZIONE STANDARD DELLA MEDIA CAMPIONARIA

ALLORA CONSIDERIAMO  $\varepsilon = \delta \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow \delta = \frac{\varepsilon \cdot \sqrt{n}}{\sigma}$

$$\Rightarrow P(|\bar{X}_n - \mu| \geq \frac{\varepsilon \cdot \sqrt{n}}{\sigma} \cdot \frac{\sigma}{\sqrt{n}}) \leq \frac{1}{\left(\frac{\varepsilon \sqrt{n}}{\sigma}\right)^2}$$

$$\Rightarrow P(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2 \cdot n} \rightarrow 0$$

$$\Rightarrow P(|\bar{X}_n - \mu| \geq \varepsilon) \rightarrow 0 \quad \text{PER } n \rightarrow +\infty \quad \square$$

### ESEMPIO (APPLICAZIONE LEGGE DEI GRANDI NUMERI)

(1)  $M = 10000$ . LANCI DI UNA MONETA;  $p = 0.5$

$$P(|\bar{X}_m - 0.5| > 0.02) = ?$$

DALLA DIMOSTRAZIONE:  $P(|\bar{X}_m - \mu| > \varepsilon) \leq \frac{\sigma^2}{m \cdot \varepsilon^2} \leftarrow \frac{1}{\delta^2}$

$$\varepsilon = \frac{\delta \sigma}{\sqrt{m}} \Rightarrow \delta = \frac{\varepsilon \cdot \sqrt{m}}{\sigma} = \frac{0.02 \cdot \sqrt{10000}}{1/2} = \frac{2}{1/2} = 4$$

$$\sigma = \sqrt{p(1-p)} = \sqrt{\frac{1}{2} \cdot \frac{1}{2}} = \sqrt{\frac{1}{4}} = \frac{1}{2}$$

$$P(|\bar{X}_{10000} - 0.5| > 0.02) \leq \frac{1}{4^2} = \frac{1}{16} = 0.0625 = \underline{\underline{6.25\%}}$$

$\bar{X}_{10000} < 0.48 \text{ o } \bar{X}_{10000} > 0.52$

(2) QUANTI LANCI DOVRANNO FARE PER OTTENERE UNA PROB. DI DISCOSTAMENTO INFERIORE A 0.0001

$$P(|\bar{X}_m - 0.5| > 0.02) \leq 0.0001$$

$$\varepsilon = \frac{\delta \sigma}{\sqrt{m}} \Rightarrow \sqrt{m} = \frac{\delta \sigma}{\varepsilon} \Rightarrow m = \frac{\delta^2 \sigma^2}{\varepsilon^2}$$

$0.0001 = \frac{1}{\delta^2} \Rightarrow \delta^2 = 10000$

$$M = \frac{10000 \cdot 1/4}{(0.02)^2} = \frac{2500}{\left(\frac{2}{100}\right)^2} = 2500 \cdot \left(\frac{10000}{4}\right) = (2500)^2 = \underline{\underline{6250000 \text{ LANCI}}}$$