UNITELMA SAPIENZA

SAPIENZA
UNIVERSITÀ DI ROMA
DIPARTIMENTO DI INFORMATICA

Probabilità

Marco Isopi

25. Legge dei grandi numeri e disuguaglianza di Chebichev.

$\{T, C\}$ # teste ~ # croci

1000 lanci $\Longrightarrow$ 500 teste e 500 croci?

990 e 50

$\#T - \#C \ll \#$ lanci

$S_n = \#$ Teste in n lanci ; $\frac{S_n}{n} \longrightarrow \frac{1}{2}$

$P(T) = p \; ; \quad 0 < p < 1$

$\frac{S_n}{n} \longrightarrow p$ $\quad S_n$ è una v.a.

$$P\left(p - \varepsilon \leq \frac{S_n}{n} \leq p + \varepsilon\right) \xrightarrow[n \to \infty]{} 1 \quad \forall_{\varepsilon > 0}$$

$$\left\{\frac{S}{n} \geq p + \varepsilon\right\} \cup \left\{\frac{S_n}{n} \leq p - \varepsilon\right\}$$

$$P\left(\frac{S_n}{n} \geq p+\varepsilon\right) \xrightarrow[n \to \infty]{} 0$$

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P\left(\frac{S_n}{n} \geq p+\varepsilon\right) = \sum_{k=m}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$m = \lceil n(p+\varepsilon) \rceil$$

$$= e^{-\lambda n \varepsilon}\left(p e^{\lambda q} + q e^{-\lambda p}\right)^n$$

$$e^x \le x + e^{x^2} \quad \forall x \in \mathbb{R}$$

$$P\left(\frac{S_n}{n} \geq p + \varepsilon\right) \le e^{-\lambda n \varepsilon}\left(p e^{\lambda^2 q^2} + q e^{\lambda^2 p^2}\right)^n$$

$$\le e^{\lambda^2 n - \lambda n \varepsilon} \qquad \lambda = \frac{1}{2}\varepsilon$$

$$P\left(\frac{S_n}{n} \geq p + \varepsilon\right) \le e^{-\frac{1}{4} n \varepsilon^2} \xrightarrow[n \to \infty]{} 0$$

Fisso $\lambda > 0$; se $K \geq m \Rightarrow \underline{e^{\lambda K} \geq e^{\lambda n(p+\varepsilon)}}$

$q = 1 - p$

$$P\left(\frac{S_n}{n} \geq p+\varepsilon\right) \leq \sum_{K=m}^{n} e^{\lambda[K - n(p+\varepsilon)]} \binom{n}{K} p^K q^{n-K}$$

$$= e^{-\lambda n \varepsilon} \sum_{K=m}^{n} \binom{n}{K} \left(p e^{\lambda q}\right)^K \left(q e^{-\lambda p}\right)^{n-K}$$

$$\leq e^{-\lambda n \varepsilon} \sum_{K=0}^{n} \binom{n}{K} \left(p e^{\lambda q}\right)^K \left(q e^{-\lambda p}\right)^{n-K}$$

$$P\left(\frac{S}{n} \leq p-\varepsilon\right) \leq e^{-\frac{1}{4} n \varepsilon^2}$$

$$P\left(p-\varepsilon \leq \frac{S_n}{n} \leq p+\varepsilon\right) \geq 1-2e^{-\frac{1}{4} n \varepsilon^2}$$

$$\xrightarrow[n \to \infty]{} 1$$

Disuguaglianza di Markov

$$X \geq 0 \; ; \quad P(X \geq t)$$

$$X \geq t \, \mathbb{1}_{\{X \geq t\}}$$

$$E(X) \geq E(t \, \mathbb{1}_{\{X \geq t\}})$$

$$E(X) \leq t \, E(\mathbb{1}_{\{X \geq t\}})$$

$$E(X) \leq t \, P(X \geq t) \; ; \quad \boxed{P(X \geq t) \leq \frac{E(X)}{t}}$$

Disuguaglianza di Chebicev

X ; $m = E(X)$; $Var(X) < \infty$ $\quad \underline{t > 0}$

$$P(|X-m| \geq t) = P(|X-m|^2 \geq t^2)$$

$$\overset{DM}{\leq} \frac{E((X-m)^2)}{t^2} = \frac{Var(X)}{t^2}$$

$$P(\quad |X-m| \geq t) \leq \frac{Var(X)}{t^2}$$

LGN $X_1, X_2, X_3, \ldots$ i.i.d.

$E(X_1) = m$; $Var(X_1) < \infty$

$$\frac{X_1 + X_2 + \ldots X_n}{n} \xrightarrow[n \to \infty]{} m$$

$$\forall \varepsilon > 0 \quad P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m\right| \geq \varepsilon\right) \xrightarrow[n \to \infty]{} 0$$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

Vogliamo che $Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \xrightarrow[n\to\infty]{} 0$

$$Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) =$$

$$\frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\left[n\, Var(X_i)\right] = \frac{Var(X_i)}{n}$$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m\right| \geq \varepsilon\right) \xrightarrow[n\to\infty]{} 0 \quad \forall \varepsilon > 0$$

Dis di Chebichev

$P(T) = p \qquad Y_n = \#$ teste in n lanci

$P\left(\left|\frac{Y_n}{n} - p\right| \le \frac{1}{10}\right) \ge \frac{99}{100} \qquad n?$

$Y_n \sim B(n, p)$; $E(Y_n) = np$; $Var(Y_n) = np(1-p)$

$E(Y_n/n) = p$; $Var(Y_n/n) = \dfrac{p(1-p)}{n}$

$$P\left(\left|Y_n/n - p\right| > \frac{1}{10}\right) \le \frac{1}{\left(\frac{1}{10}\right)^2} Var(Y_n/n) =$$

$$\frac{100}{n} p(1-p) \qquad // \qquad P\left(\left|Y_n/n - p\right| > \frac{1}{10}\right) < \frac{1}{100}$$

$$\frac{100}{n} p(1-p) < \frac{1}{100} \Longrightarrow \boxed{n > \frac{10'000}{p(1-p)}}$$

$$n > \frac{10000}{p(1-p)} \leftarrow p^* = \frac{1}{10}$$

caso peggiore

$$n > 111 \cdot 112$$

$$P\left(\left|\frac{Y_n}{n} - p\right| > \varepsilon\right) < 2e^{-\frac{1}{4}n\varepsilon^2}$$

$$\hookrightarrow < 2e^{-\frac{n}{400}}$$

vogliamo $2e^{-n/400} < \frac{1}{100}$

$$\boxed{n \geq 1565}$$