DTMC - SIMULATION!

WE START FROM A DTMC $D:(S, \Pi, P_0)$

PROBLEM: GENERATE / SAMPLE A TRAJECTORY

$s_0 s_1 s_2 s_3 \dots$

$P(s_0 s_1 s_2 s_3 \dots) = p_0 \pi_{0,1} \pi_{1,2} \pi_{2,3} \dots$

SIMULATION (D, n_{max}) // output: $s_0 s_1 \dots s_{n_{max}}$

$n \leftarrow 0$ // CURRENT STEP

$s \leftarrow$ SAMPLE (p) // CURRENT STATE

WHILE $n < n_{max}$

$s \leftarrow$ SAMPLE $(\Pi_{s, \cdot})$

$n \leftarrow n+1$

Oss: Se S è finito, SAMPLE è facile da implementare! (più o meno)

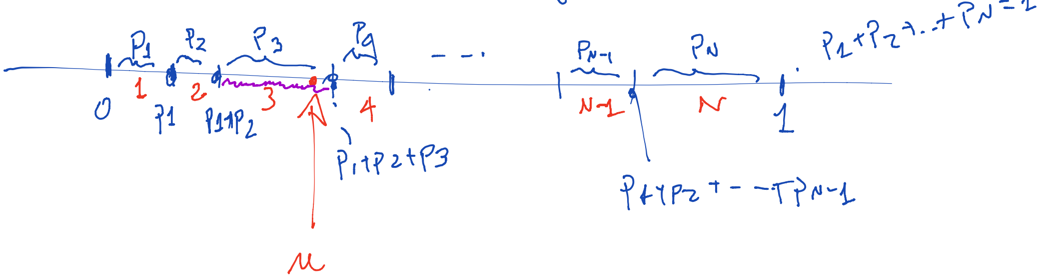
. . . .

Sia $S = \{s_1, \dots, s_N\}$ e p una distrib. di prob. su S

$p = (p_1 \dots p_N)$, $p_i \geq 0$, $\sum_i p_i = 1$

HO A DISPOSIZIONE UN PSEUDO RANDOM NUMBER GENERATOR rand()

CHE RITORNA $u \sim Unif(0,1)$



SAMPLE (P) $\quad P = (p_1, \ldots, p_N)$

$u \leftarrow$ RAND() // $u$nif $[0,1)$

$\sigma \leftarrow 0$

$i \leftarrow 0$

WHILE $\sigma < u$

$\quad i \leftarrow i+1$

$\quad \sigma \leftarrow \sigma + p_i$

RETURN i

Expected # of comparisons C

$$E[C] = \sum_i i \cdot p_i = E[P]$$

$\vec{p}$	0.1	0.2	0.3	0.4	$E[p] = 0.1 \cdot 1 + 0.2 \cdot 2 + 0.3 \cdot 3 + 0.4 \cdot 4 = 3$
$\vec{q}$	0.4	0.3	0.2	0.1	$E[q] \; 0.4 \cdot 1 + 0.3 \cdot 2 + 0.2 \cdot 3 + 0.1 \cdot 4 = 2$

ANALYSIS OF SIMULATION DATA

PROBLEMS TO SOLVE.

LET $X = (S, \Pi, P)$ A DTMC. LET $g : S \to \mathbb{R}$ state property

LET $m > 0$ GIVEN $\quad \gamma : S^{m+1} \to \mathbb{R}$, for m given Trajectory property

a) $E[g(\underbrace{X_m}_{})]$

X_m = r.v. state of X at step m

b) $E[\gamma(\underbrace{X_0 \ldots X_m}_{\text{trajectory from time 0 to } m})]$, $X_{0:m}$

IF X HAS A STEADY STATE DISTRIBUTION π, THEN

c) $E[g(X_\infty)] := E_\pi[g(s)]$

SOLUTION: PRODUCE A SAMPLE OF a), b), OR c) BY SIMULATING X, AND ANALYZE THIS SAMPLE STATISTICALLY.

Focus on a), b)

a)/b) SIM $(X, \overset{f}{g}, m)$

$s_0 \cdots s_m \leftarrow$ SIMULATE (X, m)

RETURN $g(s_m)$

$f(s_0 \cdots s_m)$

STO CAMPIONANDO da una VARIABILE $Y \sim g(X_m) / f(X_{0:m})$

$y_1 \cdots y_N$ OBSERVATIONS OF Y

PROBLEM: ESTIMATE $E[Y]$ from $y_1 \cdots y_N$

SAMPLE MEAN ESTIMATOR

$$\sim \bar{Y}_N = \frac{1}{N} \sum_{i=1}^{N} y_i$$

$y_i \sim Y_i = Y, \; Y_1, \ldots, Y_N$ iid

INDEP + IDENTICALLY DISTR AS Y

$\bar{Y}_N = \frac{1}{N} \sum_{i=1}^{N} Y_i$, IS A R.V. !

$\bar{Y}_N$ IS CONSISTENT

if $\bar{Y}_N \to E[Y]$

UNBIASED

$E[\bar{Y}_N] = E[Y]$

$$E[\bar{Y}_N] = \frac{1}{N} \sum_{i=1}^{N} E[Y_i] = \frac{1}{\cancel{N}} \cancel{N} \cdot E[Y] = E[Y]$$ UNBIASED

STRONG LAW OF LARGE NUMBER

$$\lim_{N\to\infty} \bar{Y}_N \to E[Y] \quad a.s. \qquad \left[Y_i, \bar{Y}_N : \underset{\text{ho prob } P}{\Omega} \to \mathbb{R} \right]$$

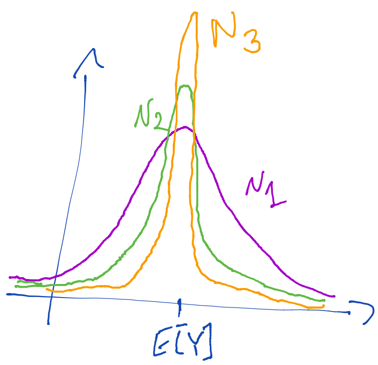
Fisso $\omega \in \Omega$

$$P\left(\left\{ \omega \mid \lim_N \bar{Y}_N(\omega) = E[Y] \right\} \right) = 1$$

CENTRAL LIMIT THEOREM

$$\sqrt{N} \frac{\bar{Y}_N - E[Y]}{\sqrt{VAR[Y]}} \xrightarrow{\text{weakly}} \mathcal{N}(0,1)$$ (gaussian with zero mean and variance 1)

$$\left[\bar{Y}_N \sim \mathcal{N}\left(E[Y], \frac{VAR[Y]}{N} \right) \right.$$



$N_1 < N_2 < N_3$!

$$VAR[\bar{Y}_N] = VAR\left[\frac{1}{N} \sum_1 Y_i \right] = \frac{1}{N^2} \sum VAR[Y_i]$$

indip.

$$= \frac{VAR[Y]}{N}$$

SAMPLE VARIANCE

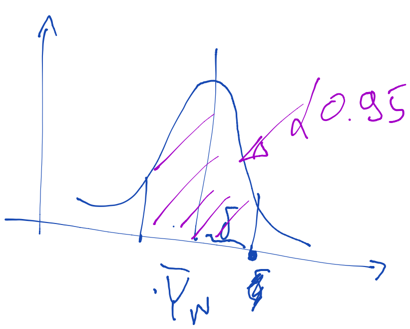
$$\bar{S}^2_N = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y}_N)^2$$

fatevi il conticino

$$= \frac{1}{N-1}\sum Y_i^2 - \frac{N}{N-1}\bar{Y}_N^2$$

$$E[\bar{S}^2_N] = \frac{N}{N-1}\left(E[Y_i^2] - E[\bar{Y}_N^2] + E[\bar{Y}]^2 - E[\bar{Y}]^2\right)$$

$$= \frac{N}{N-1}\left(VAR[Y] - \underbrace{VAR[\bar{Y}_N]}_{\frac{VAR[Y]}{N}}\right) = VAR[Y]$$



$$\mathcal{N}\left(\bar{Y}_N, \frac{\bar{S}^2_N}{N}\right) : W_N$$

TROVA δ t.c.

$$P(W_N \in [\bar{Y}_N - \delta, \bar{Y}_N + \delta]) = \alpha \quad (\alpha = 0.95)$$

$$Z \sim \mathcal{N}(0,1) \qquad \phi(z) = P(Z \le z) \ \text{cdf}$$

$$P(|Z|<z) = P(Z<z) - P(Z<-z) = \alpha$$

$$\phi(z) - (1-\phi(z)) = \alpha$$

$$\Rightarrow z_\alpha = \phi^{-1}\left(\frac{\alpha+1}{2}\right)$$

$$\delta = z_\alpha \cdot STD[\bar{Y}_N] = z \cdot \frac{\sqrt{S_N^2}}{\sqrt{N}}$$

CONFIDENCE INTERVAL

$$I_c = \left[\bar{Y}_N - \frac{1}{\sqrt{N}} \cdot z_\alpha \cdot \sqrt{S_N^2}, \bar{Y}_N + \frac{1}{\sqrt{N}} z_\alpha \sqrt{S_N^2}\right]$$

$\rightarrow$

It's a random quantity

$$P_2(E[Y] \in I_c) \simeq \alpha \quad (\text{per } N \text{ grande abbastanza } N \geq 30)$$

Nel caso b), tipically $f(s_0 \cdots s_m) \in \{0, 1\}$

In questo framework posso formalizzare REACHABILITY / ABSORP.