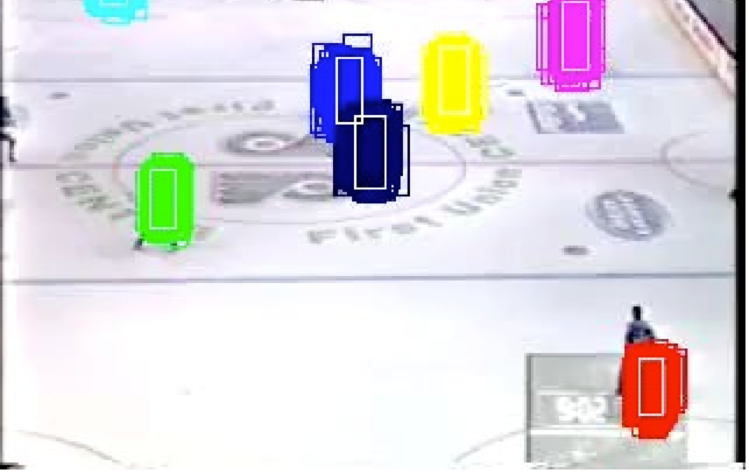
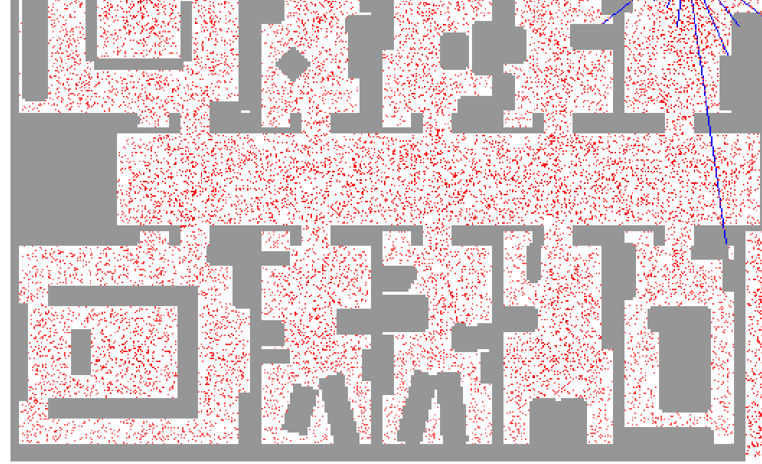
Rubén Martínez Cantín
Dpto. Informática e Ingeniería de Sistemas.

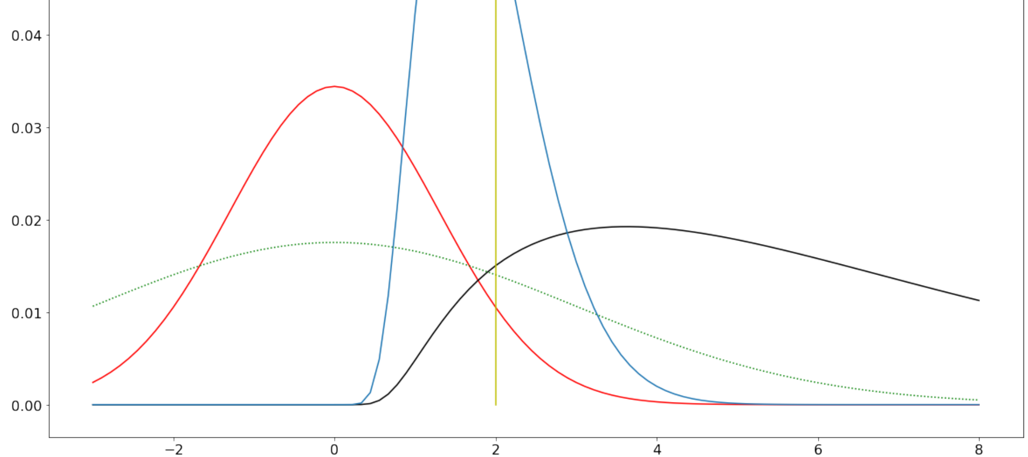
methods


First Union



- What is the posterior? What is the predictive posterior?

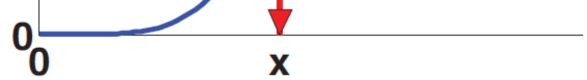
Universidad Zaragoza 1542


0.04
0.03
0.02
0.01
0.00
−2
0
2
4
6
8

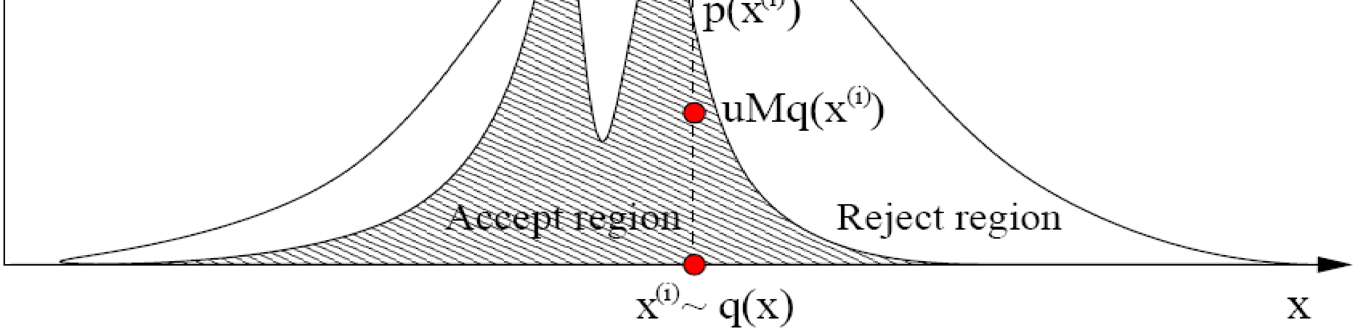
$$I_N(f) = \frac{1}{N}\sum_{i=1}^{N} f(x^{(i)}) \quad \xrightarrow[N \to \infty]{a.s.} \quad I(f) = \int_{\mathcal{X}} f(x)p(x)dx$$

Universidad Zaragoza 1542

- Solution: use a different distribution for sampling (proposal) and for evaluating (target)



Universidad Zaragoza 1542

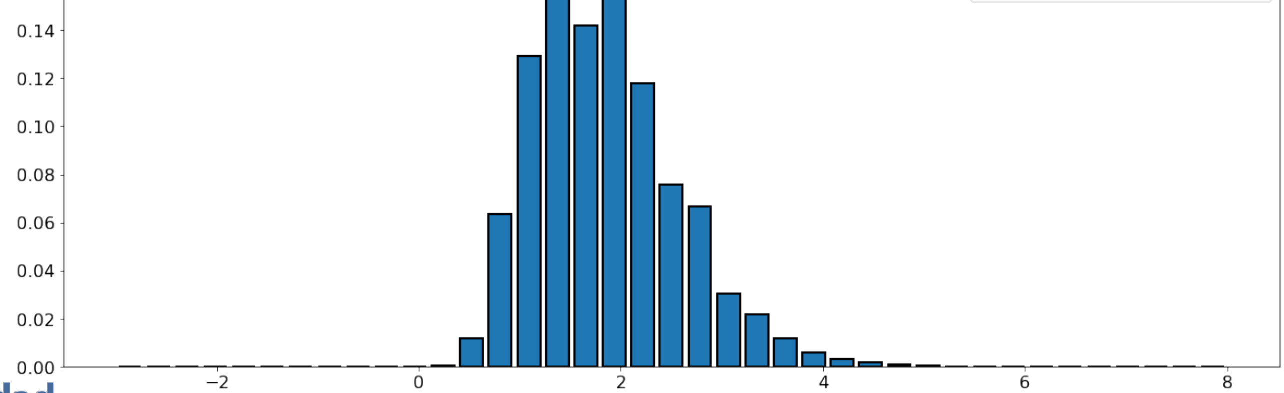


Universidad Zaragoza 1542

Video: Kevin P. Murphy

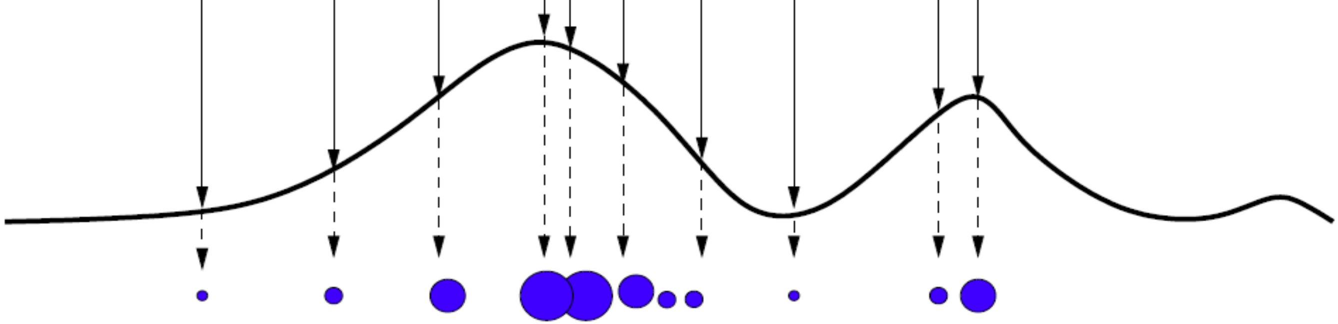
- Thus, the Monte Carlo estimate becomes

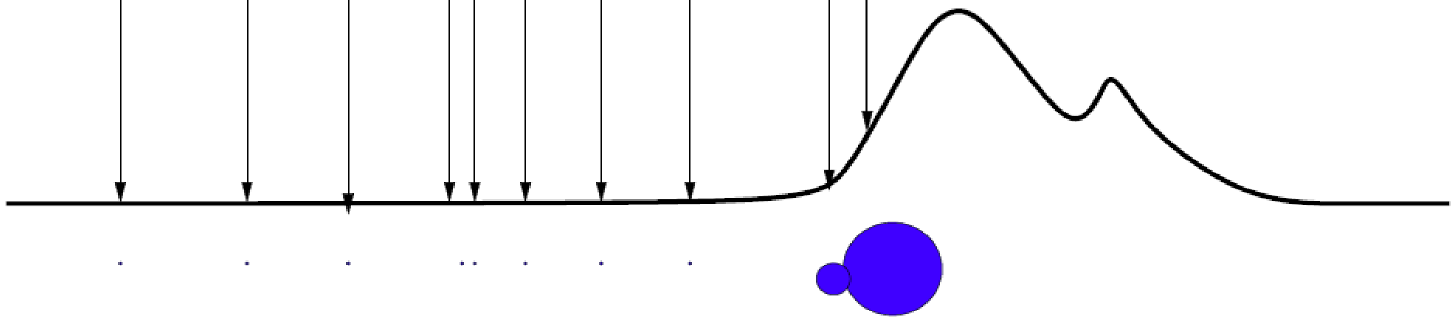
$$\hat{I}_N(f) = \sum_{i=1}^{N} f(x^{(i)}) w(x^{(i)})$$

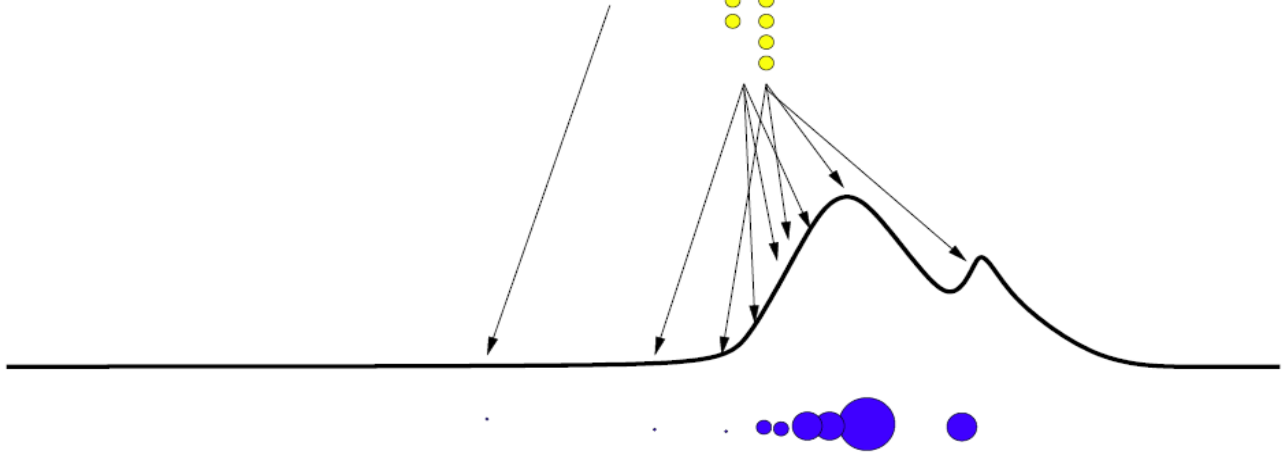

0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
−2
0
2
4
6
8

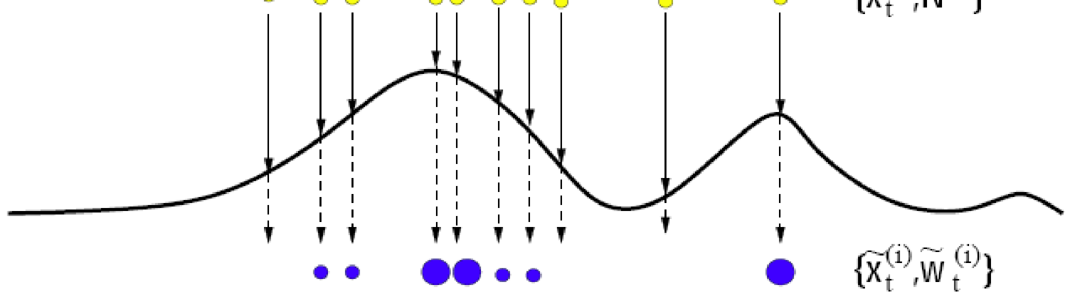
$$p(y_*|\mathbf{x}_*,\mathbf{X},\mathbf{y}) \approx \frac{\cancel{\frac{1}{N}}\sum_{i=1}^{N} w(\theta^i)p(y_*|\mathbf{x}_*,\theta^i)}{\cancel{\frac{1}{N}}\sum_{i=1}^{N} w(\theta^i)} = \sum_{i=1}^{N} \widetilde{w}^i p(y_*|\mathbf{x}_*,\theta^i)$$

$$\boxed{\widetilde{w}^i = \frac{w(\theta^i)}{\sum_{i=1}^{N} w(\theta^i)}}$$







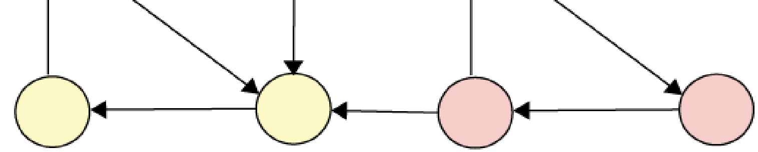

{x̃_t^(i), w̃_t^(i)}

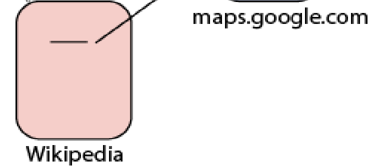
$$\nu T^t \to \pi \quad as \quad t \to \infty$$

where π is the stationary distribution (unique).

$$p(x) = (0.2, 0.4, 0.4)$$

- Aperiodicity: No traps or end states.





- How do we guarantee stability?
  - Add some random jumps

$$T = L + E$$

$i=1$

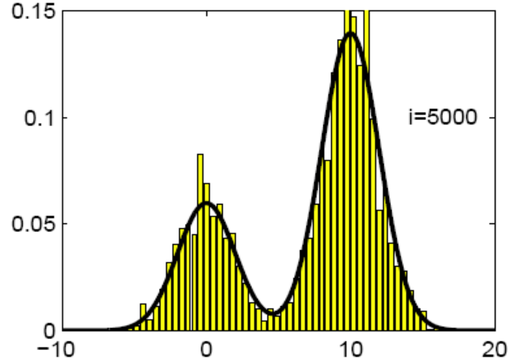
- For continuous spaces, the transition matrix becomes a kernel $p(y|x)$.

$$\int \pi(x)p(y|x)dx = \pi(y)$$

Universidad Zaragoza 1542

$$\int \pi(x_t)p(x_{t+1}|x_t)dx_t = \pi(x_{t+1})$$

- See algo Murphy2013 Sec 24.3.6



Universidad Zaragoza 1542

- Sample $u \sim U(0,1)$

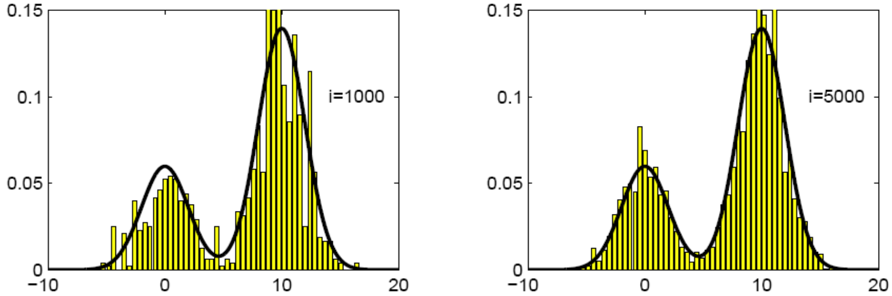
- Set new sample $x^{t+1} = \begin{cases} x' & if\ u < r \\ x^t & if\ u \geq r \end{cases}$

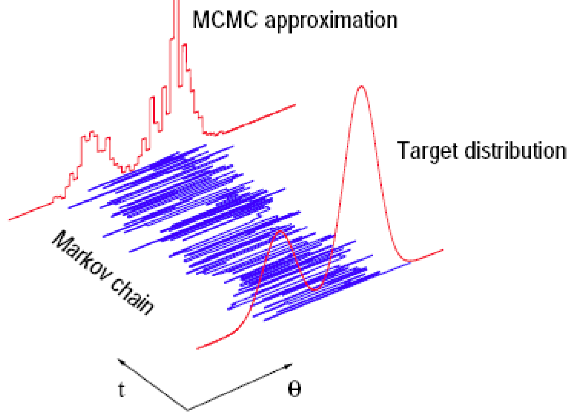
Universidad Zaragoza 1542

Jump to the same place

Jump rejected

- This has detailled balance and the stationary distribution is our target distribution (proof: Murphy2012 Sec 24.3.6)


0.15
0.1
0.05
0
−10
0
10
20
i=1000
i=5000



32

and Computer Vision