Sampling based Estimation

Particle based methods

Randomly sample instances from the distribution and use those instances as a sparse representation to estimate quantities regarding the statistics of the overall distribution

$D = \{x[1], \ldots, x[M]\}$ sampled IID from P

If $(P(X=1)=p)$

Estimator for p: $T_D = \frac{1}{M}\sum_{m=1}^{M} x[m]$

More generally for any distribution P, function f:

$E_{\hat{P}}(f) = \frac{1}{M}\sum_{m=1}^{M} f(x[m])$ f on samples — empirical expectation

$E_P(\text{indicator}) = \text{probability}$

Sampling from Discrete Distribution

$Val(X) = \{x^1, x^2, \ldots, x^k\}$ $P(x^i) = \theta^i$

0 θ^1 $\theta^1\theta^2$ 1 — uniformly in

		e.g.
X_1	θ^1	.10
X_2	θ^2	.20
X_3	$1-(\theta^1+\theta^2)$	.70

Hoeffding Bound

$$\underbrace{P_D}_{\text{prob of bad dataset}}\Big(\underbrace{T_D}_{\text{Estimator for } p} \underbrace{\notin [p-\epsilon, p+\epsilon]}_{\text{is > than } \epsilon \text{ away from p}}\Big) \le \underbrace{2e^{-2M\epsilon^2}}_{\text{bound}}$$

property of the dataset; # of samples M; epsilon

Chernoff bound

multiplicative

$$P_D(T_D \in [p(1-\epsilon), p(1+\epsilon)]) \le 2e^{-Mp\epsilon^2/3}$$

Hoeffding bound $p > 1-\delta$

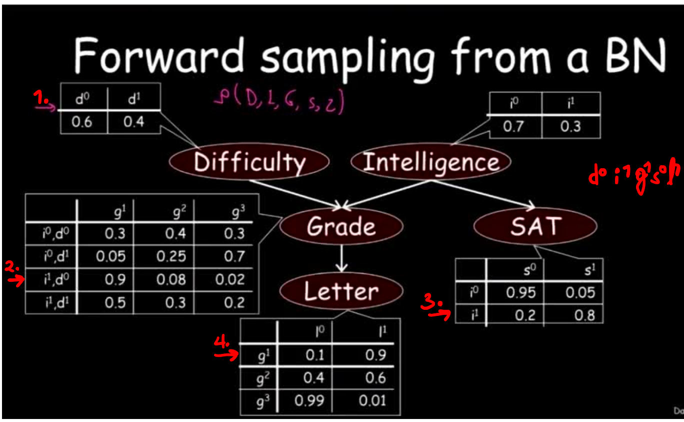
$$M > \frac{\ln(2/\delta)}{2\epsilon^2}$$

Chernoff bound $p > 1-\delta$

$$M \ge \frac{3\ln(2/\delta)}{p\epsilon^2}$$

Sample the distribution on Topological Order

From the root



d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3

	g^1	g^2	g^3
i^0,d^0	0.3	0.4	0.3
i^0,d^1	0.05	0.25	0.7
i^1,d^0	0.9	0.08	0.02
i^1,d^1	0.5	0.3	0.2

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

	l^0	l^1
g^1	0.1	0.9
g^2	0.4	0.6
g^3	0.99	0.01

Forward Samplin for Queriying

- goal: Estimate $P(Y=y)$
  - Generate samples from BN
  - Compute fraction where $Y=y$
    of samples
  - Apply bounds

- Queries with evidence
  - Goal: Estimate $P(Y=y \mid E=e)$
  - Rejection sampling
    - Generate samples
    - throw away those where $E \neq e$
    - Compute fraction where $Y=y$
    - The problem is that is not too probable to keep a sample vanishingly small

  # of samples needed grows exponentially with the number of variables observed

  - Forward sampling generally infeasible No Topological ordering