Monte Carlo Algorithms

Theorem
Let $X_1, \ldots, X_m$ be independent and identically distributed indicator random variables with $\mathbb{E}[X_i] = \mu$. If $m \geq (3 \ln(2/\delta))/(\epsilon^2 \mu)$ then

$$
\mathbb{P} \left[ \left| \frac{1}{m} \sum_{i=1}^{m} X_i - \mu \right| \geq \epsilon \mu \right] \leq \delta
$$

Simple Applications:
- Can estimate the probability $p$ a biased coin lands heads with $$(3 \ln(2/\delta))/(\epsilon^2 p_0)$$ coin tosses where $p_0$ is a known lower bound for $p$.
- Can estimate $\pi$ with $(12 \ln(2/\delta))/\epsilon^2$ assuming $\pi \geq 1$:
  - Pick random point $(A, B) \in [-1, 1]^2$. Let $X_i = 1$ if $A^2 + B^2 \leq 1$.
  - Then $\mu = \pi/4$ so a $1 + \epsilon$ approx of $\mu \geq 1/4$ gives $1 + \epsilon$ approx of $\pi$. 
DNF Counting Problem

- Problem: Given a formula in disjunctive normal form (DNF), e.g.,

\[ \phi = (x_1 \land \bar{x}_2 \land x_3) \lor (x_2 \land x_4) \lor (\bar{x}_1 \land x_3 \land x_4) \]

approximate \( c(\phi) \) = the number of satisfying assignments of \( \phi \).

- First attempt at a solution:
  - Pick a random assignment and let \( X_i = 1 \) if assignment satisfies \( \phi \).
  - \( \mathbb{E}[X_i] = \mu = c(\phi)/2^n \) where \( n \) is the number of variables.
  - Repeat \( 3 \times 2^n \ln(\delta/2))/(\epsilon^2 c(\phi)) \) times and take average.
  - Could be very time consuming since \( c(\phi) \) could be much less than \( 2^n \). (As we noted in class however, \( c(\phi)/2^n \geq 1/2^k \) if there exists a clause with at most \( k \) literals.)
Better approach: Define

\[ S_j = \text{all assignments that satisfy } j\text{th clause} \]

\[ P = \{(j, a) : j \in [t], a \in S_j\} \]

\[ Q = \{(j, a) : j \in [t], a \in S_j, a \notin S_k \text{ for all } k < j\} \]

where \( t \) is the number of clauses.

Note that \( |Q| = c(\phi) \), \( |P| = \sum_{j=1}^{t} |S_j| \), and \( |S_j| = 2^{n-\ell_j} \) where \( \ell_j \) is the number of variables in clause \( j \).

Pick random \((j, a) \in P\) and let \( X_i = 1 \) if \((j, a) \in Q\).

To pick random \((j, a) \in P\), first pick \( j \) with probability \( |S_j|/|P| \) and then pick random assignment that satisfies \( j\text{th clause} \)

\[ \mathbb{E}[X_i] = |Q|/|P| \geq 1/t \]

Repeat \((3t \ln(2/\delta))/\epsilon^2\) times and take mean.
Another Example: Counting Independent Sets

- Let $G_m = G$ and construct $G_{i-1}$ be removing a single edge from $G_i$. The $\Omega(G_i)$ be the set of independent sets of $G_i$.

- Then,
  \[
  |\Omega(G)| = \frac{|\Omega(G_m)|}{|\Omega(G_{m-1})|} \times \ldots \times \frac{|\Omega(G_1)|}{|\Omega(G_0)|} \times |\Omega(G_0)|
  \]

- Let $r_i = |\Omega(G_i)|/|\Omega(G_{i-1})|$ and note that $|\Omega(G_0)| = 2^n$:
  \[
  |\Omega(G)| = 2^n \prod_{i \in [m]} r_i
  \]

- Hence if we find estimates $\tilde{r}_i$ such that
  \[
  (1 - \epsilon/m)r_i \leq \tilde{r}_i \leq (1 + \epsilon/(2m))r_i
  \]
  then $2^n \prod_{i \in [m]} \tilde{r}_i$ is a $(1 + \epsilon)$ approximation of $|\Omega(G)|$
Estimating $r_i$

Note that

$$1 \geq r_i \geq \frac{|\Omega(G_i)|}{|\Omega(G_{i-1})|} \geq \frac{|\Omega(G_i)|}{|\Omega(G_i)| + |\Omega(G_{i-1}) \setminus \Omega(G_i)|} \geq 1/2$$

where last line follows since every set in $\Omega(G_{i-1}) \setminus \Omega(G_i)$ includes both end points of the edge removed to generate $G_{i-1}$ and so $|\Omega(G_{i-1}) \setminus \Omega(G_i)| \leq |\Omega(G_i)|$

Hence, if we sample

$$\frac{(3 \ln(2/\delta))}{((\epsilon/(2m))^2 r_i)} \leq 24m^2 \ln(2/\delta)/\epsilon^2$$

independent sets uniformly from $\Omega(G_{i-1})$ and return the fraction that are also independent sets in $G_i$, we get a $1 + \epsilon/(2m)$ approximation for $r_i$ with probability $1 - \delta$.

Setting $\delta = 1/m^2$ ensures we get good enough estimates for all $r_1, r_2, \ldots, r_m$ with probability at least $1 - 1/m$. 
But how do we sample independent sets?

- Design a Markov chain whose state space is the set of elements we want to sample and whose stationary distribution is the set of elements we want to sample from.
- Let $X_0, X_1, X_2, \ldots$, be a run of the Markov chain. For sufficiently large $r$, $X_r, X_{2r}, X_{3r}, \ldots$ are nearly independent samples from the required distribution.
- ... this is **Markov Chain Monte Carlo Method**.