EFFICIENT MONTE CARLO ESTIMATION

A universal algorithmic method and a fundamental application of Chernoff bounds

Let $\Omega$ be a population of size $N$
Let $A \subseteq \Omega$
Let $r = \frac{|A|}{|\Omega|}$

Think of $N$ as very large

Take $n$ independent uniform samples from $\Omega$
Let $X_i = 1$ if $i$-th sample was in $A$
Let $0$ if $i$-th sample was in $\Omega \setminus A$

Let $\hat{r} = \frac{\sum_{i=1}^{n} X_i}{n}$ be an estimator for $r$

We want to quantify the quality of $\hat{r}$
How can we "quantify" the "quality" of \( \tilde{r} \)?

For certain \( \varepsilon, \delta \) we want to state:

\[
\Pr \left[ |\tilde{r} - r| > \delta r \right] < \varepsilon
\]

Notice: Multiplicative error is VERY important.

For example, \( r \) could be the probability that a network fails, typically very small.

Say \( r = 0.001 \), i.e., once in ten thousand the network fails.

We want \( \tilde{r} \) to be, say, for \( \delta = \frac{1}{2} \)

\[0.0005 < \tilde{r} < 0.0015\]

indicating that the network indeed fails once in a few thousand times.

Realize that additive error \( \Pr \left[ |\tilde{r} - r| > 1\% \right] \)

would allow for \( \tilde{r} = 0 \) to be considered a good estimator.

But \( \tilde{r} = 0 \) indicates that the network never fails inadequate for every practical purpose!
THEOREM \[ \Pr \left[ \left| \bar{r} - r \right| > \delta r \right] < \varepsilon \]

for \[ n > 3 \frac{1}{\delta^2} \left( \ln \frac{2}{\varepsilon} \right) \cdot \frac{1}{r} \]

PROOF \[ \Pr \left[ \left| \bar{r} - r \right| > \delta r \right] < \varepsilon \iff \]

\[ \Pr \left[ \left| \frac{X}{n} - r \right| > \delta r \right] < \varepsilon \iff \]

\[ \Pr \left[ \left| X - rn \right| > \delta rn \right] < \varepsilon \iff \]

\[ \Pr \left[ \left| X - \mu \right| > \delta \mu \right] < \varepsilon \]

Know, by Chernoff bounds \[ \Pr \left[ \left| X - \mu \right| > \delta \mu \right] < 2e^{-\frac{\delta^2 \mu}{3}} \]

Want \[ 2e^{-\frac{\delta^2 \mu}{3}} < \varepsilon \iff \]

\[ 2e^{\frac{-\delta^2 rn}{3}} < \varepsilon \iff \]

\[ e^{\frac{-\delta^2 rn}{3}} < \frac{\varepsilon}{2} \iff \]

\[ -\frac{\delta^2 rn}{3} < \ln \frac{\varepsilon}{2} \iff \ln \frac{2}{\varepsilon} < \frac{\delta^2 rn}{3} \]

\[ \iff 3 \frac{1}{\delta^2} \left( \ln \frac{2}{\varepsilon} \right) \cdot \frac{1}{n} < n \]

QED
NOTE: Let us examine closely what the theorem says:
In order to be within $\delta r$ accuracy with probability at least $1 - \epsilon$

you need $N > \frac{3}{\delta^2} \left( \ln \frac{2}{\epsilon} \right) \frac{1}{r}$ samples

---

notice the dependency on $r$ !!!

- If $r$ is a constant, say $\frac{1}{4} = 25\% < r < 75\%$
  then $N > 12 \frac{1}{\delta^2} \left( \ln \frac{2}{\epsilon} \right)$ INDEPENDENT OF $N$

Now you understand why, for example, election exit polls get pretty good accuracy using very small sample sizes $N = 100,000,000$ $n = 5000$ is standard it is because $r = \frac{\text{BLUE}}{N}$ (also $r = \frac{\text{RED}}{N}$) is $\approx 25\%$

- If on the other hand $r = O \left( \frac{1}{N} \right)$ VERY SMALL
  you would need $n > \Omega \left( \frac{1}{\delta^2} \ln \left( \frac{2}{\epsilon} \right) \cdot N \right)$ samples
  that would be like the population size $N$ !!!