What will we cover?

overlap w/ MC

Prob 2

(1) Basic Monte Carlo + limit theorems

(2) Markov Chains

(3) Brownian + Stochastic calculus
+ computer simulation

(4) Stochastic approximation

(5) Markov jump processes
+ computer simulation

(6) Feynman-Kac equations

Main Resource:
"Applied Stochastic Analysis"
E, Li, Vanden-Eijnden

My Notes (in class and pdf on Brightspace).

Evaluation: written and programming exercises every week or two

turn those in on gradescope

I assume that you know some basic prob. (RV, independence, Expectation, Conditional expectation)

Today: Monte Carlo basics and limit theorems

Goal of Monte Carlo is to estimate

$$E[f(X)] = \int f(x)\,\pi(dx)$$

π is the "law" of X, i.e.

$$P(X \in A) = \mathbb{P}(X \in A) = \mathbb{P}(\{\omega : X(\omega) \in A\}) = \int_A \pi(dx) = \pi(A)$$

I often write $\pi[f] = E[f(X)]$

The simplest Monte Carlo estimator is

$$\bar{f}_N = \frac{1}{N}\sum_{k=0}^{N-1} f(X^{(k)}) \quad \text{where}$$

$X^{(k)}$ are $\underbrace{\text{independent and identically distributed}}_{\text{i.i.d.}}$

according to π

Notice that

$$E[\bar{f}_N] = E[f(X)] = \pi[f]$$

But what we want is

$$\bar{f}_N \to \pi[f] \quad \text{as } N \to \infty$$

$\bar{f}_N$ is a RV (i.e. a function) so we could mean different things by $\longrightarrow$

<u>Laws of Large Numbers</u> (for $\bar{f}_N$)

Suppose $\operatorname{var}(f(X)) = \sigma^2 < \infty$

$X^{(k)}$ i.i.d. from π

$$\mathrm{var}(\bar{f}_N) = E\left[(\bar{f}_N - \pi[f])^2\right] = E\left[\left(\frac{1}{N}\sum f(X^{(k)}) - \pi[f]\right)^2\right]$$

$$= \frac{1}{N^2} E\left[\sum_{k,\ell=0}^{N-1} \left(f(X^{(k)}) - \pi[f]\right)\left(f(X^{(\ell)}) - \pi[f]\right)\right]$$

$$= \frac{1}{N^2}\sum_{k=0}^{N-1} E\left[\left(f(X^{(k)}) - \pi[f]\right)^2\right]$$

$$\cancel{+ \frac{2}{N^2}\sum_{k<\ell} \mathrm{cov}\left(f(X^{(k)}), f(X^{(\ell)})\right)}$$

$$= \frac{\sigma^2}{N} \longrightarrow 0 \quad \text{as } N \to \infty$$

We'll call this convergence in mean squared error (MSE)

By Chebychev $\neq$

For any $\varepsilon > 0$ $P(|\bar{f}_N - \pi[f]| > \varepsilon) \leq \dfrac{\mathrm{Var}(\bar{f}_N)}{\varepsilon^2} = \dfrac{\sigma^2}{N\varepsilon^2}$

so $\bar{f}_N \longrightarrow \pi[f]$ "in probability"

This is called the Weak Law of Large Numbers

In fact WLLN holds for $X^{(k)}$ i.i.d if $E[|f(X)|] < \infty$

The Strong Law of Large (SLLN) says

$$\lim_{N\to\infty} \bar{f}_N(\omega) = \pi[f] \quad \forall\ \omega \in A \subset \Omega$$

"almost sure" convergence

with $P(A) = 1$

SLLN is still true for i.i.d. $X^{(i)}$ if $E[|f(X)|] < \infty$

Convergence in distribution:

The LLNs imply that for any bounded continuous function g, $E[g(\bar{f}_N)] \to g(\pi[f])$

The distribution of $\bar{f}_N$ $\left(p_N(A) = P(\bar{f}_N \in A) \right)$ converges to $\delta(x - \pi[f])$

Not so exciting since $\bar{f}_N$ is converging to a number

But consider $Z_N = \sqrt{N}\left(\bar{f}_N - \pi[f] \right)$

$$E[Z_N] = 0 \qquad Var(Z_N) = N\, var(\bar{f}_N) = \sigma^2$$

So there's a chance to make sense of the limit of the Z_N's

For example, suppose $E[e^{\lambda f(X)}] < \infty \quad \forall \lambda$

and $X^{(k)}$ are i.i.d. from π

$$E[e^{\lambda Z_N}] = E\left[e^{\frac{\lambda}{\sqrt{N}} \sum (f(X^{(k)}) - \pi[f])}\right]$$
$$= E\left[e^{\frac{\lambda}{\sqrt{N}}(f(X) - \pi[f])}\right]^N$$

expand e^x about $x=0$

$$E[e^{\lambda Z_N}] = \left(1 + \frac{\lambda^2}{2N}\sigma^2 + \mathcal{O}(N^{-3/2})\right)^N$$

$$\approx e^{\frac{\lambda^2 \sigma^2}{2}} \quad \text{(for big } N\text{)}$$

but $e^{\frac{\lambda^2\sigma^2}{2}}$ is the moment generating function of $N(0, \sigma^2)$

CLT: If $X^{(k)}$ are i.i.d. and $\sigma^2 = \text{var}(f(X)) < \infty$ then

$$Z_N = \sqrt{N}\left(\overline{f_N} - \pi[f]\right) \longrightarrow N(0, \sigma^2)$$

in distribution

The CLT tells us

$$P(\ Z_N \in I\ ) \quad \text{for large } N$$

some interval

or $P\left(\bar{f}_N - \pi[f] \in \frac{I}{\sqrt{N}} \right)$

It doesn't tell us about

$$P(\ \bar{f}_N - \pi[f] \in I\ )$$

specifically it doesn't tell us how fast

$P(\bar{f}_N - \pi[f] \in I)$ goes to 0 if $0 \notin I$

Concentration $\neq$

Suppose $E[e^{\lambda f(X)}] < \infty \quad \forall \lambda$

Chernoff's $\neq$ says

$$\begin{aligned}
P(\ \bar{f}_N - \pi[f] \geq a\ ) &= P\left(\sum f(X^{(k)}) - \pi[f] \geq Na \right) \\
&\leq e^{-\lambda Na}\, E[e^{\lambda \sum (f(X^{(k)}) - \pi[f])}] \qquad \forall \lambda > 0 \\
&= e^{-\lambda Na}\, E[e^{\lambda (f(X) - \pi[f])}]^N \qquad \forall \lambda > 0
\end{aligned}$$

$$\geq e^{-N\{\lambda a - \log E[e^{\lambda(f(X)-\pi[f])}]\}} \qquad \forall \lambda > 0$$

$$\leq e^{-N I(a)}$$

where $I(a) = \sup_{\lambda > 0} \{\lambda a - \log E[e^{\lambda(f(X)-\pi[f])}]\}$

LDP says that this decay rate is sharp

roughly:

$$-\frac{1}{N}\log P(\bar{f}_N - \pi[f] \geq a) \longrightarrow I(a)$$

$$P(\bar{f}_N - \pi[f] \geq a) = e^{-N(I(a) + o(1))}$$

$\uparrow$ LDP "Prefactor"