Introductory Remarks

What is a Monte Carlo method?

Randomness $\longrightarrow$ $\boxed{\text{MCM}}$ $\longrightarrow$ desired results

(Why the name Monte Carlo?)

Invented many times!!

1) Advent of probability theory because of gaming (also combinatorics)
2) Early 20^{th} century: to "verify" new results in probability
3) Fermi, von Neumann & Ulam to solve problem resulting from the atomic bomb

Approches to Monte Carlo

Direct Simulation (game playing)

Solution of Mathematical Problems

Many ways to present Monte Carlo

a) Mathematical
(algorithms, proofs, convergency errors)

b) Computer Science
(algorithms, implementations, simulation)

c) Applications
(physics, engineering, chemistry, finance)

Examples of Monte Carlo

(not to be covered in depth)

a) Financial instrument evaluation
(very hot!)

b) Solution of Schrödinger equation, many-body
(only known method)

c) Verification of t-distribution, etc Gosset (Student)

d) Buffon needle problem

Monte Carlo is Based on Probability

Probability <u>is</u> measure theory

Probability (measure) ~ measure $P(\Omega) = 1$

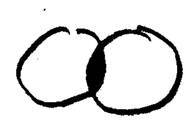
Event ~ measurable set

Random variable ~ measurable function

Expected value ~ integral

Call $A, B, C, \ldots$ events

$$P(A+B+C+\ldots) \leq P(A) + P(B) + \ldots$$

sub additivity 

$A, B, C, \ldots$ are exclusives $\leq \rightarrow =$

$A, B, C, \ldots$ are exhaustive r.h.s. $= 1$

Statistical Review

$$P(AB) = P(A/B)\ P(B)$$

$P(A|B)$ is conditional probability

if $P(A|B) = P(A)$ then
A & B are independent and

$$P(AB) = P(A)P(B)$$

$$P(ABC) = P(AB|C)\ P(C)$$

Cummulitive Distribution Function:

RV η $\qquad F(y) = P(\eta \le y)$

$$F(-\infty) = 0 \qquad F(+\infty) = 1$$

F is increasing and right-continuous

(at worst F is a step function)

Dic

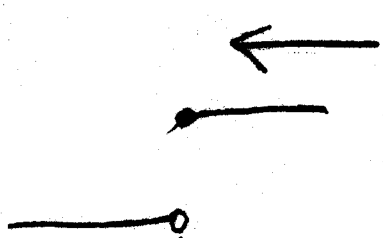
$$P[\eta = j] = \frac{1}{6}$$

$$j = 1, 2, 3, 4, 5, 6$$

$$F[y] = P(\eta \le y)$$

$$= \begin{cases} 0 & y < 1 \\ 1/6 & 1 \le y < 2 \\ 2/6 & 2 \le y < 3 \\ 3/6 & 3 \le y < 4 \\ 4/6 & 4 \le y < 5 \\ 5/6 & 5 \le y < 6 \\ 1 & y \ge 6 \end{cases}$$

$$F(y) = \int_0^y dF(x) = \frac{1}{6}\sum_{i=1}^{6} \delta(x-i)$$



$$\delta(x - x_0) = \begin{cases} +\infty & \text{at } x_0 \\ 0 & \text{elsewhere} \end{cases}$$

$$\int_0^y \delta(x - x_0)\, dx = \begin{cases} 1 & \text{if } y \ge x_0 \\ 0 & \text{if } y < x_0 \end{cases}$$

Expectation

$$E[g(\eta)] = \int g(y)\, dF(y)$$

$$F(y) = P(\eta \le y)$$

because of F this is a Stieltjes integral

If $F'(y) = f(y)$ exists then

$$E[g(\eta)] = \int g(y)\, f(y)\, dy$$

note: if η is a discrete distribution then

$$E[g(\eta)] = \sum_i g(y_i)\, f_i$$

$$F(y) = \sum_i f_i\, H(y - y_i)$$

Heaviside function

$f(y)$, f_i constitute probability d.f.

Multivariate Distributions

η, γ: $F(y, z) = P(\eta \le y, \gamma \le z)$

Marginal Distributions

$$G(y) = P(\eta \le y)$$

$$H(z) = P(\gamma \le z)$$

"integrate out the other variable"

If $F(y,z) = G(y)H(z)$, then η & γ are independent

Note: $\sum_i E[g(\eta_i)] = E[\sum_i g(\eta_i))$

hold for all r.v.'s η_i

BUT

$\prod_i E[g(\eta_i)] = E[\prod_i g(\eta_i)]$

holds only when η_i's are Independent

Note $E[g(\eta)] \neq g(E[\eta])$ generally

$$E[\eta] = \int y \, dF(y) = \mu$$

is the mean of η

$E[(\eta-\mu)^r]$ is the r^{th} central moment of η

$E[\eta^r]$ is the r^{th} moment of η

$$E[(\eta-\mu)^2] = \sigma^2 = Var[\eta]$$

~~the~~ dimensions of σ^2 & μ^2 are the same so

$\frac{\mu}{\sqrt{\sigma^2}}$ is coefficient of variation (dimensionless)

note: σ called standard deviation

$$= \sqrt{\sigma^2}$$

Multivariate "Variances"

$$Cov(z, y) = E[(y-\mu)(z-\nu)]$$

$$\mu = E[z] \qquad \nu = E[y]$$

$$Cov(z, z) = Var(z) \quad \text{by above}$$

Assume independence $E[yz] = \mu\nu$

$$Cov(z, y) = E[(yz - \mu z - y\nu + \mu\nu)]$$

$$= E[yz] - \mu\nu - \mu\nu + \mu\nu \quad = 0$$

<u>independence $\Leftrightarrow$ Cov( ) = 0</u>

$$\rho(z, y) = \frac{Cov(z, y)}{\sqrt{Var(z)\,Var(y)}}$$

$$-1 \le \rho \le +1$$

negatively correlated

positively correlated

0

uncorrelated $\longleftrightarrow$ independent

Continued

$$Var\left(\sum_i \eta_i\right) = \sum_i \sum_j Cov(\eta_i, \eta_j)$$

(left to students, it's easy!)

- $Var[g(\eta_1, \ldots, \eta_k)]$

where $\mu_i = E(\eta_i)$

$$g_i = \left.\frac{\partial g}{\partial \eta_i}\right|_{\eta_i = \mu_i}$$

$$\approx \sum_{i=1}^{k} \sum_{j=1}^{k} g_i g_j \; Cov(\eta_i, \eta_j)$$

This is like a multidimensional Taylor series expansion about the means.

Some CDFs

(PDFs left to students)

$$F(y) = \begin{cases} 0 & y < 0 \\ y & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}$$

uniform or rectangular distributi..

$U[0,1]$

$$U[a,b] \rightarrow F(y) = \begin{cases} 0 & y < a \\ \frac{y-a}{b-a} & a \le y \le b \\ 1 & y > b \end{cases}$$

Binomial $$F(y) = \sum_{t \le y} \frac{n!}{t!(n-t)!} p^t (1-p)^{n-t}$$

p is prob of single event & describes dist. of Σ n such trials

Poisson $$F(y) = \sum_{t \le y} e^{-\lambda} \frac{\lambda^t}{t!}$$

$\lambda > 0$

dist. of number "t" of events occuring at rate λ

More CDFs

exponential $F(y) = H(y)\left(1-e^{-\lambda y}\right)$

λ is exponential rate in fixed time, Poisson counts numbers of these

Gaussian $F(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$

$E[\eta] = \mu \qquad Var[\eta] = \sigma^2$

$$dF(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

$F(y)$ is called the error function

$$N(\mu, \sigma)$$

Also have

$$F(\bar{y}) = \int_{-\infty}^{\bar{y}} |2\pi \bar{\bar{V}}|^{-1/2} \exp\left[-\frac{1}{2}(\bar{t}-\bar{\mu})^T \bar{\bar{V}}(\bar{t}-\bar{\mu})\right] d\bar{t}$$

$V_{ij} = Cov(\eta_i, \eta_j)$

multivariate normal

Central Limit Theorem

Assume η_i's are independent identically distributer r.v's

with $E(\eta_i) = \mu_i$ and $Var(\eta_i) = \sigma_i^2$

$$Y_i = \frac{\eta_i - \mu_i}{\sigma_i}$$ has mean 0 and variance 1

$$P\left[\frac{1}{N}\sum_{i=1}^{N} Y_i \le y\right] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/t}\, dt$$

as $N \to \infty$

Central Limit Theorem $N \sim 10$ good
$N \sim 25$ great

Note when moments don't exist, other "stable" distributions can pop out

$$\eta \text{ has } F(y) = \frac{1}{\pi} \int_{-\infty}^{y} \frac{1}{1+t^2}\, dt$$

Cauchy distribution

Sampling

Want to sample $\eta : \eta_1 \cdots \eta_n$

sample size n

$$\bar{\eta} = \sum_{i=1}^{n} \eta_i \tfrac{1}{n} \Big/ \sum_{i=1}^{n} \tfrac{1}{n}$$

$$\text{or} \quad = \sum_{i=1}^{n} \eta_i w_i \Big/ \sum_{i=1}^{n} w_i$$

weighted sum

What about these two estimators

$t(\eta)$ an estimator of η

$E[\eta] = \mu$ then

$E[t(\eta) - \mu] = \beta$ is called the bias

$$Var[t(\eta)] = E[(t(\eta) - \mu - \beta)^2] = \sigma_t^2$$

sampling variance of t

Sampling Continued

Want t to be $\beta = 0$ (unbiased)

and σ_z^t small as possible (minimum variance)

- use minimum variance linear estimators (best weighted sum)

maximum liklihood estimates.

construct a likelihood function and choose estimate to minimize it (in ℓ^2).

-

$$\bar{\eta} = \frac{1}{n} \sum_{i=1}^{n} \eta_i \qquad \sigma_{\bar{\eta}} = \sigma / \sqrt{n}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\eta_i - \bar{\eta})^2 \qquad \sigma_{s^2} \simeq \sigma^2 / \sqrt{\tfrac{n}{2}}$$

are maximum likelihood estimates

Efficiency of Monte Carlo

$$\bar{\eta} = \frac{1}{n}\sum_{i=1}^{n} \eta_i$$

$$\frac{\bar{\eta} - \eta}{\frac{\sigma}{\sqrt{n}}}$$ is approximately $N(0,1)$

So $\eta \in \bar{\eta} \pm k\frac{\sigma}{\sqrt{n}}$

$k = 1$ $p = .68$
$k = 2$ $p = .90$
$k = 3$ $p = .995$

to get a 19% "confidence interval"

to be $1/10$ of its previous width

requires $\frac{\sigma}{\sqrt{kn}} = \frac{\sigma}{10\sqrt{n}}$ $k = 10^2$

So we say that Monte Carlo converges like $n^{-1/2}$. Thus it is prudent to find ways to make σ a small as possible (variance reduction)

Regression

line of best fit

search for $y = mx + b$ have

y_i , x_i.

$$l(m,b) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$$ likelihood function

$$\frac{\partial l}{\partial m} = \frac{\partial l}{\partial b} = 0$$

$$l_b = \sum_{i=1}^{n} 2(mx_i + b - y_i) 1$$

$$l_m = \sum_{i=1}^{n} 2(mx_i + b - y_i) x_i$$

$$\left.\begin{aligned} \sum_{i=1}^{n} y_i &= m \sum_{i=1}^{n} x_i + b \\ \sum_{i=1}^{n} y_i x_i &= m \sum_{i=1}^{n} (x_i)^2 + b \sum_{i=1}^{n} x_i \end{aligned}\right\}$$ solve these

Regression Example Continued

$$\sum_{i=1}^{n} y_i = \overline{Y} \quad \sum_{i=1}^{n} x_i = \overline{X}$$

$$\sum_{i=1}^{n} (x_i)^2 = \overline{XX} \quad \sum_{i=1}^{n} x_i y_i = \overline{XY}$$

$$\begin{pmatrix} \overline{X} & 1 \\ \overline{XX} & \overline{X} \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} \overline{Y} \\ \overline{XY} \end{pmatrix}$$

Solution (at home) give standard linear regression results.

Sampling

Fixed Sampling (n fixed)

Sequential Sampling (something other than n fixed)

Stratified Sampling break sample space into strata