Efficient Techniques, MC Revisited

Monte Carlo Method

Applicable to any system, Unbiased, Consistent

Long Execution time

Semi Analytic Technique to Accelerate MC + Method of Moments

Considerable acceleration

Needs the pdf of sufficient statistics of the symbol detection decision

Other Techniques to overcome lengthy run-time of MC

Tail Extrapolation

Curve fitting to MC simulation results

Pdf Estimators

Model based estimating the pdf of decision metric

Importance Sampling

Biasing the channel noise in a controlled way to increase error occurrence

Efficient Techniques, MC Revisited…

Tail Extrapolation

Extrapolation → Making inferences beyond the range of available data

Collecting data in higher error rates and extrapolate to low error rate cases

Higher Error Rates regimes are quick to simulate

Tail distribution assumption is class of Generalized Exponential (GE) distributions

$$f(x\,|\,\nu,\mu,\sigma)=\frac{\nu}{2\sqrt{2}\sigma\,\Gamma(1/\nu)}\exp\left\{-\left|\frac{x-\mu}{\sqrt{2}\sigma}\right|^{\nu}\right\}$$

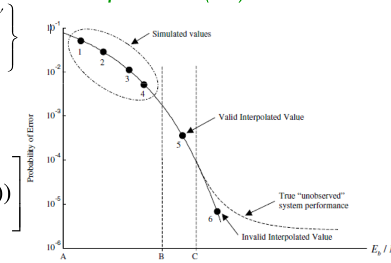
$$\Gamma(a)=\int_0^{\infty}t^{a-1}e^{-t}dt\rightarrow\Gamma(n)=(n-1)!$$

For error threshold =T:

$$\int_{T+\mu}^{\infty}f(x\,|\,\nu,\mu,\sigma)dx\approx\exp\left[-\left(\frac{T}{\sigma\sqrt{2}}\right)^{\nu}(1-\varepsilon(T))\right]$$

$$\varepsilon(T)<<1\rightarrow 1-\varepsilon(T)\approx 1$$

$$\ln(-\ln p(T))=\nu\ln(T/\sqrt{2})-\nu\ln(\sigma)$$



For M threshold values T_i, estimate $p(T_i)$ using MC

Using linear regression or LMS we can fit a line to the data pairs $(T_i, p(T_i))$

Efficient Techniques, MC Revisited…

Tail Extrapolation…

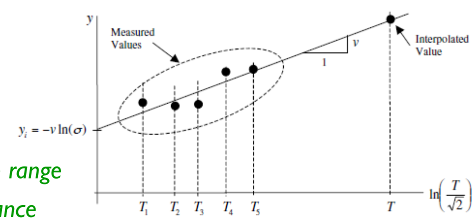
From slope → ν

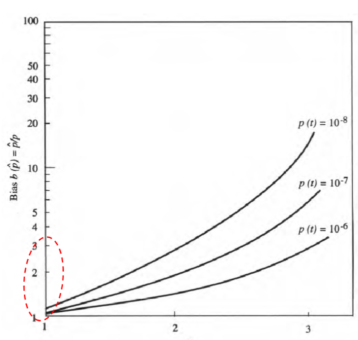
From intersect and ν → σ

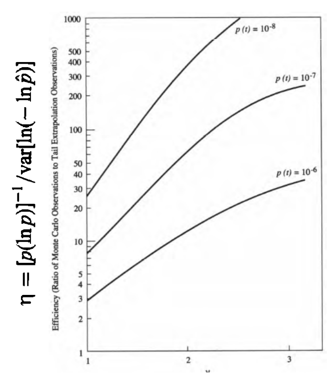
Estimation is biased

Bias can be kept in a reasonable range

Trade-off between bias and variance







Efficient Techniques, MC Revisited…

PDF Estimators

Having the pdf of decision statistics → $P_E=\int_T^{\infty}f_V(v)dv$

Histogram method is the most straight forward → biased, needs data for tails

Parametric methods are more robust

Gram-Charlier Series

a posteriori method of approximating an unknown, but nearly Gaussian, pdf with known moments: $E[Y^k]$

$$\hat{f}_Y(y)=\frac{1}{\sigma_Y\sqrt{2\pi}}\exp\left[-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right]\sum_{j=1}^{\infty}C_jH_j\left(\frac{y-\mu_Y}{\sigma_Y}\right),\quad C_k=\frac{1}{k!}\int_{-\infty}^{\infty}H_k(y)f_Y(y)dy$$

$$H_k(y)=\frac{[-\exp(-y^2/2)]^{(k)}}{\exp(-y^2/2)}\qquad C_k=\frac{1}{k!}\sum_{i=0}^{k/2}\frac{k^{2i}}{2^i\,i!}(-1)^i\mu_{k-2i}\qquad \mu_k=E\left(\frac{Y-\mu_Y}{\sigma_X}\right)^k$$

$$H_k(y)=yH_{k-1}(y)-(k-1)H_{k-2}(y),\quad k\geq 2\qquad H_0(y)=1,\quad H_1(y)=y$$

H: Hermite Polynomials If only two moments are known → Gaussian pdf

Good approximation around mean → biased, no reasonable truncation

Yields excellent results in some wireless channel conditions

Efficient Techniques, MC Revisited…

PDF Estimators…

Parzen Series

provides a smoothed estimator of a pdf

Part of a bigger class of kernel estimators

$$\hat{f}_Y(y)=\hat{f}_{Y|Y_1,\dots,Y_n}(y\mid y_1,\dots,y_n)=\frac{1}{nh(n)}\sum_{i=1}^{n} g\left(\frac{y-y_i}{h(n)}\right)$$

h: smoothing function, g:weithing function, n: data size

Asymptotically unbiased and consistent

For Gaussian like pdfs reasonable choices for g and h are:

$$h(n)=1/\sqrt{n},\qquad g(x)=\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$$

$$\rightarrow \hat{f}_Y(y)=\frac{1}{\sqrt{2\pi n}}\sum_{i=1}^{n}\exp\left(-n(y-y_i)^2/2\right)$$

Number of terms = number of samples

Has shown good results for multipath + AWGN problems

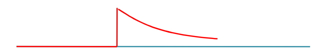
Efficient Techniques, MC Revisited…

Importance Sampling

Rare events are important! ← in MC

For low probabilities, a rare event produces a peak with 1/N vanishing rate

$$\lim_{n\to\infty}\frac{N_A}{N}=P(A)$$



Evaluation of expected values needs rare events or extreme events to happen

For instance for BER we can bias the noise to make rare events happen more frequently

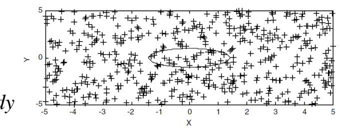
Example: Ellipse Area Estimation

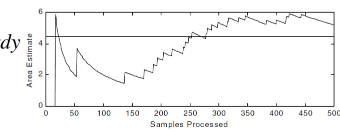
$$x^2+2y^2<2 \qquad A=\int_{-1}^{1}\int_{-\sqrt{2-2y^2}}^{\sqrt{2-2y^2}} dx\,dy=\pi\sqrt{2}=4.443$$

$$h(x,y)=0 \text{ outside the ellipse, } 1 \text{ inside}:\ A=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\,dxdy$$

$$A=A_b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)(1/A_b)\,dxdy=A_b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f_{XY}(x,y)\,dxdy$$

$$A=A_b\left(\frac{N_e}{N}\right)$$





Efficient Techniques, MC Revisited…

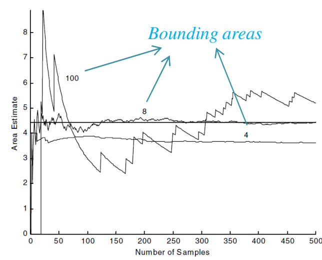
Importance Sampling…

Example: Ellipse Area Estimation…

Performing MC for different bounding sizes

Box size decreases, so does the error,

but down to a limit…



Optimal bounding region, for uniform f_{XY}

The ellipse itself! Converges to true value with just the first sample! → not a solution!

Smallest possible bounding rectangle or even a polygon → may be found analytically

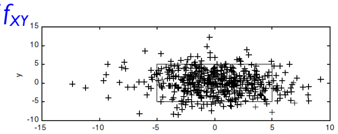
Rectangle is attractive because of the simplicity of f_{XY}

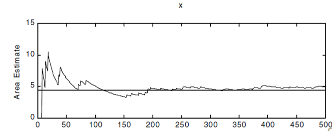
Any (almost) non-uniform pdf can be used,

$$f_{XY}(x,y)=\begin{cases}1/w(x,y), & h(x,y)=1\\ arbitrary, & h(x,y)=0\end{cases}$$

$$A=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)w(x,y)f_{XY}(x,y)\,dxdy$$

$$\hat{A}=\frac{1}{N}\sum_{i=1}^{N} h(x_i,y_i)w(x_i,y_i)\qquad (x_i,y_i)\sim f_{XY}(x,y)$$





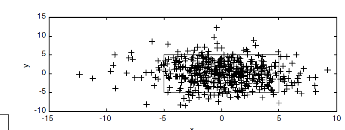
Efficient Techniques, MC Revisited…

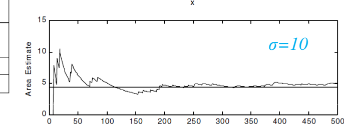
Importance Sampling…

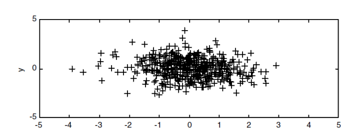
Example: Ellipse Area Estimation…

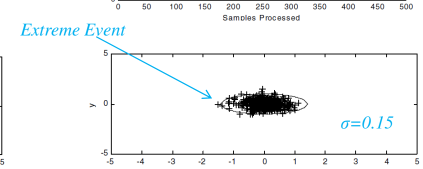
Gaussian, σ=10, 1, 0.15

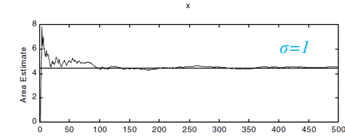
Standard Deviation, σ	Number of Samples in Ellipse	Mode of Convergence
Too small	Nearly all	Slow (exhibits extreme events)
Appropriate	Between 10% and 90%	Rapid
Too large	Very few	Slow

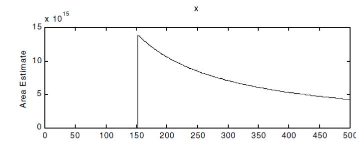










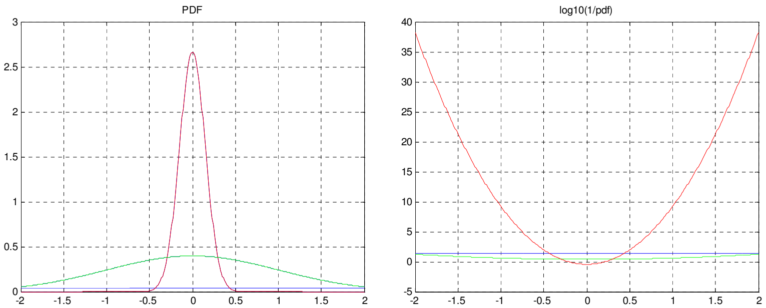


Efficient Techniques, MC Revisited…

Importance Sampling…

Gaussian, σ=10, 1, 0.15

$\sigma=10$ $\sigma=1$ $\sigma=0.15$



Standard Deviation, σ	Number of Samples in Ellipse	Mode of Convergence
Too small	Nearly all	Slow (exhibits extreme events)
Appropriate	Between 10% and 90%	Rapid
Too large	Very few	Slow

Course Notes, Simulation of Communication Systems, Sharif, EE, Iman Gholampour, imangh@sharif.edu , Fall 2011

Efficient Techniques, MC Revisited…

Importance Sampling… Sensitivity to pdf choice

If the random data comes from a physical source, like noise in communication systems

$$B=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\,f_{phy}(x,y)\,dxdy \xrightarrow{MC} \hat{B}=\frac{1}{N}\sum_{i=1}^{N}h(x_i,y_i),\quad (x_i,y_i)\sim f_{phy}, IID$$

B is like the probability that the generated points fall inside the ellipse

And we use a different noise, called biased noise, say f_{sim}

$$B=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)w(x,y)\,f_{sim}(x,y)\,dxdy \xrightarrow{MC} \hat{B}=\frac{1}{N}\sum_{i=1}^{N}h(x_i,y_i)w(x_i,y_i)$$

$$(x_i,y_i)\sim f_{sim}, IID \qquad w(x,y)=\begin{cases} f_{phy}(x,y)/f_{sim}(x,y), & h(x,y)=1 \\ arbitrary, & h(x,y)=0 \end{cases}$$

Easy to show that sufficient condition for having unbiased and consistent estimation is

$$f_{sim}(x,y)>0 \text{ for all } (x,y) \quad \text{that } h(x,y)f_{phy}(x,y)>0$$

Mode of convergence depends on w(x,y)

PDFs Inside Ellipse	Weighting Function Inside Ellipse	Rate of Convergence
$f_{sim}(x,y)=f_{phy}(x,y)$	$w(x,y)=1$	Same as simple MC
$f_{sim}(x,y)>f_{phy}(x,y)$	$w(x,y)<1$	Faster than simple MC
$f_{sim}(x,y)<f_{phy}(x,y)$	$w(x,y)>1$	Slower than simple MC

Course Notes, Simulation of Communication Systems, Sharif, EE, Iman Gholampour, imangh@sharif.edu , Fall 2011

Efficient Techniques, MC Revisited…

Importance Sampling… Applied to Communication Systems

Channel noise goes through the receiver block and its pdf changes!

Unknown or difficult to track pdf

Assume that channel generates (x,y)~f_{phy}

Receiver maps (x,y) to g(x,y)= (a,b)~f_{rec} → impractical to calculate f_{rec} from f_{phy}

$$B=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(a,b)\,f_{rec}(a,b)\,dadb \xrightarrow{MC} \hat{B}=\frac{1}{N}\sum_{i=1}^{N}h(a_i,b_i),\quad (a_i,b_i)\sim f_{rec}, IID$$

$$\hat{B}=\frac{1}{N}\sum_{i=1}^{N}h(g(x_i,y_i)) \quad (x_i,y_i)\sim f_{phy}$$

Now alter f_{phy} to f_{bias} and to remove the effect of this bias from MC estimation use a weighting function…

$$w(a,b)=\frac{f_{rec}(a,b)}{f_{rec-bias}(a,b)}=\frac{f_{phy}(x,y)}{f_{bias}(x,y)} \leftarrow (a,b)=g(x,y)$$

$$\hat{B}=\frac{1}{N}\sum_{i=1}^{N}w(g(x_i,y_i))h(x_i,y_i)$$

Not dependent on f_{rec}

Proved to be unbiased and consistent

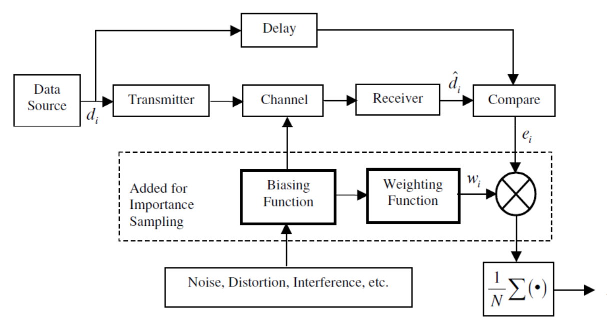
Course Notes, Simulation of Communication Systems, Sharif, EE, Iman Gholampour, imangh@sharif.edu , Fall 2011

Efficient Techniques, MC Revisited…

Importance Sampling… Applied to Communication Systems…

Mapping the MC area estimation to BER(SER) estimation

$$w_i=f_{phy}(n_i)/f_{bias}(n_i) \rightarrow \hat{p}_e=\frac{1}{N}\sum_{i=1}^{N}w_i h(n_i)$$



Ideally desired is to have a mixture of samples inside and outside of error region

Too many samples in either region will cause the convergence rate to be very slow

CIS and IIS → conventional and improved importance sampling

Efficient Techniques, MC Revisited…

Importance Sampling… Applied to Communication Systems…

CIS: changing the variance (IS10)

$$w_i = f_{phy}(n_i)/f_{bias}(n_i) = \frac{\sigma_b}{\sigma_p}\exp(-n_i^2(\sigma_p^{-2} - \sigma_n^{-2})/2)$$

IIS: Shifting the pdf (IS9)

$$w_i = f_{phy}(n_i)/f_{bias}(n_i) = \exp((2n_i c - c^2)/2\sigma^2$$

EX8:
16.1,2,6,9,10,11

AIS: Adaptive IS

$$p_e = E(g(x)) \xrightarrow{IS} \hat{p}_e = \frac{1}{N}\sum_{i=1}^{N} g(x_i)w(x_i,\theta) \quad x_i \sim f_b(x,\theta)$$

$$\mathrm{var}(\hat{p}_e) = \frac{1}{N}(I(\theta) - p_e^2) \leftarrow I(\theta) = E(g^2(x)w(x,\theta)) = E_b(g^2(x)w^2(x,\theta))$$

g is the indicator of {x>T} for BER estimation problem

θ can be optimized on each epoch to achieve the best results… (IS11)

Papers that compare several pdf classes for different applications… (IS3)

Optimum and sub optimum parameters (IS4)