

Stat 471: Lecture 12

Importance Sampling II.

Moo K. Chung
mchung@stat.wisc.edu

October 3, 2003

1. Cauchy distribution. The general Cauchy distribution with parameter θ is defined as $f(x) = \frac{c}{\pi(x^2+c^2)} \sim \text{Cauchy}(c)$. CDF is $F(x) = \frac{1}{\pi} \arctan(\frac{x}{c}) + \frac{1}{2}$ and $F^{-1}(x) = -c \cot(\pi x)$. Based on this we can generate Cauchy.

```
cauchypdf
=inline('c./(pi*(c^2+x.^2))')
>>cauchyrnd
=inline('-c*cot(pi*rand(n,1))')
cauchyrnd
=Inline function:
cauchyrnd(c,n)=-c*cot(pi*rand(n,1))
```

2. In the previous lecture we showed how to estimate

$$p = \int_{-\infty}^{-2} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$$

using simple Monte-Carlo integration technique.

```
n=1000000
X=snorm(n);
index=find(X<-2);
g=zeros(n,1);
g(index)=1;
>>[mean(g) var(g)]
ans =
    0.0228    0.0223
```

Now we apply importance sampling technique.

$$\mathbb{E}_f g(X) = \mathbb{E}_h \left[g(X) \frac{f(X)}{h(X)} \right],$$

where $X \sim f$. Let $g(x)$ to be an index function such that $g(x) = 1$ if $x \in (-\infty, -2)$ and 0 otherwise. The rule of thumb is to choose h such that f/h is bounded and has a thicker tail than f . Further $|g|f/h$ should be almost constant with finite variance. Let $X \sim N(0, 1)$. Then $\hat{p} = \sum_{i=1}^n g(X_i)/n$. Let $h \sim \text{Cauchy}(1)$. Then

```
n=1000000;
X=cauchyrnd(1,n);
srnpdf
=inline('exp(-(x.^2/2)/sqrt(2*pi))');
g=(X<-2);
Y=g.*srnpdf(X)./cauchypdf(1,X);
[mean(Y) var(Y)]
ans =
    0.0228    0.0118
```

Compare the performance

```
n=1000
n=10000
n=100000
n=1000000
[mean(g) var(g)]
ans =
    0.0290    0.0282
    0.0219    0.0214
    0.0227    0.0222
    0.0228    0.0223
[mean(Y) var(Y)]
ans =
    0.0201    0.0104
    0.0231    0.0119
    0.0232    0.0120
    0.0228    0.0118
```

3. Trivia: the term Monte-Carlo was used by Ulam and von Neumann as a Los Alamos code word for the stochastic simulations they applied to building better atomic bombs.