HMKANG@UMICH.EDU

MODULE 2.2

MONTE CARLO METHODS

Monte-Carlo estimation

- **Goal is to calculate a deterministic quantity**
- **.. using an approximation with random sampling**

- **Example**
  - Let U be a Uniform(0,1) random variable

$$\theta = \mathrm{E}[g(U)] = \int_0^1 g(u)du$$

  - Let X be a random variable following a pdf *f(x)*

$$\theta = \mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

Characterizing Monte-Carlo estimation

- **Types of Monte-Carlo methods**
  - Crude Monte-Carlo
  - Importance sampling

- **Key properties of ideal Monte-Carlo methods**
  - Consistency : the estimate converges the true value
  - Unbiasedness : $\mathrm{E}\left[\hat{\theta}\right] = \theta$
  - Smaller (or minimal) variance

A crude Monte-Carlo method

To calculate $\theta = \mathrm{E}[g(U)] = \int_0^1 g(u)du$,

1. Generate B random samples $u_1, \cdots, u_B$ from Uniform(0,1).

2. Take the average as the estimate

$$\hat{\theta} = \frac{1}{B}\sum_{i=1}^{B} g(u_i)$$

Analysis of the crude Monte-Carlo method

- **Is it consistent?**
  - By Weak law of large number,

$$\hat{\theta} \xrightarrow{P} \lim_{B\to\infty} \frac{1}{B}\sum_{i=1}^{B} g(U_i) = \mathrm{E}[g(U)] = \theta$$

- **Is it unbiased?**

$$\mathrm{E}\left[\hat{\theta}\right] = \mathrm{E}\left[\frac{1}{B}\sum_{i=1}^{B} g(u_i)\right] = \mathrm{E}[g(U)] = \theta$$

- **How small is the variance?**

$$\mathrm{Var}\left[\hat{\theta}\right] = \frac{1}{B}\mathrm{Var}[g(U)] = \frac{\mathrm{E}[g^2(U)] - \theta^2}{B}$$

An alternative Monte-Carlo estimator

To calculate $\theta = \mathrm{E}[g(U)] = \int_0^1 g(u)du$, assuming $0 < g(u) < 1$,

1. Define a rectangle R between (0,0) and (1,1)
2. Initialize counters $h = 0$, $m = 0$
3. Uniformly select a random point (x,y) from R
4. If $y < g(x)$, increase h, otherwise increase m
5. Repeat step 3-4 for B times.
6. Return $\hat{\theta} = \frac{h}{h+m}$ as the estimate.

Analysis of the hit-and-miss method

- **Consistency**

$$h \sim Binomial(B, \theta)$$
$$\hat{\theta} = \frac{h}{B} \xrightarrow{P} \theta$$

- **Bias**

$$\mathrm{E}[\hat{\theta}] = \frac{1}{B}\mathrm{E}[h] = \theta$$

- **Variance**

$$\mathrm{Var}[\hat{\theta}] = \frac{1}{B^2}\mathrm{Var}[h] = \frac{B\theta(1-\theta)}{B^2} = \frac{\theta(1-\theta)}{B}$$

Which one is better?

$$\mathrm{Var}_{hitmiss}\left[\hat{\theta}\right] - \mathrm{Var}_{crude}\left[\hat{\theta}\right] = \frac{\theta(1-\theta)}{B} - \frac{\mathrm{E}\left[g^2(U)\right] - \theta^2}{B}$$

$$= \frac{\theta - \mathrm{E}\left[g^2(U)\right]}{B} \geq \frac{\theta - \mathrm{E}[g(U)]}{B} = 0$$

$(g^2(U) < g(U)$ because $0 < g(U) < 1)$

Crude Monte-Carlo method is better than hit-and-miss in terms of variance

An example of Monte-Carlo estimation

- Consider a normal distribution

$$X \sim N(0,1)$$

- We're interested in calculating

$$\Pr(X > 10) = 7.62 \times 10^{-24}$$

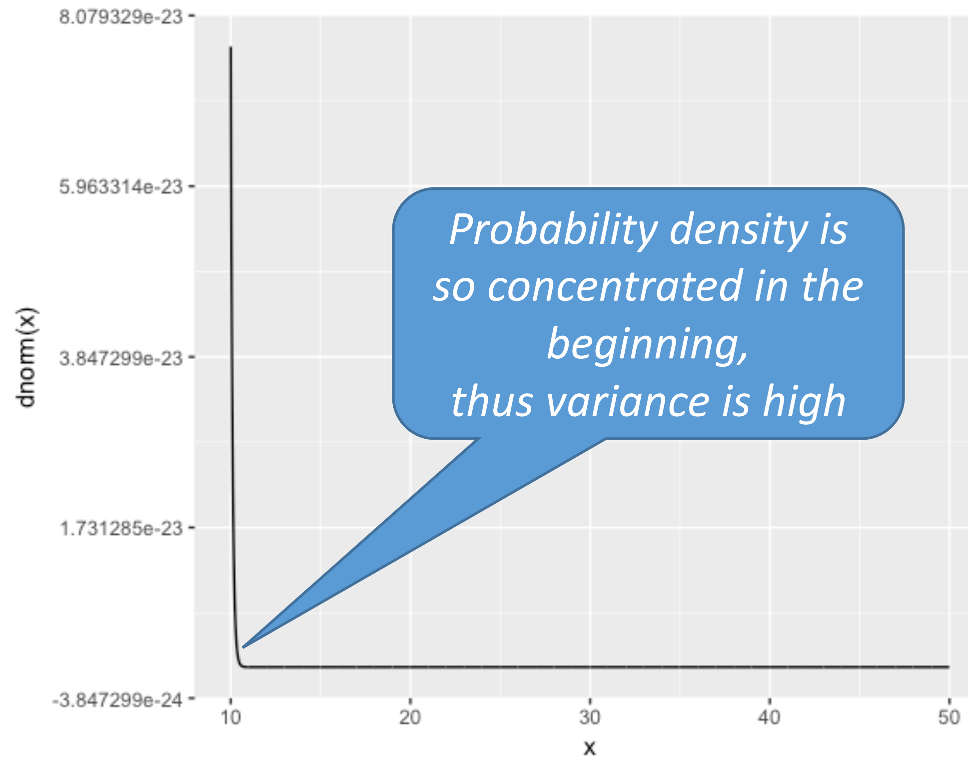
- Using Monte-Carlo methods, without using CDF

Possible solution with crude Monte-Carlo

- **Variance, when sampling from [10,10+w)** *(ignoring bias)*

$$\hat{\theta} = \frac{1}{B}\sum_{i=1}^{B} g(u)w$$

$$\mathrm{Var}[\hat{\theta}] = \frac{1}{B}\int_{10}^{10+w} [g(u)w - \theta]^2 du$$



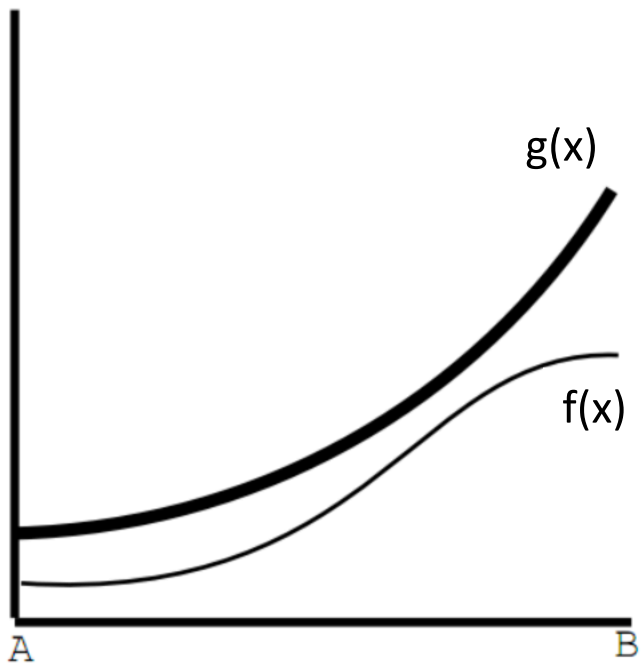
Importance sampling

- Let x_i be random variable,
- Let $f(x)$ be an arbitrary density function of the sampled random variable.

$$\theta = \int g(x)dx = \int \frac{g(x)}{f(x)} f(x)dx = \mathrm{E_f}\left[\frac{g(x)}{f(x)}\right]$$

$$\hat{\theta} = \frac{1}{B} \sum_{i=1}^{B} \frac{g(x_i)}{f(x_i)}$$

Key idea of importance sampling



- When g(x) deviates from the uniform distribution, the variance becomes very large
- The idea is to prevent sampling from high-variance distribution

Analysis of importance sampling

- **Bias**

$$E[\widehat{\theta}] = \frac{1}{B}\sum_{i=1}^{B} \mathrm{E_f}\left[\frac{g(x_i)}{f(x_i)}\right] = \frac{1}{B}\sum_{i=1}^{B}\theta = \theta$$

- **Variance**

$$\mathrm{Var}[\widehat{\theta}] = \frac{1}{B}\int\left[\frac{g(x)}{f(x)} - \theta\right]^2 f(x)dx$$

$$= \frac{\mathrm{E_f}\left[\frac{g(x_i)}{f(x_i)}\right]^2 - \theta^2}{B}$$

An importance sampling solution for *Pr(X>10)*

- Let f_0 be the pdf of standard normal distribution $N(0,1)$
- Let f_μ be the pdf of a normal distribution $N(\boldsymbol{\mu},1)$

1. Transform the problem into an unbounded integration problem

$$\theta = \Pr(X > 10) = \int_{10}^{\infty} f_0(x)dx = \int I(x > 10) f_0(x) dx$$

2. Sample B random values from $N(\boldsymbol{\mu},1)$ where $\boldsymbol{\mu}$ is close to 10
3. Estimate the probability as the weighted average

$$\hat{\theta} = \frac{1}{B}\left[I(x_i \geq 10) \frac{f_0(x_i)}{f_\mu(x_i)} \right]$$

Implementation of possible algorithms

```r
## pnormUpper() function to calculate Pr[x>t] using n random samples
pnormUpper <- function(n, t) {
  lo <- t
  hi <- t + 50 ## w=50 was used hi is a reasonably large number
  ## accept-reject sampling
  r <- rnorm(n) ## random sampling from N(0,1)
  v1 <- sum(r > t)/n  ## count how many meets the condition
  ## Monte-Carlo integration
  u <- runif(n,lo,hi) ## uniform sampling [t,t+50]
  v2 <- mean(dnorm(u))*(hi-lo) ## Monte-Carlo integration
  ## importance sampling using N(t,1)
  g <- rnorm(n,t,1) ## sample from N(t,1)
  v3 <- sum( (g > t) * dnorm(g)/dnorm(g,t,1)) / n; ## take a weighted average
  return (c(v1,v2,v3)) ## return three values
}
```

100 experiments with B=1,000

```
r <- 100    ## number of repeated experiments
B <- 1000   ## size of boostrap
t <- 10     ## Pr(t>10)
gold <- pnorm(t,lower.tail=FALSE) ## gold standard answer
emp <- matrix(nrow=r,ncol=3) ## matrix containing empirical answers
for(i in 1:r) { emp[i,] <- pnormUpper(B,t) } ## repeat r times
m <- colMeans(emp)
s <- apply(emp,2,sd)
rst <- list(gold.standard=gold,
            estimates=m,
            abs.bias=m-gold,
            abs.std=s,
            rel.bias=(m-gold)/gold,
            rel.std=s/gold)
print(rst)
```

Results

```
$gold.standard
[1] 7.619853e-24

$estimates
[1] 0.000000e+00 7.949908e-24 7.572652e-24

$abs.bias
[1] -7.619853e-24  3.300552e-25 -4.720110e-26

$abs.std
[1] 0.000000e+00 3.860388e-24 8.233907e-25

$rel.bias
[1] -1.00000000  0.04331517 -0.00619449

$rel.std
[1] 0.0000000 0.5066224 0.1080586
```

Adding results with B=10,000

```
$gold.standard
[1] 7.619853e-24

$estimates
[1] 0.000000e+00 7.949908e-24 7.572652e-24

$abs.bias
[1] -7.619853e-24  3.300552e-25 -4.720110e-26

$abs.std
[1] 0.000000e+00 3.860388e-24 8.233907e-25

$rel.bias
[1] -1.00000000  0.04331517 -0.00619449

$rel.std
[1] 0.0000000 0.5066224 0.1080586
```

```
$gold.standard
[1] 7.619853e-24

$estimates
[1] 0.000000e+00 7.551066e-24 7.587339e-24

$abs.bias
[1] -7.619853e-24 -6.878701e-26 -3.251440e-26

$abs.std
[1] 0.000000e+00 1.033607e-24 2.300938e-25

$rel.bias
[1] -1.000000000 -0.009027341 -0.004267065

$rel.std
[1] 0.00000000 0.13564662 0.03019662
```

Summary

- **Monte-Carlo estimation**
  - Use random sample to approximately calculate a complex quantitty

- **Types of Monte-Carlo estimation**
  - Crude Monte-Carlo
  - Accept-reject (hit-and-miss) Monte-Carlo estimation
  - Importance sampling