September 16, 2016

# Monte Carlo and Variance computations

*In this class:*

- *Summary of class performance so far.*
- *Review of Monte Carlo Integration + Variance computation*
- *Review of Midterm topics*
- *Review of Floating Point / general review.*

*Next class*

Review

G&C – Chapters 1, 2, 3, 5

*Next next class*

Midterm

G&C – Chapters 1, 2, 3, 5

# Monte Carlo integration

Examples in Julia!

# The big idea in math

$$\mathrm{E}[X] = \int_{\text{domain}} x p(x)\, dx$$

Let $Y$ be a r.v. where $Y = f(X)$

$$\mathrm{E}[Y] = \int_{\text{domain of } X} f(x) p(x)\, dx$$

# The variance error

Your textbook makes a critical numerical mistake!

So does Google!

Analytic quantity

$$\mathrm{Var}[X] = E[(X - E[X])^2]$$

$$\mathrm{Var}(x_1, \ldots, x_n) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mathrm{mean})^2$$

Statistical quantity

Step 1: Compute mean
Step 2: Compute variance
=> Look at data twice ☹

# The variance error

$$\begin{aligned}\mathrm{Var}[X] &= E[(X - E[X])^2] \\ &= E[X^2 - 2E[X]X + E[X]^2] \\ &= E[X^2] - 2E[X]E[X] + E[X]^2 \\ &= E[X^2] - E[X]^2\end{aligned}$$

$$\mathrm{Var}(x_1, \ldots, x_n) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\sum_{i=1}^{n} x_i\right)$$

Look at data once!

# The variance error

Your textbook makes a critical numerical mistake!

So does Google!

$$\mathrm{Var}[X] \approx \left(\mathrm{E}[X^2] - \mathrm{E}[X]^2\right)$$

```
function badvar(x)
ex2 = 0.0; ex = 0.0; n = length(x);
for i=1:n
  ex2 = ex2 + x[i]^2
  ex = ex + x[i]
end
return 1.0/(n-1)*(ex2 - (ex)^2/n)
```

# One-pass, online, streaming variance

This formula is known to be bad!

There are better alternatives!

```
function goodvar(x)
n = length(x); mean = 0.0; m2 = 0.0; N = 0;
for i=1:n
    N = N + 1
    delta = x[i] - mean
    mean = mean + delta/N
    m2 = m2 + delta*(x[i]-mean)
end
return m2/(n-1)
```

… the midterm …

(if time)

# Background I assume

Linear algebra

Calculus

Differential equations

Discrete math

Programming

Probability

I'll try to remind you what you need to know

# Topics we've covered

**Week 1**

Details of the class

History of numerical computing

The importance of numerical computing

Mathematical modeling

The XKCD raptor problem

Google's PageRank

**Week 2**

Variables and expressions in Matlab

Matrix and vector operations in Matlab

Control flow in Matlab

**Week 3**

The need for floating point

IEEE Floating point representations

General floating point systems

How to add/sub/mult with floating point

IEEE Rounding modes

Floating point guarantees

Floating point properties

IEEE Exceptions

Problematic floating point computations

Correct norms

**Week 4**

Monte Carlo methods

The Monte Hall prob

Integrating a circle.

Google's random surfer

Monte Carlo integration + variance computations

Central limit theorem & accuracy

# Homework questions

**Homework 1**

Drunkard's walk (or random walk on a line)

A simple economy

Acceleration and raptors

Writing matrices for search engines

Simple Matlab operations

Mandelbrot

**Homework 2**

Floating point representations

Converting random number generators

Nearest number

Relative roundoff error

Floating point exceptions

Fibonacci roots & floating point

Fun with floats!

Monte Carlo integration

The Birthday paradox

Random walks and birthdays