A Whirlwind Tour of Computer Simulation

Dave Goldsman

Georgia Tech

Atlanta, GA, USA

sman@gatech.edu

www.isye.gatech.edu/~sman

5/13/19

0. Getting to Know You

Goals:

* Contact Info
* Course Info

Contact Info

Instructor:

Prof. Dave Goldsman

School of ISyE, Georgia Tech

sman@gatech.edu

www.isye.gatech.edu/~sman

(404)822-8949

Office Hours: Anytime you can get me! ☺

Course Objectives

- Intro to simulation models and simulation studies.
- Organization of simulation languages. Modeling with Arena – or any comprehensive simulation package with animation capabilities.
- Statistical aspects including input analysis, random variate generation, output analysis, and variance reduction techniques.

Prerequisites

- You *ought* to know some probability and statistics.
- You should be familiar with some programming language and maybe even a spreadsheet package.
- Good News: Don't panic! I'll make the course as self-contained as possible!

Suggested Texts / Resources

- Banks, J., Carson, J.S., Nelson, B.L., and Nicol, D.M., *Discrete-Event System Simulation*, 5th ed., Prentice Hall, Upper Saddle River, 2010.
- Law, A.M., *Simulation Modeling and Analysis*, 5th ed., McGraw-Hill Education, New York, 2015.
- Kelton, W.D., Sadowski, R.P., and Zupick, N.B., *Simulation with Arena*, 6th edition, McGraw-Hill, New York, 2015.
- Free Arena software download: https://www.arenasimulation.com/academic/students

Course Notes

- I'll provide pretty extensive notes on a course website.
- This doesn't mean that you can simply print out the notes and skip the fun class…
- …so we'll also do some hands-on work.

Computer Programming

We’ll do a reasonable amount of computer programming. You’ll have some choice, but you can expect to use:

- A spreadsheet package, e.g., Excel.
- Some spreadsheet add-ons.
- A “real” language, e.g., Matlab or Python.
- A simulation language, e.g., Arena.

Syllabus – Let's Get Going!

- Introduction
- Calculus, Probability, and Statistics Boot Camp (Law, Chapter 4)
- Hand Simulations; Spreadsheet Simulations
- General Modeling Concepts (Law, 1&2)
- Verification and Validation (Law, 5)
  - Is your simulation doing what you think?

Syllabus – Fun with Arena

- Arena Basics (KSZ, Chapter 4)
- A Generic Call Center in Arena (KSZ, 5)
- An Inventory Model (KSZ, 5)
- A Manufacturing Center (KSZ, 6)
- Entity Transfers in Arena (KSZ, 7)
- More-Advanced Arena Stuff (KSZ, 8)

Syllabus – Randomness

- Random Number Generation (Law, 7)
  - generate “randomness” on a computer
- Random Variate Generation (Law, 8)
  - single random variables
  - multivariate random variables
  - random processes
  - financial models

Syllabus – Statistical Issues

- Input Analysis (Law, 6)
  - What should drive the simulation?
- Output Analysis (Law, 9)
  - Analyze what comes out of the simulation.
- Comparing Systems (Law, 10)
  - Which system is better / best?
- Variance Reduction + Other Cool Stuff

Module 1 – A Whirlwind Tour!

1. Intro to Simulation
2. Some Easy Examples
3. Generating Randomness
4. Analyzing Randomness
5. Some Bigger Examples
6. Selecting a Simulation Language

1. Intro to Simulation

- Models are high-level representations of the operation of a real-world process or system.
- Our concern will be with models that are
  - Discrete (vs. continuous)
  - Stochastic (vs. deterministic)
  - Dynamic (vs. static)
- How can you "solve" a model?
  - Analytic methods
  - Numerical methods
  - Simulation methods

Examples of Models

- Toss a stone off of a cliff. You can model its position via the usual physics equations – analytical models.
- Model the weather. Too tough for exact analytical models, so you might use numerical methods.
- Add a little randomness, and you may have to resort to a simulation model (plenty of examples coming up).

What is Simulation?

- Simulation is the imitation of a real-world process or system over time.
- Simulation involves the generation of an artificial history to draw inferences concerning the operating characteristics of the real system that is represented.

Simulation is…

- One of the top three industrial engineering / operations research / management science technologies
- Used by academics and practitioners on a wide array of theoretical and applied problems
- An indispensable problem-solving methodology

What is It Good for?

- Describe and analyze system behavior (systems can be real or conceptual)
- Ask "what if" questions about the system
- Aid in system design and optimization
- Pretty much anything can be simulated
  - Fixed systems with no "customer" arrivals (e.g., estimate stock option prices), or
  - Service systems like Queueing Centers, Manufacturing Processes, Supply Chains, Health Systems, etc.

Reasons to Simulate

- Will the system accomplish its goals?
- Current system won’t accomplish its goals. Now what?
- Need incremental improvement
- Create a specification or plan of action
- Resolve disputes
- Solve a problem, like a bottleneck
- Sell an idea

Advantages of Simulation ☺

- Can study models too complicated for analytical or numerical treatment
- Study detailed relations that might be lost in the analytical or numerical treatment
- Use as a basis for experimental studies of systems
- Use to check results and give credibility to conclusions obtained by other methods
- Reduce design blunders
- Really nice demo method
- (Sometimes) very easy
- …

Disadvantages ☹

- Sometimes not so easy
- Sometimes time consuming / costly
- Simulations give “random” output (and lots of misinterpretation of results is possible)
- To do a certain problem, better methods than simulation may exist.
- …

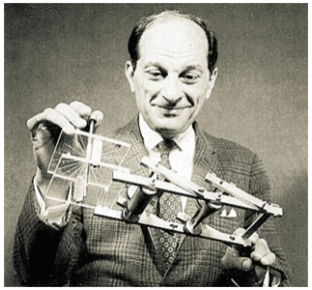
History

- 1777 – Buffon's Needle Problem

- Early 1900's – Beer and Student's t distribution

More History

- 1946 – Ulam, Metropolis, von Neumann, and the H-Bomb



- 1960's – Industrial Applications
  - Manufacturing
  - Queueing Models

Recent History

- Development of Simulation Languages
  - Easy-to-use modeling tools
  - Graphics

← Harry Markowitz

- Rigorous Theoretical Work
  - Computational algorithms
  - Probabilistic and statistical methods

Origins: Mfg/Material Handling

- Simulation is the technique of choice
  - Calculates movement of parts and interaction of system components
  - Evaluates flow of parts through the system
  - Examines conflicting demand for resources
  - Examines contemplated changes before their introduction
  - Eliminates major design blunders

Typical Questions

- What will be the throughput?
- How can we change it?
- Where are the bottlenecks?
- Which is the best design?
- What is the reliability of the system?
- What is the impact of breakdowns?

What I Used to Think Simulation Was…

$$\sum_{i=nc_{k-1}+1}^{nc_k} \sum_{j=nc_k+1}^{nc_{k+1}} h_i h_j R_{j-i}$$

$$= \left\{ \sum_{i=nc_{k-1}+1}^{n(c_k-\epsilon_k)} \sum_{j=nc_k+1}^{nc_{k+1}} + \sum_{i=n(c_k-\epsilon_k)+1}^{nc_k} \sum_{j=nc_k+1}^{n(c_k+\epsilon_k)} + \sum_{i=n(c_k-\epsilon_k)+1}^{nc_k} \sum_{j=n(c_k+\epsilon_k)+1}^{nc_{k+1}} \right\} h_i h_j R_{j-i}$$

$$= \sum_{i=nc_k-n\epsilon_k+1}^{nc_k} \sum_{j=nc_k+1}^{nc_k+n\epsilon_k} h_i h_j R_{j-i} + O(n^4 \delta^{n\epsilon_k})$$

By complete enumeration as done in Appendix A.2, we get

$$\sum_{i=nc_k-n\epsilon_k+1}^{nc_k} \sum_{j=nc_k+1}^{nc_k+n\epsilon_k} h_i h_j R_{j-i} = \sum_{i=1}^{n\epsilon_k} \sum_{j=1}^{i} h_{nc_k-j+1} h_{nc_k-j+i+1} R_i + O(n^4 \delta^{n\epsilon_k})$$

Therefore,

$$\mathrm{E}[\mathcal{A}^2_{\mathcal{C}_m}(f;n)] = \mathrm{E}[\mathcal{A}^2(f;n)] - \frac{4}{n^3} \sum_{k=1}^{m} \sum_{i=1}^{n\epsilon_k} \sum_{j=1}^{i} h_{nc_k-j+1} h_{nc_k-j+i+1} R_i + O(n\delta^{n\epsilon}) \qquad (9)$$

Actual Applications

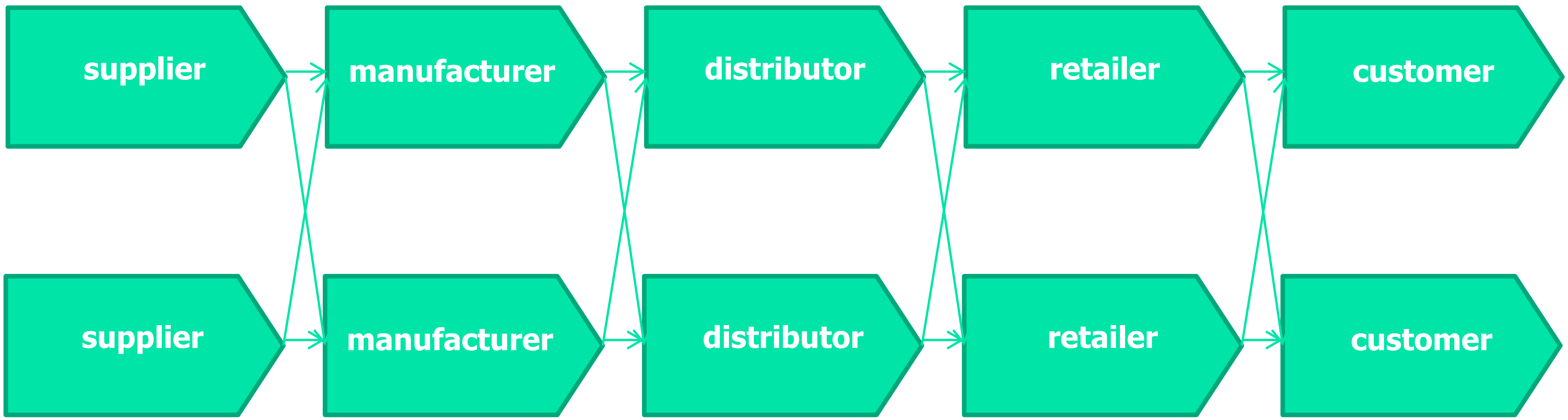
- Manufacturing
  - Automobile Production Facility
  - Carpet Production Facility
- Queueing Problems
  - Call Center Analysis
  - Fast Food Drive-Thru
  - Fast Food Drive-Thru Call Center
  - Airport Security Line

Ex: Generic Queueing System

- Combination meat processing center / fast-food restaurant / amusement park

← Denis (8) and Mina (6)

Application: Supply Chains



- Think of supply chain as a network where any node can potentially connect to any other node.
- In order to be lazy, let's have things flow from left to right. Looks a bit like a flowchart, eh?

Supply Chains (cont'd)

- 1985: Professor and CEO of SC company: "Why would anyone want to simulate a SC?"
- Now: Many folks are developing SC tools with simulation capability.
  - Use simulation to determine how much value-added a forecasting application provides.
  - Use simulation to analyze how SC randomness or model errors affect a proposed solution. Is it robust? What's the best solution?

Applications (cont'd)

- Inventory and Supply Chain Analysis
- Financial Analysis
  - Portfolio Analysis
  - Options Pricing
- Traffic Simulation
- Airspace Simulation
- Service Sector
- Health Systems

Health Systems Applications

- Patient Flow in a Hospital
- Hospital Room Allocation
- Optimization of Doctor / Nurse Scheduling
- Procurement of Supplies
- Disease Surveillance
- Propagation of Disease Spread
- Humanitarian Logistics

Surveillance Application

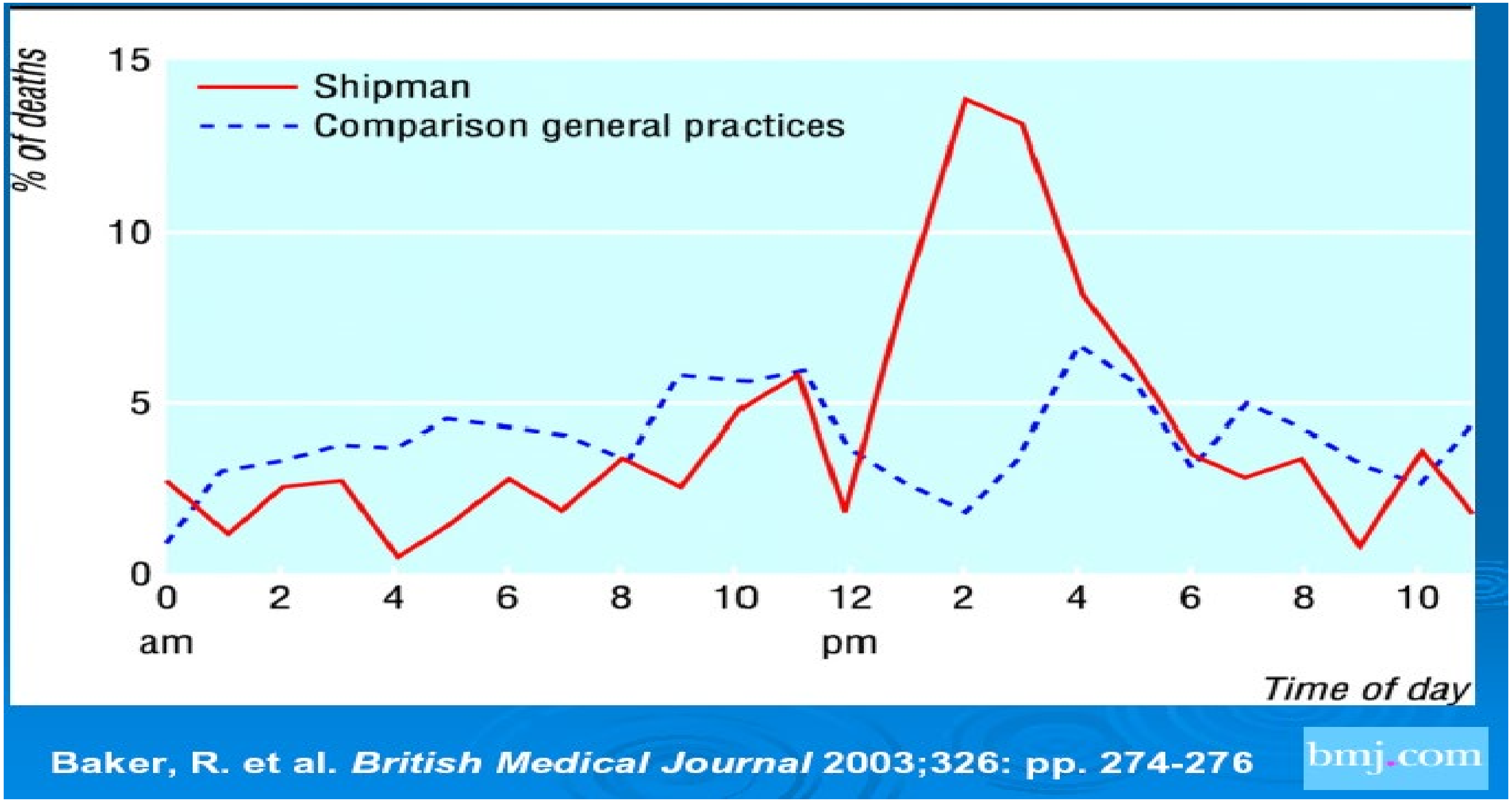
- Use simulation to monitor certain time series
- Predict issues as or before they happen
- Is a disease in the process of becoming an outbreak?
- When is something out of the ordinary occurring?
- Take advantage of HUGE data sets.

Who Is This Good-Looking Guy?

Dr. Harold Shipman

- Killed his patients using heroin overdoses (2^{nd} ranked serial killer in UK)
- Was caught after carelessly revising a patient's will, leaving all her assets to himself (don't use your own typewriter next time)
- His computer records were doctored to show his patients had needed the morphine. However, the software recorded the dates of these modifications.
- Hung himself in prison, never confessed

Dr. Harold Shipman



How Does Simulation Help?

- Surveillance relies on a series of sequential statistical hypothesis tests, where the null hypothesis is “no disease” (or “no murder”)
- The statistics involved in the test have very difficult distributions, even under the null hyp H_0
- Simulation can be used to approximate the probability distributions of the statistics under H_0
- If sampling reveals deviations from this simulated distribution, you reject H_0

Thanks to Prof. Bill Woodall for this example!

2. Some Easy Examples

- Happy Birthday
- Let's Make Some Pi
- Fun With Calculus
- Random Numbers Gone Bad
- Queues 'R Us
- Stock Market Follies

Happy Birthday

- How many people do you need in a room in order to have a 50% chance that at least two will have the same birthday?
  - 9
  - 23
  - 42
  - 183

Birthday Paradox

Birthday Paradox

1. Enter Seed:

47475

2. Click Start Button:

START

New Birthday Date:

June 22

Total # Birthdays:

24

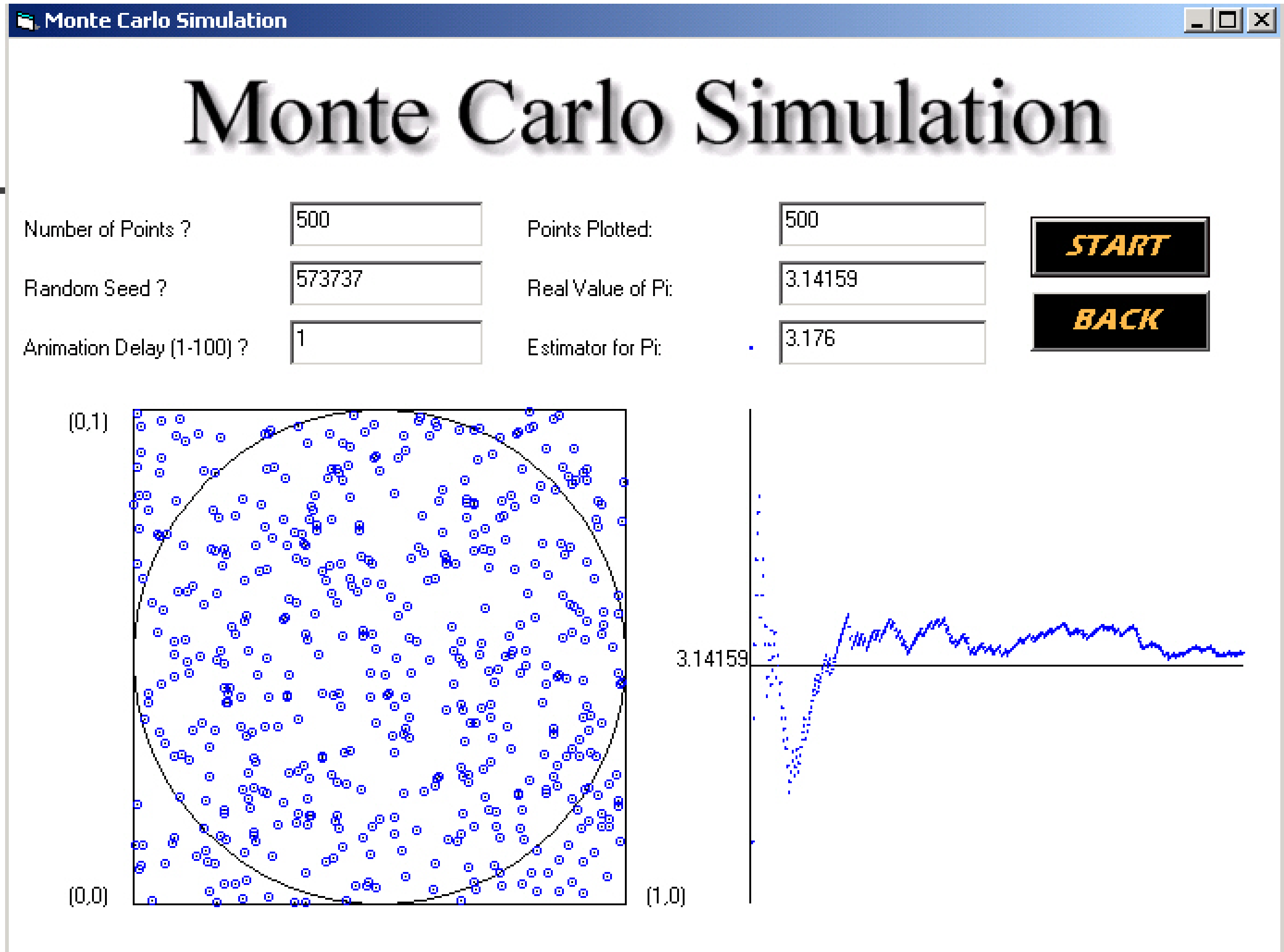
!!! MATCH !!!

BACK

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1		happy				happy			happy			
2												
3		happy			happy							
4										happy		
5										happy		
6								happy				
7		happy										
8							happy					
9												
10												
11						happy						
12												
13					happy							
14												
15		happy					happy					
16					happy							
17				happy								
18												
19							happy					
20												
21												
22						happy						
23		happy										
24				happy								
25												
26										happy		
27												
28												
29		XXXX						happy				
30	happy	XXXX										
31		XXXX		XXXX		XXXX			XXXX		XXXX	

Let's Make Some Pi

- Use Monte Carlo simulation to estimate π.
- Idea:
  - Area of a unit square is 1.
  - Area of an inscribed circle is $\pi/4$.
  - Probability that a dart thrown at the square will land in the circle is $\pi/4$.
  - Throw lots of darts. Proportion that will land in circle should approach $\pi/4$.
  - Multiply proportion by 4 to estimate π.



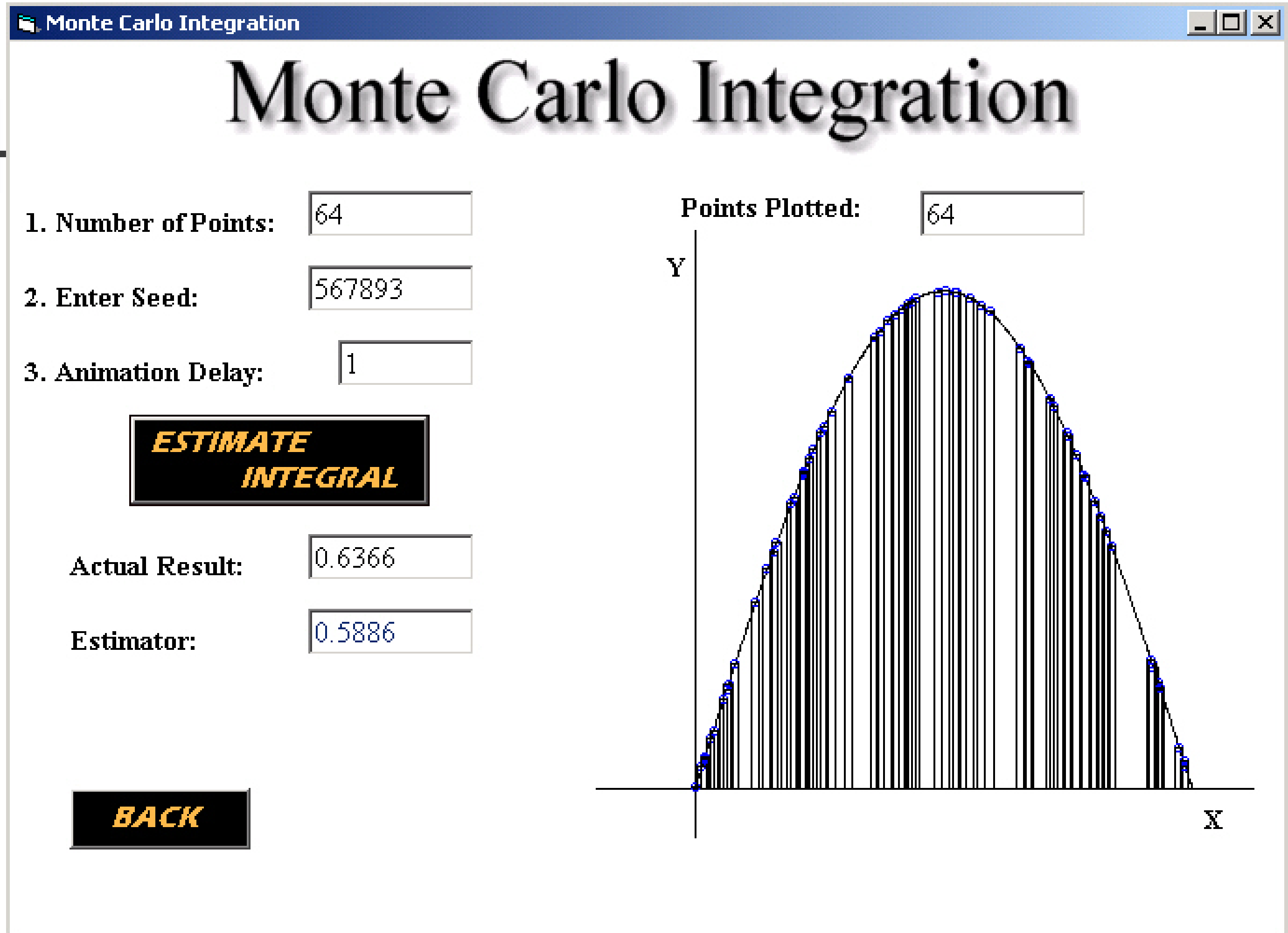
Yesterday, $\sqrt{-1}\ 2^3\ \Sigma\ \pi$, and it was really tasty!

Fun With Calculus

- Use simulation to integrate

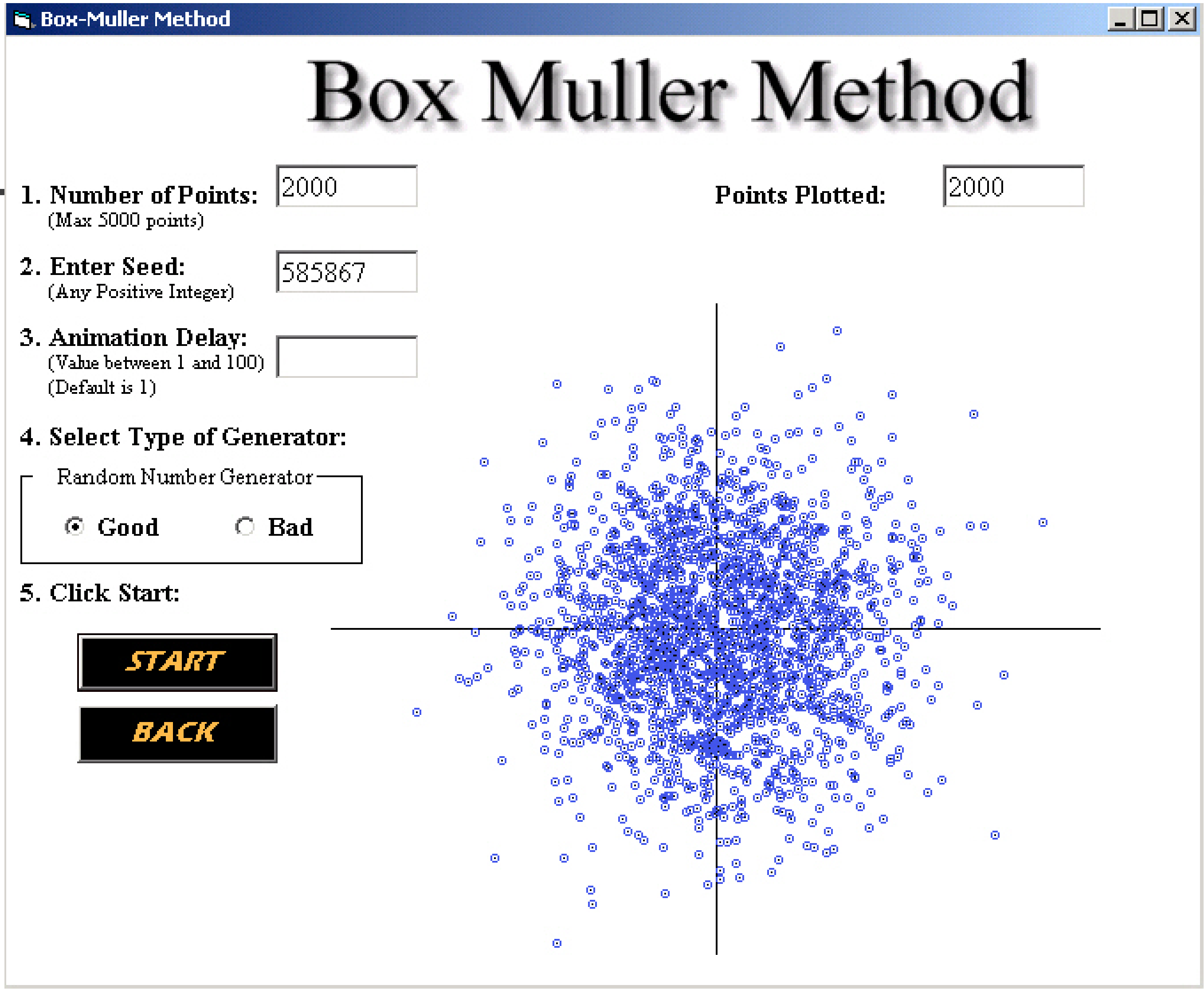
  $$f(x) = \sin(\pi x) \text{ over } [0,1].$$

- Idea:
  - Sample n rectangles.
  - Each is centered randomly on [0,1] and has width 1/n and height f(x).
  - Add up areas.
  - Make n really, really big.
  - Sum of areas approaches integral of f(x).

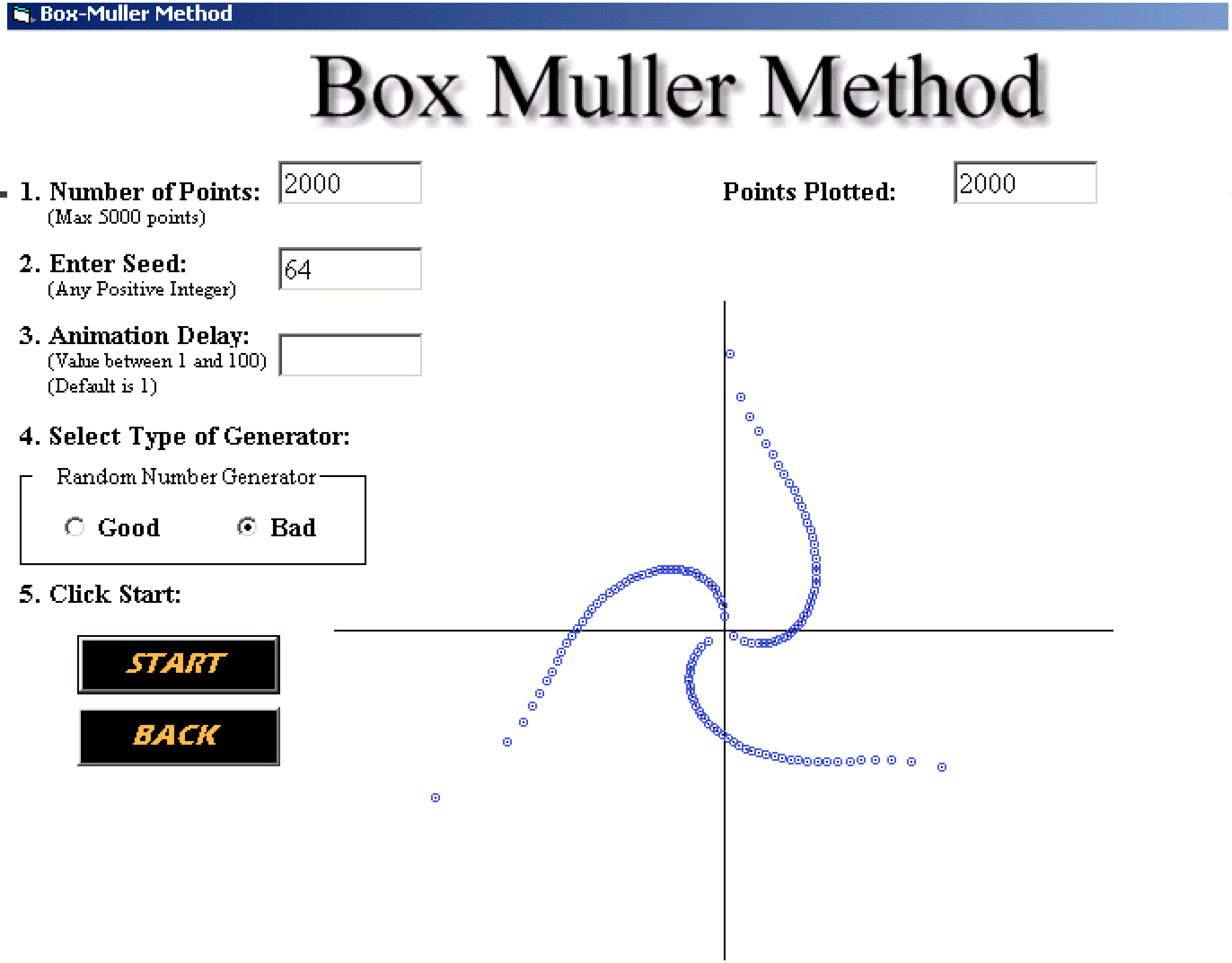

Monte Carlo Integration
Monte Carlo Integration
1. Number of Points:
64
2. Enter Seed:
567893
3. Animation Delay:
1
ESTIMATE INTEGRAL
Actual Result:
0.6366
Estimator:
0.5886
BACK
Points Plotted:
64
Y
X

Random Numbers Gone Bad

- See what happens when you use a bad random number generator.
- Idea:
  - Simulate heights vs weights.
  - Should be a 2-D bell curve (normal distribution) with most observations in the middle and some on the outside.
  - Do observations “look” random?



Box-Muller Method
Box Muller Method
1. Number of Points:
(Max 5000 points)
2000
Points Plotted:
2000
2. Enter Seed:
(Any Positive Integer)
585867
3. Animation Delay:
(Value between 1 and 100)
(Default is 1)
4. Select Type of Generator:
Random Number Generator
Good
Bad
5. Click Start:
START
BACK


Box-Muller Method
Box Muller Method
1. Number of Points:
(Max 5000 points)
2000
Points Plotted:
2000
2. Enter Seed:
(Any Positive Integer)
64
3. Animation Delay:
(Value between 1 and 100)
(Default is 1)
4. Select Type of Generator:
Random Number Generator
Good
Bad
5. Click Start:
START
BACK

Queues 'R Us

- Single-server queue at McWendys.
- Customers show up, wait in line, get served first-in-first-out.
- What happens as arrival rate approaches service rate?
  - Nothing much?
  - Line gets pretty long?
  - Hamburgers start to taste better?

Queues 'R Us (cont'd)

- Can analyze queues via simulation.
- Can analyze via numerical or exact methods.

- Fun fact: Notice anything interesting about the word "queueing"? How about "queueoid"?

M/M/1 Queue Simulation

MM1 Queue Simulation

Interarrival Mean? 4

Service Mean? 3

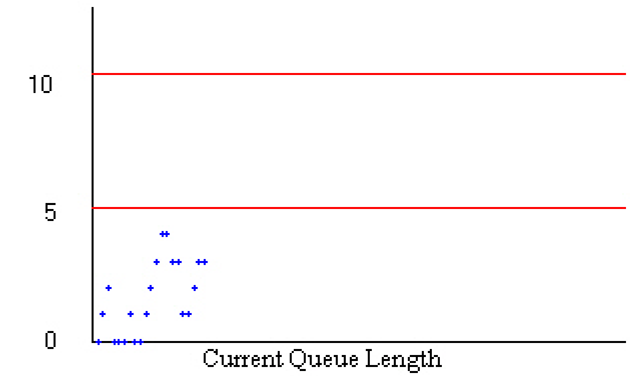
STOP

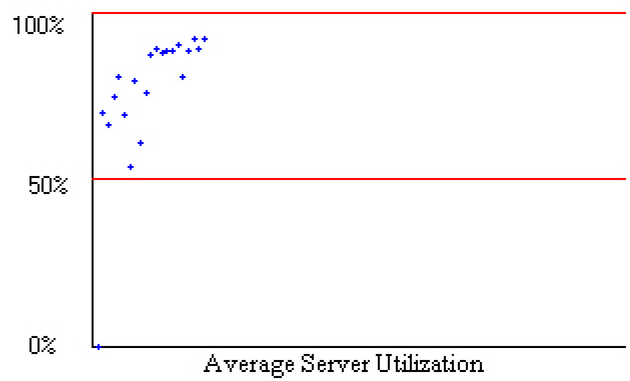
NEXT CUSTOMER

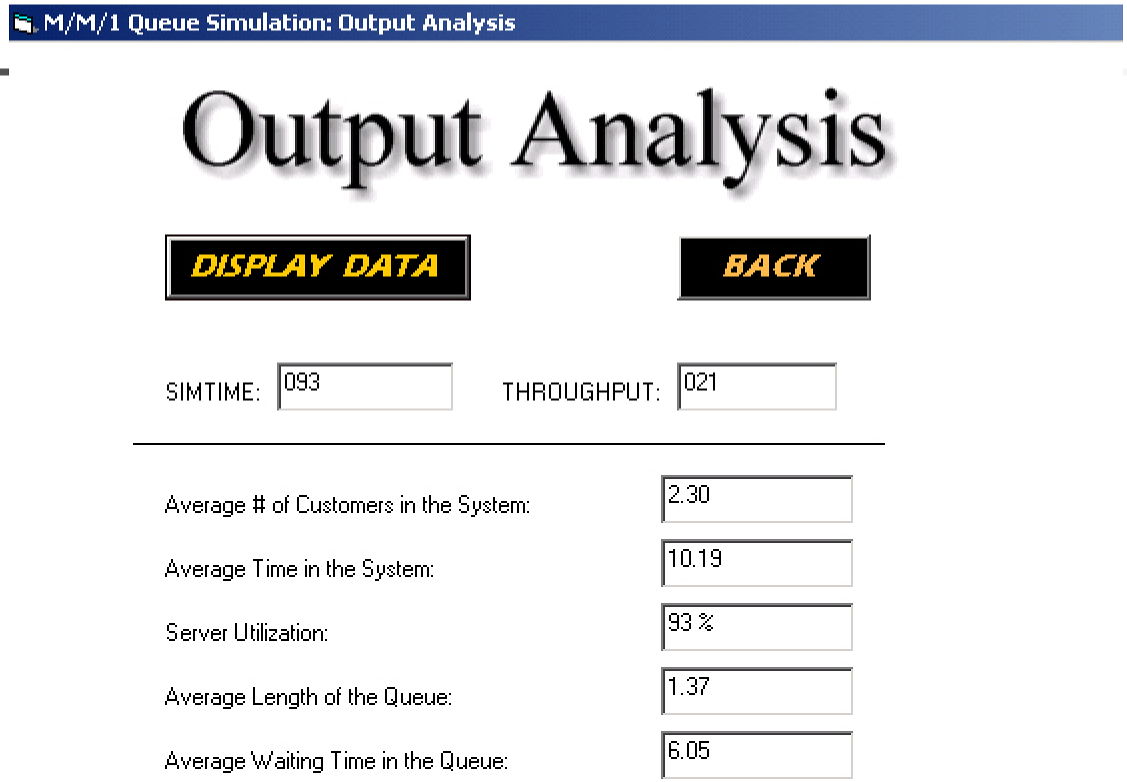
Interarrival time 6

Service Time 1

i	Arrive	Start	Service	Leave	Wait
1	0	0	7	7	0
2	2	7	3	10	5
3	5	10	5	15	5
4	18	18	2	20	0
5	20	20	1	21	0
6	24	24	2	26	0
7	25	26	11	37	1
8	37	37	2	39	0
9	39	39	15	54	0
10	43	54	9	63	11
11	47	63	2	65	16
12	50	65	1	66	15
13	51	66	2	68	15
14	61	68	2	70	7
15	65	70	2	72	5
16	66	72	1	73	6
17	72	73	10	83	1
18	74	83	4	87	9
19	76	87	1	88	11
20	77	88	4	92	11
21	83	92	1	93	9



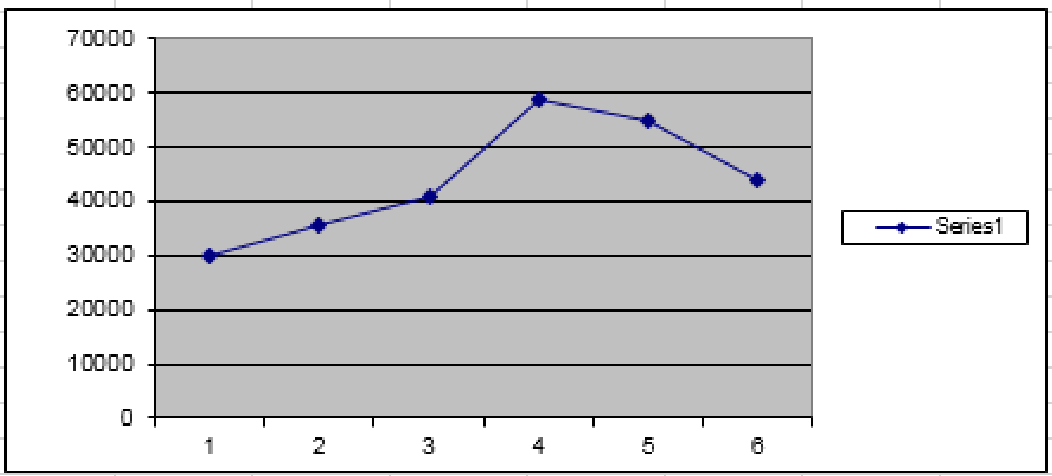



M/M/1 Queue Simulation: Output Analysis
Output Analysis
DISPLAY DATA
BACK
SIMTIME: 093
THROUGHPUT: 021
Average # of Customers in the System: 2.30
Average Time in the System: 10.19
Server Utilization: 93 %
Average Length of the Queue: 1.37
Average Waiting Time in the Queue: 6.05

Stock Market Follies

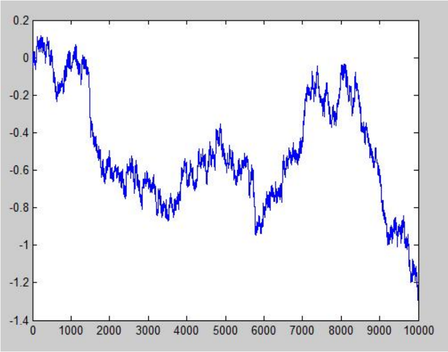
- Simulate a portfolio of various stocks
- Stock prices change randomly from year to year, with various volatilities
- Can consider different mixes for portfolio
- Simple spreadsheet application

	mean	sd	year0	year1	year2	year3	year4	year5
general year performance (so stocks are correlated)				1.18	1.14	1.14	1.05	0.89
energy	0.05	0.30	5000	5285	5155	6404	8546	8190
pharmaceuticals	0.06	0.20	5000	5221	7531	10877	9214	9417
entertainment	0.04	0.10	5000	6008	6712	7157	7635	6493
insurance	0.07	0.05	5000	6513	7317	8605	10318	9915
banking	0.06	0.30	5000	5225	5418	8557	9889	7269
computer technology	0.18	0.50	5000	7224	8873	16927	9027	2516
Totals			30000	35475.9	41005.6	58527.2	54630.4	43800.1



Taking a Random Walk

- Take a normal step up or down every time unit and plot where you are as time progresses
  - This "random walk" converges to Brownian motion
  - Einstein and Black+Scholes won Nobel Prizes for this research.



3. Generating Randomness

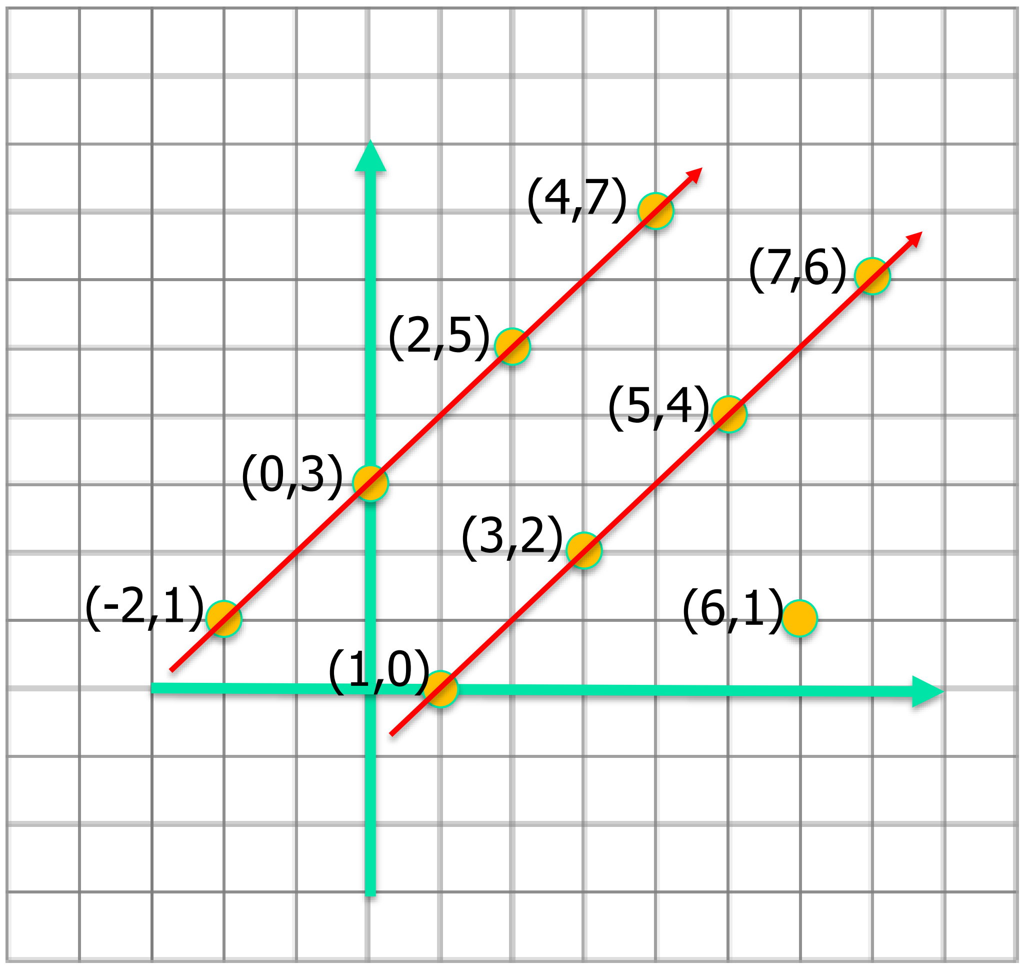
- Need random variables to run the simulation, e.g., interarrival times, service times, etc.
- Generate Unif(0,1) pseudo-random numbers (PRN's)
  - Use a deterministic algorithm
  - Not really random, but appear to be
- Generate other random variables
  - Start with Unif(0,1)'s
  - Apply certain transformations to come up with just about any other type of random variable

Unif(0,1) PRN's

- Deterministic algorithm
- Example: Linear Congruential Generator
  - Choose an integer "seed," $X(0)$
  - Set $X(i) = (a\, X(i-1) + c) \bmod(m)$, where a, c, and m are carefully chosen constants, and mod is the modulus function.
  - Set the ith PRN as $U(i) = X(i)/m$

Unif(0,1) PRN's

- Pretend Example:
  - Set $X(i) = (5\, X(i-1) + 3) \bmod(8)$, with $X(0) = 0$
  - Then $X(1) = 3 \bmod 8 = 3$
  - $X(2) = 2$, $X(3) = 5$, $X(4) = 4$, $X(5) = 7$, etc.
  - So $U(1) = X(1)/m = 3/8$
  - $U(2) = 2/8$, $U(3) = 5/8$, etc.
  - Numbers don't look particularly random.



Plot of (X(i-1) , X(i))

The random numbers fall mainly on the planes

Unif(0,1) PRN's

- Real Example
- $X(i) = 16807\ X(i-1) \bmod (2^{31} - 1)$
- $U(i) = X(i) / m$
  - This generator is used in a number of simulation languages
  - Has nice properties, including long "cycle times"
  - Better generators are out there

Generating Other Random Variates

- Start with $U(i) \sim \mathrm{Unif}(0,1)$
- Apply some appropriate transformation
- Example: $-(1/\lambda)\ \ln(U(i)) \sim \mathrm{Exp}(\lambda)$
  - Inverse transform method – can use this for various important distributions
  - Many other more-sophisticated methods available, e.g., Box-Muller method for normals

4. Analyzing Randomness

- Simulation output is nasty. Consider consecutive customer waiting times.
  - Not normally distributed – usually skewed
  - Not identically distributed – patterns change throughout the day
  - Not independent – usually correlated
- Can’t analyze via “usual” statistics methods!

Analyzing Randomness

- Two general cases to consider
  - Terminating Simulations
    - Interested in short-term behavior
    - Example: Avg customer waiting time in a bank over the course of a day
    - Example: Avg # of infected victims during a pandemic
  - Steady-State Simulations
    - Interested in long-term behavior
    - Example: Continuously running assembly line

Terminating Simulations

- Usually analyzed via Independent Replications
  - Make independent runs (replications) of the simulation, each under identical conditions
  - Sample means from each replication are assumed to be approximately i.i.d. normal
  - Use classical statistics techniques on the i.i.d. sample means (not on the original observations)

Steady-State Simulations

- First deal with initialization (start-up) bias.
  - Usually “warm up” simulation before collecting data
  - Failure to do so can ruin statistical analysis
- Many methods for dealing with steady-state data
  - Batch Means
  - Overlapping Batch Means / Spectral Analysis
  - Standardized Time Series
  - Regeneration

Steady-State Simulations

- Method of Batch Means
  - Make one long run (vs. many shorter rep's)
  - Warm up simulation before collecting data
  - Chop remaining observations into contiguous batches
  - Sample means from each batch are assumed to be approximately i.i.d. normal
  - Use classical statistics techniques on the i.i.d. batch means (not on the original observations)

Batch Means Example

- Get a confidence interval (CI) for the mean of an autoregressive process
  - This process is highly correlated
  - CI's via classical statistics ("CLT" method on next page) result in severe undercoverage
  - Look at batch means and overlapping batch means
  - BM and OBM do better than CLT.

Batch Mean Simulation

Batch Means Method

1. **Enter Seed:** 74737
(Any Positive Integer)

2. **Enter Pho:** .9
(A value < 1 and > -1)
(Default is 0)

3. **# Experiments:** 100
(100 is Default)

Batch Means

4. **Batch Size (m):** 100
(100 is Default)

5. **# of Batches (b):** 20
(50 is Default) NOTE!!!
N= m*b <= 20,000

6. **BM Conf. Level:** 90%
(90 is Default)

Overlapping BM

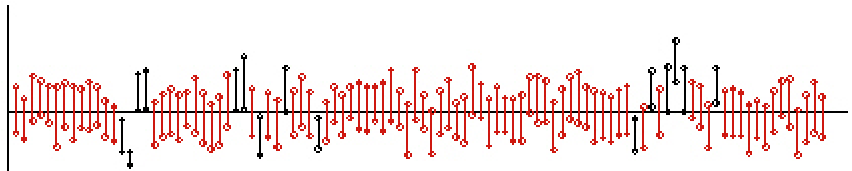
7. **Batch Size (m-o):** 100
(100 is Default) NOTE!!
N fixed from BM parameters

8. **OBM Conf Level:** 90%
(90 is Default)

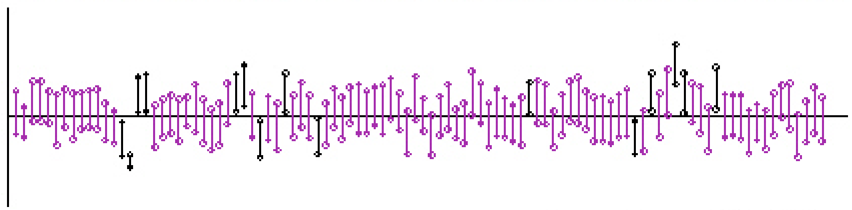
9. **Click Start:**

START BACK

Batch Means (1st 100 CIs at user specified Confidence Level - Not to Scale)



Overlapping Batch Means (1st 100 CIs at user specified Confidence Level - Not to Scale)



Method	Success Rate	MSE	Avg CI Length	Var CI Length
Batch Means Method	85.0%	246.15	.140	.470E-3
Overlapping Batch Means	85.0%	247.09	.133	.305E-3
Central Limit Theorem (Uses BM CI value)	31.0%	353.79	.032	.175E-5

5. Some Bigger Examples

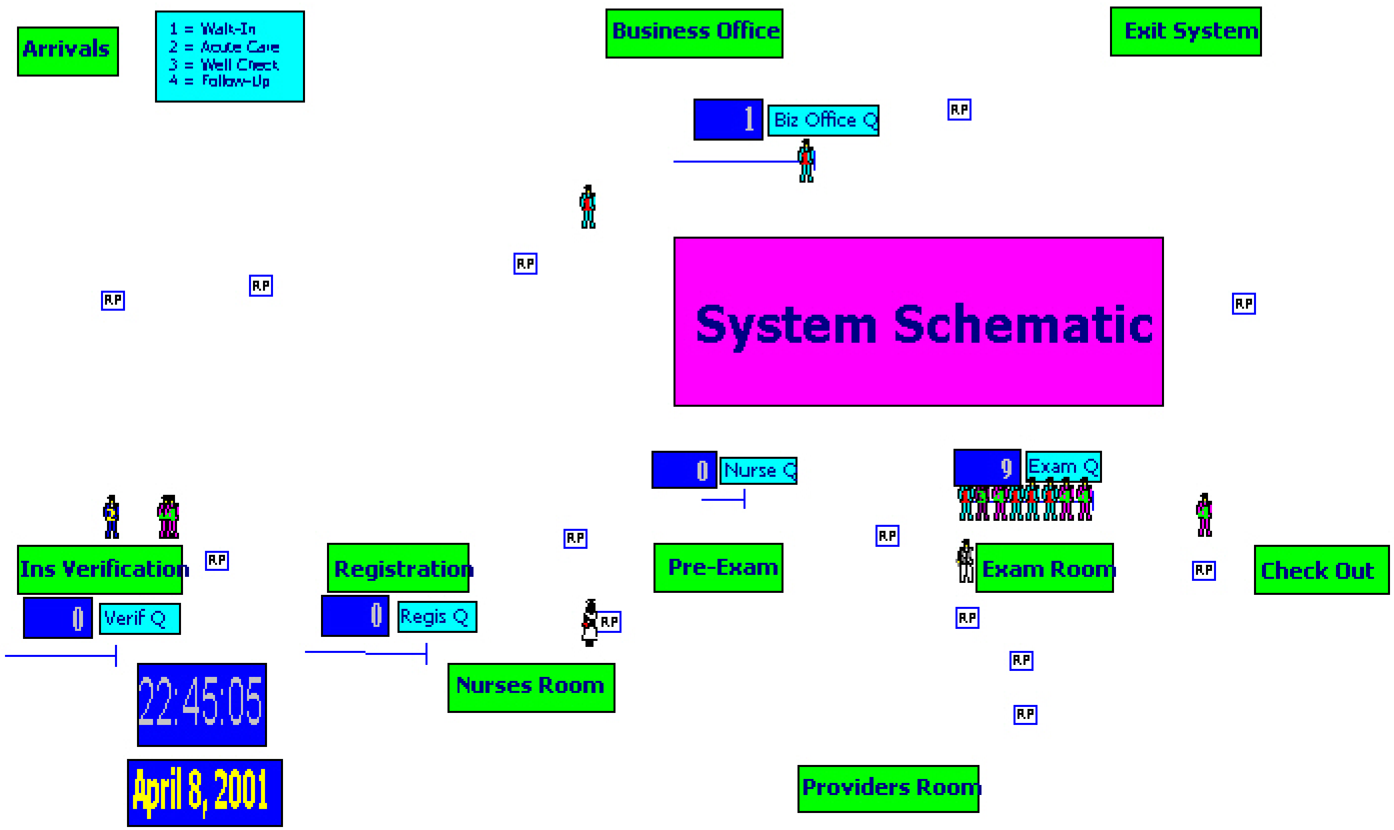
- Immunization Clinic
- Maternity Clinic
- Guinea Worm Disease
- Airline Routing Structure
- Pandemic Disease Propagation

Immunization Clinic

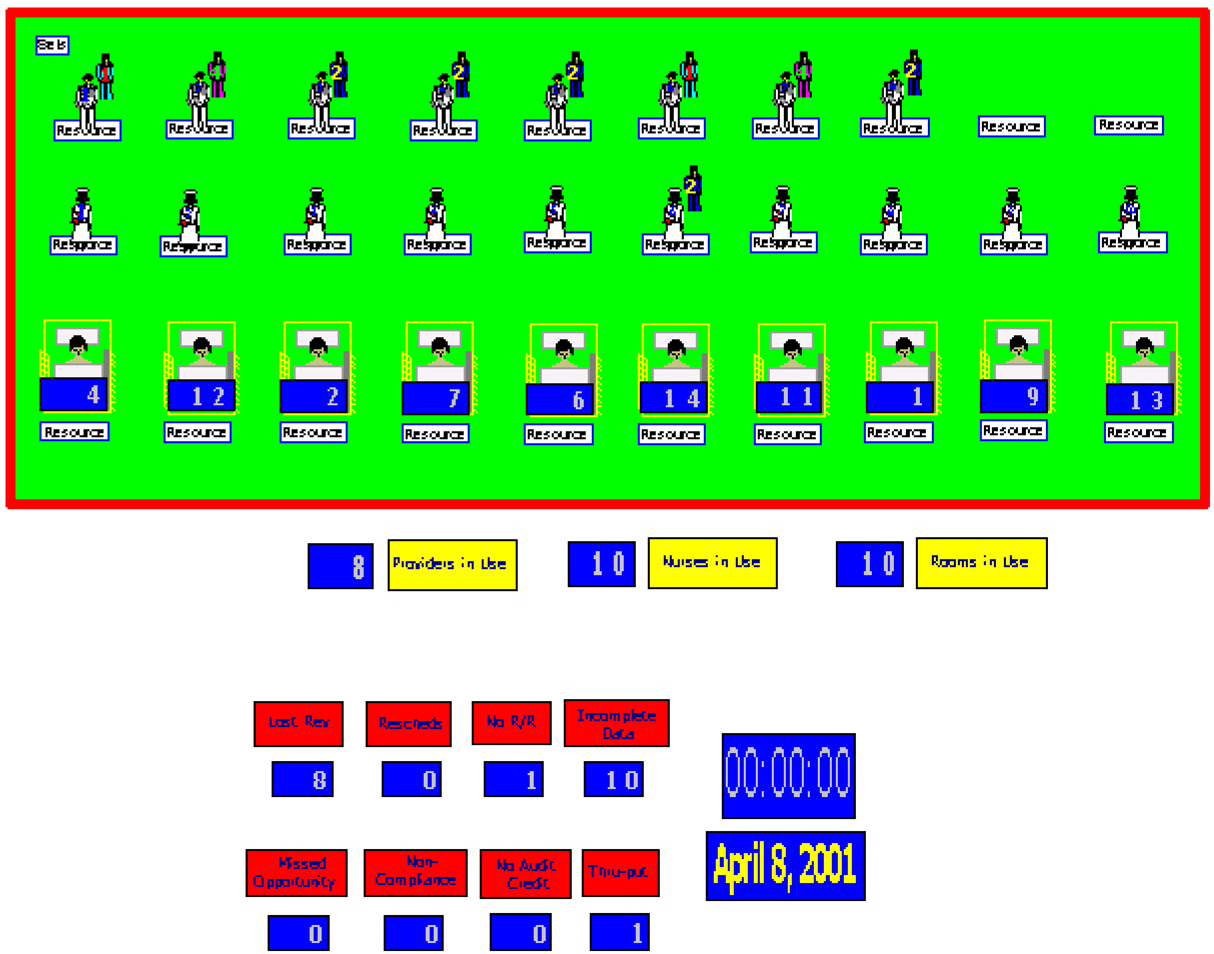
Partnership of Immunization Providers

Use simulation to

- Study operations of an immunization clinic
- Model generic clinic
- Read interarrival and service distributions from spreadsheet
- Study alternative clinic configurations



Arrivals
1 = Walk-In
2 = Acute Care
3 = Well Check
4 = Follow-Up
Business Office
Exit System
1
Biz Office Q
System Schematic
0
Nurse Q
9
Exam Q
Ins Verification
0
Verif Q
Registration
0
Regis Q
Pre-Exam
Exam Room
Check Out
Nurses Room
Providers Room
22:45:05
April 8, 2001


8
Providers in Use
10
Nurses in Use
10
Rooms in Use
Lost Rev
Rescheds
No R/R
Incomplete Data
8
0
1
10
Missed Opportunity
Non-Compliance
No Audit Credit
Thru-put
0
0
0
1
00:00:00
April 8, 2001

6. Selecting a Simulation Language

- More than 100 languages out there
  - Commercial
  - Freeware
- Maybe 5 to 10 major players
- Cost considerations
- Ease of learning
- Ease of use
- Classes, conferences