

A Glimpse of Representing Stochastic Processes

Nathaniel Osgood

CMPT 858

November 22, 2012

Dynamic Uncertainty: Stochastic Processes

- Examples of things commonly stochastically approximated
 - Stock market
 - Rainfall
 - Oil prices
 - Economic growth
- What considered “stochastic” will depend on the scope of the model
 - Detailed model: Individual behaviour, transmission, differential severity of infection, etc.
 - A meteorological model may not consider rainfall stochastic

Stochastic Processes in AnyLogic

- In AnyLogic, ABM and Discrete Event Models (“Network-Based Modeling”) are typically stochastic
 - Transitions between states
 - Event firing
 - Messages
 - (Frequent) timing of message send
 - Target of messages
 - Duration of a procedure
- As a result, there will be variation in the results from simulation to simulation

Summarizing Variability

- To gain confidence in model results, typically need to run a “Monte Carlo” ensemble of realizations
 - Deal with means, standard deviations, and empirical fractiles
 - As is seen here, there are typically still broad regularities between most runs (e.g. rise & fall)
- Need to reason over a population of realizations
 - ⇒ statistics are very valuable
 - Fractile within which historic value falls
 - Mean difference in results between interventions

Monte Carlo Methods in AnyLogic

- Monte Carlo methods draw repeated samples from distributions & stochastic processes of interest
- When running Monte Carlo method, we'd like to summarize the results of multiple runs
- One option would be to display each trajectory over time; downside: quickly gets messy
- AnyLogic's solution
 - Accumulate data regarding how many trajectories fall within given areas of value for a given interval of time using a "Histogram2D Data"
 - Display the Histogram2D Chart

MonteCarlo2D Histogram

- Divides up time into user-specified # of intervals
 - This forms a set of divisions along the horizontal (time) axis
- Divides up value axis for quantity being displayed into a user-specified # of interval
 - This forms a set of divisions along the vertical (value) axis
- Together, the divisions define a uniform (2D) grid
 - For each cell on that grid, a “Histogram2D Data” object accumulates data regarding how many trajectories include a value within that cell
 - i.e. how many trajectories have hold a range of values during a given interval of time)



Hands on Model Use Ahead



Load Sample Model:

SIR Agent Based Calibration

(Via “Sample Models” under “Help” Menu)

Monte Carlo Analysis with Fixed Parameter Values

Agent Based SIR Model - Monte Carlo Simulation

Run 100 replicat...

infectious2D

4,000
3,500
3,000
2,500
2,000

MonteCarlo2DHistogram - Parameter Variation Experiment

General Name: MonteCarlo2DHistogram Main active object class (root): Main Ignore Create Default U

Advanced

Random number generation:
 Random seed (unique simulation runs)
 Fixed seed (reproducible simulation runs) Seed Value: 1

Parameters: Varied in range Freeform Number of runs 100

Parameter	Expression
AverageI...uration	15
ContactRate	1.0
Infection...bability	0.8
AreaSide	100
TotalPopulation	10000

Model
Action
Analysis
Data Set
Statistics
Histogram Data
Histogram2D Data
Bar Chart
Stack Chart
Pie Chart
Plot
Time Plot
Time Stack Chart
Time Color Chart
Histogram
Histogram2D
Presentation
Connectivity
Enterprise Library
More Libraries...

Results of Monte Carlo Simulation



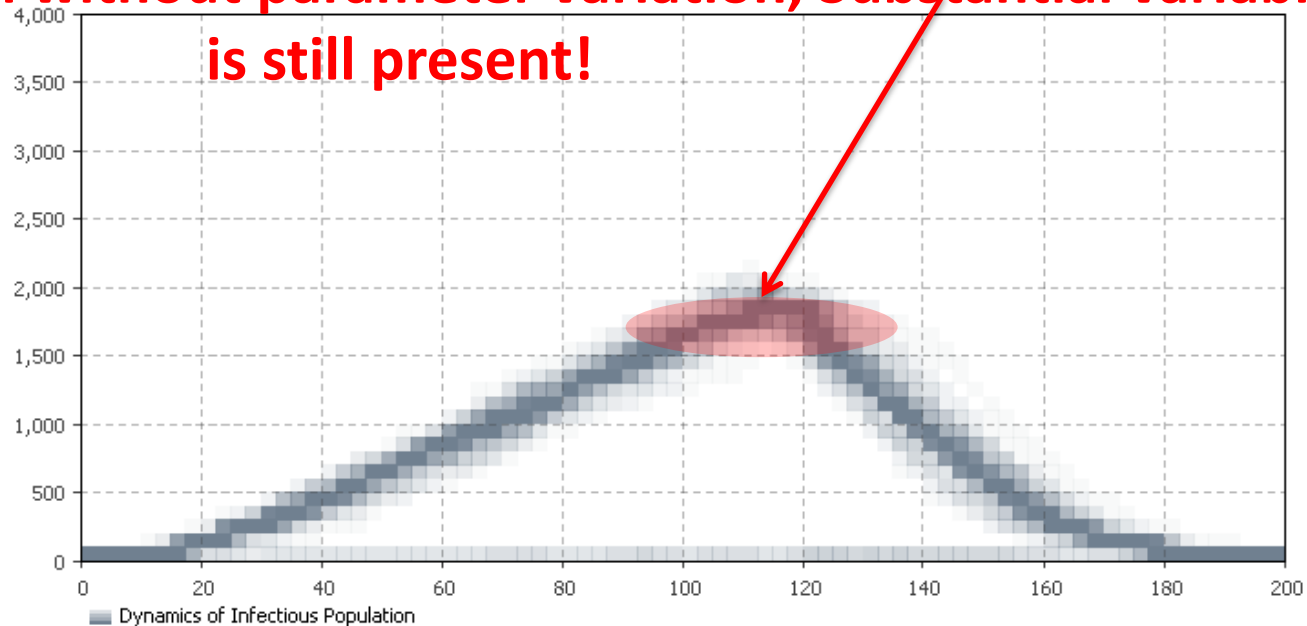
Agent Based SIR Model - Monte Carlo Simulation

Run 100 replications



AnyLogic and this model is (c) XJ Technologies, www.anylogic.com. All rights reserved.

Even without parameter variation, Substantial variability is still present!



This experiment performs multiple (100) runs of the Agent Based SIR Model with SAME (default) parameter values. As the model is essentially stochastic, each run results in a different output. In the chart above we display the summary of simulation runs (namely, the dynamics of the Infectious population size) in the form of the 2D histogram. The color intensity of a chart spot corresponds to the size of the corresponding 2D histogram bin.

Important Distinction (Declining Order of Aggregation)

- Experiment
 - Collection of simulations
- Simulation
 - Collection of replications that can yield findings across set of replications (e.g. mean value)
- Replication
 - One run of the model

Flexibility Typically Ignored

- In most AnyLogic models, an Experiment is composed of a single Simulation, which is composed of a single Replication
- In most AnyLogic models which run “ensembles” of realizations, a simulation is composed of only a single realization

Accumulating the Histogram2D dataset from other datasets

The screenshot displays the AnyLogic Advanced software interface. The main workspace shows a grid with a blue box labeled "The accumulating Histogram2D dataset is in Experiment" and a red box labeled "The source dataset is in Main". A red arrow points from the red box to the code in the "After Simulation Run" section of the "Advanced" properties panel. A blue arrow points from the blue box to the same code section. The code in the "After Simulation Run" section is `dataInfectious2D.add(root.dsInfectious);`. The "Before Simulation Run" section contains `dataInfectious2D.reset();`. The "After Iteration Code" section is empty. The "Properties" panel on the left shows the "Advanced" section selected. The "Console" panel at the bottom shows the "MonteCarlo2DHistogram - Parameter Variation Experiment" running. The "Analysis" palette on the right includes "Histogram2D Data" and "Histogram2D".

AnyLogic Advanced [EDUCATIONAL USE ONLY]

File Edit View Model Window Help

100%

Project Search

MonteCarlo2DHistogram Calibration Main MonteCarlo1stOrder Main Family ParametersVariation

Model

Action

Analysis

Data Set

Statistics

Histogram Data

Histogram2D Data

Bar Chart

Stack Chart

Pie Chart

Plot

Time Plot

Time Stack Chart

Time Color Chart

Histogram

Histogram2D

Agent Based SIR M

Run 100 replicat...

4,000

3,500

dataInfectious2D

Properties Console

MonteCarlo2DHistogram - Parameter Variation Experiment

General

Additional Class Code:

Advanced

Model Time

Initial Experiment Setup:

Presentation

Window

Replications

Description

Before Each Experiment Run:

```
dataInfectious2D.reset();
```

Before Simulation Run:

After Simulation Run:

```
dataInfectious2D.add( root.dsInfectious );
```

After Iteration Code:

Tolerance:

Presentation

Connectivity

Enterprise Library

More Libraries...

Monte Carlo Sensitivity Analyses in AnyLogic

AnyLogic Advanced [EDUCATIONAL USE ONLY]

File Edit View Model Window Help

100%

Get Support

Project Search

MonteCarlo2DHistogram Calibration Main MonteCarlo1stOrder Main Family ParametersVariation

Model

- Parameter
- Flow Aux Variable
- Stock Variable
- Event
- Dynamic Event
- Plain Variable
- Collection Variable
- Function
- Table Function
- Port
- Connector
- Entry Point
- State
- Transition
- Initial State Pointer
- Branch
- History State
- Final State
- Environment

Action

Analysis

Presentation

Connectivity

Enterprise Library

More Libraries...

Agent Based SIR Model - Monte Carlo Simulation

Run 100 replicat...

technologies
AnyLogic and this model is (c) XJ Technologies, www.anylogic.com. All rights reserved.

Choice between showing envelopes of empirical fractiles & showing counts in histogram bins

chart - Histogram2D

General Name: chart Show Name Ignore Public

Advanced

Dynamic Title: Dynamics of Infectious Populatic

Appearance Histogram: dataInfectious2D Color: slateGray

Description

Add Histogram Data

Show envelopes Show bins

Do not update automatically

Auto update after every iteration

Difference Between Chart Options

“Show envelopes”

- This option shows **envelopes of empirical fractiles**
 - These are associated with empirical fractiles defined in terms of percentages (e.g. “0.25” means boundary between lowest and 2nd lowest quartile; “0.50” means median)
 - e.g. These define envelopes of (contours) around the median within which data from different % of realizations fall
 - A “slice” through the output at a particular moment in time would be like an **extended boxplot** (showing fractiles)
- The empirical fractiles to use are themselves defined in the associated Histogram2D Data object

Reminder: 2D Histogram Data

The screenshot displays the AnyLogic Advanced software interface. The main workspace shows a 2D histogram titled "Agent Based SIR M" with a "Run 100 replicat..." button. The histogram has a vertical axis ranging from 3,000 to 4,000 and a horizontal axis with a label "dataInfectious2D".

Overlaid on the histogram is the following text in red:

Note definition of envelopes to be used in The Histogram2D Chart if "Show envelopes" is selected.

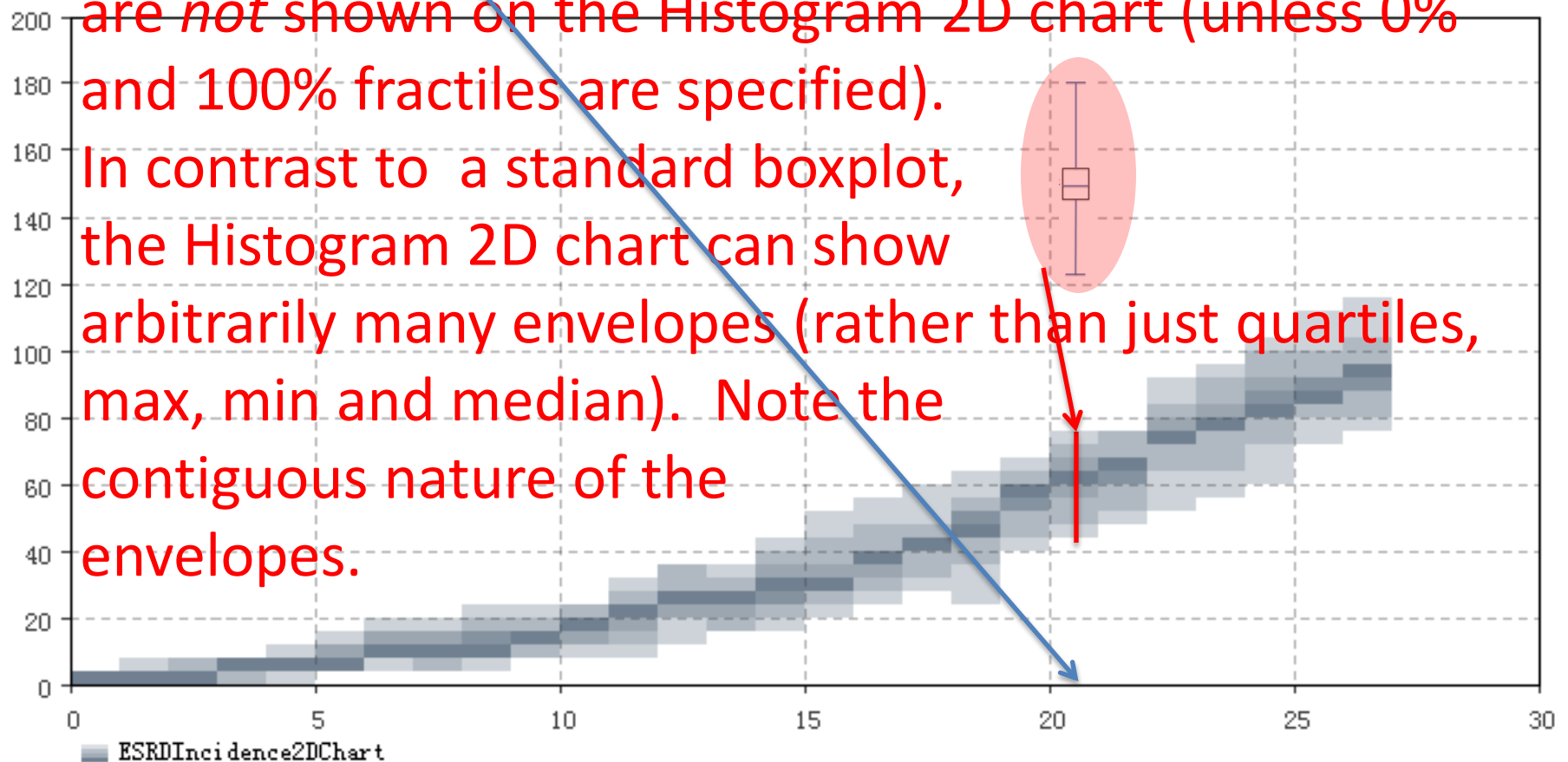
Below the histogram is the "dataInfectious2D - Histogram 2D Data" configuration panel. The "General" tab is active, showing the following settings:

- Name: dataInfectious2D
- Show Name
- Ignore
- Public
- Show At Runtime
- Horizontal axis value: []
- Vertical axis value: []
- Horizontal intervals: 100
- Range, from: 0 to: 200
- Vertical intervals: 100
- Range, from: 0 to: 4000
- Envelopes: 0.25, 0.5, 0.75
- Do not update automatically
- Auto update after every iteration

A red arrow points from the text "envelopes" to the "Envelopes" input field in the configuration panel, which is circled in red.

Example of “Show Envelopes” Output (Different Model)

A slice at **this** point in time would yield a something like a **boxplot**. Note that the “whiskers” of the boxplot are *not* shown on the Histogram 2D chart (unless 0% and 100% fractiles are specified). In contrast to a standard boxplot, the Histogram 2D chart can show arbitrarily many envelopes (rather than just quartiles, max, min and median). Note the contiguous nature of the envelopes.



Show Bins Option

The screenshot shows the AnyLogic Advanced software interface. The main window displays a 2D histogram titled "Agent Based SIR Model - Monte Carlo Simulation". The histogram shows the distribution of data points, with a grid overlay. The y-axis ranges from 1,500 to 4,000. A red arrow points from the text "The 'show bins' option is here." to the "Show bins" radio button in the "chart - Histogram2D" properties panel.

chart - Histogram2D

General Name: chart Show Name Ignore Public

Advanced

Dynamic Title: Dynamics of Infectious Populatic

Appearance Histogram: dataInfectious2D Color: slateGray

Description

Add Histogram Data

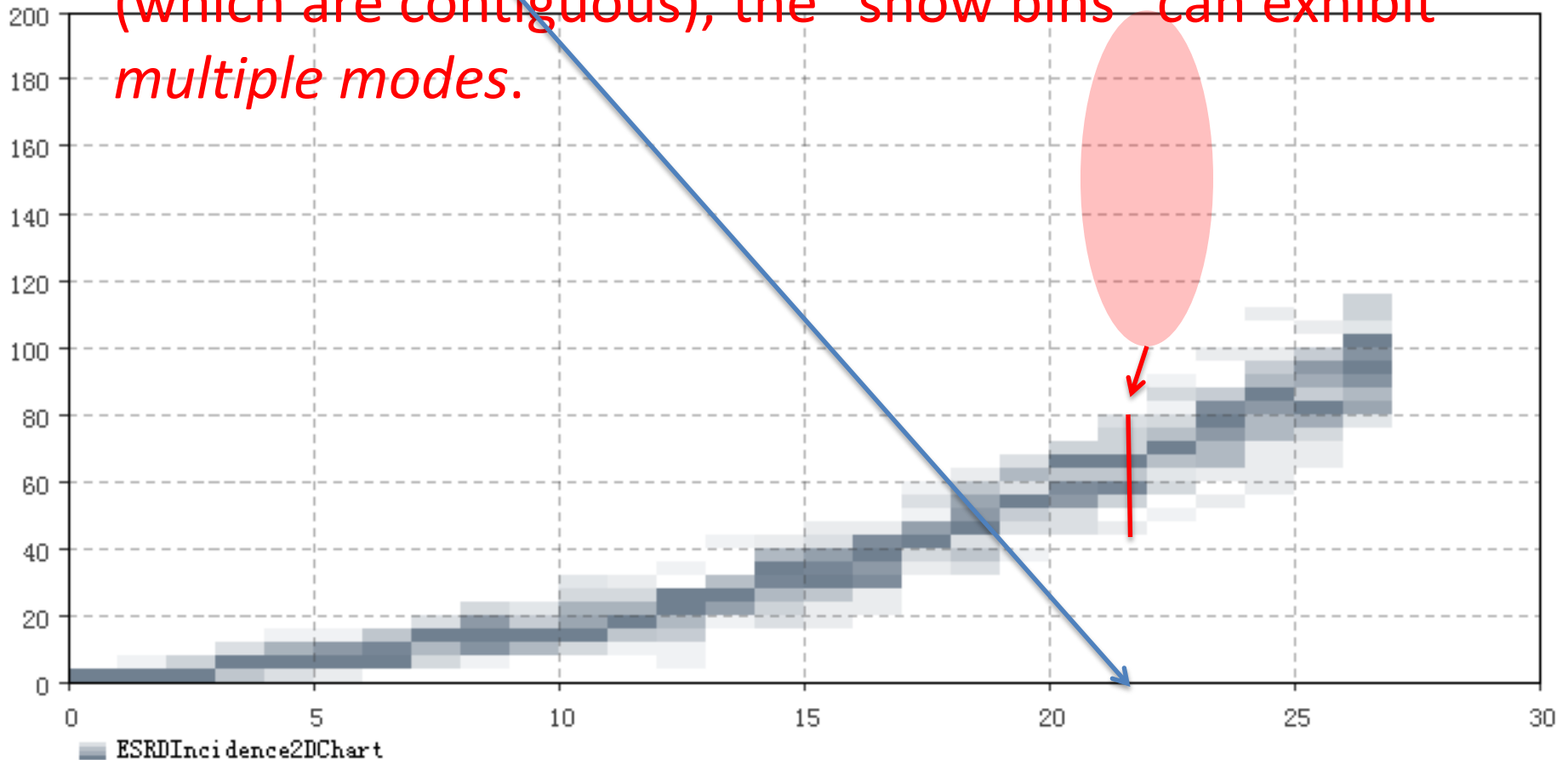
Show envelopes Show bins

Do not update automatically

Auto update after every iteration

Example of “Show Bins” Output (Different Model)

A slice at **this** point in time would yield a *histogram*.
Note: In contrast to the situation for the envelopes (which are contiguous), the “show bins” can exhibit *multiple modes*.



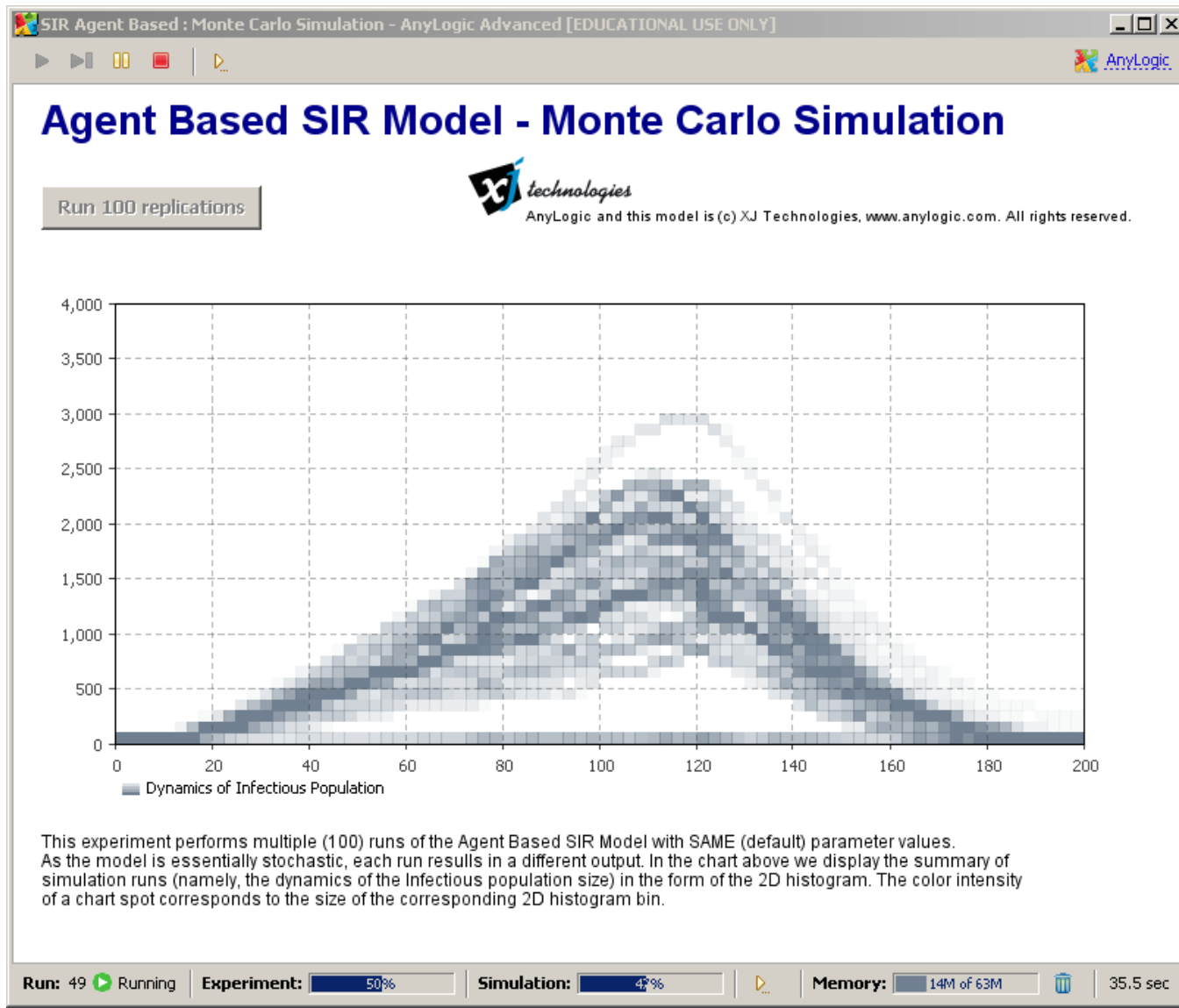
Automatic Throttling of Monte Carlo Analyses

The screenshot displays the AnyLogic Advanced software interface for an educational use. The main window is titled "Agent Based SIR Model - Monte Carlo Simulation" and shows a 2D histogram of "infectious2D" data. A button labeled "Run 100 replicat..." is visible. The bottom panel shows the configuration for the "MonteCarlo2DHistogram - Parameter Variation Experiment".

MonteCarlo2DHistogram - Parameter Variation Experiment

- Use replications
- Fixed number of replications
 - Replications per iteration: 10
- Varying number of replications (Stop replications after minimum replications, when confidence level is reached)
 - Minimum replications: 2
 - Maximum replications: 10
 - Confidence level: 80% of expression: 0
 - Error percent: 0.5

General Variety of Output



Reminder: Statistical Scaling

- Consider Taking the sample mean of n samples that vary independently around a mean
- If two samples x and y are independent samples of random variables X and Y , then $\text{Var}[x+y]=\text{Var}[X]+\text{Var}[Y]$

– So if we have n indep. samples x_i from distribution X

$$\text{Var}\left(\sum_{i=1}^n x_i\right) = n\text{Var}(X)$$

- If we scale a random variable by a factor α , the standard deviation scales by the same factor of $\alpha \Rightarrow$ the variance scales by α^2
 - i.e. $\text{StdDev}[\alpha X] = \alpha \text{StdDev}[X]$, $\text{Var}[\alpha X] = \alpha^2 \text{Var}[X]$

Statistics of Sample Mean

- Recall: Sample Mean:

$$m = \frac{\sum_{i=1}^n x_i}{n}$$

- From the preceding, variance drops as $1/n$

$$\text{Var}(m) = \text{Var}\left(\frac{\sum_{i=1}^n x_i}{n}\right) = \frac{\text{Var}\left(\sum_{i=1}^n x_i\right)}{n^2} = \frac{n\text{Var}(X)}{n^2} = \frac{\text{Var}(X)}{n}$$

- This means that standard deviation for the sample mean of n samples drops as $1/\text{sqrt}(n)$

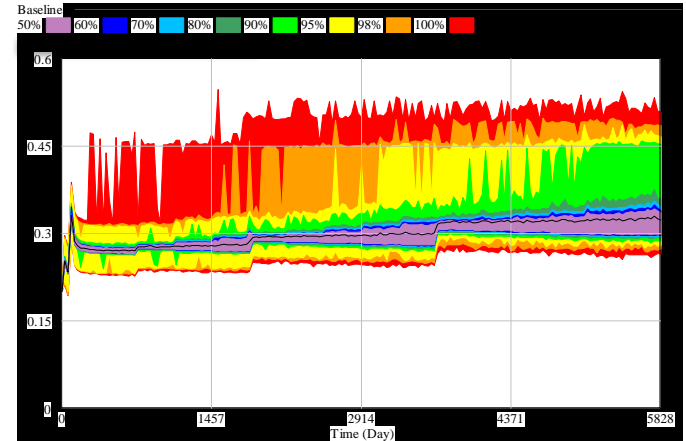
$$\text{StdDev}(m) = \sqrt{\text{Var}(m)} = \sqrt{\frac{\text{Var}(X)}{n}} = \sqrt{\frac{(\text{StdDev}(X))^2}{n}} = \frac{\text{StdDev}(X)}{\sqrt{n}}$$

- So if we wish to divide the standard deviation of the sample mean by a factor of 2, we need to take 4x the number of Monte Carlo samples

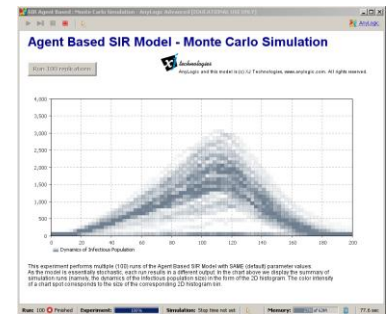
Closing Question: How can we best adapt our policies to deal with ongoing uncertainty?

- We are dealing here with making decisions in an environment that changes over time

- This uncertainty could come from
 - Stochastic variability



- Uncertainty regarding parameter values



- There is an incredibly vast # of possible policies
- **Reminder: Can successfully integrate decision analysis & simulation to neatly handle such cases**