Sensitivity Analysis

Nathaniel Osgood

Using Modeling to Prepare for Changing Healthcare Needs

Duke-NUS

April 16, 2014

Types of Sensitivity Analyses

- Variables involved
  - One-way
  - Multi-way
- Type of component being varied
  - Parameter sensitivity analysis: Parameter values
  - Structural sensitivity analysis: Examine effects of model *structure* on results
- Type of variation
  - Single alternative values
  - Monte Carlo analyses: Draws from probability distributions (many types of variations)
- Frequency of variation
  - Static (parameter retains value all through simulation)
  - Ongoing change: Stochastic process
    - Accomplished via Monte-Carlo analyses
    - Key for DES & ABM

Model Uncertainty

- Here, we are frequently examining the impact of changing
  - Our assumptions about “how the system works”
  - Our decision of how to abstract the system behaviour
- Structural sensitivity analyses
  - Vary structure of model & see impact on
    - Results
    - Tradeoffs between choices
  - Frequently recalibrate the model in this process
- Here, we are considering uncertainty about how the current state is mapped to the next state

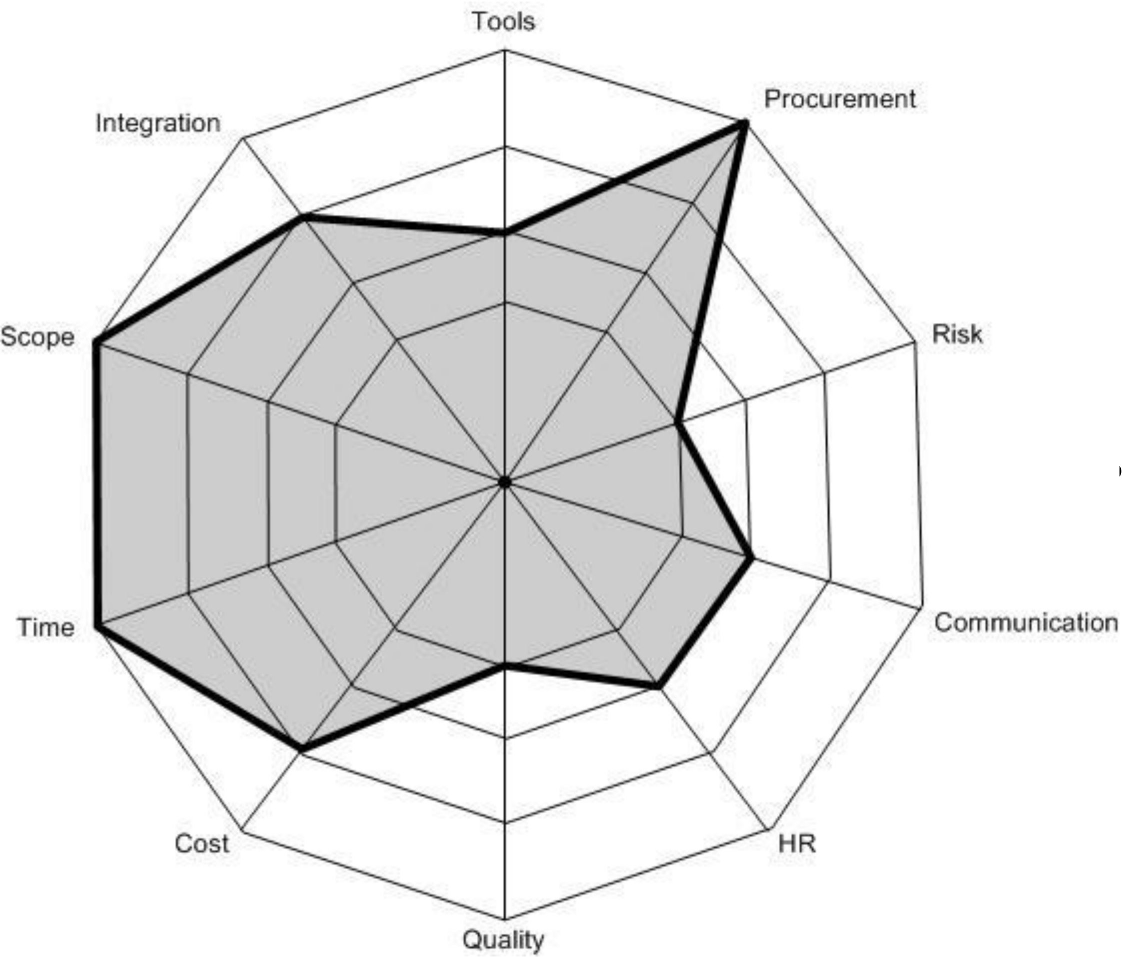
Predictor-Corrector Methods: Dealing with an Incomplete Model

- Some approaches (e.g. Kalman filter, Particle Filter) are motivated by awareness that models are incomplete
- Such approaches try to adjust model state estimates on an ongoing basis,
  - Given uncertainty about model predictions
  - New observations
- Assumption here is that the error in the model is defined by some probability distribution

Static Uncertainty
Sensitivity Analyses

- In variation, one can seek to investigate different
  - Assumptions
  - Policies
- Same relative or absolute uncertainty in different parameters may have hugely different effect on outcomes or decisions
- Help identify parameters/initial states that strongly affect
  - Key model results
  - Choice between policies
- We place more emphasis in parameter estimation & interventions into parameters exhibiting high sensitivity

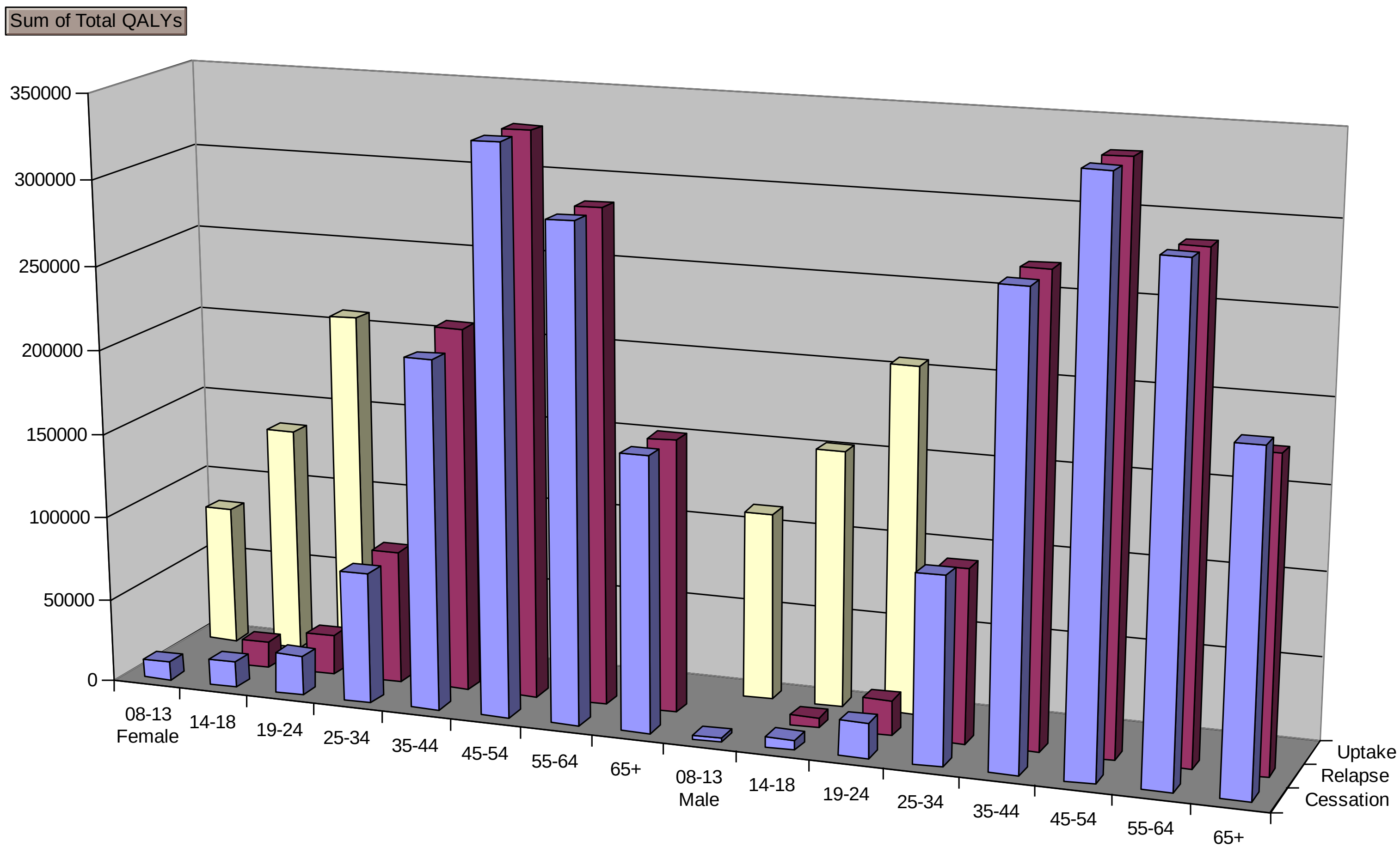
Spider Diagram



- Each axis represents a % change in a particular parameter
  - This proportional change is identical for the different parameters
- The distance assumed by the curve along that axis represents the magnitude of response to that change
  - Note that these sensitivities will depend on the state of system!

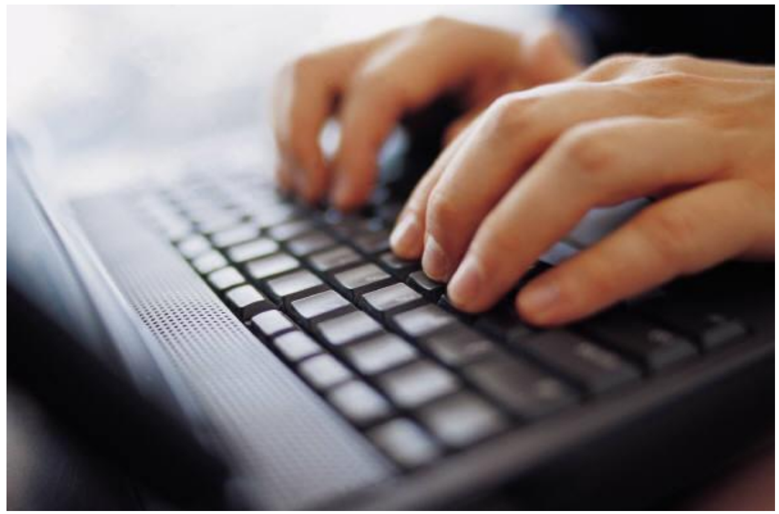
http://www.niwotridge.com/images/BLOGImages/SpiderDiagram.jpg

Systematic Examination of Policies



Tengs, Osgood, Lin

Hands on Model Use Ahead



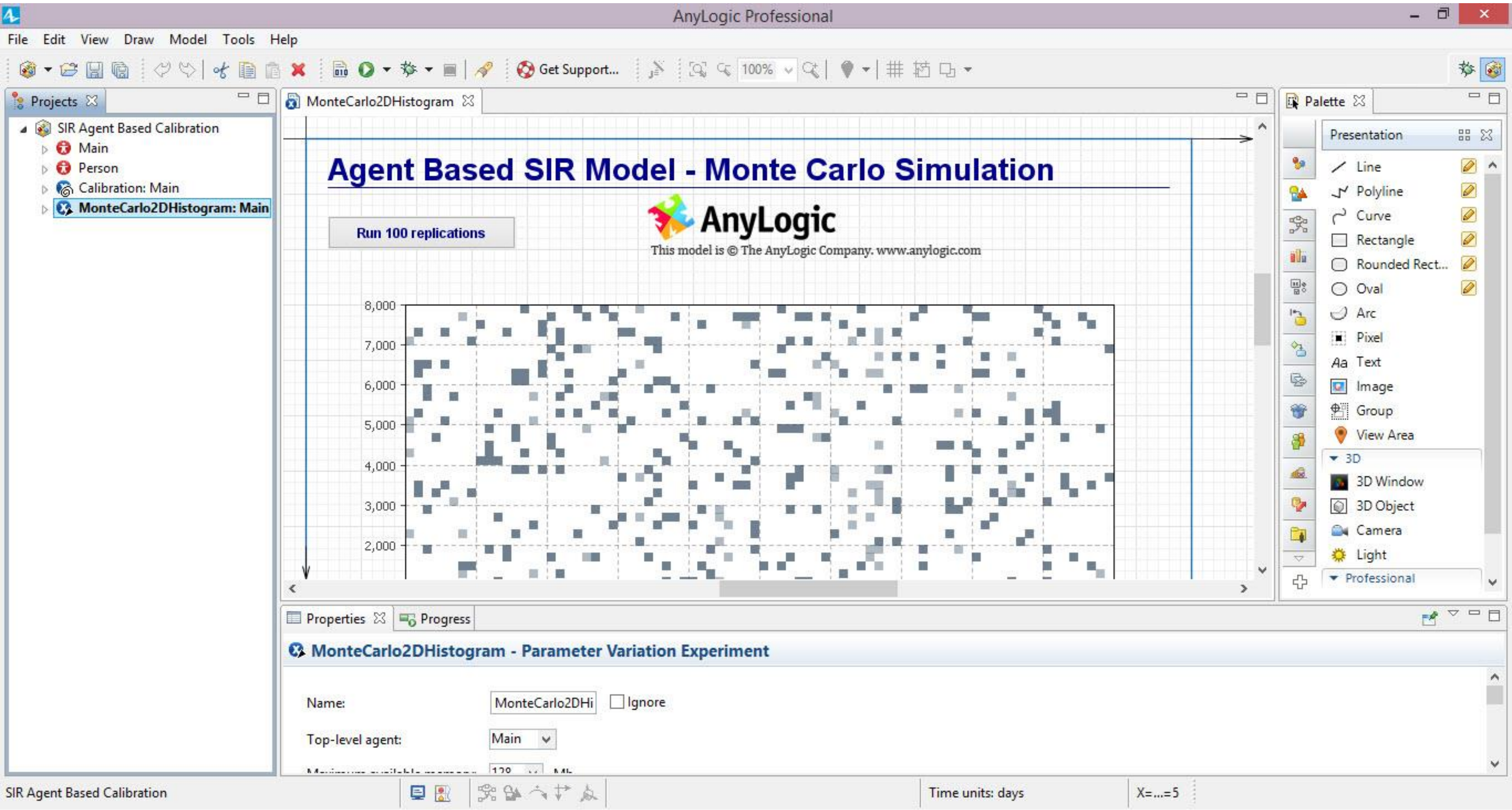
Load Sample Model:

SIR Agent Based Calibration

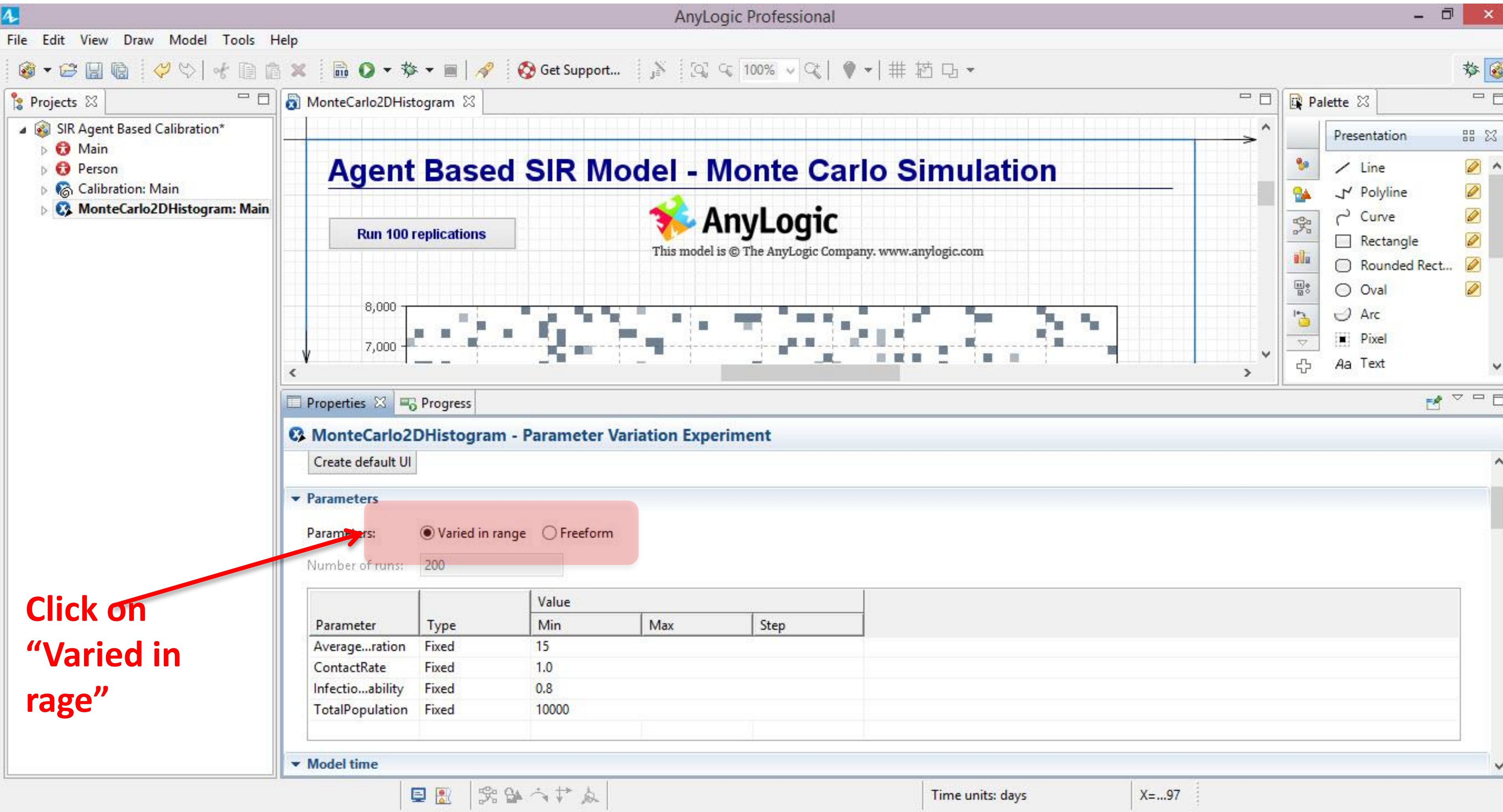
(Via “Example Models” under “Help” Menu)

Click on “MonteCarlo2DHistogram” Experiment

Setting Ranges for Parameter Variation Can Handle 1-Way or (Orthogonal) Multi-Way

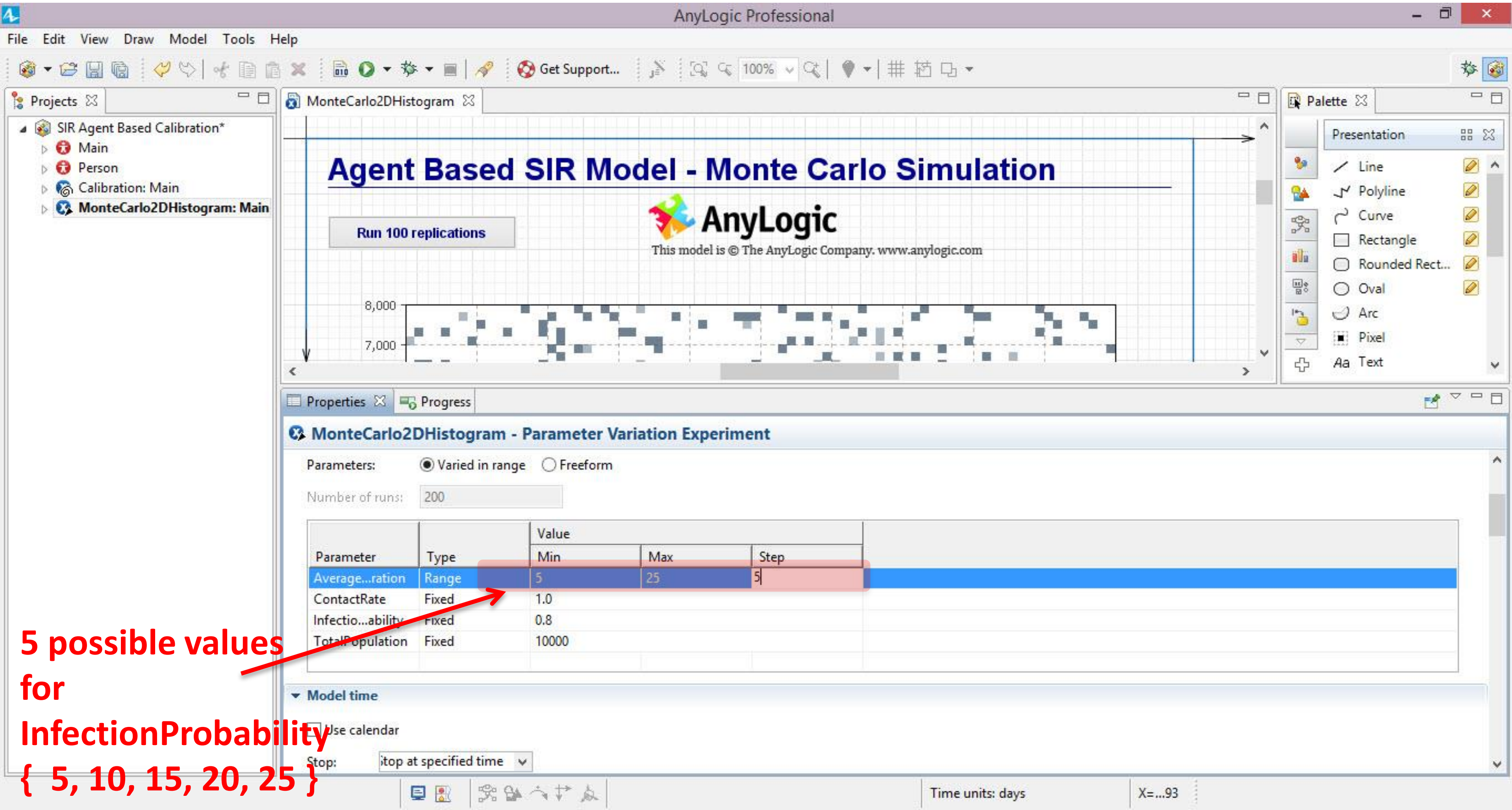


Indicating to Systematically Vary "AverageIllnessDuration"

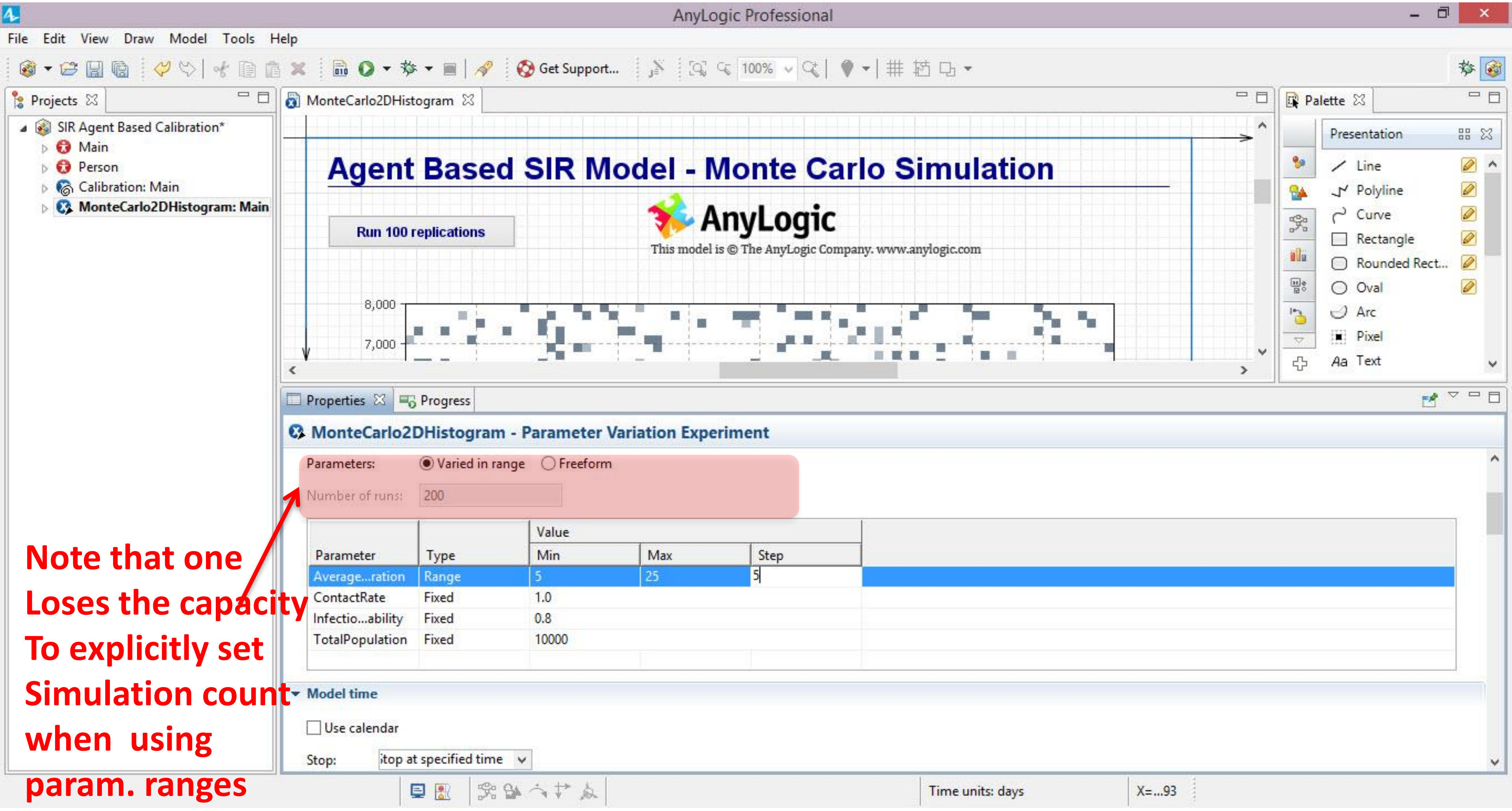


Click on "Varied in rage"

Entering Extremes & Step of the Variation

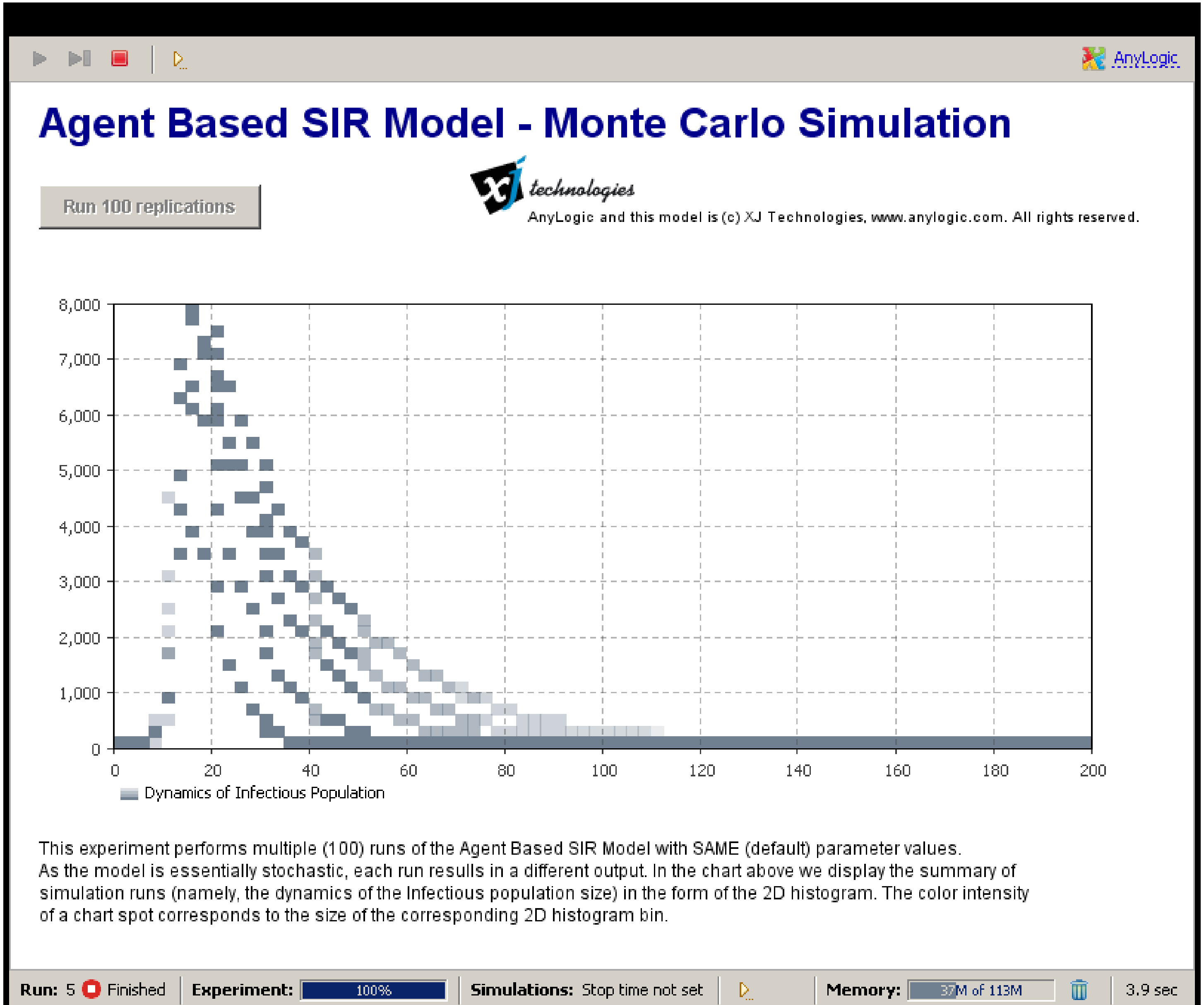


Note: Simulation Count is Defined Implicitly



Note that one
Loses the capacity
To explicitly set
Simulation count
when using
param. ranges

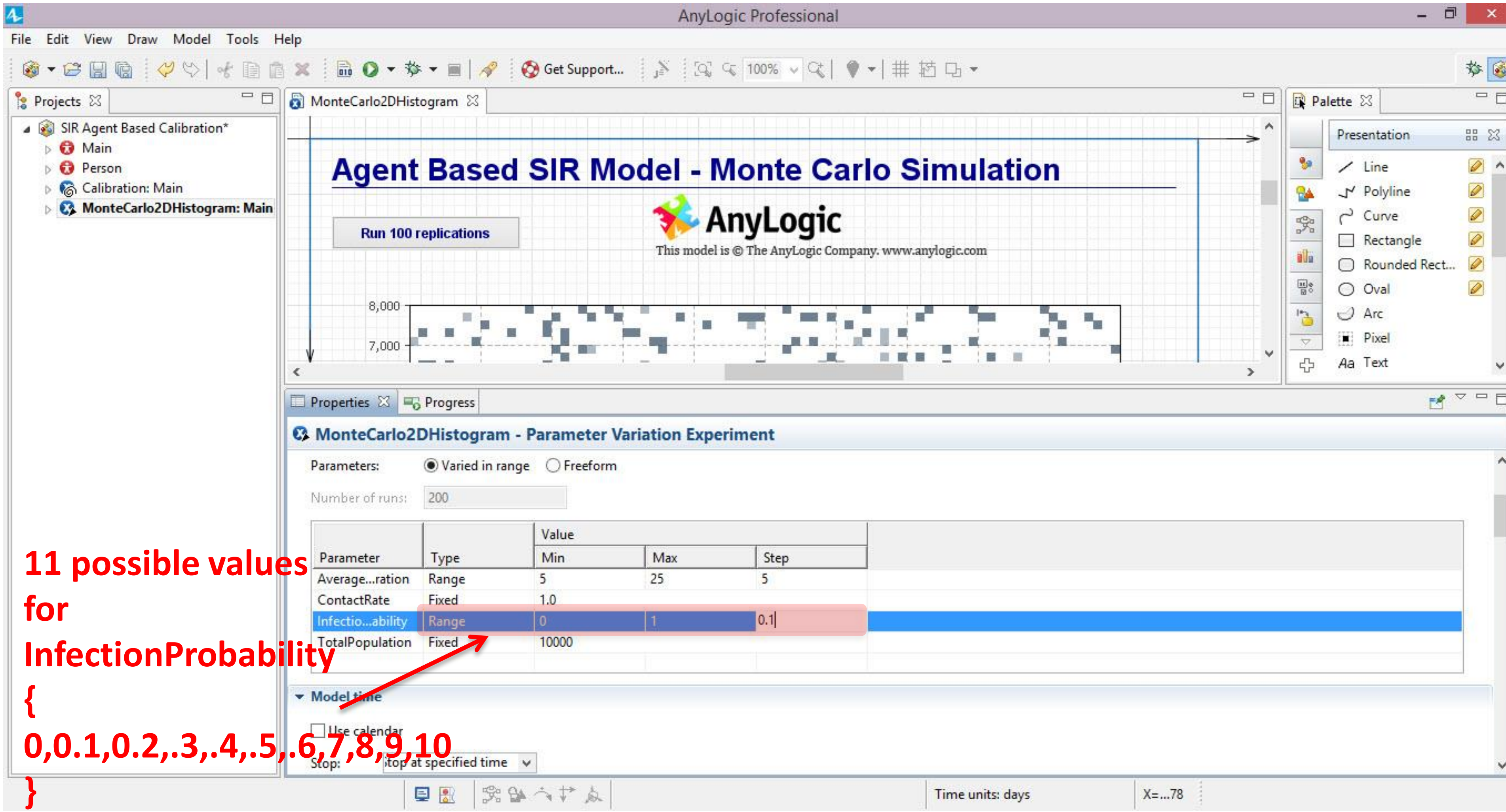
“1 Way” Sensitivity Exploration in AnyLogic



Multi-Way Sensitivity Analyses

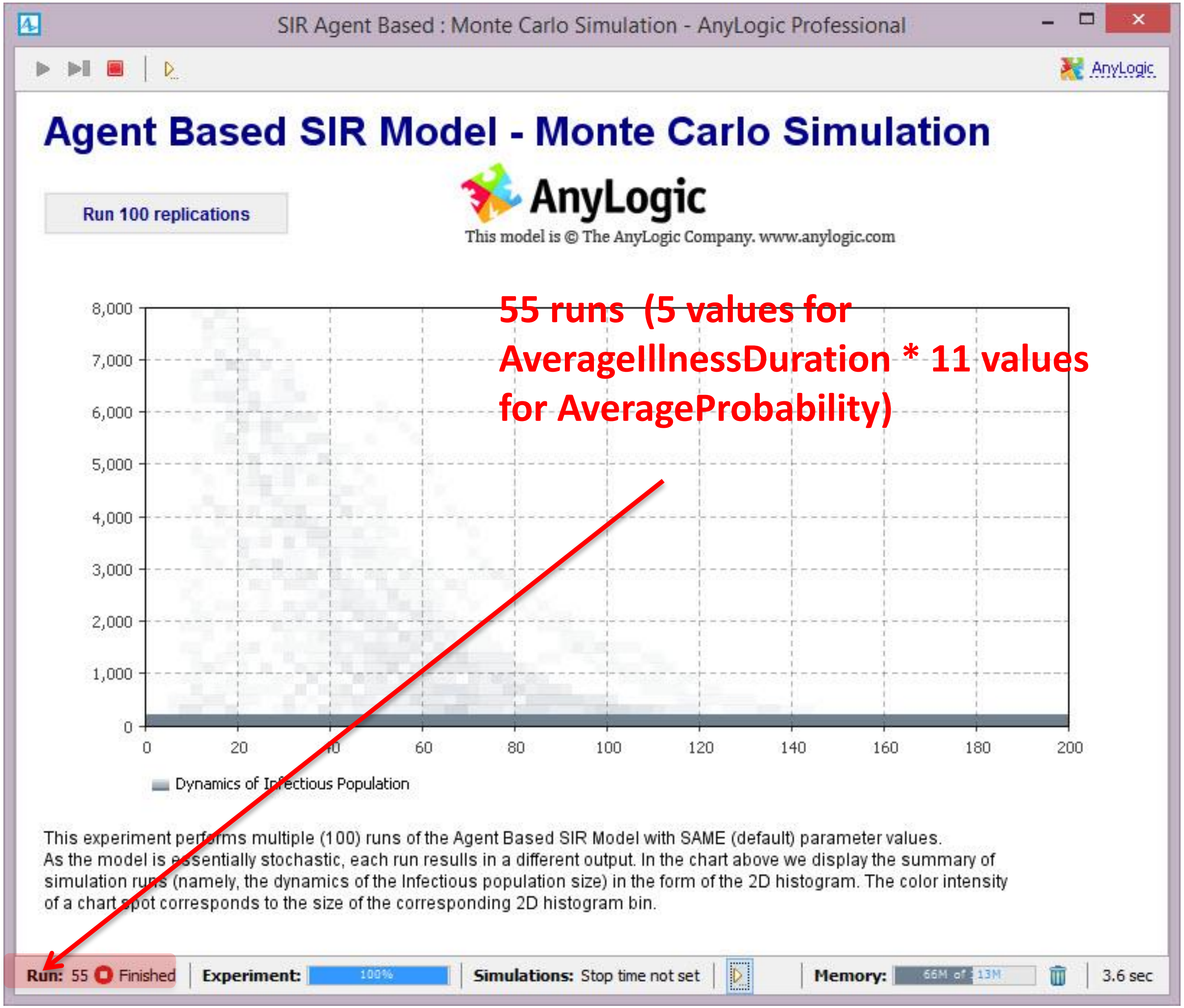
- When examining the results of changing multiple variables, need to consider how multiple variables vary together
- If this covariation reflects dependence on some underlying factor, may be able to simulate uncertainty in underlying factor

Set Range for InfectionProbability



11 possible values for InfectionProbability
{
0,0.1,0.2,.3,.4,.5,.6,7,8,9,10
}

Resulting Output



Challenge: Combinatorial Explosion

- If we have *n* distinct parameters, each varying over *c* values, we have c^n possible combinations of parameter values
- As *c* and (far more important) *n* rise, it quickly becomes infeasible to exhaustively examine such combinations
- Alternatives
  - Dimensional analysis
  - Techniques for judiciously selecting combinations count
  - Monte Carlo techniques

Challenge: Combinatorial Explosion

- If we have *n* distinct parameters, each varying over *c* values, we have c^n possible combinations of parameter values
- As *c* and (far more important) *n* rise, it quickly becomes infeasible to exhaustively examine such combinations
- Alternatives
  - Dimensional analysis
  - Techniques for judiciously selecting combinations count
  - Monte Carlo techniques

Dimensional Analysis

- Observation: Often it is not a parameter in isolation that matters, but combinations of parameters
  - Ratio of two parameters (speed of infection spread vs. speed of intervention)
  - Multiplication of two parameters (βc)
- Dimensional analysis techniques exploit the reflection that our choice of units has no impact on how the world operates to recognize that a process must be governed by **dimensionless quantities**
- We can typically reduce the parameter count by
  - Determining dimensions associated w/model quantities
  - Determining a set of dimensionless quantities

Challenge: Combinatorial Explosion

- If we have *n* distinct parameters, each varying over *c* values, we have c^n possible combinations of parameter values
- As *c* and (far more important) *n* rise, it quickly becomes infeasible to exhaustively examine such combinations
- Alternatives
  - Dimensional analysis
  - Techniques for judiciously selecting combinations count
  - Monte Carlo techniques

Judiciously Selecting Set of Parameter Combinations Being Examined: Discrete

- Latin hypercubes
  - Test each value of each parameter at least once
  - i.e. for each possible value of a given parameter p, we are guaranteed to explore at least one parameter vector where parameter p holds that value
- Orthogonal arrays
  - Test each pair of parameter values at least once
  - i.e. for each possible pair of values of given parameters p_1 and p_2, we are guaranteed to explore at least one parameter vector where parameters p_1 and p_2 hold those values

Latin Squares

- Each value (or equivalence class) of each field occurs exactly once
- This is a "minimal set" that exercises all values (equivalence classes) of each field

http://en.wikipedia.org/wiki/File:Fisher-stainedglass-gonville-caius.jpg#filelinks

Orthogonal Arrays ("Pair-wise Testing", "All Pairs Testing")

- Key characteristic: Every *pair* of variable values occurs at least once (but not all combinations!)
  - NB:contrast with Latin Squares, where every value is tried
- Exploits fact that *compatibility* problems typically are pairwise
  - If problem is not revealed in just by single value or equivalence class, it is likely to by no more than a *pair*
- The count of all combinations of values would typically be vastly larger
  - All Combinations: $\prod_{i=1}^{n} |X_i|$ Rises as least as 2^n
  - Pairwise: $\max(|X_i|)\max_{j,j\neq i}(|X_j|)$ (?)
- => we can do much less work than required for all combinations, but still secure most of value!

Example Case Reduction

S.No	Parameters and Values	100% coverage	reduced to
1	$3^2 2^2 1^1$	36	9
2	$10^4\ 7^4\ 8^3\ 6^2\ 3^2\ 5^2\ 4^2\ 9^1$	14338695168000000	144
3	10^{15}	1000000000000000	199
4	$10^5 5^6 3^5 2^5$	243000000000000	142
5	$3^3 4^2 5^2 6^2 7^1 8^2 10^1 9^1$	15676416000	96

From http://www.testersdesk.com/pairwse_testersdesk.html

Judiciously Selecting Set of Parameter Combinations Being Examined: Continuous

- Grid
- Grid adaptation of
  - Latin Hypercube
  - Orthogonal array
- While not directly sampled by AnyLogic, these can be easily implemented by
  - Creating vectors of the appropriate values to use successively for a given parameter value
  - For a given iteration, finding the appropriate value for a particular parameter by indexing into the vector for that parameter

Challenge: Combinatorial Explosion

- If we have n distinct parameters, each varying over c values, we have c^n possible combinations of parameter values
- As c and (far more important) n rise, it quickly becomes infeasible to exhaustively examine such combinations
- Alternatives
  - Dimensional analysis
  - Techniques for judiciously selecting combinations count
  - Monte Carlo techniques

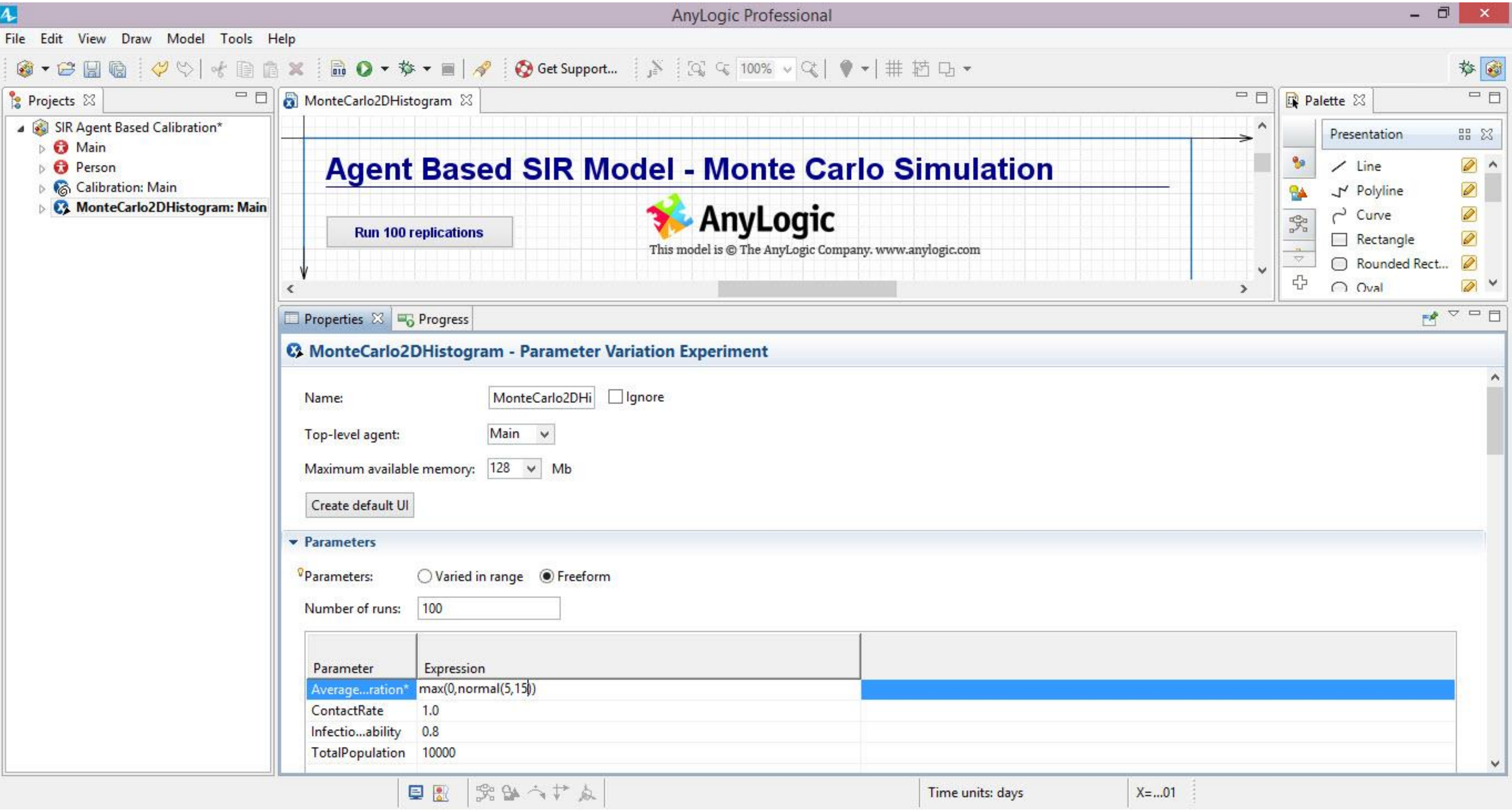
Imposing a Probability Distribution Monte Carlo Analysis

- We feed in probability distributions to reflect our uncertainty about one or more parameters
- The model is run many, many times (realizations)
  - For each realization, the model uses a different draw from those probability distribution
- What emerges is resulting probability distribution for model outputs

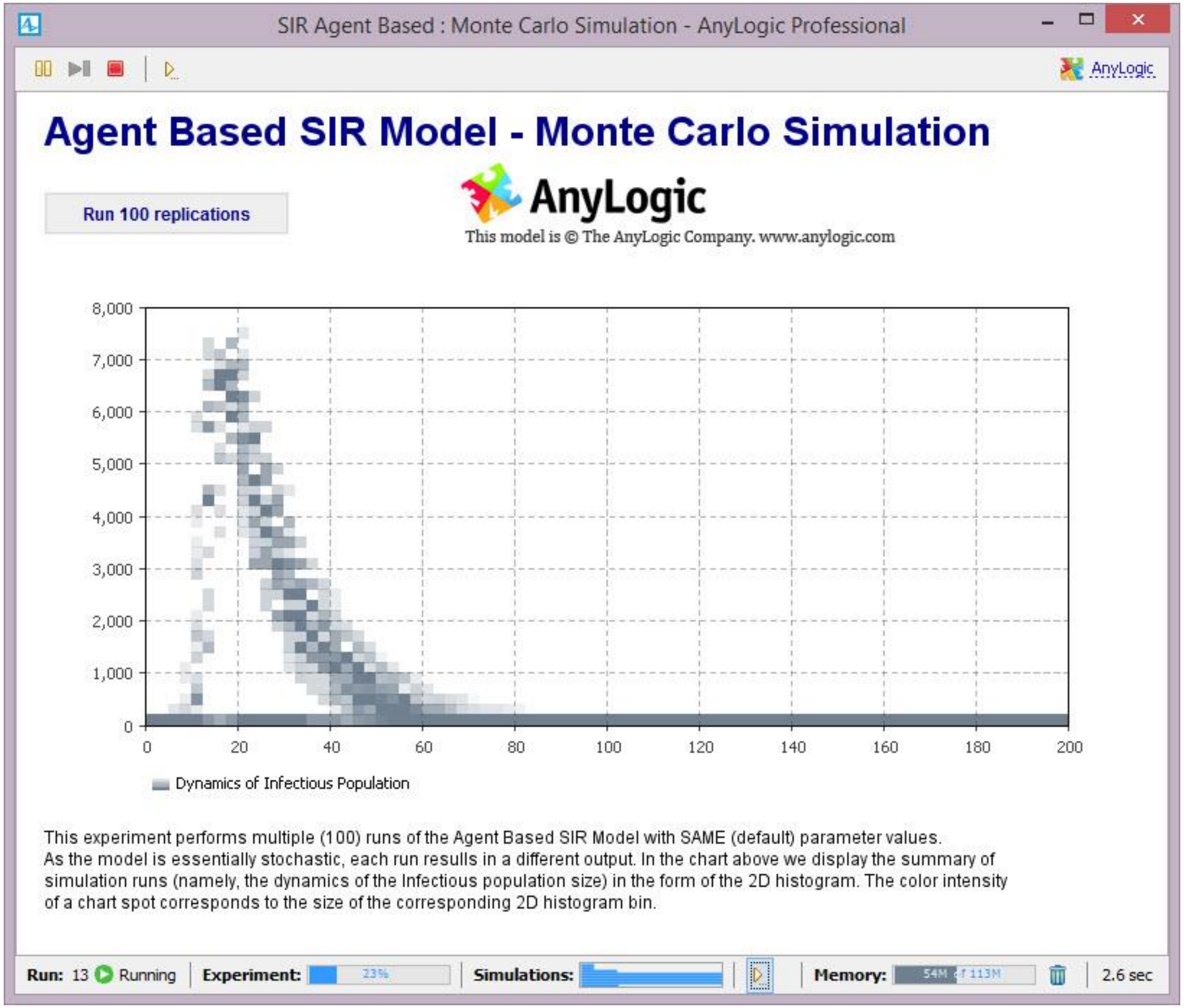
Monte Carlo Techniques (Continuous Parameters)

- For a given amount of computational effort, Monte Carlo techniques typically exhibit far more favourable properties as the count of dimensions rises than does downsampling the grid of values
- Effort Scaling
  - Grid: $\theta\left(\prod_{i=1}^{D} d_i\right)$
  - Monte Carlo: $\theta(N)$ (N: count of Monte Carlo samples)
- Error (eg when determining mean) given limited eff.
  - Coarser grid: $\theta\left(c^{-\frac{2}{D}}\right) = \theta\left(\frac{1}{c^{\frac{2}{D}}}\right)$ Spends much effort in low-probability areas of space
  - Monte Carlo: $\theta\left(\frac{1}{\sqrt{N}}\right)$ Naturally concentrates samples in highly probable regions (because sampled)

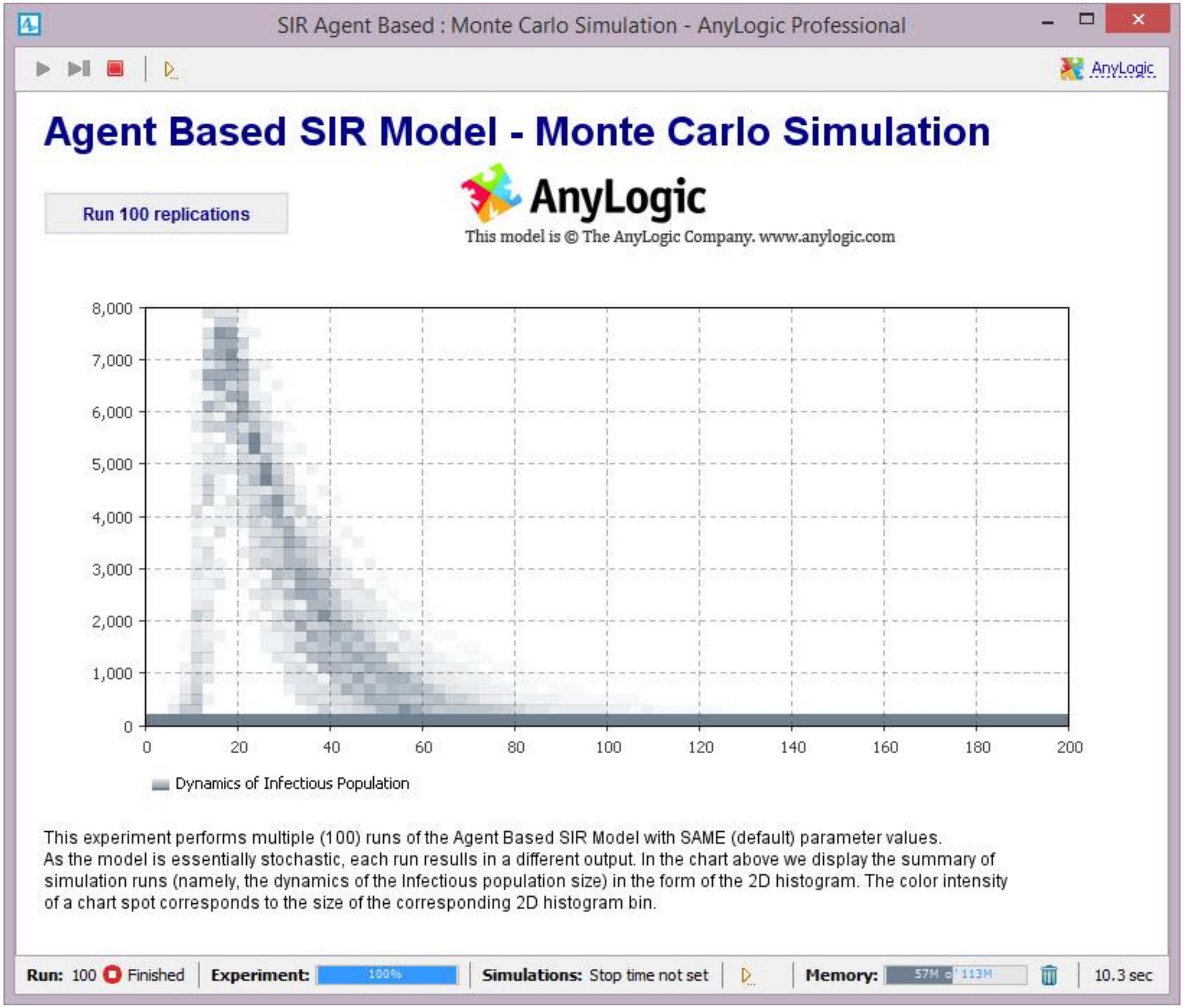
Monte Carlo Analyses in AnyLogic: Specifying Distributions for Parameters



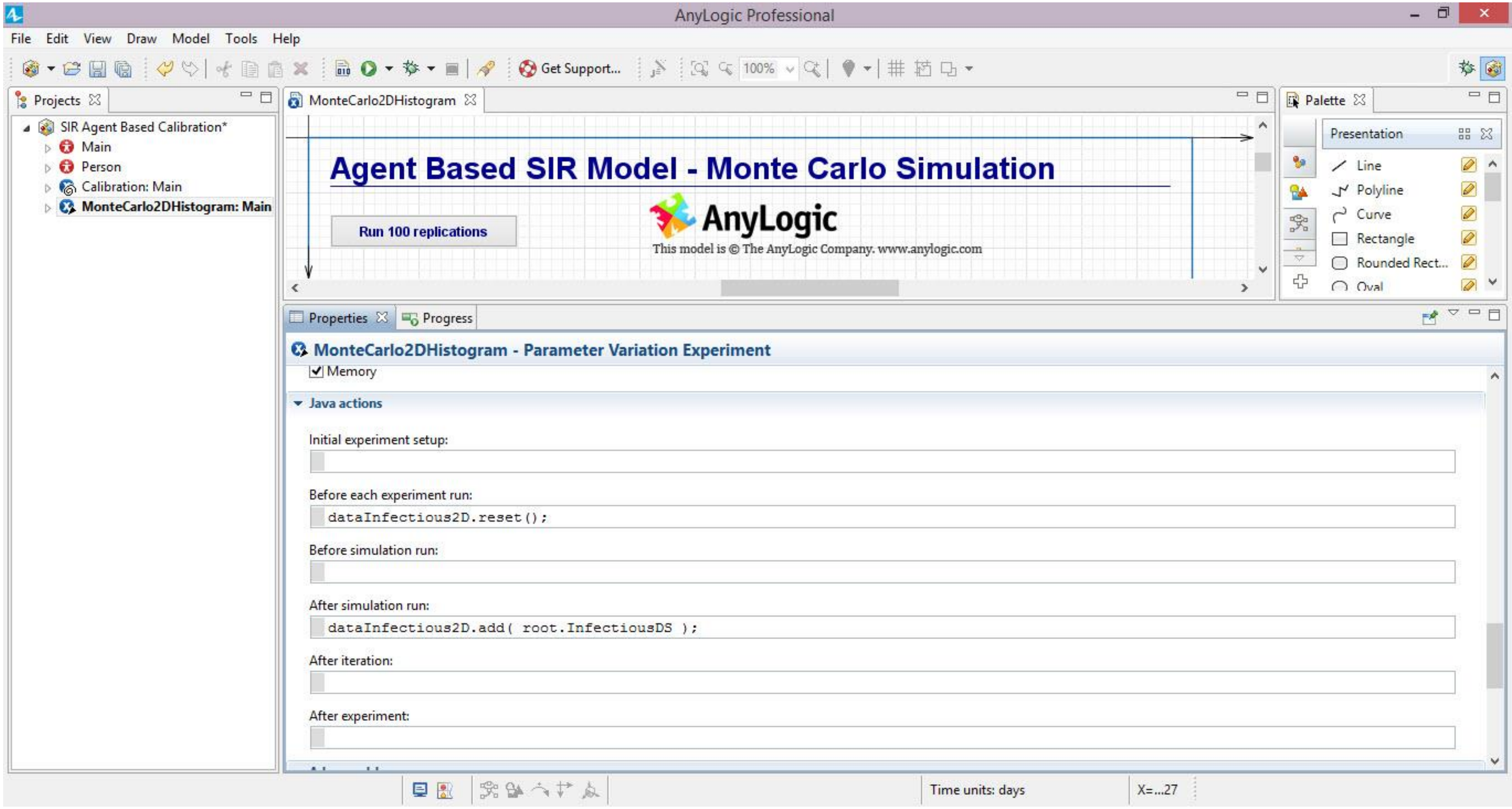
Monte Carlo Output After Some Runs



Monte Carlo Output After All Runs



Populating the 2D Dataset



Sensitivity in Initial States

- Frequently we don’t know the exact state of the system at a certain point in time
- A very useful type of sensitivity analysis is to vary the initial model state
- In aggregate models, this can be accomplished by
  – Varying the number of people in the stock via a parameter to adjust
- In an agent-based model, state has far larger dimensionality
  – Can modify different numbers of people with characteristic, location of people with characteristic, etc.