

Object Oriented Modeling and Simulation

Input Analysis and Specification

José M. Garrido C.

Department of Computer Science
College of Computing and Software Engineering
Kennesaw State University

Spring 2017

Includes the specification of:

- The types of input variables, deterministic or random
 - The probability distributions that random input variables follow
 - Estimation of the input parameters.
-
- The input data should be collected and analyzed and the model inputs can be specified in the model in a valid and realistic manner.
 - If the real system does not yet exist or no data is available, careful and reasonable assumptions need to be taken

- Determine what data is required, which depends on the model scope and the detail level decided on previously.
- Then data is collected by reviewing existing files and documentation, conducting interviews, direct observations, and taking assumptions, and other activities.

Analysis and Interpretation of Data

- Deciding their suitability of use in the simulation model then developing a representation of the data.
- The tests on data are performed for data independence (randomness), homogeneity, and stationarity.
- After data have been tested, it should be converted to a form that can be useful for simulation.

- The two basic types of input quantities are deterministic and random quantities.
- Sensitivity analysis can help to decide whether a quantity is deterministic or random.

Random Values

- When an input quantity in a real system exhibits some variability, then it is modeled as a random variable.
- The behavior of a random variable is defined by a probability distribution.
- A probability distribution describes the range of possible values that a random variable can attain and the probability that the value of the random variable is within any (measurable) subset of that range.

Representing Data

- There are several techniques to represent the data for use in the simulation model.
- The recommended technique is to select a theoretical probability distribution that best fits the data.
- Another technique is to use an empirical distribution that characterizes the data.

Histograms

- A histogram is a chart of tabulated frequencies, shown as bars.
- It shows what proportion of cases fall into each of several groups.
- The groups are usually specified as non-overlapping intervals of some variable.
- The categories (bars) must be adjacent.
- A histogram of the frequency distribution of the data is usually constructed first.
- A fitness test is next carried out to check whether one of the theoretical distributions fits the data.

Illustration of a Histogram

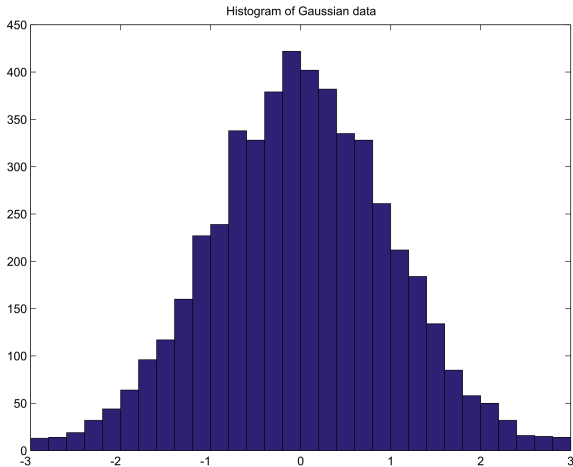


Figure: A histogram of a frequency distribution

Discrete

A probability distribution is called discrete if its cumulative distribution function only increases in jumps, i.e., a probability distribution is discrete if there is a finite or countable set whose probability is 1. A discrete distribution has a finite set of values that can be drawn from a range.

Continuous

A continuous distribution has an infinite set of possible values within a range. A continuous distribution describes probabilistic properties of a random variable which takes on a continuous (not countable) set of values.

Common Probability Distributions

- *Randint*
- *NegExp*
- *Poisson*
- *Normal*
- *Uniform*
- *Binomial*
- *Bernoulli*
- *Erlang*
- *Gamma*
- *Geometric*
- *HyperGeometric*
- *Triangular*
- *Weibull*