ANALYSIS OF SIMULATION DATA:

INPUT MODELING

Lecture # 27

Lecture by Engr. Sidra

Input Modeling

- Input models provide the driving force for a simulation model.
- Faulty models of inputs will lead to output whose interpretation could give rise to misleading recommendations.
- There are four steps in the development of a useful model of input data:

1. Collect data from the real system of interest.
   - *This often requires a substantial time and resource commitment. Unfortunately, in some situations it is not possible to collect data*
   - *for example, when time is extremely limited, when the input process does not yet exist, or when laws or rules prohibit the collection of data.*
   - *When data are not available, expert opinion and knowledge of the process must be used to make educated guesses.*

Input Modeling

2. Identify a probability distribution to represent the input process.
   - *When data are available, this step typically begins with the development of a frequency distribution, or histogram, of the data.*
   - *Given the frequency distribution and a structural knowledge of the process, a family of distributions is chosen.*
3. Choose parameters that determine a specific instance of the distribution family.
   - *When data are available, these parameters may be estimated from the data.*

Input Modeling

4. Evaluate the chosen distribution and the associated parameters for goodness of fit.
   - *Goodness of fit may be evaluated informally, via graphical methods, or formally, via statistical tests.*
   - *The chi-square and the Kolmogorov-Smirnov tests are standard goodness-of-fit tests.*
   - *If not satisfied that the chosen distribution is a good approximation of the data, then the analyst returns to the second step, chooses a different family of distributions, and repeats the procedure.*
   - *If several iterations of this procedure fail to yield a fit between an assumed distributional form and the collected data, the empirical form of the distribution may be used.*

1. Data Collection

- Data collection is one of the biggest tasks in solving a real problem.
- It is one of the most important and difficult problems in simulation.
- And even when data are available, they have rarely been recorded in a form that is directly useful for simulation input modeling.
- **Pitfalls**
  - *Stale Data*
  - *Unexpected Data*
  - *Time Varying Data*
  - *Dependent Data*