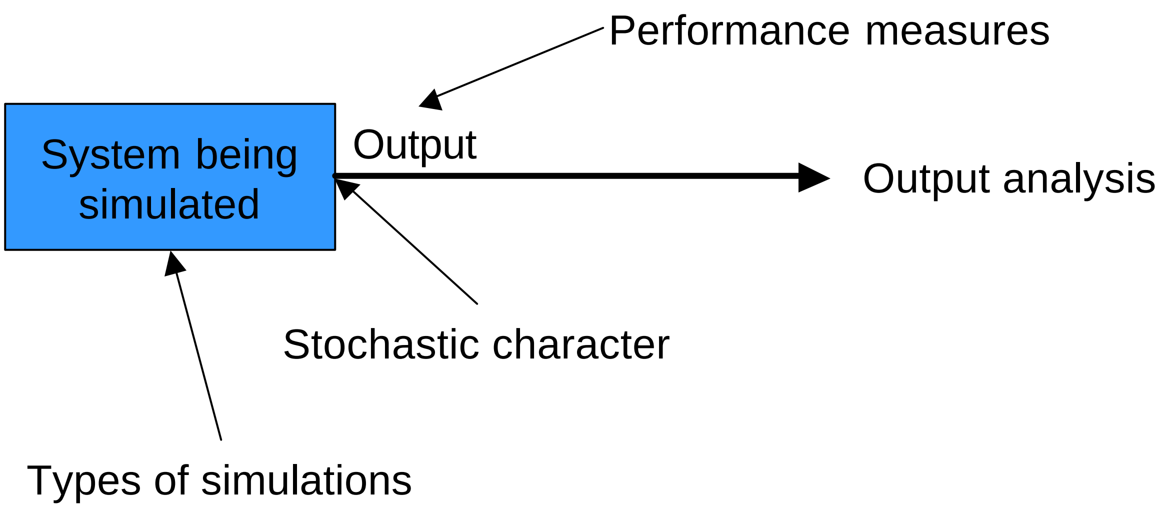
EE/CpE 345

Modeling and Simulation

Fall 2003

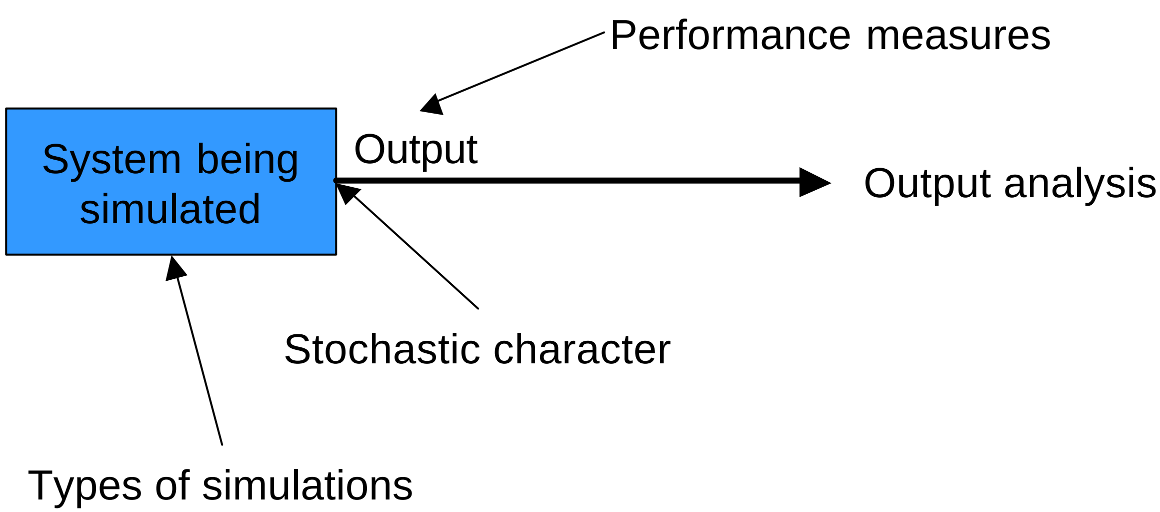
Class 11

Output Analysis for a Single Model



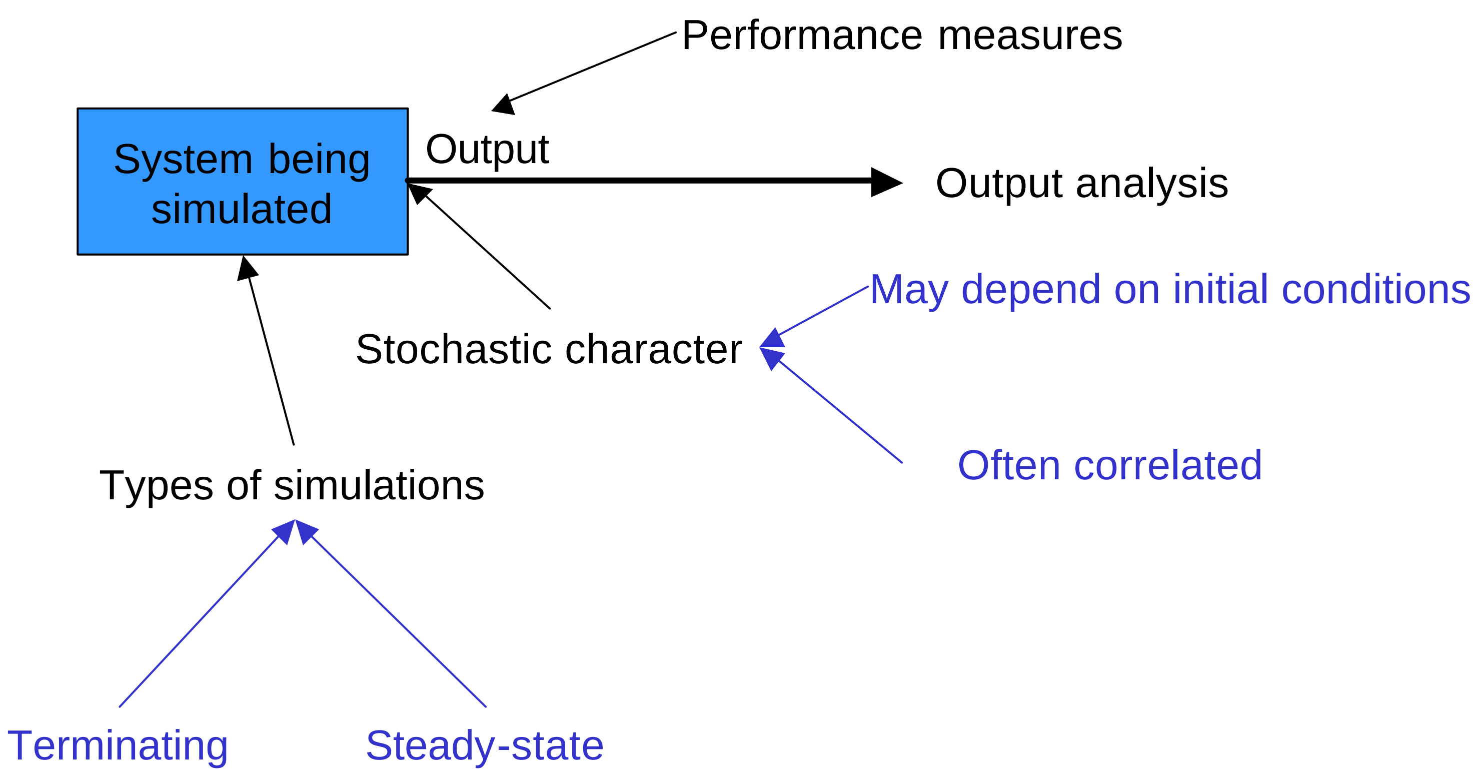
- Output analysis
  - Examination of data generated by simulation
    - To predict performance, or
    - To compare performance of two or more designs

Output Analysis for a Single Model

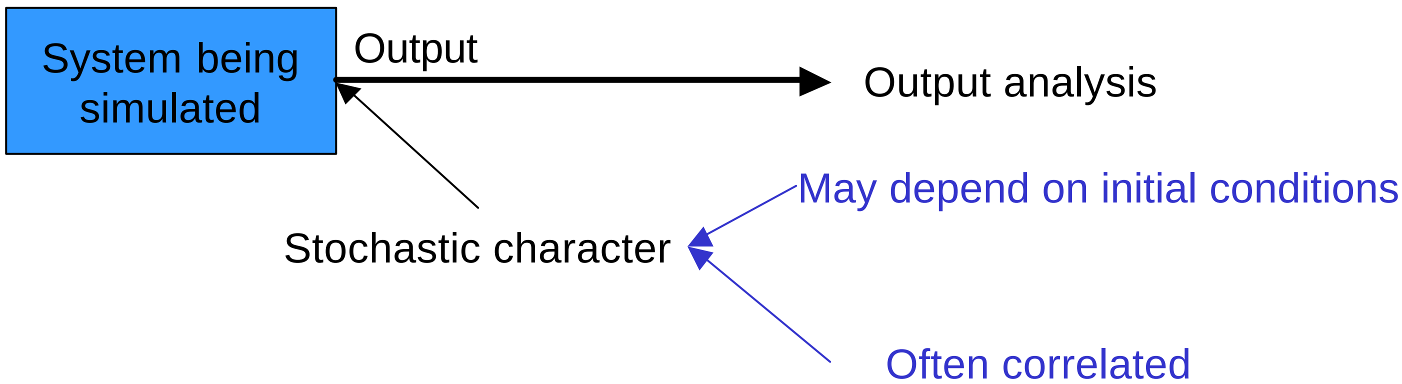


- Output analysis
  - Examination of data generated by simulation
    - **To predict performance**, or
    - To compare performance of two or more designs (Chapter 12)

Output Analysis for a Single Model

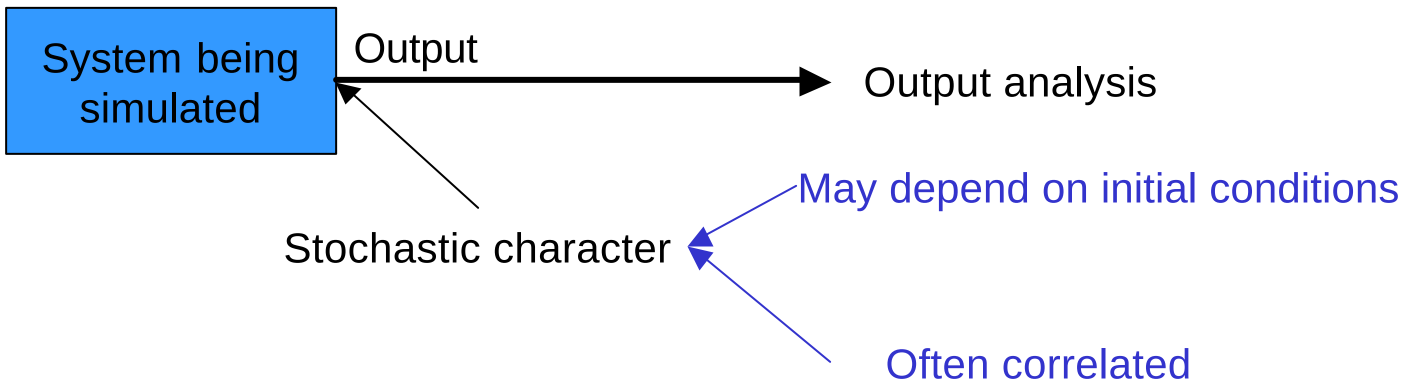


Output Analysis for a Single Model



- Inputs are generated randomly:
  - Outputs derived from inputs -> stochastic outputs
  - Outputs may depend on initial conditions of inputs and system
  - System behavior often nonlinear function of input

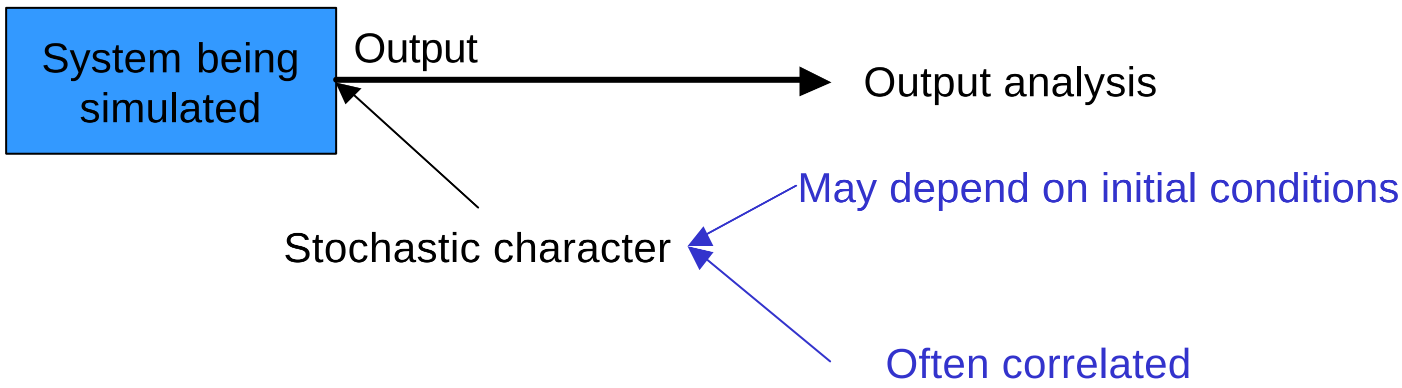
Output Analysis for a Single Model



- Inputs are generated randomly:
  - Outputs derived from inputs -> stochastic outputs
  - Outputs may depend on initial conditions of inputs and system
  - System behavior often nonlinear function of input

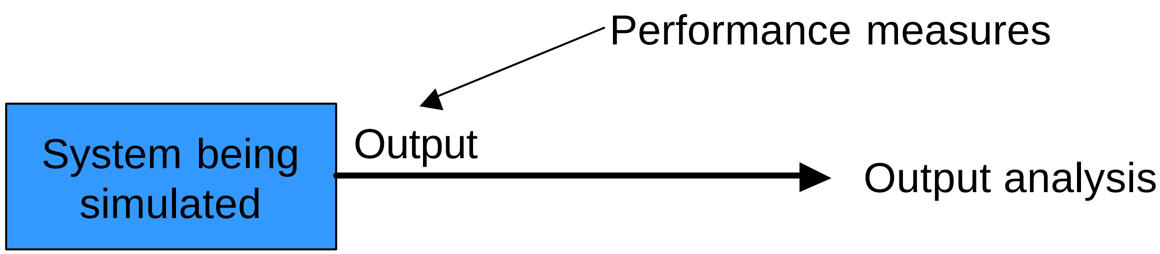
} Definition of a chaotic system

Output Analysis for a Single Model



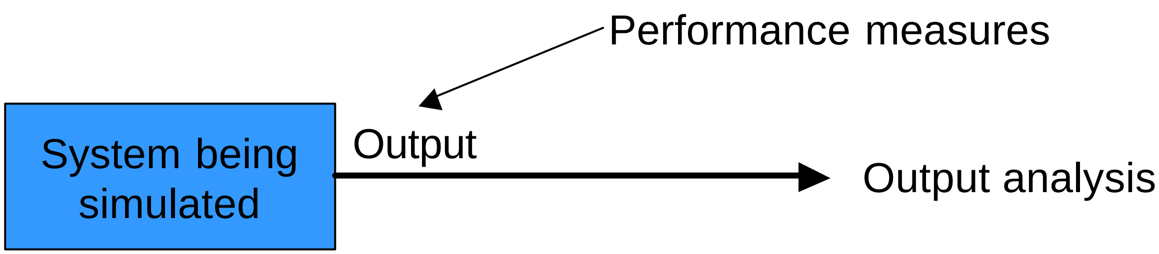
- Correlated outputs:
  - Consider time intervals: $T_1 = [t_1,t_2)$, $T_2 = [t_2, t_3)$, …
  - E.g., if average queue length was long in T_i, starting conditions in T_{i+1} will bias that interval to longer than average queue length

Output Analysis for a Single Model



θ: System performance measure

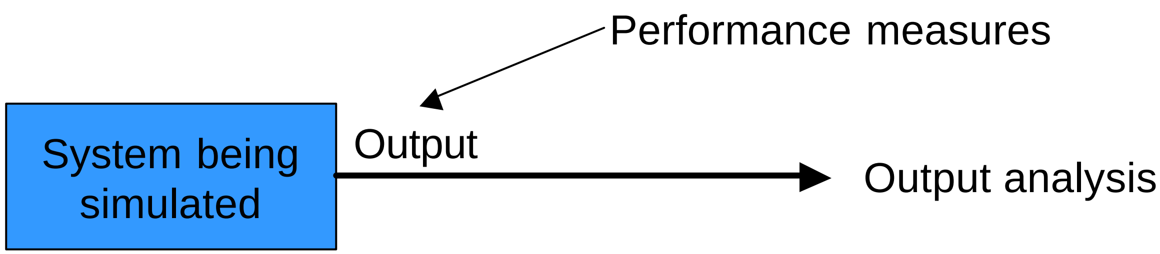
Output Analysis for a Single Model



θ: System performance measure

$\hat{\theta}$: Estimator of θ

Output Analysis for a Single Model

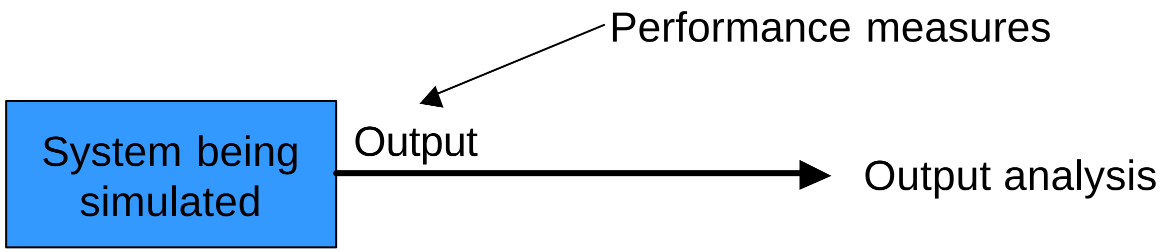


θ: System performance measure

$\hat{\theta}$: Estimator of θ

var($\hat{\theta}$): measure of precision of $\hat{\theta}$

Output Analysis for a Single Model



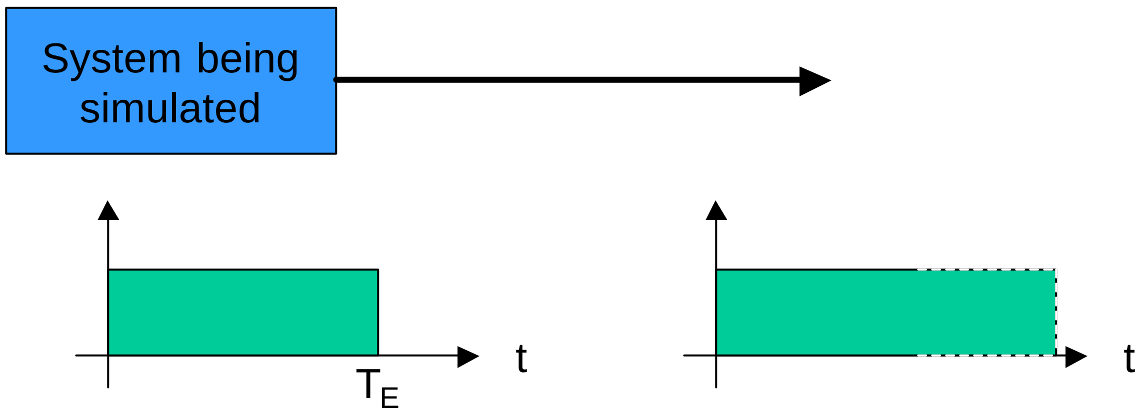
θ: System performance measure

$\hat{\theta}$: Estimator of θ

var($\hat{\theta}$): measure of precision of $\hat{\theta}$

Estimate var($\hat{\theta}$),
number of observations needed for given precision

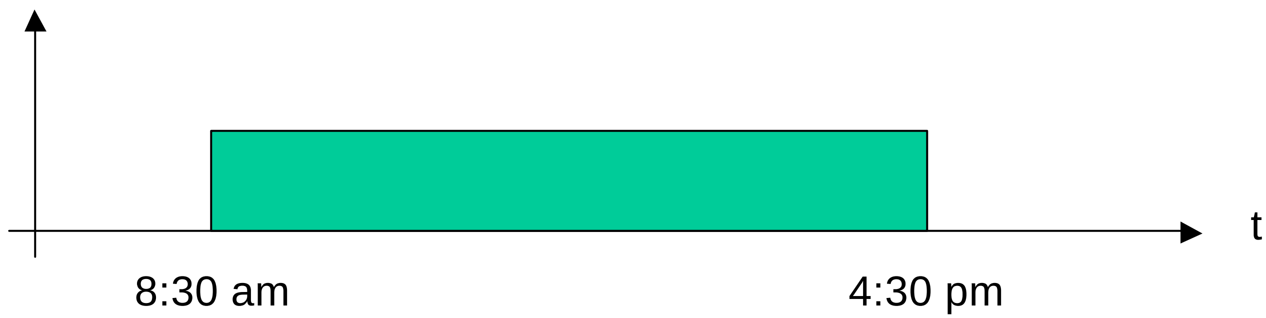
Types of Simulations with Respect to Output Analysis



- Terminating or transient simulation:
  - Runs until end time, T_E
  - Initial conditions at t=0 specified
  - Stop time, T_E, or stop event, E, must be specified

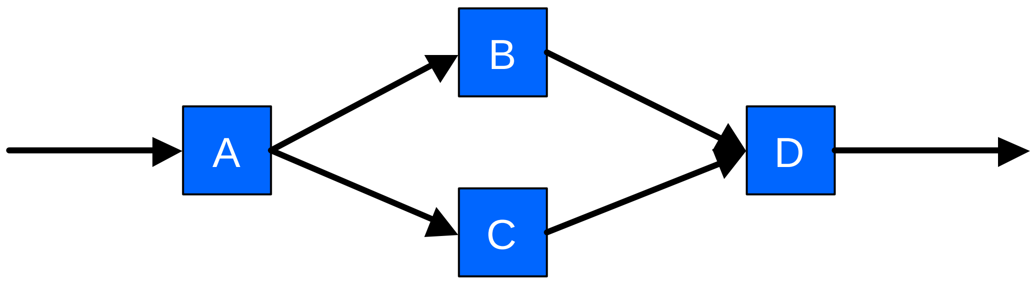
- Steady-state simulation:
  - Runs continuously or over a long period
  - Properties are not influenced by initial conditions

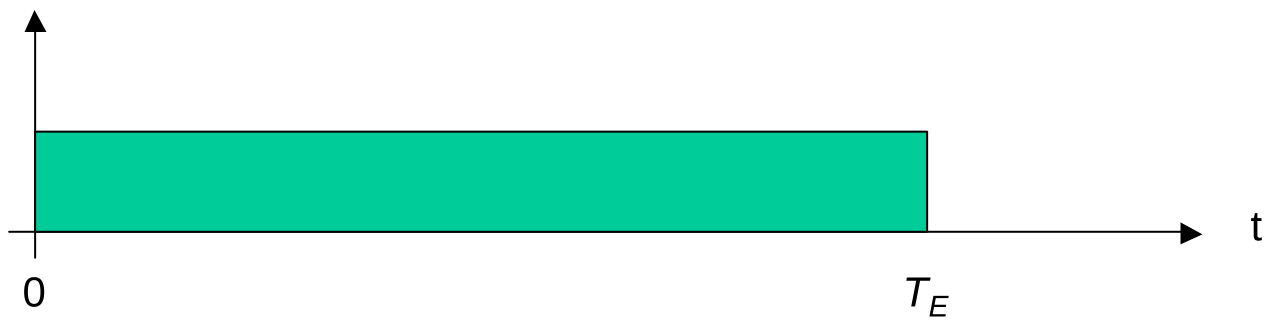
Terminating Simulations



- Example 11.1:
  - Bank opens at 8:30 am with 0 customers and 8 tellers
  - Bank closes at 4:30 pm
  - Termination event set at start of simulation

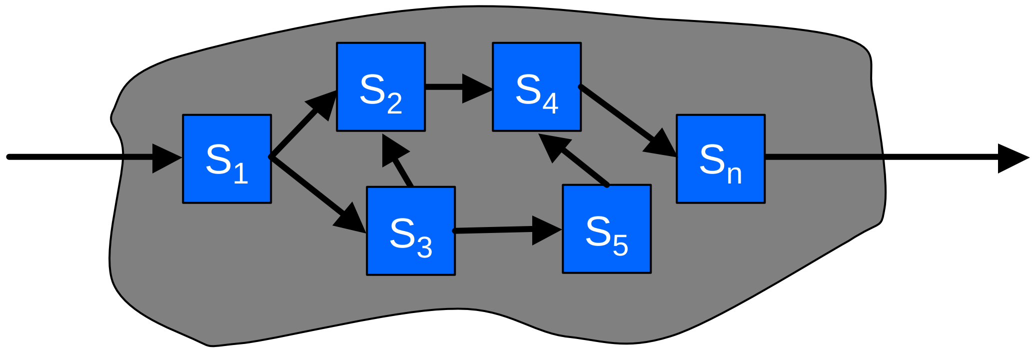
Terminating Simulations

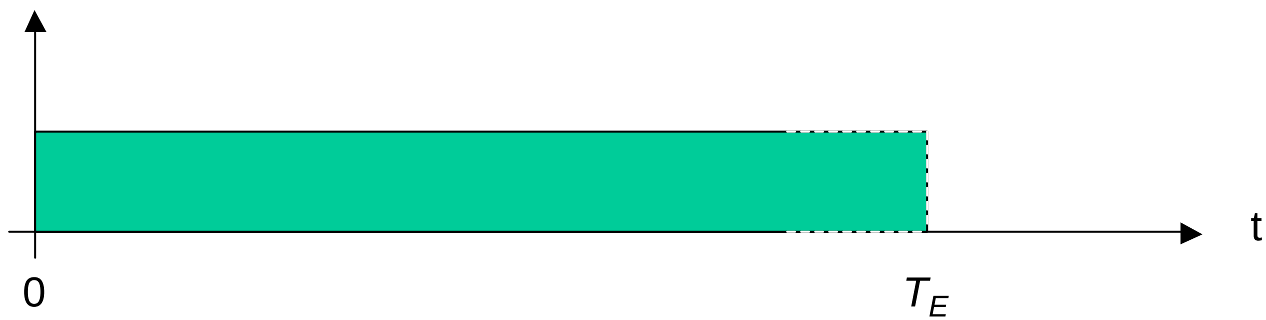




- Example 11.3:
  - System starts at t=0 with new components
  - Simulation continues until A or D or (B and C) fail
  - Termination time not known in advance

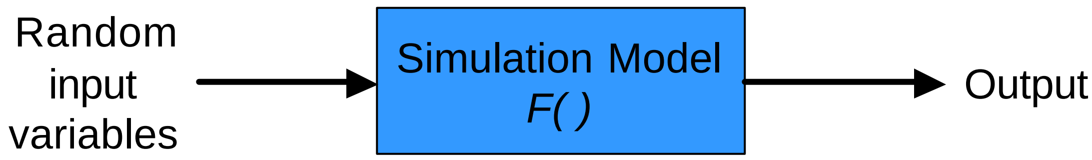
Steady-State Simulations





- Example 11.6:
  - Simulation ends:
    - at a arbitrarily specified T_E or
    - when output parameters are determined to specified precision

Stochastic Nature of Output Data

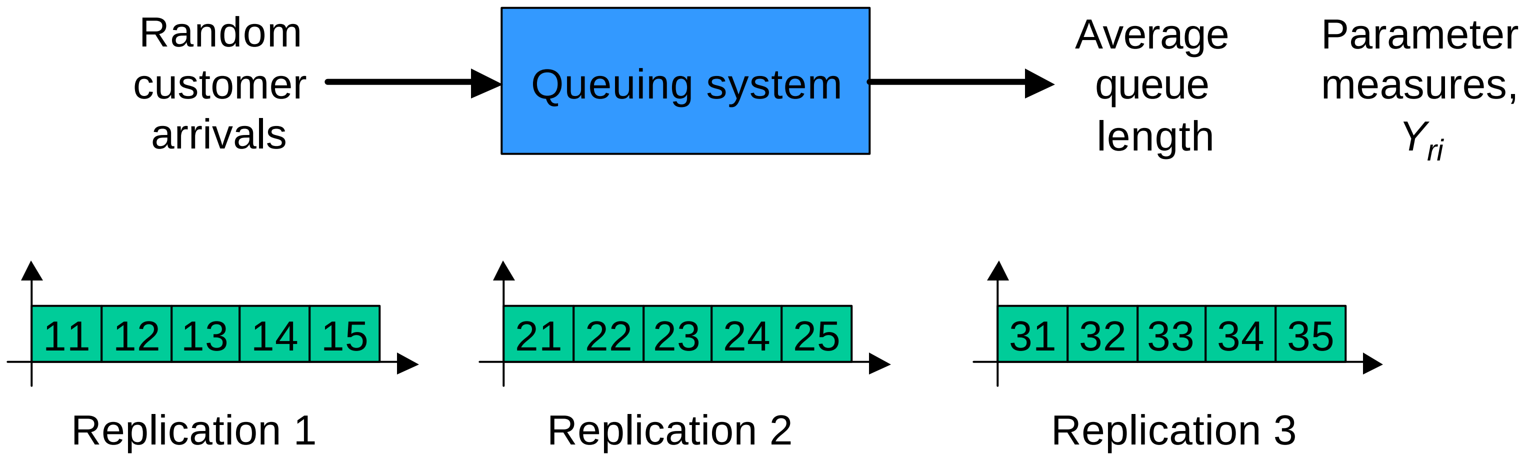


- Consider one run of a simulation model over [0,*T*]
- Model is an input-output transformation

- Run simulation model *n* times with different sets of input data
- Observe parameter $\hat{q}_r$, $r = 1,..,n$

- Two questions to be addressed by statistical analysis:
  – Estimation of true parameter $q = E(\hat{q}_r)$, “point estimate”
  – Estimation of error in point estimate - as standard error or confidence interval

Bias in Estimating Parameters
Example 11.8



- Run three independent trials (replications)
- Break each trial into 5 batches

- Batches with a replication are correlated and will give a bias in estimating parameters

Measures of Performance and Their Estimation

- Discrete time data: $q = f(\{Y_1, Y_2, \ldots, Y_n\})$, q is an ordinary mean
- Continuous time data: $f = f(\{Y(t), 0 <= t <= T_E\})$, f is a time-weighted mean

Point Estimation - Discrete Time

- Point estimator of q based on $\{Y_1, Y_2, \ldots, Y_n\}$:

$$\hat{q} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

- Estimator $\hat{q}$ is unbiased if:

$$E(\hat{q}) = q$$

- Estimator bias is:

$$E(\hat{q}) - q$$

Point Estimation - Continuous Time

- Point estimator of f based on $\{Y(t), 0 <= t <= T_E\}$:

$$\hat{f} = \frac{1}{T_E} \int_0^{T_E} Y(t)dt$$

- Estimator $\hat{f}$ is generally biased

Interval Estimation

- Estimation of the variance of the point estimator

- The true variance of the point estimator is:

$$s^{2}(\hat{q}) = \mathrm{var}(\widehat{q})$$

- The estimate of the variance of point estimator is:

$$\hat{s}^{2}(\hat{q}) = \widehat{\mathrm{var}}(\widehat{q})$$

  – Which is generally biased

Interval Estimation

- If the variance of the point estimator is nearly unbiased, the statistic, t

$$t = \frac{\hat{q} - q}{\hat{s}(\hat{q})}$$

- is approximately t-distributed with f degrees of freedom and can be bound by a 100(1-a)% confidence interval:

$$\hat{q} - t_{a/2,f}\,\hat{s}(\hat{q}) \le q \le \hat{q} + t_{a/2,f}\,\hat{s}(\hat{q})$$

- $t_{a/2,f}$ is the 100(1-a/2) percentage point of a t-distribution with f degrees of freedom: $P(t >= t_{a/2,f}) = a/2$

Output Analysis for Terminating Simulations

- Simulation runs over a time interval $[0, T_E]$
- Output observations: $\{Y_1, Y_2, \ldots, Y_n\}$; sample size n

- We want to estimate q:

$$q = E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)$$

- Repeat the simulation R times, with random initial conditions, independent random number streams from run to run

- Y_{ri} is the i^{th} observation within replication r
- Y_{r1}, Y_{r2} are correlated, but for different replications r and s, Y_{ri} and Y_{sj} are independent for all i and j
- Define a sample mean:

$$\hat{q}_r = \frac{1}{n_r}\sum_{i=1}^{n_r} Y_{ri}, \qquad r = 1, \ldots, R$$

Output Analysis for Steady-State Simulation

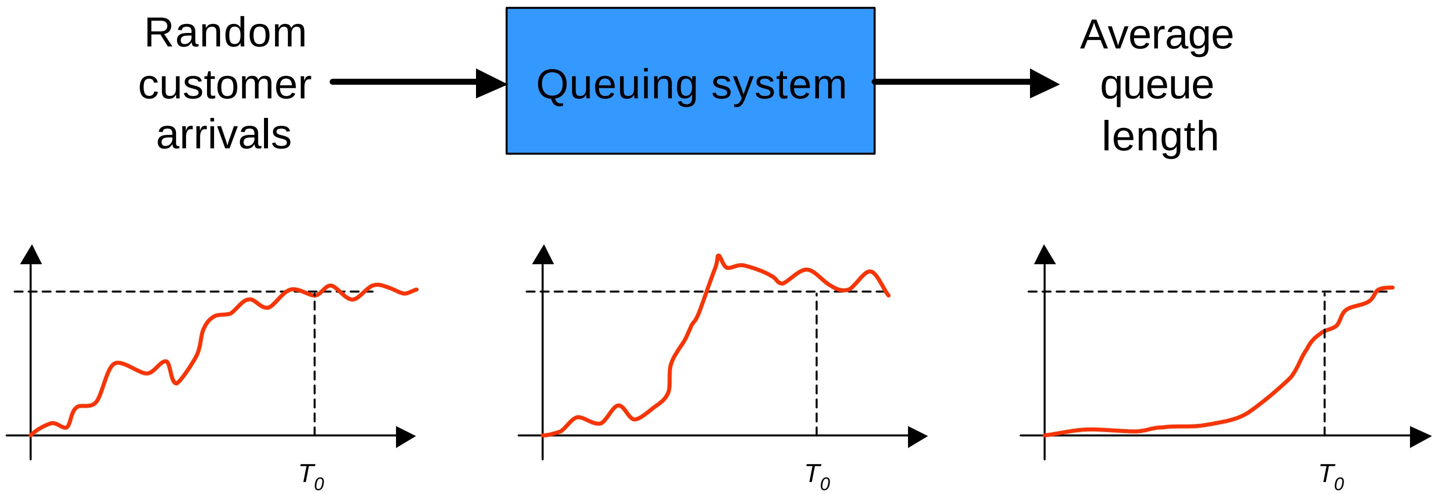
- A single run of a simulation model
- Goal is to estimate a steady-state parameter of system

- The run produces observations Y_1, Y_2, … which are samples of a correlated time series.
- Estimate:

$$q = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} Y_i$$

- Since simulation can't run forever, how do you choose *n* to terminate simulation? Base it on:
  - Bias in point estimator due to inappropriate initial conditions. This is more severe for short simulations. Averages out for long runs
  - Desired precision of point estimator - measure point estimator variance
  - Time or computer budget available.

Initialization Bias in Steady-State Simulations



- Simulation takes some time to stabilize, in this case to reach the average queue length expected for level of utilization
- Rather than initializing simulation with no customers in queue, consider starting system with starting state near expected steady-state condition. Base this initial condition on:
  - observations of real system
  - analysis or simulation of simpler system
- Alternative: Don't collect data until some time, T_0, after system has had time to stabilize

Homework

- Complete your projects, due next class

- Project report should (at least) address:
  - **Introduction/background**: What was the problem you studied?
  - **Assumptions**: How did you develop your model?
  - **Observations of physical system**: Representative input data you collected
  - **Simulation program**: Listing of all code needed to build your simulation
  - **Simulation results**: Representative simulation execution output**s**
  - **Validation/Verification**: What leads you to believe that your model & simulation are meaningful?
  - **Conclusions**: What did your simulation show?
  - **Recommendations**: How would you modify the physical system or how you use it to improve performance?
  - **Future work**: What follow on activities would be appropriate if you were to continue this simulation work?
  - **References**: Cite previous results or information you based your work on