Output Analysis for Simulations

Yu Wang

Dept of Industrial Engineering
University of Pittsburgh

Feb 16, 2009
Why output analysis is needed

• Simulation includes randomness => random output
• Statistical techniques must be used to analyze the results
• Typical assumption is not true
  – Independent and identically distributed (i.i.d.)
  – Normal distribution
Performance Measures

- Transient performance measures
  - Terminating or finite-horizon measures
  - Evaluate the system's evolution over a finite time horizon
  - Transient simulation

- Steady-state performance measures
  - Long-run or infinite-horizon measures
  - How the system evolves over an infinite time horizon.
  - Steady-state simulation
**Demo - Problem Statement**

- **A simple inventory system:**
  - Demands: Independent Poisson random variables with mean $\lambda$ (for a given product on successive days are)
  - $X_j$ is the stock level at the beginning of day $j$
  - $D_j$ is the demand
  - Sales: $\min(D_j , X_j)$
  - Lost Sales: $\max(0,D_j − X_j)$
  - Stock at the end of the day is $Y_j = \max(0,X_j − D_j)$
  - Revenue for each sale: $c$
  - Holding cost for each unsold item: $h$
Demo - Problem Statement (cont'd)

- Inventory control: \((s, S)\) policy:
  - If \(Y_j < s\), order \(S - Y_j\) items, Else, do not order

- Order is made in the evening:
  - with prob \((p)\): it arrives during the night and can be used for the next day,
  - with prob \((1-p)\): it never arrives (in which case a new order will have to be made the next evening)

- For arrived order:
  - Fixed cost \(K\) + Marginal cost of \(k / item\).

- The stock at the beginning of the first day
  - \(X_0 = S\).
Output Analysis for Transient Simulation
Transient Performance Measures

- Terminating simulation:
  - A “natural” event B specifies the length of time in which one is interested for the system.
  - “Initial conditions” can have a large impact

- Example: Inventory system
  - $E[Z]$, expected value of profit per day (1 year)
  - $P\{Z\geq 50\} = E[I(Z\geq 50)]$, profit is large than $50 in a day
Output Analysis for Transient Simulation

- **Point estimator:**

  \[
  E(X) \approx \bar{X}(n) = \frac{1}{n} \sum_{i=1}^{n} X_i
  \]

  \[
  \sigma^2 \approx S^2(n) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}(n))^2
  \]

- **By Central Limit Theorem (CLT)**

  \[
  \frac{(\bar{X}(n) - \mu)}{S(n) / \sqrt{n}} \overset{D}{\approx} N(0, 1)
  \]

  
  \[
  1 - \delta = P\{-z \leq N(0, 1) \leq z\} \approx P\{-z \leq \frac{\sqrt{n}}{S(n)}(\bar{X}(n) - \mu) \leq z\}
  \]

  \[
  = P\{\mu \in [\bar{X}(n) \pm \frac{zS(n)}{\sqrt{n}}]\},
  \]

  \[
  P\left\{\mu \in \left[\bar{X}(n) - \frac{zS(n)}{\sqrt{n}}, \bar{X}(n) + \frac{zS(n)}{\sqrt{n}}\right]\right\} \approx 1 - \delta
  \]
Demo - Transient Simulation

- Simulate Parameters:
  - Demand (Poisson) mean $\lambda = 100.0$
  - Sale price $c = 2.0$
  - Holding cost $h = 0.1$
  - Fixed ordering cost $K = 10.0$
  - Marginal ordering cost (per item) $k = 1.0$
  - Probability that an order arrives $p = 0.95$

- Simulate 365 days (1 year):
  - Control policy $(s, S) = (80, 200)$
  - Replicate the simulation 100 times
**Demo - Fixed Sample Size Output (n=100)**

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>84.61</td>
<td>84.62</td>
<td>85.16</td>
<td>84.78</td>
<td>86.26</td>
<td>86.57</td>
<td>85.58</td>
<td>83.47</td>
<td>84.59</td>
<td>84.33</td>
</tr>
<tr>
<td>86.42</td>
<td>85.77</td>
<td>84.46</td>
<td>86.41</td>
<td>84.89</td>
<td>86.02</td>
<td>86.43</td>
<td>85.91</td>
<td>85.98</td>
<td>85.13</td>
</tr>
<tr>
<td>84.71</td>
<td>85.14</td>
<td>85.76</td>
<td>84.55</td>
<td>84.35</td>
<td>85.04</td>
<td>87.06</td>
<td>86.50</td>
<td>85.13</td>
<td>83.76</td>
</tr>
<tr>
<td>86.16</td>
<td>85.67</td>
<td>84.84</td>
<td>84.51</td>
<td>85.15</td>
<td>86.01</td>
<td>84.54</td>
<td>84.79</td>
<td>83.81</td>
<td>85.73</td>
</tr>
<tr>
<td>85.20</td>
<td>85.73</td>
<td>85.53</td>
<td>85.48</td>
<td>85.16</td>
<td>85.20</td>
<td>84.53</td>
<td>86.25</td>
<td>85.70</td>
<td>85.44</td>
</tr>
<tr>
<td>84.60</td>
<td>85.31</td>
<td>84.98</td>
<td>86.31</td>
<td>85.10</td>
<td>84.39</td>
<td>84.95</td>
<td>85.28</td>
<td>84.82</td>
<td>85.29</td>
</tr>
<tr>
<td>86.00</td>
<td>84.29</td>
<td>84.87</td>
<td>83.85</td>
<td>85.96</td>
<td>85.69</td>
<td>84.56</td>
<td>86.40</td>
<td>85.14</td>
<td>85.41</td>
</tr>
<tr>
<td>85.04</td>
<td>84.27</td>
<td>85.07</td>
<td>85.39</td>
<td>86.00</td>
<td>85.81</td>
<td>86.02</td>
<td>84.81</td>
<td>86.30</td>
<td>83.26</td>
</tr>
<tr>
<td>86.11</td>
<td>85.83</td>
<td>84.54</td>
<td>86.93</td>
<td>85.45</td>
<td>83.83</td>
<td>86.35</td>
<td>84.85</td>
<td>85.86</td>
<td>85.05</td>
</tr>
<tr>
<td>85.13</td>
<td>85.08</td>
<td>83.72</td>
<td>85.10</td>
<td>86.30</td>
<td>86.17</td>
<td>85.45</td>
<td>86.01</td>
<td>84.96</td>
<td>85.77</td>
</tr>
</tbody>
</table>

Annual Average Profit

\[ \bar{X}(n) = \frac{1}{n} \sum_{i=1}^{n} X_i = 85.285 \]

\[ S^2(n) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}(n))^2 = (0.788)^2 \]

For \( \delta = 0.05 \Rightarrow z = 1.96 \) \( \left( \text{by } P \left\{ \mu \in \left[ \bar{X}(n) - \frac{zS(n)}{\sqrt{n}}, \bar{X}(n) + \frac{zS(n)}{\sqrt{n}} \right] \right\} \approx 1 - \delta \right) \)

95.0% confidence interval: [85.131, 85.439]
Pre-specifying C.I. Width

- Fixed-sample size methods – sample size is fixed prior to running any simulations

- Want to get an estimator with a small Pre-specified error $\varepsilon$, $\bar{X}(n) \pm \varepsilon$ at $100(1 - \delta)\%$ C.I.

- Half width:

\[ H_n = \frac{zS(n)}{\sqrt{n}} = \varepsilon \rightarrow n = \left( \frac{zS(n)}{\varepsilon} \right)^2 \]
Pre-specifying C.I. Width (cont'd)

• Two-stage procedure:
  
  – 1\textsuperscript{st} stage: take \( n_0(\geq 50) \) runs to estimate variance \( S_1^2(n_0) \)
  
  – Calculate needed runs: \( N_\alpha(\varepsilon) = \left( \frac{zS_1(n_0)}{\varepsilon} \right)^2 \)
  
  – 2\textsuperscript{nd} stage: generate \( N_\alpha(\varepsilon) \) independent runs and calculate estimator & C.I.
    \[
    \left[ \bar{X}(\varepsilon) - \frac{z\tilde{S}(\varepsilon)}{\sqrt{N_\alpha(\varepsilon)}}, \bar{X}(\varepsilon) + \frac{z\tilde{S}(\varepsilon)}{\sqrt{N_\alpha(\varepsilon)}} \right]
    \]
  
• For relative precision \( \varepsilon \) (e.g., ±5\%) \( N_r(\varepsilon) = \left( \frac{z\tilde{S}_1(n_0)}{\bar{X}(n_0)\varepsilon} \right)^2 \)
Demo - Pre-specifying C.I. Width (absolute)

- Want half width: 
  \[ H_n = \frac{zS(n)}{\sqrt{n}} = \varepsilon = 0.01 \]

- Two-stage procedure:
  - 1\textsuperscript{st} stage: \( n_0 = 100 \quad S_1(n_0) = 0.788 \)
  - Needed runs:
    \[ N_a(\varepsilon) = \left\lfloor \left( \frac{zS_1(n_0)}{\varepsilon} \right)^2 \right\rfloor = \left\lfloor \left( \frac{1.96 \times 0.788}{0.01} \right)^2 \right\rfloor = 23855 \]
  - 2\textsuperscript{nd} stage: generate 23855 independent runs.
    \[ \hat{X}(\varepsilon) = 85.117 \quad \hat{S}(\varepsilon) = 0.782 \]

95.0% confidence interval:
\[
\left[ \hat{X}(\varepsilon) - \frac{z\hat{S}(\varepsilon)}{\sqrt{N_a(\varepsilon)}}, \hat{X}(\varepsilon) + \frac{z\hat{S}(\varepsilon)}{\sqrt{N_a(\varepsilon)}} \right] = [85.108, 85.127] \]
Demo - Pre-specifying C.I. Width (relative)

- Want C.I. to be ±0.5% of the estimator: \( \varepsilon = 0.005 \bar{X}(n_0) \)

- Two-stage procedure:
  - 1\(^{\text{st}}\) stage: \( n_0 = 100 \quad \bar{X}_1(n_0) = 85.285 \quad S_1(n_0) = 0.788 \)
  - Needed runs:
    \[
    N_r(\varepsilon) = \left[ \left( \frac{zS_1(n_0)}{X_1(n_0)\varepsilon} \right)^2 \right] = \left[ \left( \frac{1.96 \times 0.788}{85.285 \times 0.005} \right)^2 \right] = 14
    \]
  - 2\(^{\text{nd}}\) stage: generate 14 independent runs.
    \( \tilde{X}(\varepsilon) = 85.217 \quad \tilde{S}(\varepsilon) = 0.958 \)

95.0% confidence interval:
\[
\left[ \bar{X}(\varepsilon) - \frac{z\tilde{S}(\varepsilon)}{\sqrt{N_r(\varepsilon)}}, \bar{X}(\varepsilon) + \frac{z\tilde{S}(\varepsilon)}{\sqrt{N_r(\varepsilon)}} \right] = [84.715, 85.719]
Output Analysis for
Steady-State Simulation
Steady-State Performance Measures

- Output is a (discrete-time) stochastic process: \( \mathbf{Y} = (Y_1, Y_2, Y_3, \ldots) \)
- Define \( F_i(y \mid \mathcal{C}) = P(Y_i \leq y \mid \mathcal{C}) \) as the distribution function of \( Y_i \) given the initial conditions \( \mathcal{C} \)
- If \( F_i(y \mid \mathcal{C}) \rightarrow F(y) \) as \( i \rightarrow \infty \), then \( F(y) \) is the steady-state distribution of the process \( \mathbf{Y} \), or “\( Y_i \) converges in distribution to \( \mathbf{Y} \)”
Steady-State Performance Measures (cont'd)

- $Y$ is a random variable with $F$, then $E(Y)$ can be used as a steady-state performance measure.

$F_i(y \mid C) \approx F(y)$, for all $y$
Output Analysis for Steady-State Simulation

• Discrete-time Process: \( Y = (Y_i : i = 1, 2, \ldots) \)

• Continuous-time Process: \( Y = (Y(s) : s \geq 0) \)

\[
\frac{1}{m} \sum_{i=1}^{m} Y_i \to \nu \quad \frac{1}{s} \int_{0}^{s} Y(u) du \to \nu
\]

\[
\frac{\sqrt{m}}{\bar{\sigma}} (\bar{Y}(m) - \nu) \overset{\mathcal{D}}{\approx} N(0, 1)
\]

• Difficulty in estimating \( \bar{\sigma} \),

• Not satisfy i.i.d. assumption
Multiple Replications

- Replace one long replication (length=m) by \( r \) (10\(\leq r \leq 30\)) i.i.d. replications (length \( k = m/r \))
  - Independence: non-overlapping streams of random number for different replications
  - Identical distributed:
    - Same initial conditions
    - Same system dynamics
- Get sample variance across the replications
Multiple Replications (cont'd)

- Output from \( r \) replications (length = \( k \))

\[
\begin{align*}
Y_{1,1} & & Y_{1,2} & & Y_{1,3} & & \cdots & & Y_{1,k}, \\
Y_{2,1} & & Y_{2,2} & & Y_{2,3} & & \cdots & & Y_{2,k}, \\
\vdots & & \vdots & & \vdots & & \cdots & & \vdots \\
Y_{r,1} & & Y_{r,2} & & Y_{r,3} & & \cdots & & Y_{r,k}.
\end{align*}
\]

\[
X'_j = \frac{1}{k} \sum_{i=1}^{k} Y_{j,i}
\]

\[
\bar{X}'(r) = \frac{1}{r} \sum_{i=1}^{r} X'_i \\
S'^2(r) = \frac{1}{r - 1} \sum_{j=1}^{r} (X'_{j} - \bar{X}'(r))^2
\]

\[
\left[ \bar{X}'(r) - \frac{tS'(r)}{\sqrt{r}}, \bar{X}'(r) + \frac{tS'(r)}{\sqrt{r}} \right]
\]

Where \( t \equiv t_{r-1,1-\delta/2} \) is chosen such that \( P\{T_{r-1} \leq t\} = 1 - \delta/2 \)
Demo - Steady-state Simulation

• Multiple Replication:
  - Output from \( r = 30 \) replications (length \( k = 1000 \))

\[
\bar{X}'(r) = \frac{1}{r} \sum_{i=1}^{r} X'_j = 85.151
\]

\[
S'^2(r) = \frac{1}{r - 1} \sum_{j=1}^{r} (X'_j - \bar{X}'(r))^2 = (0.478)^2
\]

For \( \delta = 0.05 \Rightarrow t \equiv t_{r-1,1-\delta/2} = t_{29,0.975} = 2.045 \)

95.0% confidence interval:

\[
\left[ \bar{X}'(r) - \frac{tS'(r)}{\sqrt{r}}, \bar{X}'(r) + \frac{tS'(r)}{\sqrt{r}} \right] = [84.973, 85.330]
\]

(C.I. Width) /2 = 0.1785
Multiple Replications (cont'd)

• Major problem:
  $\bar{x}'(r)$ may be significantly biased by the initial conditions

• Solution: initial-data deletion

\[
X_j = \frac{1}{k - c} \sum_{i = c + 1}^{k} Y_{j, i}
\]

using $X_j$ to compute the estimator $\bar{X}(r)$ and C.I.
Demo - Steady-state Simulation (cont'd)

- Long-run Average Profit:
Demo - Steady-state Simulation (cont'd)

- Multiple Replication:
  - Initial-data deletion: delete first 15 observations
  - Output from \( r = 30 \) replications (length \( k = 1000-15 = 985 \))

\[
\bar{X}(r) = 85.083 \quad S^2(r) = (0.463)^2
\]

95.0% confidence interval:

\[
\left[ \bar{X}'(r) - \frac{tS'(r)}{\sqrt{r}}, \bar{X}'(r) + \frac{tS'(r)}{\sqrt{r}} \right] = [84.910, 85.255]
\]

\[
\text{(C.I. Width) /2} \quad 0.1729
\]
Single-Replicate Methods

- Typically, two obs. \( Y_i \) and \( Y_{i+p} \) are almost independent when \( p \) is large

\[
\begin{align*}
\text{Batch 1:} & \quad Y_1 Y_2 \ldots Y_b \\
\text{Batch 2:} & \quad Y_{b+1} Y_{b+2} \ldots Y_{2b} \\
\text{Batch 3:} & \quad Y_{2b+1} Y_{2b+2} \ldots Y_{3b} 
\end{align*}
\]

- If \( b \) is large, most obs. in one batch are almost independent to those in other batch (except for the adjacent one)

- The sample mean of each of the batches:
  - Almost independent
  - Close to normal distribution
Single-Replicate Methods

- Total Run Length: \( m \)
- Number of Batches: \( n \) \((10 \leq n \leq 30)\)
- Batch Size: \( b = m / n \)

- Batch mean \( \bar{Y}_j(b) = \frac{1}{b} \sum_{l=(j-1)b+1}^{jb} Y_l \)

- Estimator
  \[
  \bar{Y}(n, b) = \frac{1}{n} \sum_{j=1}^{n} \bar{Y}_j(b) = \frac{1}{m} \sum_{i=1}^{m} Y_i
  \]
  \[
  S^2(n, b) = \frac{1}{n-1} \sum_{j=1}^{n} \left( \bar{Y}_j(b) - \bar{Y}(n, b) \right)^2
  \]

- C.I.
  \[
  \left[ \bar{Y}(n, b) - \frac{tS(n, b)}{\sqrt{n}}, \bar{Y}(n, b) + \frac{tS(n, b)}{\sqrt{n}} \right]
  \]
**Demo - Steady-state Simulation (cont'd)**

- Single-Replicate (Batch Mean):
  - With initial-data deletion: delete first 15 observations
  - Parameters:
    - Batch number $n = 30$, Batch size $b = 985$
    - Total runs length $m = 30 \times 985 + 15 = 29565$

\[
\bar{Y}(n, b) = 85.121 \\
S^2(n, b) = (0.444)^2
\]

95.0% confidence interval:

\[
\left[ \bar{Y}(n, b) - \frac{tS(n, b)}{\sqrt{n}}, \bar{Y}(n, b) + \frac{tS(n, b)}{\sqrt{n}} \right] = [84.956, 85.287]
\]

(C.I. Width) /2 | 0.1658
Other Methods for Steady-State Output Analysis

- Spectral (Anderson 1994)
- Regenerative (Crane and Iglehart 1975, Shedler 1993)
- Standardized time series methods (Schruben 1983)
Multiple Performance Measures

- Different measures: $\mu_1, \mu_2, \mu_3, \ldots$
- Joint Confidence Level:
  - Bonferroni's inequality
    \[ P\{\mu_s \in I_s, \text{for all } s = 1, 2, \ldots, q\} \geq 1 - \sum_{s=1}^{q} \delta_s \]
  - Joint confidence level is less than the C.L. for any individual
- Compare different systems for the “Best”
  - Selection procedures
  - Multiple comparison procedures
Other Useful Methods

- Variance-reduction techniques
  - Common random numbers
  - Antithetic variates
  - Control variates
  - Importance sampling
  - Stratified sampling
  - Conditional Monte-Carlo
  - Splitting
  - Etc.
Conclusion

- Introduce techniques for statistically analyzing the output from a simulation
  - Transient
    - Fixed-sample size
    - Pre-specifying C.I. width
  - Steady-state
    - Multiple replications
    - Batch mean
    - Other methods
  - Multiple Performance Measures

- All methods are *Asymptotically* Valid:
  - Large run lengths are needed
Thank You!