Introduction to Modeling and Generating Probabilistic Input Processes for Simulation

Emily K. Lada
SAS Institute Inc.

Natalie M. Steiger
University of Maine

Mary Ann Wagner
SAIC

James R. Wilson
NC State University

<www.ie.ncsu.edu/jwilson>

December 6, 2005
OVERVIEW

I. Introduction

II. Univariate Input Models
   A. Generalized Beta Distribution Family
   B. Johnson Translation System of Distributions
   C. Bézier Distribution Family

III. Multivariate Input Models
   A. Multivariate Johnson Translation System
   B. Matching Exactly a Given Mean Vector and Covariance Matrix

IV. Time-Dependent Arrival Processes

V. Conclusions and Recommendations
I. Introduction

- Stochastic simulations require valid input models—e.g., probability distributions that accurately mimic the random input processes driving the target system.
• Problems in using many conventional probability models:

1. They cannot adequately represent real-world behavior, e.g. in the tails of the underlying distribution.

2. Parameter estimation based on sample data or subjective information (expert opinion) is often troublesome.

3. Fine-tuning the fitted model is difficult; e.g., many conventional probability distributions have the following drawbacks—
   (a) A limited number of parameters available to control the fitted distribution, and
   (b) No effective mechanism for directly manipulating the fitted distribution while simultaneously updating its parameter estimates.
• Conventional approach to identifying an input model uses sample data to select from a list of well-known alternatives based on
  1. informal graphical techniques such as probability plots, \( Q-Q \) plots, histograms, empirical frequency distributions, or box-plots; and
  2. statistical goodness-of-fit tests such as the Kolmogorov-Smirnov, chi-squared, and Anderson-Darling tests.
• Drawbacks of conventional input modeling

1. Visual comparison of a histogram to a fitted probability density function (p.d.f.) depends on the (arbitrary) layout of the histogram.

2. Problems with statistical goodness-of-fit tests include:
   (a) In small samples, low power to detect lack of fit results in an inability to reject any alternatives.
   (b) In large samples, practically insignificant fit discrepancies result in rejection of all alternatives.
• Problems in estimating the parameters of the selected input model from sample data:
  – Matching the mean and standard deviation of the fitted distribution with that of the sample often fails to capture relevant shape characteristics.
  – Some estimation methods, such as maximum likelihood and percentile matching, may simply fail to estimate some parameters.
  – Users lack a comprehensive basis for selecting the “best-fitting” model.
• Problems with parameter estimation based on subjective information (expert opinion):
  – Subjective estimates of moments such as the mean and standard deviation can be unreliable and depend critically on the units of measurement.
  – Subjective estimates of extreme quantiles (e.g., lower and upper limits of the fitted distribution) are unreliable.

• Practitioners lack definitive procedures for identifying and estimating valid input models; thus, output analysis is often based on incorrect input processes.

• We focus on methods for input modeling that alleviate many of these problems.
II. Univariate Input Models

A. Generalized Beta Distribution Family

If $X$ is a continuous random variable with lower limit $a$ and upper limit $b$ whose distribution is to be approximated and subsequently sampled in a simulation, then often we can model the behavior of $X$ using a generalized beta distribution.

- Generalized beta p.d.f.

$$f_X(x) = \frac{\Gamma(\theta_1 + \theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)(b - a)^{\theta_1+\theta_2}} (x - a)^{\theta_1-1} (b - x)^{\theta_2-1} \text{ for } a \leq x \leq b,$$

where $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, dt$ (for $z > 0$) denotes the gamma function.
• The beta p.d.f. can accommodate a wide variety of shapes, including
  ► symmetric and positively or negatively skewed unimodal p.d.f.’s;
  ► \( J \)- and \( U \)-shaped p.d.f.’s;
  ► left- and right-triangular p.d.f.’s; and
  ► uniform p.d.f.’s.

• Some examples illustrating the range of distributional shapes achievable with the beta p.d.f. follow.
Positively and Negatively Skewed Unimodal Beta Densities
U-shaped Beta Densities
$J$-shaped and Left-triangular Beta Densities

$\theta_1 = 1, \theta_2 = 2$

$\theta_1 = 0.8, \theta_2 = 2$

$\theta_1 = 0.2, \theta_2 = 2$
Symmetric and Uniform Beta Densities

\[ \theta_1 = \theta_2 = 5 \]

\[ \theta_1 = \theta_2 = 2 \]

\[ \theta_1 = \theta_2 = 1 \]
• Cumulative distribution function (c.d.f.) of beta variate $X$,

$$F_X(x) = \Pr\{X \leq x\} = \int_{-\infty}^{x} f_X(w) \, dw \quad \text{for all real } x,$$

has no convenient analytical expression.
• Mean and variance of $X$ are given by

$$
\begin{align*}
\mu_X &= E[X] = \frac{\theta_1 b + \theta_2 a}{\theta_1 + \theta_2}, \\
\sigma^2_X &= E[(X - \mu_X)^2] = \frac{(b - a)^2 \theta_1 \theta_2}{(\theta_1 + \theta_2)^2 (\theta_1 + \theta_2 + 1)}.
\end{align*}
$$

(2)

• Provided $\theta_1, \theta_2 > 1$ so that the p.d.f. (1) is unimodal, the mode is given by

$$
m = \frac{(\theta_1 - 1)b + (\theta_2 - 1)a}{\theta_1 + \theta_2 - 2}.
$$

(3)

• The key distributional characteristics (2) and (3) are simple functions of $a$, $b$, $\theta_1$, and $\theta_2$; and this facilitates rapid input modeling.
Fitting Beta Distributions to Data or Subjective Information

Given the data set \( \{X_i : i = 1, \ldots, n\} \) of size \( n \), we let

\[
X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}
\]

denote the order statistics; and we compute the sample statistics

\[
\begin{align*}
\hat{a} &= 2X_{(1)} - X_{(2)}, \\
\hat{b} &= 2X_{(n)} - X_{(n-1)}, \\
\bar{X} &= n^{-1} \sum_{i=1}^{n} X_i, \\
S^2 &= (n - 1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.
\end{align*}
\]
• Moment-matching estimates of $\theta_1$, $\theta_2$ are computed from

$$
\hat{\theta}_1 = \frac{d_1^2(1 - d_1)}{d_2^2} - d_1, \quad \hat{\theta}_2 = \frac{d_1(1 - d_1)^2}{d_2^2} - (1 - d_1),
$$

where

$$
d_1 = \frac{\bar{X} - \hat{a}}{\hat{b} - \hat{a}} \quad \text{and} \quad d_2 = \frac{S}{\hat{b} - \hat{a}}.
$$
BetaFit (AbouRizk, Halpin, and Wilson 1994) is a Windows-based package for fitting the beta distribution to sample data by computing $\hat{a}$, $\hat{b}$, $\hat{\theta}_1$, and $\hat{\theta}_2$ using the following estimation methods:

- moment matching;
- feasibility-constrained moment matching (so that the feasibility conditions $\hat{a} < X_{(1)}$ and $X_{(n)} < \hat{b}$ are always satisfied);
- maximum likelihood (assuming $a$ and $b$ are known and thus are not estimated); and
- ordinary least squares (OLS) and diagonally weighted least squares (DWLS) estimation of the c.d.f.

BetaFit is in the public domain and is available on the Web via <www.ie.ncsu.edu/jwilson/page3>.
Application of BetaFit to a Sample of $n = 9,980$ Observations of End-to-End Chain Lengths (in Angströms) of Nafion, an Ionic Polymer Used As a “Smart Material,” Based on the Method of Moment Matching
Introduction to Modeling and Generating Probabilistic Input Processes for Simulation

Sample Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>20.537639</td>
<td>57.440907</td>
<td>2.267328</td>
<td>10.652630</td>
<td>9.481540</td>
<td>71.404654</td>
</tr>
<tr>
<td>Variance</td>
<td>57.440907</td>
<td>0.973585</td>
<td>3.792337</td>
<td>3.792337</td>
<td>9.481540</td>
<td>71.404654</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.267328</td>
<td>0.973585</td>
<td>3.792337</td>
<td>3.792337</td>
<td>9.481540</td>
<td>71.404654</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>10.652630</td>
<td>3.792337</td>
<td>3.792337</td>
<td>3.792337</td>
<td>9.481540</td>
<td>71.404654</td>
</tr>
<tr>
<td>Max</td>
<td>71.404654</td>
<td>71.404654</td>
<td>71.404654</td>
<td>71.404654</td>
<td>71.404654</td>
<td>71.404654</td>
</tr>
</tbody>
</table>

Fitted Beta distribution parameters

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>param-a</th>
<th>param-b</th>
<th>KS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>20.537639</td>
<td>57.440907</td>
<td>2.267328</td>
<td>10.652630</td>
<td>1.569526</td>
<td>7.221156</td>
<td>4.291073</td>
</tr>
<tr>
<td>Variance</td>
<td>57.440907</td>
<td>0.973585</td>
<td>3.792337</td>
<td>3.792337</td>
<td>1.569526</td>
<td>7.221156</td>
<td>4.291073</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.267328</td>
<td>0.973585</td>
<td>3.792337</td>
<td>3.792337</td>
<td>1.569526</td>
<td>7.221156</td>
<td>4.291073</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>10.652630</td>
<td>3.792337</td>
<td>3.792337</td>
<td>3.792337</td>
<td>1.569526</td>
<td>7.221156</td>
<td>4.291073</td>
</tr>
<tr>
<td>param-a</td>
<td>1.569526</td>
<td>1.569526</td>
<td>1.569526</td>
<td>1.569526</td>
<td>1.569526</td>
<td>1.569526</td>
<td>1.569526</td>
</tr>
</tbody>
</table>

Sample and Fitted CDF

Sample CDF

Fitted Beta CDF

Method: Matching Mean/Var & Sample End Points.
Result of Applying BetaFit to Nafion Data Set Using Maximum Likelihood Estimation
### Sample Statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>20.537639</td>
</tr>
<tr>
<td>Variance</td>
<td>57.440907</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.267328</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>10.652630</td>
</tr>
<tr>
<td>Min</td>
<td>3.481640</td>
</tr>
<tr>
<td>Max</td>
<td>71.404654</td>
</tr>
</tbody>
</table>

### Fitted Beta Distribution Parameters

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>20.914326</td>
</tr>
<tr>
<td>Variance</td>
<td>49.658112</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.865371</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.629277</td>
</tr>
<tr>
<td>Min</td>
<td>9.197191</td>
</tr>
<tr>
<td>Max</td>
<td>73.546794</td>
</tr>
<tr>
<td>param-a</td>
<td>2.079227</td>
</tr>
<tr>
<td>param-b</td>
<td>9.339726</td>
</tr>
<tr>
<td>KS</td>
<td>5.737301</td>
</tr>
</tbody>
</table>

---

**Sample And Fitted CDF**

- **Method:** Maximum Likelihood Estimates.
Result of Applying BetaFit to Nafion Data Set Using Ordinary Least Squares Estimation of the C.d.f.
**Sample Statistics** | **Fitted Beta distribution parameters**
---|---
Mean | 20.537639 | Mean | 20.292192 |
Variance | 57.440907 | Variance | 57.863376 |
Skewness | 2.267328 | Skewness | 0.959338 |
Kurtosis | 10.652630 | Kurtosis | 3.768994 |
Min | 3.481640 | Min | 9.015046 |
Max | 71.404654 | Max | 71.959110 |
param-a | 1.624904 | |
param-b | 7.444597 | |
KS | 1.449050 | |

**Sample And Fitted CDF**

Method: Ordinary Least Square Minimization.
• Rapid input modeling with subjective estimates $\hat{a}$, $\hat{m}$, and $\hat{b}$ of the minimum, mode, and maximum, respectively, of the target distribution:

$$\hat{\theta}_1 = \frac{d^2 + 3d + 4}{d^2 + 1} \quad \text{and} \quad \hat{\theta}_2 = \frac{4d^2 + 3d + 1}{d^2 + 1},$$

(4)

where

$$d = \frac{\hat{b} - \hat{m}}{\hat{m} - \hat{a}}.$$

The mode of the fitted beta distribution will differ from $\hat{m}$ by at most 4.4%; in practice the error is usually at most 1%.
VIBES (AbouRizk, Halpin, and Wilson 1991) is a Windows-based package for fitting the beta distribution to subjective estimates of:

1. the endpoints \( a \) and \( b \); and

2. any of the following combinations of distributional characteristics—
   - the mean \( \mu_X \) and the variance \( \sigma^2_X \),
   - the mean \( \mu_X \) and the mode \( m \),
   - the mode \( m \) and the variance \( \sigma^2_X \),
   - the mode \( m \) and an arbitrary quantile \( x_p = F_X^{-1}(p) \) for \( p \in (0, 1) \),
     or
   - two quantiles \( x_p \) and \( x_q \) for \( p, q \in (0, 1) \).
• Advantages of the beta distribution as an input-modeling tool:
  – sufficient flexibility to represent with reasonable accuracy a wide diversity of distributional shapes; and
  – convenient estimation of parameters from sample data or subjective information.

• Disadvantages of the beta distribution as an input-modeling tool:
  – difficult to explain; and
  – difficult to sample—some popular beta variate generators break down when \( \theta_1 > 10 \) or \( \theta_2 > 10 \).
Application of Beta Distributions to Pharmaceutical Manufacturing

Pearlswig (1995) developed a simulation of a proposed facility for manufacturing effervescent tablets.

- For each operation, he obtained three time estimates ($\hat{a}$, $\hat{m}$, and $\hat{b}$) from the process engineers.
- Extremely conservative estimates given for upper limits (so that $\hat{b} \gg \hat{m}$).
- With triangular distributions to model processing times, bottlenecks resulted in excessively low simulation estimates of annual production.
- Using (4), Pearlswig fitted beta distributions to all operation times; and then the simulation results conformed to production levels of similar plants elsewhere.
B. Johnson Translation System of Distributions

To fit a distribution to the continuous random variable \( X \), Johnson (1949a) proposed finding a “translation” of \( X \) to a standard normal random variable \( Z \) with mean 0 and variance 1 so that \( Z \sim N(0, 1) \).

The proposed normalizing translations have the general form

\[ Z = \gamma + \delta \cdot g \left( \frac{X - \xi}{\lambda} \right), \quad (5) \]

where \( \gamma \) and \( \delta \) are shape parameters, \( \lambda \) is a scale parameter, \( \xi \) is a location parameter, and the function \( g(\cdot) \) defines the four distribution families in the Johnson translation system,

\[
g(y) = \begin{cases} 
\ln(y), & \text{for } S_L \text{ (lognormal) family}, \\
\ln\left(y + \sqrt{y^2 + 1}\right), & \text{for } S_U \text{ (unbounded) family}, \\
\ln[y/(1 - y)], & \text{for } S_B \text{ (bounded) family}, \\
y, & \text{for } S_N \text{ (normal) family}.
\end{cases}
\]
• Johnson c.d.f.

If (5) is an exact normalizing translation of $X$ to a standard normal random variable, then the c.d.f. of $X$ is given by

$$F_X(x) = \Phi \left[ \gamma + \delta \cdot g \left( \frac{x - \xi}{\lambda} \right) \right] \quad \text{for all } x \in \mathcal{H},$$

where: $\Phi(z) = (2\pi)^{-1/2} \int_{-\infty}^{z} \exp \left( -\frac{1}{2} w^2 \right) dw$ is the standard normal c.d.f.; and the space of $X$ is

$$\mathcal{H} = \begin{cases} 
[\xi, +\infty), & \text{for } S_L \text{ (lognormal) family}, \\
(\infty, +\infty), & \text{for } S_U \text{ (unbounded) family}, \\
[\xi, \xi + \lambda], & \text{for } S_B \text{ (bounded) family}, \\
(\infty, +\infty), & \text{for } S_N \text{ (normal) family}.
\end{cases}$$
• Johnson p.d.f. is

\[
    f_X(x) = \frac{\delta}{\lambda (2\pi)^{1/2}} g' \left( \frac{x - \xi}{\lambda} \right) \exp \left\{ -\frac{1}{2} \left[ \gamma + \delta \cdot g \left( \frac{x - \xi}{\lambda} \right) \right]^2 \right\}
\]

for all \( x \in \mathcal{H} \), where

\[
    g'(y) = \begin{cases} 
        1/y, & \text{for } S_L \text{ (lognormal) family,} \\
        1/\sqrt{y^2 + 1}, & \text{for } S_U \text{ (unbounded) family,} \\
        1/[y/(1 - y)], & \text{for } S_B \text{ (bounded) family,} \\
        1, & \text{for } S_N \text{ (normal) family.}
    \end{cases}
\]

Following are examples illustrating all the distributional shapes in the Johnson system.
Symmetric Bimodal and Nearly Uniform Johnson $S_B$ Densities
Nearly $J$-shaped and Symmetric Unimodal Johnson $S_B$ Densities
\[ \gamma = 1, \delta = 1 \]
\[ \sqrt{\beta_1} = 0.728, \beta_2 = 2.91 \]
\[ \gamma = 1, \delta = 2 \]
\[ \sqrt{\beta_1} = 0.282, \beta_2 = 2.77 \]

Positively Skewed and Symmetric Unimodal Johnson $S_B$ Densities
\[
\begin{align*}
\gamma &= 0, \delta = 2 \\
\sqrt{\beta_1} &= 0, \beta_2 = 4.51
\end{align*}
\]

Case (a)

\[
\begin{align*}
\gamma &= 1, \delta = 2 \\
\sqrt{\beta_1} &= -0.872, \beta_2 = 5.59
\end{align*}
\]

Case (b)

\[
\begin{align*}
\gamma &= 1, \delta = 1 \\
\sqrt{\beta_1} &= -5.37, \beta_2 = 93.4
\end{align*}
\]

Case (c)

Symmetric and Negatively Skewed Johnson \( S_U \) Densities
Fitting Johnson Distributions to Sample Data

We select an estimation method and the desired translation function $g(\cdot)$ and then obtain estimates of $\gamma$, $\delta$, $\lambda$, and $\xi$.

The Johnson system has the flexibility to match—

(a) any feasible combination of values for the mean $\mu_X$, variance $\sigma_X^2$, skewness

$$\alpha_X = E\left[\frac{(X - \mu_X)^3}{\sigma_X^3}\right] \quad \text{often denoted by } \sqrt{\beta_1},$$

and kurtosis

$$\beta_X = E\left[\frac{(X - \mu_X)^4}{\sigma_X^4}\right] \quad \text{often denoted by } \beta_2;$$

or

(b) sample estimates of the moments $\mu_X$, $\sigma_X^2$, $\alpha_X$, and $\beta_X$. 


FITTR1 (Swain, Venkatraman, and Wilson 1988) is a software package for fitting Johnson distributions to sample data using the following estimation methods:

- OLS and DWLS estimation of the c.d.f.;
- minimum $L_1$ and $L_\infty$ norm estimation of the c.d.f.;
- moment matching; and
- percentile matching.
• VISIFIT (DeBrota et al. 1989b) is a Windows-based software package for fitting Johnson $S_B$ distributions to subjective information, possibly combined with sample data. The user must provide estimates of $a$, $b$, and any two of the following characteristics:
  
  ▶ the mode $m$;
  
  ▶ the mean $\mu_X$;
  
  ▶ the median $x_{0.5}$;
  
  ▶ arbitrary quantile(s) $x_p$ or $x_q$ for $p, q \in (0, 1)$;
  
  ▶ the width of the central 95% of the distribution; or
  
  ▶ the standard deviation $\sigma_X$.

Venkatraman, Swain and Wilson (1988), DeBrota et al. (1989b), FITTR1, and VISIFIT are available on the Web via

<www.ie.ncsu.edu/jwilson/more_info>.
Generating Johnson Variates by Inversion

[1] Generate \( Z \sim N(0, 1) \).

[2] Apply to \( Z \) the inverse translation

\[
X = \xi + \lambda \cdot g^{-1}\left( \frac{Z - \gamma}{\delta} \right),
\]

(6)

where for all real \( z \) we define the inverse translation function

\[
g^{-1}(z) = \begin{cases} 
  e^z, & \text{for } S_L \text{ (lognormal) family}, \\
  (e^z - e^{-z})/2, & \text{for } S_U \text{ (unbounded) family}, \\
  1/(1 + e^{-z}), & \text{for } S_B \text{ (bounded) family}, \\
  z, & \text{for } S_N \text{ (normal) family}.
\end{cases}
\]

(7)
Application of Johnson Distributions to Smart Materials Research

• Matthews et al. (2005) and Weiland et al. (2005) use a multiscale modeling approach to predict material stiffness of a certain class of smart materials called ionic polymers.

• Material stiffness depends on effective length of the polymer chains comprising the material.

• In a case study of the ionic polymer Nafion, Matthews et al. (2005) develop a simulation of polymer-chain conformation on a nanoscopic level so as to generate a large number of end-to-end chain lengths.

• The chain-length p.d.f. is estimated and used as input to a macroscopic-level mathematical model to predict material stiffness.
Johnson $S_U$ C.d.f. Fitted to $n = 9,980$ Nafion Chain Lengths Using DWLS Estimation
Johnson $S_U$ P.d.f Fitted to $n = 9,980$ Nafion Chain Lengths Using DWLS Estimation
• Matthews et al. (2005) and Weiland et al. (2005) obtain more accurate and intuitively appealing fits to Nafion chain-length data with Johnson p.d.f.’s than with other distributions.
  – Material stiffness is computed from the second derivative $f''_X(x)$ of the fitted p.d.f.
  – There is a relatively simple relationship between the Johnson parameters and material stiffness.
Application of Johnson Distributions to Healthcare

• To model arrival patterns of patients who have scheduled appointments at community healthcare clinics in San Diego, Alexopoulos et al. (2005) estimate the distribution of patient tardiness—that is, deviation from the scheduled appointment time.

• Alexopoulos et al. (2005) perform an exhaustive analysis of 18 continuous distributions, concluding that the $S_U$ distribution provided superior fits to the available data.
C. Bézier Distribution Family

Definition of Bézier Curves

- A Bézier curve is often used to approximate a smooth function on a bounded interval by forcing the Bézier curve to pass in the vicinity of selected control points

\[ \{ \mathbf{p}_i \equiv (x_i, z_i)^T : i = 0, 1, \ldots, n \} \]

in two-dimensional Euclidean space.
A Bézier curve of degree $n$ with control points $\{p_0, p_1, \ldots, p_n\}$ is given parametrically by

\[
P(t) = \begin{bmatrix} P_x(t; n, x), & P_z(t; n, z) \end{bmatrix}^T = \sum_{i=0}^{n} B_{n,i}(t) p_i \quad \text{for} \ t \in [0, 1],
\]

where $x \equiv (x_0, x_1, \ldots, x_n)^T$ and $z \equiv (z_0, z_1, \ldots, z_n)^T$, and where the blending function, $B_{n,i}(t) \equiv \frac{n!}{i!(n-i)!} t^i(1-t)^{n-i}$ for $t \in [0, 1]$, is the $i$th Bernstein polynomial for $i = 0, 1, \ldots, n$. 


Bézier Distribution and Density Functions

• If $X$ is a continuous random variable on $[a, b]$ with c.d.f. $F_X(\cdot)$ and p.d.f. $f_X(\cdot)$, then we can approximate $F_X(\cdot)$ arbitrarily closely using a Bézier curve of the form (8) by taking a sufficient number $(n + 1)$ of control points with appropriate coordinates

$$p_i = (x_i, z_i)^T$$

for the $i$th control point, where $i = 0, \ldots, n$. 
• If $X$ is Bézier, then the c.d.f. of $X$ is given parametrically by

$$
P(t) = \left[ P_x(t; n, x), P_z(t; n, z) \right]^T
$$

\[
= \{ x(t), F_X[x(t)] \}^T \text{ for } t \in [0, 1],
\]

where

\[
x(t) = P_x(t; n, x) = \sum_{i=0}^{n} B_{n,i}(t)x_i,
\]

\[
F_X[x(t)] = P_z(t; n, z) = \sum_{i=0}^{n} B_{n,i}(t)z_i
\]

for $t \in [0, 1]$. (11)

For a detailed discussion of Bézier distributions, see


<ftp.ncsu.edu/pub/eos/pub/jwilson/wagner96iie.pdf>
• If $X$ is Bézier with c.d.f. $F_X(\cdot)$ given by (10), then the p.d.f. $f_X(x)$ is

\[
P^*(t) = \left[ P^*_x(t; n, x), P^*_z(t; n, x, z) \right]^T
\]

\[
= \left\{ x(t), f_X[x(t)] \right\}^T \text{ for } t \in [0, 1],
\]

where $x(t) = P^*_x(t; n, x) = P_x(t; n, x)$ as in (11) and

\[
f_X[x(t)] = P^*_z(t; n, x, z)
\]

\[
= \frac{P_z(t; n - 1, \Delta z)}{P_x(t; n - 1, \Delta x)} = \frac{\sum_{i=0}^{n-1} B_{n-1,i}(t) \Delta z_i}{\sum_{i=0}^{n-1} B_{n-1,i}(t) \Delta x_i},
\]

where

\[
\Delta x \equiv (\Delta x_0, \ldots, \Delta x_{n-1})^T \text{ and } \Delta z \equiv (\Delta z_0, \ldots, \Delta z_{n-1})^T,
\]

with

\[
\Delta x_i \equiv x_{i+1} - x_i \text{ and } \Delta z_i \equiv z_{i+1} - z_i \text{ for } i = 0, 1, \ldots, n - 1.
\]
Generating Bézier Variates by Inversion

[1] Generate a random number $U \sim \text{Uniform}[0, 1]$.

[2] Find $t_U \in [0, 1]$ such that

$$F_X[x(t_U)] = \sum_{i=0}^{n} B_{n,i}(t_U)z_i = U. \quad (12)$$

[3] Deliver the variate

$$X = x(t_U) = \sum_{i=0}^{n} B_{n,i}(t_U)x_i.$$

Codes to implement this approach are available on the Web via

<www.ie.ncsu.edu/jwilson/page3>
Using PRIME to Model Bézier Distributions

- **PRIME** (Wagner and Wilson 1996a) is a Windows-based system for fitting Bézier distributions to data or subjective information.
- **PRIME** is available on the previously mentioned Web site.
- Control points appear as indexed black squares that can be manipulated with the mouse and keyboard.
  - Each control point exerts on the c.d.f. a “magnetic” attraction whose strength is given by the associated Bernstein polynomial (9).
  - Moving a control point causes the displayed c.d.f. to be updated (nearly) instantaneously.
**PRIME Windows Showing the Bézier C.d.f. (Left Panel) with Its Control Points and the P.d.f. (Right Panel)**
PRIME includes the following methods for fitting Bézier distributions to sample data:

- OLS estimation of the c.d.f.;
- minimum $L_1$ and $L_\infty$ norm estimation of the c.d.f.;
- maximum likelihood estimation (assuming $a$ and $b$ are known);
- moment matching; and
- percentile matching.

Bézier Distribution Fitted to $n = 9,980$ Nafion Chain Lengths Using OLS Estimation of the C.d.f.
Advantages of the Bézier distribution family:

- It is extremely flexible and can represent a wide diversity of distributional shapes, including multiple modes and mixed distributions.

- If data are available, then the likelihood ratio test of Wagner and Wilson (1996b) can be used with any of the available estimation methods to find automatically both the number and location of the control points.

- In the absence of data, PRIME can be used to determine the conceptualized distribution based on known quantitative or qualitative information.

- As the number \((n + 1)\) of control points increases, so does the flexibility in fitting Bézier distributions.
III. Multivariate Input Models

A. Multivariate Johnson Translation System

To model the distribution of $\mathbf{X} = (X_1, X_2)^T$, Johnson (1949b) proposed

- matching the first four moments (i.e., the mean, variance, skewness, and kurtosis) for each component $X_i$ separately ($i = 1, 2$); and then

- attempting to match approximately $\text{Corr}(X_1, X_2)$, the correlation between $X_1$ and $X_2$.


<www.informs-cs.org/wsc97papers/0047.PDF>.
To model the continuous random vector \( \mathbf{X} = (X_1, \ldots, X_\nu)^T \) with \( \nu \) components, we seek a normalizing translation such that

\[
\mathbf{Z} = \mathbf{y} + \delta g \left[ \lambda^{-1} (\mathbf{X} - \xi) \right] \sim N_\nu(0_\nu, \Sigma),
\]

so that \( \mathbf{Z} = (Z_1, \ldots, Z_\nu)^T \) has a \( \nu \)-dimensional standard normal distribution with null mean vector \( 0_\nu \) and covariance matrix \( \Sigma \) whose \( (i, j) \) entry is \( \text{Corr}(Z_i, Z_j) \).
Based on a random sample \( \{ \mathbf{X}_\ell = (X_{1,\ell}, X_{2,\ell}, \ldots, X_{\nu,\ell})^T : \ell = 1, \ldots, n \} \), the vector-valued normalizing translation (13) is accomplished as follows:

[1] Identify the transformation

\[
\mathbf{g} \left[ (y_1, \ldots, y_\nu)^T \right] \equiv \left[ g_1(y_1), \ldots, g_\nu(y_\nu) \right]^T
\]

so that for the \( i \)th component \( X_i \) of \( \mathbf{X} \) (where \( i = 1, \ldots, \nu \)), the translation function \( g_i(\cdot) \) is selected to match the skewness and kurtosis computed from \( \{ X_{i,\ell} : \ell = 1, \ldots, n \} \).
[2] Estimate the matrices of shape parameters,

\[ \gamma \equiv (\gamma_1, \ldots, \gamma_v)^T, \]

\[ \delta \equiv \text{diag}(\delta_1, \ldots, \delta_v) = \begin{bmatrix}
\delta_1 & 0 & \cdots & 0 \\
0 & \delta_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \delta_v
\end{bmatrix}, \]

and the matrices of the respective location and scale parameters,

\[ \xi \equiv (\xi_1, \ldots, \xi_v)^T, \]

\[ \lambda \equiv \text{diag}(\lambda_1, \ldots, \lambda_v), \]

so that for the each component \( X_i \) of \( X \), not only the estimated function \( \hat{g}_i(\cdot) \) but also the parameter estimates \( \hat{\gamma}_i, \hat{\delta}_i, \hat{\lambda}_i, \) and \( \hat{\xi}_i \) are determined so as to match the sample mean, variance, skewness, and kurtosis of the random sample \( \{X_{i,\ell} : \ell = 1, \ldots, n\} \).
[3] Estimate correlation matrix $\Sigma$ by

(a) inserting each sample value $\{X_\ell : \ell = 1, \ldots, n\}$ into the estimated normalizing translation (13) to obtain the corresponding sample $\{Z_\ell : \ell = 1, \ldots, n\}$ of estimated standard normal random vectors,

$$Z_\ell = \gamma + \delta g \left[ \lambda^{-1} (X_\ell - \xi) \right] \quad \text{for} \quad \ell = 1, \ldots, n;$$

(b) computing the sample correlation matrix of the $\{Z_j\}$ as the approximate moment-matching estimator of $\Sigma$,

$$\hat{\Sigma}_{ij} = \frac{\sum_{\ell=1}^{n} (Z_{i,\ell} - \overline{Z}_i)(Z_{j,\ell} - \overline{Z}_j)}{\sqrt{\sum_{\ell=1}^{n} (Z_{i,\ell} - \overline{Z}_i)^2} \sqrt{\sum_{\ell=1}^{n} (Z_{j,\ell} - \overline{Z}_j)^2}} \quad \text{for} \quad i, j = 1, \ldots, \nu$$
Generating Johnson Random Vectors

- To generate a Johnson random vector $\mathbf{X}$ with translation function $g(\cdot)$ and parameters $\nu, \delta, \lambda, \xi,$ and $\Sigma$, we must first compute the “square root” matrix $\mathbf{Q}$ of $\Sigma$ so that we have the Cholesky decomposition

$$\Sigma = \mathbf{Q} \mathbf{Q}^T.$$
• Then $X$ is generated as follows:

[1] Take

$$U = (U_1, \ldots, U_\nu)^T,$$

where $\{U_\ell : \ell = 1, \ldots, \nu\} \overset{\text{i.i.d.}}{\sim} N(0, 1)$

so that

$$Z = QU \sim N_\nu(0, \Sigma).$$

[2] Apply to $Z$ the inverse translation

$$X = \xi + \lambda g^{-1} \left[ \delta^{-1}(Z - \nu) \right],$$

where

$$g^{-1} \left[ (z_1, \ldots, z_\nu)^T \right] \equiv [g_1^{-1}(z_1), \ldots, g_\nu^{-1}(z_\nu)]^T.$$
Limitations of the Procedure Based on (13)–(14)

• This method will generate random vectors with exactly the same marginal moments as the original sample data \( \{X_\ell : \ell = 1, \ldots, n\} \) from which the sample estimators \( \hat{\gamma}, \hat{\delta}, \hat{\lambda}, \hat{\xi}, \) and \( \hat{\Sigma} \) are computed (at least to the limits of machine accuracy).

• However, if some of the empirical marginal distributions of the original sample data (or the corresponding underlying theoretical marginals) possess marked skewness, then the correlation matrix of the fitted multivariate Johnson distribution will not match the sample correlation matrix of the original data set.
Application of Multivariate Johnson Distributions to Welfare Policy Analysis

- McDaniel, Sullivan, and Wilson (1988) develop ADSSIM, a simulation of the Institutional Care and Community Care programs of the Texas Department of Human Services (TDHS), to estimate the effects of changes in program eligibility criteria on client loads, costs, etc., for new and existing services forecast over a multiyear planning horizon.
• Trivariate Johnson $S_{B_{BB}}$ distributions were fitted to a random sample from the joint distribution of the following quantities for each individual in each target population defined by a combination of age cohort and service region:
  ► monthly income;
  ► dollar value of countable resources; and
  ► functional disability score.

• ADSSIM was used by TDHS to analyze budget requests and to redesign sections of its biennial survey so that the modeling approach could be applied to all Community Care programs.
B. Matching Exactly a Given Mean Vector and Covariance Matrix

Alternative Multivariate Distributions with Univariate Johnson Marginals

Sometimes we seek to fit a multivariate distribution having a given mean vector $\mu_X$ and covariance matrix $\Sigma_X$ such that for $i = 1, \ldots, \nu$, the $i$th marginal is a univariate Johnson distribution with the same first four moments as $X_i$, the $i$th component of the target random variable $X$.

We present the method of Stanfield et al. (1996); see


• Suppose that $\mathbf{X}$ has correlation matrix

$$
C_{\mathbf{X}} = \begin{bmatrix} \text{Corr}(X_i, X_j) \end{bmatrix}
$$

with lower-triangular “square root” matrix

$$
\Theta_{\mathbf{X}} \equiv \begin{bmatrix} \theta_{ij} \end{bmatrix} = C_{\mathbf{X}}^{1/2}
$$

(so that we have the Cholesky decomposition $C_{\mathbf{X}} = \Theta_{\mathbf{X}} \Theta_{\mathbf{X}}^T$) together with the matrix of standard deviations

$$
\sigma_{\mathbf{X}} \equiv \text{diag}[\text{Var}^{1/2}(X_1), \ldots, \text{Var}^{1/2}(X_\nu)].
$$
• If $\mathbf{Y} = (Y_1, \ldots, Y_\nu)^T$ consists of independent standardized Johnson variates so that $Y_i$ has a Johnson distribution with mean 0 and variance 1, then

$$\mathbf{W} \equiv \mu_\mathbf{X} + \sigma_\mathbf{X} \Theta_\mathbf{X} \mathbf{Y}$$

has the mean vector $\mu_\mathbf{X}$ and covariance matrix $\Sigma_\mathbf{X}$.

• For $i = 1, \ldots, \nu$, we assign $g_i(\cdot), \gamma_i, \delta_i, \lambda_i, \text{ and } \xi_i$ for $Y_i$ so that the $i$th component $W_i$ of $\mathbf{W}$ has the same skewness $\alpha_{X_i}$ and kurtosis $\beta_{X_i}$ as $X_i$ has.
• Let $\alpha_X$ and $\beta_X$ be $\nu \times 1$ vectors whose $i$th elements are the skewness $\alpha_{X_i}$ and kurtosis $\beta_{X_i}$ of the random variable $X_i$, respectively, for $i = 1, \ldots, \nu$. Similarly, let $\alpha_Y$ and $\beta_Y$ denote the skewness and kurtosis vectors for $Y$.

• Define the auxiliary matrix

$$\Theta^{(k)}_X \equiv [\theta^{k}_{ij}] \text{ for } k = 3, 4$$

together with the auxiliary vector

$$\Psi_X \equiv (\psi_1, \ldots, \psi_\nu)^T,$$

where

$$\psi_i = 6 \sum_{j=1}^{\nu} \sum_{\ell=j+1}^{\nu} \theta^2_{ij} \theta^2_{i\ell} \text{ for } i = 1, \ldots, \nu.$$
• If $X$ were generated according to $X = \mu_X + \sigma_X \Theta_X Y$, then the skewness and kurtosis vectors $\alpha_X$ and $\beta_X$ for $X$ would be given by

\[
\begin{align*}
\alpha_X &= \Theta_X^{(3)} \alpha_Y \\
\beta_X &= \Theta_X^{(4)} \beta_Y + \Psi_X
\end{align*}
\]

(16)

• Given estimates or exact values of $\alpha_X$ and $\beta_X$ for the target random vector $X$, we solve the moment-matching equations (16) to yield the skewness and kurtosis vectors $\alpha_Y$ and $\beta_Y$ required for $Y$:

\[
\begin{align*}
\alpha_Y &= \left[ \Theta_X^{(3)} \right]^{-1} \alpha_X \\
\beta_Y &= \left[ \Theta_X^{(4)} \right]^{-1} (\beta_X - \Psi_X)
\end{align*}
\]

(17)
• If we compute $\Theta_X$, $\Theta_X^{(3)}$, $\Theta_X^{(4)}$, and $\Psi_X$ from known or estimated matrices $\mu_X$, $\sigma_X$, $C_X$, $\alpha_X$, and $\beta_X$, then $W$ given by (15) has the same mean vector and covariance matrix as $X$ has; moreover for $i = 1, \ldots \nu$, the skewness of $W_i$ matches the skewness $\alpha_{X_i}$ of $X_i$.

• For $i = 1, \ldots \nu$, the kurtosis of $W_i$ matches the kurtosis $\beta_{X_i}$ of $X_i$ with some exceptions detailed below.

• This computing effort consists mainly of: (i) determining $4\nu$ marginal moments and $\nu(\nu - 1)/2$ correlation values; (ii) inverting two $\nu \times \nu$ matrices; (iii) computing $\psi_i$ for $i = 1, \ldots, \nu$; and (iv) fitting $\nu$ univariate Johnson distributions by the method of moments.
Generating Random Vectors with Given Mean and Covariance Matrix

[1] Generate $Z = (Z_1, \ldots, Z_\nu)^T$, where $\{Z_\ell : \ell = 1, \ldots, \nu\}$ i.i.d. $\sim N(0, 1)$.

[2] Deliver

$$X = \mu_X + \sigma_X \Theta_X \{\xi + \lambda g^{-1}[\delta^{-1}(Z - \gamma)]\}, \quad (18)$$

where $g(\cdot)$, $\gamma$, $\delta$, $\lambda$, and $\xi$ were computed one component at a time from (17).
Limitations of the Procedure Based on (15)–(18)

• If a standardized Johnson random variable $Y_i$ has skewness $\alpha_{Y_i}$ and kurtosis $\beta_{Y_i}$, then we must have

$$\beta_{Y_i} \geq \alpha_{Y_i}^2 + 1 \text{ for } i = 1, \ldots, \nu.$$  \hfill (19)

• The solution (17) to the moment-matching equations (16) is not guaranteed to satisfy (19) even though we have $\beta_{X_i} \geq \alpha_{X_i}^2 + 1$ for each component $X_i$ of the target random vector $\mathbf{X}$.

• The family of multivariate distributions formulated in equations (15)–(17) may not be able to match exactly all available information about the marginal kurtosis of the components of $\mathbf{X}$. 
Application of Procedure Based on (15)–(18) to Remanufacturing

• Stanfield, Wilson, and King (2004) used the approach of (15)–(18) to develop a simulation model of the operation of a NAVAIR depot, an aircraft remanufacturing facility.

• Operation times for a workpiece at different workstations are correlated, with marginal distributions having a variety of shapes and often exhibiting substantial skewness.

• In practical simulations of product-reuse facilities, there are frequently hundreds of operations so that $\nu \geq 100$; and then the approach of (15)–(18) has substantial computational advantages.
IV. Time-Dependent Arrival Processes

• Many simulation applications require high-fidelity input models of arrival processes with arrival rates that depend strongly on time.

• Nonhomogeneous Poisson processes (NHPPs) have been used successfully to model complex time-dependent arrival processes in many applications.

• An NHPP \( \{N(t) : t \geq 0\} \) is a counting process such that
  
  \begin{itemize}
    \item \( N(t) \) is the number of arrivals in the time interval \( (0, t] \);
    \item \( \lambda(t) \) is the instantaneous arrival rate at time \( t \), and \( \lambda(t) \) satisfies the Poisson postulates; and
    \item the (cumulative) mean-value function is given by
  \end{itemize}

  \[ \mu(t) \equiv \mathbb{E}[N(t)] = \int_0^t \lambda(z) \, dz \quad \text{for all} \quad t \geq 0. \]  

(20)
• We discuss the nonparametric approach of Leemis (1991, 2000, 2004) for modeling and simulation of NHPPs; see


The context is a recent application to modeling and simulating unscheduled patient arrivals to a community healthcare clinic (Alexopoulos et al. 2005).

Suppose we have a time interval \((0, S]\) over which we observe several independent replications (realizations) of a stream of unscheduled patient arrivals constituting an NHPP with arrival rate \(\lambda(t)\) for \(t \in (0, S]\).

For example, \((0, S]\) might represent the time period on each weekday during which unscheduled patients may walk into a clinic—say, between 9 A.M. and 5 P.M. so that \(S = 480\) minutes.
• Suppose $k$ realizations of the arrival stream over $(0, S]$ have been recorded so that we have
  
  - $n_i$ patient arrivals in the $i$th realization for $i = 1, 2, \ldots, k$; and
  
  - $n = \sum_{i=1}^{k} n_i$ patient arrivals accumulated over all realizations.
• Let $\{t(i) : i = 1, \ldots, n\}$ denote the overall set of arrival times for all unscheduled patients expressed as an offset from the beginning of $(0, S]$ and then sorted in increasing order.

For example, if we observed $n = 250$ patient arrivals over $k = 5$ days, each with an observation interval of length $S = 480$ minutes, then

- $t(1) = 2.5$ minutes means that over all 5 days, the earliest arrival occurred 2.5 minutes after the clinic opened on one of those days; and

- $t(2) = 4.73$ minutes means that the second-earliest arrival occurred 4.73 minutes after the clinic opened on one of those days.

- $t(n) = 478.5$ minutes means that the latest arrival occurred 478.5 minutes after the clinic opened on one of those days.
• We estimate the mean-value function $\mu(t)$ as follows.
  ▶ We take $t(0) \equiv 0$ and $t(n+1) \equiv S$.
  ▶ For $t(i) < t \leq t(i+1)$ and $i = 0, 1, \ldots, n$, we take
    $$\hat{\mu}(t) = \frac{in}{(n+1)k} + \left\{ \frac{n[t - t(i)]}{(n+1)k[t(i+1) - t(i)]} \right\}.$$  (21)

• Equation (21) provides a basis for modeling and simulating unscheduled patient-arrival streams when the arrival rate exhibits a strong dependence on time.
Nonparametric Estimator of Mean Value Function
Goodness-of-fit Testing for the Fitted Mean-Value Function

- In addition to the realizations of the target arrival process that were used to compute the estimated mean-value function $\hat{\mu}(t)$, suppose we observe one additional realization

\[ \{ A'_i : i = 1, 2, \ldots, n' \} \]

independently of the previously observed realizations, with the $i$th patient arriving at time $A'_i$ for $i = 1, \ldots, n'$.

- If the target arrival stream is an NHPP with mean-value function $\mu(t)$ for $t \in (0, S]$, then the transformed arrival times

\[ \{ B'_i = \mu(A'_i) : i = 1, 2, \ldots, n' \} \]

constitute a homogeneous Poisson process with an arrival rate of 1.
• If the target arrival stream is an NHPP with mean-value function $\mu(t)$ for $t \in (0, S]$, then the corresponding transformed interarrival times

$$\{X'_i = B'_i - B'_{i-1} : i = 1, 2, \ldots, n'\}$$

(with $B'_0 \equiv 0$) constitute a random sample from an exponential distribution with a mean of 1.

• To test the adequacy of the fitted mean-value function $\hat{\mu}(t)$ as an approximation to $\mu(t)$, apply the Kolmogorov-Smirnov test to the data set

$$\{X''_i = \hat{\mu}(A'_i) - \hat{\mu}(A'_{i-1}) : i = 1, 2, \ldots n'\}$$

(with $A'_0 \equiv 0$), where the hypothesized c.d.f. in the goodness-of-fit test is

$$F_{X''_i}(x) = 1 - e^{-x} \quad \text{for all} \quad x \geq 0.$$
Generating Realizations of the Fitted NHPP

[1] Set $i \leftarrow 1$ and $N \leftarrow 0$.
[2] Generate $U_i \sim \text{Uniform}(0, 1)$.
[3] Set $B_i \leftarrow -\ln(1 - U_i)$.
[4] While $B_i < n/k$ do

Begin

Set $m \leftarrow \left\lfloor \frac{(n + 1)kB_i}{n} \right\rfloor$;

Set $A_i \leftarrow t(m) + \left\{t(m+1) - t(m)\right\}\left\{\left(\frac{n + 1)kB_i}{n}\right) - m\right\}$;

Set $N \leftarrow N + 1$; Set $i \leftarrow i + 1$;

Generate $U_i \sim \text{Uniform}(0, 1)$;

Set $B_i \leftarrow B_{i-1} - \ln(1 - U_i)$.

End

NHPP Simulation Procedure of Leemis (1991)
Advantages of Leemis’s Nonparametric Approach to Modeling and Simulation of NHPPs

- It does not require the assumption of any particular form for arrival rate \( \lambda(t) \) as a function of \( t \).

- It provides a strongly consistent estimator of mean-value function—that is,\[
\lim_{k \to \infty} \hat{\mu}(t) = \mu(t) \quad \text{for all} \quad t \in (0, S] \quad \text{with probability 1.}
\]

- The simulation algorithm given above, which is based on inversion of \( \hat{\mu}(t) \) so that
\[
A_i = \hat{\mu}^{-1}(B_i) \quad \text{for} \quad i = 1, \ldots, N,
\]
is also asymptotically valid as \( k \to \infty \).
Application to Organ Transplantation Policy Analysis

• The United Network for Organ Sharing (UNOS) applied a simplified variant of this approach in the development and use of the UNOS Liver Allocation Model (ULAM) for analyzing the cadaveric liver-allocation system in the U.S. (see Harper et al. 2000).

• ULAM incorporated models of

  (a) the streams of liver-transplant patients arriving at 115 transplant centers, and

  (b) the streams of donated organs arriving at 61 organ procurement organizations in the United States.

  Virtually all these arrival streams exhibited long-term trends as well as strong dependencies on the time of day, the day of the week, and the geographic location of the arrival stream.
Handling Arrival Processes Having Trends and Cyclic Effects

- Kuhl, Sumant, and Wilson (2005) develop a “semiparametric” method for modeling and simulating arrival processes that may exhibit a long-term trend or nested periodic phenomena (such as daily and weekly cycles), where the latter effects might not necessarily possess the symmetry of sinusoidal oscillations.

- See


  <ftp.ncsu.edu/pub/eos/pub/jwilson/kuhl04joc.pdf>
Fitted Rate Function over 100 Replications of a Test Process with One Cyclic Rate Component and Long-term Trend
Fitted Mean-Value Function over 100 Replications of a Test Process with One Cyclic Rate Component and Long-term Trend
Web-based Input Modeling Software

<www.rit.edu/~kuhl1/simulation>
V. Conclusions and Recommendations

• The common thread running through this tutorial is the focus on robust input models that are
  ▶ computationally tractable and
  ▶ sufficiently flexible to represent adequately many of the probabilistic phenomena that arise in many applications of discrete-event stochastic simulation.

• Notably absent is a discussion of Bayesian input-modeling techniques—a topic that will receive increasing attention in the future.

• Additional material on input modeling is available via
  
  <www.ie.ncsu.edu/jwilson/more_info>.