## Ch 6 Selecting input probability distribution

<table>
<thead>
<tr>
<th>Type of system</th>
<th>Sources of randomness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>Processing times, machine times to failure, machine repair times</td>
</tr>
<tr>
<td>Defense-related</td>
<td>Arrival times and payloads of missiles or airplanes, outcome of an engagement, miss distances for munitions</td>
</tr>
<tr>
<td>Communications</td>
<td>Interarrival times of messages, message types, message lengths</td>
</tr>
<tr>
<td>Transportation</td>
<td>Ship-loading times, interarrival times of customers to a subway</td>
</tr>
</tbody>
</table>
A single-server queueing system has exponential interarrival times with a mean of 1 minute. Suppose that 200 service times are available from the system, but their underlying probability distribution is unknown.

<table>
<thead>
<tr>
<th>Service-time distribution</th>
<th>Average delay in queue</th>
<th>Average number in queue</th>
<th>Proportion of delays $\geq 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>6.71</td>
<td>6.78</td>
<td>0.064</td>
</tr>
<tr>
<td>Gamma</td>
<td>4.54</td>
<td>4.6</td>
<td>0.019</td>
</tr>
<tr>
<td>Weibull</td>
<td>4.36</td>
<td>4.41</td>
<td>0.013</td>
</tr>
<tr>
<td>Logormal</td>
<td>7.19</td>
<td>7.3</td>
<td>0.078</td>
</tr>
<tr>
<td>Normal</td>
<td>6.04</td>
<td>6.13</td>
<td>0.045</td>
</tr>
</tbody>
</table>
Specify a distribution

- Approach 1: the data values themselves are used directly in the simulation
- Approach 2: the data values themselves are used an empirical distribution function
- Approach 3: standard techniques of statistical inference are used to “fit” a theoretical distribution form to the data and to perform hypothesis tests to determine the goodness of fit
• Drawback of approach 1
  • The simulation can only reproduce what has happened historically
  • There is seldom enough data to make all the desired simulation runs
• Approach 2 avoids these shortcomings (for continuous data)
If a theoretical distribution can be found that fits the observed data reasonably well, then this will generally be preferable to using an empirical distribution for the following reasons:

- An empirical distribution function may have certain “irregularities,” particularly if only a small number of data values are available.
- If empirical distributions are used in the usual way, it is not possible to generate values outside the range of the observed data in the simulation.
- There may be a compelling physical reason in some situations for using a certain theoretical distribution form as a model for a particular input random variable.
- A theoretical distribution is a compact way of representing a set of data values.
- A theoretical distribution is easier to change.
6.2 Useful probability distribution

- The purpose is to discuss a variety of distribution that have been found to be useful in simulation modeling and to provide a unified listing of relevant properties of these distributions.
  - Continuous distribution (CD)
  - Discrete distribution (DD)
6.2.1 Parameterization of CD

- Three types of parameters
  - **Location** parameter \( \gamma \) specifies an abscissa (x axis) location point of a distribution’s range of values. (usually \( \gamma \) is the midpoint or lower endpoint of the distribution’s range)
    - \( Y = X + \gamma \)
  - **Scale** parameter \( \beta \) determines the scale (or unit) of measurement of the values in the range of the distribution.
    - \( Y = \beta X \)
  - **Shape** parameter \( \alpha \) determines, distinct from location and scale, the basic form or shape of a distribution within the general family of distributions of interest.
6.2.2 Continuous distribution

- Uniform
- Exponential
- Gamma
- Weibull
- Normal
- Longnormal
- Beta
- Pearson type V, VI
- Log-logistic
- Johnson $S_B$, $S_U$
- Triangular
Uniform

Possible applications: Used as a “first” model for a quantity that is felt to be randomly varying between \(a\) and \(b\) but about which little else is known. The \(U(0, 1)\) distribution is essential in generating random values from all other distributions (see Chaps. 7 and 8).

Density (See Fig. 6.5)

\[
f(x) = \begin{cases} 
  \frac{1}{b-a} & \text{if } a \leq x \leq b \\
  0 & \text{otherwise}
\end{cases}
\]

Distribution

\[
F(x) = \begin{cases} 
  0 & \text{if } x < a \\
  \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\
  1 & \text{if } b < x
\end{cases}
\]

Parameters: \(a\) and \(b\) real numbers with \(a < b\), \(a\) is a location parameter, \(b - a\) is a scale parameter

Range: \([a, b]\)

Mean

\[
\frac{a + b}{2}
\]

Variance

\[
\frac{(b - a)^2}{12}
\]

Mode: Does not uniquely exist

MLE: \(d = \min \{X_i\}, \ b = \max \{X_i\}\)

Comments:
1. The \(U(0, 1)\) distribution is a special case of the beta distribution (when \(a = 1\) or \(b = 1\)).
2. If \(X \sim U(0, 1)\) and \([x, x + \Delta x]\) is a subinterval of \([0, 1]\) with \(\Delta x \geq 0\),

\[
f(X \in [x, x + \Delta x]) = \int_x^{x+\Delta x} dy = (x + \Delta x) - x = \Delta x
\]

which justifies the name “uniform.”

\[1/(b-a)\]

FIGURE 6.5

\(U(a, b)\) density function.
The exponential distribution is a special case of both the gamma and Weibull distributions (for shape parameter $\alpha = 1$ and scale parameter $\beta$ in both cases).

1. If $X_1, X_2, \ldots, X_n$ are independent $\text{expo}(\beta)$ random variables, then $X_1 + X_2 + \cdots + X_n \sim \text{gamma}(m, \beta)$, also called the $m$-Erlang distribution.

2. The exponential distribution is the only continuous distribution with the memoryless property (see Prob. 4.26).
Gamma Distribution

\[ f(x) = \begin{cases} \frac{\beta^\alpha x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \]

where \( \Gamma(\alpha) \) is the gamma function, defined by \( \Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt \) for any real number \( \alpha > 0 \). Some properties of the gamma function: \( \Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \) for any \( \alpha > 0 \), \( \Gamma(k + 1) = k! \) for any nonnegative integer \( k \), \( \Gamma(k + \frac{1}{2}) = \sqrt{\pi} \cdot 1 \cdot 3 \cdot 5 \cdots (2k - 1)/2^k \) for any positive integer \( k \), \( \Gamma(1/2) = \sqrt{\pi} \).

Possible Applications
- Time to complete some task, e.g., customer service or machine repair

Comments
- The \( \text{expo}(\beta) \) and \( \text{gamma}(1, \beta) \) distributions are the same.
- For a positive integer \( m \), the \( \text{gamma}(m, \beta) \) distribution is called the \( m \)-\text{Erlang}(\beta) distribution.
- The chi-square distribution with \( k \) df is the same as the \( \text{gamma}(k/2, 2) \) distribution.
- If \( X_1, X_2, \ldots, X_n \) are independent random variables with \( X_i \sim \text{gamma}(\alpha_i, \beta) \), then \( X_1 + X_2 + \cdots + X_n \sim \text{gamma}(\alpha_1 + \alpha_2 + \cdots + \alpha_n, \beta) \).
- If \( X_1 \) and \( X_2 \) are independent random variables with \( X_1 \sim \text{gamma}(\alpha_1, \beta) \), then \( X_1/(X_1 + X_2) \sim \beta(\alpha_1, \alpha_2) \).
- \( X \sim \text{gamma}(\alpha, \beta) \) if and only if \( Y = 1/X \) has a Pearson type V distribution with shape and scale parameters \( \alpha \) and \( 1/\beta \), denoted \( \text{PV}(\alpha, 1/\beta) \).

\[ \lim_{x \to \infty} f(x) = \begin{cases} \infty & \text{if } \alpha < 1 \\ \frac{1}{\beta} & \text{if } \alpha = 1 \\ 0 & \text{if } \alpha > 1 \end{cases} \]
Weibull, Weibull($\alpha, \beta$)

**Possible applications**
- Time to complete some task
- Time to failure of a piece of equipment

**Density**
\[ f(x) = \begin{cases} \alpha \beta^{-\alpha} x^{\alpha-1} e^{-\frac{x^\alpha}{\beta}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \]

**Distribution**
\[ F(x) = \begin{cases} 1 - e^{-\frac{x^\alpha}{\beta}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \]

**Parameters**
- Shape parameter $\alpha > 0$
- Scale parameter $\beta > 0$

**Range**
\([0, \infty)\)

**Mean**
\[ \frac{\beta}{\alpha} \Gamma\left(\frac{1}{\alpha}\right) \]

**Variance**
\[ \frac{\beta^2}{\alpha^2} \left( \frac{1}{\alpha^2} \Gamma\left(\frac{2}{\alpha}\right) - \frac{1}{\alpha} \left( \frac{1}{\alpha-1} \right) \right) \]

**Mode**
\[ \left( \frac{\alpha - 1}{\alpha} \right)^{\frac{1}{1/\alpha}} \text{ if } \alpha \geq 1 \]
\[ 0 \text{ if } \alpha < 1 \]

**MLE**
The following two equations must be satisfied:
\[ \frac{\sum_{i=1}^{n} x_i^\alpha \ln x_i}{\sum_{i=1}^{n} x_i^\alpha} - \frac{1}{\alpha} = \frac{\sum_{i=1}^{n} \ln x_i}{n}, \]
\[ \hat{\beta} = \left( \frac{\sum_{i=1}^{n} x_i^\alpha}{n} \right)^{\frac{1}{\alpha}} \]

The first can be solved for $\hat{\alpha}$ numerically by Newton’s method, and the second equation then gives $\hat{\beta}$ directly. The general recursive step for the Newton iterations is
\[ \hat{\alpha}_{k+1} = \frac{\hat{\alpha}_k + \frac{A + 1/\hat{\alpha}_k - C \beta}{1/\hat{\alpha}_k^2 + (B \hat{\beta} - C^2)/\beta}} \]
Mean: $\mu$

Variance: $\sigma^2$

Mode: $\mu$

MLE: $\hat{\mu} = \bar{X}(n)$, $\sigma = \left[ \frac{1}{n-1} \sum (x_i - \mu)^2 \right]^{1/2}$

Comments:
1. If two jointly distributed normal random variables are uncorrelated, they are also independent. For distributions other than normal, this implication is not true in general.
2. Suppose that the joint distribution of $X_1, X_2, \ldots, X_n$ is multivariate normal and let $\mu_i = \text{E}(X_i)$ and $\Sigma = \text{Cov}(X_i, X_j)$. Then for any real numbers $a_1, a_2, \ldots, a_n$, the random variable $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ has a normal distribution with mean $\mu = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$ and variance $\sigma^2 = \sum a_i^2 \sigma_i^2$.

Note that we need not assume independence of the $X_i$'s, if the $X_i$'s are independent, then $\sigma^2 = \sum a_i^2 \text{Var}(X_i)$.

3. The $N(0, 1)$ distribution is often called the standard or unit normal distribution.

4. If $X_1, X_2, \ldots, X_n$ are independent standard normal random variables, then $X_1^2 + X_2^2 + \cdots + X_n^2$ has a chi-square distribution with $n$ df, which is also the gamma($n/2, 1/2$) distribution.

5. If $X \sim N(\mu, \sigma^2)$, then $e^X$ has the lognormal distribution with parameters $\mu$ and $\sigma$, denoted LN($\mu, \sigma^2$).

6. If $X \sim N(0, 1)$, if $Y$ has a chi-square distribution with $k$ df, and if $X$ and $Y$ are independent, then $X^2 + Y$ has a t distribution with $k$ df (sometimes called Student's t distribution).

7. If the normal distribution is used to represent a nonnegative quantity (e.g., time), then its density should be truncated at $x = 0$ (see Sec. 6.8).

8. As $\sigma \to 0$, the normal distribution becomes degenerate at $\mu$.

![Figure 6.9](image)

N(0, 1) density function.
Lognormal

\( \text{LN}(\mu, \sigma^2) \)

Possible applications

Time to perform some task (density takes on shapes similar to gamma(\( \alpha, \beta \)) and Weibull(\( \alpha, \beta \)) densities for \( \alpha > 1 \), but can have a large "spike" close to \( x = 0 \) that is often useful); quantities that are the product of a large number of other quantities (by virtue of central limit theorems).

Density

\[
f(x) = \begin{cases} 
\frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(\frac{-\ln(x - \mu)^2}{2\sigma^2}\right) & \text{if } x > 0 \\
0 & \text{otherwise}
\end{cases}
\]

Distribution

No closed form

Parameters

Shape parameter \( \sigma > 0 \), scale parameter \( \mu \in (-\infty, \infty) \)

Range

\([0, \infty)\)

Mean

\(e^{\mu + \frac{\sigma^2}{2}}\)

Variance

\(e^{2\mu + \sigma^2}(e^\sigma - 1)\)

Mode

\(e^{\mu - \frac{\sigma^2}{2}}\)

MLE

\[
\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \ln X_i, \quad \hat{\sigma} = \left[ \frac{1}{n} \sum_{i=1}^{n} (\ln X_i - \hat{\mu})^2 \right]^{1/2}
\]

Comments

1. \( X \sim \text{LN}(\mu, \sigma^2) \) if and only if \( \ln X \sim N(\mu, \sigma^2) \). Thus, if one has data \( X_1, X_2, \ldots, X_n \) that are thought to be lognormal, the logarithms of the data points \( \ln X_1, \ln X_2, \ldots, \ln X_n \) can be treated as normally distributed data for purposes of hypothesizing a distribution, parameter estimation, and goodness-of-fit testing.

2. As \( \sigma \to 0 \), the lognormal distribution becomes degenerate at \( e^\mu \). Thus, lognormal densities for small \( \sigma \) have a sharp peak at the mode.

3. \( \lim_{x \to 0} f(x) = 0 \), regardless of the parameter values.
**Beta**

**Beta**\( (\alpha_1, \alpha_2) \)

**Possible applications**

Used as a rough model in the absence of data (see Sec. 6.11); distribution of a random proportion, such as the proportion of defective items in a shipment time to complete a task, e.g., in a PERT network.

**Density**

(see Fig. 6.11)

\[
    f(x) = \begin{cases} 
        \frac{x^{\alpha_1-1}(1-x)^{\alpha_2-1}}{B(\alpha_1, \alpha_2)} & \text{if } 0 < x < 1 \\
        0 & \text{otherwise}
    \end{cases}
\]

where \( B(\alpha_1, \alpha_2) \) is the beta function, defined by

\[
    B(z_1, z_2) = \int_0^1 t^{z_1-1}(1-t)^{z_2-1} \, dt
\]

for any real numbers \( z_1 > 0 \) and \( z_2 > 0 \). Some properties of the beta function:

\[
    B(z_1, z_2) = B(z_2, z_1), \quad B(z_1, z_2) = \frac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1 + z_2)}
\]

**Distribution**

No closed form, in general. If either \( \alpha_1 \) or \( \alpha_2 \) is a positive integer, a binomial expansion can be used to obtain \( P(x) \), which will be a polynomial in \( x \), and the powers of \( x \) will be, in general, positive real numbers ranging from 0 through \( \alpha_1 + \alpha_2 - 1 \).

**Parameters**

Shape parameters \( \alpha_1 > 0 \) and \( \alpha_2 > 0 \)

**Range**

\([0, 1]\)

**Mean**

\[
    \frac{\alpha_1}{\alpha_1 + \alpha_2}
\]

**Variance**

\[
    \frac{\alpha_1\alpha_2}{(\alpha_1 + \alpha_2)^2(\alpha_1 + \alpha_2 + 1)}
\]

**Mode**

0 and 1

6. The beta(1, 2) density is a left triangle, and the beta(2, 1) density is a right triangle.

7. \[
    \lim_{x \to 0} f(x) = \begin{cases} 
        \infty & \text{if } \alpha_1 < 1 \\
        \alpha_2 & \text{if } \alpha_1 = 1 \\
        0 & \text{if } \alpha_1 > 1
    \end{cases}
\]

8. The density is symmetric about \( x = \frac{1}{2} \) if and only if \( \alpha_1 = \alpha_2 \). Also, the mean and the mode are equal if and only if \( \alpha_1 = \alpha_2 \).
**Pearson type V**

<table>
<thead>
<tr>
<th>Possible applications</th>
<th>Density (see Fig. 6.12)</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to perform some task (density takes on shapes similar to lognormal, but can have a larger &quot;spike&quot; close to $x = 0$)</td>
<td>$f(x) = \begin{cases} \frac{x^{(\alpha-1)}e^{-x/\beta}}{\beta \Gamma(\alpha)} &amp; \text{if } x &gt; 0 \ 0 &amp; \text{otherwise} \end{cases}$</td>
<td>$F(x) = \begin{cases} 1 - F_{\gamma}(\frac{1}{x}) &amp; \text{if } x &gt; 0 \ 0 &amp; \text{otherwise} \end{cases}$</td>
</tr>
</tbody>
</table>

where $F_{\gamma}(x)$ is the distribution function of a gamma($\alpha$, $1/\beta$) random variable.

**Parameters**
- Shape parameter $\alpha > 0$, scale parameter $\beta > 0$.

**Range** $[0, \infty)$

**Mean** $\frac{\beta}{\alpha - 1}$ for $\alpha > 1$

**Variance** $\frac{\beta^2}{(\alpha - 1)(\alpha - 2)}$ for $\alpha > 2$

**Mode** $\frac{\beta}{\alpha + 1}$

**MLE**

If one has data $X_1, X_2, \ldots, X_n$ then fit a gamma($\alpha, \beta$) distribution to $1/X_1, 1/X_2, \ldots, 1/X_n$ resulting in the maximum-likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$.

Then the maximum-likelihood estimators for the PTS($\alpha, \beta$) are $\hat{\alpha} = \hat{\alpha}$ and $\hat{\beta} = 1/\hat{\beta}$ (see comment 1 below).

**Comments**
1. $X \sim \text{PTS}(\alpha, \beta)$ if and only if $Y = 1/X \sim \text{gamma}(\alpha, 1/\beta)$. Thus, the Pearson type V distribution is sometimes called the inverted gamma distribution.
2. Note that the mean and variance exist only for certain values of the shape parameter.

**FIGURE 6.12**

PTS($\alpha, 1$) density functions.
PT6(α₁, α₂, β)

 Possible applications: Time to perform some task

 Density (see Fig. 6.13): 

 \[ f(x) = \begin{cases} \frac{(x/\beta)^{\alpha_1-1}}{\beta B(\alpha_1, \alpha_2)(1 + (x/\beta)^{\alpha_1-\alpha_2})} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \]

 Distribution: 

 \[ F(x) = \begin{cases} F_B\left(\frac{x}{x + \beta}\right) & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \]

 where \( F_B(x) \) is the distribution function of a beta(α₁, α₂) random variable

 Parameters:

 - Range: \( [0, \infty) \)
 - Mean: \( \frac{\beta \alpha_1}{\alpha_2 - 1} \) for \( \alpha_2 > 1 \)
 - Variance: \( \frac{\beta^2 \alpha_1 (\alpha_1 + \alpha_2 - 1)}{(\alpha_2 - 1)^2(\alpha_2 - 2)} \) for \( \alpha_2 > 2 \)
 - Mode: \( \begin{cases} \beta (\alpha_1 - 1) & \text{if } \alpha_1 \geq 1 \\ \alpha_2 + 1 & \text{otherwise} \end{cases} \)

 If one has data \( X_1, X_2, \ldots, X_n \) that are thought to be PT6(α₁, α₂, 1), then fit a beta(α₁, α₂) distribution to \( X_i/(1 + X_i) \) for \( i = 1, 2, \ldots, n \), resulting in the maximum-likelihood estimators \( \hat{\alpha}_1 \) and \( \hat{\alpha}_2 \). Then the maximum-likelihood estimators for the PT6(α₁, α₂, 1) (note that \( \beta = 1 \)) distribution are also \( \hat{\alpha}_1 \) and \( \hat{\alpha}_2 \) (see comment 1 below).

 Comments:

 1. \( X \sim \text{PT6}(\alpha_1, \alpha_2, 1) \) if and only if \( Y = X/(1 + X) \sim \text{beta}(\alpha_1, \alpha_2) \).
 2. If \( X_1 \) and \( X_2 \) are independent random variables with \( X_1 \sim \text{gamma}(\alpha_1, \beta) \) and \( X_2 \sim \text{gamma}(\alpha_2, 1) \), then \( Y = X_1/X_2 \sim \text{PT6}(\alpha_1, \alpha_2, \beta) \) (see Prob. 6.3).
 3. Note that the mean and variance exist only for certain values of the shape parameter \( \alpha_2 \).

 FIGURE 6.13
 PT6(α₁, α₂, 1) density functions.
<table>
<thead>
<tr>
<th>Log-logistic</th>
<th>LL(α, β)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Possible applications</strong></td>
<td>Time to perform some task</td>
</tr>
<tr>
<td><strong>Density</strong> (see Fig. 6.14)</td>
<td>$f(x) = \begin{cases} \frac{\alpha(x/\beta)^{\alpha-1}}{\beta(1 + (x/\beta)^\alpha)} &amp; \text{if } x &gt; 0 \ 0 &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td><strong>Distribution</strong></td>
<td>$F(x) = \begin{cases} 1 - \frac{1}{1 + (x/\beta)^\alpha} &amp; \text{if } x &gt; 0 \ 0 &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td><strong>Parameters</strong></td>
<td>Shape parameter $\alpha &gt; 0$, scale parameter $\beta &gt; 0$</td>
</tr>
</tbody>
</table>

Comment:

$X \sim LL(\alpha, \beta)$ if and only if $\ln X$ is distributed as a logistic distribution (see Prob. 8.1) with location parameter $\ln \beta$ and scale parameter $1/\alpha$. Thus, if one has data $X_1, X_2, \ldots, X_n$ that are thought to be log-logistic, the logarithms of the data points $\ln X_1, \ln X_2, \ldots, \ln X_n$ can be treated as having a logistic distribution for purposes of hypothesizing a distribution, parameter estimation, and goodness-of-fit testing.
<table>
<thead>
<tr>
<th>Johnson S₄</th>
<th>JSB((\alpha_1, \alpha_2, a, b))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Density</strong></td>
<td>(f(x) = \begin{cases} \frac{a_d(b - a)}{(x - a)(b - x)\sqrt{2\pi}} e^{-\frac{1}{2}\left[n_1\Phi(x) - \frac{x + a_1}{b - x}\right]^2} &amp; \text{if } a &lt; x &lt; b \ 0 &amp; \text{otherwise} \end{cases} )</td>
</tr>
<tr>
<td><strong>Distribution</strong></td>
<td>(F(x) = \begin{cases} 0 &amp; \text{if } a &lt; x &lt; b \ \frac{\Phi(n_1 + \alpha_2 \ln \frac{x - a}{b - x})}{n_1} &amp; \text{otherwise} \end{cases} )</td>
</tr>
</tbody>
</table>

where \(\Phi(x)\) is the distribution function of a normal random variable with \(\mu = 0\) and \(\sigma^2 = 1\).

**Parameters**
- Location parameter \(a \in (-\infty, \infty)\), scale parameter \(b - a\) (\(b > a\)), shape parameters \(\alpha_1 \in (-\infty, \infty)\) and \(\alpha_2 > 0\)

**Range**
\([a, b]\)

**Mean**
All moments exist but are extremely complicated [see Johnson, Kotz, and Balakrishna (1994, p. 35)].

**Mode**
The density is bimodal when \(\alpha_2 < \frac{1}{\sqrt{2}}\) and
\[
|\alpha_1| < \frac{\sqrt{1 - 2\alpha_2^2}}{\alpha_2} - 2\alpha_2 \tanh^{-1}\left(\sqrt{1 - 2\alpha_2^2}\right)
\]
[tanh\(^{-1}\) is the inverse hyperbolic tangent]; otherwise, the density is unimodal. The equation satisfied by any mode \(x\), other than at the endpoints of the range, is
\[
\frac{2(x - a)}{b - a} = 1 + \alpha_2 \alpha_1 \ln \left(\frac{x - a}{b - a}\right)
\]

**Comments**
1. \(X \sim \text{JSB}(\alpha_1, \alpha_2, a, b)\) if and only if
\[
\alpha_1 + \alpha_2 \ln \left(\frac{x - a}{b - x}\right) \sim N(0, 1)
\]
2. The density function is skewed to the left, symmetric, or skewed to the right if \(\alpha_1 > 0\), \(\alpha_1 = 0\), or \(\alpha_1 < 0\), respectively.
3. \(\lim_{x \to a^+} f(x) = \lim_{x \to b^-} f(x) = 0\) for all values of \(\alpha_1\) and \(\alpha_2\).
4. The four parameters may be estimated using a number of methods [see, for example, Swain, Venkatraman, and Wilson (1988) and Stoffer and Shapiro (1980)].

**Figure 6.15**
JSB(\(\alpha_1, \alpha_2, 1\)) density functions.
Johnson $S_N$

\[ f(x) = \frac{a_0}{\sqrt{2\pi}} \left( \frac{a_1 + a_2}{\beta} \right) \phi \left( \sqrt{\frac{(x - \gamma)^2}{\beta^2} + 1} \right) \]

for \(-\infty < x < \infty\)

**Distribution**

\[ F(x) = \Phi \left( a_1 + a_2 \frac{x - \gamma}{\beta} \right) \]

for \(-\infty < x < \infty\)

**Parameters**

- Location parameter $\gamma \in (-\infty, \infty)$, scale parameter $\beta > 0$, shape parameters $a_1, a_2 \in (-\infty, \infty)$ and $a_3 > 0$

**Range**

$(-\infty, \infty)$

**Mean**

$\gamma - \beta \tanh^{-1} \left( \frac{a_1}{a_3} \right)$, where tanh is the hyperbolic tangent

**Mode**

The equation satisfied by the mode, other than at the endpoints of the range, is

$\gamma + \beta \ln \left( 1 + e^{\gamma/\beta} \right) = 0$

**Comments**

1. $X \sim JSU(a_1, a_2, \gamma, \beta)$ if and only if

$\frac{X - \gamma}{\beta} \sim \Phi \left( a_1 + a_2 \frac{x - \gamma}{\beta} \right)$

2. The density function is skewed to the left, symmetric, or skewed to the right if $a_1 > 0$, $a_1 = 0$, or $a_1 < 0$, respectively.

3. The four parameters may be estimated by a number of methods [see, for example, Swain, Vonnegut, and Wilson (1948) and Stufler and Shapiro (1980)].

---

**FIGURE 4.16**

$JSU(\alpha_1, \alpha_2, \gamma, \beta)$ density functions.
<table>
<thead>
<tr>
<th>Triangular</th>
<th>$\text{triang}(a, b, c)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possible applications</td>
<td>Used as a rough model in the absence of data (see Sec. 6.11)</td>
</tr>
<tr>
<td>Density (see Fig. 6.17)</td>
<td>$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} &amp; \text{if } a \leq x \leq c \ \frac{2(b-x)}{(b-a)(c-a)} &amp; \text{if } c &lt; x \leq b \ 0 &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td>Distribution</td>
<td>$F(x) = \begin{cases} 0 &amp; \text{if } x &lt; a \ \frac{(x-a)^2}{(b-a)(c-a)} &amp; \text{if } a \leq x \leq c \ 1 - \frac{(b-x)^2}{(b-a)(c-a)} &amp; \text{if } c &lt; x \leq b \ 1 &amp; \text{if } b &lt; x \end{cases}$</td>
</tr>
<tr>
<td>Parameters</td>
<td>$a$, $b$, and $c$ real numbers with $a &lt; c &lt; b$. $a$ is a location parameter, $b - a$ is a scale parameter, $c$ is a shape parameter</td>
</tr>
<tr>
<td>Ranges</td>
<td>$[a, b]$</td>
</tr>
<tr>
<td>Mean</td>
<td>$\frac{a + b + c}{3}$</td>
</tr>
</tbody>
</table>

**Variance**

\[ \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18} \]

**Mode**

$c$

**MLE**

The limiting cases as $c \to b$ and $c \to a$ are called the right triangular and left triangular distributions, respectively, and are discussed in Prob. 8.7. For $a = 0$ and $b = 1$, both the left and right triangular distributions are special cases of the beta distribution.

**Comment**

Our use of the triangular distribution, as described in Sec. 6.11, is as a rough model when there are no data. Thus, MLEs are not relevant.

**Figure 6.17**

$\text{triang}(a, b, c)$ density function.
Discrete distribution

- Bernoulli
- Discrete uniform
- Binomial
- Geometric
- Negative binormal
- Poisson
<table>
<thead>
<tr>
<th>Bernoulli(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Possible applications</strong></td>
</tr>
<tr>
<td><strong>Mass (see Figs. 6.18)</strong></td>
</tr>
<tr>
<td><strong>Distribution</strong></td>
</tr>
<tr>
<td><strong>Parameter</strong></td>
</tr>
<tr>
<td><strong>Range</strong></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td><strong>Variance</strong></td>
</tr>
<tr>
<td><strong>Mode</strong></td>
</tr>
<tr>
<td><strong>MLE</strong></td>
</tr>
<tr>
<td><strong>Corrections</strong></td>
</tr>
<tr>
<td>1. A Bernoulli(p) random variable $X$ can be thought of as the outcome of an experiment that either “hits” or “misses.” If the probability of success is $p$, and we let $X = 0$ if the experiment fails and $X = 1$ if it succeeds, then $X \sim \text{Bernoulli}(p)$. Such an experiment, often called a <em>Bernoulli trial</em>, provides a convenient way of relating several other discrete distributions to the Bernoulli distribution.</td>
</tr>
<tr>
<td>2. If $x$s is a positive integer and $X_1, X_2, \ldots, X_n$ are independent Bernoulli(p) random variables, then $X_1 + X_2 + \cdots + X_n$ has the binomial distribution with parameters $n$ and $p$. Thus, a binomial random variable can be thought of as the number of successes in a fixed number of independent Bernoulli trials.</td>
</tr>
<tr>
<td>3. Suppose we begin making independent replications of a Bernoulli trial with probability $p$ of success on each trial. Then the number of failures before observing the first success has a geometric distribution with parameter $p$. For a positive integer $k$, the number of failures before observing the $k$th success has a negative binomial distribution with parameters $k$ and $p$.</td>
</tr>
</tbody>
</table>

**Figure 6.18** Bernoulli(p) mass function ($p > 0.5$ here).
Discrete uniform \( DU(i, j) \)

Possible applications: Random occurrence with several possible outcomes, each of which is equally likely; used as a "first" model for a quantity that is varying among the integers \( i \) through \( j \) but about which little else is known.

Mass (see Fig. 6.19)

\[
p(x) = \begin{cases} \frac{1}{j-i+1} & \text{if } x \in \{i, i+1, \ldots, j\} \\ 0 & \text{otherwise} \end{cases}
\]

Distribution

\[
P(x) = \begin{cases} 0 & \text{if } x < i \\ \frac{|x-i|+1}{j-i+1} & \text{if } i \leq x \leq j \\ 1 & \text{if } j < x \end{cases}
\]

where \( \lfloor x \rfloor \) denotes the largest integer \( \leq x \)

Parameters: \( i \) and \( j \) integers with \( i \leq j \); \( i \) is a location parameter, \( j - i \) is a scale parameter.

Range

\( \{i, i+1, \ldots, j\} \)

Mean

\[
\frac{i + j}{2}
\]

Variance

\[
\frac{(j-i+1)^2 - 1}{12}
\]

Mode

Does not uniquely exist

MLE

\[
i = \min X, \quad j = \max X
\]

Comment

The \( DU(0, 1) \) and \( Bernoulli(\frac{1}{2}) \) distributions are the same.

![DU(i, j) mass function](image)
The binomial distribution is used to model the number of successes in a fixed number of independent Bernoulli trials, each with a constant success probability $p$. In this context, a Bernoulli trial is a trial where there are exactly two possible outcomes, often referred to as success or failure. The number of trials is denoted by $n$, and the number of successes is denoted by $X$.

### Parameters

- **$n$**: Number of trials.
- **$p$**: Probability of success in each trial.
- **$X$**: Number of successes.

### Mass Function

The mass function of the binomial distribution is given by:

$$ P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} $$

where $\binom{n}{x}$ is the binomial coefficient, defined as

$$ \binom{n}{x} = \frac{n!}{x!(n-x)!} $$

### Mean

$$ \mu = np $$

### Variance

$$ \sigma^2 = np(1-p) $$

### Mode

- If $np$ is an integer, the mode is $\text{mode} = np$.
- If $np$ is not an integer, the mode is $\text{mode} = \lfloor np \rfloor$ and $\lfloor np \rfloor + 1$.

### MLE

If $X$ is known, then $\hat{p} = \frac{X}{n}$. If both $n$ and $p$ are unknown, then $\hat{n}$ and $\hat{p}$ exist if and only if $\bar{X} = \frac{X}{n} > (n-1)^{\frac{1}{n}}$. If these conditions are satisfied, the MLE for $p$ is

$$ \hat{p} = \frac{\bar{X}}{\bar{X} + (n-\bar{X})} $$

### Comments

1. If $Y_1, Y_2, \ldots, Y_n$ are independent Bernoulli($p$) random variables, then $Y_1 + Y_2 + \cdots + Y_n \sim \text{Bin}(n, p)$.
2. If $X_1, X_2, \ldots, X_n$ are independent random variables and $X_i \sim \text{Bin}(t, p_i)$, then $X_1 + X_2 + \cdots + X_n \sim \text{Bin}(t_1 + t_2 + \cdots + t_n, p_1 + p_2 + \cdots + p_n)$.
3. The bin$(t, p)$ distribution is symmetric if and only if $p = \frac{1}{2}$.
4. $X \sim \text{Bin}(t, p)$ if and only if $X \sim \text{Bin}(1, p)$.
5. The bin$(1, p)$ and Bernoulli$(p)$ distributions are the same.

Note also that $g(t, \bar{X}/t)$ is a unimodal function of $t$. The figure illustrates the probability mass function (PMF) for different values of $n$ and $p$. The bars represent the probability of each possible outcome for $X$. The mean and variance are also indicated on the graph.
<table>
<thead>
<tr>
<th>Geometric</th>
<th>geom(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possible applications</td>
<td>Number of failures before the first success in a sequence of independent Bernoulli trials with probability p of success on each trial; number of items inspected before encountering the first defective item; number of items in a batch of random size; number of items demanded from an inventory</td>
</tr>
</tbody>
</table>

**Mass (see Fig. 6.21)**

\[
p(x) = \begin{cases} 
  p(1 - p)^x & \text{if } x \in \{0, 1, \ldots\} \\
  0 & \text{otherwise}
\end{cases}
\]

**Distribution**

\[
F(x) = \begin{cases} 
  1 - (1 - p)^{x+1} & \text{if } x \geq 0 \\
  0 & \text{otherwise}
\end{cases}
\]

**Parameter**

\(p \in (0, 1)\)

**Range**

\(\{0, 1, \ldots\}\)

**Mean**

\[
\frac{1 - p}{p}
\]
### Negative Binomial

#### Parameters
- **$s$**: a positive integer, $s \in (0, 1)
- **$p$**: $s \in (0, 1)$

#### Range
- $0, 1, \ldots$

#### Mean
- $\frac{s-1}{p}$

#### Variance
- $\frac{s(p+1)}{p^2}$

#### Mode
- Let $y = \lfloor (1 - p) - 1/\lfloor y + 1 \rfloor \rfloor$ if $y$ is an integer otherwise $\lfloor y + 1 \rfloor$.

### MLE

If $s$ is known, then $\hat{p} = s/(\log(s) + 1)$. If both $s$ and $p$ are unknown, then $\hat{s} = \lfloor (s - \log(s)) - 1/\lfloor y + 1 \rfloor \rfloor$. Let $\hat{X}_1, \ldots, \hat{X}_n$ be the number of $X_1, \ldots, X_n$ which exceed the maximum observed value of $n$. Let $\hat{X} = \max\{\hat{X}_1, \ldots, \hat{X}_n\}$.

### Comments
1. If $Y_1, Y_2, \ldots, Y_s$ are independent geometric($p$) random variables, then $Y_1 + Y_2 + \cdots + Y_s \sim \text{negbin}(s, p)$.
2. If $Y_1, Y_2, \ldots, Y_s$ is a sequence of independent Bernoulli($p$) random variables and $X = \text{max}(\sum_{i=1}^{s} Y_i - s, 0)$, then $X \sim \text{negbin}(s, p)$.
3. If $X_1, X_2, \ldots, X_s$ are independent random variables and $X_1 \sim \text{negbin}(s, p)$, then $X_1 + X_2 + \cdots + X_s \sim \text{negbin}(s, p)$.
4. The negbin($s, p$) and geometric($p$) distributions are the same.
**Poisson**

**Poisson(λ)**

<table>
<thead>
<tr>
<th>Possible applications</th>
<th>Number of events that occur in an interval of time when the events are occurring at a constant rate (see Sec. 6.12); number of items in a batch of random size; number of items demanded from an inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass (see Fig. 6.23)</td>
<td>$p(x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!} &amp; \text{if } x \in {0, 1, \ldots} \ 0 &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td>Distribution</td>
<td>$P(x) = \begin{cases} 0 &amp; \text{if } x &lt; 0 \ \frac{1}{e} \frac{\lambda^x}{x!} &amp; \text{if } x \geq 0 \end{cases}$</td>
</tr>
</tbody>
</table>

**Parameter**
- $\lambda > 0$

**Range**
- $\{0, 1, \ldots\}$

**Mean**
- $\lambda$

**Variance**
- $\lambda$

**Mode**
- $\lceil \frac{\lambda - 1}{\lambda} \rceil$ if $\lambda$ is an integer
- $\lfloor \lambda \rfloor$ otherwise

**MLE**
- $\hat{\lambda} = \bar{X}(n)$.

**Comments**
1. Let $Y_1, Y_2, \ldots$ be a sequence of nonnegative IID random variables and let $X = \max\{i: \sum_{j=1}^{i} Y_j \leq 1\}$. Then the distribution of the $Y_i$'s is $\exp(1/\lambda)$ if and only if $X \sim \text{Poisson}(\lambda)$. Also, if $X' = \max\{i: \sum_{j=1}^{i} Y_j = \lambda\}$, then the $Y_i$'s are $\exp(1)$ if and only if $X' \sim \text{Poisson}(\lambda)$ (see also Sec. 6.12).
2. If $X_1, X_2, \ldots, X_n$ are independent random variables and $X_i \sim \text{Poisson}(\lambda_i)$, then $X_1 + X_2 + \cdots + X_n \sim \text{Poisson}(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$.

**Figure 6.23**
Poisson(λ) mass functions.

\[ \lambda = 0.5 \]

\[ \lambda = 0.1 \]

\[ \lambda = 1 \]

\[ \lambda = 2 \]

\[ \lambda = 6 \]
Continuous, piecewise-linear empirical distribution function from original data.

\[
F(x) = \begin{cases} 
0 & \text{if } x < X_{(1)} \\
\frac{i - 1}{n - 1} + \frac{x - X_{(0)}}{(n - 1)(X_{(i+1)} - X_{(i)})} & \text{if } X_{(i)} \leq x < X_{(i+1)} \\
1 & \text{for } i = 1, 2, \ldots, n - 1 \\
1 & \text{if } X_{(n)} \leq x
\end{cases}
\]

\[
G(x) = \begin{cases} 
0 & \text{if } x < a_0 \\
G(a_{j-1}) + \frac{x - a_{j-1}}{a_j - a_{j-1}} [G(a_j) - G(a_{j-1})] & \text{if } a_{j-1} \leq x < a_j \\
1 & \text{for } j = 1, 2, \ldots, k \\
1 & \text{if } a_k \leq x
\end{cases}
\]
**Figure 6.25**
Continuous, piecewise-linear empirical distribution function from grouped data.

**Figure 6.26**
Typical density function experienced in practice.
6.3 Techniques for assessing sample independence
FIGURE 6.27
Correlation plot for independent exponential data.

FIGURE 6.28
Scatter diagram for independent exponential data.
FIGURE 6.29
Correlation plot for correlated queueing data.
FIGURE 6.30
Scatter diagram for correlated queueing data.