

---

## **AIMMS Language Reference - Available Distributions and Statistical Operators**

This file contains only one chapter of the book. For a free download of the complete book in pdf format, please visit [www.aimms.com](http://www.aimms.com) or order your hard-copy at [www.lulu.com/aimms](http://www.lulu.com/aimms).

Copyright © 1993–2011 by Paragon Decision Technology B.V. All rights reserved.

Paragon Decision Technology B.V.	Paragon Decision Technology Inc.	Paragon Decision Technology Pte.
Schipholweg 1	500 108th Avenue NE	Ltd.
2034 LS Haarlem	Ste. # 1085	80 Raffles Place
The Netherlands	Bellevue, WA 98004	UOB Plaza 1, Level 36-01
Tel.: +31 23 5511512	USA	Singapore 048624
Fax: +31 23 5511517	Tel.: +1 425 458 4024	Tel.: +65 9640 4182
	Fax: +1 425 458 4025	

Email: [info@aimms.com](mailto:info@aimms.com)  
WWW: [www.aimms.com](http://www.aimms.com)

AIMMS is a registered trademark of Paragon Decision Technology B.V. IBM ILOG CPLEX and sc CPLEX is a registered trademark of IBM Corporation. GUROBI is a registered trademark of Gurobi Optimization, Inc. KNITRO is a registered trademark of Ziena Optimization, Inc. XPRESS-MP is a registered trademark of FICO Fair Isaac Corporation. MOSEK is a registered trademark of Mosek ApS. WINDOWS and EXCEL are registered trademarks of Microsoft Corporation.  $\text{T}_{\text{E}}\text{X}$ ,  $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ , and  $\text{A}_{\text{M}}\text{S}-\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$  are trademarks of the American Mathematical Society. LUCIDA is a registered trademark of Bigelow & Holmes Inc. ACROBAT is a registered trademark of Adobe Systems Inc. Other brands and their products are trademarks of their respective holders.

Information in this document is subject to change without notice and does not represent a commitment on the part of Paragon Decision Technology B.V. The software described in this document is furnished under a license agreement and may only be used and copied in accordance with the terms of the agreement. The documentation may not, in whole or in part, be copied, photocopied, reproduced, translated, or reduced to any electronic medium or machine-readable form without prior consent, in writing, from Paragon Decision Technology B.V.

**Paragon Decision Technology B.V. makes no representation or warranty with respect to the adequacy of this documentation or the programs which it describes for any particular purpose or with respect to its adequacy to produce any particular result. In no event shall Paragon Decision Technology B.V., its employees, its contractors or the authors of this documentation be liable for special, direct, indirect or consequential damages, losses, costs, charges, claims, demands, or claims for lost profits, fees or expenses of any nature or kind.**

**In addition to the foregoing, users should recognize that all complex software systems and their documentation contain errors and omissions. The authors, Paragon Decision Technology B.V. and its employees, and its contractors shall not be responsible under any circumstances for providing information or corrections to errors and omissions discovered at any time in this book or the software it describes, whether or not they are aware of the errors or omissions. The authors, Paragon Decision Technology B.V. and its employees, and its contractors do not recommend the use of the software described in this book for applications in which errors or omissions could threaten life, injury or significant loss.**

This documentation was typeset by Paragon Decision Technology B.V. using  $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$  and the LUCIDA font family.

---

# Appendices

## Appendix A

### Description of available distributions and statistical operators

This chapter provides a more elaborate description of the distributions and distribution and sample operators listed in Tables 6.5, 6.6 and 6.7. You can use this information when you want to set up an experiment around your (optimization-based) AIMMS model.

*This chapter*

For each of the available distributions we describe

*Description of distributions*

- its parameters, mean and variance,
- the unit relationship between its parameters and result,
- its shape, and
- its typical use in applications.

Such information may be useful in the selection and use of a distribution to describe the particular statistical behavior of input data of experiments that you want to perform on top of your model. However, a general guideline for choosing the right might be in order and is provided in the next paragraph.

Whenever your experiment counts a number of occurrences, you should first make a distinction between experiments with replacement (i.e. throwing dice), experiments without replacement (i.e. drawing cards from a deck), or experiments in which independent occurrences take place at random moments (i.e. customers appearing at a desk). Having made this distinction, Table A.1 will help you to select the right distribution for your experiment. In any other case the Normal distribution should be considered first. Although this distribution is unbounded, it is declining so rapidly that it can often be used even when the result should be bounded. If the Normal distribution does not suffice, the primary selection criterium is existence of bounds: AIMMS provides the user with distributions with no bounds, one (lower) bound and two (upper and lower) bounds. See section A.2 (continuous distributions) for details.

*Choosing the right distribution*

For each of the available distribution and sample operators we provide

- the interpretation of its result, and
- the formula for the computation of the operator.

Such information may be useful when you want to perform an analysis of the results of your experiments.

All distribution operators that are listed in Section A.3 have been introduced in AIMMS 3.4, although the `DistributionCumulative` and `DistributionInverseCumulative` operator were already available under the names `CumulativeDistribution` and `InverseCumulativeDistribution`, respectively. Furthermore, in order to obtain a consistent set of distribution functions the prototype for some of them has been slightly adapted. Section A.2 discusses the function prototype of the continuous distribution functions in full detail. Both the old and the new function prototypes are discussed in the AIMMS Function Reference. To make sure that models using distribution functions and developed in an older version of AIMMS are working correctly, you should set the option `Distribution_compatibility` to 'AIMMS 3.0'.

*Description of  
distribution  
operators*

*Option for  
backward  
compatibility*

---

## A.1 Discrete distributions

We start our discussion with the discrete distributions available in AIMMS. They are

- the `Binomial` distribution,
- the `HyperGeometric` distribution,
- the `Poisson` distribution,
- the `Negative Binomial` distribution, and
- the `Geometric` distribution.

The `Binomial`, `HyperGeometric` and `Poisson` distributions describe the number of times that a particular outcome (referred to as "success") occurs. In the `Binomial` distribution, the underlying assumption is a fixed number of trials and a constant likelihood of success. In the `HyperGeometric` distribution, the underlying assumption is "sampling without replacement": A fixed number of trials are taken from a population. Each element of this population denotes a success or failure and cannot occur more than once. In the `Poisson` distribution the number of trials is not fixed. Instead we assume that successes occur independently of each other and with equal chance for all intervals with the same duration.

*Discrete  
distributions*

*Discrete  
distributions  
describing  
successes*

	with replacement	without replacement	independent occurrences at random moments
example	throwing dice	drawing cards	serving customers
# trials until first success / time until first occurrence	Geometric	not supported in AIMMS	Exponential (continuous)
# trials until $n$ -th success / time until $n$ -th occurrence	Negative Binomial	not supported in AIMMS	Gamma (continuous)
# successes in fixed # trials / # successes in fixed time	Binomial	Hypergeometric	Poisson

Table A.1: Overview of discrete distributions in AIMMS

The Negative Binomial distribution describes the number of failures before a specified number of successes have occurred. It assumes a constant chance of success for each trial, so it is linked to the Binomial distribution. Similarly, the distribution linked to Poisson distribution that describes the amount of time until a certain number of successes have occurred is known as the Gamma distribution and is discussed in Section A.2. The Negative Binomial distribution is a special case of the Geometric distribution and describes the number of failures before the first success occurs. Similarly, the Exponential distribution is a special case of the Gamma distribution and describes the amount of time until the first occurrence.

*Distributions describing trials*

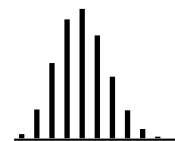
Table A.1 shows the relation between the discrete distributions. The continuous Exponential and Gamma distribution naturally fit in this table as they represent the distribution of the time it takes before the first/ $n$ -th occurrence (given the average time between two consecutive occurrences).

*Discrete distributions overview*

The Binomial( $p, n$ ) distribution:

- **Input parameters** : Probability of success  $p$  and number of trials  $n$
- **Input check** : integer  $n > 0$  and  $0 < p < 1$
- **Permitted values** :  $\{i \mid i = 0, 1, \dots, n\}$
- **Formula** :  $P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}$
- **Mean** :  $np$
- **Variance** :  $np(1 - p)$
- **Remarks** : Binomial( $p, n$ ) = HyperGeometric( $p, n, \infty$ )

*Binomial distribution*



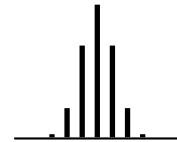
A typical example for this distribution is the number of defectives in a batch of manufactured products where a fixed percentage was found to be defective in previously produced batches. Another example is the number of persons in a group voting yes instead of no, where the probability of yes has been determined on the basis of a sample.

The HyperGeometric( $p, n, N$ ) distribution:

- **Input parameters** : Known initial probability of success  $p$ , number of trials  $n$  and population size  $N$
- **Input check** : integer  $n, N : 0 < n \leq N$ , and  $p \in \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}$
- **Permitted values** :  $\{i \mid i = 0, 1, \dots, n\}$
- **Formula** :  $P(X = i) = \frac{\binom{Np}{i} \binom{N(1-p)}{n-i}}{\binom{N}{n}}$
- **Mean** :  $np$
- **Variance** :  $np(1-p)\frac{N-n}{N-1}$

As an example of this distribution, consider a set of 1000 books of which 30 are faulty. When considering an order containing 50 books from this set, the HyperGeometric(0.03, 50, 1000) distribution shows the probability of observing  $i$  ( $i = 0, 1, \dots, n$ ) faulty books in this subset.

HyperGeometric distribution

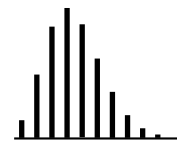


The Poisson( $\lambda$ ) distribution:

- **Input parameters** : Average number of occurrences  $\lambda$
- **Input check** :  $\lambda > 0$
- **Permitted values** :  $\{i \mid i = 0, 1, \dots\}$
- **Formula** :  $P(X = i) = \frac{\lambda^i}{i!} e^{-\lambda}$
- **Mean** :  $\lambda$
- **Variance** :  $\lambda$
- **Remarks** :  $\text{Poisson}(\lambda) = \lim_{p \rightarrow 0} \text{Binomial}(p, \lambda/p)$

The Poisson distribution should be used when there is a constant chance of a 'success' over time or (as an approximation) when there are many occurrences with a very small individual chance of 'success'. Typical examples are the number of visitors in a day, the number of errors in a document, the number of defects in a large batch, the number of telephone calls in a minute, etc.

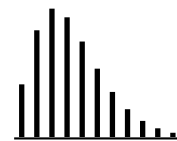
Poisson distribution



The NegativeBinomial( $p, r$ ) distribution:

- **Input parameters** : Success probability  $p$  and number of successes  $r$
- **Input check** :  $0 < p < 1$  and  $r = 1, 2, \dots$
- **Permitted values** :  $\{i \mid i = 0, 1, \dots\}$
- **Formula** :  $P(X = i) = \binom{r+i-1}{i} p^r (1-p)^i$
- **Mean** :  $r/p - r$
- **Variance** :  $rp^2(1-p)$

Negative Binomial distribution

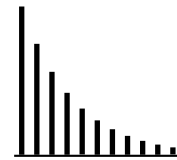


Whenever there is a repetition of the same activity, and you are interested in observing the  $r$ -th occurrence of a particular outcome, then the Negative Binomial distribution might be applicable. A typical situation is going from door-to-door until you have made  $r$  sales, where the probability of making a sale has been determined on the basis of previous experience. Note that the NegativeBinomial distribution describes the number of *failures* before the  $r$ -th success. The distribution of the number of *trials*  $i$  before the  $r$ -th success is given by  $P_{\text{NegativeBinomial}(p,r)}(X = i - r)$ .

The Geometric( $p$ ) distribution:

- **Input parameters** : Probability of success  $p$
- **Input check** :  $0 < p < 1$
- **Permitted values** :  $\{i \mid i = 0, 1, \dots\}$
- **Formula** :  $P(X = i) = (1 - p)^i p$
- **Mean** :  $1/p - 1$
- **Variance** :  $p^2(1 - p)$
- **Remarks** :  $\text{Geometric}(p) = \text{NegativeBinomial}(p, 1)$

Geometric  
distribution



The Geometric distribution is a special case of the NegativeBinomial distribution. So it can be used for the same type of problems (the number of visited doors before the first sale). Another example is an oil company drilling wells until a producing well is found, where the probability of success is based on measurements around the site and comparing them with measurements from other similar sites.

---

## A.2 Continuous distributions

In this section we discuss the set of continuous distributions available in AIMMS.

Continuous  
distributions

The three distributions with both lower and upper bound are

- the Uniform distribution,
- the Triangular distribution, and
- the Beta distribution.

The five distributions with only a lower bound are

- the LogNormal distribution,
- the Exponential distribution,
- the Gamma distribution,
- the Weibull distribution, and
- the Pareto distribution.

The three unbounded distributions are

- the Normal distribution,

- the Logistic distribution, and
- the Extreme Value distribution.

Every parameter of a continuous distributions can be characterized as either a *shape* parameter  $\beta$ , a *location* parameter  $l$ , or a *scale* parameter  $s$ . While the presence and meaning of a shape parameter is usually distribution-dependent, location and scale parameters find their origin in the common transformation

$$x \mapsto \frac{x - l}{s}$$

to shift and stretch a given distribution. By choosing  $l = 0$  and  $s = 1$  the standard form of a distribution is obtained. If a certain distribution has  $n$  shape parameters ( $n \geq 0$ ), these shape parameters will be passed as the first  $n$  parameters to AIMMS. The shape parameters are then followed by two optional parameters, with default values 0 and 1 respectively. For double-bounded distributions these two optional parameters can be interpreted as a lower and upper bound (the value of the location parameter  $l$  for these distributions is equal to the lower bound and the value of the scale parameter  $s$  is equal to the difference between the upper and lower bound). For single-bounded distributions the bound value is often used as the location parameter  $l$ . In this section, whenever the location parameter can be interpreted as a mean value or whenever the scale parameter can be interpreted as the deviation of a distribution, these more meaningful names are used to refer to the parameters. Note that the LogNormal, Gamma and Exponential distributions are distributions that will mostly be used with location parameter equal to 0.

*Parameters of continuous distributions*

When transforming a distribution to standard form, distribution operators change. Section A.5 (scaling of statistical operators) gives the relationships between distribution operators working on random variables  $X(l, s)$  and  $X(0, 1)$ .

*Transformation to standard form*

When a random variable representing some real-life quantity with a given unit of measurement (see also Chapter 30) is distributed according to a particular distribution, some parameters of that distribution are also naturally expressed in terms of this same unit while other parameters are expected to be unitless. In particular, the location and scale parameters of a distribution are measured in the same unit of measurement as the corresponding random variable, while shape parameters (within AIMMS) are implemented as unitless parameters.

*Units of measurement*

When you use a distribution function, AIMMS will perform a unit consistency check on its parameters and result, whenever your model contains one or more QUANTITY declarations. In the description of the continuous distributions below, the expected units of the distribution parameters are denoted in square brackets. Throughout the sequel,  $[x]$  denotes that the parameter should have the same unit of measurement as the random variable  $X$  and  $[-]$  denotes that a parameter should be unitless.

*Unit notation in this appendix*

In practice, the Normal distribution is used quite frequently. Such widespread use is due to a number of pleasant properties:

*A commonly used distribution*

- the Normal distribution has no shape parameters and is symmetrical,
- random values are more likely as they are closer to the mean value,
- it can be directly evaluated for any given mean and standard deviation because it is fully specified through the mean and standard deviation parameter,
- it can be used as a good approximation for distributions on a finite interval, because its probability density is declining fast enough (when moving away from the mean),
- the mean and sum of any number of uncorrelated Normal distributions are Normal distributed themselves, and thus have the same shape, and
- the mean and sum of a large number of uncorrelated distributions are always approximately Normal distributed.

For random variables that have a known lower and upper bound, AIMMS provides three continuous distributions on a finite interval: the Uniform, Triangular and Beta distribution. The Uniform (no shape parameters) and Triangular (one shape parameter) distributions should be sufficient for most experiments. For all remaining experiments, the user might consider the highly configurable Beta (two shape parameters) distribution.

*Distributions for double bounded variables*

When your random variable only has a single bound, you should first check whether the Gamma distribution can be used or whether the Normal distribution is accurate enough. The LogNormal distribution should be considered if the most likely value is near but not at the bound. The Weibull or Gamma distribution ( $\beta > 1$ ), or even the ExtremeValue distribution are alternatives, while the Weibull or Gamma distribution ( $\beta \leq 1$ ) or Pareto distribution should be considered if the bound is the most likely value.

*Distributions for single bounded variables*

The Gamma (and as a special case thereof the Exponential) distribution is widely used for its special meaning. It answers the question: how long does it take for a success to occur, when you only know the average number of occurrences (like in the Poisson distribution). The Exponential distribution gives the time to the first occurrence, and its generalization, the Gamma( $\beta$ ) distribution gives the time to the  $\beta$ -th occurrence. Note that the sum of a Gamma( $\beta_1, l_1, s$ ) and Gamma( $\beta_2, l_2, s$ ) distribution has a Gamma( $\beta_1 + \beta_2, l_1 + l_2, s$ ) distribution.

*The Gamma distribution*

If you assume the logarithm of a variable to be Normal distributed, the variable itself is LogNormal-distributed. As a result, it can be shown that the chance of an outcome in the interval  $[x \cdot c_1, x \cdot c_2]$  is equal to the chance of an outcome in the interval  $[x/c_2, x/c_1]$  for some  $x$ . This might be a reasonable assumption in price developments, for example.

*The LogNormal distribution*

The Uniform( $min, max$ ) distribution:

- **Input parameters** :  $min [x], max [x]$
- **Input check** :  $min < max$
- **Permitted values** :  $\{x \mid min \leq x \leq max\}$
- **Standard density** :  $f_{(0,1)}(x) = 1$
- **Mean** :  $1/2$
- **Variance** :  $1/12$

Uniform  
distribution

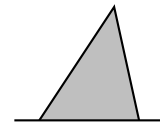


In the Uniform distribution all values of the random variable occur between a fixed minimum and a fixed maximum with equal likelihood. It is quite common to use the Uniform distribution when you have little knowledge about an uncertain parameter in your model except that its value has to lie anywhere within fixed bounds. For instance, after talking to a few appraisers you might conclude that their single appraisals of your property vary anywhere between a fixed pessimistic and a fixed optimistic value.

The Triangular( $\beta, min, max$ ) distribution:

- **Input parameters** : shape  $\beta [-], min [x], max [x]$
- **Input check** :  $min < max, 0 < \beta < 1$
- **Permitted values** :  $\{x \mid min \leq x \leq max\}$
- **Standard density** :  $f_{(\beta,0,1)}(x) = \begin{cases} 2x/\beta & \text{for } 0 \leq x \leq \beta \\ 2(1-x)/(1-\beta) & \text{for } \beta < x \leq 1 \end{cases}$
- **Mean** :  $(\beta + 1)/3$
- **Variance** :  $(1 - \beta + \beta^2)/18$
- **Remarks** : The shape parameter  $\beta$  indicates the position of the peak in relation to the range, i.e.  $\beta = \frac{peak-min}{max-min}$

Triangular  
distribution



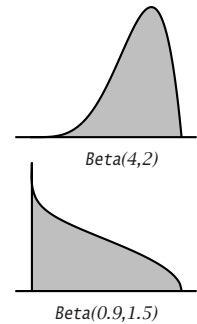
In the Triangular distribution all values of the random variable occur between a fixed minimum and a fixed maximum, but not with equal likelihood as in the Uniform distribution. Instead, there is a most likely value, and its position is not necessarily in the middle of the interval. It is quite common to use the Triangular distribution when you have little knowledge about an uncertain parameter in your model except that its value has to lie anywhere within fixed bounds and that there is a most likely value. For instance, assume that a few appraisers each quote an optimistic as well as a pessimistic value of your property. Summarizing their input you might conclude that their quotes provide not only a well-defined interval but also an indication of the most likely value of your property.

The Beta( $\alpha, \beta, min, max$ ) distribution:

- **Input parameters** : shape  $\alpha$  [-], shape  $\beta$  [-],  $min$  [ $x$ ],  $max$  [ $x$ ]
- **Input check** :  $\alpha > 0, \beta > 0, min < max$
- **Permitted values** :  $\{x \mid min < x < max\}$
- **Standard density** :  $f_{(\alpha, \beta, 0, 1)}(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}$   
 where  $B(\alpha, \beta)$  is the Beta function
- **Mean** :  $\alpha / (\alpha + \beta)$
- **Variance** :  $\alpha\beta(\alpha + \beta)^{-2}(\alpha + \beta + 1)^{-1}$
- **Remarks** : Beta(1,1, $min, max$ )=Uniform( $min, max$ )

The Beta distribution is a very flexible distribution whose two shape parameters allow for a good approximation of almost any distribution on a finite interval. The distribution can be made symmetrical, positively skewed, negatively skewed, etc. It has been used to describe empirical data and predict the random behavior of percentages and fractions. Note that for  $\alpha < 1$  a singularity occurs at  $x = min$  and for  $\beta < 1$  at  $x = max$ .

Beta distribution



The LogNormal( $\beta, min, s$ ) distribution:

- **Input parameters** : shape  $\beta$  [-], lowerbound  $min$  [ $x$ ] and scale  $s$  [ $x$ ]
- **Input check** :  $\beta > 0$  and  $s > 0$
- **Permitted values** :  $\{x \mid min < x < \infty\}$
- **Standard density** :  $f_{(\beta, 0, 1)}(x) = \frac{1}{\sqrt{2\pi x \ln(\beta^2 + 1)}} e^{-\frac{(\ln(x^2(\beta^2+1)))}{2\ln(\beta^2+1)}}$
- **Mean** : 1
- **Variance** :  $\beta^2$

If you assume the logarithm of the variable to be Normal( $\mu, \sigma$ )-distributed, then the variable itself is LogNormal( $\sqrt{e^{\sigma^2}-1}, 0, e^{\mu-\sigma^2/2}$ )-distributed. This parameterization is used for its simple expressions for mean and variance. A typical example is formed by real estate prices and stock prices. They all cannot drop below zero, but they can grow to be very high. However, most values tend to stay within a particular range. You usually can form some expected value of a real estate price or a stock price, and estimate the standard deviation of the prices on the basis of historical data.

LogNormal distribution



The Exponential( $min, s$ ) distribution:

- **Input parameters** : lowerbound  $min$  [ $x$ ] and scale  $s$  [ $x$ ]
- **Input check** :  $s > 0$
- **Permitted values** :  $\{x \mid min \leq x < \infty\}$
- **Standard density** :  $f_{(0,1)}(x) = \lambda e^{-x}$
- **Mean** : 1
- **Variance** : 1
- **Remarks** : Exponential( $min, s$ ) = Gamma(1, $min, s$ )  
 Exponential( $min, s$ ) = Weibull(1, $min, s$ )

Exponential distribution



Assume that you are observing a sequence of independent events with a constant chance of occurring in time, with  $s$  being the average time between occurrences. (in accordance with the Poisson distribution) The  $\text{Exponential}(0, s)$  distribution gives answer to the question: how long a time do you need to wait until you observe the first occurrence of an event. Typical examples are time between failures of equipment, and time between arrivals of customers at a service desk (bank, hospital, etc.).

The  $\text{Gamma}(\beta, \text{min}, s)$  distribution:

- **Input parameters** : shape  $\beta$  [-], lowerbound  $\text{min}$  [ $x$ ] and scale  $s$  [ $x$ ]
- **Input check** :  $s > 0$  and  $\beta > 0$
- **Permitted values** :  $\{x \mid \text{min} < x < \infty\}$
- **Standard density** :  $f_{(\beta,0,1)}(x) = x^{\beta-1} e^{-x} / \Gamma(\beta)$   
where  $\Gamma(\beta)$  is the Gamma function
- **Mean** :  $\beta$
- **Variance** :  $\beta$

Gamma  
distribution



The Gamma distribution gives answer to the question: how long a time do you need to wait until you observe the  $\beta$ -th occurrence of an event (instead of the first occurrence as in the Exponential distribution). Note that it is possible to use non-integer values for  $\beta$  and a location parameter. In these cases there is no natural interpretation of the distribution and for  $\beta < 1$  a singularity exists at  $x = \text{min}$ , so one should be very careful in using the Gamma distribution this way.

The  $\text{Weibull}(\beta, \text{min}, s)$  distribution:

- **Input parameters** : shape  $\beta$  [-], lowerbound  $\text{min}$  [ $x$ ] and scale  $s$  [ $x$ ]
- **Input check** :  $\beta > 0$  and  $s > 0$
- **Permitted values** :  $\{x \mid \text{min} \leq x < \infty\}$
- **Standard density** :  $f_{(\beta,0,1)}(x) = \beta x^{\beta-1} e^{-x^\beta}$
- **Mean** :  $\Gamma(1 + 1/\beta)$
- **Variance** :  $\Gamma(1 + 2/\beta) - \Gamma^2(1 + 1/\beta)$

Weibull  
distribution

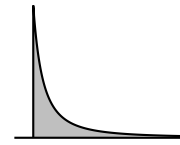


The Weibull distribution is another generalization of the Exponential distribution. It has been successfully used to describe failure time in reliability studies, and the breaking strengths of items in quality control testing. By using a value of the shape parameter that is less than 1, the Weibull distribution becomes steeply declining and could be of interest to a manufacturer testing failures of items during their initial period of use. Note that in that case there is a singularity at  $x = \text{min}$ .

The Pareto( $\beta, l, s$ ) distribution:

- **Input parameters** : shape  $\beta$  [-], location  $l$  [x] and scale  $s$  [x]
- **Input check** :  $s > 0$  and  $\beta > 0$
- **Permitted values** :  $\{x \mid l + s < x < \infty\}$
- **Standard density** :  $f_{(\beta,0,1)}(x) = \beta/x^{\beta+1}$
- **Mean** : for  $\beta > 1$  :  $\beta/(\beta - 1), \infty$  otherwise
- **Variance** : for  $\beta > 2$  :  $\beta(\beta - 1)^{-2}(\beta - 2)^{-1}, \infty$  otherwise

Pareto  
distribution



The Pareto distribution has been used to describe the sizes of such phenomena as human population, companies, incomes, stock fluctuations, etc.

The Normal( $\mu, \sigma$ ) distribution:

- **Input parameters** : Mean  $\mu$  [x] and standard deviation  $\sigma$  [x]
- **Input check** :  $\sigma > 0$
- **Permitted values** :  $\{x \mid -\infty < x < \infty\}$
- **Standard density** :  $f_{(0,1)}(x) = e^{-x^2/2}/\sqrt{2\pi}$
- **Mean** : 0
- **Variance** : 1
- **Remarks** : Location  $\mu$ , scale  $\sigma$

Normal  
distribution



The Normal distribution is frequently used in practical applications as it describes many phenomena observed in real life. Typical examples are attributes such as length, IQ, etc. Note that while the values in these examples are naturally bounded, a close fit between such data values and normally distributed values is quite common in practice, because the likelihood of extreme values away from the mean is essentially zero in the Normal distribution.

The Logistic( $\mu, s$ ) distribution:

- **Input parameters** : mean  $\mu$  [x] and scale  $s$  [x]
- **Input check** :  $s > 0$
- **Permitted values** :  $\{x \mid -\infty < x < \infty\}$
- **Standard density** :  $f_{(0,1)}(x) = (e^x + e^{-x} + 2)^{-1}$
- **Mean** : 0
- **Variance** :  $\pi^2/3$

Logistic  
distribution

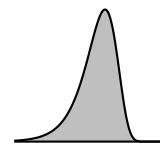


The Logistic distribution has been used to describe growth of a population over time, chemical reactions, and similar processes. Extreme values are more common than in the somewhat similar Normal distribution.

The Extreme Value( $l, s$ ) distribution:

- **Input parameters** : Location  $l$  [x] and scale  $s$  [x]
- **Input check** :  $s > 0$
- **Permitted values** :  $\{x \mid -\infty < x < \infty\}$
- **Standard density** :  $f_{(0,1)}(x) = e^x e^{-e^x}$

Extreme Value  
distribution



- **Mean** :  $\gamma = 0.5772\dots$  (Euler's constant)
- **Variance** :  $\pi^2/6$

Extreme Value distributions have been used to describe the largest values of phenomena observed over time: water levels, rainfall, etc. Other applications include material strength, construction design or any other application in which extreme values are of interest. In literature the Extreme Value distribution that is provided by AIMMS is known as a type 1 Gumbel distribution.

---

### A.3 Distribution operators

The distribution operators discussed in this section can help you to analyze the results of an experiment. For example, it is expected that the sample mean of a sequence of observations gets closer to the mean of the distribution that was used during the observations as the number of observations increases. To compute statistics over a sample, you can use the sample operators discussed in Section A.4 or you can use the histogram functions that are explained in Section 19.3 of The User's Guide. The following distribution operators are available in AIMMS:

*Distribution operators*

- the `DistributionCumulative(distr,x)` operator,
- the `DistributionInverseCumulative(distr, $\alpha$ )` operator,
- the `DistributionDensity(distr,x)` operator,
- the `DistributionInverseDensity(distr, $\alpha$ )` operator,
- the `DistributionMean(distr)` operator,
- the `DistributionDeviation(distr)` operator,
- the `DistributionVariance(distr)` operator,
- the `DistributionSkewness(distr)` operator, and
- the `DistributionKurtosis(distr)` operator.

`DistributionCumulative(distr,x)` computes the probability that a random variable  $X$  drawn from the distribution *distr* is less or equal than  $x$ . Its inverse, `DistributionInverseCumulative(distr, $\alpha$ )`, computes the smallest  $x$  such that the probability that a variable  $X$  is greater than or equal to  $x$  does not exceed  $\alpha$ .

*Cumulative distributions ...*

The `DistributionDensity(distr,x)` expresses the expected density around  $x$  of sample points drawn from a *distr* distribution and is in fact the derivative of `DistributionCumulative(distr,x)`. The `DistributionInverseDensity(distr, $\alpha$ )` is the derivative of `DistributionInverseCumulative(distr, $\alpha$ )`. Given a random variable  $X$ , the `DistributionInverseDensity` can be used to answer the question of how much a given value  $x$  should be increased such that the probability  $P(X \leq x)$  is increased with  $\alpha$  (for small values of  $\alpha$ ).

*... and their derivatives*

For continuous distributions  $distr$ ,  $\alpha \in [0, 1]$ , and  $x = \text{DistributionInverseCumulative}(distr, \alpha)$  it holds that

... for discrete distributions

$$\begin{aligned}\text{DistributionDensity}(distr, x) &= \partial\alpha/\partial x \\ \text{DistributionInverseDensity}(distr, \alpha) &= \partial x/\partial\alpha\end{aligned}$$

Note that the above two relations make it possible to express `DistributionInverseDensity` in terms of `DistributionDensity`. Through this relation the `DistributionInverseDensity` is also defined for discrete distributions.

The operators `DistributionMean`, `DistributionDeviation`, `DistributionVariance`, `DistributionSkewness` and `DistributionKurtosis` provide the mean, standard deviation, variance, skewness and kurtosis of a given distribution. Note that the values computed using the sample operators converges to the values computed using the corresponding distribution operators as the size of the sample increases (the law of large numbers).

*Distribution statistics*

---

## A.4 Sample operators

The statistical sample operators discussed in this section can help you to analyze the results of an experiment. The following operators are available in AIMMS:

*Sample operators*

- the Mean operator,
- the GeometricMean operator,
- the HarmonicMean operator,
- the RootMeanSquare operator,
- the Median operator,
- the SampleDeviation operator,
- the PopulationDeviation operator,
- the Skewness operator,
- the Kurtosis operator,
- the Correlation operator, and
- the RankCorrelation operator.

The results of the Skewness, Kurtosis, Correlation and RankCorrelation operator are unitless. The results of the other sample operators listed above should have the same unit of measurement as the expression on which the statistical computation is performed. Whenever your model contains one or more QUANTITY declarations, AIMMS will perform a unit consistency check on arguments of the statistical operators and their result.

*Associated units*

The following mean computation methods are supported: (arithmetic) mean or average, geometric mean, harmonic mean and root mean square (RMS). The first method is well known and has the property that it is an unbiased estimate of the expectation of a distribution. The geometric mean is defined as the  $N$ -th root of the product of  $N$  values. The harmonic mean is the reciprocal of the arithmetic mean of the reciprocals. The root mean square is defined as the square root of the arithmetic mean of the squares. It is mostly used for averaging the measurements of a physical process. *Mean*

- **Operator** :  $\text{Mean}(\text{domain}, \text{expression})$
- **Formula** :  $\frac{1}{n} \sum_{i=1}^n x_i$
- **Operator** :  $\text{GeometricMean}(\text{domain}, \text{expression})$
- **Formula** :  $n \sqrt[n]{\prod_{i=1}^n x_i}$
- **Operator** :  $\text{HarmonicMean}(\text{domain}, \text{expression})$
- **Formula** :  $\frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$
- **Operator** :  $\text{RootMeanSquare}(\text{domain}, \text{expression})$
- **Formula** :  $\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$

The median is the middle value of a sorted group of values. In case of an odd number of values the median is equal to the middle value. If the number of values is even, the median is the mean of the two middle values. *Median*

- **Operator** :  $\text{Median}(\text{domain}, \text{expression})$
- **Formula** :  $\text{median} = \begin{cases} x_{\frac{N+1}{2}} & \text{if } N \text{ is odd} \\ \frac{1}{2} (x_{\frac{N}{2}} + x_{\frac{N+2}{2}}) & \text{if } N \text{ is even} \end{cases}$

The standard deviation is a measure of dispersion about the mean. It is defined as the root mean square of the distance of a set of values from the mean. There are two kinds of standard deviation: the standard deviation of a sample of a population, also known as  $\sigma_{n-1}$  or  $s$ , and the standard deviation of a population, which is denoted by  $\sigma_n$ . The relation between these two standard deviations is that the first kind is an unbiased estimate of the second kind. This implies that for large  $n$   $\sigma_{n-1} \approx \sigma_n$ . The standard deviation of an sample of a population can be computed by means of *Standard deviation*

- **Operator** :  $\text{SampleDeviation}(\text{domain}, \text{expression})$

$$\blacksquare \text{ Formula : } \sqrt{\frac{1}{n-1} \left( \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 \right)}$$

whereas the standard deviation of a population can be determined by

■ Operator : `PopulationDeviation(domain,expression)`

$$\blacksquare \text{ Formula : } \sqrt{\frac{1}{n} \left( \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 \right)}$$

The skewness is a measure of the symmetry of a distribution. Two kinds of skewness are distinguished: positive and negative. A positive skewness means that the tail of the distribution curve on the right side of the central maximum is longer than the tail on the left side (skewed "to the right"). A distribution is said to have a negative skewness if the tail on the left side of the central maximum is longer than the tail on the right side (skewed "to the left"). In general one can say that a skewness value greater than 1 or less than  $-1$  indicates a highly skewed distribution. Whenever the value is between 0.5 and 1 or  $-0.5$  and  $-1$ , the distribution is considered to be moderately skewed. A value between  $-0.5$  and 0.5 indicates that the distribution is fairly symmetrical.

*Skewness*

■ Operator : `Skewness(domain,expression)`

$$\blacksquare \text{ Formula : } \frac{\sum_{i=1}^n (x_i - \mu)^3}{\sigma_{n-1}^3}$$

where  $\mu$  denotes the mean and  $\sigma_{n-1}$  denotes the standard deviation.

The kurtosis coefficient is a measure for the peakedness of a distribution. If a distribution is fairly peaked, it will have a high kurtosis coefficient. On the other hand, a low kurtosis coefficient indicates that a distribution has a flat peak. It is common practice to use the kurtosis coefficient of the standard Normal distribution, equal to 3, as a standard of reference. Distributions which have a kurtosis coefficient less than 3 are considered to be platykurtic (meaning flat), whereas distributions with a kurtosis coefficient greater than 3 are leptokurtic (meaning peaked). Be aware that in literature also an alternative definition of kurtosis is used in which 3 is subtracted from the formula used here.

*Kurtosis*

■ Operator : `Kurtosis(domain,expression)`

$$\blacksquare \text{ Formula : } \frac{\sum_{i=1}^n (x_i - \mu)^4}{\sigma_{n-1}^4}$$

where  $\mu$  denotes the mean and  $\sigma_{n-1}$  denotes the standard deviation.

The correlation coefficient is a measurement for the relationship between two variables. Two variables are positive correlated with each other when the correlation coefficient lies between 0 and 1. If the correlation coefficient lies between  $-1$  and 0, the variables are negative correlated. In case the correlation coefficient is 0, the variables are considered to be unrelated to one another. Positive correlation means that if one variable increases, the other variable increases also. Negative correlation means that if one variable increases, the other variable decreases.

*Correlation coefficient*

■ **Operator** : Correlation(*domain, x\_expression, y\_expression*)

$$n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i$$

■ **Formula** :

$$\frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{\sqrt{\left( n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 \right) \left( n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right)}}$$

If one wants to determine the relationship between two variables, but their distributions are not equal or the precision of the data is not trusted, one can use the rank correlation coefficient to determine their relationship. In order to compute the rank correlation coefficient the data is ranked by their value using the numbers  $1, 2, \dots, n$ . These rank numbers are used to compute the rank correlation coefficient.

*Rank correlation*

■ **Operator** : RankCorrelation(*domain, x\_expression, y\_expression*)

$$6 \sum_{i=1}^n (\text{Rank}(x_i) - \text{Rank}(y_i))^2$$

■ **Formula** :  $1 - \frac{\sum_{i=1}^n (\text{Rank}(x_i) - \text{Rank}(y_i))^2}{n(n^2 - 1)}$

---

## A.5 Scaling of statistical operators

Shifting and scaling distribution has an effect on the distribution operators, and on sample operators when the samples are from a specified distribution. Location and scale parameters find their origin in a common transformation

*Transforming distributions*

$$x \mapsto \frac{x - l}{s}$$

to shift and stretch a given distribution. By choosing  $l = 0$  and  $s = 1$  one obtains the standard form of a given distribution, and the relation of operators working on the general and standard form of distributions is as follows:

$$\begin{aligned}
X(l, s) &= l + sX(0, 1) \\
\text{DistributionDensity}(x; l, s) &= \frac{1}{s} \text{DistributionDensity}\left(\frac{x-l}{s}; 0, 1\right) \\
\text{DistributionInversDensity}(\alpha; l, s) &= s \cdot \text{DistributionInversDensity}(\alpha; 0, 1) \\
\text{DistributionCumulative}(x; l, s) &= \text{DistributionCumulative}\left(\frac{x-l}{s}; 0, 1\right) \\
\text{DistributionInverseCumulative}(\alpha; l, s) &= l + s \cdot \text{DistributionInverseCumulative}(\alpha; 0, 1) \\
\text{Mean}(X(l, s)) &= l + s \cdot \text{Mean}(X(0, 1)) \\
\text{Median}(X(l, s)) &= l + s \cdot \text{Median}(X(0, 1)) \\
\text{Deviation}(X(l, s)) &= s \cdot \text{Deviation}(X(0, 1)) \\
\text{DistributionVariance}(X(l, s)) &= s^2 \cdot \text{DistributionVariance}(X(0, 1)) \\
\text{Skewness}(X(l, s)) &= \text{Skewness}(X(0, 1)) \\
\text{Kurtosis}(X(l, s)) &= \text{Kurtosis}(X(0, 1)) \\
\text{(Rank)Correlation}(X(l, s), Y) &= \text{(Rank)Correlation}(X(0, 1), Y)
\end{aligned}$$

The transformation formula for the Mean holds for both the `DistributionMean` and the Mean of a sample. However, for the `GeometricMean`, the `HarmonicMean` and the `RootMeanSquare`, only the scale factor can be propagated easily during the transformation. Thus, for a sample taken from a distribution  $X$  and a any mean operator  $M$  from the `GeometricMean`, `HarmonicMean` or `RootMeanSquare`, it holds that

*Transformation  
of the mean*

$$M(X(l, s)) = s \cdot M(X(l, 1))$$

but in general

$$M(X(l, s)) \neq l + M(X(0, s))$$

The transformation formula for the deviation is valid for the `DistributionDeviation`, the `SampleDeviation` and `PopulationDeviation`, while the transformation formulae for the `Skewness` and `Kurtosis` hold for both the distribution and sample operators.

*Transformation  
of the other  
moments*