

# **Statistics Toolbox Release Notes**

---



<b>Summary by Version</b> .....	<b>1</b>
<b>Version 6.2 (R2008a) Statistics Toolbox™ Software</b> ....	<b>3</b>
<b>Version 6.1 (R2007b) Statistics Toolbox™ Software</b> ....	<b>5</b>
<b>Version 6.0 (R2007a) Statistics Toolbox™ Software</b> ....	<b>9</b>
<b>Version 5.3 (R2006b) Statistics Toolbox™ Software</b> ....	<b>13</b>
<b>Version 5.2 (R2006a) Statistics Toolbox™ Software</b> ....	<b>17</b>
<b>Version 5.1 (R14SP3) Statistics Toolbox™ Software</b> ...	<b>21</b>
<b>Version 5.0.2 (R14SP2) Statistics Toolbox™ Software</b> ..	<b>25</b>
<b>Compatibility Summary for Statistics Toolbox Software</b> .....	<b>26</b>



## Summary by Version

This table provides quick access to what's new in each version. For clarification, see “Using Release Notes” on page 1.

Version (Release)	New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
<b>Latest Version V6.2 (R2008a)</b>	Yes Details	Yes Summary	Bug Reports Includes fixes	Printable Release Notes: PDF  Current product documentation
V6.1 (R2007b)	Yes Details	Yes Summary	Bug Reports Includes fixes	No
V6.0 (R2007a)	Yes Details	Yes Summary	Bug Reports Includes fixes	No
V5.3 (R2006b)	Yes Details	Yes Summary	Bug Reports Includes fixes	No
V5.2 (R2006a)	Yes Details	No	Bug Reports Includes fixes	No
V5.1 (R14SP3)	Yes Details	No	No	No
V5.0.2 (R14SP2)	Yes Details	No	Bug Reports Includes fixes	No

### Using Release Notes

Use release notes when upgrading to a newer version to learn about:

- New features
- Changes
- Potential impact on your existing files and practices

Review the release notes for other MathWorks™ products required for this product (for example, MATLAB® or Simulink®) for enhancements, bugs, and compatibility considerations that also might impact you.

If you are upgrading from a software version other than the most recent one, review the release notes for all interim versions, not just for the version you are installing. For example, when upgrading from V1.0 to V1.2, review the release notes for V1.1 and V1.2.

## What's in the Release Notes

### New Features and Changes

- New functionality
- Changes to existing functionality

### Version Compatibility Considerations

When a new feature or change introduces a reported incompatibility between versions, the **Compatibility Considerations** subsection explains the impact.

Compatibility issues reported after the product is released appear under Bug Reports at the MathWorks Web site. Bug fixes can sometimes result in incompatibilities, so you should also review the fixed bugs in Bug Reports for any compatibility impact.

### Fixed Bugs and Known Problems

The MathWorks offers a user-searchable Bug Reports database so you can view Bug Reports. The development team updates this database at release time and as more information becomes available. This includes provisions for any known workarounds or file replacements. Information is available for bugs existing in or fixed in Release 14SP2 or later. Information is not available for all bugs in earlier releases.

Access Bug Reports using your MathWorks Account.

## Version 6.2 (R2008a) Statistics Toolbox™ Software

This table summarizes what's new in Version 6.2 (R2008a):

New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
Yes Details below	Yes Summary	Bug Reports Includes fixes	Printable Release Notes: PDF  Current product documentation

New features and changes introduced in this version are organized by these topics:

- “Descriptive Statistics” on page 3
- “Model Assessment” on page 4
- “Multivariate Methods” on page 4
- “Probability Distributions” on page 4
- “Regression Analysis” on page 4
- “Statistical Visualization” on page 4

### Descriptive Statistics

Bootstrap confidence intervals computed by `bootci` are now more accurate for lumpy data.

### Compatibility Considerations

The formula for `bootci` confidence intervals of type 'bca' or 'cper' involves the proportion of bootstrap statistics less than the observed statistic. The formula now takes into account cases where there are many bootstrap statistics exactly equal to the observed statistic.

## **Model Assessment**

Two new cross-validation functions, `cvpartition` and `crossval`, partition data and assess models in regression, classification, and clustering applications.

## **Multivariate Methods**

A new sequential feature selection function, `sequentialfs`, selects predictor subsets that optimize user-defined prediction criteria.

The new `nnmf` function performs nonnegative matrix factorization (NMF) for dimension reduction.

## **Probability Distributions**

The new `sobolset` and `haltonset` functions produce quasi-random point sets for applications in Monte Carlo integration, space-filling experimental designs, and global optimization. Options allow you to skip, leap over, and scramble the points. The `grandstream` function provides corresponding quasi-random number streams for intermittent sampling.

## **Regression Analysis**

The new `plsregress` function performs partial least-squares regression for data with correlated predictors.

## **Statistical Visualization**

The `normspec` function now shades regions of a normal density curve that are either inside or outside specification limits.



## Version 6.1 (R2007b) Statistics Toolbox™ Software

This table summarizes what's new in Version 6.1 (R2007b):

New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
Yes Details below	Yes Summary	Bug Reports Includes fixes	No

New features and changes introduced in this version are organized by these topics:

- “Cluster Analysis” on page 5
- “Design of Experiments” on page 6
- “Hypothesis Tests” on page 6
- “Probability Distributions” on page 6
- “Regression Analysis” on page 7
- “Statistical Visualization” on page 8

### Cluster Analysis

The new `gmdistribution` class represents Gaussian mixture distributions, where random points come from different multivariate normal distributions with certain probabilities. The `gmdistribution` constructor creates mixture models with specified means, covariances, and mixture proportions, or by fitting a mixture model with a specified number of components to data. Methods for the class include:

- `fit` — Distribution fitting function
- `pdf` — Probability density function
- `cdf` — Cumulative distribution function
- `random` — Random number generator
- `cluster` — Data clustering

- `posterior` — Cluster posterior probabilities
- `mahal` — Mahalanobis distance

The `cluster` function for hierarchical clustering now accepts a vector of cutoff values, and returns a matrix of cluster assignments, with one column per cutoff value.

### **Compatibility Considerations**

The `kmeans` function now returns a vector of cluster indices of length  $n$ , where  $n$  is the number of rows in the input data matrix  $X$ , even when  $X$  contains NaN values. In the past, rows of  $X$  with NaN values were ignored, and the vector of cluster indices was correspondingly reduced in size. Now the vector of cluster indices contains NaN values where rows have been ignored, consistent with other toolbox functions.

### **Design of Experiments**

A new option in the  $D$ -optimal design function `candexch` specifies fixed design points in the row-exchange algorithm. A similar feature is already available for the `daugment` function, which uses the coordinate-exchange algorithm.

### **Hypothesis Tests**

The `kstest` function now uses a more accurate method to calculate the  $p$ -value for a single-sample Kolmogorov-Smirnov test.

### **Compatibility Considerations**

`kstest` now compares the computed  $p$ -value to the desired cutoff, rather than comparing the test statistic to a table of values. Results may differ from those in previous releases, especially for small samples in two-sided tests where an asymptotic formula was used in the past.

### **Probability Distributions**

A new fitting function, `copulafit`, has been added to the family of functions that describe dependencies among variables using copulas. The function fits parametric copulas to data, providing a link between models of marginal distributions and models of data correlations.

A number of probability functions now have improved accuracy, especially for extreme parameter values. The functions are:

- `betainv` — More accurate for probabilities in  $P$  near 1.
- `binocdf` — More efficient and less likely to run out of memory for large values in  $X$ .
- `binopdf` — More accurate when the probabilities in  $P$  are on the order of `eps`.
- `fcdf` — More accurate when the parameter ratios  $V2./V1$  are much less than the values in  $X$ .
- `ncx2cdf` — More accurate in some extreme cases that previously returned 0.
- `poisscdf` — More efficient and less likely to run out of memory for large values in  $X$ .
- `tcdf` — More accurate when the squares of the values in  $X$  are much less than the parameters in  $V$ .
- `tinvs` — More accurate when the probabilities in  $P$  are very close to 0.5 and the outputs are very small in magnitude.

Function-style syntax for `paretotails` objects has been removed.

## Compatibility Considerations

The changes to the probability functions listed above may lead to different, but more accurate, outputs than in previous releases.

In previous releases, syntax of the form `obj(x)` for a `paretotails` objects `obj` invoked the `cdf` method. This syntax now produces a warning. To evaluate the cumulative distribution function, use the syntax `cdf(obj,x)`.

## Regression Analysis

The new `corr cov` function converts a covariance matrix to the corresponding correlation matrix.

The `mvregress` function now supports an option to force the estimated covariance matrix to be diagonal.

### **Compatibility Considerations**

In previous releases the `mvregress` function, when using the `'cwlsl'` algorithm, estimated the covariance of coefficients `COVB` using the estimated, rather than the initial, covariance of the responses `SIGMA`. The initial `SIGMA` is now used, and `COVB` differs to a degree dependent on the difference between the initial and final estimates of `SIGMA`.

### **Statistical Visualization**

The `boxplot` function has a new `'compact'` plot style suitable for displaying large numbers of groups.

## Version 6.0 (R2007a) Statistics Toolbox™ Software

This table summarizes what's new in Version 6.0 (R2007a):

New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
Yes Details below	Yes Summary	Bug Reports Includes fixes	No

New features and changes introduced in this version are organized by these topics:

- “Data Organization” on page 9
- “Hypothesis Testing” on page 10
- “Multivariate Statistics” on page 10
- “Probability Distributions” on page 10
- “Regression Analysis” on page 11
- “Statistical Visualization” on page 12
- “Other Improvements” on page 12

### Data Organization

New categorical and dataset arrays are available for organizing and processing statistical data.

- Categorical arrays facilitate the use of nominal and ordinal categorical data.
- Dataset arrays provide a natural way to encapsulate heterogeneous statistical data and metadata, so that it can be accessed and manipulated using familiar methods analogous to those for numerical matrices.
- Categorical and dataset arrays are supported by a variety of new functions for manipulating the encapsulated data.
- Categorical arrays are now accepted as input arguments in all Statistics Toolbox™ functions that make use of grouping variables.

## Hypothesis Testing

Expanded options are available for linear hypothesis testing.

- The new `linhyptest` function performs linear hypothesis tests on parameters such as regression coefficients. These tests have the form  $H*b = c$  for specified values of  $H$  and  $c$ , where  $b$  is a vector of unknown parameters.
- The `covb` output from `regstats` and the `SIGMA` output from `nlinfit` are suitable for use as the covariance matrix input argument required by `linhyptest`. The following functions have been modified to return a `covb` output for use with `linhyptest`: `coxphfit`, `glmfit`, `mnrfit`, `robustfit`.

## Multivariate Statistics

The new `cholcov` function computes a Cholesky-like decomposition of a covariance matrix, even if the matrix is not positive definite. Factors are useful in many of the same ways as Cholesky factors, such as imposing correlation on random number generators.

The `classify` function for discriminant analysis has been improved.

- The function now computes the coefficients of the discriminant functions that define boundaries between classification regions.
- The output of the function is now of the same type as the input grouping variable `group`.

## Compatibility Considerations

The `classify` function now returns outputs of different type than it did in the past. If the input argument `group` is a logical vector, output is now converted to a logical vector. In the past, output was returned as a cell array of 0s and 1s. If `group` is numeric, the output is now converted to the same type. For example, if `group` is of type `uint8`, the output will be of type `uint8`.

## Probability Distributions

New `paretotails` objects are available for modeling distributions with an empirical cdf or similar distribution in the center and generalized Pareto distributions in the tails.

- The `paretotails` function converts a data sample to a `paretotails` object. The objects are useful for generating random samples from a distribution similar to the data, but with tail behavior that is less discrete than the empirical distribution.
- Objects from the `paretotails` class are supported by a variety of new methods for working with the piecewise distribution.
- The `paretotails` class provides function-like behavior, so that `p(x)` evaluates the cdf of `p` at values `x`.

## Regression Analysis

The new `mvregresslike` function is a utility related to the `mvregress` function for fitting regression models to multivariate data with missing values. The new function computes the objective (log likelihood) function, and can also compute the estimated covariance matrix for the parameter estimates.

New `classregtree` objects are available for creating and analyzing classification and regression trees.

- The `classregtree` function fits a classification or regression tree to training data. The objects are useful for predicting response values from new predictors.
- Objects from the `classregtree` class are supported by a variety of new methods for accessing information about the tree.
- The `classregtree` class provides function-like behavior, so that `t(X)` evaluates the tree `t` at predictor values in `X`.
- The following functions now create or operate on objects from the new `classregtree` class: `treefit`, `treedisp`, `treeeval`, `treefit`, `treeprune`, `treetest`.

## Compatibility Considerations

Objects from the `classregtree` class are intended to be compatible with the structure arrays that were produced in previous versions by the classification and regression tree functions listed above. In particular, `classregtree` supports dot indexing of the form `t.property` to obtain properties of the object `t`. The class also provides function-like behavior through parenthesis indexing, so that `t(x)` uses the tree `t` to classify or compute fitted values for

predictors `x`, rather than index into `t` as a structure array as it did in the past. As a result, cell arrays should now be used to aggregate `classregtree` objects.

## Statistical Visualization

The new `scatterhist` function produces a scatterplot of 2D data and illustrates the marginal distributions of the variables by drawing histograms along the two axes. The function is also useful for viewing properties of random samples produced by functions such as `copularnd`, `mvnrnd`, and `lhsdesign`.

## Other Improvements

- The `mvtrnd` function now produces a single random sample from the multivariate  $t$  distribution if the `cases` input argument is absent.
- The `zscore` function, which centers and scales input data by mean and standard deviation, now returns the means and standard deviations as additional outputs.



## Version 5.3 (R2006b) Statistics Toolbox™ Software

This table summarizes what's new in Version 5.3 (R2006b):

New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
Yes Details below	Yes Summary	Bug Reports Includes fixes	No

New features and changes introduced in this version are organized by these topics:

- “Demos” on page 13
- “Design of Experiments” on page 13
- “Hypothesis Tests” on page 14
- “Multinomial Distribution” on page 14
- “Regression Analysis” on page 15
- “Statistical Process Control” on page 15

### Demos

The following demo has been updated:

- Selecting a Sample Size — Modified to highlight the new `sampsizepwr` function

### Design of Experiments

The following visualization functions, commonly used in the design of experiments, have been added:

- `interactionplot` — Two-factor interaction plot for the mean
- `maineffectplot` — Main effects plot for the mean
- `multivarichart` — Multivari chart for the mean

## Hypothesis Tests

The following functions for hypothesis testing have been added or improved:

- `jbtest` — Replaces the chi-square approximation of the test statistic, which is asymptotic, with a more accurate algorithm that interpolates  $p$ -values from a table of quantiles. A new option allows you to run Monte Carlo simulations to compute  $p$ -values outside of the table.
- `lillietest` — Uses an improved version of Lilliefors' table of quantiles, covering a wider range of sample sizes and significance levels, with more accurate values. New options allow you to test for exponential and extreme value distributions, as well as normal distributions, and to run Monte Carlo simulations to compute  $p$ -values outside of the tables.
- `runstest` — Adds a test for runs up and down to the existing test for runs above or below a specified value.
- `sampsizepwr` — New function to compute the sample size necessary for a test to have a specified power. Options are available for choosing a variety of test types.

## Compatibility Considerations

If the significance level for a test lies outside the range of tabulated values, [0.001, 0.5], then both `jbtest` and `lillietest` now return an error. In previous versions, `jbtest` returned an approximate  $p$ -value and `lillietest` returned an error outside a smaller range, [0.01, 0.2]. Error messages suggest using the new Monte Carlo option for computing values outside the range of tabulated values.

If the data sample for a test leads to a  $p$ -value outside the range of tabulated values, then both `jbtest` and `lillietest` now return, with a warning, either the smallest or largest tabulated value. In previous versions, `jbtest` returned an approximate  $p$ -value and `lillietest` returned NaN.

## Multinomial Distribution

The multinomial distribution has been added to the list of almost 50 probability distributions supported by the toolbox.

- `mnpdf` — Multinomial probability density function

- `mnrnd` — Multinomial random number generator

## Regression Analysis

### Multinomial Regression

Support has been added for multinomial regression modeling of discrete multi-category response data, including multinomial logistic regression. The following new functions supplement the regression models in `glmfit` and `glmval` by providing for a wider range of response values:

- `mnrfit` — Fits a multinomial regression model to data
- `mnrval` — Computes predicted probabilities for the multinomial regression model

### Multivariate Regression

The new `mvregress` function carries out multivariate regression on data with missing response values. An option allows you to specify how missing data is handled.

### Survival Analysis

`coxphfit` — A new option allows you to specify the values at which the baseline hazard is computed.

## Statistical Process Control

The following new functions consolidate and expand upon existing functions for statistical process control:

- `capability` — Computes a wider range of probabilities and capability indices than the `capable` function found in previous releases
- `controlchart` — Displays a wider range of control charts than the `ewmplot`, `schart`, and `xbarplot` functions found in previous releases
- `controlrules` — Supplements the new `controlchart` function by providing for a wider range of control rules (Western Electric and Nelson)

- `gagerr` — Performs a gage repeatability and reproducibility study on measurements grouped by operator and part

### **Compatibility Considerations**

The `capability` function subsumes the `capable` function that appeared in previous versions of Statistics Toolbox™ software, and the `controlchart` function subsumes the functions `ewmaplot`, `schart`, and `xbarplot`. The older functions remain in the toolbox for backwards compatibility, but they are no longer documented or supported.

## Version 5.2 (R2006a) Statistics Toolbox™ Software

This table summarizes what's new in Version 5.2 (R2006a):

New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
Yes Details below	No	Bug Reports Includes fixes	No

New features and changes introduced in this version are organized by these topics:

- “Analysis of Variance” on page 17
- “Bootstrapping” on page 17
- “Demos” on page 18
- “Design of Experiments” on page 18
- “Hypothesis Tests” on page 18
- “Multivariate Distributions” on page 19
- “Random Number Generation” on page 19
- “Robust Regression” on page 20
- “Statistical Process Control” on page 20

### Analysis of Variance

Support for nested and continuous factors has been added to the `anovan` function for  $N$ -way analysis of variance.

### Bootstrapping

The following functions have been added to supplement the existing `bootstrp` function for bootstrap estimation:

- `bootci` — Computes confidence intervals of a bootstrapped statistic. An option allows you to choose the type of the bootstrap confidence interval.

- `jackknife` — Draws jackknife samples from a data set and computes statistics on each sample

## Demos

The following demos have been added to the toolbox:

- Bayesian Analysis for a Logistic Regression Model
- Time Series Regression of Airline Passenger Data

The following demo has been updated to demonstrate new features:

- Random Number Generation

## Design of Experiments

The new `fracfactgen` function finds a set of fractional factorial design generators suitable for fitting a specified model.

The following functions for  $D$ -optimal designs have been enhanced:

- `cordexch`, `daugment`, `dcovary`, `rowexch` — New options specify the range of values and the number of levels for each factor, exclude factor combinations, treat factors as categorical rather than continuous, control the number of iterations, and repeat the design generation process from random starting points
- `candexch` — New options control the number of iterations and repeat the design generation process from random starting points
- `candgen` — New options specify the range of values and the number of levels for each factor, and treat factors as categorical rather than continuous
- `x2fx` — New option treats factors as categorical rather than continuous

## Hypothesis Tests

The new `dwtest` function performs a Durbin-Watson test for autocorrelation in linear regression.

## Multivariate Distributions

Two new functions have been added to compute multivariate cdfs. These supplement existing functions for pdfs and random number generators for the same distributions.

- `mvncdf` — Cumulative distribution function for the multivariate normal distribution
- `mvtcdf` — Cumulative distribution function for the multivariate  $t$  distribution

## Random Number Generation

### Copulas

New functions have been added to the toolbox that allow you to use copulas to model correlated multivariate data and generate random numbers from multivariate distributions.

- `copulacdf` — Cumulative distribution function for a copula
- `copulaparam` — Copula parameters as a function of rank correlation
- `copulapdf` — Probability density function for a copula
- `copularnd` — Random numbers from a copula
- `copulastat` — Rank correlation for a copula

### Markov Chain Monte Carlo Methods

The following functions generate random numbers from nonstandard distributions using Markov Chain Monte Carlo methods:

- `mhsample` — Generate random numbers using the Metropolis-Hasting algorithm
- `slicesample` — Generate random numbers using a slice sampling algorithm

## **Pearson and Johnson Systems of Distributions**

Support has been added for random number generation from Pearson and Johnson systems of distributions.

- `pearsrnd` — Random numbers from a distribution in the Pearson system
- `johnsrnd` — Random numbers from a distribution in the Johnson system

## **Robust Regression**

To supplement the `robustfit` function, the following functions now have options for robust fitting:

- `nlinfit` — Nonlinear least-squares regression
- `nlparci` — Confidence intervals for parameters in nonlinear regression
- `nlpredci` — Confidence intervals for predictions in nonlinear regression

## **Statistical Process Control**

The following control chart functions now support time-series objects:

- `xbarplot` — Xbar plot
- `schart` — Standard deviation chart
- `ewmaplot` — Exponentially weighted moving average plot



## Version 5.1 (R14SP3) Statistics Toolbox™ Software

This table summarizes what's new in Version 5.1 (R14SP3):

New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
Yes Details below	No	No	No

New features and changes introduced in this version are organized by these topics:

- “Demos” on page 21
- “Descriptive Statistics” on page 22
- “Hypothesis Tests” on page 22
- “Probability Distributions” on page 23
- “Regression Analysis” on page 24
- “Statistical Visualization” on page 24

### Demos

The following demos have been added to the toolbox:

- Curve Fitting and Distribution Fitting
- Fitting a Univariate Distribution Using Cumulative Probabilities
- Fitting an Orthogonal Regression Using Principal Components Analysis
- Modelling Tail Data with the Generalized Pareto Distribution
- Pitfalls in Fitting Nonlinear Models by Transforming to Linearity
- Weighted Nonlinear Regression

The following demo has been updated:

- Modelling Data with the Generalized Extreme Value Distribution

## Descriptive Statistics

The new `partialcorr` function computes the correlation of one set of variables while controlling for a second set of variables.

The `grpstats` function now computes a wider variety of descriptive statistics for grouped data. Choices include the mean, standard error of the mean, number of elements, group name, standard deviation, variance, confidence interval for the mean, and confidence interval for new observations. The function also supports the computation of user-defined statistics.

## Hypothesis Tests

### Chi-Square Goodness-of-Fit Test

The new `chi2gof` function tests if a sample comes from a specified distribution, against the alternative that it does not come from that distribution, using a chi-square test statistic.

### Variance Tests

Three functions have been added to test sample variances:

- `vartest` — One-sample chi-square variance test. Tests if a sample comes from a normal distribution with specified variance, against the alternative that it comes from a normal distribution with a different variance.
- `vartest2` — Two-sample  $F$ -test for equal variances. Tests if two independent samples come from normal distributions with the same variance, against the alternative that they come from normal distributions with different variances.
- `vartestn` — Bartlett multiple-sample test for equal variances. Tests if multiple samples come from normal distributions with the same variance, against the alternative that they come from normal distributions with different variances.

### Ansari-Bradley Test

The new `ansaribradley` function tests if two independent samples come from the same distribution, against the alternative that they come from distributions that have the same median and shape but different variances.

## Tests of Randomness

The new `runstest` function tests if a sequence of values comes in random order, against the alternative that the ordering is not random.

## Probability Distributions

Support has been added for two new distributions:

- “Generalized Extreme Value Distribution” on page 23
- “Generalized Pareto Distribution” on page 23

### Generalized Extreme Value Distribution

The Generalized Extreme Value distribution combines the Gumbel, Frechet, and Weibull distributions into a single distribution. It is used to model extreme values in data.

The following distribution functions have been added:

- `gevcdf` — Cumulative distribution function
- `gevfit` — Parameter estimation function
- `gevinv` — Inverse cumulative distribution function
- `gevlike` — Negative log-likelihood function
- `gevpdf` — Probability density function
- `gevrnd` — Random number generator
- `gevstat` — Distribution statistics

### Generalized Pareto Distribution

The Generalized Pareto distribution is used to model the tails of a data distribution.

The following distribution functions have been added:

- `gpcdf` — Cumulative distribution function
- `gpfit` — Parameter estimation function

- `gpinv` — Inverse cumulative distribution function
- `gplike` — Negative log-likelihood function
- `gppdf` — Probability density function
- `gprnd` — Random number generator
- `gpstat` — Distribution statistics

## Regression Analysis

- The new `coxphfit` function fits Cox's proportional hazards regression model to data.
- The new `invpred` function estimates the inverse prediction intervals for simple linear regression.
- The `polyconf` function has new options to let you specify the confidence interval computed.

## Statistical Visualization

Both the `ecdf` and `ksdensity` functions now produce plots when no output arguments are specified.

## Version 5.0.2 (R14SP2) Statistics Toolbox™ Software

This table summarizes what's new in Version 5.0.2 (R14SP2):

New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
Yes Details below	No	Bug Reports Includes fixes	No

New features and changes introduced in this version are organized by this topic:

### **Multivariate Statistics**

The cophenet function now returns cophenetic distances as well as the cophenetic correlation coefficient.

## Compatibility Summary for Statistics Toolbox Software

This table summarizes new features and changes that might cause incompatibilities when you upgrade from an earlier version, or when you use files on multiple versions. Details are provided in the description of the new feature or change.

Version (Release)	New Features and Changes with Version Compatibility Impact
<p><b>Latest Version</b> <b>V6.2 (R2008a)</b></p>	<p>See the <b>Compatibility Considerations</b> subheading for this change:</p> <ul style="list-style-type: none"> <li>• “Descriptive Statistics” on page 3</li> </ul>
<p>V6.1 (R2007b)</p>	<p>See the <b>Compatibility Considerations</b> subheading for each of these new features and changes:</p> <ul style="list-style-type: none"> <li>• “Cluster Analysis” on page 5</li> <li>• “Hypothesis Tests” on page 6</li> <li>• “Probability Distributions” on page 6</li> <li>• “Regression Analysis” on page 7</li> </ul>
<p>V6.0 (R2007a)</p>	<p>See the <b>Compatibility Considerations</b> subheading for each of these new features and changes:</p> <ul style="list-style-type: none"> <li>• “Multivariate Statistics” on page 10</li> <li>• “Regression Analysis” on page 11</li> </ul>

<b>Version (Release)</b>	<b>New Features and Changes with Version Compatibility Impact</b>
V5.3 (R2006b)	See the <b>Compatibility Considerations</b> subheading for each of these new features and changes: <ul style="list-style-type: none"><li>• “Hypothesis Tests” on page 14</li><li>• “Statistical Process Control” on page 15</li></ul>
V5.2 (R2006a)	None
V5.1 (R14SP3)	None
V5.0.2 (R14SP2)	None