Statistics Toolbox Release Notes

Contents

Summary by Version

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

Version (Release)	New Features and Changes	Version Compatibility Considerations	Fixed Bugs and Known Problems	Related Documentation at Web Site
Latest Version V5.3 (R2006b)	Yes Details	Yes Summary	Bug Reports Includes fixes	Printable Release Notes: PDF Current product documentation
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

About Release Notes

Use release notes when upgrading to a newer version to learn about new features and changes, and the potential impact on your existing files and practices. Release notes are also beneficial if you use or support multiple versions.

If you are not upgrading from the most recent previous version, review 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 New Features and Changes, Version Compatibility Considerations, and Bug Reports for V1.1 and V1.2.

New Features and Changes

These include

- New functionality

- Changes to existing functionality

- Changes to system requirements (complete system requirements for the current version are at the MathWorks Web site)

- Any version compatibility considerations associated with each new feature or change

Version Compatibility Considerations

When a new feature or change introduces a known incompatibility between versions, its description includes a **Compatibility Considerations** subsection that details the impact. For a list of all new features and changes that have compatibility impact, see the "Compatibility Summary for Statistics Toolbox" on page 19.

Compatibility issues that become known after the product has been released are added to Bug Reports at the MathWorks Web site. Because bug fixes can sometimes result in incompatibilities, also review fixed bugs in Bug Reports for any compatibility impact.

Fixed Bugs and Known Problems

MathWorks Bug Reports is a user-searchable database of known problems, workarounds, and fixes. The MathWorks updates the Bug Reports database as new problems and resolutions become known, so check it as needed for the latest information.

Access Bug Reports at the MathWorks Web site using your MathWorks Account. If you are not logged in to your MathWorks Account when you link to Bug Reports, you are prompted to log in or create an account. You then can view bug fixes and known problems for R14SP2 and more recent releases.

The Bug Reports database was introduced for R14SP2 and does not include information for prior releases. You can access a list of bug fixes made in prior versions via the links in the summary table.

Related Documentation at Web Site

Printable Release Notes (PDF). You can print release notes from the PDF version, located at the MathWorks Web site. The PDF version does not support links to other documents or to the Web site, such as to Bug Reports. Use the browser-based version of release notes for access to all information.

Product Documentation. At the MathWorks Web site, you can access complete product documentation for the current version and some previous versions, as noted in the summary table.

Version 5.3 (R2006b) Statistics Toolbox

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	Printable Release Notes: PDF Current product documentation

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

- "Demos" on page 5
- "Design of Experiments" on page 5
- "Hypothesis Tests" on page 5
- "Multinomial Distribution" on page 6
- "Regression Analysis" on page 6
- "Statistical Process Control" on page 7

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

- `maineffectsplot` — 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 ewmaplot, 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 the Statistics Toolbox, 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

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 9
- "Bootstrapping" on page 9
- "Demos" on page 9
- "Design of Experiments" on page 9
- "Hypothesis Tests" on page 10
- "Multivariate Distributions" on page 10
- "Random Number Generation " on page 10
- "Robust Regression" on page 11
- "Statistical Process Control" on page 11

Analysis of Variance

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

Bootstrapping

To 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

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 13
- "Descriptive Statistics" on page 13
- "Hypothesis Tests" on page 13
- "Probability Distributions" on page 14
- "Regression Analysis" on page 15
- "Statistical Visualization" on page 16

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:

13

- `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 14
- "Generalized Pareto Distribution" on page 15

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

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" on page 18

Multivariate Statistics

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

Compatibility Summary for Statistics Toolbox

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
Latest Version V5.3 (R2006b)	See the **Compatibility Considerations** subheading for each of these new features and changes: • "Hypothesis Tests" on page 5 • "Statistical Process Control" on page 7
V5.2 (R2006a)	None
V5.1 (R14SP3)	None
V5.0.2 (R14SP2)	None