

# New Journal: Annals of Applied Statistics

Bradley Efron, Stanford University, USA

The Annals of Applied Statistics, newest journal from the Institute of Mathematical Statistics, began

accepting papers on November 1, 2006. Computational statistics is an area of particular interest. Please visit the journal website at <http://www.imstat.org/aoas> for further information.

## Forthcoming Events: useR! 2007

by Dianne Cook

The first North American useR! will be held at Iowa State University, Ames, Iowa, August 8–10, 2007, which will be a week after JSM'07.

This follows successful meetings in Vienna, Austria, in 2006 and 2004, and also Directions in Statistical Computing (DSC) meetings in Auckland, NZ (Feb 2007), Seattle (2005), and Vienna (1999, 2001, 2003). Information about the meeting can be found at <http://www.useR2007.org/> or by emailing [useR2007@iastate.edu](mailto:useR2007@iastate.edu).

The important dates to remember are:

- April 23, 2007: Deadline for paper submission,

with referees comments returned by April 30.

- May 1, 2007: Early registration ends
- June 30: Poster abstract submission ends
- August 8, 2007: useR! 2007 begins!

We look forward to meet you in Ames!

On behalf of the organizing committee:

Diane Cook  
Department of Statistics  
Iowa State University  
Ames, Iowa, U.S.A.  
[dicook@iastate.edu](mailto:dicook@iastate.edu)

## Changes in R 2.5.0

by the R Core Team

### User-visible changes

- `apropos(x)` and `find(x)` now both only work for character `x`, and hence drop all non-standard evaluation behaviour.
- Data frames can have 'automatic' row names which are not converted to `dimnames` by `as.matrix()`. (Consequently, e.g., `t()` for such data frames has `NULL` column names.) This change leads to memory reductions in several places, but can break code which assumes character `dimnames` for data frames derived from matrices.  
No existing R object is regarded as having 'automatic' row names, and it may be beneficial to recreate such objects via `read.table()` or `data.frame()`.
- Using `$` on an atomic vector now raises a warning, as does `use` on an S4 class for which a method has not been defined.
- The Unix-alike readline terminal interface now does command-completion for R objects, incor-

porating the functionality formerly in package **rcompletion** by Deepayan Sarkar. This can be disabled by setting the environment variable `R_COMPLETION` to `FALSE` when starting R (e.g., in `'~/Renvirom'`).

### New features

- `abbreviate()` no longer has an 8191 byte limit on the size of strings it can handle.
- `abs(x)` now returns integer for integer or logical arguments.
- `apropos()` has a new argument `ignore.case` which defaults to `TRUE`, potentially matching more than previously, thanks to a suggestion by Seth Falcon.
- `args()`, `str()` and `print()` now give the argument lists of primitive functions.
- `as.matrix()` gains the `...` argument that several packages have assumed it always had (and S-PLUS has).
- Manipulation of integers as roman numerals via `as.roman()` in package **utils**.

- `attr()` no longer treats `name = NA_character_` as meaning `name = "NA"`.
- `binom.test()` now allows a 'fuzz' for calculated integer values in its `x` and `n` arguments.
- `boxplot(*, notch = TRUE)` now warns when notches are outside hinges; related to PR#7690.
- New function `callCC()` providing a downward-only version of Scheme's `call` with current continuation.
- `capabilities()` now has a `profmem` entry indicating whether R has been compiled with memory profiling.
- `colnames<-()` and `rownames<-()` now handle data frames explicitly, so calling `colnames<-` on a data frame no longer alters the representation of the row names.
- `commandArgs()` has a new `trailingOnly` argument to be used in conjunction with `'--args'`.
- `contour()` now passes graphical parameters in `...` to `axis()` and `box()`.
- New data set `crimtab` on Student (1908)'s 3000 criminals.
- `cut.default()` has a new argument `ordered_result`.
- `.deparseOpts()` has two new options: `keepNA` to ensure that different types (logical, integer, double, character and complex) of NAs are distinguished, and `S_compatible` to suppress the use of R-specific features such as 123L and to deparse integer values of a double vector with a trailing decimal point.

The `keepInteger` option now uses the suffix `L` rather than `as.integer()` where possible (unless all entries are `NA` or `S_compatible` is also set).

Other deparse options can now be added to `all` (which has not for some time actually switched on all options).

Integer sequences `m:n` are now deparsed in that form.

- `deparse()` and `dput()` now include `keepInteger` and `keepNA` in their defaults for the `control` argument.
- `detach()` now takes another argument, `unload`, which indicates whether or not to unload the package and then only cleans up the `S4` methods if the package successfully unloads.
- There are new constants `NA_integer_`, `NA_real_`, `NA_complex_` and `NA_character_` to denote NAs of those types, and they will be used in deparsing in place of `as.integer(NA)` etc unless `.deparseOpts()` includes `S_compatible`.
- `dev.print()` now recognizes screen devices as all those with an enabled display list, rather than a hard-coded set.
- Objects of class `difftime` are now handled more flexibly. The units of such objects can now be accessed via a `units()` function, which also has a replacement form, and there are conversion methods to and from numeric, which also allow the specification of units. Objects of this class can also be stored in data frames now. A `format()` method has been added, and the `print` method was revised.
- New function `environmentName()` to give the print name of environments such as `'namespace:base'`. This is now used by `str()`.
- New function `env.profile()` provides R level access to summary statistics on environments. In a related patch, `new.env()` now allows the user to specify an initial size for a hashed environment.
- `file()` can read the X11 clipboard selection as `"X11_clipboard"` on suitable X11-using systems.
- `file("stdin")` is now recognized, and refers to the process's `stdin` file stream whereas `stdin()` refers to the console. These may differ, for example for a GUI console, an embedded application of R or if `'--file='` has been used.
- `file_test()` is now also available in package **utils**. (It is now private in package **tools**.)
- `file.show()` gains an encoding argument.
- New functions `formatUL()` and `formatOL()` in package **utils** for formatting unordered (`itemize`) and ordered (`enumerate`) lists.
- The statistics reported when `gcinfo(TRUE)` are now of the amounts used (in Mb) and not of the amounts free (which are not really relevant when there are no hard limits, only gc trigger points).
- New function `get_all_vars()` to retrieve all the (untransformed) variables that the default method of `model.frame()` would use to create the model frame.
- `interaction()` has a new argument `lex.order`.

- `initialize()` (in **methods**) now tries to be smarter about updating the new instance in place, thereby reducing copying.
- `install.packages(dependencies = NA)` is a new default, which is to install essential dependencies when installing from repositories to a single library. As a result of this change, `update.packages()` will install any new dependencies of the packages it is updating (alongside the package in the same library tree).

If `lib` is not specified or is specified of length one and the chosen location is not a writable directory, `install.packages()` offers to create a personal library directory for you if one does not already exist, and to install there.

- `is.atomic`, `is.call`, `is.character`, `is.complex`, `is.double (= is.real)`, `is.environment`, `is.expression`, `is.function`, `is.integer`, `is.list`, `is.logical`, `is.null`, `is.object`, `is.pairlist`, `is.recursive`, `is.single` and `is.symbol (= is.name)` are no longer internally S3 generic, nor can S4 methods be written for them.
- The factor methods of `is.integer` and `is.numeric` have been replaced by internal code.
- Added `is.raw()` for completeness.
  - `l10n_info()` also reports if the current locale is Latin-1.
  - `levels<-()`, `names()` and `names<-()` now dispatch internally for efficiency and so no longer have S3 default methods.
  - `.libPaths()` now does both tilde and glob expansion.
  - Functions `lm()`, `glm()`, `loess()`, `xtabs()` and the default method of `model.frame()` coerce their formula argument (if supplied) to a formula.
  - `max()`, `min()` and `range()` now work with character vectors.
  - `message()` has a new argument `appendLF` to handle messages with and without new-lines. There is a new message class `packageStartupMessage` that can be suppressed separately.
  - A new function, `method.skeleton()` writes a skeleton version of a call to `setMethod()` to a file, with correct arguments and format, given the name of the function and the method signature.

- `mode<-` and `storage.mode<-` do slightly less copying.
- `nls.control(*, printEval = FALSE, warnOnly = FALSE)` are two new options to help better analyze (non-)convergence of `nls()`, thanks to Kate Mullen.  
`nls()` and `summary(nls())` now contain more information and also print information about convergence.
- `options(device = )` now accepts a function object as well as the name of a function.
- `pdf()` supports new values for `paper` of `"US"` (same as `"letter"`), `"a4r"` and `"USr"` (the latter two meaning rotated to landscape). `postscript()` also accepts `paper = "US"`.
- `persp()` now respects the graphical pars `cex.axis`, `cex.lab`, `font.axis` and `font.lab`.
- New faster internal functions `pmax.int()` and `pmin.int()` for inputs which are atomic vectors without classes (called by `pmax/pmin` where applicable).  
`pmin/pmax` are now more likely to work with classed objects: they work with `POSIXlt` date-times, for example.
- `postscript()` now by default writes grey colors (including black and white) via `setgray`, which gives more widely acceptable output. There are options to write pure RGB, CMYK or gray via the new argument `colormodel`.
- `rbind.data.frame()` now ignores all zero-row inputs, as well as zero-column inputs (which it used to do, undocumented). This is because `read.table()` can create zero-row data frames with `NULL` columns, and those cannot be extended.
- `readChar()` and `writeChar()` can now work with a raw vector.
- `read.table()`, `write.table()` and allies have been moved to package **utils**.
- `rgb()` now accepts the red, green and blue components in a single matrix or data frame.
- New utility function `RShowDoc()` in package **utils** to find and display manuals and other documentation files.
- New `.row_names_info()` utility function finds the number of rows efficiently for data frames; consequently, `dim.data.frame()` has become very fast for large data frames with 'automatic' row names.
- `RSiteSearch()` now also allows to search postings of the 'R-devel' mailing list.

- `screepplot()` is now (S3) generic with a default method, thanks to a patch from Gavin Simpson.
- Experimental `verbose` argument for `selectMethod()`. Might be replaced later by a better interface for method selection inspection.
- Added links to source files to the parsing routines, so that `source()` can now echo the original source and comments (rather than deparising). This affects `example()` and `Sweave()` as well.
- `stack()` and `unstack()` have been moved to package **utils**.
- `strptime()` now sets the 'tzone' attribute on the result if `tz != ""`.
- `str.default()` typically prints fewer entries of logical vectors.
- The `RweaveLatex` driver for `Sweave()` now supports two new options: `expand=FALSE`, to show chunk references in the output, and `concordance=TRUE`, to output the concordance between input and output lines.
- `system()` now takes the same set of arguments on all platforms, with those which are not applicable being ignored with a warning. Unix-alikes gain `input` and `wait`, and Windows gains `ignore.stderr`.
- `system.time()` and `proc.time()` now return an object of class `proc_time` with a `print()` method that returns a POSIX-like format with names.
- `Sys.getenv()` has a new argument `unset` to allow `unset` and `set` to "" to be distinguished (if the OS does). The results of `Sys.getenv()` are now sorted (by name).
- New function `Sys.glob()`, a wrapper for the POSIX.2 function `glob(3)` to do wildcard expansion (on systems which have it, plus an emulation on Windows).
- `Sys.setenv()` is a new (and preferred) synonym for `Sys.putenv()`. The internal C code uses the POSIX-preferred `setenv` rather than `putenv` where the former is available.
- New function `Sys.unsetenv()` to remove environment variables (on systems where `unsetenv` is implemented or `putenv` can remove variables, such as on Windows).
- `text()`, `mtext()`, `strheight()`, `strwidth()`, `legend()`, `axis()`, `title()`, `pie()`, `grid.text()` and `textGrob()` all attempt to coerce non-language annotation objects (in the sense of `is.object`) to character vectors. This is principally intended to cover factors and POSIXt and Date objects, and is done via the new utility function `as.graphicsAnnot()` in package **grDevices**.
- `tk_select.list()` in package **tcltk** now chooses the width to fit the widest item.
- `retracemem()` and `untracemem()` are now primitives for efficiency and so migrate from **utils** to **base**.
- `union()`, `intersect()`, `setdiff()` and `setequal()` now coerce their arguments to be vectors (and they were documented only to apply to vectors).
- `uniroot()` now works if the zero occurs at one of the ends of the interval (suggestion of Tamas Papp).
- There is a new function `View()` for viewing matrix-like objects in a spreadsheet, which can be left up whilst R is running.
- New function `withVisible()` allows R level access to the visibility flag.
- `zip.file.extract()` has been moved to package **utils**.
- A few more cases of subassignment work, e.g., `raw[] <- list` and `vector[] <- expression`, with suitable coercion of the LHS.
- There is a warning if a '\ ' is used unnecessarily in a string when being parsed., e.g., '\ .' where probably '\\ .' was intended. ('\ .' is valid, but the same as '.'). Thanks to Bill Dunlap for the suggestion.
- Introduced the suffix `L` for integer literals to create integer rather than numeric values, e.g., `100L`, `0x10L`, `1e2L`.
- Set the parser to give verbose error messages in case of syntax errors.
- The class `LinearMethodsList` has been extended and will be used to create list versions of methods, derived from the methods tables (environments). The older recursive `MethodsList` class will be deprecated (by the release of 2.5.0 if possible).
- There are more flexible ways to specify the default library search path. In addition to `R_LIBS` and `.Library`, there are `.Library.site` (defaults to 'R\_HOME/site-library') and `R_LIBS_USER` (defaults to a platform- and version-specific directory in '~ / R'). See `?libPaths` for details.

- LAPACK has been updated to version 3.1.0. This should cause only small changes to the output, but do remember that the sign of eigenvectors (and principal components) is indeterminate.
- PCRE has been updated to version 7.0.
- Several functions handle row names more efficiently:
  - `read.table()` and `read.DIF()` make use of integer row names where appropriate, and avoid at least one copy in assigning them.
  - `data.frame()` and the standard `as.data.frame()` methods avoid generating long dummy row names and then discarding them.
  - `expand.grid()` and `merge()` generate compact 'automatic' row names.
  - `data.matrix()` and `as.matrix.data.frame()` have a new argument 'rownames.force' that by default drops 'automatic' row names.
- `data_frame[i, j]` is substantially more memory-efficient when only a small part of the data frame is selected, especially when (part of) a single column is selected.
- Command-line R (and 'Rterm.exe' under Windows) accepts the options '`-f filename`', '`--file=filename`' and '`-e expression`' to follow other script interpreters. These imply '`--no-save`' unless '`--save`' is specified.
- Invalid bytes in character strings in an MBCS now deparse/print in the form '`\xc1`' rather than '`<c1>`', which means they can be parsed/scanned.
- Printing functions (without source attributes) and expressions now preserves integers (using the L suffix) and NAs (using `NA_real_` etc where necessary).
- The 'internal' objects `.helpForCall`, `.tryHelp` and `topicName` are no longer exported from **utils**.
- The internal regex code has been upgraded to glibc 2.5 (from 2.3.6).
- Text help now attempts to display files which have an encoding section in the specified encoding via `file.show()`.
- R now attempts to keep track of character strings which are known to be in Latin-1 or UTF-8 and print or plot them appropriately in other locales. This is primarily intended to

make it possible to use data in Western European languages in both Latin-1 and UTF-8 locales. Currently `scan()`, `read.table()`, `readLines()`, `parse()` and `source()` allow encodings to be declared, and console input in suitable locales is also recognized.

New function `Encoding()` can read or set the declared encodings for a character vector.

- There have been numerous performance improvements to the data editor on both Windows and X11. In particular, resizing the window works much better on X11.

## Deprecated & defunct

- `symbol.C()` and `symbol.For()` are defunct, and have been replaced by wrappers that give a warning.
- Calling a builtin function with an empty argument is now always an error.
- The autoloading of `ts()` is defunct.
- The undocumented reserved word `GLOBAL.ENV` has been removed. (It was yet another way to get the value of the symbol `.GlobalEnv`.)
- The deprecated behaviour of `structure()` in adding a class when specifying with `tsp` or `levels` attributes is now defunct.
- `unix()` is now finally defunct, having been deprecated for at least seven years.
- `Sys.putenv()` is now deprecated in favour of `Sys.setenv()`, following the POSIX recommendation.
- Building R with '`--without-iconv`' is deprecated.
- Using `$` on an atomic vector is deprecated (it was previously valid and documented to return `NULL`).
- The use of `storage.mode<-` for other than standard types (and in particular for value "single") is deprecated: use `mode<-` instead.

## Installation

- A suitable `iconv` (e.g., from glibc or GNU `libiconv`) is required. For 2.5.x only you can build R without it by configuring using '`--without-iconv`'.
- There is support again for building on AIX (tested on 5.2 and 5.3) thanks to Ei-ji Nakama.

- Autoconf 2.60 or later is used to create configure. This makes a number of small changes, and incorporates the changes to the detection of a C99-compliant C compiler backported for 2.4.1.
- Detection of a Java development environment was added such that packages don't need to provide their own Java detection.  
R CMD javareconf was updated to look for the corresponding Java tools as well.
- Added workaround for reported non-POSIX sh on OSF1. (PR#9375)
- make install-strip now works, stripping the executables and also the shared libraries and modules on platforms where libtool knows how to do so.
- Building R as a shared library and standalone nmath now installs pkg-config files 'libR.pc' and 'libRmath.pc' respectively.
- Added test for insufficiently complete implementation of sigaction.

## C-level facilities

- Functions str2type, type2char and type2str are now available in 'Rinternals.h'.
- Added support for Objective C in R and packages (if available).
- R\_ParseVector() has a new 4th argument SEXP srcfile allowing source references to be attached to the returned expression list.
- Added ptr\_R\_WriteConsoleEx callback which allows consoles to distinguish between regular output and errors/warnings. To ensure backward compatibility it is only used if ptr\_R\_WriteConsole is set to NULL.

## Utilities

- Additional Sweave() internal functions are exported to help writing new drivers, and RweaveLatexRuncode() is now created using a helper function (all from a patch submitted by Seth Falcon).
- The following additional flags are accessible from R CMD config: OBJC, OBJCFLAGS, JAR, JAVA, JAVAC, JAVAH, JAVA\_HOME, JAVA\_LIBS
- R CMD build now takes the package name from the 'DESCRIPTION' file and not from the directory. (PR#9266)

- checkS3methods() (and hence R CMD check) now checks agreement with primitive internal generics, and checks for additional arguments in methods where the generic does not have a ... argument.

codoc() now knows the argument lists of primitive functions.

- R CMD INSTALL and R CMD REMOVE now use as the default library (if '-l' is not specified) the first library that would be used if R were run in the current environment (and they run R to find it).
- There is a new front-end 'Rscript' which can be used for #! scripts and similar tasks. See help("Rscript") and 'An Introduction to R' for further details.
- R CMD BATCH (not Windows) no longer prepends 'invisible(options(echo = TRUE))' to the input script. This was the default unless '--slave' is specified and the latter is no longer overridden.

On all OSes it makes use of the '-f' argument to R, so file("stdin") can be used from BATCH scripts.

On all OSes it reports proc.time() at the end of the script unless q() is called with options to inhibit this.

- R CMD INSTALL now prepends the installation directory (if specified) to the library search path.
- Package installation now re-encodes R files and the 'NAMESPACE' file if the 'DESCRIPTION' file specifies an encoding, and sets the encoding used for reading files in preparing for LazyData. This will help if a package needs to be used in (say) both Latin-1 and UTF-8 locales on different systems.
- R CMD check now reports on non-ASCII strings in datasets. (These are a portability issue, which can be alleviated by marking their encoding: see 'Writing R Extensions'.)
- Rdiff now converts CRLF endings in the target file, and converts UTF-8 single quotes in either to ASCII quotes.
- New recommended package **codetools** by Luke Tierney provides code-analysis tools. This can optionally be used by R CMD check to detect problems, especially symbols which are not visible.
- R CMD config now knows about LIBnn.

- New recommended package **rcompgen** by Deepayan Sarkar provides support for command-line completion under the Unix terminal interface (provided readline is enabled) and the Windows Rgui and Rterm front ends.

## Bug fixes

- `gc()` can now report quantities of 'Vcells' in excess of 16Gb on 64-bit systems (rather than reporting NA).
- Assigning class factor to an object now requires it has integer (and not say double) codes.
- `structure()` ensures that objects with added class factor have integer codes.
- The formula and outer attributes of datasets `ChickWeight`, `CO2`, `DNase`, `Indometh`, `Loblolly`, `Orange` and `Theoph` now have an empty environment and not the environment used to dump the datasets in the package.
- Dataset `Seatbelts` now correctly has class `c("mts", "ts")`.
- `str()` now labels classes on data frames more coherently.
- Several 'special' primitives and `.Internals` could return invisibly if the evaluation of an argument led to the visibility flag being turned off. These included `as.character()`, `as.vector()`, `call()`, `dim()`, `dimnames()`, `lapply()`, `rep()`, `seq()` and `seq_along()`. Others (e.g., `dput()` and `print.default()`) could return visibly when this was not intended.
- Several primitives such as `dim()` were not checking the number of arguments supplied before method dispatch.
- Tracing of primitive functions has been corrected. It should now be the case that tracing either works or is not allowed for all primitive functions. (Problems remain if you make a primitive into a generic when it is being traced. To be fixed later.)
- `max.col()` now omits infinite values in determining the relative tolerance.
- R CMD Sweave and R CMD Stangle now respond to `'--help'` and `'--version'` like other utilities.
- `.libPaths()` adds only existing directories (as it was documented to, but could add non-directories).
- `setIs()` and `setClassUnion()` failed to find some existing subclasses and produced spurious warnings, now fixed.
- `data.frame()` ignored `row.names` for 0-column data frames, and no longer treats an explicit `row.names=NULL` differently from the default value.
- `identical()` looked at the internal structure of the `row.names` attribute, and not the value visible at R level.
- `abline(reg)` now also correctly works with intercept-only `lm` models, and `abline()` warns more when it's called illogically.
- `warning()` was truncating messages at `getOption("warning.length") - 1` (not as documented), with no indication. It now appends `[... truncated]`.
- `Stangle/Sweave` were throwing spurious warnings if options `result` or `strip.white` were unset.
- `all.equal()` was ignoring `check.attributes` for list and expression targets, and checking only attributes on raw vectors. Logical vectors were being compared as if they were numeric, (with a mean difference being quoted).
- Calculating the number of significant digits in a number was itself subject to rounding errors for digits  $\geq 16$ . The calculation has been changed to err on the side of slightly too few significant digits (but still at least 15) rather than far too many. (An example is `print(1.001, digits=16)`.)
- `unlink()` on Unix-alikes failed for paths containing spaces.
- `substr()` and friends treated NA `start` or `stop` incorrectly.
- `merge(x, y, all.y = TRUE)` would sometimes incorrectly return logical columns for columns only in `y` when there were no common rows.
- `read.table(fn, col.names=)` on an empty file returned NULL columns, rather than `logical(0)` columns (which is what results from reading a file with just a header).
- `grid.[xy]axis(label=logical(0))` failed.
- `expression()` was unnecessarily duplicating arguments.
- `as.expression(list)` returned a single-element expression vector, which was not compatible with S: it now copies lists element-by-element.

- `supsmu(periodic = TRUE)` could segfault. (PR#9502, detection and patch by Bill Dunlap.)
- `pmax/pmin` called with only logical arguments did not coerce to numeric, although they were documented to do so (as `max/min` do).
- `methods()` did not know that `cbind()` and `rbind()` are internally generic.
- `dim(x) <- NULL` removed the names of `x`, but this was always undocumented. It is not clear that it is desirable but it is S-compatible and relied on, so is now documented.
- `which(x, arr.ind = TRUE)` did not return a matrix (as documented) if `x` was an array of length 0.
- C-level `duplicate()` truncated CHARSXPs with embedded nuls.
- Partial matching of attributes was not working as documented in some cases if there were more than two partial matches or if 'names' was involved.
- `data(package=character(0))` was not looking in './data' as documented.
- `summary.mlm()` failed if some response names were "" (as can easily happen if `cbind()` is used).
- The `postscript()` and `pdf()` drivers shared an encoding list but used slightly different formats. This caused problems if both were used with the same non-default encoding in the same session. (PR#9517)
- The data editor was not allowing Inf, NA and NaN to be entered in numerical columns. It was intended to differentiate between empty cells and NAs, but did not do so: it now does so for strings.
- `supsmu()` could segfault if all cases had non-finite values. (PR#9519)
- `plnorm(x, lower.tail=FALSE)` was returning the wrong tail for  $x \leq 0$ . (PR#9520)
- `which.min()` would not report a minimum of +Inf, and analogously for `which.max()`. (PR#9522)
- R CMD check could fail with an unhelpful error when checking Rd files for errors if there was only one file and that had a serious error. (PR#9459)
- `try()` has been reimplemented using `tryCatch()` to solve two problems with the original implementation: (i) `try()` would run non-NULL options("error") expressions for errors within a try, and (ii) `try()` would catch user interrupts.
- `str(obj)` could fail when `obj` contained a dendrogram.
- Using `data_frame[, last_column] <- NULL` failed (PR#9565).
- `choose(n, k)` could return non-integer values for integer `n` and small `k` on some platforms.
- `nclass.scott(x)` and `nclass.FD(x)` no longer return NaN when `var(x)` or `IQR(x)` (respectively) is zero.  
`hist()` now allows `breaks = 1` (which the above patch will return), but not `breaks = Inf` (which gave an obscure error).
- `strptime("%j")` now also works for the first days of Feb-Dec. (PR#9577)
- `write.table()` now recovers better if `file` is an unopened connection. (It used to open it for both the column names and the data.)
- Fixed bug in `mosaicplot(sort=)` introduced by undocumented change in R 2.4.1 (changeset r39655).
- `contr.treatment(n=0)` failed with a spurious error message. (It remains an error.)
- `as.numeric()` was incorrectly documented: it is identical to `as.double()`.
- `jitter(rep(-1, 3))` gave NaNs. (PR#9580)
- `max.col()` was not random for a row of zeroes. (PR#9542)
- `ansari.test(conf.int=TRUE, exact=FALSE)` failed.
- `trace()` now works on S3 registered methods, by modifying the version in the S3 methods table.
- `rep(length=1, each=0)` segfaulted.
- `postscript()` could overflow a buffer if used with a long command argument.
- The internal computations to copy complete attribute lists did not copy the flag marking S4 objects, so the copies no longer behaved like S4 objects.
- The C code of `nllminb()` was altering a variable without duplicating it. (This did not affect `nllminb()` but would have if the code was called from a different wrapper.)
- `smooth(kind = "3RS3R")` (the current default) used `.C(DUP = FALSE)` but altered its input argument. (This was masked by duplication in `as.double()`.)



- The signature for the predefined S4 method for `as.character()` was missing . . . .
- `readBin(raw_vector)` could read beyond the end of the vector when size-changing was involved.
- The C entry point `PrintValue` (designed to emulate auto-printing) would not find `show()` for use on S4 objects, and did not have the same search path (for `show()`, `print()` and `print()`

methods) as auto-printing. Also, auto-printing and `print()` of S4 objects would fail to find `show()` if the methods namespace was loaded but the package was not attached (or otherwise not in the search path).

- `print()` (and auto-printing) now recognize S4 objects even when `methods` is not loaded, and print a short summary rather than dump the internal structure.

## Changes on CRAN

by Kurt Hornik

### New contributed packages

**ARES** Allelic richness estimation, with extrapolation beyond the sample size. Generates an allelic richness accumulation curve. This curve shows the expected number of unique alleles in a population when taking a sample of individuals. The function `aresCalc` takes a binary data matrix as input, showing the presence of alleles per individual, and gives an accumulation curve (mean with 95% confidence bounds) back. The function `aresPlot` can be used to plot the output from `aresCalc`. By Emiel van Loon and Scott Davis.

**BayHaz** Bayesian Hazard Rate Estimation: a suite of R functions for Bayesian estimation of smooth hazard rates via Compound Poisson Process (CPP) priors. By Luca La Rocca.

**BiasedUrn** Biased Urn model distributions. Statistical models of biased sampling in the form of univariate and multivariate non-central hypergeometric distributions, including those of Wallenius and Fisher (also called extended hypergeometric distribution). By Agner Fog.

**BootCL** Bootstrapping test for chromosomal localization. By Eun-Kyung Lee, Samsun Sung, and Heeal Kim.

**Broddingnag** Very large numbers in R. Real numbers are held using their natural logarithms, plus a logical flag indicating sign. The package includes a vignette that gives a step-by-step introduction to using S4 methods. By Robin K. S. Hankin.

**CCA** Canonical correlation analysis. Provides a set of functions that extend the `cancor` function with new numerical and graphical outputs. It also include a regularized extension of the canonical correlation analysis to deal with data

sets with more variables than observations. By Ignacio González and Sébastien Déjean.

**CreditMetrics** Functions for calculating the CreditMetrics risk model. By Andreas Wittmann.

**DAAGxtras** Data sets and functions additional to **DAAG**, used in additional exercises for the book "Data Analysis and Graphics Using R" by J. H. Maindonald and W. J. Brain (2007, 2nd edn.), and for laboratory exercises prepared for a 'Data Mining' course. By John Maindonald.

**GenABEL** genome-wide association analysis between quantitative or binary traits and SNPs. By Yurii Aulchenko.

**GeoXp** Interactive exploratory spatial data analysis. A tool for researchers in spatial statistics, spatial econometrics, geography, ecology etc., allowing to link dynamically statistical plots with elementary maps. By Christine Thomas-Agnan, Yves Aragon, Anne Ruiz-Gazen, Thibault Laurent, and Laurianne Robidou.

**HydroMe** Estimation of the parameters in infiltration and water retention models by curve-fitting method. The models considered are those that are commonly used in soil science. By Christian Thine Omuto.

**IPSUR** Data sets and functions accompanying the book "Introduction to Probability and Statistics Using R" by G. Andy Chang and G. Jay Kerns (in progress). By G. Jay Kerns with contributions by Theophilus Boye, adapted from the work of John Fox et al.

**InfNet** Simulation of epidemics in a network of contacts. The simulations consider SIR epidemics with events in continuous time (exponential inter-event times). It can consider a structure of local networks and has an option to visualize it with the animator called SoNIA (<http://www.stanford.edu/group/sonia/>). By Lilia Ramirez Ramirez and Mary Thompson.